Forum of Mathematics, Sigma (2025), **Vol. 13**:e17 1–71
doi:10.1017/fms.2024.152

CAMBRIDGE UNIVERSITY PRESS

RESEARCH ARTICLE

Large N limit of the Yang–Mills measure on compact surfaces II: Makeenko–Migdal equations and the planar master field

Antoine Dahlqvist [1] and Thibaut Lemoine [2]

[1]University of Sussex, School of Mathematical and Physical Sciences, Pevensey 3 Building, Brighton, UK;
E-mail: a.dahlqvist@sussex.ac.uk (corresponding author).
[2]Collège de France, 3, rue d'Ulm, Paris, 75005, France; E-mail: thibaut.lemoine@college-de-france.fr.

Received: 19 January 2024; **Revised:** 22 August 2024; **Accepted:** 23 November 2024

2020 Mathematical Subject Classification: *Primary* – 60B15, 60B20, 81T13, 81T32; *Secondary* – 46L54, 57M10, 57R56

Abstract

This paper considers the large N limit of Wilson loops for the two-dimensional Euclidean Yang–Mills measure on all orientable compact surfaces of genus larger or equal to 1, with a structure group given by a classical compact matrix Lie group. Our main theorem shows the convergence of all Wilson loops in probability, given that it holds true on a restricted class of loops, obtained as a modification of geodesic paths. Combined with the result of [20], a corollary is the convergence of all Wilson loops on the torus. Unlike the sphere case, we show that the limiting object is remarkably expressed thanks to the master field on the plane defined in [3, 39], and we conjecture that this phenomenon is also valid for all surfaces of higher genus. We prove that this conjecture holds true whenever it does for the restricted class of loops of the main theorem. Our result on the torus justifies the introduction of an interpolation between free and classical convolution of probability measures, defined with the free unitary Brownian motion but differing from t-freeness of [5] that was defined in terms of the liberation process of Voiculescu [67]. In contrast to [20], our main tool is a fine use of Makeenko–Migdal equations, proving uniqueness of their solution under suitable assumptions, and generalising the arguments of [21, 33].

Contents

1 Introduction **2**
- 1.1 Yang–Mills measure and master field, statement of results 4
- 1.2 Atiyah–Bott–Goldman measure 7
- 1.3 Noncommutative distribution and master field on the torus: an interpolation between free and classical convolution 9
  - 1.3.1 Noncommutative probability and free independence 9
  - 1.3.2 Free unitary Brownian motion and t-freeness 10
  - 1.3.3 A matrix approximation for another interpolation from classical to free convolution 11
- 1.4 Strategy of proof via Makeenko–Migdal deformations 12

2 Homology and homotopy on embedded graphs **15**
- 2.1 Four equivalence relations on paths and loops on maps 15
- 2.2 Discrete homology, winding function and Makeenko–Migdal vectors 20
- 2.3 Regular polygon tilings of the universal cover, tiling-length of a tame loop and geodesic loops 26
- 2.4 Shortening homotopy sequence 29

2.5 Nested and marked loops 33
2.6 Pull and twist moves 34
2.7 Vertex desingularisation and complexity 35
3 Yang–Mills measure and Makeenko–Migdal equations 37
3.1 Metric and heat kernel on classical groups 37
3.2 Area weighted maps, Yang–Mills measure and area continuity 37
3.3 Planar master field, main results and conjecture 42
3.4 Invariance in law and Wilson loop expectation 46
3.5 Makeenko–Migdal equations, existence and uniqueness problem 47
4 Proof of the main result, stability of convergence under deformation 51
4.1 Non-null homology loops 51
4.2 Null homology loops 53
5 Proof of convergence after surgery 58
6 Interpolation between regular representations 59
6.1 State extension and interpolation 59
6.2 Master field on the torus and t-freeness 62
7 Appendix 64
7.1 Casimir element and trace formulas 64
7.2 Makeenko–Migdal equations 65
References 69

1. Introduction

The two-dimensional Yang–Mills measure is a probability model originating from Euclidean quantum field theory in the setting of pure gauge theory. It describes a generalised random connection on a principle bundle over a two-dimensional manifold, with a compact Lie group as structure group, making rigorous the path integral over connections for the so-called Yang–Mills action. Different equivalent mathematical definitions have been given in two dimensions and are due to [31, 23, 56, 32, 1, 2, 42], or more recently[1] to [10, 15]. The work of [71] brought to light many special features of the Yang–Mills measure in two dimensions, including its partial integrability, used as a way to perform exact volume computations for the Atiyah–Bott–Goldman measure [4, 27] on the space of flat connections [46, 8, 57].

When a compact Lie group G and a surface Σ are given, the Yang–Mills measure can be mathematically understood as a random matrix model which assigns to any loop[2] of the surface a random matrix so that concatenation and reversion of loops are compatible with the group operations. In [39], it is shown that it gives rise to a random homomorphism from the group of rectifiable reduced loops of the surface to the chosen group G.

We consider here a closed, connected, orientable surface Σ of genus $g \geq 1$ and a group G belonging to a series of classical compact matrix groups. We are primarily interested in the traces of these matrices, called *Wilson loops,* when the rank of G goes to infinity. We ask whether Wilson loops converge in probability under the Yang–Mills measure, towards a deterministic function.

Let us try to give a brief historical account of this problem. In physics, a motivation for the focus on Wilson loops is due to K. Wilson's work [69] related to quarks confinement. The idea of studying the large rank regime in gauge theories, known as large N limit, was first initiated by t'Hooft [65] on QCD. This lead to many articles in theoretical physics in the 1980's studying the question in two dimensions, a partial list being [36, 37, 49, 51, 70, 29, 28, 30]. In mathematics, this problem was advertised by I. Singer in [62], where the candidate limit of Wilson loops was called *master field*, following the physics

[1] See [14] for a review of the stochastic quantisation approach and [60] for a review focusing on its application to large N limit problems. See also [11] for recent progress in three dimensions for the Yang-Mills measure coupled to Higgs fields

[2] with enough regularity

literature. The case of the plane and the sphere have been respectively proved in [72, 3, 43] and[3] [21]. The case of general compact surfaces has been first investigated by [33], where loops contained in topological discs can be considered under small area constraints and when the convergence is assumed for simple loops. The study of similar questions in the plane for analogs of the Yang–Mills measure has been treated in [9]. In higher dimension, an analog[4] of this question for a lattice model has also been considered [12]. Very recently and independently from the current work, it was shown in [47, 48] that under the Atiyah–Bott–Goldman measure, which can be understood as the weak limit of the Yang–Mills measure when the area of the surface vanishes, the expectation of Wilson loops converges and has a $\frac{1}{N}$ expansion when the group belongs to the series of special unitary matrices and the surface is closed, orientable and of genus $g \geq 2$. For further details and references on the motivations of this problem, we refer to [20, Sec. 1] and [44, Sec. 2.5.].

In this article, we give a complete answer in the case of the torus and a conjecture and a partial result for all surfaces with genus $g \geq 2$. It is the sequel of [20] where we have shown the convergence for a large[5] but incomplete class of loops. Let us recall that in the case of the plane, the master field can be described thanks to free probability and more specifically in terms of free unitary Brownian motion [3, 39]. The case of the sphere involves a different noncommutative stochastic process called the free unitary Brownian bridge [21]. In contrast, for the torus, we show that after lifting loops to the universal cover, the master field is also described by the planar master field, and we conjecture that the same holds true for any surface of higher genus. In the torus case, the master field provides an interpolation parametrised by the total area of the torus, between the free and the classical convolution of two Haar unitaries built with the free unitary Brownian motion, which differs from the t-freeness introduced by [5] using the liberation process of [67].

One of the main technical contributions of this article is to strengthen the stability of Wilson loops convergence under homotopy equivalences established in [21, 33] considering all topologies, and removing the small area constraint. The main tool at stake is a set of recursive equations named after Makeenko and Migdal [49]. When a loop is deformed in a specific way – that we call a Makeenko–Migdal deformation – these equations relate the differential of the expected Wilson loops with the expectation of a product of Wilson loops having a smaller number of intersection points. These equations can be understood as a remarkable instance and a simplification of Schwinger–Dyson equations used in random matrix theory.[6] They were first inferred heuristically in [49] as an integration by parts for the path integral over the space of connections. A first rigorous proof was given in the case of the plane in[7] [39] and was later tremendously simplified and generalised in [26, 25] in a local way that applies to any surface. Makeenko–Migdal equations were crucial to [21, 33] leading to an induction argument on the number of intersection points that reduced the convergence of all Wilson loops on the sphere to the case of simple loops. In the case of other surfaces, the very same strategy fails a priori, as some loops cannot be deformed to simpler loops without raising the number of intersection points, while some homotopy classes do not contain any loop for which the convergence is known to hold. We show here that the first hurdle can be overcome, allowing to reduce the problem, completely in the torus case and partially when $g \geq 2$, to the class of loops considered in [20]. We leave the completion of this program for all compact surfaces to a future work.

[3]See also [33], where a conditional result was obtained implying the case of the sphere, given the convergence for simple loops.

[4]Though in this case, there is at the time of writing, no construction of the continuous Yang–Mills measure in dimension 3 and higher.

[5]Informally described as all simple loops or iteration of simple loops, and all loops which do not visit one handle of the surface.

[6]See, for instance, [16], and [12, 61] for Schwinger-Dyson type equations in lattice gauge theory with Wilson action.

[7]See also [18, sect. 7] for a variation of this proof and [39, section 0] for the heuristics of the original proof of [49] based on an integration by parts in infinite dimension. See also [24] for a proof closer in spirits to the original geometric argument of [49], in an axial gauge setting.

Organisation of the paper

The first four following sections of the introduction give respectively an informal definition of the Yang–Mills measure and of the main results, a discussion on the relation with the Atiyah–Bott–Goldman measure and the work [48, 47], a consequence of the result on the torus in noncommutative probability, and lastly, a sketch of the strategy of the main proofs. Section 2 recalls and adapts some combinatorial notions of discrete homotopy and homology of loops in embedded graph instrumental to the proof. Section 3 gives the definition of the Yang–Mills measure and a statement of the Makeenko–Migdal equations and states the main results of the article. Section 4 consists of the proof of our main technical result, which is Proposition 3.22. Section 5 describes the behaviour of Wilson loops when one performs surgery on the underlying surface. Section 6 shows how the master field on the torus provides a new interpolation between classical and free convolution of Haar unitaries, and proves an analog of this statement for other surfaces. In an appendix, for the sake of completeness, we recall and prove several results on Makeenko–Migdal equations that are quite standard in the literature for unitary groups but not necessarily for all classical groups.

1.1. Yang–Mills measure and master field, statement of results

We shall first give a heuristic definition of the Yang–Mills measure in its geometric setting and state informally the main results of the current article. Proper definitions and statements are respectively given in Sections 3.2 and 3.3.

Let Σ be either a compact, connected, closed orientable surface of genus $g \geq 1$ endowed with a Riemannian metric – we shall call it a *compact surface of genus g* in the sequel – or the Euclidean plane $\mathbb{R}^2$ with its standard inner product. Let G_N be a classical compact matrix Lie group of size N (i.e., viewed as a compact subgroup of $\mathrm{GL}_N(\mathbb{C})$). We assume that the Lie algebra $\mathfrak{g}_N$ of G_N is endowed with the following Ad-invariant inner product:

$$\langle X, Y\rangle = \frac{\beta N}{2}\mathrm{Tr}(X^*Y), \quad \forall X, Y \in \mathfrak{g}_N, \tag{1}$$

where β is equal to 1 if $G_N = \mathrm{SO}(N)$, 2 if $G_N = \mathrm{U}(N)$ or $\mathrm{SU}(N)$, and 4 if $G_N = \mathrm{Sp}(N)$. Given a G_N-principal bundle (P, π, Σ), a connection is a 1-form ω on M valued in adjoint fibre bundle $\mathrm{ad}(P)$; its curvature is the $\mathrm{ad}(P)$-valued 2-form $\Omega = d\omega + \frac{1}{2}[\omega \wedge \omega]$. The *Yang–Mills action* of a connection ω on a G_N-principal bundle (P, π, Σ) is defined by

$$S_{\mathrm{YM}}(\omega) = \frac{1}{2}\int_\Sigma \langle \Omega \wedge \star\Omega\rangle, \tag{2}$$

where $\star$ denotes the Hodge operator. An important feature of dimension 2 is that whenever Ψ is a diffeomorphism of Σ preserving its volume form,

$$S_{\mathrm{YM}}(\Psi_*\omega) = S_{\mathrm{YM}}(\omega). \tag{3}$$

The *Euclidean Yang–Mills measure* is the formal Gibbs measure

$$d\mu_{\mathrm{YM}}(\omega)\text{``}=\text{''}\frac{1}{Z}e^{-S_{\mathrm{YM}}(\omega)}\mathcal{D}\omega, \tag{4}$$

where $\mathcal{D}\omega$ plays the role of a formal Lebesgue measure on the space of connections over an arbitrary principal bundle[8] and Z is a normalisation constant supposed to ensure the total mass to be 1. We

[8]There is here an apparent additional issue with this vague definition. A slightly less dubious state space could be obtained by fixing a representant of each principal bundle equivalence class over Σ and by considering instead the set of pairs of a principal bundle belonging to this family together with a connection on it. When Σ is a contractible space or if G_N is simply connected, there is only one equivalence class of G_N-principal bundles over Σ and this issue disappears. We shall not discuss further the question of the type of the principal bundle under the Yang–Mills in this text. For more details and rigorous results, we refer to [40].

choose here not to include a parameter in front of the action, as it can be included in the volume form of Σ.

The space $\mathcal{A}(P)$ being infinite-dimensional, the latter equation has no mathematical meaning. Though at first stance, as the Yang–Mills action of ω can be seen as the L^2-norm of the curvature Ω, an analogy with Gaussian measures can be hoped. However, when G_N is not abelian, Ω depends nonlinearly on ω, which prevents any direct construction of μ_{YM} using a Gaussian measure. In two dimensions, this nonlinearity can be compensated by the so-called gauge symmetry of S_{YM}, which allows to bypass this problem. This enabled the constructions of [31, 23, 56] based on stochastic calculus. See also [13] for a recent approach defining further a random, distribution-valued connection on trivial bundles over the two-dimensional torus. We follow here instead the approach of [42], which focuses on the holonomy of a connection, whose law can be directly defined using the heat kernel on G_N. The definition we are using is recalled in Section 3.2; it agrees with the construction of [31, 23, 56] thanks to the so-called Driver–Sengupta formula. An important feature of this measure is suggested by (3). For any two-dimensional Riemannian manifold Σ' diffeomorphic to Σ, and for any diffeomorphism $\Psi : \Sigma \to \Sigma'$, there is an induced measure $\Psi_*(\mathrm{YM}_\Sigma)$ on connections of $(P, \Psi \circ \pi, \Sigma')$. If Ψ preserves the area, then

$$\Psi_*(\mathrm{YM}_\Sigma) = \mathrm{YM}_{\Sigma'}.$$

We shall call this property the *area-invariance* of the Yang–Mills measure. Moreover, for any relatively compact, contractible, open subset U of Σ, the restriction to U induces a measure $\mathcal{R}_*^U(\mathrm{YM}_\Sigma)$ on connections of the bundle $(\pi^{-1}(U), \pi, U)$. When Σ is the Euclidean plane $\mathbb{R}^2$ or the Poincaré disc $\mathbb{D}_{\mathfrak{h}}$, with its usual (hyperbolic) metric, it satisfies[9] $\mathcal{R}_*^U(\mathrm{YM}_\Sigma) = \mathrm{YM}_U$, where U is endowed with the metric of Σ.

Let ω be a connection on a G_N-principal bundle (P, π, Σ), and U be an open subset of Σ where $\pi : \pi^{-1}(U) \to U$ can be[10] trivialised. When such a trivialisation has been fixed, its *holonomy* is a function $\gamma \mapsto \mathrm{hol}(\omega, \gamma)$ mapping paths[11] $\gamma : [0, 1] \to U$ to elements of the group G_N such that

$$\mathrm{hol}(\omega, \gamma_1\gamma_2) = \mathrm{hol}(\omega, \gamma_2)\mathrm{hol}(\omega, \gamma_1)$$

for any paths γ_1 and γ_2 such that the endpoint of γ_1 coincides with the starting point of γ_2, while for any path γ,

$$\mathrm{hol}(\omega, \gamma^{-1}) = \mathrm{hol}(\omega, \gamma)^{-1},$$

where $\gamma_1\gamma_2$ and γ^{-1} denote the concatenation and reversion of the paths.

When G_N is a group of matrices of size N and ℓ is a loop of U, the *Wilson loop* associated to ℓ is the function

$$W_\ell(\omega) = \mathrm{tr}(\mathrm{hol}(\omega, \ell)),$$

where $\mathrm{tr} = \frac{1}{N}\mathrm{Tr}$, with Tr the usual trace of matrices. This function can be shown to be independent of the choice of local trivialisation of (P, π, Σ) and is therefore only a function of ω and ℓ.

Our primary source of interest is the study of the random variables $W_\ell := W_\ell(\omega)$, for loops of Σ, when ω is sampled according to YM_Σ. We are interested in the large N limit of W_ℓ, when the scalar product $\langle \cdot, \cdot \rangle$ is chosen as (1) and the volume form of the surface is fixed. The paper [62] seems to be the first mathematical article addressing this question, and it motivates the following conjecture, also suggested by [43, 25, 33].

---

[9]Compact surfaces do not have this property, but there is still absolute continuity in place of equality. This was instrumental in [20].

[10]The tubular neighbourhood of a smooth loop or of an embedded graph could be such an open set.

[11]In this section, the space of paths is not specified and could be taken as the space of piecewise smooth paths with constant speed and transverse intersections. A loop is a path with starting point equal to its endpoint.

https://doi.org/10.1017/fms.2024.152 Published online by Cambridge University Press

Conjecture 1.1. *Let G_N be a classical compact matrix Lie group of size N, endowed with the metric of Section 3.1 and denote by Σ a compact surface of genus $g \geq 0$, the Euclidean plane $\mathbb{R}^2$ or the Poincaré disc $\mathbb{D}_{\mathfrak{h}}$. For any loop ℓ of Σ, there is a constant $\Phi_\Sigma(\ell)$ such that under YM_Σ,*

$$W_\ell \to \Phi_\Sigma(\ell) \text{ in probability as } N \to \infty. \tag{5}$$

The functional Φ_Σ is called the master field on Σ.

The case of plane was first proved in [72, 3] for $G_N = \mathrm{U}(N)$. In [43], the above statement was proved simultaneously to [3] for all groups mentioned and for a large family of loops given by loops of finite length. Moreover, motivated by the physics articles [49, 51, 37], Lévy proved in [43] recursion relations giving a way to compute explicitly $\Phi_{\mathbb{R}^2}$ for all loops with finitely many intersections.

By area invariance and restriction property, the result on the hyperbolic plane can be deduced directly from these latter works as follows. According to a theorem of Moser [54], any relatively compact open disc U of $\mathbb{D}_{\mathfrak{h}}$ with hyperbolic volume t can be mapped to the open Euclidean disc D_t of $\mathbb{R}^2$ centered at 0 and of area t, by a diffeomorphism $\Psi : U \to D_t$ sending the restriction of the hyperbolic volume form on U to the restriction of Euclidean volume form on D_t. By area-invariance, $\mathcal{R}^U_*(\mathrm{YM}_{\mathbb{D}_{\mathfrak{h}}}) = \mathrm{YM}_U = \Psi^{-1}_*(\mathrm{YM}_{D_t})$, so that the conjecture holds true for $\mathbb{D}_{\mathfrak{h}}$ with

$$\Phi_{\mathbb{D}_{\mathfrak{h}}}(\ell) = \Phi_{\mathbb{R}^2}(\Psi \circ \ell)$$

for any loop ℓ with range included in U.

For $\Sigma = \mathbb{S}^2$, the conjecture was proved in [21] for all loops of finite length and $G_N = \mathrm{U}(N)$, while [33] gave a conditional result on $\mathbb{S}^2$ based on an argument similar to [21], as well as a conditional result for other surfaces for loops included in a topological disc, given convergence of Wilson loops for simple loops. In [20], we gave an alternative argument proving a generalisation of the results by Hall in [33] on compact surfaces while relaxing his assumptions; see Section 1.4. The current article was written with the aim to strengthen the argument common to [21] and [33] in order to address the conjecture on all compact manifolds. This led to the following theorem and conjecture.

Theorem 1.2. *When $\mathbb{T}_T$ is a torus of volume $T > 0$, conjecture 1.1 is valid. Moreover, considering $\mathbb{T}_T$ as the quotient of the Euclidean plane $\mathbb{R}^2$ by $\sqrt{T}.\mathbb{Z}^2$,*

$$\Phi_{\mathbb{T}_T}(\ell) = \begin{cases} \Phi_{\mathbb{R}^2}(\tilde{\ell}) & \text{if } \ell \text{ is contractible,} \\ 0 & \text{otherwise,} \end{cases}$$

where for any continuous loop ℓ in $\mathbb{T}_T$, $\tilde{\ell}$ is a lift of ℓ to $\mathbb{R}^2$, that is a smooth loop of $\mathbb{R}^2$, whose projection on $\mathbb{R}^2/\sqrt{T}.\mathbb{Z}^2$ is ℓ.

We discuss an interpretation of this result in terms of noncommutative probability in Section 1.3. For compact surfaces of higher genus, a natural candidate is given as follows. Recall that for any compact surface Σ of volume $T > 0$ and genus $g \geq 2$, there is a covering map $p : \mathbb{D}_{\mathfrak{h}} \to \Sigma$ mapping the hyperbolic metric of $\mathbb{D}_{\mathfrak{h}}$ to the metric of Σ.

Conjecture 1.3. *For any compact surface Σ of genus $g \geq 2$, with universal cover $p : D_{\mathfrak{h}} \to \Sigma$, the conjecture 1.1 is valid with*

$$\Phi_\Sigma(\ell) = \begin{cases} \Phi_{\mathbb{D}_{\mathfrak{h}}}(\tilde{\ell}) & \text{if } \ell \text{ is contractible,} \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

The conjecture 1.3 is also justified by the main result of [20], which implies the following corollary (together with Lemma 3.11 that we will state later). Recall that a simple loop ℓ of Σ is separating if the set $\Sigma \setminus \ell$, where ℓ also denotes the range of the loop, has two connected components $\Sigma_{1,\ell}, \Sigma_{2,\ell}$.

Corollary 1.4. *If ℓ is a separating loop of compact surface Σ of genus $g \geq 1$ and $\Sigma_{2,\ell}$ is not a disc, then under YM_Σ, the convergence (5) holds true with the limit (6).*

In the present paper, we also obtain two conditional results that seem in line with the above conjecture.

Proposition 1.5. *For any compact surface of genus $g \geq 2$, when G_N is a classical compact matrix group of size N, assume that for any geodesic loop ℓ of Σ with nonzero homology, under YM_Σ,*

$$W_\ell \to 0 \text{ in probability as } N \to \infty. \tag{7}$$

Then (7) also holds true for all loops with nonzero homology.

This proposition will follow straightforwardly from Proposition 3.22.

Assume $g \geq 2$ and Γ_g is a discrete subgroup of isometries acting freely, properly on $\mathbb{D}_\mathfrak{h}$ and that $\mathbb{D}_\mathfrak{h}/\Gamma_g$ is a compact surface of genus g with finite total volume $T > 0$. There is a fundamental domain for this action given by a $4g$ hyperbolic polygon D of volume T, centred at 0.

Theorem 1.6. *The conjecture 1.3 holds true if (7) is true for every noncontractible loop ℓ of Σ such that its lift $\tilde{\ell}$ to $\mathbb{D}_\mathfrak{h}$ can be written $\tilde{\ell} = \gamma_1\gamma_2$, where γ_2 is a geodesic, and γ_1 is smooth, included in $\overline{D}$ and intersecting ∂D at most once, transversally at its endpoint.*

A more precise statement is given in Theorem 3.14. Besides, the recent results of [47] are consistent with the above statement as discussed in the next subsection. An outline of the proofs of Theorems 1.2, 1.6, and Proposition 1.5 will be given in subsection 1.4, and the proofs themselves will be the subject of Sections 4 and 5.

1.2. Atiyah–Bott–Goldman measure

Another measure on connections is due to Atiyah, Bott and Goldman [4, 27] when $g \geq 2$. Recently, the limit of Wilson loops under this measure has been investigated by [48, 47]; we discuss the relation with our result.

Let G be a compact connected semisimple[12] Lie group G, $\mathfrak{g}$ its Lie algebra, endowed with an invariant inner product, and $Z(G)$ its center. For any $g \geq 2$, let $K_g : G^{2g} \to G$ be the product of commutators:

$$K_g(a_1, b_1, \ldots, a_g, b_g) = [a_1, b_1] \cdots [a_g, b_g].$$

The space

$$\mathcal{M}_g = K_g^{-1}(e)/G$$

is called the *moduli space of flat G-connections* over a compact surface of genus $g \geq 2$, where G acts by diagonal conjugation, as

$$h.(z_1, \ldots, z_{2g}) = (hz_1h^{-1}, \ldots, hz_{2g}h^{-1}), \ \forall z \in G^{2g}, g \in G.$$

For any $z \in G^{2g}$, its isotropy group is $Z_z = \{h \in G, h.z = z\}$. The set $\mathcal{M}_g^0 = \{z \in G^{2g} : Z_z = Z(G)\}$ can be shown to be a manifold [27, 58] of dimension $(2g-2)\dim(G)$, endowed with a symplectic form ω_A with finite total volume. Besides, using the holonomy map along a suitable $2g$–tuple $\ell_1, \ldots, \ell_{2g}$ of loops, $\mathcal{M}_g^0$ can be identified with a subset of smooth connections ω on a G-principal bundle over Σ such

[12] Mind that this excludes $\mathrm{U}(N)$.

https://doi.org/10.1017/fms.2024.152 Published online by Cambridge University Press

that $S_{\mathrm{YM}}(\omega) = 0$. This subset is a manifold with a symplectic structure [4], equal to the push-forward of ω_A. The Atiyah–Bott–Goldman measure is the volume form on $\mathcal{M}_g^0$ associated to ω_A, given by

$$\mathrm{vol}_g = \frac{\omega_A^{\frac{1}{2}\dim \mathcal{M}_g^0}}{(\frac{1}{2}\dim \mathcal{M}_g^0)!}. \tag{8}$$

Let us denote by $\mu_{ABG,g}$ the probability measure on $\mathcal{M}_g^0$ obtained by normalising vol_g. It appeared in [71] that integrating against the Yang–Mills measure on a compact surface of total area T and letting T tend to 0 allows to obtain formulas for integrals against $\mu_{ABG,g}$. This convergence was proved rigorously by Sengupta in [58]. Using the holonomy mapping of the Yang–Mills measure, the convergence can be understood as follows. Consider a heat kernel $(p_t)_{t>0}$ on G, when its Lie algebra $\mathfrak{g}$ is endowed with its Killing form $\langle \cdot, \cdot \rangle$.

Theorem 1.7 (Symplectic limit of Yang–Mills measure). *Let $f : G^{2g} \to \mathbb{C}$ be a continuous G-invariant function and $\tilde{f} : \mathcal{M}_g^0 \to \mathbb{C}$ be the induced function on the moduli space. Then*

$$\lim_{T \downarrow 0} \int_{G^{2g}} f(x) p_T(K_g(x)) dx = \frac{\mathrm{vol}(G)^{2-2g}}{|Z|} \int_{\mathcal{M}_g^0} \tilde{f} d\mathrm{vol}_g. \tag{9}$$

For any word w in the variables $a_1, \dots, b_g$ and their inverses, setting

$$W_w(z) = \frac{1}{N}\mathrm{Tr}(w(z_1, z_1^{-1}, \dots, z_{2g}, z_{2g}^{-1})), \ \forall z \in G^{2g}$$

defines also a function on $\mathcal{M}_g^0$. Denoting it also by W_w and considering the loop ℓ_w obtained by the concatenation $w(\ell_1, \ell_1^{-1}, \dots, \ell_{2g}, \ell_{2g}^{-1})$, the last statement can be reformulated as

$$\lim_{T \downarrow 0} \mathbb{E}_{\mathrm{YM}_{\Sigma_T}}[W_{\ell_w}] = \int_{\mathcal{M}_g^0} W_w d\mu_{ABG,g}.$$

Given the surface group

$$\Gamma_g = \langle a_1, b_1, \dots, a_g, b_g | [a_1, b_1] \dots [a_g, b_g] \rangle,$$

consider the equivalence relation $\sim$ on the set of words with $2g$ letters and their inverses, such that $w \sim w'$ iff $w(a_1, \dots, b_g)$ and $w'(a_1, \dots, b_g)$ are equal in Γ_g. Thanks to the defining relation of $\mathcal{M}_g$, for any word w, the function W_w depends only on the equivalence class of w. When $\gamma \in \Gamma_g$ is the evaluation of w in Γ_g, denote this function by W_γ. In [47], Magee obtained the following analog of asymptotic freeness of Haar unitary random matrices.

Theorem 1.8 ([47] Corollary 1.2). *Consider the group $G = \mathrm{SU}(N)$. For any $\gamma \in \Gamma_g$,*

$$\lim_{N \to \infty} \mathbb{E}_{\mu_{ABG,g}}[W_\gamma] = \begin{cases} 1 & \textit{if } \gamma = 1, \\ 0 & \textit{otherwise.} \end{cases}$$

Since for any word with evaluation $\gamma \in \Gamma_g$, it can be shown that $\gamma = 1$ if and only if the loop ℓ_w is contractible, the above statement can be understood as the $T = 0$ case of Conjecture 1.3, with a weaker convergence given in expectation instead of in probability. In [48], it is also shown that $\mathbb{E}_{\mu_{ABG,g}}[W_\ell]$ admits an asymptotic expansion in powers of $\frac{1}{N}$.

Let us discuss the main differences between the approach of [47] and ours:

- Although both approaches use the convergence of the partition function of the model, we use in [20] the Markov property of the Yang–Mills holonomy field in order to prove the convergence for simple loops. Then we use the Makeenko–Migdal equations to induce the convergence on a larger class of loops; the latter is actually not needed in the zero volume case.
- We only consider the limit of Wilson loops, whereas Magee obtains a $\frac{1}{N}$ expansion.
- We prove a convergence in probability whereas Magee gets a convergence in expectation.
- We also consider a larger family of matrix groups, whereas he only treats the unitary case.
- In the case $g = 1$, the Atiyah–Bott–Goldman measure is ill-defined; hence, Magee's paper cannot handle it, but we still find a result when $T > 0$, which gives a matrix approximation of an interpolation between classical and free convolution of Haar unitaries.

1.3. Noncommutative distribution and master field on the torus: an interpolation between free and classical convolution

We discuss here the noncommutative distribution associated to the master field on the torus, leading to Corollary 1.12 below, obtained by specialising Theorem 1.2 to projection of loops restrained to the lattice $\sqrt{T}.\mathbb{Z}^2$.

1.3.1. Noncommutative probability and free independence

Let us give an extremely brief account of these notions. We refer to [68, 52] for more details. A *noncommutative probability space*[13] is the data of a tuple $(\mathcal{A}, *, 1, \tau)$, where $(\mathcal{A}, *, 1)$ is a unital $*$-algebra over $\mathbb{C}$, and τ is a positive, tracial state – that is, a linear map $\tau : \mathcal{A} \to \mathbb{C}$ with

$$\tau(aa^*) \geq 0 \text{ and } \tau(ab) = \tau(ba),\ \forall a, b \in \mathcal{A},$$

with furthermore $\tau(1) = 1$ and $\tau(a^*) = \overline{\tau(a)},\ \forall a \in \mathcal{A}$. We shall often leave as implicit the choice of unit and $*$, and denote a noncommutative probability space simply as a pair $(\mathcal{A}, \tau)$.

Example 1.9. For $N \geq 1$, the tuple $(M_N(\mathbb{C}), *, \mathrm{Id}_N, \mathrm{tr})$, where $\mathrm{tr} = \frac{1}{N}\mathrm{Tr}$, gives such a space. Consider the group $\mathrm{U}(N)$ of unitary complex matrices of size N and a group Γ with unit element 1. Let $(\mathbb{C}[\Gamma], *)$ be the group algebra of Γ endowed with the skew-linear idempotent defined by $\gamma^* = \gamma^{-1},\ \forall \gamma \in \Gamma$. Then, whenever $\rho : \Gamma \to \mathrm{U}_N(\mathbb{C})$ is a unitary representation of Γ, setting $\tau_\rho = \mathrm{tr} \circ \rho$, the tuple $(\mathbb{C}[\Gamma], *, 1, \tau_\rho)$ is a noncommutative probability space.

Let $(\mathcal{A}_1, \mathcal{A}_2)$ be unital subalgebras of a noncommutative probability space $\mathcal{A}_1$.

- They are *classically independent* if $\forall a_1, \ldots, a_n \in \mathcal{A}_1, b_1, \ldots, b_n \in \mathcal{A}_2$,

$$\tau(a_1 b_1 a_2 \ldots a_n b_n) = \tau(a_1 \ldots a_n)\tau(b_1 \ldots b_n).$$

- They are *freely independent* if for any $n \in \mathbb{N}$, for any $\{i_1, \ldots, i_n\} \in \{1, 2\}^n$ such that $i_1 \neq i_2, \ldots, i_{n-1} \neq i_n$ and for any $a_k \in \mathcal{A}_{i_k}$,

$$\tau(a_k) = 0,\ \forall 1 \leq k \leq n \Longrightarrow \tau(a_1 \cdots a_n) = 0.$$

These definitions can be generalised to any number of subalgebras, and a family of elements $(a_i)_{i \in I}$ of a noncommutative probability space $(\mathcal{A}, \tau)$ is said to be independent (resp. free) if the family $(\mathcal{A}_i)_{i \in I}$ is independent (resp. free), where for all $i \in I$, $\mathcal{A}_i$ is the subalgebra generated by a_i and a_i^*. We shall then say that $(a_i)_{i \in I}$ are resp. independent and free under τ.

When I is an arbitrary set, let us denote by $\mathbb{C}\langle X_i, X_i^*, i \in I\rangle$ the unital $*$-algebra of noncommutative polynomials in the variables $X_i, X_i^*, \in I$, with $*$ mapping X_i to X_i^* for all $i \in I$. When $(\mathcal{A}, *, 1, \tau)$ is a

[13] Sometimes denoted NCPS.

https://doi.org/10.1017/fms.2024.152 Published online by Cambridge University Press

noncommutative probability space and $\mathbf{a} = (a_i)_{i \in I}$ is a family of elements of $\mathcal{A}$, its *noncommutative distribution* is the positive, tracial, state on $\mathbb{C}\langle X_i, X_i^*, i \in I\rangle$ given by

$$\tau_{\mathbf{a}}(P) = \tau(P(a_i, i \in I)), \ \forall P \in \mathbb{C}\langle X_i, X_i^*, i \in I\rangle,$$

where $P(a_i, i \in I) \in \mathcal{A}$ denotes the evaluation of P replacing X_i and X_i^* by a_i and a_i^*. Likewise, when $\mathcal{A}$ and $\mathcal{B}$ are subalgebras of a same noncommutative probability space $(\mathcal{C}, \tau)$, we call the state $\tau_{\langle \mathcal{A}, \mathcal{B}\rangle}$ on $\mathbb{C}\langle X_a, Y_b, a \in \mathcal{A}, b \in \mathcal{B}\rangle$ given by

$$\tau_{\langle \mathcal{A}, \mathcal{B}\rangle}(P(X_a, Y_b; a \in \mathcal{A}, b \in \mathcal{B})) = \tau(P(a, b; a \in \mathcal{A}, b \in \mathcal{B})),$$

the joint distribution of $(\mathcal{A}, \mathcal{B})$ in $(\mathcal{C}, \tau)$.

When a, b are two elements of noncommutative probability spaces with respective noncommutative distribution τ_a and τ_b, there are unique states $\tau_a \star \tau_b$ and $\tau_a *_c \tau_b$ on $\mathbb{C}\langle X, Y, X^*, Y^*\rangle$ such that $\tau_X = \tau_a$ and $\tau_Y = \tau_b$ both under and $\tau_a *_c \tau_b$ and $\tau_a \star \tau_b$, while the joint distribution (X, Y) under $\tau_a \star \tau_b$ and $\tau_a *_c \tau_b$, are respectively freely and classically independent. The states $\tau_a \star \tau_b$ and $\tau_a *_c \tau_b$ are resp. called the *free* and the classical *convolution* of τ_a and τ_b. We define likewise the free and classical convolution of two states on $\tau_{\mathcal{A}}, \tau_{\mathcal{B}}$ of NCPS $(\mathcal{A}, \tau_{\mathcal{A}}), (\mathcal{B}, \tau_{\mathcal{B}})$ as states $\tau_{\mathcal{A}} \star \tau_{\mathcal{B}}$ and $\tau_{\mathcal{A}} *_c \tau_{\mathcal{B}}$ on $\mathbb{C}\langle X_a, Y_b, a \in \mathcal{A}, b \in \mathcal{B}\rangle$.

Let us recall the following result of *asymptotic freeness* due to Voiculescu [66], and for the considered group series by [17], see also [43, Sect. I-3].

Theorem 1.10 [66, 17, 43]. *Let A and B be two deterministic matrices of size N with respective noncommutative distribution satisfying for all fixed $P \in \mathbb{C}\langle X, X^*\rangle$,*

$$\tau_A(P) \to \tau_a(P), \tau_B(P) \to \tau_b(P), \ \text{ as } N \to \infty,$$

for some state τ_a, τ_b on $\mathbb{C}\langle X, X^*\rangle$. Consider U and V two independent Haar unitary matrices on a group G_N and $\rho_N : \mathbb{C}[\mathbb{F}_2] \to G_N$ the associated unitary representation of the free group of rank 2.

Then for any $\gamma \in \mathbb{F}_2$ and $P \in \mathbb{C}\langle X, Y, X^*, Y^*\rangle$, the following limit holds in probability as $N \to \infty$,

$$\tau_{\rho_N}(\gamma) \to \begin{cases} 1 & \text{if } \gamma = 1, \\ 0 & \text{if } \gamma \in \mathbb{F}_2 \setminus \{1\} \end{cases} \tag{10}$$

and

$$\tau_{A, UBU^*}(P) \to \tau_a \star \tau_b(P). \tag{11}$$

On the one hand, the first convergence (10) can be proved to be a special case of (11) when A and B are themselves independent Haar unitary random variables. On the other hand, when A and B are unitary or Hermitian with uniformly bounded spectrum, (11) can be deduced from (10) by functional calculus.

One of the motivations of the current article was to understand an analog of (10) when (U, V) are sampled according to a different law with correlation, as discussed in Section 1.3.3.

1.3.2. Free unitary Brownian motion and t-freeness

We refer here to [6, 67, 5] for more details. Consider a noncommutative probability space $(\mathcal{A}, \tau, *, 1)$. An element $u \in \mathcal{A}$ is called unitary when $uu^* = u^*u = 1$. It is Haar unitary if for any integer $n > 0$, $\tau(u^n) = \tau((u^*)^n) = 0$. The *free unitary Brownian motion* on a $*$-probability space $(\mathcal{A}, \tau, *, 1)$ is a family $(u_t)_{t \geq 0}$ of unitary elements of $\mathcal{A}$ such that the increments $u_{t_1}u_0^*, \ldots, u_{t_n}u_{t_{n-1}}^*$ are free for all $0 \leq t_1 \leq \cdots \leq t_n$, and for any $k \in \mathbb{Z}^*$ and $0 < s < t$,

$$\tau((u_t u_s^*)^k) = \tau(u_{t-s}^k),$$

while $\tau(u_t^k) = \nu_t(|k|)$ is C^1 with for all $m \geq 0$,

$$\frac{d}{dt}\nu_t(m) = -\frac{m}{2}\nu_t(m) - \frac{m}{2}\sum_{l=1}^{m}\nu_t(l)\nu_t(m-l),\ \forall t \geq 0,\ \nu_0(m) = 1. \tag{12}$$

Let us set $\nu_t = \tau_{u_t}$. It follows from the above expression that as t tends respectively to 0 and $+\infty$, the distribution ν_t converges pointwise to the one of, respectively, 1 and a Haar unitary. In view of (11), it is also natural to introduce the following deformation of free convolution.

Theorem 1.11 [67]. *Let $(\mathcal{A}, \tau_{\mathcal{A}})$ and $(\mathcal{B}, \tau_{\mathcal{B}})$ be two noncommutative probability spaces, and $t > 0$ be a fixed real number. Then there exists a noncommutative probability space $(\mathcal{C}^{(t)}, \tau_{\mathcal{C}^{(t)}})$ such that*

1. *$\mathcal{A}$ and $\mathcal{B}$ can be identified with two independent subalgebras of $(\mathcal{C}^{(t)}, \tau_{\mathcal{C}^{(t)}})$ with*

$$\tau_{\mathcal{C}^{(t)}}(a) = \tau_{\mathcal{A}}(a) \text{ and } \tau_{\mathcal{C}^{(t)}}(b) = \tau_{\mathcal{B}}(b),\ \forall (a,b) \in \mathcal{A}\times\mathcal{B}.$$

2. *There is a unitary element $u_t \in \mathcal{C}^{(t)}$ free with the subalgebra of $\mathcal{C}^{(t)}$ generated by $\mathcal{A}$ and $\mathcal{B}$, such that u_t has distribution ν_t.*

The t-free convolution product of $\tau_{\mathcal{A}}$ and $\tau_{\mathcal{B}}$ is then the joint distribution $\tau_{\mathcal{A}} \star_t \tau_{\mathcal{B}}$ of $(\mathcal{A}, u_t\mathcal{B}u_t^*)$ in the noncommutative probability space $(\mathcal{C}^{(t)}, \tau_{\mathcal{C}^{(t)}})$. It does not depend on the choice of $(\mathcal{C}^{(t)}, \tau_{\mathcal{C}^{(t)}})$ satisfying 1) and 2).

The above construction was introduced more generally[14] by Voiculescu [67] in his study of free entropy and free Fisher information via the liberation process. For any $t > 0$, two subalgebras $\mathcal{A}$ and $\mathcal{B}$ of a same noncommutative probability space $(\mathcal{C}, \tau)$ with respective distribution $\tau_{\mathcal{A}}$ and $\tau_{\mathcal{B}}$ are said to be *t-free* if their joint distribution under τ is given by $\tau_{\mathcal{A}} \star_t \tau_{\mathcal{B}}$. It can be shown ([67, 5]) that the following limits hold pointwise:

$$\lim_{t\downarrow 0} \tau_{\mathcal{A}} \star_t \tau_{\mathcal{B}} = \tau_{\mathcal{A}} *_c \tau_{\mathcal{B}} \text{ and } \lim_{t\to+\infty} \tau_{\mathcal{A}} \star_t \tau_{\mathcal{B}} = \tau_{\mathcal{A}} \star \tau_{\mathcal{B}}.$$

1.3.3. A matrix approximation for another interpolation from classical to free convolution

Let us present an application of Theorem 1.2. Consider a heat kernel $(p_t)_{t>0}$ on a classical compact matrix Lie group G_N endowed with the metric defined by (1), and for any $T > 0$, define a probability measure setting[15]

$$d\mu_{N,T}(A,B) = Z_T^{-1} p_T([A,B])dAdB \tag{13}$$

on G_N^2, where $dAdB$ denotes the Haar measure on G_N^2 and Z_T is the partition function

$$Z_T = \int_{G_N^2} p_T([A,B])dAdB.$$

As the limits $\lim_{T\downarrow 0} p_T(U)dU = \delta_{\mathrm{Id}_N}$ and $\lim_{T\to\infty} p_T(U)dU = dU$ hold weakly, we can think about μ_T as a model of random matrices interpolating between commuting and noncommuting settings. In [20, Thm 2.15], we have proved that though A and B are not Haar distributed for N fixed, as $N \to \infty$, they converge individually to Haar unitaries. Moreover, we also saw that under $\mu_{N,T}$, $[A,B]$ converges in noncommutative distribution, with limit given by ν_T, a free unitary Brownian motion at time T. In view of (10), it is then natural to investigate the possible limit of the joint law, hoping for a nontrivial

[14] Not necessarily with the assumption of classical independence for the initial state.

[15] Consider three independent random variables (A, B, U_T) on G_N, where A, B are Haar distributed, and U_T is Brownian motion at time T, that is with distribution $p_T(U)dU$. It is can be shown using a suitable definition of conditioning that $\mu_{N,T}$ is the law of the first two marginals of the triple (A, B, U_T) conditioned to satisfy $[A, B] = U_T$.

https://doi.org/10.1017/fms.2024.152 Published online by Cambridge University Press

coupling of Haar unitaries. Note that analog models with potentials[16] have been investigated in [16]. A challenge appearing in the setting of [16] is that these general results are limited to weak coupling regimes.[17] A consequence of our work is that $\mu_{N,T}$ has a noncommutative limit for all $T > 0$, leading to an interpolation between independent and free Haar unitaries. Denote by τ_u the distribution of a Haar unitary.

Corollary 1.12. *For any* $T > 0$*, there is a state* Φ_T *on* $\mathcal{A} = \mathbb{C}\langle X, X^*, Y, Y^*\rangle$*, such that for any* $P \in \mathcal{A}$*, under* $\mu_{N,T}$*,*

$$\mathrm{tr}(P(A,B)) \to \Phi_T(P) \text{ in probability as } N \to \infty$$

with

$$\lim_{T\downarrow 0} \Phi_T(P) = \tau_u *_c \tau_u(P) \text{ and } \lim_{T\to+\infty} \Phi_T(P) = \tau_u \star \tau_u(P).$$

Besides, for all $T, t > 0$*,*

$$\Phi_T \neq \tau_u \star_t \tau_u, \tag{14}$$

while

$$\Phi_T((XYX^*Y^*)^n) = \nu_T(n) = \tau_u \star_{\frac{T}{4}} \tau_u((XYX^*Y^*)^n),\ \forall n \in \mathbb{Z}^*.$$

We prove in Section 6.2 the above corollary together with a few other properties of Φ_T. Let us mention that the interpolation provided by Corollary 1.12 is not the only possible interpolation, even if we exclude the t-free convolution; for instance, another interpolation was proposed in [50] using rank one Harish–Chandra–Itzykson–Zuber integrals. Let us also mention that there are variations of freeness for family of algebras which are partly commuting [53, 63].

1.4. Strategy of proof via Makeenko–Migdal deformations

An important property, formally inferred by integration by parts from (4) in [49] and rigorously proved in [43] based on the Driver–Sengupta formula, is a family of equations almost characterising the function Φ_Σ when Σ is the plane. Other proofs have been given in [18, 26]. The proofs of [26] were much shorter and local, and it was possible to adapt them to all compact surfaces [25]. See also [24] for a different approach based on the construction of the Yang–Mills measure via white noise and [55] for a proof based on the representation of Wilson loop expectations as surface sums.

These equations can be described informally as follows. Consider a smooth loop ℓ with a transverse intersection at a point v. Assume that $(\ell_\varepsilon)_\varepsilon$ is a deformation of ℓ in a neighborhood of v such that the areas of the four corners adjacent to v are modified as in Figure 1. Then the Makeenko–Migdal equation

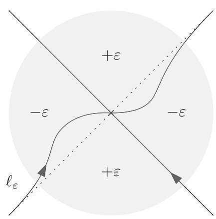

Figure 1. *Makeenko–Migdal deformation near an intersection point.*

[16]Though the class of potentials considered in [16] does not cover the heat kernel.
[17]Meaning that the parameter of the potential responsible for the non-independence of A and B needs to be small enough.

https://doi.org/10.1017/fms.2024.152 Published online by Cambridge University Press

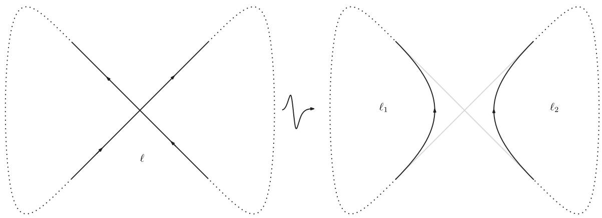

Figure 2. *Desingularisation at a simple intersection point.*

at v for a master field Φ_Σ is given by

$$\frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} \Phi_\Sigma(\ell_\varepsilon) = \Phi_\Sigma(\ell_1)\Phi_\Sigma(\ell_2), \tag{15}$$

where ℓ_1, ℓ_2 are two loops obtained by desingularising ℓ at v as on Figure 2. The works [43, 21, 33] can be understood as a study of existence and uniqueness of variants of equation (15). Our strategy here is to extend these results to all compact surfaces of genus $g \geq 1$.

A motivation of [43] for proving these relations was to compute explicitly the planar master field by induction on the number of intersections and to characterise it through differential equations. It was realised there that for the plane, there is no uniqueness for the Makeenko–Migdal equations alone, but there is if they are completed by an additional family of equations.[18] In [21, 33], the authors are interested in a perturbation of (15) arising from finite N analogs of (15) in view of proving the convergence of Wilson loops. The same lack of uniqueness occurs but is dealt with differently, adding in some sense boundary conditions, specifying the value of the master field[19] for simple loops. With this boundary condition, both [21, 33] are able to deduce the convergence of Wilson loops[20] on the sphere by induction on the number of intersection points. To complete the proof of Wilson loops convergence, it is then necessary to prove the convergence for boundary conditions via other means: this was done in [21] using a representation through a discrete[21] β-ensemble.

In [33], the author applied the same argument on all compact surfaces with a boundary condition given by simple loops within a disc and a uniquess or convergence result for loops within a disc. See the introduction of [20] for a more detailed discussion. In [20], we were able, using an independent argument, to prove the same result but without any boundary condition and making a relation with the planar master field.

Theorem 1.13 ([20], Theorem 2.16, 2.17). *Let ℓ be a loop in a compact connected orientable Riemann surface Σ of genus $g \geq 1$ with area measure* vol.

1. *If ℓ is topologically trivial and included in a disc U such that* $\mathrm{vol}(U) < \mathrm{vol}(\Sigma)$*, then as N tends to infinity, under* YM_Σ*,*

$$W_\ell \to \tilde{\Phi}(\psi \circ \ell) \text{ in probability,}$$

where $\tilde{\Phi}$ denotes the master field in the planar disc $\psi(U)$ where $\psi : U \to \psi(U) \subset \mathbb{R}^2$ is an area-preserving diffeomorphism.

2. *If ℓ is simple and noncontractible, then for any $n \in \mathbb{Z}^*$, as N tends to infinity,*

$$W_{\ell^n} \to 0 \text{ in probability}.$$

[18] Associated to each face adjacent to an infinite face.
[19] Or the convergence of Wilson loops
[20] This argument is valid for loops with finitely many transverse intersections. An additional step which is not considered in [33] is to extend it to loops with finite length.
[21] As suggested in [33], another route here could be to relate Wilson loops for simple loops on the sphere to the Dyson Brownian bridge on the unit circle, which has been studied recently at another scale in [45].

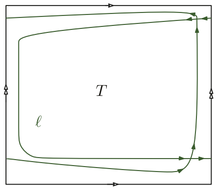

Figure 3. *In this example, it is impossible to change the area around any intersection point, respecting the constraint of Makeenko–Migdal given in Figure 1, without raising the number of intersection points.*

A first remark is that evaluating the planar master field at lift of contractible loops to the universal cover of Σ, as in Conjecture 1.3, gives a solution to Makeenko–Migdal equations. Our main focus will therefore be to study uniqueness of the Makeenko–Migdal equations or its deformation arising for finite N.

The general strategy of this article is to use Theorem 1.13 as a boundary condition to prove Proposition 1.5 and Theorem 1.6. For the torus, any nontrivial closed geodesic is either simple or the iteration of a simple closed loop; Proposition 1.5 together with Theorem 1.6 yield Theorem 1.2. For surfaces of genus $g \geq 2$, the results of [20] do not include all the loops in the assumption of Theorem 1.6, and there are then loops whose homotopy class does not include any simple loop, or any loop obtained by iterating a simple loop [7] (moreover, most geodesics have intersection points).

Let us now discuss how this strategy is implemented here. When applying the argument of [21] or [33], it is difficult to prove a result better than Theorem 1.13, which, given point 1 of Theorem 1.13, makes the use of Makeenko–Migdal equations pointless – a first obstacle being, for instance, a loop like in Figure 3, where it does not seem possible to apply Makeenko–Migdal equations at any vertex to deform the loop into a simpler loop.

To improve on [33], a first step is to characterise for surfaces of genus $g \geq 1$, the allowed deformations in Makeenko–Migdal equations. Viewing the evaluation at a regular loop of the master field as a function of faces area, we wonder along which deformation of loops the derivative of the master field is a linear combination of area derivatives such as the one involved in the left-hand side of (15). This was understood first in the plane by [43]. This is achieved here for surfaces in Section 2.2 with the following conclusions:

- When a loop has nonzero homology, then any reasonable deformation is allowed;
- When a loop has zero homology, then it is possible to define the winding number of the loop with respect to each face, and that a deformation is allowed if and only if it preserves the algebraic area, which is the area of each face weighted by the winding number of the loop with respect to it.

This observation allows to consider the simpler case of loops with nonzero homology separately. In this case, it is possible to argue as follows by induction, showing at each step that the derivative along a suitable deformation is bounded by induction assumption. First, considering the lift of a loop with nonzero homology to the universal cover, by induction on the number of intersections, it is possible to reduce the problem to loops with nonzero homology such that each strand of the lift going through a fundamental region has[22] no intersection point. Then Proposition 1.5 can be proved by induction on the number of fundamental domains visited. A key remark in this case is that at each intersection point, the two loops obtained by desingularisation have both nonzero homology and visit strictly less fundamental domains. This programme is carried out in section 4.1.

A second step is to overcome the difficulty met in Figure 3. This loop has vanishing homology. The cause of the obstruction becomes clearer thanks to the first step: it is not possible to decrease the area of the central face as it is a strict maximum of the winding number function. An idea is to deform the loop in a face that we want to 'inflate' so that the algebraic area remains preserved, as suggested on the following figure.

[22]We shall call below these loops proper loops.

https://doi.org/10.1017/fms.2024.152 Published online by Cambridge University Press

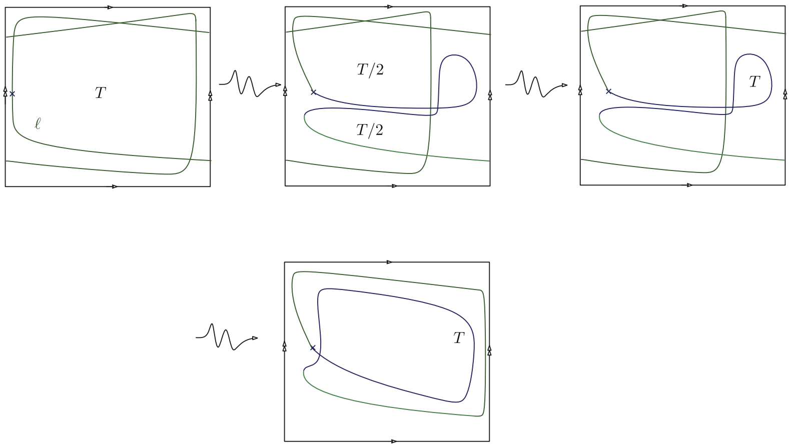

Figure 4. *Discrete homotopy towards a loop included in a disc preserving the algebraic area. Faces are labelled by their area. Faces without label have area* 0*.*

An apparent issue with this argument is that the number of intersections of the loops involved in the different steps may increase, preventing a direct induction on the number of intersections as in [43, 21, 33]. To solve this issue, we consider a family of 'marked' loops, consisting of two paths whose concatenation ℓ is a loop, where the second path is generic, while the first one has a specific form.[23] In particular, we require that a loop obtained by desingularisation at an intersection point of the first part is either in a fundamental domain or is the contraction of the initial loop ℓ along some faces bounded by the perturbed part. Because of the nested part of the perturbation part, it becomes possible[24] to argue by induction determining a complexity function on marked loops adapted[25] to the boundary conditions considered. The choice of complexity is done in Section 2.4; the induction is then performed in Theorem 4.5.

Lastly, it remains to extend our convergence result to a wider family of loops. This is first done using the property of continuity and compatibility on closed simplices of areas for loops with finitely many transverse intersections.[26] Then a more general argument introduced in [9, 21], building on the construction of [42], allows to consider all loops with finite length.[27]

2. Homology and homotopy on embedded graphs

2.1. Four equivalence relations on paths and loops on maps

We recall briefly here standard notions and define some notations of topological discretisation of a surface.

Definition 2.1. A *graph* $\mathbb{G} = (V, E, I)$ is a triple consisting of two sets V and E and an incidence relation $I \subset V \times E$ such that for any $e \in E$, the cardinal of $\{v \in V : (v, e) \in I\}$ is 1 or 2. The elements of V (resp. E) are called *vertices* (resp. *edges*).

This definition might seem very abstract at first sight, but it is actually simple: it merely says that a graph is made of edges and vertices, and that each edge is incident to either 1 vertex (the edge is then called a loop) or 2 vertices. Let $\mathbb{G} = (V, E, I)$ be a graph, and $e_1, e_2 \in E$ be two distinct edges.

[23] We call this part *nested*; see Section 2.5.
[24] See p. 52 where this idea is illustrated to prove the uniqueness of Makeenko–Migdal equations for the example of Figure 4.
[25] We believe there is a lot of flexibility here in the argument. We choose here a combinatorial approach, but it would be interesting to use instead a continuous functional on loops.
[26] See Lemma 3.4.
[27] This second step is not needed to consider projection of loops on a lattice.

https://doi.org/10.1017/fms.2024.152 Published online by Cambridge University Press

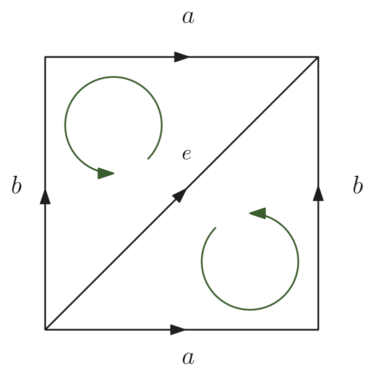

Figure 5. *A map embedded in the torus, with $|V| = 1$, $|E_+| = 3$ and $|F_+| = 2$. The edges named a are glued together, and same for the edges named b. There are two positively oriented faces, with respective boundaries $ea^{-1}b$ and $ab^{-1}e^{-1}$; the orientations are represented by the counterclockwise green arrows. Euler's formula is indeed satisfied.*

1. If there is $v \in V$ such that $(v, e_1) \in I$ and $(v, e_2) \in I$, then e_1 and e_2 are called *adjacent*.
2. If there are $v_1, v_2 \in V$ such that $(v_i, e_j) \in I$ for all $1 \leq i, j \leq 2$, then e_1 and e_2 form a *double edge*.

More generally, if n edges share the same incidence vertices, they form a multiple edge, and $\mathbb{G}$ is called a *multigraph*. A *topological map M* on a surface Σ is a finite multigraph $\mathbb{G} = (V, E, I)$ together with an embedding $\theta : \mathbb{G} \to \Sigma$ such that

- the images of two distinct vertices $v_1, v_2 \in V$ by θ are distinct points of Σ,
- for any edge $e \in E$, there is an edge $e^{-1} \in E$ such that $\theta_{e^{-1}} = \theta_e^{-1}$,
- the images of edges $e \in E$ are continuous curves $\theta_e : [0, 1] \to \Sigma$ with endpoints $\underline{e} = \theta_e(0)$ and $\overline{e} = \theta_e(1)$ such that θ_e and $\theta_{e'}$ can only intersect at their endpoints (unless $e' \in \{e, e^{-1}\}$),
- the complement of the skeleton $\mathrm{Sk}(\mathbb{G}) = \bigcup_{e \in E} \theta_e$ of $\mathbb{G}$ in Σ is split in one or several connected components that are all homeomorphic to discs and represent the faces of the map.

An *orientation* of the map is the choice of a subset $E_+ \subset E$ such that for any $e \in E$, $|\{e, e^{-1}\} \cap E_+| = 1$. The orientation of Σ also induces an orientation of the faces as follows: a face f is *positively oriented* if its boundary ∂f is represented by $e_1 \cdots e_n$, where $e_1, \ldots, e_n$ are the edges constituting ∂f in positive order. It is *negatively oriented* if its boundary is represented by $e_n^{-1} \cdots e_1^{-1}$. We denote by F (resp. F_+) the set of all faces with both orientations (resp. the positively oriented faces), and for any $f \in F_+$, we denote by $f^{-1} \in F$ the same face with reverse orientation.

Remark 2.2. With our conventions, each unoriented edge and unoriented face is counted twice; therefore, Euler's formula reads

$$|V| - \tfrac{1}{2}|E| + \tfrac{1}{2}|F| = |V| - |E_+| + |F_+| = 2 - 2g$$

if $\mathbb{G}$ is embedded in a surface of genus g. See Figure 5 for an illustration.

From now on, we will denote by $\mathbb{G} = (V, E, F)$ a topological map, and θ and Σ will be implied. Given $\mathbb{G} = (V, E, F)$, one can describe a CW-complex corresponding to the map: its vertices are 0-cells, its edges 1-cells and its faces 2-cells. Besides, the skeleton of the map is exactly the skeleton of the complex. We will describe the corresponding chain and cochain complexes in the next section, as well as their (co)homology. In this section, we rather focus on topological features of maps and algebraic properties of loops in a map.

A map with boundary is a map (V, E, F) together with a proper subset B of F such that the closures of 2-cells associated to any pair of distinct elements of B do not intersect. A *path* in $\mathbb{G}$[28] is either a single vertex or a finite string of edges $e_1 \dots e_n$ with $n \geq 1$ such that for all $k \in \{1, \dots, n-2\}$, $\underline{e}_{k+1} = \overline{e}_k$. We say it is constant in the first case and set $|\gamma| = 0$, while in the second, we denote by $\overline{\gamma} = \overline{e}_n$ and $\underline{\gamma} = \underline{e}_1$ its endpoint and starting point, and by $|\gamma| = n$ its *length*. A loop of $\mathbb{G}$ is a path γ with $\underline{\gamma} = \overline{\gamma}$. A loop ℓ is *based* at a vertex v when $\underline{\ell} = v$. We say it is *simple* when all vertices of ℓ occur only once ℓ, but $\underline{\ell}$ which occurs exactly twice. We write, respectively, $\mathrm{P}(\mathbb{G})$ and $\mathrm{L}(\mathbb{G})$ for the set of paths and loops of $\mathbb{G}$. The respective sets of paths starting from a vertex $v \in V$ are denoted by $\mathrm{P}_v(\mathbb{G})$ and $\mathrm{L}_v(\mathbb{G})$. Whenever α and β are two paths with $\overline{\alpha} = \underline{\beta}$, $\alpha\beta$ denotes their *concatenation*, while α^{-1} is the path run in reverse direction, with the convention that $\gamma_1\alpha\gamma_2 = \alpha$ when γ_1 and γ_2 are constant paths at $\underline{\alpha}$ and $\overline{\alpha}$. We say that β is a *subpath* of $\delta \in \mathrm{P}(\mathbb{G})$ and write $\beta \prec \delta$, if there are paths α and γ with $\delta = \alpha\beta\gamma$.

Homeomorphic loops: When two maps $\mathbb{G}, \mathbb{G}'$ yield homeomorphic CW complexes, they induce a bijection between cells of same dimension. Denote by $\Phi : E \to E'$ the associated bijection between edges of $\mathbb{G}$ and $\mathbb{G}'$ and the associated bijection between $\mathrm{P}(\mathbb{G})$ and $\mathrm{P}(\mathbb{G}')$. Consider two paths α and β within maps $\mathbb{G}_1$ and $\mathbb{G}_2$. We say that α and β are *homeomorphic* and write

$$\alpha \sim_\Sigma \beta$$

if there are maps $\mathbb{G}$ and $\mathbb{G}'$ finer than, respectively, $\mathbb{G}_1$ and $\mathbb{G}_2$ such that $\mathbb{G}$ and $\mathbb{G}'$ are homeomorphic, with induced bijection $\Phi : \mathrm{P}(\mathbb{G}) \to \mathrm{P}(\mathbb{G}')$ such that

$$\Phi(\alpha) = \beta.$$

Cyclically equivalent loops: We say that two loops are *cyclically equivalent* when one can be obtained from the other by cyclically permuting its edges. By convention, two constant loops are cyclically equivalent if they have equal base-point. This defines an equivalence relation $\sim_c$ on $\mathrm{L}(\mathbb{G})$. An element of the quotient $\mathrm{L}_c(\mathbb{G}) = \mathrm{L}(\mathbb{G})/{\sim_c}$ is called an *unrooted loop*.

Reduced loops: A path γ' is obtained by insertion of an edge in a path γ if $\gamma = \gamma_1\gamma_2$ and $\gamma' = \gamma_1 e e^{-1}\gamma_2$ with γ_1, γ_2 two subpaths of γ and e an edge such that $\overline{\gamma_1} = \underline{e} = \underline{\gamma_2}$. Vice versa, we say in this situation that γ is obtained by erasing of an edge of γ'. Two paths are said to have the same *reduction* if a finite sequence of erasures and insertions of edges transforms one into the other. This defines an equivalence relation $\sim_r$ on $\mathrm{P}(\mathbb{G})$, and we write $\mathrm{RP}(\mathbb{G}) = \mathrm{P}(\mathbb{G})/{\sim_r}$, $\mathrm{RP}_v(\mathbb{G}) = \mathrm{P}_v(\mathbb{G})/{\sim_r}$ and $\mathrm{RL}_v(\mathbb{G}) = \mathrm{L}_v(\mathbb{G})/{\sim_r}$ for any $v \in V$. The reduction of a path $\gamma \in \mathrm{P}(\mathbb{G})$ is the unique path of minimal length in its $\sim_r$-equivalence class. We say that two loops are $\sim_{r,c}$-equivalent if one can be obtained from the other by iterated cyclic permutations, insertions and erasures of edges.

Lassos and discrete homotopy: For any face $f \in F$, its boundary can be identified with an unrooted loop ∂f. When $r \in V$ is a vertex of ∂f, we write $\partial_r f$ for the loop in the $\sim_c$-class of ∂f with $\underline{\partial_r f} = r$. When F_* is a subset of F, a *F_*-lasso* is a loop of the form $\alpha\partial_r f\alpha^{-1}$, where f is an oriented face belonging up to orientation to F_* and $\alpha \in \mathrm{P}(\mathbb{G})$ is a path such that $r = \overline{\alpha}$ is a vertex of ∂f. When $\gamma \in \mathrm{P}(\mathbb{G})$, γ' is obtained by *lasso insertion* from γ if $\gamma = \gamma_1\gamma_2$ for some paths $\gamma_1, \gamma_2 \in \mathrm{P}(\mathbb{G})$ and $\gamma' = \gamma_1\ell\gamma_2$, where ℓ is a lasso with $\overline{\gamma}_1 = \underline{\ell} = \underline{\gamma}_2$. Conversely, γ' is said to be obtained from γ by *lasso erasure*. We say that two paths are *discrete homotopic* if there is a finite sequence of lassos or edge erasures and insertions transforming one into the other. This defines an equivalence relation $\sim_h$ on $\mathrm{P}(\mathbb{G})$ which is also well defined on $\mathrm{RP}(\mathbb{G})$. Moreover, two paths of $\mathbb{G}$ are discrete homotopic if and only[29] if their images in $\Sigma_{\mathbb{G}}$

[28]Here, $\mathbb{G}$ denotes a map, but the definition of a path can also be applied *mutatis mutandis* to any graph or CW complex. In particular, it will also make sense later for universal covers of maps.

[29]Let us sketch an argument for this equivalence. When F_* is a subset of faces, using reduction, two paths $\gamma, \gamma' \in \mathrm{P}(\mathbb{G})$ with same endpoints can be obtained by F_*-lasso or edge erasure or insertion if and only if $\gamma' = w\gamma$ with a word w in F_*-lassos based at $\underline{\gamma}$. Also, the images of γ and γ' are homotopic with fixed endpoints if and only if $\gamma' = w\gamma$, where $w \in \mathrm{L}_{\underline{\gamma}}(\mathbb{G})$ is a loop whose image in $\Sigma_{\mathbb{G}}$ is contractible. It is therefore enough to show that for any $v \in V$, a loop $\ell \in \mathrm{L}_v(\mathbb{G})$ has a contractible image in $\Sigma_{\mathbb{G}}$ if and only if it is a word in lassos based at v. In order to prove this latter fact, choose a neighbourhood deformation retract U of the embedding of $\mathbb{G}$ in $\Sigma_{\mathbb{G}}$. Applying inductively Van Kampen's theorem to the covering by open sets given by the union of U and the images of the faces of $\mathbb{G}$ in $\Sigma_{\mathbb{G}}$ implies that $\pi_1(\Sigma_{\mathbb{G}})$ is a quotient of $\pi_1(U)$ by a group generated by lassos.

https://doi.org/10.1017/fms.2024.152 Published online by Cambridge University Press

are homotopic with fixed endpoints. For any $v \in V$, we denote the quotient $\mathrm{P}_v(\mathbb{G})/\sim_h$ and $\mathrm{L}_v(\mathbb{G})/\sim_h$ by $\tilde{V}_v$ and $\pi_{1,v}(\mathbb{G})$. When $F_* \subset F$, we say that two paths of $\mathbb{G}$ are *F_*-homotopic* if there is a finite sequence of F_*-lassos or edge erasures and insertions transforming one into the other. This defines an equivalence relation on $\mathrm{P}(\mathbb{G})$ denoted by $\sim_{F_*}$. When K is a closed, compact, contractible subset of $\Sigma_{\mathbb{G}}$ given by the closure of the union of images of F_*, for any pair of paths $\gamma_1, \gamma_2 \in \mathrm{P}(\mathbb{G})$ whose image in $\Sigma_{\mathbb{G}}$ is included in K and with same endpoints,[30] $\gamma_1 \sim_{F_*} \gamma_2$.

The group of reduced loops and the fundamental group: For any vertex $v \in V$, we define a group by endowing $\mathrm{RL}_v(\mathbb{G})$ with the multiplication given by concatenation and the inverse map given by reversing the orientation of loops. The group $\pi_{1,v}(\mathbb{G})$ is the quotient of $\mathrm{RL}_v(\mathbb{G})$ by the normal subgroup generated by lassos based at v. Since two loops of $\mathbb{G}$ are discrete homotopic if and only if their images in $\Sigma_{\mathbb{G}}$ are homotopic, the group $\pi_{1,v}(\mathbb{G})$ is isomorphic to the fundamental group of the surface $\Sigma_{\mathbb{G}}$. For any group G, let us write $[a,b] = aba^{-1}b^{-1}$, $\forall a, b \in G$. Then $\pi_{1,v}(\mathbb{G})$ is isomorphic to the surface group

$$\Gamma_g = \langle x_1, y_1, \ldots, x_g, y_g | [x_1, y_1] \ldots [x_g, y_g] \rangle.$$

We will also consider, for $r \geq 1$, the group

$$\Gamma_{r,g} = \langle z_1, \ldots, z_r, x_1, y_1, \ldots, x_g, y_g \mid z_1 \ldots z_r = [x_1, y_1] \ldots [x_g, y_g] \rangle.$$

Lemma 2.3 [42]. *For any map $\mathbb{G}$, the following assertions hold:*

1. *The group* $\mathrm{RL}_v(\mathbb{G})$ *is free of rank* $|E_+| - |V| + 1 = |F_+| + 2g - 1$.
2. *Assume that $g \geq 0$ and $|F_+| = r$. For any $v \in V$, there are lassos $(\ell_i, 1 \leq i \leq r)$ based at v, with faces in bijection with F_+, and loops $a_1, b_1, \ldots, a_g, b_g \in \mathrm{L}_v(\mathbb{G})$ such that the application*

$$\Theta : \Gamma_{r,g} \to \mathrm{RL}_v(\mathbb{G})$$

that maps z_i to ℓ_i for all $1 \leq i \leq r$, x_m (resp. y_m) to a_m (resp. b_m) for all $1 \leq m \leq g$ is an isomorphism.[31] *The diagram*

$$\begin{array}{ccccccc} 1 & \to & \Gamma_{r,g} & \to & \mathrm{RL}_v(\mathbb{G}) & \to & 1 \\ & & \downarrow & & \downarrow & & \\ 1 & \to & \Gamma_g & \to & \pi_{1,v}(\mathbb{G}) & \to & 1 \end{array}$$

is then commutative, where the left downwards morphism is the group morphism mapping z to $1 \in \Gamma_{r,g}$ for any $z \in \{z_1, \ldots, z_r\}$, and $t \in \Gamma_{r,g}$ to $t \in \Gamma_g$ for any $t \in \{x_1, y_1, \ldots, x_g, y_g\}$.

Refining maps: When $\mathbb{G}' = (V', E', F')$ and $\mathbb{G} = (V, E, F)$ are two maps, $\mathbb{G}'$ is *finer* than $\mathbb{G}$ if (V, E) is a subgraph of (V', E') and $\Sigma_{\mathbb{G}'} = \Sigma_{\mathbb{G}}$, so that we can identify V and E with subsets of, respectively, V' and $\mathrm{P}(\mathbb{G})$, while any face of $\mathbb{G}$ is the union of faces of $\mathbb{G}'$. Conversely, we say that $\mathbb{G}$ is *coarser* then $\mathbb{G}'$.

Dual map: When $\mathbb{G} = (V, E, F)$ is a map of genus g with surface $\Sigma_{\mathbb{G}}$, we define its *dual map* as follows: we put a vertex f^* inside each face $f \in F$, and for each edge $e \in E$ that separates two faces f_1 and f_2, we draw a new edge e^* that intersects it in its midpoint and connects the vertices f_1^* and f_2^*. There is a bijection $V^* \simeq F$, $E^* \simeq E$ and $F^* \simeq V$, and a dual edge inherits the orientation from the edges it crosses as follows: if e^* crosses $e \in E_+$ from the right,[32] then $e^* \in E_+^*$. See Figure 6 for an illustration. In particular, we see that if $e = (\underline{e}, \overline{e})$ is an edge and $e^* = (\underline{e^*}, \overline{e^*})$ is the dual edge, then we have the following facts:

$$e^* \in \partial \underline{e},\ (e_*^{-1}) \in \partial \overline{e},\ e \in \partial \overline{e^*},\ e^{-1} \in \partial \underline{e^*}. \tag{16}$$

[30]This fact can be proven by an argument similar to the one used in the previous footnote, and its proof is left to the reader.
[31]Denoting here abusively the $\sim_r$ class of a loop by the same symbol as the loop.
[32]Formally, it means that the dual edge (f_1, f_2) with $f_1, f_2 \in F_+$ is in E_+^* if the edge $e \in E_+$ it crosses satisfies $e \in \partial f_2$ and $e^{-1} \in \partial f_1$.

https://doi.org/10.1017/fms.2024.152 Published online by Cambridge University Press

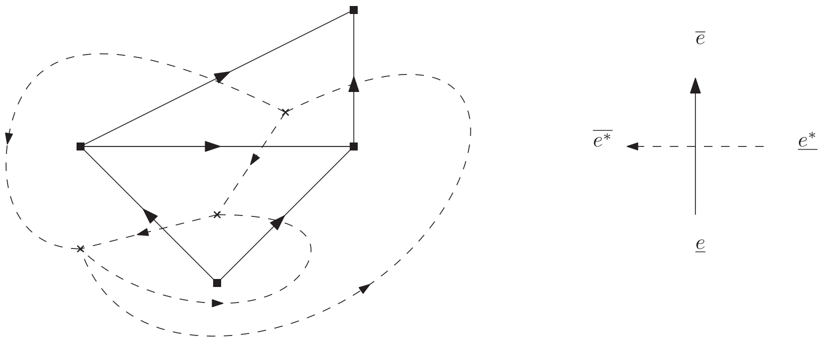

Figure 6. *On the left: a map embedded in the sphere (in plain lines), and its dual map (in dashed lines). On the right: the orientation convention of an edge and its dual. We have* $\underline{e}, \overline{e} \in V = F^*$ *and* $\underline{e^*}, \overline{e^*} \in F = V^*$.

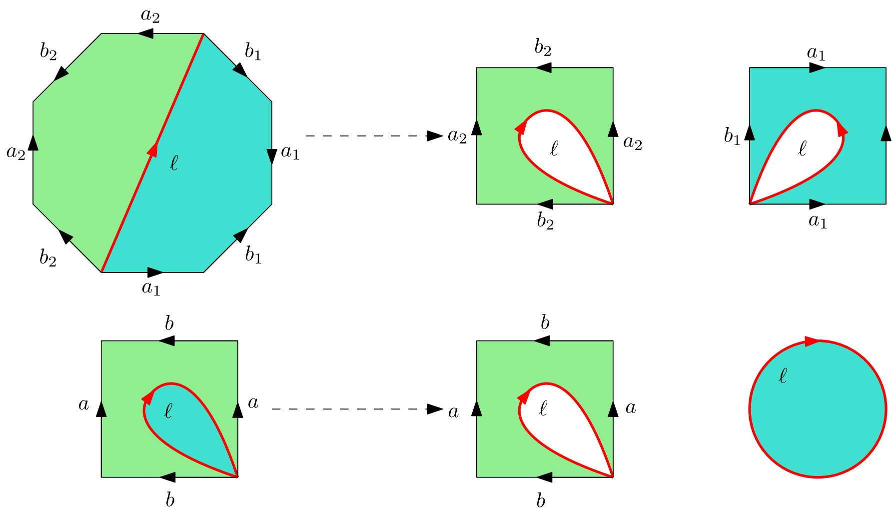

Figure 7. *Two examples of map cuts. In the first line, a genus 2 map is cut along* ℓ *into two genus 1 maps with boundary* ℓ*; hence, it is an essential cut. In the second line, a genus 1 map is cut into a genus 1 and a genus 0 maps, with boundary* ℓ*. It is therefore not essential.*

Cut of a map: When $\mathbb{G} = (V, E, F)$ is a map and ℓ is a simple loop of $\mathbb{G}$, with dual edges E^*_ℓ, we say that ℓ is separating if the graph $(F, E^* \setminus E^*_\ell)$ has exactly two connected components (F_1, E^*_1) and (F_2, E^*_2). Consider $i \in \{1, 2\}$. Denote by E_i the union of E_ℓ with the set of edges of $\mathbb{G}$ dual to E^*_i, and by V_i the vertices of $\mathbb{G}$ endpoints of edges in E_i. We then define a map with one boundary component by setting $\mathbb{G}_i = (V_i, E_i, F_i \sqcup \{f_{i,\infty}\})$, where $\{f_i, \infty\}$ is the label of a boundary face with boundary ℓ. We say that the pair of maps with boundary $(\mathbb{G}_1, \{f_{1,\infty}\}), (\mathbb{G}_2, \{f_{2,\infty}\})$ is the *cut* of $\mathbb{G}$ along ℓ. We say that the cut is *essential* if ℓ is not contractible. A cut is essential if and only if the maps $\mathbb{G}_1$ and $\mathbb{G}_2$ have genus larger or equal to 1. Both cases are illustrated in Figure 7 below. When a map is cut, the lemma 2.3 can be specified as follows.

Lemma 2.4. *Assume that* $(\mathbb{G}_1, \{f_{1,\infty}\}), (\mathbb{G}_2, \{f_{2,\infty}\})$ *is the cut of a map* $\mathbb{G} = (V, E, F)$ *of genus* $g \geq 0$*, along a simple loop* $\ell \in \mathrm{L}_v(\mathbb{G})$*. Denote by* g_1 *and* g_2 *the genus of* $\mathbb{G}_1$ *and* $\mathbb{G}_2$ *and by* r_1 *and* r_2 *their number of non-boundary faces, so that* $\mathbb{G}$ *has genus* $g = g_1 + g_2$ *and* $r = r_1 + r_2$ *faces. Then the following holds true.*

https://doi.org/10.1017/fms.2024.152 Published online by Cambridge University Press

1. *The group* $\mathrm{RL}_v(\mathbb{G})$ *is isomorphic to* $\mathrm{RL}_v(\mathbb{G}_1) * \mathrm{RL}_v(\mathbb{G}_2)$.
2. *There are lassos* $(\ell_i, 1 \leq i \leq r)$ *based at* v *and loops* $a_1, b_1, \ldots, a_g, b_g \in \mathrm{L}_v(\mathbb{G})$ *as in Lemma 2.4 with the additional property that* $\ell_1, \ldots, \ell_{r_1}, a_1, \ldots, b_{g_1} \in \mathrm{L}_v(\mathbb{G}_1)$ *and* $\ell_{r_1+1}, \ldots, \ell_r, a_{g_1+1}, \ldots, b_g \in \mathrm{L}_v(\mathbb{G}_2)$. *The group* $\mathrm{RL}_v(\mathbb{G}_1)$ *is then free over the basis* $\ell_1, \ldots, \ell_{r_1}, a_1, \ldots, b_{g_1}$.

2.2. Discrete homology, winding function and Makeenko–Migdal vectors

We recall here some elementary results about the homology of topological maps and discuss their relation to Makeenko–Migdal vectors introduced in [43, 21, 33]. It will lead us to a construction of the winding function, as well as a characterisation of the Makeenko–Migdal vectors, which, as we recall, encode the deformations that are allowed by Makeenko–Migdal equations. In the sequel, R will denote a ring that is either $\mathbb{R}$ or $\mathbb{Z}$, unless specified otherwise. We shall start with a general property of finitely generated free modules.

Proposition 2.5. *Let* A *be a finitely generated free* R*-module, and* $B = (e_1, \ldots, e_n)$ *be a free basis of* A.

1. *There exists a nondegenerate bilinear form* $\langle \cdot, \cdot \rangle$ *on* A *such that* B *is an orthonormal basis.*
2. *There is a canonical isomorphism* $A \cong \mathrm{Hom}(A, R)$ *expressed through the bilinear form* $\langle \cdot, \cdot \rangle$.

Proof. The first point is obvious: we set $\langle e_i, e_j \rangle = \delta_{ij}$, and we extend the form by bilinearity. For the second point, set

$$\Phi : \begin{cases} A \longrightarrow \mathrm{Hom}(A, R) \\ x \longmapsto (y \mapsto \langle x, y \rangle) \end{cases}$$

and notice that it is indeed an isomorphism. □

We can present the homology of a topological map from several equivalent ways, and we will present three of them.

Definition 2.6. Let $\mathbb{G} = (V, E, F)$ be a topological map. We define its associated *chain complex* by the sequence

$$0 \longrightarrow C_2(\mathbb{G}; R) \xrightarrow{\partial} C_1(\mathbb{G}; R) \xrightarrow{\partial} C_0(\mathbb{G}; R) \longrightarrow 0,$$

where $C_0(\mathbb{G}; R)$ (resp. $C_1(\mathbb{G}; R)$, $C_2(\mathbb{G}; R)$) is the free R-module generated[33] by V (resp. E, F). The boundary operator is defined by linear extension of the boundary operator in the underlying surface, defined by

$$\partial e = \overline{e} - \underline{e}, \ \forall e \in E,$$
$$\partial f = \sum_{e \in \partial f} e, \ \forall f \in F.$$

Let $\mathbb{G} = (V, E, F)$ be a topological map, and let $\mathbb{G}^* = (V^*, E^*, F^*)$ be its dual map. For any $v \in V = F^*$, we define its *boundary* ∂v as the cycle $e_1^* \cdots e_n^*$ of dual edges constituting the positively oriented boundary of the face v.

Definition 2.7. Let $\mathbb{G} = (V, E, F)$ be a topological map. Its associated *cochain complex* is defined by the sequence

$$0 \longleftarrow \Omega^2(\mathbb{G}, R) \xleftarrow{d} \Omega^1(\mathbb{G}, R) \xleftarrow{d} \Omega^0(\mathbb{G}, R) \longleftarrow 0,$$

[33]Remark that E and F define generating families but not free families. A free basis of $C_1(\mathbb{G}, R)$ (resp. $C_2(\mathbb{G}; R)$) is given by E_+ (resp. F_+).

https://doi.org/10.1017/fms.2024.152 Published online by Cambridge University Press

where $\Omega^k(\mathbb{G}, R) = \mathrm{Hom}(C_k(\mathbb{G}; R), R)$ for any $0 \le k \le 2$, and d is the dual of the boundary operator:

$$\begin{aligned} df(e) =& f(\partial e) = f(\overline{e}) - f(\underline{e}), \ \forall e \in E, \ \forall f \in \Omega^0(\mathbb{G}, R), \\ d\omega(f) =& \omega(\partial f) = \sum_{e \in \partial f} \omega(e), \ \forall f \in F, \ \forall \omega \in \Omega^1(\mathbb{G}, R). \end{aligned}$$

The elements of $\Omega^k(\mathbb{G}, R)$ are called R-valued *k-forms* on $\mathbb{G}$.

Proposition 2.8. *For any topological map $\mathbb{G}$, its cochain complex is isomorphic to the chain complex of the dual map $\mathbb{G}^*$.*

Proof. Let us first note that, as free modules, we have indeed canonical isomorphisms $\Phi : C_k(\mathbb{G}^*; R) \cong \Omega^{2-k}(\mathbb{G}, R)$ for $0 \le k \le 2$. They are explicitely given as follows:

$$\begin{aligned} f_v &: v' \in V \mapsto \delta_{vv'}, \ \forall v \in F_+^* \simeq V, \\ \omega_e &: e' \in E_+ \mapsto \delta_{ee'}, \ \forall e^* \in E_+^* \simeq E_+, \\ \mu_f &: f' \in F_+ \mapsto \delta_{ff'}, \ \forall f \in V^* \simeq F_+. \end{aligned}$$

Using (16) in conjunction with the definitions of ∂ and d, one can easily find that the diagram

$$\begin{array}{ccccccccc} 0 & \longrightarrow & C_2(\mathbb{G}^*; R) & \xrightarrow{\partial} & C_1(\mathbb{G}^*; R) & \xrightarrow{\partial} & C_0(\mathbb{G}^*; R) & \longrightarrow & 0 \\ & & \| & & \| & & \| & & \\ 0 & \longrightarrow & \Omega^0(\mathbb{G}, R) & \xrightarrow{d} & \Omega^1(\mathbb{G}, R) & \xrightarrow{d} & \Omega^2(\mathbb{G}, R) & \longrightarrow & 0 \end{array}$$

commutes, which proves the isomorphism. □

We shall denote by μ_* the constant 2-form defined by $\mu_*(f) = 1$ for all $f \in F_+$. Thanks to Proposition 2.5, there is for any $0 \le k \le 2$ a canonical isomorphism $\Phi : C_k(\mathbb{G}; R) \xrightarrow{\cong} \Omega^k(\mathbb{G}; R)$, represented by the applications $v \mapsto f_v$, $e \mapsto \omega_e$ and $f \mapsto \mu_f$ used in the proof of Proposition 2.8.

We define d^* as the adjoint of d on the cochain complex of $\mathbb{G}$, meaning that

$$\begin{aligned} d^*\omega &= \sum_{e \in E_+} \omega(e) f_{\partial e} = \sum_{e \in E_+} \sum_{v \in \partial e} \omega(e) f_v, \ \forall \omega \in \Omega^1(\mathbb{G}, R), \\ d^*\mu &= \sum_{f \in F_+} \mu(f) \omega_{\partial f} = \sum_{f \in F_+} \sum_{e \in \partial f} \mu(f) \omega_e, \ \forall \mu \in \Omega^2(\mathbb{G}, R). \end{aligned}$$

In particular, $d^* : \Omega^1(\mathbb{G}, R) \to \Omega^0(\mathbb{G}, R)$ is the *divergence operator* in R. Kenyon's terminology [38]. Let $e \in E_+$ be an oriented edge, and $e^* \in E_+^*$ be the dual edge (i.e., the faces $f, f' \in F_+$ such that $e \in \partial f$ and $e^{-1} \in \partial f$ satisfy $f' = \underline{e^*}$ and $f = \overline{e^*}$). Then we have, for any $\mu \in \Omega^2(\mathbb{G}, R) \cong C_0(\mathbb{G}^*; R)$,

$$d^*\mu(e) = \mu(f) - \mu(f') = \langle \mu, \partial e^* \rangle.$$

We obtain an isomorphism of chain complexes given by the following commutative diagram:

$$\begin{array}{ccccccccc} 0 & \longrightarrow & C_2(\mathbb{G}; R) & \xrightarrow{\partial} & C_1(\mathbb{G}; R) & \xrightarrow{\partial} & C_0(\mathbb{G}; R) & \longrightarrow & 0 \\ & & \| & & \| & & \| & & \\ 0 & \longrightarrow & \Omega^2(\mathbb{G}, R) & \xrightarrow{d^*} & \Omega^1(\mathbb{G}, R) & \xrightarrow{d^*} & \Omega^0(\mathbb{G}, R) & \longrightarrow & 0 \end{array}$$

Equipped with these chain complexes, we can then do some (basic) homology. Let us introduce a few notations, following, for instance, the conventions of [35]. We set

https://doi.org/10.1017/fms.2024.152 Published online by Cambridge University Press

- $\diamond_1 = \ker(d^* : \Omega^1(\mathbb{G}, R)) \simeq \ker(\partial : C_1(\mathbb{G}; R) \to C_2(\mathbb{G}; R)$ the module of cycles,
- $\bigstar_1^* = d^*(\Omega^2(\mathbb{G}, R)) \simeq \partial(C_2(\mathbb{G}; R))$ the module of boundaries,
- $\bigstar_1 = d(\Omega^0(\mathbb{G}, R)) \simeq \partial(C_2(\mathbb{G}^*; R))$ the module of coboundaries.

All these modules are equipped with the bilinear form $\langle \cdot, \cdot \rangle$ from Proposition 2.5, associated to a given basis.

Definition 2.9. Let $\mathbb{G}$ be a topological map. Its *first homology module* is defined as the R-module

$$H_1(\mathbb{G}; R) = \diamond_1 / \bigstar_1^*.$$

When ℓ is a loop of $\mathbb{G}$, its R-homology $[\ell]_R$ is the image of the element ω_ℓ in $H_1(\mathbb{G}; R)$. For any $n \geq 2$, its $\mathbb{Z}_n$-homology $[\ell]_{\mathbb{Z}_n}$ is the element $1 \otimes [\ell]_{\mathbb{Z}} \in H_1(\mathbb{G}; \mathbb{Z}_n) = \mathbb{Z}_n \otimes_{\mathbb{Z}} H_1(\mathbb{G}; \mathbb{Z})$.

Note that by the universal coefficient theorem for homology, the change of ring commutes with the homology

$$H_1(\mathbb{G}; R) = R \otimes_{\mathbb{Z}} H_1(\mathbb{G}; \mathbb{Z}),$$

even if we take $R = \mathbb{Z}_n$.

Proposition 2.10. *Denote by $\diamond_0$ the R-module spanned by ω_ℓ for all loops ℓ in $\mathbb{G}$.*

1. *We have the following equality of R-modules:*

$$\diamond_0 = \diamond_1 = \bigstar_1^\perp.$$

2. *We have the following direct sum decomposition into orthogonal subspaces:*

$$\Omega^1(\mathbb{G}, R) = \bigstar_1 \oplus \diamond_1. \tag{17}$$

Proof. Let us start by showing that $\diamond_1 = \bigstar_1^\perp$. If $\omega \in \diamond_1$, then for any $f \in \Omega^0(\mathbb{G}, R)$, we have

$$\langle \omega, df \rangle = \langle d^*\omega, f \rangle = 0$$

and $\omega \in \bigstar_1^\perp$. Conversely, remark that a free basis of $\bigstar_1$ is given by $(df_v, v \in V)$, so that for any $\omega \in \bigstar_1^\perp$, we have

$$\langle d^*\omega, f_v \rangle = \langle \omega, df_v \rangle = 0,$$

and $\omega \in \diamond_1$.

Now let us prove that $\diamond_0 = \diamond_1$. If $\ell = e_1 \cdots e_n$ is a loop in $\mathbb{G}$, then

$$d^*\omega_\ell = \sum_{i=1}^n \sum_{e \in E_+} \omega_{e_i}(e) f_{\partial e} = \sum_{i=1}^n f_{\partial e_i} = 0,$$

so that $\diamond_0 \subset \diamond_1$. Let $\omega \in \diamond_0^\perp$ be a 1-form. We define a 0-form $f_\omega \in \Omega^0(\mathbb{G}, R)$ by setting

$$f_\omega(v) = \sum_{i=1}^n \omega(e_i),$$

where $e_1 \cdots e_n$ is a path in $\mathbb{G}$ starting from a given reference vertex v_0 and ending at v. The fact that it does not depend on the path follows from the fact that $\omega \perp \omega_\ell$ for any loop ℓ. We see that for any $e \in E$, $df_\omega(e) = \omega(e)$; therefore, we have the inclusion $\diamond_0^\perp \subset \bigstar_1 = \diamond_1^\perp$. We finally get that $\diamond_0 = \diamond_1$. The direct sum decomposition follows from the standard decomposition of a module into a submodule and

https://doi.org/10.1017/fms.2024.152 Published online by Cambridge University Press

its orthogonal, provided that the bilinear form is not degenerate on this submodule, which is trivially the case here. □

Proposition 2.11. *The R-module*

$$\mathcal{H}_1 = (\bigstar_1^*)^\perp \cap \diamond_1$$

is isomorphic to $H_1(\mathbb{G}; R)$.

Proof. Recall that $\bigstar_1^* \subset \diamond_1$ thanks to the property of the chain complex. It follows from the direct sum decomposition

$$\diamond_1 = \bigstar_1^* \oplus \mathcal{H}_1$$

that for any $\omega \in \diamond_1$, there is a unique couple $(\omega_0, \mu) \in \mathcal{H}_1 \times \Omega^2(\mathbb{G}, R)$ such that $\omega = \omega_1 + d^*\mu$. It is then straightforward to check that the map $[\omega] \mapsto \omega_0$ is the isomorphism that we were looking for. □

The winding number of a planar loop $\ell = e_1 \cdots e_n$ around a point is an integer that counts how many times the loop cycles around the point; in particular, we see that in the case of a topological map, it defines a function $n_\ell \in \Omega^2(\mathbb{G}, \mathbb{Z})$ that counts how many times the loop cycles around each face. One can see that it is equivalent to require that $d^*n_\ell(e_i) = 1$ for any i such that $e_i \in E_+$, and -1 for any i such that $e_i^{-1} \in E_+$. It sums up as

$$d^*n_\ell = \omega_\ell.$$

Is it possible to get such a construction for compact orientable surfaces? The general answer is *not exactly*, because 'bad' things can happen when ℓ has a nontrivial homology, but it is still possible when $[\ell] = 0$, as stated by the following lemma.

Lemma 2.12. *Assume that* $\mathbb{G}$ *is embedded in an orientable surface of genus* g.

1. $H_1(\mathbb{G}; R)$ *is free of rank* $2g$, *and there are* $2g$ *simple loops* $(a_i, b_i)_{1 \le i \le g}$ *of* $\mathbb{G}$ *such that* $[a_1]_R, [b_1]_R, \ldots, [a_g]_R, [b_g]_R$ *is a free basis of* $H_1(\mathbb{G}; R)$. *Equivalently,* $\omega_{a_1}, \omega_{b_1}, \ldots, \omega_{a_g}, \omega_{b_g}$ *is a free basis of* $\mathcal{H}_1$.
2. *When* $g \ge 1$ *and* $v \in V$, $(\ell_i, 1 \le i \le r)$ *and* $a_1, b_1, \ldots, a_g, b_g \in \mathrm{L}_v(\mathbb{G})$ *are as in Lemma 2.3, the map*

$$\Gamma_g = \langle x_1, y_1, \ldots, x_g, y_g \mid [x_1, y_1] \ldots [x_g, y_g] \rangle \longrightarrow H_1(\mathbb{G}; \mathbb{Z})$$

that maps x_m *to* $[a_m]_\mathbb{Z}$ *and* y_m *to* $[b_m]_\mathbb{Z}$ *is well-defined, onto morphism, with kernel given by the commutator group* $[\Gamma_g, \Gamma_g]$.
3. *For any loop* ℓ *of* $\mathbb{G}$ *such that* $[\ell]_R = 0$, *there is a unique*[34] $n_\ell \in \Omega^2(\mathbb{G}, R)$ *such that*

$$\omega_\ell = d^*n_\ell.$$

We call the 2-form n_ℓ *the winding function of* ℓ, *and we shall identify it to an element of* $\{\mu_*\}^\perp$.

Proof. Points 1 and 2 are standard results, and their proof can be found in Chapter 2 of [64]. We shall prove the last point. Recall that

$$\diamond_1 = \bigstar_1^* \oplus \mathcal{H}_1,$$

and that for any loop ℓ, the 1-form ω_ℓ is in $\diamond_1$. Hence, there is a unique pair (h_ℓ, n_ℓ) with $h_\ell \in \mathcal{H}_1$ and $n_\ell \in \Omega^2(\mathbb{G}, R)$ such that $\omega_\ell = h_\ell + d^*n_\ell$. If $[\ell]_R = 0$, it means that $h_\ell = 0$ and $\omega_\ell = d^*n_\ell$, as expected. □

An example of winding number is depicted in Figure 8.

[34]Up to a constant.

https://doi.org/10.1017/fms.2024.152 Published online by Cambridge University Press

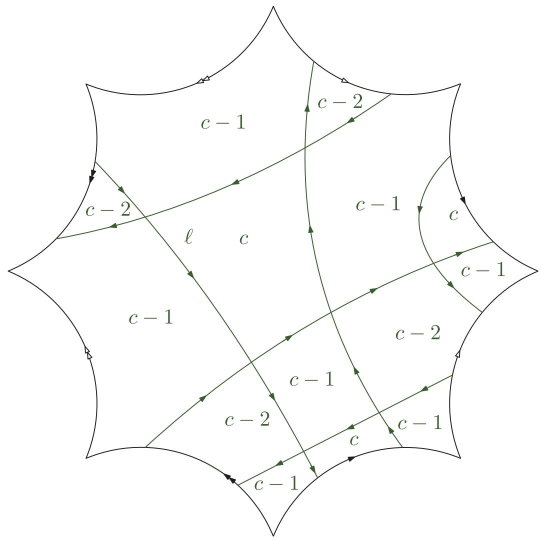

Figure 8. *A representant of the winding number function with $c \in R$, for a loop ℓ of null homology, on a map of genus 2. The loop is drawn in green, and the value on each positively oriented face is displayed on each 2-cell.*

Let ℓ be a based loop of a topological map $\mathbb{G} = (V, E, F)$ which uses each non-oriented edge at most once and each vertex at most twice. We denote by E_ℓ the subset of edges $e \in E$ such that ℓ runs through e or e^{-1}. Whenever a vertex v is visited twice, the four outgoing edges at v visited by ℓ can be ordered e_1, e_2, e_3, e_4 respecting the counterclockwise, cyclic ordering of the orientation of the map. There are six possible configurations, described by the order of appearance of the edges in ℓ. For instance, if ℓ is based at $\overline{e_1}$ and starts with e_1^{-1}, these configurations correspond to the orders $(e_1^{-1}, e_2, e_3, e_4^{-1})$, $(e_1^{-1}, e_2, e_4^{-1}, e_3)$, $(e_1^{-1}, e_3, e_2^{-1}, e_4)$, $(e_1^{-1}, e_3, e_4^{-1}, e_2)$, $(e_1^{-1}, e_4, e_2^{-1}, e_3)$ and $(e_1^{-1}, e_4, e_3^{-1}, e_2)$. Up to rotations, they fall into the three generic configurations illustrated in Figure 9. We say that ℓ is *tame* if it is only of the first type of Figure 9 – that is, if it can be split into two loops that meet at the intersection point. See Figure 10 for an example.

The set V_ℓ of vertices visited twice by ℓ is then called the *(transverse) intersection points* of ℓ.

Definition 2.13. Let ℓ be a tame loop in a map $\mathbb{G}$. The Makeenko–Migdal vector at $v \in V_\ell$ is the 2-form

$$\mu_v = d(\omega_{e_1}) + d(\omega_{e_3}) = -d(\omega_{e_2}) - d(\omega_{e_4}). \tag{18}$$

We denote by $\mathfrak{m}_\ell$ the $\mathbb{R}$-vector space generated by $\{\mu_v, v \in V_\ell\}$ and $\{d\omega_e, e \notin E_\ell\}$.

The Makeenko–Migdal vectors are a way to encode algebraically the previously defined Makeenko–Migdal deformations; see, in particular, Figure 1.

Lemma 2.14. *Let ℓ be a tame loop of a map $\mathbb{G}$. Then*

$$\mathfrak{m}_\ell = \begin{cases} \{\alpha \in \Omega^2(\mathbb{G}, R) : \langle \alpha, \mu_* \rangle = 0\} & \text{if } [\ell]_R \neq 0, \\ \{\alpha \in \Omega^2(\mathbb{G}, R) : \langle \alpha, \mu_* \rangle = \langle \alpha, n_\ell \rangle = 0\} & \text{if } [\ell]_R = 0. \end{cases}$$

Remark 2.15. The signification of the conditions given in the characterisation of $\mathfrak{m}_\ell$ is the following: $\langle \alpha, \mu_* \rangle = 0$ means that the deformation corresponding to α preserves the total area of the graph, whereas $\langle \alpha, n_\ell \rangle = 0$ means that the deformation preserves the algebraic area of the graph, which corresponds to multiplying the area of each face by its winding number (with respect to the loop ℓ).

https://doi.org/10.1017/fms.2024.152 Published online by Cambridge University Press

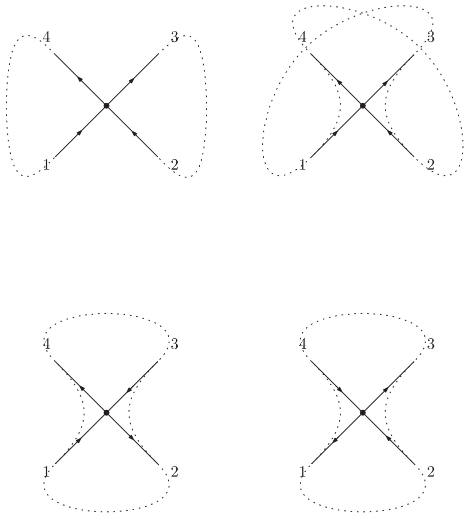

Figure 9. *The three main types of transverse simple intersections. In each case, the dotted paths might be arbitrarily complicated and have multiple intersections outside* $\{e_1, e_2, e_3, e_4\}$. *Only the first case corresponds to a tame loop.*

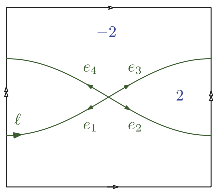

Figure 10. *A tame loop in a graph with one vertex and* 2 *faces. The value of* μ_v *is displayed on each face in blue.*

Proof of Lemma 2.14. Let us first remark that the above construction is invariant by the following appropriate subdivisions. Let us call subdivision of an oriented face f_*, the operation of adding two new vertices on its boundary and adding an edge e connecting them; the new map $\mathbb{G}'$ has 2 new vertices, 1 more edge and 1 more face, with in place of f_*, two faces f_1 and f_2 with the same orientation induced from f_*, while any other face is identified with a face of $\mathbb{G}$. The map $\mathbb{G}'$ being finer than $\mathbb{G}$, ℓ can be identified with a tame loop of $\mathbb{G}'$ that we denote by the same letter. Consider the map $P : \Omega^2(\mathbb{G}', R) \longrightarrow \Omega^2(\mathbb{G}, R)$ with $P(\varphi)(f) = \varphi(f')$ whenever a face f of $\mathbb{G}$ is identified with a face of $\mathbb{G}'$ and $\varphi(f_1) + \varphi(f_2)$ when $f = f_*$. On one hand, $P(d\omega_e) = 0$, and P maps all other vectors of the defining generating family of $\mathfrak{m}'_\ell$ to the generating family of $\mathfrak{m}_\ell$. Therefore, $P(\mathfrak{m}'_\ell) = \mathfrak{m}_\ell$. As $P : \{d\omega_e\}^\perp \longrightarrow \Omega^2(\mathbb{G}, R)$ is an isometry, while $d\omega_e \in \mathfrak{m}'_\ell \cap \ker(P)$, $P(\mathfrak{m}'^{\perp}_\ell) = \mathfrak{m}^\perp_\ell$. On the other hand, $P(\mu'_*) = \mu_*$, and when $[\ell]_R = 0$, $P(n'_\ell) = n_\ell$. We conclude that it is enough to prove the claim for any subdivision of $\mathbb{G}$.

We can then w.l.o.g. assume that ℓ and the paths $a_1, b_1, \ldots, a_g, b_g$ of Lemma 2.12 do not share any edge in common. Under this assumption, let us set $\mathcal{S} = \{\ell, a_1, b_1, \ldots, a_g, b_g\}$ and denote by $\mathrm{T}(\mathcal{S})$ the set of oriented edges e such that an element of $\mathcal{S}$ runs through e or e^{-1}. Let η be the permutation of the edges E such that $\eta(e^{-1}) = \eta(e)^{-1}$ for any edge $e \in E$, with $2 + 4g$ nontrivial cycles associated to elements of $\mathcal{S}$ forgetting the base point. More precisely, for each $\gamma \in \{\ell, a_1, b_1, \ldots, a_g, b_g\}$ with $\gamma = e_1 \ldots e_n$, $(e_1, \ldots, e_n)$ and $(e_1^{-1}, \ldots, e_n^{-1})$ are cycles of η, whereas $\eta(e) = e$ for any $e \notin \mathrm{T}(\mathcal{S})$. For any $\omega \in \Omega^1(\mathbb{G}, R)$, setting

$$\eta.\omega = \omega \circ \eta^{-1}$$

https://doi.org/10.1017/fms.2024.152 Published online by Cambridge University Press

defines a 1-form. We claim that for any oriented edge $e \in \mathrm{T}(\mathcal{S})$,

$$\alpha_e = d\omega_e - d(\eta.\omega_e) \in \mathfrak{m}_\ell.$$

Indeed, it is nonzero only when $\gamma \in \mathcal{S}$ runs through e or e^{-1}, in which case, it follows from (18) that α_e is a Makeenko–Migdal vector at, respectively, $\overline{e}$ or $\underline{e}$.

Let us now consider $\beta \in \mathfrak{m}_\ell^\perp \cap \{\mu_*\}^\perp$. Then

$$\langle \beta, \alpha_e \rangle = \langle \beta, (d - d \circ \eta)\omega_e \rangle = 0, \ \forall e \in \mathrm{T}(\mathcal{S}),$$

whereas $\langle \beta, d\omega_e \rangle = 0, \ \forall e \notin \mathrm{T}(\mathcal{S})$, so that

$$d^*\beta = (d \circ \eta)^*(\beta) = \eta^{-1} \circ d^*\beta \text{ and } \langle d^*\beta, \omega_e \rangle = 0, \ \forall e \notin \mathrm{T}(\mathcal{S}).$$

It follows that

$$d^*\beta = c\omega_\ell + \sum_{i=1}^g (a_i\omega_{a_i} + b_i\omega_{b_i}), \text{ for some } c, a_1, b_1, \ldots, a_g, b_g \in \mathbb{R}.$$

Using the decomposition $\diamond_1 = \bigstar_1^* \oplus \mathcal{H}_1$, we find

$$d^*\beta = cd^*n_\ell \text{ and } ch_\ell + \sum_{i=1}^g (a_i\omega_{a_i} + b_i\omega_{b_i}) = 0.$$

Since $\beta \in \hat{\mathfrak{m}}_\ell^\perp$, for any edge e such that e, e^{-1} do not belong to ℓ, we obtain that $\langle d^*\beta, \omega_e \rangle = 0$. In particular, $a_i = b_i = 0$ for all i and $d^*\beta = c\omega_\ell$. Since $\beta \in \mu_*^\perp$, it follows that either $[\ell]_R = 0$ and $\beta = cn_\ell$ or $c = 0$ and $\beta = 0$. We conclude that either $[\ell]_R = 0$ and $\mathfrak{m}_\ell^\perp \cap \{\mu_*\}^\perp = \mathbb{R}.n_\ell$, or $[\ell]_R \neq 0$ and $\mathfrak{m}_\ell^\perp \cap \{\mu_*\}^\perp = \{0\}$. □

2.3. Regular polygon tilings of the universal cover, tiling-length of a tame loop and geodesic loops

To simplify the presentation, we shall work only with surfaces of genus g obtained by a standard quotient of $4g$ polygons. We fix here notations and definitions relative to the universal cover of such maps. We refer to [7] for more details.

Regular maps and regular loops: A $2g$-*bouquet map* is a map (V, E, F) with 1 vertex v, 1 face and $2g$ edges, so that for $f \in F$, there are $2g$ oriented edges $a_1, b_1, \ldots, a_g, b_g \in E$ corresponding to distinct edges, with $\partial_v f = [a_1, b_1] \ldots [a_g, b_g]$. A $2g$-bouquet map can be obtained by labelling the edges of a $4g$-polygon counterclockwise $e_1, \ldots, e_{4g}$ and gluing e_{i+4k} to e_{i+4k+1} for all $0 \leq k \leq g-1$ $i \in \{1, 2\}$,; cf. Figure 11.

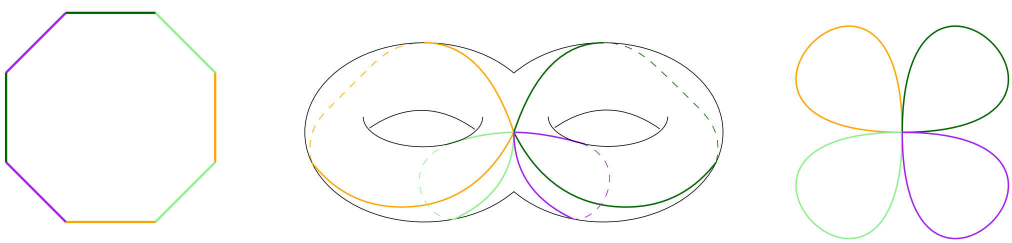

Figure 11. *Three representations of a 4-bouquet. From left to right: as a polygon whose sides are identified pairwise (and whose vertices are all identified), as a graph embedded in a surface of genus 2, and as the skeleton of the corresponding CW complex.*

https://doi.org/10.1017/fms.2024.152 Published online by Cambridge University Press

A *regular map* is a pair given by a map $\mathbb{G} = (V, E, F)$ and a $2g$-bouquet map $\mathbb{G}_g$, such that $\mathbb{G}$ is finer than $\mathbb{G}_g$. Each edge of $\mathbb{G}_g$ is uniquely decomposed as a concatenation of edges of $\mathbb{G}$. Let $\partial E \subset E$ be the set of edges appearing in these concatenations. We then denote by ∂V the set of endpoints of edges of ∂E and $\mathring{V} = V \setminus \partial V$. When $(\mathbb{G}, \mathbb{G}_g)$ is a regular map, we refine the notion of tame loops defined in the previous section as follows. A *loop* $\ell \in \mathrm{L}(\mathbb{G})$ is *regular* whenever it is tame and none of its edges belong to ∂E and $\underline{\ell} \in \mathring{V}$. In particular, its intersection points satisfy $V_\ell \subset \mathring{V}$.

Universal cover of a regular map: Let $(\mathbb{G}, \mathbb{G}_g)$ be a regular map with $\mathbb{G} = (V, E, F)$. When $g = 1$, consider the closed square P_1 with vertices coordinates in $\{-\frac{1}{2}, \frac{1}{2}\}$ and the tiling of $\mathbb{R}^2$ by translation of P_1 by $\mathbb{Z}^2$. When $g \geq 2$, consider a tiling of the Poincaré hyperbolic disc $\mathbb{H}$ by a family of closed regular $4g$-polygons of $\mathbb{H}$ whose sides do not intersect 0, and denote by P_1 the polygon among them enclosing 0. The group Γ_g can be identified with $\mathbb{Z}^2$ when $g = 1$ and with a subgroup of Möbius transformations that acts properly by isometry on $\mathbb{H}$ when $g \geq 2$. The group Γ_g acts freely on the set of tiles, and for each $h \in \Gamma_g$, there is a unique tile P_h with $h \cdot 0$ belonging to the interior of P_h. Let us define $\tilde{\Sigma}_{\mathbb{G}}$ as $\mathbb{R}^2$ when $g = 1$ and $\mathbb{H}$ when $g \geq 2$. The quotient of $\tilde{\Sigma}_{\mathbb{G}}$ by Γ_g is homeomorphic to $\Sigma_{\mathbb{G}}$, and we denote by $p : \tilde{\Sigma}_{\mathbb{G}} \to \Sigma_{\mathbb{G}}$ the quotient mapping. There is a unique CW decomposition of $\tilde{\Sigma}_{\mathbb{G}}$ such that the restriction of p to the interior of each cell of $\tilde{\Sigma}_{\mathbb{G}}$ is a homeomorphism onto the interior of a cell of $\Sigma_{\mathbb{G}}$ labelled by an element of V, E or F. We denote a labelling of the cells of this CW complex by $\tilde{\mathbb{G}} = (\tilde{V}, \tilde{E}, \tilde{F})$ and call $\tilde{\mathbb{G}}$ a *universal cover* of $\mathbb{G}$. There is a natural map from $\tilde{V}, \tilde{E}, \tilde{F}$ to, respectively, V, E and F that we also denote by p. The map $\tilde{\mathbb{G}}$ is finer than the universal cover $\tilde{\mathbb{G}}_g = (\tilde{V}_g, \tilde{E}_g, \tilde{F}_g)$ of $\mathbb{G}_g$, where faces $\tilde{F}_g$ can be identified with polygons $(P_g)_{g\in\Gamma_g}$, and Γ_g acts free transitively on $\tilde{V}_g$. As for maps, the pair $(\tilde{V}, \tilde{E})$ can be identified with a graph, and we denote by $\mathrm{P}(\tilde{\mathbb{G}})$ its set of paths. For each path $\gamma = e_1 \dots e_n \in \mathrm{P}(\mathbb{G})$ and $\tilde{v} \in p^{-1}(v_0)$, the lift of γ from $\tilde{v}$ is the unique path $\tilde{\gamma} = \tilde{e}_1 \dots, \tilde{e}_n \in \mathrm{P}(\tilde{\mathbb{G}})$ with $\underline{\tilde{\gamma}} = \tilde{v}$ and $p(\tilde{e}_k) = e_k$ for all $1 \leq k \leq n$. Viceversa, when $\tilde{\gamma} = (\tilde{e}_1, \dots, \tilde{e}_n) \in \mathrm{P}(\tilde{\mathbb{G}})$, its projection is the path $p(\gamma) = (p(\tilde{e}_1), \dots, p(\tilde{e}_n)) \in \mathrm{P}(\mathbb{G})$. Its image in $\mathrm{RP}(\mathbb{G})$ does not depend on the $\sim_r$ equivalence class $[\gamma]$ of γ; we denote it by $p([\gamma]) \in \mathrm{RP}(\mathbb{G})$. When $\tilde{v} \in \tilde{V}$ and $v = p(\tilde{v})$, the group $\mathrm{RL}_{\tilde{v}}(\tilde{\mathbb{G}})$ of reduced loop of $(\tilde{V}, \tilde{E})$ based at $\tilde{v}$ allows to complete the diagram of Lemma 2.3 in the following way. The proof is standard and left to the reader.

Lemma 2.16. *Let $(\mathbb{G}, \mathbb{G}_g)$ be a regular map, the following assertions hold:*

1. *The sequence*

$$1 \to \mathrm{RL}_{\tilde{v}}(\tilde{\mathbb{G}}) \xrightarrow{p} \mathrm{RL}_v(\mathbb{G}) \to \pi_{1,v}(\mathbb{G}) \to 1$$

   is a short exact sequence.
2. *Denote by Γ_c the kernel of the morphism $\Gamma_{r,g} \to \Gamma_g$ considered in Lemma 2.3 and let $s : \Gamma_g \to \Gamma_{r,g}$ be an injective right-inverse map with $s(\Gamma_g) = \Gamma_{top}$, where Γ_{top} the subgroup of $\Gamma_{r,g}$ generated $S_{top} = \{x_1, y_1, \dots, x_g, y_g\}$, built as follows. Consider a spanning tree $\mathcal{T}$ of the Cayley graph of Γ_g generated by $x_1, \dots, y_g$. Identifying Γ_{top} with paths of the Cayley graph of Γ_g starting from 1, set for any $\gamma \in \Gamma_g$, $s(\gamma) \in \Gamma_{top}$ to be the unique path of $\mathcal{T}$ from 1 to γ. Then Γ_c is free of infinite countable rank with free basis $\{s(\gamma) z_i s(\gamma)^{-1}, \gamma \in \Gamma_g\}$.*
3. *Assume $\tilde{v} \in \tilde{V}_g$ and that $(\ell_i, 1 \leq i \leq r)$, $a_1, b_1, \dots, a_g, b_g \in \mathrm{L}_v(\mathbb{G})$ and $\Theta : \Gamma_{r,g} \to \mathrm{RL}_v(\mathbb{G})$ are as in Lemma 2.3. Denote by $\mathrm{RL}_{top}(\mathbb{G})$ the subgroup of $\mathrm{RL}_v(\mathbb{G})$ generated by $a_1, \dots, b_g$. Then the restrictions of Θ to Γ_{top} and Γ_c yield isomorphisms $\Theta : \Gamma_{top} \to \mathrm{RL}_{top}(\mathbb{G})$ and $\Theta : \Gamma_c \to p(\mathrm{RL}_{\tilde{v}}(\tilde{\mathbb{G}}))$. Denoting by $\tilde{\Theta} : \Gamma_c \to \mathrm{RL}_{\tilde{v}}(\tilde{\mathbb{G}})$ the morphism with $p \circ \tilde{\Theta} = \Theta$, the diagram*

$$\begin{array}{ccccccccc} 1 & \to & \Gamma_c & \to & \Gamma_{r,g} & \to & \Gamma_g & \to & 1 \\ & & \downarrow \tilde{\Theta} & & \downarrow \Theta & & \downarrow & & \\ 1 & \to & \mathrm{RL}_{\tilde{v}}(\tilde{\mathbb{G}}) & \xrightarrow{p} & \mathrm{RL}_v(\mathbb{G}) & \to & \pi_{1,v}(\mathbb{G}) & \to & 1 \end{array}$$

https://doi.org/10.1017/fms.2024.152 Published online by Cambridge University Press

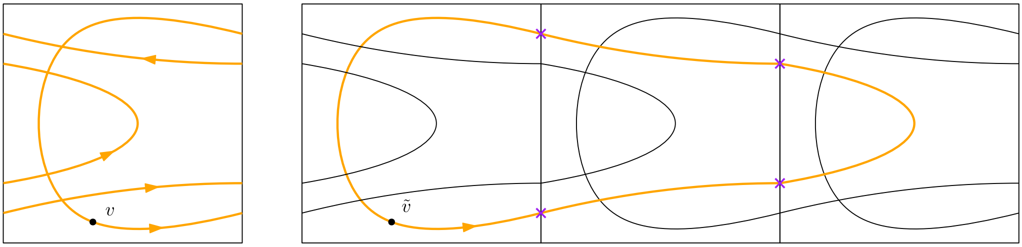

Figure 12. *On the left: a regular loop with base v in a regular map of genus 1. On the right: its lift on the universal cover, with base ṽ. It has a tiling length of 4, as it crosses four times the boundaries of tiles (at the purple crosses).*

is commutative and exact. Consider a spanning tree $\mathcal{T}$ of $\tilde{V}_g$, and for any $x \in \tilde{V}_g$, denote by γ_x the unique path of $\mathcal{T}$ from $\tilde{v}$ to x. Then $\mathrm{RL}_{\tilde{v}}(\tilde{\mathbb{G}})$ is free of infinite rank, with free basis

$$\left\{\widetilde{\gamma_x \ell_i \gamma_x^{-1}}, x \in \tilde{V}_g, 1 \leq i \leq r\right\}.$$

Tile decomposition: For all $h \in \Gamma_g$, we denote by $D_h \subset \tilde{V}, D_h^* \subset \tilde{F}$ and $\mathring{D}_h \subset \tilde{V}$ the subsets of vertices and faces of $\mathbb{G}$, whose image in $\tilde{\Sigma}_{\mathbb{G}}$ is included, respectively, in P_h and its interior $\mathring{P}_h$. The projection $\mathring{D}$ of $\mathring{D}_h$ does not depend on $h \in \Gamma$. When $U \subset \tilde{F}$ and $E_c \subset \tilde{E}$, we denote by $U \setminus E_c$ the subgraph of the graph of $\tilde{\mathbb{G}}^*$, where all faces from $\tilde{F} \setminus U$ and all edges dual to E_c are removed. Let us consider the oriented graph with vertices Γ_g such that there is an edge between a and b if and only if P_a and P_b share a side. The action of Γ_g on $\mathbb{H}$ induces a free, transitive, isometric action on this graph, and we denote by $|h|_{\Gamma_g}$ the *distance* between any $h \in \Gamma_g$ and 1. For any nonconstant regular loop $\ell = (e_1, \ldots, e_n) \in \mathrm{L}(\mathbb{G})$, we call

$$|\ell|_D = n - 1 - \#\left\{1 \leq i \leq n-1 : \exists h \in \Gamma_g \text{ with } \{\underline{e_i}, \overline{e_i}, \overline{e_{i+1}}\} \subset D_h\right\}$$

the *tiling length* of[35] ℓ. See Figure 12.

There is then a unique tuple $\gamma_0, \ldots, \gamma_{|\ell|_D}$ of paths of $\mathbb{G}$, such that for any lift $\tilde{\ell}$ of ℓ, there are lifts $\tilde{\gamma}_0, \ldots, \tilde{\gamma}_{|\gamma|_D}$ of $\gamma_0, \ldots, \gamma_{|\ell|_D}$ such that

$$\tilde{\ell} = \tilde{\gamma}_0 \ldots \tilde{\gamma}_{|\gamma|_D}, \tag{19}$$

and for all $0 \leq k \leq |\ell|_D$, there are $h_0, h_1, \ldots, h_{|\ell|_{|D|}} \in \Gamma_g$ such that all[36] vertices of $\tilde{\gamma}_k$ belong to D_{h_k}, while $\ell_\Gamma = (h_0, \ldots, h_{|\ell|_D})$ is a path in Γ_g. We call $\ell_D = \gamma_{|\ell|_D} \gamma_0$ the *initial strand* of ℓ. We call ℓ_Γ the *tiling path* of ℓ and set

$$|\ell|_\Gamma = |h_{|\ell|_D}|_{\Gamma_g}.$$

A loop ℓ_1 of $(\mathbb{G}, \mathbb{G}_g)$ is called an *inner loop* of ℓ if ℓ_1 is regular and included in $\mathring{D}$, and $\ell_1 \prec \ell$. We then say that $\underline{\ell_1}$ is a *contractible intersection point* of ℓ and denote by $V_{c,\ell}$ the set of such points. A *proper loop* is a regular loop ℓ with $\#V_{c,\ell} = 0$.

A path $\gamma \in \mathrm{P}(\mathbb{G})$ is said to be *geodesic* when its embedding in the surface is the restriction of a geodesic of the surface.[37] A *path* in Γ_g is *geodesic* if it is the tiling path of a geodesic path of a regular map.

[35]Since the loop is regular, it is also understood as the number of pair consecutive edges of ℓ crossing the boundary of a polygon.
[36]This latter claim does not hold if ℓ is not regular.
[37]Mind that we also consider the power of a geodesic to be a geodesic.

https://doi.org/10.1017/fms.2024.152 Published online by Cambridge University Press

2.4. Shortening homotopy sequence

We define here operations on regular loops allowing to decrease their tiling length. We say that a sequence $\ell_1, \ldots, \ell_n$ is a *shortening homotopy sequence* from ℓ_1 to ℓ_n if $\ell_1, \ldots, \ell_n$ are regular loops such that $|\ell_1|_D \geq \ldots \geq |\ell_n|_D$, and for all $1 \leq l < n$,

$$\#V_{c,l} = \#V_{c,l+1} = 0 \text{ or } \#V_{c,l} > \#V_{c,l+1},$$

while there is a regular map (V, E, F) with $\ell_l, \ell_{l+1} \in \mathrm{P}(\mathbb{G})$ and a subset of faces $K_l \subsetneq F$, with

$$\ell_l \sim_{K_l} \ell_{l+1}.$$

The aim of this section is to prove the following.

Proposition 2.17. *For any proper loop ℓ, there is a shortening homotopy sequence $\ell_1, \ldots, \ell_m$, a geodesic loop ℓ' and a path η within the same map $\mathbb{G} = (V, E, F)$ as ℓ_m, such that $\ell_m \sim_K \eta\ell'\eta^{-1}$ for some $K \subset F$ with $K \neq F$. The path η can be chosen simple, within a fundamental domain and crossing ℓ_m and ℓ' only at their endpoints.*

We need two additional notions for this proof.

Bulk of a loop: Consider a regular map $(\mathbb{G}, \mathbb{G}_g)$ with $\mathbb{G} = (V, E, F)$ and a contractible loop ℓ of $\mathbb{G}$ whose lift is a loop $\tilde{\ell}$ of $\tilde{\mathbb{G}}$. Let E_c be the set of edges used by $\tilde{\ell}$ and let O_ℓ be the unbounded component of $\tilde{\mathbb{G}}^* \setminus E_c$. The *bulk* of ℓ is then $K_\ell = p(\tilde{F} \setminus O_\ell)$. Since E_ℓ is connected, the image of O_ℓ in $\tilde{\Sigma}_{\mathbb{G}}$ is a surface with one boundary, and the image $\tilde{X}_\ell$ of $\tilde{F} \setminus O_\ell$ in $\tilde{\Sigma}_{\mathbb{G}}$ is a contractible set. The image of ℓ is then contractible within $X_\ell = p(\tilde{X}_\ell)$ and

$$\ell \sim_{K_\ell} \ell_*,$$

where ℓ_* is the constant loop at $\underline{\ell}$.

Adding a rim to a regular map: When $(\mathbb{G}, \mathbb{G}_g)$ is a regular map, let us define a map $\mathbb{G}_r$ finer than $\mathbb{G}$ in the following way. First, add exactly one vertex to each edge[38] of $E \setminus \partial E$ with one endpoint in ∂V and exactly two when both endpoints belong to ∂V. Each new vertex is paired uniquely with a vertex of ∂V, and their set inherits the cyclic order of vertices of ∂V. Second, add an edge for each consecutive new vertices. We denote by $\mathbb{G}_r$ the new map defined thereby and call the set $\partial_r E$ of edges added in the second step the *rim* of $\mathbb{G}$. Each face of the new map, whose boundary has an edge in ∂E, has exactly four adjacent edges with exactly one in ∂E_r. We denote this set of faces by F_r. We denote all other faces of $\mathbb{G}_r$ by F_i. For any $f \in F$, either its boundary has no edge in ∂E and it is identified to a face of F_i, or it is the union of faces of $\mathbb{G}_r$ with exactly one in F_i that we abusively also denote by *f*. For any oriented edge e of $\mathbb{G}_r$ belonging to ∂E, its right retract is the oriented edge of ∂E_r belonging to the face of F_r on the right of e. When γ is a path with edges in ∂E, its *right retraction* is the concatenation of the right retraction of its edges. The left retraction is defined likewise.

We can now prove the existence of shortening homotopy sequence starting from any regular loop, using a 5 type of operations.

Step 1–*Deleting contraction points:* Consider a regular loop ℓ such that $\#V_{c,\ell} > 0$ of a regular map with faces set F. Any lift $\tilde{\alpha}$ of an inner loop $\alpha \prec \ell$ is a loop, and we can consider its bulk. Denote by K the union of bulks for all inner loops. Any face bordering ∂E does not belong to K so that $K \subsetneq F$, while ℓ is $\sim_K$-equivalent to the regular loop ℓ' with all inner loops erased.

Step 2–*Backtrack erasure:* Assume that ℓ is a regular loop of a regular map $(\mathbb{G}, \mathbb{G}_g)$ such that there is $1 < i < |\ell|_D$ with $h_{i-1} = h_{i+1}$, where $(h_1, \ldots, h_{|\ell|_D})$ is the tiling path of ℓ. Consider the decomposition of $\tilde{\ell}$ as in (19). Let $\mathbb{G}'$ be the map $(\mathbb{G}, \mathbb{G}_g)$ with a rim added. Denote by e_i and e_o the last and first edge of γ_{i-1} and γ_{i+1}. Then $\overline{e}_i$ and $\underline{e}_o$ belong to the same edge e of $\mathbb{G}_g$. Let $\beta' \in \mathbb{P}(\mathbb{G}')$ be the reduced path using only edges of the rim with $\underline{\beta'} = \overline{e}_i$ and $\overline{\beta}' = \underline{e_o}$. Denote by γ'_{i-1} and γ'_{i+1} the reduction of $\gamma_{i-1}e_i^{-1}$

[38]Recall the definition of ∂E and ∂V for a regular map in Section 2.3.

https://doi.org/10.1017/fms.2024.152 Published online by Cambridge University Press

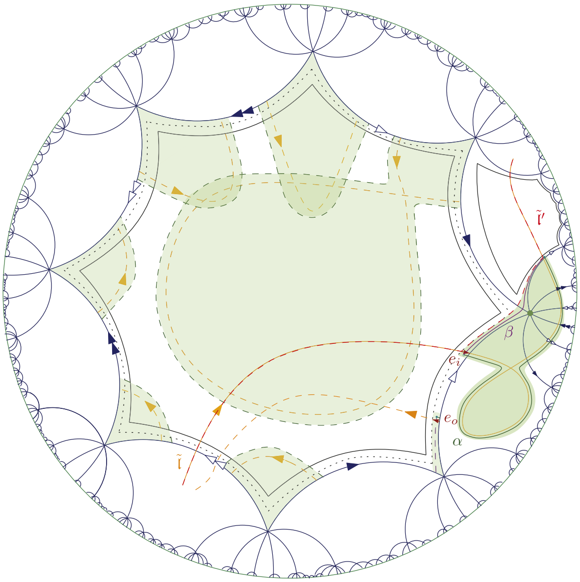

Figure 13. *Discrete homotopy at a left turn of ℓ when $g = 2$ and $k = 7$. The latter vertex is shown as a green dot; contracted faces are shown in green. The second rim is displayed with dotted lines. A lift $\tilde{\ell}$ of the initial loop in displayed in plain orange line, while a lift $\tilde{\ell}'$ of the terminal loop is displayed in dashed red line.*

and $e_o^{-1}\gamma_{i+1}$. The *backtrack erasure* for the backtracking (h_{i-1}, h_i, h_{i+1}) of ℓ is the regular loop

$$\ell' = \gamma_1 \dots \gamma_{i-2}\gamma'_{i-1}\beta'\gamma'_{i+1}\gamma_{i+2} \dots \gamma_{|\ell|_D}.$$

It can be obtained from ℓ by the following discrete homotopy. Since a lift of the paths β' and $e_i\gamma_i e_o$ starting in $D_{h_{i-1}}$ both ends in D_{h_i}, the loop of $e_i\gamma_i e_o\beta'^{-1}$ is contractible. Denote by K_{bt} its bulk. Then

$$\ell \sim_{F_{bt}} \ell'.$$

Since γ_i only intersect the rim of $\mathbb{G}'$ through the edge e, any face belonging to the rim whose boundary intersects two different edges of $\mathbb{G}_b$ is not in K_{bt}. It follows that $K_{bt} \neq F'$.

Step 3-*Vertex switch:* Let ℓ be a regular loop of a regular map $(\mathbb{G}, \mathbb{G}_g)$ and consider its decomposition as in (19). A *half turn* of ℓ is a sequence $\gamma_l, \dots, \gamma_{l+k}$ such that $2g \leq k \leq 4g-1$, and $D_{h_l}, D_{h_{l+1}}, \dots, D_{h_{l+k}}$ runs around a common vertex $v \in \mathbb{G}_g$. Consider such a long turn and let $\mathbb{G}' = (V', E', F')$ be the map obtained from $\mathbb{G}$ by adding twice a rim as described in the last paragraph. See Figure 13 for an example. Let e_i and e_o be, respectively, the last and the first edge of γ_l and γ_{l+k} in $\mathbb{G}'$. Besides, let $\beta_p \in \mathrm{P}(\mathbb{G}'^*)$ be the shortest reduced path from a face adjacent of e_i to a face adjacent of e_o that crosses first e_i and uses only faces of F_r so that its lift starting from $D^*_{h_l}$ goes through $D^*_{h_l} \cup D^*_{h_{l+k}}$ and ends in $D^*_{h_{l+k}}$. Let $\beta' \in \mathrm{P}(\mathbb{G}')$ be the reduced path from $\underline{e_i}$ to $\underline{e_o}$, such that each edge of β' is bordering a face of β_p. Denote by γ'_l and γ'_{k+l} the reduction of $\gamma_l \overline{e_i}^{-1}$ and $e_o^{-1}\gamma_{k+l}$. The *vertex switch* of ℓ for the considered half turn is the regular loop

$$\ell' = \gamma_0\gamma_1 \dots \gamma_{l-1}\gamma'_l\beta'\gamma'_{k+l}\gamma_{k+l+1} \dots \gamma_{|\ell|_D}.$$

It can be obtained from ℓ by the following discrete homotopy. Consider the loop $e_i\gamma_{l+1} \dots \gamma_{l+k-1}e_o\beta'^{-1}$. Since a lift of β' starting in D_{h_l} ends in $D_{h_{l+k}}$, it follows that $e_i\gamma_{l+1} \dots \gamma_{l+k-1}e_o\beta'^{-1}$ is contractible. Denote by K_{sw} its bulk.

https://doi.org/10.1017/fms.2024.152 Published online by Cambridge University Press

Then,

$$e_i\gamma_{l+1}\dots\gamma_{l+k-1}e_o\sim_{K_{sw}}\beta'$$

and

$$\ell\sim_{K_{sw}}\gamma_0\dots\gamma_{l-1}\gamma_l'e_i\gamma_{l+1}\dots\gamma_{l+k-1}e_o\gamma_{k+l+1}\dots\gamma_{|\ell|_D}\sim_{K_{sw}}\ell'.$$

Besides, $F_{sw}\neq F'$. Indeed, consider the map $\mathbb{G}_1$ obtained by adding a single rim to $\mathbb{G}$, so that $\mathbb{G}'$ is finer than $\mathbb{G}_1$. Let $F_{cr},\tilde{F}_{cr}$ be the set of of faces of $\mathbb{G}_1$ neighbouring, respectively, $p(v)$ and v. The restriction of p to $\tilde{F}_{cr}$ is a homeomorphism onto F_{cr}. Since $k<4g$, there is at least one face $\tilde{f}_{cr}$ of $\tilde{F}_{cr}$ that does not belong to $p^{-1}(F_{sw})$. Since β uses only faces of F_r', any face of $F'\setminus F_r$ included in $f_{cr}=p(\tilde{f}_{cr})$ does not belong to K_{sw}.

The following lemma reformulates a result due to [7] relating $|\ell|_D$ to long turns of ℓ when $g\geq 2$.

Lemma 2.18. *Let ℓ be a regular loop of a regular map $(\mathbb{G},\mathbb{G}_b)$. There is a finite sequence $\ell_1,\dots,\ell_n$ or regular loops obtained by vertex switches or backtracking erasures such that $\ell_1=\ell$, $|\ell_1|_D\geq|\ell_2|_D\dots\geq|\ell_n|_D$ and*

$$|\ell_n|_D=|\ell|_{\Gamma_g}.$$

Proof. The case $g=1$ is elementary. An argument goes as follows. The path in $\Gamma_1=\mathbb{Z}^2$ associated to ℓ can be assumed up to axial symmetries that h_ℓ has nonnegative coordinates. A backtracking of ℓ_Γ can be erased by a backtracking erasure of ℓ. A path is geodesic if and only if all of its increments coordinates are nonnegative. There are two consecutive increments with a negative followed by a positive sign. This pair corresponds to a backtrack or a half turn of ℓ if one or two coordinates change. Applying to a backtrack erasure or a switch at the half turn, the new loop has one less pair of increments with coordinates changing sign.

When $g\geq 2$, the result follows from [7, Lemma 2.5]. In the setting of [7], a half turn of ℓ is a half cycle of the path in Γ_g associated to ℓ. A switch at a half turn corresponds to a replacement of a half cycle with its complementary. Moreover, in the setting of [7], replacing a long chain by its complementary chain can be obtained by successively replacing a long cycle by its complementary cycle. □

Step 4–From minimal tiling length to geodesic tiling paths: We say that a regular path γ of a regular map has *minimal tiling length* when $|\gamma|_D=|\gamma|_\Gamma$. When $g\geq 2$, the following is a consequence of [7, Thm 2.8].

Lemma 2.19. *If ℓ is a regular loop of a regular map, there is a sequence of regular loops $\ell_1,\dots,\ell_n$ with minimal tiling length equal to $|\ell_1|_\Gamma$ obtained by switches and backtrack erasure, such that $\ell_1=\ell$ while the tiling path of ℓ_n is geodesic.*

Proof. When $g\geq 2$, in the setting of [7], our condition for a tiling path to be geodesic is equivalent for it to be a shortest path. Since switches at half turns imply switches for half cycles of the tiling path in the setting of [7], the result follows from point (c) of [7, Thm 2.8].

When $g\geq 1$, for any regular loop with minimal tiling length, we can assume w.l.o.g. that both coordinates of the endpoint (a,b) of ℓ_Γ are nonnegative. When γ is a path of $\mathbb{Z}^2$ with only positive coordinates, a corner swap of γ is the path obtained by replacing a sequence of the form $(x,y),(x+1,y),(x+1,y+1)$ with $(x,y),(x,y+1),(x+1,y+1)$ or vice versa. Any other path of $\mathbb{Z}^2$ with the same endpoints can be obtained by corner swaps and backtrack erasure. Since a switch at a half turn of ℓ implies a corner swap of its tiling path and that tiling paths with positive coordinates have minimal length in $\mathbb{Z}^2$, the claim follows. □

Step 5–From geodesic tiling paths to geodesic paths: Assume that ℓ is a regular loop such that ℓ_Γ is geodesic and set $n=|\ell|_D=|\ell|_\Gamma$. Let $\ell^{(*)}$ be a geodesic loop with $\ell_\Gamma^{(*)}=\ell_\Gamma$. Up to translation of the geodesic associated to $\ell^{(*)}$, we can assume that $\ell^{(0)}$ and $\ell^{(*)}$ are regular paths of a same regular map $(\mathbb{G}^{(0)}, \mathbb{G}_g^{(0)})$. Let $\eta \in P(\mathbb{G}^{(0)})$ that does not cross the boundary of the polygon, while $\underline{\eta} = \underline{\ell}$ and $\overline{\eta} = \underline{\ell}^{(*)}$, without using any edge of $\ell^{(*)}$. Denote by $(\mathbb{G}, \mathbb{G}_g)$ the regular map obtained by adding a rim to $(\mathbb{G}^{(0)}, \mathbb{G}_g^{(0)})$. Using the same notation as in (19), consider the tile paths decompositions of ℓ and $\ell^{(*)}$ adding an upper-script $(*)$ for the second decomposition. For any $0 \le k \le n-1$, let e_k and $e_k^{(*)}$ be the last edges of, respectively, γ_k and $\gamma_k^{(*)}$, denote by β_k the reduced path with edges in ∂E_r from $\underline{e_k^{(*)}}$ to $\underline{e_k}$ and define $\ell^{(k)}$ as the reduction of

$$\eta\gamma_0^{(*)} \dots \gamma_k^{(*)} {e_k^{(*)}}^{-1} \beta_k e_k \gamma_{k+1} \dots \gamma_n.$$

Let us set $\ell^{(n)} = \eta \ell^{(*)} \eta^{-1}$ and $\ell^{(-1)} = \ell$. Let α_k be the reduction of the loop $\eta^{-1}\gamma_0 e_0^{-1} \beta_0^{-1} e_0^{(*)} {\gamma_0^{(*)}}^{-1}$ when $k=0$, ${e_{k-1}^{(*)}}^{-1} \beta_{k-1} e_{k-1} \gamma_k {e_k}^{-1} \beta_k^{-1} {\gamma_k^{(*)}}^{-1}$ when $0 < k < n$ and ${e_{n-1}^{(*)}}^{-1} \beta_{n-1} e_{k-1} \gamma_n \eta {\gamma_n^{(*)}}^{-1}$ when $k = n$. With this notation,

$$\ell \sim_r \eta\alpha_0\gamma_0^* \alpha_1 \gamma_1^* \dots \alpha_k \gamma_k^{(*)} {e_k^{(*)}}^{-1} \beta_k e_k \gamma_{k+1} \dots \gamma_n \text{ for } 0 \le k < n$$

and

$$\ell \sim_r \eta\alpha_0\gamma_0^* \alpha_1 \gamma_1^* \dots \alpha_n \gamma_n^{(*)} \eta^{-1}.$$

Therefore, for all $0 \le k \le n$

$$\ell^{(k-1)} = \alpha\beta \text{ and } \ell^{(k)} = \alpha\alpha_k\beta, \tag{20}$$

for some paths $\alpha, \beta \in \mathrm{P}(\mathbb{G})$. For all $0 \le k \le n$, α_k is contractible. Denoting by K_k its associated bulk, (20) yields

$$\ell^{(k)} \sim_{K_k} \ell^{(k-1)} \text{ for all } 0 \le k \le n.$$

Besides, since α_k intersects at most two edges of $\mathbb{G}_g$, any face within the rim $f \in F_r$, which borders a different edge of $\mathbb{G}_g$, does not belong to K_k. Therefore, $K_k \neq F$.

Proof of Proposition 2.17. For any regular loop ℓ, the claimed shortening homotopy sequence can be obtained by applying first the deletion of contraction points, followed by Lemma 2.18, 2.19 and lastly a shortening homotopy sequence from a loop with geodesic tiling path to a loop conjugated to a geodesic loop. □

The following lemma is not necessary for our main argument and can be skipped at first reading. Let us note that it is also possible to do the vertex switch operation (Step 3) before deleting contraction points (Step 1) thanks to the following.

Lemma 2.20. *Consider ℓ is a regular loop within a regular map $(\mathbb{G}, \mathbb{G}_g)$ with faces set F. Denote, respectively, by K and E_{in} the union of bulks and the set of edges of its initial strand ℓ_D. Then $F \setminus K$ is connected in $\mathbb{G}^* \setminus (\partial E \cup E_{in})$.*

Proof. Since ℓ is regular, any edge crossing ∂E does not belong to E_{in}, and faces adjacent to ∂E belong to the same connected component X of $F \setminus K$ in $\mathbb{G}^* \setminus (\partial E \cup E_{in})$. Denote by $\tilde{X}$ the lift of X in D_1^*. Assume that $F \setminus K$ is not connected in $\mathbb{G}^* \setminus (\partial E \cup E_{in})$ and consider a connected component K' different from X. Then all edges of $\partial K'$ belong to E_{in}. Since the infinite connected component of $\tilde{\mathbb{G}}^* \setminus E_{in}$ is given by $\tilde{F} \setminus D^* \cup X$, the lift of K' in D_1^* is included in the bounded connected component of $\tilde{\mathbb{G}}^* \setminus E_{in}$, where we identified E_{in} with the set of edges of the lift of ℓ_D starting from D_1. It follows that K' is included in K, which is a contradiction. □

https://doi.org/10.1017/fms.2024.152 Published online by Cambridge University Press

2.5. *Nested and marked loops*

Nested loop: We say that a loop ℓ of a regular map with n transverse intersection points is *nested* if it is regular and if there are subloops $\ell_1 \prec \ell_2 \prec \ldots \prec \ell_n$ with a strictly increasing number of intersection points. By convention, a constant loop is a nested loop. A regular loop is nested if and only if its transverse intersection points can be labelled $v_1, v_2, \ldots, v_n$ so that it visits them in the order $(v_1 v_2 \ldots v_{n-1} v_n v_n v_{n-1} \ldots v_2 v_1)$. See Figure 14.

Remark 2.21. A nested loop is an example of a splittable loop as defined in [21, Section 6.5], originally introduced in [36] and called therein planar loops. Note that the right example in Figure 14 is splittable but not nested.

Marked loops: A *marked loop* is a couple (ℓ, γ_{nest}) of a regular loop and a regular path within a regular map $\mathbb{G}$ such that

1. When $(\gamma_0, \ldots, \gamma_{|\ell|_D})$ denotes the tiling decomposition of ℓ, $\gamma_0 = \gamma_{nest}\gamma'$, for some path γ'.
2. The path γ_{nest} is nonconstant and of the form $\alpha\ell_{nest}\beta$ where ℓ_{nest} is a nested loop and α, β are simple paths, such that the only intersections between α, β and ℓ_{nest} are at $\overline{\alpha}$ and $\underline{\beta}$.
3. The path γ_{nest} does not intersect transversally the two components of the initial strand ℓ_D.
4. The path γ_{nest} does not intersect any inner loop of $\alpha\beta\gamma'\gamma_1 \ldots \gamma_{|\ell|_D}$.
5. The bulk F_{nest} of the contractible loop ℓ_{nest} has exactly $\#V_\ell$ faces of $\mathbb{G}$, and there is exactly one face f_o of $\mathbb{G}$ adjacent to F_{nest} in $\mathbb{G}^*$.

See Figure 15 for an example. We call the loop and the path defined by $(\ell, \gamma_{nest})^\wedge = \alpha\beta\gamma'\gamma_1 \ldots \gamma_{|\ell|_D}$ and $(\ell, \gamma_{nest})^{\wedge_*} = \gamma'\gamma_1 \ldots \gamma_{|\ell|_D}$ the *pruning* and the *cut* of (ℓ, γ_{nest}). We shall often denote them abusively simply by $\ell^\wedge$ and $\ell^{\wedge_*}$. We call f_o the *outer face* of (ℓ, γ_{nest}) and the simple sub-loop of ℓ with length 1 the *central* loop of (ℓ, γ_{nest}). Being a sub-loop of ℓ_{nest}, it is contractible; faces belonging to its bulk are called *central*.

A *moving edge* is an edge e of ℓ with the following property:

- When ℓ_{nest} is constant, e is any edge of γ_{nest}.
- Otherwise, e bounds a central face of ℓ_{nest}.

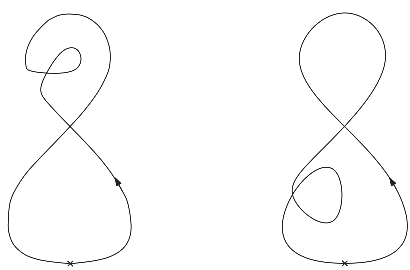

Figure 14. *Left: a nested loop. Right: this is not a nested loop.*

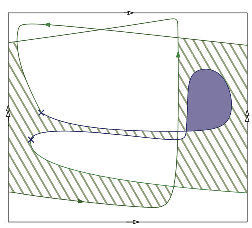

Figure 15. *A marked loop. Its nested part is drawn in blue. There are exactly one central face coloured in blue and one outer face filled with dashed green lines.*

https://doi.org/10.1017/fms.2024.152 Published online by Cambridge University Press

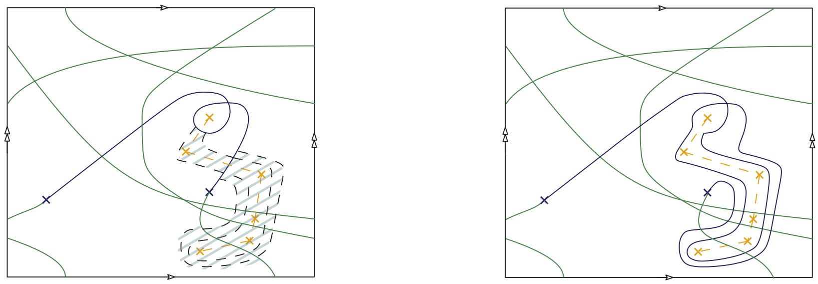

Figure 16. *Left: A marked loop with the nested part drawn in blue. New edges of the modified regular map are drawn with dashed lines. The union of faces of F_{stem} is a stroke with dashed lines. Right: Pull of the left marked loop along the path of the dual drawn in orange.*

Remark 2.22. For any nested loop ℓ included in a fundamental domain, it is easily shown by induction on $n = \#V_\ell$ that the dual graph $\mathbb{G}^*$ with the edges of ℓ removed has exactly $n+1$ connected components. The fifth condition above can be removed considering regular maps finer than $\mathbb{G}$.

The following is then a simple variation of Proposition 2.17.

Lemma 2.23. *For any marked loop (ℓ, γ_{nest}) with $\ell^\wedge$ proper, there is a shortening homotopy sequence $\ell_1, \ldots, \ell_m$ such that*

1. $\ell_1 \sim_c \ell$,
2. *There is a common nested subpath γ_{nest} of $\ell_1, \ldots, \ell_n$, such that (ℓ_k, γ_{nest}) is a marked loop for all $k \geq 1$, and $\ell_k^\wedge$ is proper for $k \geq 2$.*
3. *There are proper subsets $K_1, \ldots, K_m$ of faces, such that $\ell_k^{\wedge *} \sim_{K_k} \ell_{k+1}^{\wedge *}$ for all $1 \leq k < m$.*
4. *There is a marked loop (ℓ', γ'_{nest}), such that $\ell_m \sim_\Sigma \ell'$ and $\ell'^\wedge$ is geodesic.*

2.6. Pull and twist moves

We introduce here two operations on loops in order to later modify shortening homotopy sequences to satisfy the constraint imposed by Makeenko–Migdal equations – namely, to keep constant the algebraic area of loops introduced in Section 2.2. This type of operation shall be required only when considering loops with vanishing homology.

Pull move: Consider a nonconstant marked loop (ℓ, γ_{nest}) in a regular map $(\mathbb{G}, \mathbb{G}_g)$ with $\ell^\wedge$ has no inner loops that are subpaths of ℓ_D. Then the graph obtained from the dual graph $\mathbb{G}^*$ by removing all edges crossing ∂E or ℓ_D but edges of γ_{nest} is connected. For any face f of $\mathbb{G}$ and any moving edge e that does not bound f, there is therefore a simple path $\gamma^* = a_1^* \ldots a_m^*$ in the dual graph $\mathbb{G}^*$ with endpoint f and first edge a_1^* dual to e that crosses neither ∂E nor ℓ_D but possibly at γ_{nest}. Let us define inductively a new map $\mathbb{G}'$ finer than $\mathbb{G}$, a new marked loop (ℓ', γ'_{nest}), as well as a subset F_{stem} of faces of $\mathbb{G}'$. An example of the result in displayed in Figure 16. Let us first set $F_{stem} = \emptyset$. Denote by $a_1, \ldots, a_m$ the edges dual to $a_1^*, \ldots, a_m^*$. Let $k \geq 1$ be the largest k such that a_k^* is dual to an edge of γ_{nest}.

1. Add two new vertices to all edges dual to $a_k^*, \ldots, a_m^*$. For all $l \geq k$, when $a_l = a_{l,0}a_{l,1}a_{l,2}$ is the edge decomposition of a_l in the new map, replace a_l by $a_{l,1}$.
2. Cut all faces visited by $a_k^* \ldots a_m^*$ but f into three faces adding two noncrossing edges such that endpoints of a new edge do not belong to the same initial edge. Add to F_{stem} all new faces bounded by 2 new edges.
3. Cut the face f into two faces, adding an edge connecting the two new vertices on the edge dual to e_m^* introduced in Step 2. Add to F_{stem} the new face included in f whose boundary has only two edges.
4. Denote by η the simple path using only edges added in Step 2 and 3 such that $\underline{\eta} = \underline{a_k}$ and $\overline{\eta} = \overline{a_k}$. Transform ℓ and γ_{nest}, replacing the occurrence of the edge a_k and a_k^{-1} by, respectively, η and η^{-1}.

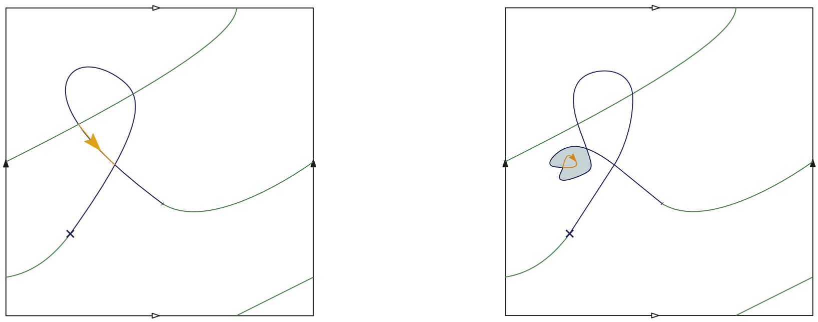

Figure 17. *Left: A marked loop with the nested part drawn in blue. The chosen moving edge is drawn in orange. Right: n-twist of the left marked loop, with $n = -2$ and the chosen moving edge. The new moving edge is displayed in orange.*

5. When $k = 1$, stop the procedure. Otherwise, repeat this operation for the nested loop obtained in Step 4 and the path $a_1^* \dots a_m^*$.

The last marked loop produced is called the *pull of* (ℓ, γ_{nest}) *along* γ^*.

Twist move: Consider a marked loop (ℓ, γ_{nest}) with a moving edge e.

Let us refine a regular and marked loop as follows. Add a vertex to e and cut the face left of e into two faces, adding an oriented edge e' with both endpoints equal to the new vertex, such that e' is the boundary a positively oriented face. The initial moving edge reads $e = e_1 e_2$ in the new map. The *left twist* of (ℓ, γ_{nest}) is the marked loop obtained by replacing the occurrence of e by $e_1 e' e_2$ in both ℓ and γ_{nest}. The new marked loop has then e' has unique moving edge. We denote by F_{tw} the face bounded by e'. The right twist of (ℓ, γ_{nest}) is defined similarly considering the right face and a negative orientation. When n is, respectively, positive or negative, the n-twist of a marked loop is obtained by applying, respectively, n left twists or $-n$ right-twists. We denote then by F_{tw} the $|n|$ faces of the new map bounded solely by newly added edges. See Figure 17 for an example.

2.7. Vertex desingularisation and complexity

Consider a regular map $\mathbb{G}$. Assume that ℓ is a regular loop and $v \in V_\ell$ is an intersection point. We denote by ℓ_1 and ℓ_2 the two sub-loops of ℓ based at v such that $\ell \sim_c \ell_1 \ell_2$. We then set

$$\delta_v \ell = \ell_1 \otimes \ell_2 \in \mathbb{C}[\mathrm{L}_c(\mathbb{G})]^{\otimes 2}, \tag{21}$$

with the convention that ℓ_1 is left of ℓ_2 at v as displayed on Figure 2. By definition of Makeenko-Migdal vectors given in Section 2.2, there are[39] linear forms $(\alpha_v)_{v \in V_\ell}$ and $(\beta_e)_{e \in E \setminus E_\ell}$ on $\mathfrak{m}_\ell$ such that

$$X = \sum_{v \in V_\ell} \alpha_v(X) \mu_v + \sum_{e \in E \setminus E_\ell} \beta_e(X) d\omega_e, \ \forall X \in \mathfrak{m}_\ell.$$

We then set

$$\delta_X \ell = \sum_{v \in V_\ell} \alpha_v(X) \delta_v \ell.$$

[39]We fix them arbitrarily, for instance, using the pseudo-inverse of the Gram matrix of the spanning family $(\alpha_v)_{v \in V_\ell}$ and $(\beta_e)_{e \in E \setminus E_\ell}$.

https://doi.org/10.1017/fms.2024.152 Published online by Cambridge University Press

Let us define a complexity on loops that strictly decreases after such operations. Let us set

$$\mathcal{C}(\ell) = |\ell|_D + \#V_{c,\ell} \tag{22}$$

when ℓ is a regular loop and

$$\mathcal{C}^{\mathfrak{m}}(x) = |\ell|_D + \#V_{c,\ell^\wedge} \tag{23}$$

when $x = (\ell, \gamma_{nest})$ is a marked loop.

Lemma 2.24.

1. *For any regular loop ℓ, $v \in V_\ell$, if $\delta_v \ell = \ell_1 \otimes \ell_2$, then*

$$\mathcal{C}(\ell_1), \mathcal{C}(\ell_2) < \mathcal{C}(\ell).$$

Moreover, if $[\ell] \neq 0$, then $[\ell_1]$ or $[\ell_2] \neq 0$.

2. *For any marked loop x, $\mathcal{C}^{\mathfrak{m}}(x)$ only depends on $x^{\wedge_*}$. Moreover, when $y = (\ell, \gamma_{nest})$ is a marked loop with $y^{\wedge_*} = x^{\wedge_*}$, for $v \in V_{\ell^\wedge}$, if $\delta_v \ell = \ell_1 \otimes \ell_2$, then there are ℓ_1', ℓ_2', with $\ell_i' \sim_c \ell_i$ and subpaths γ_1, γ_2 of ℓ_1', ℓ_2', such that $x_1 = (\ell_1', \gamma_1), x_2 = (\ell_2', \gamma_2)$ are marked loops with*

$$\mathcal{C}^{\mathfrak{m}}(x_1), \mathcal{C}^{\mathfrak{m}}(x_2) < \mathcal{C}^{\mathfrak{m}}(x).$$

Proof. Consider a regular loop ℓ. When $v \in V_{c,\ell}$, one can assume that ℓ_2 is an inner loop that is $|\ell_2|_D = 0$, and

$$\#V_{c,\ell_1} + \#V_{c,\ell_2} + 1 = \#V_{c,\ell}.$$

Otherwise, both ℓ_1 and ℓ_2 are regular loops both crossing ∂E at least twice so that $|\ell_1|_D, |\ell_2|_D > 0$. Moreover,

$$|\ell_1|_D + |\ell_2|_D = |\ell|_D$$

since both count the number of edges of ∂E crossed by ℓ. Therefore,

$$|\ell_1|_D, |\ell_2|_D < |\ell|_D. \tag{24}$$

Moreover, $\omega_\ell = \omega_{\ell_1} + \omega_{\ell_2}$, $[\ell] = [\ell_1] + [\ell_2]$. In particular, if $[\ell] \neq 0$, $[\ell_1] \neq 0$ or $[\ell_2] \neq 0$. This concludes the proof of the first point. Consider now two marked loops $x = (\ell', \gamma_{nest}'), y = (\ell, \gamma_{nest})$ with $y^{\wedge_*} = x^{\wedge_*}$. Then

$$|\ell^\wedge|_D = |\ell'^\wedge|_D \text{ and } \#V_{c,\ell^\wedge} = \#V_{c,\ell'^\wedge},$$

so that $\mathcal{C}^{\mathfrak{m}}(x) = \mathcal{C}^{\mathfrak{m}}(y)$. Assume that $v \in V_{\ell^\wedge}$ and $\delta_v \ell = \ell_1 \otimes \ell_2$ such that γ_{nest} is a subpath of ℓ_2. Consider e the first edge of ℓ_1 and $\ell_2' \sim_c \ell_2$ with $\underline{\ell_2'} = \underline{\gamma_{nest}}$. Then $(\ell_1, e), (\ell_2', \gamma_{nest})$ are marked loops. If $v \in V_{c,\ell^\wedge}$,

$$\#V_{c,\ell_1^\wedge} + \#V_{c,\ell_2'^\wedge} + 1 = \#V_{c,\ell}.$$

Otherwise, $|\ell_1^\wedge|_D, |\ell_2^\wedge|_D > 0$, and the proof of 2 follows as for the first point. ◻

Let us fix a choice for x_1, x_2 used in the above lemma. Consider a marked loop $x = (\ell, \gamma_{nest})$, $v \in V_\ell$ and assume $\delta_v \ell = \ell_1 \otimes \ell_2$. When $v \in V_{\ell^\wedge}$, exactly one loop, say ℓ_1, has γ_{nest} as subpath, and we set $x_1 = (\ell_1', \gamma_{nest})$ and $x_2 = (\ell_2, e)$, where e is the first edge of ℓ_2, and $\ell_1' \sim_c \ell_1$ with $\underline{\ell_1'} = \underline{\gamma_{nest}}$. When $v \in V_{\ell_{nest}}$, exactly one loop, say ℓ_1, is a sub-loop of ℓ_{nest}, set $x_1 = (\ell_1, \ell_1)$ and $x_2 = \overline{(\ell', \gamma_{nest})}$, where

https://doi.org/10.1017/fms.2024.152 Published online by Cambridge University Press

(ℓ', γ'_{nest}) is obtained from (ℓ, γ_{nest}) by erasing the edges of ℓ_1 (then $\ell' \sim_c \ell_2$). Otherwise, we set $x_1 = (\ell_1, e_1), x_2 = (\ell_2, e_2)$, where e_i is the first edge of ℓ_i. We then write

$$\delta_v x = x_1 \otimes x_2. \tag{25}$$

3. Yang–Mills measure and Makeenko–Migdal equations

3.1. Metric and heat kernel on classical groups

We recall here briefly the definition and main properties of the heat kernel on classical groups that will be needed to define the discrete Yang–Mills measure. These results are quite standard and can also be found, for instance, in [43, Section 1]. In this text, for any $N \geq 1$, we denote by G_N a compact classical group of rank N – that is, $\mathrm{U}(N), \mathrm{SU}(N), \mathrm{SO}(N)$ or $\mathrm{Sp}(N)$ – following the same conventions as in Section 2.1.2 of [20].

For any compact Lie group G, its Lie algebra $\mathfrak{g}$ is endowed with an invariant inner product $\langle \cdot, \cdot \rangle$. Setting

$$\mathcal{L}_X f(g) = \frac{d}{dt}\Big|_{t=0} f(ge^{tX}),\ \forall f \in C^\infty(G) \text{ and } g \in G,$$

the Laplacian associated to $\langle \cdot, \cdot \rangle$ is the operator defined by

$$\Delta_G f = \sum_{1 \leq i \leq d} \mathcal{L}_{X_i} \circ \mathcal{L}_{X_i}(f),\ \forall f \in C^\infty(G),$$

where $(X_i)_{1 \leq i \leq d}$ is an arbitrary orthonormal basis.

Definition 3.1. The *heat kernel* on G is the solution $p : (0, \infty) \times G \to \mathbb{R}_+, (t, g) \mapsto p_t(g)$ of the heat equation, with $p_t \in C^\infty(G)$ for all $t > 0$ and

$$\begin{cases} \partial_t p_t(g) & = \Delta_G p_t(g),\ \forall g \in G,\ \forall t > 0, \\ \lim_{t \downarrow 0} p_t(g) dg & = \delta_{I_N}, \end{cases} \tag{26}$$

where the convergence in the second line holds weakly.

It defines a semigroup for the convolution product; that is,

$$p_t * p_s = p_{t+s},\ \forall t, s > 0. \tag{27}$$

It inherits the following properties from the conjugation invariance of the scalar product: for all $g, h \in G$ and $t > 0$,

$$p_t(hgh^{-1}) = p_t(g) \tag{28}$$

and

$$p_t(g^{-1}) = p_t(g). \tag{29}$$

When G_N is a compact classical group of rank N, we choose (1) as an invariant inner product.

3.2. Area weighted maps, Yang–Mills measure and area continuity

We recall here a definition of the discrete and continuous Yang–Mills measure in two dimensions on arbitrary surfaces, with a focus on the former.

https://doi.org/10.1017/fms.2024.152 Published online by Cambridge University Press

Area vectors and area-weighted maps: When $\mathbb{G} = (V, E, F)$ is a topological map, an area vector is a function $a : F \to \mathbb{R}_+$. We say that $(\mathbb{G}, a)$ is an *area-weighted map* with volume $\sum_{f \in F} a_f$. When K is a subset of faces of $\mathbb{G}$, we then write $a(K) = \sum_{f \in K} a(f)$ its volume. When $m = (\mathbb{G}, a)$ and $m' = (\mathbb{G}', a')$ are area weighted maps with faces set F and F', m' is finer than m if $\mathbb{G}'$ is finer than $\mathbb{G}$ and $a_f = \sum_{f' \in F' : f' \subset F} a'_{f'}$. When $T > 0$, we denote by

$$\Delta_{\mathbb{G}}(T) = \{a : F \to \mathbb{R}_+ : \sum_{f \in F} a_f = T\}$$

the closed simplex of area vectors of fixed volume T and its interior by

$$\Delta^o_{\mathbb{G}}(T) = \{a \in \Delta_{\mathbb{G}}(T) : a(f) > 0, \ \forall f \in F\}.$$

Its faces are given as follows. For any subset $K \subsetneq F$, we set

$$\Delta_{K,\mathbb{G}}(T) = \{a \in \Delta_{\mathbb{G}}(T) : a(f) = 0, \ \forall f \in K\}$$

and

$$\Delta^o_{K,\mathbb{G}}(T) = \{a \in \Delta_{K,\mathbb{G}}(T) : a(f) > 0, \ \forall f \in F \setminus K\}.$$

The tangent space at any point of $\Delta^o_{\mathbb{G}}(T)$ is canonically identified with the space of 2-forms $X \in \Omega^2(\mathbb{G}, \mathbb{R})$ with $\sum_{f \in F} X(f) = 0$. For such a two form X, we shall also denote by X the associated constant vector field on $\Delta^o_{\mathbb{G}}(T)$ and write $X.\Psi$ for the derivative of a function $\Psi \in C^1(\Delta^o_{\mathbb{G}}(T))$ along X.

When $(\mathbb{G}, B)$ is a map with boundary faces B, we set

$$\Delta_{\mathbb{G},B}(T) = \{a : F \setminus B \to \mathbb{R}_+ : \sum_{f \in F \setminus B} a_f = T\}$$

and

$$\Delta^o_{\mathbb{G},B}(T) = \{a \in \Delta_{\mathbb{G},B}(T) : a(f) > 0, \ \forall f \in F\}.$$

When $\mathbb{G}' = (V', E')$ is finer than $\mathbb{G}$, any face F of $\mathbb{G}$ can be identified with a subset of faces of $\mathbb{G}'$, and for any $a \in \Delta_{\mathbb{G}'}(T)$, we denote $r^{\mathbb{G}'}_{\mathbb{G}}(a)$ or simply $r_{\mathbb{G}}(a) \in \Delta_{\mathbb{G}}(T)$ the associated area vector of $\mathbb{G}$ summing the values of a on any given subset of faces. We then say that the area weighted map $(\mathbb{G}', a)$ is finer than $(\mathbb{G}, r_{\mathbb{G}}(a))$.

Multiplicative functions and Wilson loops: Given a map $\mathbb{G} = (V, E, F)$ and a compact group G, we say that a function $h : \mathrm{P}(\mathbb{G}) \to G$ is *multiplicative* if for any pair of paths γ_1, γ_2 with $\overline{\gamma}_1 = \underline{\gamma}_2$,

$$h_{\gamma_1 \gamma_2} = h_{\gamma_2} h_{\gamma_1}. \tag{30}$$

We denote their set by $\mathcal{M}(\mathrm{P}(\mathbb{G}), G)$. Endowing it with pointwise multiplication, it is a compact group, and fixing an orientation of the edges, the evaluation on these edges defines an isomorphism

$$\mathcal{M}(\mathrm{P}(\mathbb{G}), G) \simeq G^E.$$

The Haar measure on $\mathcal{M}(\mathrm{P}(\mathbb{G}), G)$ can be identified via this isomorphism to the tensor product of the Haar measure on G; we denote it simply by dh.

When $\mathbb{G}'$ is a map finer than $\mathbb{G}$, the restriction from $\mathrm{P}(\mathbb{G}')$ to $\mathrm{P}(\mathbb{G})$ defines a map

$$\mathcal{R}^{\mathbb{G}'}_{\mathbb{G}} : \mathcal{M}(\mathrm{P}(\mathbb{G}'), G) \to \mathcal{M}(\mathrm{P}(\mathbb{G}), G).$$

https://doi.org/10.1017/fms.2024.152 Published online by Cambridge University Press

A *Wilson loop* is a function of the form

$$\begin{aligned} \mathcal{M}(\mathrm{P}(\mathbb{G}), G) &\longrightarrow \mathbb{C} \\ h &\longmapsto \chi(h_\ell), \end{aligned}$$

where $\chi : G \to \mathbb{C}$ is a function invariant by conjugation and $\ell \in \mathrm{L}(\mathbb{G})$. By centrality, the value $\chi(h_\ell)$ depends on ℓ only through its $\sim_c$-equivalence class l, and we denote it by $\chi(h_l)$. When G_N is a compact classical group, for any loop $\ell \in \mathrm{L}(\mathbb{G})$, we shall focus on the Wilson loop W_ℓ obtained considering as central function

$$\chi = \mathrm{tr}_N,$$

where $\mathrm{tr}_N = d_N^{-1}\mathrm{Tr}$ is the standard trace Tr in the natural matrix representation normalised by the size d_N of the matrix – that is, $2N$ in the symplectic case and N otherwise.

Discrete Yang–Mills measure, nonsingular case on closed surfaces: When $T > 0$, $\mathbb{G}$ is a map with boundary faces B and $a \in \Delta^o_{\mathbb{G},B}(T)$, the Yang–Mills measure is the probability measure $\mathrm{YM}_{\mathbb{G},B,a}$ on the compact group $\mathcal{M}(\mathrm{P}(\mathbb{G}), G)$ with density

$$Z^{-1}_{\mathbb{G},B,a} \prod_{f \in F \setminus B} p_{a_f}(h_{\partial f})$$

with respect to the Haar measure on $\mathcal{M}(\mathrm{P}(\mathbb{G}), G)$, where $Z_{\mathbb{G},B,a} = 1$ if $B \neq \emptyset$ and

$$Z_{\mathbb{G},a} = \int_{\mathcal{M}(\mathrm{P}(\mathbb{G}),G)} \prod_{f \in F} p_{a_f}(h_{\partial f}) dh$$

otherwise. In the above formula, ∂f is the boundary of the face for some arbitrary choice of root and orientation. This does not change the value of $p_{a_f}(h_{\partial f})$ thanks to (28) and (29). The fact that this density defines a probability measure when $B \neq \emptyset$ follows, for instance, from Lemma 3.3 below. We denote $\mathrm{YM}_{\mathbb{G},\emptyset,a}$ simply by $\mathrm{YM}_{\mathbb{G},a}$. The following lemma is standard and follows easily from the definition of the discrete Yang–Mills measure.

Lemma 3.2.

1. *For any $a \in \Delta^o_{\mathbb{G}}(T)$, the constant $Z_{\mathbb{G},a}$ depends only on T and the genus g of $\mathbb{G}$; we denote it by $Z_{g,T}$.*
2. *When $m' = (\mathbb{G}', a')$, $m = (\mathbb{G}, a)$ are two area weighted maps with m' finer than m and $a' \in \Delta^o_{\mathbb{G}'}(T)$, then*

$$\mathcal{R}^{\mathbb{G}'}_{\mathbb{G}\,*}(\mathrm{YM}_{\mathbb{G}',a'}) = \mathrm{YM}_{\mathbb{G},a}.$$

Uniform continuity and compatibility: The Yang–Mills measure is also well defined on the faces on the simplex of area vectors. For any $r, g \geq 1$ let us consider the set $\mathrm{Hom}(\Gamma_{g,r}, G)$ of group morphisms. When endowed with pointwise multiplication, it is a compact group, and thanks to the presentation of Lemma 2.3,

$$\mathrm{Hom}(\Gamma_{g,r}, G) \simeq G^{r+2g-1}.$$

Moreover, this presentation allows to write the following integration formula.

Lemma 3.3 [42]. *Assume $(\mathbb{G}, a)$ is an area weighted map with r faces, and $(\ell_i, 1 \leq i \leq r)$ and $a_1, b_1, \ldots, a_g, b_g$ are as in Lemma 2.3. For any $1 \leq i \leq r$, denote by a_i the area of the face of ℓ_i. Then*

https://doi.org/10.1017/fms.2024.152 Published online by Cambridge University Press

for any continuous function $\chi : G^{2g+r} \to \mathbb{C}$ *and any* $a \in \Delta^o_{\mathbb{G}}(T)$ *and* $1 \leq k \leq r$*,*

$$\mathbb{E}_{\mathrm{YM}_{\mathbb{G},a}}(\chi(h_{\ell_1},\ldots,h_{\ell_r},h_{a_1},\ldots,h_{b_g}))$$
$$= Z^{-1}_{g,T}\int_{G^{2g+r-1}}\chi(z_1,\ldots,z_r,x_1,\ldots,y_g)p_{a_k}(z_k)\prod_{i=1,i\neq k}^{r}p_{a_i}(z_i)dz_i\prod_{l=1}^{g}dx_l dy_l,$$

where we set $z_k = (z_1\ldots z_{k-1})^{-1}[a_1,b_1]\ldots[a_g,b_g](z_{k+1}\ldots z_r)^{-1}$*. When* B *is a nonempty subset of faces of* $\mathbb{G}$ *and lassos with faces in its complement have labels* $i_1,\ldots,i_p$*,*

$$\mathbb{E}_{\mathrm{YM}_{\mathbb{G},B,a}}(\chi(h_{\ell_{i_1}},\ldots,h_{\ell_{i_p}},h_{a_1},\ldots,h_{b_g}))$$
$$= \int_{G^{2g+p}}\chi(z_1,\ldots,z_p,x_1,\ldots,y_g)\prod_{i=1}^{p}p_{a_i}(z_i)dz_i\prod_{l=1}^{g}dx_l dy_l.$$

The above expression yields the following continuity in the area parameter. For any vertex v of a map $\mathbb{G}$, the restriction of a multiplicative function to loops based at v depends only on the $\sim_r$-class of a loop, and the restriction operation defines a map $\mathcal{R}_v : \mathcal{M}(\mathrm{P}(\mathbb{G}),G) \to \mathrm{Hom}(\mathrm{RL}_v(\mathbb{G},G))$. For any $a \in \Delta^o_{\mathbb{G}}(T)$, we set $\mathrm{YM}_{a,\mathbb{G},v} = \mathcal{R}_{v*}(\mathrm{YM}_{\mathbb{G},a})$. Using the weak convergence of the heat kernel (26), we directly deduce the following result.

Lemma 3.4. *The family of measures* $(\mathrm{YM}_{a,\mathbb{G},v}, a \in \Delta^o_{\mathbb{G}}(T))$ *on the set* $\mathrm{Hom}(\mathrm{RL}_v(\mathbb{G}),G)$ *has a weakly continuous extension to* $\Delta_{\mathbb{G}}(T)$*. It has the following properties.*

1. *Consider* $K \subset F$ *with* $K \neq F$*. Let* $S \subset \{1,\ldots,r\}$ *be the labels of the lassos with faces in* $F \setminus K$*. Then for any* $a \in \Delta^o_{K,\mathbb{G}}(T)$ *and any continuous function* $\chi : G^{2g+r} \to \mathbb{C}$*,*

$$\mathbb{E}_{\mathrm{YM}_{\mathbb{G},a,v}}(\chi(h_{\ell_1},\ldots,h_{\ell_r},h_{a_1},\ldots,h_{b_g}))$$
$$= \frac{1}{Z_{g,T}}\int_{G^{2g+r-1}}\chi(z_1,\ldots,z_r,x_1,\ldots,y_g)p_{a_k}(z_k)\prod_{i\in S,i\neq k}p_{a_i}(z_i)dz_i\prod_{l=1}^{g}dx_l dy_l,$$

   where we set $z_k = (z_1\ldots z_{k-1})^{-1}[a_1,b_1]\ldots[a_g,b_g](z_{k+1}\ldots z_r)^{-1}$ *for* $k \in S$ *arbitrary and* $z_i = 1$ *for all* $i \notin S$*.*
2. *Consider a weighted map* $(\mathbb{G}',a')$ *finer than* $(\mathbb{G},a)$ *and denote the restriction map* $\mathcal{R}^{\mathbb{G}'}_{\mathbb{G}}$: $\mathrm{Hom}(\mathrm{RL}_v(\mathbb{G}',G)) \to \mathrm{Hom}(\mathrm{RL}_v(\mathbb{G},G))$*. Then,*

$$\mathcal{R}^{\mathbb{G}'}_{\mathbb{G}\,*}(\mathrm{YM}_{\mathbb{G}',a',v}) = \mathrm{YM}_{\mathbb{G},a,v}.$$

3. *Consider* $K \subset F$ *with* $K \neq F$ *and* $a \in \Delta_{K,\mathbb{G}}(T)$*. Then for any loops* $\ell,\ell' \in \mathrm{RL}_v(\mathbb{G})$ *with* $\ell \sim_K \ell'$*,* h_ℓ *and* $h_{\ell'}$ *have same law under* $\mathrm{YM}_{\mathbb{G},a,v}$*.*

Continuous Yang–Mills measure: Thanks to the invariance by subdivision of the discrete Yang–Mills measure, given a Riemannian metric, it is possible to take the projective limit of measures defined on graphs embedded in Σ whose edges are piecewise geodesic. It allows to define a multiplicative random process $(H_\gamma)_\gamma$ indexed by all piecewise geodesic paths, whose marginals are given by the discrete Yang-Mills measure.

This was done in [42], where the author is furthermore able to show a weak convergence result allowing to define uniquely the distribution of a multiplicative function $(H_\gamma)_{\mathrm{P}(\Sigma)}$ indexed by all path of finite length. Let us recall this result.

Denote by $\mathrm{P}(\Sigma)$ the set of Lipschitz functions $\gamma : [0,1] \to \Sigma$ with speed bounded from above and from below, considered up to bi-Lipshitz re-parametrisations of $[0,1]$. The set $\mathrm{P}(\Sigma)$ is endowed with the starting and endpoint maps, $\gamma \mapsto \underline{\gamma},\overline{\gamma}$ and of the operations of concatenation and reversion as

https://doi.org/10.1017/fms.2024.152 Published online by Cambridge University Press

above. A path of Σ is an element of $\gamma \in \mathrm{P}(\Sigma)$. It is *simple* if for any parametrisation $p : [0,1] \to \Sigma$, $p : [0,1) \to \Sigma$ is injective. We consider then the set

$$\mathcal{M}(\mathrm{P}(\Sigma), G)$$

of multiplicative functions as in (30). It is a compact subset of $G^{\mathrm{P}(\Sigma)}$ when the latter is endowed with the product topology. A loop is a path $\ell \in \mathrm{P}(\Sigma)$ such that $\underline{\ell} = \overline{\ell}$. We denote their set by $\mathrm{L}(\Sigma)$. For any $x, y \in \Sigma$, we endow $\mathrm{P}_{x,y}(\Sigma) = \{\gamma \in \mathrm{P}(\Sigma) : \underline{\gamma} = x, \overline{\gamma} = y\}$ with a metric setting for any $\gamma_1, \gamma_2 \in \mathrm{P}_{x,y}(\Sigma)$,

$$d(\gamma_1, \gamma_2) = \inf_{p_1, p_2} \|p_1 - p_2\|_\infty + |\mathscr{L}(\gamma_1) - \mathscr{L}(\gamma_2)|,$$

where the infimum is taken over all parametrisations p_1, p_2 of γ_1, γ_2, and for any $\gamma \in \mathrm{P}(\Sigma)$, $\mathscr{L}(\gamma)$ denotes the Riemannian length of γ. Endowing $\mathcal{M}(\mathrm{P}(\Sigma), G)$ with the cylindrical sigma field $\mathcal{B}_{\Sigma,G}$, we denote by $(H_\gamma)_{\gamma \in \mathrm{P}(\Sigma)}$ the canonical process. When $G = G_N$ is a classical compact matrix Lie group of size N, we write for any path $\gamma \in \mathrm{P}(\Sigma)$,

$$W_\gamma = \mathrm{tr}_N(H_\gamma).$$

When $(\mathbb{G}, a)$ is an area weighted map of genus $g \geq 0$, a *Riemannian embedding* of $(\mathbb{G}, a)$ in a Riemann surface with volume vol is a collection of simple paths $(\gamma_e)_{e \in E}$ in $P(\Sigma)$ indexed by edges of $\mathbb{G}$ that do not cross but at their endpoints with the following properties:

1. The ranges of all paths $(\gamma_e)_{e \in E}$ form the 1-cells of a CW complex isomorphic to the CW complex of $\mathbb{G}$.
2. Fixing such an isomorphism, each 2-cell of the complex associated to $(\gamma_e)_{e \in E}$ is a subset of Σ of Riemannian volume $a(f)$ whenever it is identified with a face f of $\mathbb{G}$.

When Σ is the Euclidean plane or the hyperbolic disc, while $\mathbb{G}$ is a map of genus 0, f_∞ is a face of $\mathbb{G}$ and $a \in \Delta_{\mathbb{G},\{f_\infty\}}(T)$, and an embedding in Σ of the area weighted map $(\mathbb{G}, \{f_\infty\}, a)$ with one boundary component is a collection of simple paths $(\gamma_e)_{e \in E}$ in $\mathrm{P}(\mathbb{R}^2)$ indexed by edges of $\mathbb{G}$ that do not cross but at their endpoints with the following properties:

1. The ranges of all paths $(\gamma_e)_{e \in E}$ form the 1-cells of a CW complex isomorphic to the CW complex of $\mathbb{G}$, such that the unique unbounded 2-cell is mapped to f_∞.
2. Fixing such an isomorphism, each bounded 2-cell of the complex associated to $(\gamma_e)_{e \in E}$ is a subset of Σ of Riemannian volume $a(f)$ whenever it is identified with a face f of $\mathbb{G}$.

In each case, we say that $\mathbb{G}$ is embedded in Σ if there is an area vector a satisfying the property 2.

When $\mathbb{G} = (V, E, F)$ is a map, $\ell \in \mathrm{L}(\mathbb{G})$, Σ is a two-dimensional Riemannian manifold and $l \in \mathrm{L}(\Sigma)$, we say that l is a *drawing* of $\ell = e_1 \dots e_n$ if there is a Riemannian embedding $(\gamma_e)_{E \in E}$ of $\mathbb{G}$ into Σ, such that l is the concatenation $\gamma_{e_1} \dots \gamma_{e_n}$. The next two theorems are due to Lévy [42].

Theorem 3.5. *Let Σ be a compact Riemannian surface with area measure* vol*, G a fixed compact Lie group such that $\mathfrak{g}$ is endowed with a G-invariant inner product. There exists a unique measure* YM_Σ *on $(\mathcal{M}(\mathrm{P}(\Sigma), G), \mathcal{B}_\Sigma, G)$, with the following properties.*

1. *If $(\gamma_e)_{e \in E}$ is a Riemannian embedding in Σ of an area-weighted map $(\mathbb{G}, a)$ with edges E, the distribution of $(H_{\gamma_e})_{e \in E}$ is the discrete Yang–Mills measure* $\mathrm{YM}_{\mathbb{G},a}$*.*
2. *For any $x, y \in \Sigma$, if $(\gamma_n)_{n \geq 1}$ is a sequence of paths of* $\mathrm{P}_{x,y}(\Sigma)$ *with* $\lim_{n \to \infty} d(\gamma_n, \gamma) = 0$ *for some $\gamma \in \mathrm{P}(\Sigma)$, then under* YM_Σ*, the sequence of random variables $(H_{\gamma_n})_{n \geq 1}$ converges in probability to H_γ.*

The process $(H_\gamma)_{\gamma \in \mathrm{P}(\Sigma)}$ is called the Yang–Mills holonomy process.

Theorem 3.6. *Let Σ be a Euclidean plane $\mathbb{R}^2$ or the hyperbolic disc $D_{\mathfrak{h}}$, endowed with their area measure* vol*, G a fixed compact Lie group such that $\mathfrak{g}$ is endowed with a G-invariant inner product. There exists a measure* YM_Σ *on $(\mathcal{M}(\mathrm{P}(\Sigma), G), \mathcal{B}_\Sigma, G)$, with following properties.*

https://doi.org/10.1017/fms.2024.152 Published online by Cambridge University Press

1. *If* $(\gamma_e)_{e\in E}$ *is a Riemannian embedding in* Σ *of an area-weighted map of genus* 0 *with one boundary* $(\mathbb{G}, \{f_\infty\}, a)$ *and edge set E, the distribution of* $(H_{\gamma_e})_{e\in E}$ *is the discrete Yang–Mills measure* $\mathrm{YM}_{\mathbb{G},a}$.
2. *For any* $x, y \in \Sigma$, *if* (γ_n) *is a sequence of paths of* $\mathrm{P}_{x,y}(\Sigma)$ *with* $d(\gamma_n, \gamma) \underset{n\to\infty}{\to} 0$ *for some* $\gamma \in \mathrm{P}(\Sigma)$, *then under* YM_Σ, *the sequence of random variables* $(H_{\gamma_n})_{n\in\mathbb{N}}$ *converges in probability to* H_γ.

The process $(H_\gamma)_{\gamma\in\mathrm{P}(\Sigma)}$ *is called the Yang–Mills holonomy process.*

The first author showed with Cébron, Gabriel and Norris in [9, 21] that the proof of the above theorem can be adapted to yield the following extension result when G is allowed to vary. Let us denote by $A(\Sigma)$ the subset of paths of $\mathrm{P}(\Sigma)$ with a piecewise geodesic bi-Lipschitz parametrisation.

Proposition 3.7. *Let* $(G_N)_N$ *be a sequence of compact classical groups. Assume the following two properties.*

1. *For any* $\gamma \in A(\Sigma)$, $\Phi(\gamma) = \lim_{N\to\infty} W_\gamma$, *where the convergence holds in probability under* YM_Σ *and* $\Phi(\gamma)$ *is constant.*
2. *There is a constant* $K > 0$ *independent of* N, *such that for any simple contractible loop* $\ell \in \mathrm{L}(\Sigma)$ *bounding an area* $t > 0$,

$$\mathbb{E}_{\mathrm{YM}_\Sigma}[1 - \Re(W_\ell)] \leq Kt.$$

Then $\Phi : A(\Sigma) \to \mathbb{C}$ *has a unique extension to* $\mathrm{P}(\Sigma)$ *such that for all* $x, y \in \Sigma$, $\Phi : \mathrm{P}_{x,y}(\Sigma) \to \mathbb{C}$ *is continuous, and for any* $\gamma \in \mathrm{L}(\Sigma)$, W_γ *converges in probability towards* $\Phi(\gamma)$ *as* $N \to \infty$.

The argument given in Section 5 of [21] for the sphere applies verbatim on any compact surface Σ to yield the above statement; we will not repeat it in the current version. The same applies for the following lemma.

Lemma 3.8.

1. *For any map* $\mathbb{G}$, *there is a regular map* $\mathbb{G}'$ *finer than* $\mathbb{G}$.
2. *For any* $\gamma \in A(\Sigma)$, *there is a graph* $\mathbb{G}$ *with Riemannian embedding in* Σ, *such that* γ *is the drawing of a path of* $\mathbb{G}$.
3. *For any area weighted map* $(\mathbb{G}, a)$ *and* $\gamma \in \mathrm{P}(\mathbb{G})$, *there is a regular area weighted map* $(\mathbb{G}', a')$ *finer than* $(\mathbb{G}, a)$, γ' *a regular path of* $\mathbb{G}'$ *and K a subset of faces of* $\mathbb{G}'$, *such that*

$$\gamma \sim_K \gamma'.$$

4. *For any compact Lie group G and any* $\gamma \in A(\Sigma)$, *there is a regular path* $\mathfrak{p}$ *in a regular map* $\mathbb{G}$ *and* $a \in \Delta_{\mathbb{G}}(T)$ *such that under* YM_Σ, W_γ *has same law as* $W_{\mathfrak{p}}$ *under* $\mathrm{YM}_{\mathbb{G},a}$.

Together with the last proposition, this lemma reduces the study of Wilson loops for all loops of finite length to the case of regular loops.

3.3. Planar master field, main results and conjecture

In the above setting, the following was proved in [43] and[40] [33]; see [72, 3] for a weaker statement with a smaller class of loops and of groups G_N. Recall the definitions of the desingularisation operation in Section 2.7, of Makeenko-Migdal vectors (18) and of area weighted maps in Section 3.2.

Theorem 3.9. *Assume that* G_N *is a compact classical group of rank N. Assume that* $(\mathbb{G}, \{f_\infty\}, a)$ *is any area weighted map of genus* 0, *with one boundary component and* $\ell \in \mathrm{L}(\mathbb{G})$, *or that* $l \in \mathrm{L}(\mathbb{R}^2)$.

[40] In [43], to get uniqueness, (b) is replaced by an additional set of differential equations

https://doi.org/10.1017/fms.2024.152 Published online by Cambridge University Press

Then the following convergences hold in probability,[41] *and the limits are constant and independent of the type of series of* G_N*:*

$$\Phi_\ell^{f_\infty}(a) = \lim_{N\to\infty} W_\ell \text{ under } \mathrm{YM}_{\mathbb{G},\{f_\infty\},a}$$

and

$$\Phi_{\mathbb{R}^2}(l) = \lim_{N\to\infty} W_l \text{ under } \mathrm{YM}_{\mathbb{R}^2}.$$

The function $\Phi_{\mathbb{R}^2}$ *is characterised by the following properties:*

1. *For any* $x \in \mathbb{R}^2$, $\Phi_{\mathbb{R}^2} : \mathrm{P}_{x,x}(\mathbb{R}^2) \to \mathbb{C}$ *is continuous.*
2. *Whenever* $l \in \mathrm{L}(\mathbb{R}^2)$ *is a drawing of a loop* ℓ *of an area weighted map of genus* 0 *with one boundary component* $(\mathbb{G}, \{f_\infty\}, a)$,

$$\Phi_{\mathbb{R}^2}(l) = \Phi_{\ell,f_\infty}(a).$$

3. *For any map of genus* 0 *with one boundary component* $(\mathbb{G}, \{f_\infty\})$, $T > 0$, *and any loop* $\ell \in \mathrm{L}(\mathbb{G})$, Φ_ℓ *is uniformly continuous on* $\Delta_{\mathbb{G},\{f_\infty\}}(T)$ *and differentiable on* $\Delta^o_{\mathbb{G},\{f_\infty\}}(T)$ *such that*
   (a) *if* $\mathbb{G}$ *is regular,* ℓ *is a tame loop and* $v \in V_\ell$ *is a transverse intersection with* $\delta_v \ell = \ell_1 \otimes \ell_2$,

$$\mu_v.\Phi_{\ell,f_\infty} = \Phi_{\ell_1,f_\infty}\Phi_{\ell_2,f_\infty} \text{ in } \Delta^o_{\mathbb{G},\{f_\infty\}}(T).$$

   (b) *Whenever* l *is the boundary of a topological disc of area t,*

$$\Phi_{\mathbb{R}^2}(l) = e^{-\frac{t}{2}}.$$

See the appendix of [43] for a table of values of $\Phi_{\mathbb{R}^2}$. Alternatively, the master field can be characterised using free probability as follows. For any real $t \geq 0$ and any integer $n \geq 0$, set

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

It is known since the work of Biane [6] that these quantities are related to the limits of the moments of Brownian motions on $\mathrm{U}(N)$, and Lévy proved in [43] that it is still the case for the other compact classical matrix Lie groups. Let us reformulate slightly a result of [43] in our notations.

Lemma 3.10 (Proposition 6.1.2 of [43]). *Consider an area weighted map* $(\mathbb{G}, \{f_\infty\}, a)$ *of genus* 0 *with one boundary component. Assume* $\mathbb{G} = (V, E, F)$, $\#F = r + 1$, $F = \{f_1, \dots, f_r, f_\infty\}$ *and* $v \in V$. *For any* $\ell \in \mathrm{L}_v(\mathbb{G})$, $\Phi_{\ell,f_\infty}(a)$ *depends on* ℓ *only through its* $\sim_r$ *class. Setting*

$$\tau_v(\ell) = \Phi_{\ell,f_\infty}(a) \text{ and } \ell^* = \ell^{-1},\ \forall \ell \in \mathrm{RL}_v(\mathbb{G})$$

and extending these maps linearly and sesquilinearly defines a noncommutative probability space $(\mathbb{C}[\mathrm{RL}_v(\mathbb{G})], \tau_v, *)$. *Assume that* $\ell_1, \dots, \ell_r, \ell_\infty$ *is a family of lassos as in Lemma 2.3 with* ℓ_i *bounding* f_i *for* $1 \leq i \leq r$ *and* ℓ_∞ *for* f_∞. *Then* τ_v *is the unique state on* $(\mathbb{C}[\mathrm{RL}_v(\mathbb{G})], *)$ *such that*

1. *for all* $n \in \mathbb{Z}^*$, $\tau_v(\ell_i^n) = \nu_{a(f_i)}(n)$,
2. $\ell_1, \dots, \ell_r$ *are freely independent under* τ_v.

Similarly, the following lemma follows from the classical result of [6] and Lemma 3.3. It shows that the conclusion of the former one is valid when the genus condition is dropped.

[41] It is also shown in [43] that the following convergences are almost sure.

https://doi.org/10.1017/fms.2024.152 Published online by Cambridge University Press

Lemma 3.11. *Consider an area weighted regular map of genus $g \geq 1$ and with boundary $(\mathbb{G}, \{f_\infty\}, a)$. Assume $\mathbb{G} = (V, E, F)$, $\#F = r + 1$ with $F = \{f_1, \ldots, f_r, f_\infty\}$ and $v \in V$. Assume that $a_1, \ldots, b_g$ and $\ell_1, \ldots, \ell_{r+1}$ are $2g$ simple loops and $r + 1$ lassos as in Lemma 2.3, with ℓ_i bounding f_i for $1 \leq i \leq r$ and f_∞ for $i = r + 1$. Assume that G_N is a sequence of compact classical matrix Lie groups of size N. Then for any $T > 0$, $a \in \Delta_{\mathbb{G},\{f_\infty\}}(T)$ and $\ell \in \mathrm{RL}_v(\mathbb{G})$,*

$$W_\ell \to \Phi_\ell^{1,g}(a) \text{ under } \mathrm{YM}_{\mathbb{G},\{f_\infty\},a},$$

where $\Phi_\ell^{1,g}(a)$ is constant. Moreover, there is a constant $K > 0$ independent of $\mathbb{G}$ and $N \geq 1$, such that for any face $f \in F \setminus \{f_\infty\}$,

$$\mathbb{E}[1 - \Re(W_{\partial f})] \leq K a(f). \tag{*}$$

The $*$-algebra $(\mathbb{C}[\mathrm{RL}_v(\mathbb{G})], *)$ is endowed with a unique state τ_v satisfying

$$\tau_v(\ell) = \Phi_\ell^{1,g}(a), \ \forall \ell \in \mathrm{RL}_v(\mathbb{G}).$$

Moreover, τ_v is characterised by the following three properties:

1. *$\ell_1, \ldots, \ell_r, a_1, \ldots, b_g$ are freely independent under τ_v.*
2. *under τ_v, $a_1, \ldots, b_g$ are $2g$ Haar unitaries.*
3. *for any $1 \leq i \leq r$ and $n \in \mathbb{Z}^*$,*

$$\tau_v(\ell_i^n) = \nu_{a(f_i)}(n).$$

A sketch of the proof is given in Section 5.

From Lemma 3.11 and the absolute continuity result of [20] follows Corollary 1.4 for loops avoiding at least one handle. Let us give now a discrete reformulation of Corollary 1.4. Its proof is given below in Section 5. Let us recall the definition of the universal cover $\tilde{\mathbb{G}} = (\tilde{V}, \tilde{E}, \tilde{F})$ of a regular map $(\mathbb{G}, \mathbb{G}_b)$ given in Section 2.3, with a canonical covering map $p : \tilde{F} \to F$. When $a \in \Delta_{\mathbb{G}}(T)$, let us set $\tilde{a} = a \circ p : \tilde{F} \to [0, T]$.

Theorem 3.12. *Assume that $(\mathbb{G}, a)$ is an area weighted map cut along a simple loop $\ell \in \mathrm{L}(\mathbb{G})$ given by $(\mathbb{G}_1, \{f_{1,\infty}\})$ and $(\mathbb{G}_2, \{f_{2,\infty}\})$, with the same convention as in Section 2.1. Assume that $\mathbb{G}_2$ has genus $g_2 \geq 1$. Then, for any loop $\ell \in \mathrm{L}(\mathbb{G}_1)$ and $a \in \Delta_{\mathbb{G}}(T)$ with $0 < \sum_{f \in F_2} a(f) < T$,*

$$W_\ell \underset{N\to\infty}{\to} \Phi_\ell(a) = \begin{cases} \Phi_{\tilde{\ell}}(\tilde{a}) & \text{if } \ell \sim_h c_{\underline{\ell}}, \\ 0 & \text{if } \ell \nsim_h c_{\underline{\ell}}, \end{cases} \quad \text{in probability under } \mathrm{YM}_{\mathbb{G},a}, \tag{31}$$

where $\tilde{\ell}$ is a lift of ℓ in $\tilde{\mathbb{G}}$. Moreover, when $g_2 \geq 2$, the convergence holds true uniformly in $a \in \Delta_{\mathbb{G}}(T)$. Besides, there is a constant $K > 0$ independent of $\mathbb{G}$ and $N \geq 1$, and depending only on $a(F_2) \in (0, T)$ such that for any face $f \in F_1$,

$$\mathbb{E}[1 - \Re(W_{\partial f})] \leq K a(f). \tag{32}$$

When $\mathbb{G}$ has genus 1, the above result gives information about loops included in a topological disc but does not say anything about other loops – for instance, contractible loops obtained by concatenation of simple loops of nontrivial homology. A more satisfying answer is then given by the following theorem.

Theorem 3.13. *Consider a compact classical group G_N of rank N, a torus $\mathbb{T}_T$ of volume $T > 0$ obtained as a quotient of the Euclidean plane $\mathbb{R}^2$ by the lattice $\sqrt{T}\mathbb{Z}^2$. Then, the following convergence holds in probability under $\mathrm{YM}_{\mathbb{T}_T}$:*

$$W_l \underset{N\to\infty}{\longrightarrow} \Phi_{\mathbb{T}_T}(l) = \begin{cases} \Phi_{\mathbb{R}^2}(\tilde{l}) & \text{if } l \text{ is contractible,} \\ 0 & \text{otherwise,} \end{cases}$$

where for any loop $l \in \mathrm{L}(\mathbb{T}_T)$, $\tilde{l} \in \mathrm{P}(\mathbb{R}^2)$ is a finite length path with projection to $\mathbb{T}_T$ given by l. Besides, $\Phi_{\mathbb{T}_T} : \mathrm{L}(\mathbb{T}_T) \to \mathbb{C}$ is the unique function satisfying

1. *For any $x \in \mathbb{T}_T$, $\Phi_{\mathbb{T}_T} : \mathrm{L}_x(\mathbb{T}_T) \to \mathbb{C}$ is continuous for the length metric d.*
2. *For any regular loop ℓ in a regular map $\mathbb{G}$ of genus 1, there is a differentiable function*

$$\Phi_\ell : \Delta_{\mathbb{G}}(T) \to \mathbb{C}$$

such that for any transverse intersection $v \in V_\ell$, with $\delta_v \ell = \ell_1 \otimes \ell_2$,

$$\mu_v.\Phi_\ell = \Phi_{\ell_1}\Phi_{\ell_2}, \tag{33}$$

and such that $\Phi_\ell(a) = \Phi_{\mathbb{T}_T}(l)$ whenever $l \in \mathrm{L}(\mathbb{T}_T)$ is a drawing of ℓ with associated area measure on $\mathbb{T}_T$ given by a.

3. *For any loop $l \in \mathrm{L}(\mathbb{T}_T)$ obtained by projection of a loop $\tilde{l} \in \mathrm{L}(\mathbb{R}^2)$ included in a fundamental domain of $\mathbb{T}_T$,*

$$\Phi_{\mathbb{T}_T}(l) = \Phi_{\mathbb{R}^2}(\tilde{l}).$$

4. *For any noncontractible simple loop $l \in \mathrm{L}(\mathbb{T}_T^2)$ and $n \in \mathbb{Z}^*$,*

$$\Phi_{\mathbb{T}_T}(l^n) = 0.$$

When $g \geq 2$, we are unable to show a satisfying version of Conjecture 1.3, but we are able to prove the following conditional results.

Theorem 3.14. *Consider a compact classical group G_N of rank N, $g \geq 2$ and $T > 0$. Assume that for any regular area weighted map $(\mathbb{G}, a)$ of genus g,*

$$W_\ell \underset{N\to\infty}{\longrightarrow} \Phi_{\tilde{\ell}}(\tilde{a}) \text{ in probability under } \mathrm{YM}_{\mathbb{G},a}, \tag{34}$$

whenever $\ell \in \mathrm{L}(\mathbb{G})$ such that

1. *any lift $\tilde{\ell} \in \mathrm{L}(\tilde{\mathbb{G}})$ of ℓ is included in a fundamental domain, or*
2. *$\ell = \gamma_{nest}\gamma$, where γ_{nest} is a nested loop and γ is a geodesic path.*[42]

Then for any regular map $\mathbb{G}$ of genus g, (34) holds true for all $\ell \in \mathrm{L}(\mathbb{G})$.

Besides, the following weaker statement can be proved independently.

Proposition 3.15. *Consider a compact classical group G_N of rank N and $g \geq 2$. Assume that for any regular area weighted map $(\mathbb{G}, a)$ of genus g,*

$$W_\ell \underset{N\to\infty}{\longrightarrow} 0 \text{ in probability under } \mathrm{YM}_{\mathbb{G},a}, \tag{35}$$

whenever $\ell \in \mathrm{L}(\mathbb{G})$ is a geodesic loop with non zero-homology. Then for any regular map $\mathbb{G}$ of genus g, (35) holds true for all $\ell \in \mathrm{L}(\mathbb{G})$ with nonzero homology.

[42]See Figure 14 and Section 2.3.

https://doi.org/10.1017/fms.2024.152 Published online by Cambridge University Press

Remark 3.16. The above statements may give the impression that any possible master field is expressed in terms of the planar case. This is nonetheless not the case as the Wilson loops on the sphere converge to different limits [21]. See also the discussion in [20, Section 2.5].

The proofs of Theorems 3.13, 3.14 and Proposition 3.15 are provided in the end of Section 3.5.

3.4. Invariance in law and Wilson loop expectation

Before proceeding to the main part of this paper, let us give a partial result that only holds in expectation but relies on a simpler argument: the invariance in law by an action of the center of the structure group G_N. Consider a regular map $\mathbb{G} = (V, E, F)$ with r faces, $v \in V$ and a basis $\ell_1, \ldots, \ell_r, a_1, \ldots, b_g$ of the free group $\mathrm{RL}_v(\mathbb{G})$ as in Lemma 2.3. For any $h \in G^{2g}$ and $\phi \in \mathrm{Hom}(\mathrm{RL}_v(\mathbb{G}), G)$, let us denote by $h.\phi \in \mathrm{Hom}(\mathrm{RL}_v(\mathbb{G}), G)$ the unique group morphism with

$$h.\phi(\ell_i) \mapsto \phi(\ell_i), \text{ for } 1 \le i \le r$$

and

$$h.\phi(a_i) = h_{2i-1}\phi(a_i) \text{ and } h.\phi(b_i) = h_{2i}\phi(b_i) \text{ for } 1 \le i \le g.$$

Let us denote by Z the center of G. When $h \in Z^{2g}$, it follows easily from point 2 of Lemma 2.12 that

$$h.\phi(\ell) = \phi_h([\ell]_{\mathbb{Z}})\phi(\ell), \ \forall \ell \in \mathrm{RL}_v(\mathbb{G}), \tag{36}$$

where $\phi_h \in \mathrm{Hom}(H_1(d^*, \mathbb{Z}), Z)$ is the unique group morphism such that

$$\phi_h([a_i]_{\mathbb{Z}}) = h_{2i-1} \text{ and } \phi_h([b_i]_{\mathbb{Z}}) = h_{2i} \text{ for } 1 \le i \le g.$$

Lemma 3.17. *Let $\mathbb{G}$ be regular map, $T > 0$, $a \in \Delta_{\mathbb{G}}(T)$. Denoting by $(H_\ell)_{\ell \in \mathrm{RL}_v(\mathbb{G})}$ the canonical G-valued random variable on $\mathrm{Hom}(\mathrm{RL}_v(\mathbb{G}), G)$, the following assertions hold true.*

1. *The measure $\mathrm{YM}_{a,\mathbb{G},v}$ on $\mathrm{Hom}(\mathrm{RL}_v(\mathbb{G}), G)$ is invariant under the action of Z^{2g}.*
2. *Assume that $\chi : G \to \mathbb{C}$ is continuous and $\alpha : Z \to \mathbb{C}$ is such that $\chi(z.h) = \alpha_\chi(z)\chi(h)$, $\forall (z, h) \in Z \times G$. Then*
   (a) *for any $h \in Z^{2g}$ and $\ell \in \mathrm{RL}_v(\mathbb{G})$,*
   $$\mathbb{E}_{\mathrm{YM}_{a,\mathbb{G},v}}[\chi(H_\ell)] = \alpha_\chi \circ \phi_h([\ell]_{\mathbb{Z}})\mathbb{E}_{\mathrm{YM}_{a,\mathbb{G},v}}[\chi(H_\ell)].$$
   (b) *if there is $\phi \in \mathrm{Hom}(H_1(d^*, \mathbb{Z}), Z)$ with $\phi([\ell]_{\mathbb{Z}}) \neq 0$, then*
   $$\mathbb{E}_{\mathrm{YM}_{a,\mathbb{G},v}}[\chi(H_\ell)] = 0.$$
3. *When G is a classical compact matrix Lie group, for any $\ell \in \mathrm{RL}_v(\mathbb{G})$, $\mathbb{E}[W_\ell] = 0$ if one of the following conditions is satisfied:*
   (a) *$G = \mathrm{U}(N)$ and $[\ell]_{\mathbb{Z}} \neq 0$.*
   (b) *$G = \mathrm{SU}(N)$ and $[\ell]_{\mathbb{Z}_n} \neq 0$*
   (c) *$G = \mathrm{SO}(2N)$ and $[\ell]_{\mathbb{Z}_2} \neq 0$.*

Proof. The implication $2.a) \Rightarrow 2.b) \Rightarrow 3$ are elementary. Thanks to (36), $1 \Rightarrow 2.a)$. Lastly, consider 1. Denote by $d\phi$ the Haar measure on $\mathrm{Hom}(\mathrm{RL}_v(\mathbb{G}), G)$ endowed with pointwise multiplication, and

$$R_g : (\ell_1, \ldots, \ell_r, x_1, y_1, \ldots, x_g, y_g) \mapsto [x_1, y_1] \cdots [x_g, y_g]$$

the word map on $\mathrm{RL}_v(\mathbb{G})$ corresponding to the commutator of the generators of the fundamental group. By Lemma 3.4, it is enough to consider $a \in \Delta^o_{\mathbb{G}}(T)$ and denote by $a_1, \ldots a_r$ the area enclosed by the

https://doi.org/10.1017/fms.2024.152 Published online by Cambridge University Press

meanders of $\ell_1, \dots, \ell_r$ and set $a_{r+1} = T - \sum_{i=1}^r a_i$. For any continuous function $\chi : \mathrm{Hom}(\mathrm{RL}_v(\mathbb{G}), G) \to \mathbb{C}$ and $h \in Z^{2g}$, $d\phi$ is invariant by the action of Z^{2g} and

$$\begin{aligned}
&\int_{\mathrm{Hom}(\mathrm{RL}_v(\mathbb{G}),G)} \chi(h^{-1}.\phi) d\mathrm{YM}_{a,\mathbb{G},v}(\phi) \\
&\quad = \int_{\mathrm{Hom}(\mathrm{RL}_v(\mathbb{G}),G)} \chi(h^{-1}.\phi) p_{a_{r+1}}(\phi((\ell_1 \dots \ell_r)^{-1} R_g)) \prod_{i=1}^r p_{a_i}(\phi(\ell_i)) d\phi \\
&\quad = \int_{\mathrm{Hom}(\mathrm{RL}_v(\mathbb{G}),G)} \chi(\phi) p_{a_{r+1}}(h.\phi((\ell_1 \dots \ell_r)^{-1} R_g)) \prod_{i=1}^r p_{a_i}(h.\phi(\ell_i)) d\phi \\
&\quad = \int_{\mathrm{Hom}(\mathrm{RL}_v(\mathbb{G}),G)} \chi(\phi) p_{a_{r+1}}(h.\phi((\ell_1 \dots \ell_r)^{-1} R_g)) \prod_{i=1}^r p_{a_i}(h.\phi(\ell_i)) d\phi,
\end{aligned}$$

where in the last line we used that $h.\phi([x_i, y_i]) = [\phi(x_i)h_{2i-1}, \phi(y_i)h_{2i-1}]) = \phi([x_i, y_i])$ for $1 \le i \le g$ and $h.\phi(\ell_j) = \phi(\ell_j)$, for $1 \le j \le r$. □

3.5. Makeenko–Migdal equations, existence and uniqueness problem

The main tool of the current article is approximate versions of equations (33), satisfied on any surface when G_N is a compact classical group and $N \to \infty$. Let us introduce a setting to prove existence and uniqueness of these equations.

For any regular map $\mathbb{G}$ and any vertex v of $\mathbb{G}$, let $\mathcal{A}_v(\mathbb{G})$ be the algebra with elements in $\mathbb{C}[\mathrm{L}_v(\mathbb{G})]$ endowed with the multiplication given by concatenation, with unit 1_v, or simply 1, given by the constant loop at v, and setting $\ell^* = \ell^{-1}$ for all $\ell \in \mathrm{L}_v(\mathbb{G})$ and extending it skew-linearly. When w is another vertex, $\mathcal{A}_{v,w}(\mathbb{G})$ denotes the usual tensor product of the $*$-algebras $\mathcal{A}_v(\mathbb{G})$ and $\mathcal{A}_w(\mathbb{G})$. Its elements belong to $\mathbb{C}[\mathrm{L}_v(\mathbb{G})] \otimes \mathbb{C}[\mathrm{L}_w(\mathbb{G})]$, and multiplication and $*$-operation are defined for all $(x_i, y_i) \in \mathcal{A}_v(\mathbb{G}) \times \mathcal{A}_w(\mathbb{G})$ by

$$(x_1 \otimes y_1).(x_2 \otimes y_2) = (x_1 x_2) \otimes (y_1 y_2) \text{ and } (x_1 \otimes y_1)^* = x_1^* \otimes y_1^*.$$

Let us fix $g \ge 1$ and $T > 0$. Recall the notion of refinement for area weighted maps in Section 3.2. A *Wilson loop system* is a family of continuous functions $\phi_{\ell_1}, \phi_{\ell_1 \otimes \ell_2} : \Delta_{\mathbb{G}}(T) \to \mathbb{C}$ given for each map $\mathbb{G}$ of genus g and each pair of loops $\ell_1, \ell_2 \in \mathrm{L}(\mathbb{G})$, with the following properties:

1. For any constant loop c,

$$\phi_{\ell_1 \otimes c} = \phi_{\ell_1} \text{ and } \phi_c = 1.$$

2. For any pair of loops ℓ_1, ℓ_2 within a same map of genus g,

$$\phi_{\ell_1 \otimes \ell_2} = \phi_{\ell_2 \otimes \ell_1}$$

   depend on ℓ_1, ℓ_2 only through their $\sim_{r,c}$ equivalence class.
3. If $\mathbb{G}'$ is finer than $\mathbb{G}$ of genus g, then for all loops $\ell, \ell_1, \ell_2 \in \mathrm{L}(\mathbb{G})$,

$$\phi_\ell \circ r_{\mathbb{G}}^{\mathbb{G}'} = \phi_\ell \text{ and } \phi_{\ell_1 \otimes \ell_2} \circ r_{\mathbb{G}}^{\mathbb{G}'} = \phi_{\ell_1 \otimes \ell_2},$$

   where loops are identified in the right-hand sides with elements of $\mathrm{L}(\mathbb{G}')$.
4. If $\mathbb{G}'$ is isomorphic to $\mathbb{G}$ of genus g, $a \in \Delta_{\mathbb{G}}(T)$ is mapped to $a' \in \Delta_{\mathbb{G}'}(T)$, while $\ell_1', \ell_2' \in \mathrm{L}(\mathbb{G})$ with $\ell_1 \sim_\Sigma \ell_1', \ell_2 \sim_\Sigma \ell_2'$, through the same isomorphism map, then

$$\phi_{\ell_1}(a) = \phi_{\ell_1'}(a') \text{ and } \phi_{\ell_1 \otimes \ell_2}(a) = \phi_{\ell_1' \otimes \ell_2'}(a').$$

https://doi.org/10.1017/fms.2024.152 Published online by Cambridge University Press

5. If $\mathbb{G} = (V, E, F)$ is a map of genus g, $\ell_1, \ell_1', \ell_2 \in \mathrm{L}(\mathbb{G})$, $K \subset F$ with $\ell_1 \sim_K \ell_1'$, then

$$\phi_{\ell_1 \otimes \ell_2}(a) = \phi_{\ell_1' \otimes \ell_2}(a),\ \forall a \in \Delta_{K,\mathbb{G}}(T).$$

6. For any map $\mathbb{G}$ of genus g with vertex v, for any $a \in \Delta_{\mathbb{G}}(T)$, extending $\ell \in \mathrm{L}_v(\mathbb{G}) \mapsto \phi_\ell(a)$ linearly defines a nonnegative state $\phi_{a,v}$ on $(\mathcal{A}_v(\mathbb{G}), 1_v, *)$, while for any $x \in \mathcal{A}_v(\mathbb{G})$,

$$\phi_{x \otimes x^*} \geq 0.$$

Our motivation for considering the above definition[43] is the following example.

Example 3.18. Whenever G_N is a compact classical group, from the above definition of the Yang-Mills measure, the collection

$$a \in \Delta_{\mathbb{G}}(T) \mapsto (\mathbb{E}_{\mathrm{YM}_{\mathbb{G},a}}[W_\ell], \mathbb{E}_{\mathrm{YM}_{\mathbb{G},a}}[W_{\ell_1} W_{\ell_2}])$$

for all regular maps $\mathbb{G}$ of genus g and loops $\ell, \ell_1, \ell_2 \in \mathrm{L}(\mathbb{G})$ is a Wilson loop system.

Recall that it follows from the first part of 6 that

$$\phi_{x^* y} = \overline{\phi_{y^* x}} \tag{37}$$

for all $x, y \in \mathcal{A}_v(\mathbb{G})$ and from 1 that ℓ has a unitary distribution in $(\mathcal{A}_v(\mathbb{G}), 1, *, \phi_{v,a})$, for any vertex v and $\ell \in \mathrm{L}_v(\mathbb{G})$; that is,

$$\phi_{v,a}((\ell\ell^* - 1)(\ell\ell^* - 1)^*) = \phi_{v,a}((\ell^*\ell - 1)(\ell^*\ell - 1)^*) = 0.$$

When ϕ is a Wilson loop system, for any map $\mathbb{G}$ and any loop $\ell \in \mathrm{L}(\mathbb{G})$, (37), the second part of point 6 and point 1 yield[44]

$$\mathscr{V}_{\phi,\ell} = \phi_{\ell \otimes \ell^{-1}} - |\phi_\ell|^2 = \phi_{\ell \otimes \ell^{-1}} - \phi_\ell \phi_{\ell^{-1}} \geq 0,$$

since for any $a \in \Delta_{K,\mathbb{G}}(T)$, $\mathscr{V}_{\phi,\ell}(a) = \phi_{x_\ell(a) \otimes x_\ell(a)^*}(a)$, where $x_\ell(a) = \ell - \phi_\ell(a) 1_v$.

Recall definition (21). We say that a Wilson loop system ϕ is an *exact solution* of Makeenko–Migdal equations if

1. For any tame loop ℓ within map $\mathbb{G}$ of genus g, $\phi \in C^1(\Delta^o_{\mathbb{G}}(T))$ and for any $v \in V_\ell$,

$$\mu_v.\phi_\ell = \phi_{\delta_v \ell}.$$

2. For any pair of regular loops within the same map, $\phi_{\alpha \otimes \beta} = \phi_\alpha \phi_\beta$.
3. For any regular loop ℓ with $\ell \nsim_h c_{\underline{\ell}}$, $\phi_\ell = 0$.

We say that a sequence $(\phi^N)_{N \geq 1}$ of Wilson loop systems is an *approximate solution of Makeenko–Migdal equations* if for any map $\mathbb{G}$ of genus g, any loop ℓ in $\mathrm{L}(\mathbb{G})$, ϕ^N_ℓ and $\mathscr{V}^N_{\phi,\ell}$ are in $C^1(\Delta^o_{\mathbb{G}}(T))$, there is a constant $C > 0$ independent of ℓ and $N \geq 1$, such that for any intersection point $v \in V_\ell$,

$$|\mu_v.\phi^N_\ell - \phi^N_{\delta_v(\ell)}| \leq \frac{C}{N}, \tag{38}$$

$$|\mu_v.\mathscr{V}_{\phi^N,\ell}| \leq \mathscr{V}_{\phi^N,\ell} + \mathscr{V}_{\phi^N,\ell_1} + \mathscr{V}_{\phi^N,\ell_2} + \frac{C}{N} \tag{39}$$

[43] Similar functions associated to the next example have been used in [21, Sect. 4.3].
[44] This function is the variance of a Wilson loop in the canonical example of a Wilson loop system associated to the Yang-Mills measure.

https://doi.org/10.1017/fms.2024.152 Published online by Cambridge University Press

and

$$|\mu_v.\mathscr{V}_{\phi^N,\ell}| \le \sqrt{\mathscr{V}_{\phi^N,\ell_1}\mathscr{V}_{\phi^N,\ell_2}} + |\phi^N_{\ell_1}|\sqrt{\mathscr{V}_{\phi^N,\ell_2}} + |\phi^N_{\ell_2}|\sqrt{\mathscr{V}_{\phi^N,\ell_1}} + \frac{C}{N}, \tag{40}$$

where $\ell_1 \otimes \ell_2 = \delta_v \ell$.

Remark 3.19. Note that it follows from point 3 that if ϕ is a Wilson loop system and ℓ, ℓ_1, ℓ_2 are a regular loops of a map $\mathbb{G} = (V, E, F)$ with $e \in E^o \setminus (E^o_\ell \cup E^o_{\ell_1} \cup E^o_{\ell_2})$, then

$$d\omega_e.\phi_\ell = d\omega_e.\phi_{\ell_1\otimes\ell_2} = 0. \tag{41}$$

Consequently, for any regular loop ℓ, using the same linear forms as in Section 2.7, if ϕ^∞ and (ϕ^N) are, respectively, exact and approximate solutions of Makeenko–Migdal equations, for any regular loop ℓ and $X \in \mathfrak{m}_\ell$,

$$X.\phi^\infty_\ell = \phi^\infty_{\delta_X \ell} \text{ and } |X.\phi^N_\ell - \phi^N_{\delta_X\ell}| \le \frac{\|X\|C}{N}, \tag{42}$$

while

$$|X.\mathscr{V}_{\phi^N,\ell}| \le C\|X\|\left(\sum_{v\in V_\ell}\left(\sqrt{\mathscr{V}_{\phi^N,\ell_1}\mathscr{V}_{\phi^N,\ell_2}} + |\phi^N_{\ell_1}|\sqrt{\mathscr{V}_{\phi^N,\ell_2}} + |\phi^N_{\ell_2}|\sqrt{\mathscr{V}_{\phi^N,\ell_1}}\right) + \frac{1}{N}\right) \tag{43}$$

and

$$|X.\mathscr{V}_{\phi^N,\ell}| \le \|X\|C\left(\mathscr{V}_{\phi^N,\ell} + \sum_{v\in V_\ell}(\mathscr{V}_{\phi^N,\ell_{v,1}} + \mathscr{V}_{\phi^N,\ell_{1,2}}) + \frac{1}{N}\right), \tag{44}$$

where for any $v \in V_\ell$, we wrote $\partial_v \ell = \ell_{1,v} \otimes \ell_{2,v}$.

The existence problem of these equations is a consequence of [25] and [43] for the approximate solutions and, given Theorem 3.9, of a simple computation for the exact ones.

Lemma 3.20. *Consider $g \ge 1, T > 0$.*

1. *Assume that G_N is a compact classical group of rank N. Then setting for all map $\mathbb{G}$, $a \in \Delta_{\mathbb{G}}(T)$ and all loops $\ell, \ell_1, \ell_2 \in \mathrm{L}(\mathbb{G})$,*

$$\phi^N_\ell(a) = \mathbb{E}_{\mathrm{YM}_{\mathbb{G},a}}[W_\ell],\ \phi^N_{\ell_1\otimes\ell_2}(a) = \mathbb{E}_{\mathrm{YM}_{\mathbb{G},a}}[W_{\ell_1}W_{\ell_2}]$$

   defines an approximate solution of the Makeenko–Migdal equations.
2. *Denoting by c_v the constant loop at a vertex v, setting for any map $\mathbb{G}$, $a \in \Delta_{\mathbb{G}}(T)$ and $\ell \in \mathrm{L}(\mathbb{G})$,*

$$\phi_\ell(a) = \begin{cases} \Phi_{\tilde\ell}(\tilde a) & \text{if } \ell \sim_h c_{\underline{\ell}}, \\ 0 & \text{if } \ell \nsim_h c_{\underline{\ell}}, \end{cases}$$

   defines an exact solution of the Makeenko–Migdal equations.

Proof. Point 1 is a direct consequence of Proposition 7.5 below, together with Cauchy–Schwarz or arithmetic-geometric mean inequality to get (43) and (44). For point 2, we shall only check that the Makeenko–Migdal equations are satisfied and leave the other points to the reader. Consider a map $\mathbb{G}$ of genus g with $\ell \in \mathrm{L}(\mathbb{G})$ and $v \in V_\ell$. Consider $\delta_v \ell = \ell_1 \otimes \ell_2$ and let us show that $\mu_v \phi_\ell = \phi_{\ell_1}\phi_{\ell_2}$. If $\ell \nsim_h c_{\underline{\ell}}$, then the rerooting ℓ' at v of ℓ satisfies $\ell' \nsim_h c_v$. Therefore, $\ell_1 \nsim_h c_v$ or $\ell_2 \nsim_h c_v$, and we

https://doi.org/10.1017/fms.2024.152 Published online by Cambridge University Press

conclude that $\phi_\ell = \phi_{\ell'} = 0 = \phi_{\ell_1}\phi_{\ell_2}$. Assume now $\ell \sim_h c_{\underline{\ell}}$. Consider the universal cover $\tilde{\mathbb{G}} = (\tilde{V}, \tilde{E}, \tilde{F})$ of $\mathbb{G}$ with projection map p. For all $a \in \Delta^o_{\mathbb{G}}(T)$,

$$\mu_v.\phi_\ell(a) = \mu_v.(\Phi_{\tilde{\ell}}(\tilde{a})) = \sum_{\tilde{v} \in p^{-1}(v) \cap T_{\tilde{\ell}}} (\mu_{\tilde{v}}.\Phi_{\tilde{\ell}})(\tilde{a}),$$

where $T_{\tilde{\ell}}$ is the set of vertices of $\tilde{\mathbb{G}}$ visited by $\tilde{\ell}$. Since ℓ is regular, either $\# p^{-1}(v) \cap T_{\tilde{\ell}} = 2$ and $\#(V_{\tilde{\ell}} \cap p^{-1}(v)) = 0$, or $\#(p^{-1}(v) \cap T_{\tilde{\ell}}) = \#(p^{-1}(v) \cap V_{\tilde{\ell}}) = 1$.

In the first case, $\ell_1 \nsim_h c_v$ and $\ell_2 \nsim_h c_v$, so that $\phi_{\ell_1} = \phi_{\ell_2} = 0$. Moreover, for any $\tilde{v} \in p^{-1}(v) \cap T_{\tilde{\ell}}$ and $e_1, \ldots, e_4 \in \tilde{E}^o$ four cyclically ordered, outgoing edges at $\tilde{v}$, we may assume that $\tilde{\ell}$ uses e_1^{-1} and e_3, while $e_2, e_4 \notin E_{\tilde{\ell}}$. Therefore, $d\omega_{e_2}.\Phi_{\tilde{\ell}} = d\omega_{e_4}.\Phi_{\tilde{\ell}} = 0$, and as $\mu_{\tilde{v}} = \pm(d\omega_{e_2} + d\omega_{e_4})$, $(\mu_{\tilde{v}}.\Phi_{\tilde{\ell}})(\tilde{a}) = 0 = \phi_{\ell_1}(a)\phi_{\ell_2}(a)$.

In the second case, for $\tilde{v} \in V_{\tilde{\ell}} \cap p^{-1}(v) = T_{\tilde{\ell}} \cap p^{-1}(v)$, by definition of the universal cover, $\ell_1 \sim_h c_v \sim_h \ell_2$. Then $\delta_{\tilde{v}}\tilde{\ell}' = \tilde{\ell}_1 \otimes \tilde{\ell}_2$, where $\tilde{\ell}_1, \tilde{\ell}_2$ are lift with initial condition $\tilde{v}$, so that using 3a) of Theorem 3.9, we get

$$(\mu_{\tilde{v}}.\Phi_{\tilde{\ell}})(\tilde{a}) = \Phi_{\tilde{\ell}_1}(\tilde{a})\Phi_{\tilde{\ell}_2}(\tilde{a}) = \phi_{\ell_1}(a)\phi_{\ell_2}(a).$$

□

The main technical result of this article is the proof of the following uniqueness statements. Denote by $\mathfrak{L}_g$ the space of regular loops of regular maps of genus $g \geq 1$. Let us say that a subset $\mathcal{F}$ of $\mathfrak{L}_g$ is a *good boundary condition of the Makeenko–Migdal equations* if for any pair ϕ^∞ and $(\phi^N)_{N\geq 1}$ made of an exact and an approximate solutions of Makeenko–Migdal equations,

$$\lim_{N\to\infty} \|\phi_\ell^N - \phi_\ell^\infty\|_\infty + \|\mathscr{V}_{\phi^N,\ell}\|_\infty = 0, \ \forall \ell \in \mathcal{F} \tag{45}$$

implies

$$\lim_{N\to\infty} \|\phi_\ell^N - \phi_\ell^\infty\|_\infty + \|\mathscr{V}_{\phi^N,\ell}\|_\infty = 0, \ \forall \ell \in \mathfrak{L}_g. \tag{46}$$

Setting

$$\Psi_\ell^N = \phi^N_{(\ell - \phi_\ell^\infty c)\otimes(\ell-\phi_\ell^\infty c)^*} = \mathscr{V}_{\phi^N,\ell} + |\phi_\ell^N - \phi_\ell^\infty|^2, \tag{47}$$

where c is the constant loop at $\underline{\ell}$, this is equivalent to

$$\lim_{N\to\infty} \|\Psi_\ell^N\|_\infty = 0, \ \forall \ell \in \mathcal{F} \Rightarrow \lim_{N\to\infty} \|\Psi_\ell^N\|_\infty = 0, \ \forall \ell \in \mathfrak{L}_g.$$

Proposition 3.21. *For any genus $g \geq 1$ and total volume $T > 0$, the family of loops $\ell \in \mathfrak{L}_g$ with a subpath γ, such that (ℓ, γ) is a marked loop and $(\ell, \gamma)^\wedge$ is geodesic, is a good boundary condition.*

Denote by $\mathfrak{L}_g^*$ the subset of $\mathfrak{L}_g$ of loops ℓ with $[\ell]_{\mathbb{Z}} \neq 0$. Let us say that a subset $\mathcal{F}^*$ of $\mathfrak{L}_g^*$ is a *good boundary condition in homology* if for any pair ϕ^∞ and $(\phi^N)_{N\geq 1}$ made of an exact and an approximate solution of Makeenko–Migdal equations, using the same notation as in (47),

$$\lim_{N\to\infty} \|\Psi_\ell^N\|_\infty = 0, \ \forall \ell \in \mathcal{F}^* \Rightarrow \lim_{N\to\infty} \|\Psi_\ell^N\|_\infty = 0, \ \forall \ell \in \mathfrak{L}_g^*.$$

The following can be proven independently from Proposition 3.21.

Proposition 3.22. *For any genus $g \geq 1$ and total volume $T > 0$, the family of geodesic loops in $\mathfrak{L}_g^*$ is a good boundary condition in homology.*

When $g = 1$, for any loop $\ell \in \mathfrak{L}_g$, $\ell \sim_h c_{\underline{\ell}}$ if and only if $[\ell]_{\mathbb{Z}} = 0$ and any geodesic loop is of the form s^d, where s is a simple loop and $d \geq 1$. Therefore, Proposition 3.22 and 3.21 have the following consequence.

Corollary 3.23. *Consider $g = 1$, $T > 0$, the set of regular loops $\ell \in \mathfrak{L}_g$ such that $|\ell|_D = 0$ or $\ell = s^d$ for some simple loop s and some integer $d \geq 1$ is a good boundary condition.*

Proof of Theorem 3.14 and Proposition 3.15. Since L^2 convergence implies convergence in probability, both statements follow from Lemma 3.20 and of, respectively, proposition 3.21 and 3.22. □

Proof of Theorem 3.13. Using the solutions given by 1 and 2 of Lemma 3.20, Theorem 3.12 implies that the boundary conditions of Corollary 3.23 are satisfied. Therefore, the convergence in probability holds true for any regular loops. Using Lemma 3.8, it follows that the convergence holds for all $\gamma \in A(\Sigma) \cap \mathrm{L}(\Sigma)$. When $\gamma \in A(\Sigma) \setminus \mathrm{L}(\Sigma)$, under YM_Σ, W_γ is Haar distributed, so that $\mathbb{E}_{\mathrm{YM}_\Sigma}[|W_\gamma|^2] \to 0$ as $N \to \infty$ by [22]. To prove the convergence in probability for any path of finite length, it is now enough to combine the area bound (32) with Proposition 3.7. The uniqueness claim is proved identically considering in place of the above approximate solution, a constant sequence given by an exact solution. □

4. Proof of the main result, stability of convergence under deformation

In this section, we give a proof first of Proposition 3.22, then of Proposition 3.21. We consider exact and approximate solutions ϕ^∞ and $(\phi^N)_{N\geq 1}$ of Makeenko–Migdal equations in genus $g \geq 1$ and volume $T > 0$, define Ψ^N as in (47) and consider the subset $\mathfrak{B}_g \subset \mathfrak{L}_g$ of loops ℓ with map $\mathbb{G}$, satisfying

$$\Psi^N_\ell \underset{N\to\infty}{\longrightarrow} 0 \text{ uniformly on } \Delta_{\mathbb{G}}(T). \tag{48}$$

Our aim is to find a small subset $\mathfrak{C}_g$ of loops in $\mathfrak{L}_g$, such that $\mathfrak{C}_g \subset \mathfrak{B}_g$ implies $\mathfrak{B}_g = \mathfrak{L}_g$. In the first and second second sections, we shall use, respectively, the following bounds. Thanks to (42), (43) and (44), using the same notation, for any $\ell \in \mathfrak{L}_g$ and $X \in \mathfrak{m}_\ell$,

$$|X.\Psi^N_\ell| \leq \|X\| C'_\ell \left(\sum_{v\in V_\ell} \left(\sqrt{\Psi^N_{\ell_{v,1}}} + |\phi^\infty_{\ell_{v,1}}| \right) \left(\sqrt{\Psi^N_{\ell_{v,2}}} + |\phi^\infty_{\ell_{v,2}}| \right) + \frac{1}{N} \right) \tag{49}$$

and

$$|X.\Psi^N_\ell| \leq \|X\| C'_\ell \left(\Psi^N_\ell + \sum_{v\in V_\ell} \left(\Psi^N_{\ell_{v,1}} + \Psi^N_{\ell_{v,2}} \right) + \frac{1}{N} \right), \tag{50}$$

where $C'_\ell > 0$ is a constant independent of $N \geq 1$.

4.1. Non-null homology loops

Let us denote by $\mathfrak{B}^*_g$ the subset $\mathfrak{B}_g \cap \mathfrak{L}^*_g$. The purpose of this section is to prove Proposition 3.22. It is equivalent to the following statement.

Theorem 4.1. *Denote by $\mathfrak{C}^*_g$ the subset of $\mathfrak{L}^*_g$ of regular loops with nonzero homology which are geodesic. If $\mathfrak{C}^*_g \subset \mathfrak{B}^*_g$, then $\mathfrak{B}^*_g = \mathfrak{L}^*_g$.*

The proof of this Theorem hinges on the following application of Makeenko–Migdal equations, similarly to the argument of [21, 33].

https://doi.org/10.1017/fms.2024.152 Published online by Cambridge University Press

Lemma 4.2. *Let $\ell, \ell' \in \mathfrak{L}_g^*$ be two loops of a regular map $\mathbb{G}$ with faces set F, such that there is $K \subset F$ with $K \neq F$ and $\ell \sim_K \eta\ell'\eta^{-1}$, where η is a path with $\underline{\eta} = \underline{\ell'}$ and $\overline{\eta} = \underline{\ell}$. Assume that $\ell' \in \mathfrak{B}_g^*$ and that for any $v \in V_\ell$,*

$$(\delta_v(\ell) = \ell_1 \otimes \ell_2) \Rightarrow (\ell_1 \text{ or } \ell_2 \text{ belongs to } \mathfrak{B}_g^*). \tag{51}$$

Then $\ell \in \mathfrak{B}_g^*$.

Proof. Setting

$$a'(f) = \begin{cases} \frac{T}{\#F - \#K} & \text{if } f \notin K, \\ 0 & \text{if } f \in K \end{cases} \tag{52}$$

defines an element of $\Delta_K(T)$. According to the compatibility condition 3 of Lemma 3.4 and using that $\ell' \in \mathfrak{B}_g^*$,

$$\Psi_\ell^N(a') = \Psi_{\eta\ell'\eta^{-1}}^N(a') = \Psi_{\ell'}^N(a') \longrightarrow 0 \text{ as } N \to +\infty. \tag{*}$$

Now since $[\ell] \neq 0$ and $a, a' \in \Delta_{\mathbb{G}}(T)$, according to Lemma 2.14, $X = a - a' \in \mathfrak{m}_\ell$.

Using the assumption (51) and the inequality (49), each term of the summand vanishes uniformly on $\Delta_{\mathbb{G}}(T)$ as $N \longrightarrow \infty$, and for any $t \in (0, 1)$,

$$|\partial_t \Psi_\ell^N(a + tX)| = |X.\Psi_\ell^N(a + tX)| \leq C_\ell \|X\| \varepsilon_N \leq C_\ell(\|a\| + \|a'\|)\varepsilon_N, \tag{**}$$

where $\varepsilon_N \longrightarrow 0$. Thanks to the boundary condition (*), we conclude that

$$\Psi_\ell^N(a) = \Psi_{\ell'}^N(a') + \int_0^1 \partial_t \Psi_\ell(a' + tX)dt$$

converges to 0 uniformly in $a \in \Delta_{\mathbb{G}}(T)$, as $N \longrightarrow \infty$; that is, $\ell \in \mathfrak{B}_g^*$. □

We split the proof Theorem 4.1 into two steps. The first one allows to contract inner loops; the second allows to follow a shortening sequence from proper loops to loops conjugated to a geodesic. Denote by $\mathfrak{P}_g^*$ the subset of $\mathfrak{L}_g^*$ of loops which are proper or included in a fundamental domain. Theorem 4.1 is a direct consequence of the following.

Proposition 4.3.

a) *If $\mathfrak{P}_g^* \subset \mathfrak{B}_g^*$, then $\mathfrak{B}_g^* = \mathfrak{L}_g^*$.*
b) *If $\mathfrak{C}_g^* \subset \mathfrak{B}_g^*$, then $\mathfrak{P}_g^* \subset \mathfrak{B}_g^*$.*

Proof. Let us recall the definition of $\mathcal{C}$ above Lemma 2.24. Let us prove first point a). Assume $\mathfrak{P}_g^* \subset \mathfrak{B}_g^*$ and introduce for any $n \geq 0$ the subset $\ell_{n,g}^*$ of loops $\ell \in \mathfrak{L}_g^*$ with $\mathcal{C}(\ell) \leq n$. By assumption, $\ell_{0,g}^* \subset \mathfrak{P}_g^* \subset \mathfrak{B}_g^*$.

Consider $n > 0$ and assume $\ell_{n-1,g}^* \subset \mathfrak{B}_g^*$. Consider $\ell \in \ell_{n,g}$ with $\#V_{c,\ell} > 0$. According to Lemma 2.24, for all $v \in V_\ell$ with $\delta_v\ell = \ell_1 \otimes \ell_2$, $\mathcal{C}(\ell_1), \mathcal{C}(\ell_2) < n$ and $[\ell_1]$ or $[\ell_2] \neq 0$. Hence, ℓ_1 or ℓ_2 belongs to $\ell_{n-1,g}^*$. Choosing K as the bulk of an inner loop α of ℓ, and ℓ' the loop obtained from ℓ by erasing the edges of α, $\ell' \sim_K \ell$, $\ell' \in \ell_{n-1,g}^*$ and Lemma 4.2 implies $\ell \in \mathfrak{B}_g^*$. Point a) follows by induction.

Let us now prove b) Assume that $\mathfrak{C}_g^* \subset \mathfrak{B}_g^*$ and introduce for any $n \geq 0$ the subset $\mathfrak{P}_{n,g}^*$ of proper loops $\ell \in \mathfrak{P}_g^*$ with $|\ell|_D \leq n$.

By assumption, $\mathfrak{P}_{0,g}^* \subset \mathfrak{C}_g^* \subset \mathfrak{B}_g^*$. Assume that $n > 0$ and $\mathfrak{P}_{n-1,g}^* \subset \mathfrak{B}_g^*$, and consider $\ell \in \mathfrak{P}_{n,g}^*$. According to Proposition 2.17, there is a geodesic loop $\ell' \in \mathfrak{C}_g^*$ and shortening homotopy sequence

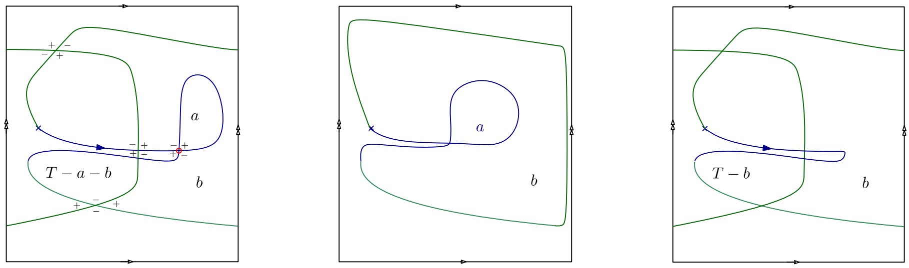

Figure 18. *Faces are labelled by their area. Faces without label have area* 0*. In the left figure,* ± *symbols stand for the area change involved in the decomposition of* $\delta_a F$ *as a signed sum of Makeenko–Migdal vectors at four vertices acting on* ψ*. Here, only the vertex highlighted with a red circle yields a desingularisation with only null-homology loops.*

$\ell_1, \ldots, \ell_m$ of proper loops with $\ell_1 = \ell$ and $\ell_m \sim_K \eta\ell'\eta^{-1}$ for some path η and proper subset of faces K. By assumption, $\ell' \in \mathfrak{B}_g^*$. Using Lemma 2.24 and Lemma 4.2, by induction on m, $\ell \in \mathfrak{B}_g^*$.

This concludes the proof of b) by induction on n. □

Remark 4.4. In the above proof, if we furthermore assume simple loops with nonvanishing homology to be included in $\mathfrak{B}_g^*$, it is also possible to argue by induction on the number of vertices.

4.2. Null homology loops

The purpose of this subsection is to prove Proposition 3.21. It is equivalent to the following statement.

Theorem 4.5. *Denote by* $\mathfrak{C}_g^\vee$ *the subset of* $\mathfrak{L}_g$ *of regular loops* ℓ*, such that there is a nested subpath* γ_{nest} *of* ℓ *making* (ℓ, γ_{nest}) *a marked loop on a map of genus g and with* $\ell^\wedge$ *geodesic. If* $\mathfrak{C}_g^\vee \subset \mathfrak{B}_g$*, then* $\mathfrak{B}_g = \mathfrak{L}_g$.

To prove this theorem, we shall use the following lemma, formally analog to Lemma 4.2. Though, unlike Lemma 4.2, due to the new constraint on the Makeenko–Migdal vectors, we work here with marked loops and change the nested part in order to keep the constraint satisfied while performing the required homotopy. This will break the induction on the number of intersection points or the complexity $\mathcal{C}$ on regular loops.

The main idea to address this problem hinges on the observation, applied in Step 3 of the proof below, that loops obtained by desingularisation at the intersection points of the nested part of a marked loops yield either inner loops or a contraction of faces bounded by inner loops of the nested part. The Makeenko-Migdal equation leads then to a Grönwall inequality that allows to use an induction on the complexity $\mathcal{C}^{\mathfrak{m}}$ of marked loops.

Uniqueness of Makeenko–Migdal equations; example of Figure 4: Let us illustrate the main idea used in the lemma by a simple example related to the deformation considered in Figure 4. Consider $\Delta = \{(a, b) \in \mathbb{R}_+^2 : a + b \le T\}$ and a function $F \in C^1(\Delta)$ associated to a solution ψ of the Makeenko–Migdal equations for the loop illustrated on the left of Figure 18. Assume that ψ vanishes on loops of non-null homology and matches with the planar master field for loops included in a fundamental domain. Then the restrictions $F_{|a+b=T}$ and $F_{|a=0}$ are, respectively, associated to the loops on the middle and on the right of Figure 18. Since the first loop is included in a fundamental domain, using the value of the planar master field (see, for instance, Table 1 of [43, Appendix]),

$$F(a, T-a) = (1-a)e^{-\frac{a+T}{2}}, \ \forall a \in [0, T]. \qquad (\clubsuit)$$

https://doi.org/10.1017/fms.2024.152 Published online by Cambridge University Press

Writing the vector field ∂_a as the signed sum of Makeenko–Migdal vectors at four vertices displayed in the left-hand side of Figure 18, as ψ vanishes on loops with non-null homology, only the term associated to the circled vertex contributes to the right-hand side of Makeenko-Migdal equations with a negative sign. Moreover,[45] the desingularisation at this vertex yields a simple contractible loop bounding an area a and a loop obtained by contracting the face with the area parameter a to a point, as displayed on the right of Figure 18 but with area parameter $a+b$ in place of b. Therefore, using, respectively, Table 1 of [43, Appendix] and the restriction property, ψ maps these two loops to $e^{-\frac{a}{2}}$ and $F(0,a+b)$. All in all,

$$\partial_a F(a,b) = -e^{-\frac{a}{2}} F(0,a+b),\ \forall (a,b) \in \Delta. \qquad (\spadesuit)$$

Now the equation ($\spadesuit$) with boundary condition ($\clubsuit$) has a unique solution. Indeed, denoting by G the difference of two solutions, and setting $H(t) = \sup_{a\in[0,t]} |G(a,t-a)|$,

$$H(t) \le \int_t^T H(s)ds,\ \forall t \in [0,T].$$

By Grönwall's inequality, $H(t) = 0$ for all $t \in [0,T]$. Since the following right-hand side satisfies ($\spadesuit$) and ($\clubsuit$), we conclude that[46]

$$F(a,b) = (1-a)e^{-\frac{2T-b}{2}},\ \forall (a,b) \in \Delta.$$

Let us return to the proof of Theorem 4.5. Denote by $\mathfrak{L}_g^{\mathrm{m}}$ the set of marked loops on a regular map of genus g and by $\mathfrak{B}_g^{\mathrm{m}}$ the set of $(\ell,\gamma_{nest}) \in \mathfrak{L}_g^{\mathrm{m}}$ such that $\ell' \in \mathfrak{B}_g$, whenever $(\ell',\gamma'_{nest}) \in \mathfrak{L}_g^{\mathrm{m}}$ with $\ell'^{\wedge_*} = \ell^{\wedge_*}$. Recall the notation (25) for the desingularisation of a marked loop.

Lemma 4.6. *Assume that for any regular loop ℓ with $|\ell|_D = 0$, $\ell \in \mathfrak{B}_g$. Let $x = (\alpha,\alpha_{nest}), y = (\beta,\beta_{nest}) \in \mathfrak{L}_g^m$ be two marked loops on a same regular map $\mathbb{G}$ and K a proper subset of faces of $\mathbb{G}$, such that $\alpha_{nest} = \beta_{nest}$ with a moving edge that is not adjacent to any face of K, $\alpha^{\wedge_*} \sim_K \beta^{\wedge_*}$ and $y \in \mathfrak{B}_g^{\mathrm{m}}$, while*

$$\forall v \in V_{\alpha^\wedge}, \delta_v(x) = x_1 \otimes x_2 \text{ with } x_1, x_2 \in \mathfrak{B}_g^{\mathrm{m}}. \qquad (53)$$

Then $\alpha \in \mathfrak{B}_g$.

Proof of Lemma 4.6. Since $\alpha \sim_K \beta$ and $\beta \in \mathfrak{B}_g$,

$$\Psi_\alpha^N = \Psi_\beta^N \longrightarrow 0 \text{ uniformly on } \Delta_{\mathbb{G},K}(T). \qquad (54)$$

Let us prove that Ψ_α^N converges uniformly to zero on $\Delta_{\mathbb{G},K}(T)$.

Step 1: Let us first show that it is enough to show the latter convergence on one half of the simplex $\Delta_{\mathbb{G}}(T)$ depending on the winding[47] of α. Thanks to Theorem 4.1, we can assume that $[\alpha] = 0$. Recall from Lemma 2.12 that since $[\alpha] = [\alpha^\wedge] = 0$, we can define winding number functions $n_\alpha, n_{\alpha^\wedge} \in \Omega^2(\mathbb{G})$ for α and $\alpha^\wedge$, and that they are unique up to the choice of an additive constant. Let F_{nest} be the bulk of the nested part of (α,α_{nest}) and let f_o be its outer face. We fix $n_\alpha, n_{\alpha^\wedge}$ setting $n_\alpha(f_o) = n_{\alpha^\wedge}(f_o) = 0$. Now define

$$\Delta_\pm(T) = \{a \in \Delta_{\mathbb{G}}(T) : \pm\langle n_{\alpha^\wedge}, a\rangle \ge 0\}$$

[45] We use here the property at works in the proof below.
[46] Besides, using again Table 1 of [43, Appendix], the reader can also check that the right-hand side is indeed the value of the master field at the lift to $\mathbb{R}^2$ of the loop on the left-hand side of Figure 18.
[47] This will inform in which direction to twist α in the next step.

https://doi.org/10.1017/fms.2024.152 Published online by Cambridge University Press

and choose two faces $f_-, f_+ \in F_{nest} \cup \{f_o\}$ such that

$$n_\alpha(f_-) = \min_{f \in F_{nest} \cup \{f_o\}} n_\alpha(f_-) = n_- \text{ and } n_\alpha(f_+) = \max_{f \in F_{nest} \cup \{f_o\}} n_\alpha(f_+) = n_+.$$

Since $\Delta_+(T) \cup \Delta_-(T) = \Delta_{\mathbb{G}}(T)$ and $\Psi^N_{\alpha^{-1}} = \Psi^N_\alpha$, it is enough to show that as $N \to \infty$, $\Psi^N_\alpha \to 0$ uniformly on $\Delta_+(T)$.

Step 2: Let us now modify α to have a face disjoint from K and with high enough winding number, so that changing its area allows to put all areas of faces of K to zero without changing the algebraic area. The new area vector will be denoted by a' below; it will vanish on a set K_* of faces that covers K, while satisfying the algebraic constraint (56).

Consider $\lambda = 2 \max_{f \in F} |n_\alpha(f)|$ and define (ℓ, γ_{nest}) as the λ-twist of (α, α_{nest}). Denote by $\mathbb{G}' = (V', E', F')$ the associated map finer than $\mathbb{G}$ and by F_{tw} the subset of λ faces of F' associated to the twist move such that $\ell \sim_{F_{tw}} \alpha$. Denote by f_l the face of $\mathbb{G}$ left of the moving edge and, respectively, by f_l' and f_c' the unique face of $\mathbb{G}'$ adjacent to F_{tw} and the central face of (ℓ, γ_{nest}). Faces of $F \setminus \{f_l\}$ are not changed by the twist and can be identified with $F' \setminus \left(F_{tw} \cup \{f_l'\}\right)$. In particular, faces of K can and will be identified with faces of $\mathbb{G}'$. We shall write $f_-' = f_l'$ when $f_- = f_l$, and $f_-' = f_-$ otherwise.

Recall that $[\ell] = [\alpha] = 0$ and denote by n_ℓ the winding number function of ℓ with $n_\ell(f_o') = 0$. It satisfies

$$n_\ell(f_c') = \lambda + n_\alpha(f_l), 1 \leq n_\ell(f) - n_\alpha(f_l) \ \leq \lambda - 1, \ \forall f \in F_{tw} \setminus \{f_c\}$$

while

$$n_\ell(f) = n_\alpha(f), \ \forall f \in F' \setminus \left(F_{tw} \cup \{f_o'\}\right) \text{ and } n_\ell(f_l') = n_\alpha(f_l).$$

It follows that

$$n_\ell(f_c') = \max_{f \in F'} n_\ell(f). \tag{55}$$

Recall that $\alpha^\wedge = \ell^\wedge$ viewed as loops in $\mathbb{G}'$ and denote

$$\Delta_+'(T) = \{a \in \Delta_{\mathbb{G}'}(T) : \langle n_{\ell^\wedge}, a \rangle \geq 0\}.$$

Since the restriction map from $\Delta_+'(T)$ to $\Delta_+(T)$ is surjective, it is enough to show that $\Psi^N_\ell \to 0$ uniformly on $\Delta_+'(T)$.

For any $a \in \Delta_+'(T)$, thanks to (55) and since $n_\ell(f) \geq n_-$ for all $f \in F_{nest}$,

$$n_\ell(f_-')T \leq \langle n_{\ell^\wedge}, a \rangle + a(F_{nest}) n_- \leq \langle n_\ell, a \rangle \leq n_\ell(f_c')T.$$

Hence, setting $K_* = F' \setminus \{f_c', f_-'\}$, there is a vector $a' \in \Delta_{K^*}(T)$ with

$$\langle n_\ell, a' \rangle = \langle n_\ell, a \rangle, \tag{56}$$

and thanks to Lemma 2.14, $X = a' - a \in \mathfrak{m}_\ell$. Moreover, since $n_{\ell^\wedge}$ vanishes on $\{f_l', f_-'\}$, $\langle n_{\ell^\wedge}, a' \rangle = 0$ and $a' \in \Delta_+'(T)$.

Step 3: Let us show that Makeenko–Migdal equations applied along the vector X allow to apply a Grönwall inequality to prove the uniform convergence of Ψ_ℓ on $\Delta_+'(T)$; according to Step 2, this will then conclude the proof. Therefore, we shall need to bound $\delta_v \Psi_\ell$ for all $v \in V_\ell$. For vertices not belonging to the nested part of ℓ, we shall use the assumption (53). For the remaining ones, it will be achieved thanks to the restriction identity (58) below, which will be achieved similarly to (♠) in the simple example introducing the proof. Similarly to the uniqueness argument in that example, we then deduce a Grönwall type inequality in (59) from which the claim follows.

https://doi.org/10.1017/fms.2024.152 Published online by Cambridge University Press

i) Consider first vertices outside of the nested part of ℓ. Note that $V_{\ell^\wedge} = V_{\alpha^\wedge}$. Writing $z = (\ell, \gamma_{nest})$, for all $v \in V_{\ell^\wedge} = V_{\alpha^\wedge}$, if $\delta_v(x) = x_1 \otimes x_2$, then $\delta_v(z) = z_1 \otimes z_2$, where one marked loop, say z_1, is identical to or obtained from x_i by λ- twist at the moving edge e, whereas the other satisfies $z_2 = x_2$. In particular, $z_i^\wedge = x_i^\wedge$, and using the assumption (53), $z_i \in \mathfrak{B}_g^{\mathfrak{m}}$ and $\ell_{v,i} \in \mathfrak{B}_g$ for $i \in \{1, 2\}$.

ii) Consider now $V_{\ell_{nest}}$. We denote by $v_1, \ldots, v_n$ the intersection points of the nested part of ℓ, ordering them so that $\ell_{nest} = (v_1 \ldots v_n v_n \ldots v_1)$ and by F'_{nest} the bulk of ℓ_{nest}. For all $1 \leq k \leq n$, w.l.o.g., $\delta_{v_k}(\ell) = \alpha_k \otimes \ell_k$, where α_k is a nested loop with $|\alpha_k|_D = 0$; hence, $\alpha_k \in \mathfrak{B}_g$, and ℓ_k is a sub-loop of ℓ, with $\ell_1 = \alpha$ and $\ell_k \sim_{F_{nest}} \ell$ for all $1 \leq k \leq n$. Denote by F_k the minimal subset of F_{nest} with $\ell_k \sim_{F_k} \ell$. Since $X \in \mathfrak{m}_\ell$, using inequality (50), we find

$$|X.\Psi_\ell^N| \leq C\left(\varepsilon_N + \Psi_\ell^N + \sum_{k=1}^{n} \Psi_{\ell_k}^N\right), \tag{57}$$

where $C > 0$ is a constant independent of N and

$$\varepsilon_N = \frac{1}{N} + \sup_{1 \leq k \leq n_+} \|\Psi_{\alpha_k}^N\|_\infty + \sup_{v \in V_{\ell^\wedge}} \left(\|\Psi_{\ell_{v,1}}^N\|_\infty + \|\Psi_{\ell_{v,2}}\|_\infty\right).$$

Thanks to i), and as $|\alpha_k|_D = 0$, for all k, $\ell_{v,1}, \ell_{v,2}, \alpha_k \in \mathfrak{B}_g$ for all $v \in V_{\ell^\wedge}$ and $1 \leq k \leq n$, so that $\lim_{N \to \infty} \varepsilon_N = 0$. Let us now bound each term of the sum in the right-hand side of (57) in terms of Ψ_ℓ. Consider for all $t \in [0, 1]$,

$$\Delta_{in}(t) = \{a \in \Delta_{\mathbb{G}'}(tT) : a(f) = 0, \forall f \notin F_{nest} \cup \{f'_o\}\},$$

and for all $a \in \Delta_{F_{tw},+}(T)$ fixed, set

$$H_a^N(t) = \sup_{b \in \Delta_{in}(1-t)} \Psi_\ell(ta + b), \ \forall 0 \leq t \leq 1.$$

On the one hand, for any $t \in (0, 1)$ and $b \in \Delta_{in}(1 - t)$,

$$\partial_s \Psi_\ell^N(sa + (t - s)a' + b) = X.\Psi_\ell^N(sa + (t - s)a' + b), \ \forall s \in (0, t).$$

On the other hand, for all $s \in (0, t)$ and any k, as $a(F_k) = 0$ and $\ell_k \sim_{F_k} \ell$, there are $b_1, \ldots, b_n \in \Delta_{in}(1 - s) \cap \Delta_{F_k}((1 - s)T)$ (see Figure 19) such that

$$\Psi_{\ell_k}(sa + (t - s)a' + b) = \Psi_\ell(sa + b_k), \ \forall 1 \leq k \leq n. \tag{58}$$

Combining the last two equalities with the bound (57), we find

$$H_a^N(t) \leq H_a^N(0) + \varepsilon_N C + (n + 1)C \int_0^t H_a^N(s)ds, \ \forall t \in [0, 1], a \in \Delta'_+(T).$$

By Grönwall's inequality,

$$H_a^N(t) \leq (H_a^N(0) + \varepsilon_N C) \exp((n + 1)Ct), \ \forall t \in [0, 1]. \tag{59}$$

Since $\Delta_{in}(1) \subset \Delta_K(T)$, by (54),

$$\sup_{a \in \Delta'_+(T)} H_a^N(0) \leq \sup_{x \in \Delta_{\mathbb{G},K}(T)} \Psi_\ell(x)$$

https://doi.org/10.1017/fms.2024.152 Published online by Cambridge University Press

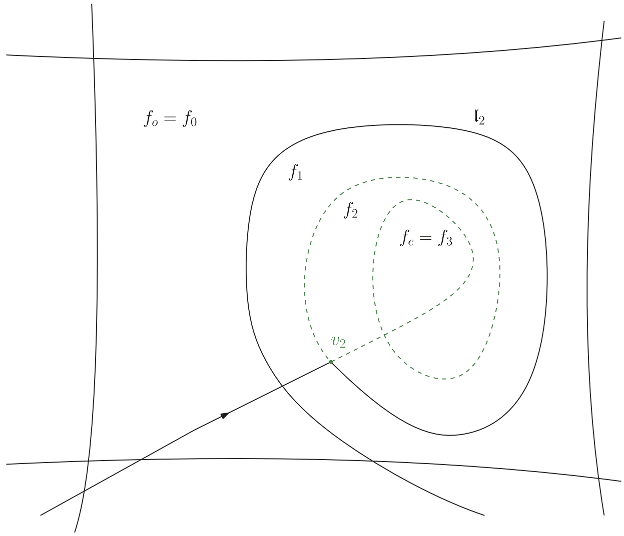

Figure 19. *Example of a n-left twist with* $n = 3$*. We consider here* $k = 2$*; the area of* F_2 *needs to be 'moved' into* f_1*. We have* $a(f_1) = a(f_2) = a(f_3) = 0 = a'(f_1) = a'(f_2)$*. For all* $0 < s < t < 1$*, define* b_2 *setting* $b_2(f_1) = b(F_1) + (t - s)a'(F_1)$ *and* 0 *for other faces. Denote* $a_{s,t} = sa + (t - s)a' + b$ *and* $\tilde{a}_{s,t} = sa + b_2$*. On the one hand, for any face* $f \notin F_1$*,* $a_{s,t}(f) = a'_{s,t}(f)$ *while* $a_{s,t}(F_1) = a'_{s,t}(F_1)$*; therefore,* $\Psi^N_{\ell_2}(a_{s,t}) = \Psi^N_{\ell_2}(\tilde{a}_{s,t})$*. On the other hand,* $\tilde{a}_{s,t}(F_2) = 0$ *so that* $\Psi^N_{\ell_2}(\tilde{a}_{s,t}) = \Psi^N_{\ell}(\tilde{a}_{s,t})$*.*

vanishes as $N \to \infty$. Since $\varepsilon_N \to 0$ as $N \to \infty$, from (59),

$$\Psi^N_\ell(a) = H^N_a(1) \to 0$$

uniformly in $a \in \Delta'_+(T)$.

□

Using this lemma, the rest of the proof is a refinement of the null-homology case. Denote by $\mathfrak{C}^{\mathfrak{m}}_g, \mathfrak{P}^{\mathfrak{m}}_g$ the set of marked loops $(\ell, \ell_{nest}) \in \mathfrak{L}^{\mathfrak{m}}_g$ with $|\ell|_D = 0$, or, respectively, $\ell^\wedge \in \mathfrak{C}_g$ and $\ell^\wedge$ proper. Theorem 4.5 is then a direct consequence of the following Proposition.

Proposition 4.7.

a) *If* $\mathfrak{P}^{\mathfrak{m}}_g \subset \mathfrak{B}^{\mathfrak{m}}_g$*, then* $\mathfrak{B}^{\mathfrak{m}}_g = \mathfrak{L}^{\mathfrak{m}}_g$.
b) *If* $\mathfrak{C}^{\mathfrak{m}}_g \subset \mathfrak{B}^{\mathfrak{m}}_g$*, then* $\mathfrak{P}^{\mathfrak{m}}_g \subset \mathfrak{B}^{\mathfrak{m}}_g$.

Proof. Let us recall the definition of $\mathcal{C}^{\mathfrak{m}}$ above Lemma 2.24. Let us prove first point a). Assume $\mathfrak{P}^{\mathfrak{m}}_g \subset \mathfrak{B}^{\mathfrak{m}}_g$ and introduce for any $n \geq 0$ the subset $\ell^{\mathfrak{m}}_{n,g}$ of marked loops $x \in \mathfrak{L}^{\mathfrak{m}}_g$ with $\mathcal{C}^{\mathfrak{m}}(x) \leq n$. By assumption, $\ell^{\mathfrak{m}}_{0,g} \subset \mathfrak{P}^{\mathfrak{m}}_g \subset \mathfrak{B}^{\mathfrak{m}}_g$.

Consider $n > 0$ and assume $\ell^{\mathfrak{m}}_{n-1,g} \subset \mathfrak{B}^{\mathfrak{m}}_g$. Consider $x = (\alpha, \alpha_{nest}) \in \ell_{n,g}$ with $\#V_{c,x^\wedge} > 0$. According to Lemma 2.24, for all $v \in V_\ell$, $\delta_v x = x_1 \otimes x_2$, with $\mathcal{C}^{\mathfrak{m}}(\ell_1), \mathcal{C}^{\mathfrak{m}}(\ell_2) < n$. Hence, $x_1, x_2 \in \ell^{\mathfrak{m}}_{n-1,g}$. Thanks to Proposition 4.3, we can assume $[\alpha] = 0$. Choosing K as the bulk of an inner loop ℓ of $x^\wedge$, and $y = (\beta, \beta_{nest})$ the marked loop obtained from x by erasing the edges of ℓ, $\alpha \sim_K \beta$, $y \in \ell^{\mathfrak{m}}_{n-1,g}$. Since α_{nest} do not intersect inner loops of $\alpha^\wedge$, the moving edge of x is not adjacent to any face of K. Lemma 4.6 applies and yields $\ell \in \mathfrak{B}^{\mathfrak{m}}_g$. Point a) follows by induction.

Consider now b). Assume that $\mathfrak{C}^{\mathfrak{m}}_g \subset \mathfrak{B}^{\mathfrak{m}}_g$ and introduce for any $n \geq 0$ the subset $\mathfrak{P}^{\mathfrak{m}}_{n,g}$ of marked loops $x \in \mathfrak{P}^{\mathfrak{m}}_g$ with $|x^\wedge|_D \leq n$. By assumption, $\mathfrak{P}^{\mathfrak{m}}_{0,g} \subset \mathfrak{C}^{\mathfrak{m}}_g \subset \mathfrak{B}^{\mathfrak{m}}_g$. Assume that $n > 0$ and $\mathfrak{P}^{\mathfrak{m}}_{n-1,g} \subset \mathfrak{B}^{\mathfrak{m}}_g$, and consider $x = (\alpha, \alpha_{nest}) \in \mathfrak{P}^{\mathfrak{m}}_{n,g}$. According to Proposition 2.17, there is a geodesic loop $\ell' \in \mathfrak{C}^{\mathfrak{m}}_g$ and shortening homotopy sequence $x_1,\dots,x_m$ of marked loops with $x_i^\wedge$ proper, $x_1 = x$ and $x_m = (\ell_m,\gamma_m)$ such that $\ell_m \sim_K \eta\ell'\eta^{-1}$ for some path η and proper subset of faces K. By assumption, $\ell' \in \mathfrak{B}_g^{\mathfrak{m}}$. Consider the first proper set of faces K_1 with $x_1^{\wedge *} \sim_{K_1} x_2^{\wedge *}$. Denote by x', y' the pull of x_1 and x_2 to a face that does not belong to K_1. Lemma 4.6 applies to x', y'. Since $\mathcal{C}^{\mathfrak{m}}(x_i)$ is nonincreasing, we conclude by induction on m that $x \in \mathfrak{B}_g^{\mathfrak{m}}$. This concludes the proof of b) by induction on n. □

5. Proof of convergence after surgery

We give here the main arguments to prove Theorem 3.12.

Proof of Lemma 3.11. Let us start by applying the second part of Lemma 3.3. It follows that, under $\mathrm{YM}_{\mathbb{G},\{f_\infty\},a}$, $(h_{\ell_1},\dots,h_{\ell_r},h_{a_1},\dots,h_{b_g})$ are independent random variables on G_N, such that for all $1 \le i \le g$, h_{a_i}, h_{b_i} are Haar distributed, while for any $1 \le k \le r$, h_{ℓ_k} has the same law as a Brownian motion at time $a(f_k)$. It is now standard (see [43, Section 3]) that as $N \to \infty$, this tuple of matrices is asymptotically freely independent, and its joint noncommutative distribution converges towards τ_v, satisfying the properties (*), 1,2 and 3. □

Let us use the same notation as in Theorem 3.12. In what follows, we will denote by $\mathbb{E}$ (resp. $\mathbb{E}_i$, $\mathbb{E}'_i$) the expectation with respect to $\mathrm{YM}_{\mathbb{G},a}$ (resp. $\mathrm{YM}_{\mathbb{G}_i,a_i}$, $\mathrm{YM}_{\mathbb{G}'_i,a}$). In a previous paper, we proved that the restriction to $\mathbb{G}'_1$ of $\mathrm{YM}_{\mathbb{G},a}$ is absolutely continuous with respect to $\mathrm{YM}_{\mathbb{G}'_1,a}$.

Proposition 5.1 ([20], Corollary 4.3). *Let $\ell \in \mathrm{RL}_v(\mathbb{G}'_1)$. For any $f : G_N \to \mathbb{C}$ bounded, measurable and central,*

$$\mathbb{E}[f(H_\ell)] = \mathbb{E}'_1[f(H_\ell)I(H_{\ell_0}^{-1})], \tag{60}$$

where $I : G_N \to \mathbb{C}$ is a bounded measurable function such that

$$\|I\|_\infty \le \frac{Z_{g_2,a(F_2)}}{Z_{g,T}}.$$

Note that the bound in the previous proposition ensures that I is uniformly bounded because for any considered sequence $(G_N)_N$, the corresponding sequences of partition functions converge towards a nonzero limit.[48]

Proof of Theorem 3.12. Without loss of generality, we can assume that $\mathbb{G}$ is a regular map, with $v = p(\tilde{v})$, where $\tilde{v} \in \tilde{V}_g$. Let ℓ be a loop in $\mathrm{L}_v(\mathbb{G}_1)$. According to Proposition 5.1,

$$\mathbb{E}[W_\ell] = \mathbb{E}'_1[W_\ell I(H_{\ell_0^{-1}})],$$

where I is uniformly bounded in N. From Lemma 3.11, W_ℓ converges in probability towards $\Phi_\ell^{1,g_1}(a_1)$ under $\mathrm{YM}_{\mathbb{G}_1,a}$. Because I is uniformly bounded in N, this convergence holds true as well under $\mathrm{YM}_{\mathbb{G},a}$. It remains to identify $\Phi_\ell^{1,g_1}(a_1)$ with $\Phi_\ell(a)$.

Consider a free basis $\ell_1,\dots,\ell_r,a_1,b_1,\dots,a_g,b_g$ of $\mathrm{RL}_v(\mathbb{G})$ as in Lemma 2.4 and let us identify $\mathrm{RL}_v(\mathbb{G}_1)$ as a subgroup of $\mathrm{RL}_v(\mathbb{G})$. Denote by $\tilde{\tau}$ the linear functional on $(\mathbb{C}[\mathrm{RL}_v(\mathbb{G})],*)$ that satisfies for all $\ell \in \mathrm{RL}_v(\mathbb{G})$,

$$\tilde{\tau}(\ell) = \Phi_\ell(a).$$

It is enough to show that the restriction of $\tilde{\tau}$ to $\mathbb{C}[\mathrm{RL}_v(\mathbb{G}_1)]$ satisfies 1, 2 and 3 of Lemma 3.11.

Point 3 follows from point 1 of Lemma 3.10. Consider point 2. For any $\ell \in S_{top} = \{a_1,b_1,\dots,a_{g_1},b_{g_1}\}$ and $k \in \mathbb{Z}^*$, ℓ^k is not contractible, and therefore, $\tilde{\tau}(\ell^k) = 0$. Let us now prove

[48] Besides, when $g_2 \ge 2$ and $G_N \ne \mathrm{U}(N)$, it remains bounded uniformly in $a \in \Delta_{\mathbb{G}}(T)$, which allows then to drop the condition $a(F_2) > 0$ in Theorem 3.12.

https://doi.org/10.1017/fms.2024.152 Published online by Cambridge University Press

point 1. Note that $\ell_1, \dots, \ell_{r_1}$ have the same joint distribution under τ_v and $\tilde{\tau}$. Hence, thanks to point 2 of Lemma 3.10, $\ell_1, \dots, \ell_{r_1}$ are freely independent under $\tilde{\tau}$.

Since $g_2 \geq 1$, according to Lemma 2.3, identifying $\pi_{1,v}(\mathbb{G})$ with $\pi_{1,v}(\mathbb{G})$, the images of $a_1, b_1, \dots, a_{g_1}, b_{g_1}$ in $\pi_{1,v}(\mathbb{G}) \simeq \Gamma_g$ span a free subgroup $\Gamma^\#$ of Γ_g of rank $2g_1$, isomorphic to the group $\mathrm{RL}_{top,1}$ generated by $a_1, \dots, b_{g_1}$ in $\mathrm{RL}_v(\mathbb{G})$. Therefore, $\tilde{V}_L = \widetilde{\Gamma^\#.V_g}$ is included in a spanning tree $\mathcal{T}$ of $\tilde{V}_g$. Choosing $(\gamma_x)_{x\in\tilde{V}_g}$ as in Lemma 2.16, $(\gamma_x \ell_i \gamma_x^{-1})_{x\in\tilde{V}_g, 1\leq i\leq r}$ is a free basis of lassos of $\mathrm{RL}_{\tilde{v}}(\tilde{\mathbb{G}})$. Therefore, thanks to Lemma 3.10, $(\gamma_x \ell_i \gamma_x^{-1})_{x\in\tilde{V}_L, 1\leq i\leq r_1}$ are freely independent under $\tilde{\tau}$. For any $\gamma \in \mathrm{RL}_{top,1}$, denote by $\mathcal{A}_\gamma$ the subalgebra generated by $\gamma\ell_1\gamma^{-1}, \dots, \gamma\ell_{r_1}\gamma^{-1}$. We infer in particular that the subalgebras $(\mathcal{A}_\gamma)_{\gamma\in\mathrm{RL}_{top,1}}$ are freely independent under $\tilde{\tau}$.

Now since $\Gamma^\#$ is free over the image of S_{top}, for any alternated word w in $a_1, \dots, b_{g_1}$, the image in $\Gamma^\#$ is not trivial and the associated loop $\ell_w \in \mathrm{RL}_{top,1}$ is not contractible; hence, $\tilde{\tau}(w) = 0$. To conclude, it remains to show that the subalgebra $\mathcal{A}_\mathfrak{f}$ and $\mathcal{A}_{top}$ of $\mathbb{C}[\mathrm{RL}_v(\mathbb{G}_1)]$ spanned, respectively, by $S_\mathfrak{f} = \{\ell_1, \dots \ell_{r_1}\}$ and S_{top} are freely independent under $\tilde{\tau}$. Since $\tilde{\tau}$ is tracial and unital, it is enough to show

$$\tilde{\tau}(w_1\alpha_1 w_2 \dots w_n \alpha_n w_{n+1}) = 0$$

whenever $w_1, \dots, w_n \in \mathrm{RL}_{top,1} \setminus \{c_v\}, w_{n+1} \in \mathrm{RL}_{top,1}$ and $\alpha_1, \dots, \alpha_n \in \mathcal{A}_\mathfrak{f}$ with $\tilde{\tau}(\alpha_1) = \dots = \tilde{\tau}(\alpha_n) = 0$. Denote by $G_\mathfrak{f}$ the subgroup of $\mathrm{RL}_v(\mathbb{G}_1)$ generated by $S_\mathfrak{f}$. Since $\mathrm{RL}_{top,v}$ is isomorphic to $\Gamma^\#$, if $w_1 \dots w_{n+1}$ does not reduce to the constant loop, then for any $x_1, \dots, x_n \in G_\mathfrak{f}, w_1 x_1 w_2 \dots w_n x_n w_{n+1} \sim_h w_1 \dots w_{n+1}$ is not contractible, and the claim follows. Otherwise, we have $w_1 \dots w_{n+1} = 1 \in \mathrm{RL}_{top,1}$ and

$$w_1\alpha_1 w_2 \dots w_n \alpha_n w_{n+1} = \gamma_1 \alpha_1 \gamma_1^{-1} \gamma_2 \alpha_1 \gamma_2^{-1} \dots \gamma_n \alpha_1 \gamma_n^{-1}, \tag{61}$$

where $\gamma_i = w_1 \dots w_i$ for all $1 \leq i \leq n$. Now for all $1 \leq i < n$, since $w_{i+1} \neq 1 \in \mathrm{RL}_{top,1}$, $\gamma_i \neq \gamma_{i+1}$ and it follows that (61) is an alternated word in centered elements of $(\mathcal{A}_g)_{g\in\mathrm{RL}_{top,1}}$. Since these subalgebras are free under $\tilde{\tau}$, the claim follows. □

6. Interpolation between regular representations

6.1. State extension and interpolation

In this section, we remark that the maps considered in Conjecture 1.3 have a positivity property and can be seen as states of a noncommutative probability space.

Lemma 6.1. *Consider two groups G, Γ, a surjective morphism $\pi : G \to \Gamma$, and τ a unital state on $(\mathbb{C}[K], 1_G, *)$, where $K = \ker(\pi)$ and 1_G denote the neutral element of G. For any $g \in G$, set*

$$\tilde{\tau}(g) = \begin{cases} \tau(g) & \text{if } \pi(g) = 1_\Gamma, \\ 0 & \text{otherwise.} \end{cases} \tag{62}$$

Assume that for any $(g,k) \in G \times K$,

$$\tau(gkg^{-1}) = \tau(k). \tag{63}$$

Then $\tilde{\tau}$ extends linearly to a unital state on $(\mathbb{C}[G], 1, *)$.

Proof. Let us check that $\tilde{\tau}$ is tracial. For any $a, b \in G$, if $\pi(a) \neq \pi(b)^{-1}$, then $\pi(ab), \pi(ba) \neq 1_\Gamma$ and $\tilde{\tau}(ab) = \tilde{\tau}(ba) = 0$. Otherwise, thanks to (63), $\tilde{\tau}(ab) = \tau(ab) = \tau(babb^{-1}) = \tau(ba) = \tilde{\tau}(ba)$. Let us check now the positivity condition. Since π is surjective, there is a right-inverse map $s : \Gamma \to G$

https://doi.org/10.1017/fms.2024.152 Published online by Cambridge University Press

satisfying $\pi \circ s(\gamma) = \gamma$ for all $\gamma \in \Gamma$. Consider $x = \sum_{g \in G} \alpha_g g$ for some finitely supported sequence $(\alpha_g)_{g \in G}$. Then

$$\begin{aligned}\tilde{\tau}(xx^*) &= \sum_{a,b \in G} \alpha_a \overline{\alpha}_b \tilde{\tau}(ab^{-1}) = \sum_{a,b \in G : \pi(a)=\pi(b)} \alpha_a \overline{\alpha}_b \tau(ab^{-1}) \\ &= \sum_{\gamma \in \Gamma} \sum_{a,b \in K:} \alpha_{as(\gamma)} \overline{\alpha}_{bs(\gamma)} \tau\left(ab^{-1}\right) \\ &= \sum_{\gamma \in \Gamma} \tau\left(y_\gamma y_\gamma^*\right) \geq 0,\end{aligned}$$

where we set for any $\gamma \in \Gamma$, $y_\gamma = \sum_{a \in K} \alpha_{as(\gamma)} a$. □

When G is a group, let us denote by τ_{reg_G} and τ_{triv_G} the regular and the trivial states on $(\mathbb{C}[G], 1_G, *)$ defined by

$$\tau_{reg_G}(g) = \begin{cases} 1 & \text{if } g = 1_G, \\ 0 & \text{otherwise,} \end{cases} \text{ and } \tau_{triv_G}(g) = 1, \ \forall g \in G.$$

The following lemma is straightforward and gives states interpolating between regular representations of G and K.

Lemma 6.2. *Consider G, Γ, π and K as in Lemma 6.1 and $(\tau_T)_{T>0}$ a family of states on $(\mathbb{C}[K], 1, *)$ satisfying (63), such that for any $k \in K$,*

$$\lim_{T \to 0} \tau_T(k) = \tau_{triv_K}(k) \text{ and } \lim_{T \to \infty} \tau_T(k) = \tau_{reg_K}(k). \tag{64}$$

Then for any $g \in G$,

$$\lim_{T \to 0} \tilde{\tau}_T(g) = \tau_{reg_\Gamma} \circ \pi(g) \text{ and } \lim_{T \to \infty} \tilde{\tau}_T(g) = \tau_{reg_G}(g).$$

Let us consider two examples of extensions of the surface group Γ_g.

Extensions to the free group of even rank: Consider the free group $\mathbb{F}_{2g}$ in $2g$ generators $a_1, b_1, \ldots, a_g, b_g$ and the morphism

$$\pi : \mathbb{F}_{2g} \to \Gamma_g = \langle x_1, y_1 \ldots, x_g, y_g \mid [x_1, y_1] \ldots [x_g, y_g] \rangle$$

with $\pi(a_i) = x_i, \pi(b_i) = y_i$, $\forall i$ and $K = \ker(\pi)$. Identifying $\mathbb{F}_{2g}$ with $\Gamma_{1,g}$, this morphism coincides with $\Gamma_{1,g} \to \Gamma_g$ considered in 3 of Lemma 2.16, and accordingly, there is a right-inverse $s : \Gamma_g \to \mathbb{F}_{2g}$ such that K is free over

$$(w_\gamma)_{\gamma \in \Gamma_g} = (s(\gamma)[a_1, b_1] \ldots [a_g, b_g] s(\gamma)^{-1})_{\gamma \in \Gamma_g}.$$

Assume that $(\mu_T)_{T>0}$ is a family of measures on the unit circle such that for any integer $n \neq 1$,

$$\lim_{T \to 0} \int_{\mathbb{U}} \omega^n \mu_T(d\omega) = 1 \text{ and } \lim_{T \to \infty} \int_{\mathbb{U}} \omega^n \mu_T(d\omega) = 0.$$

Denote by $1 \in \mathbb{F}_{2g}$ the empty word and consider the unique state τ_T on $(\mathbb{C}[K], 1, *)$ such that under τ_T, $(w_\gamma)_{\gamma \in \Gamma_g}$ are freely independent and identically distributed with distribution μ_T.

Proposition 6.3. *For any $g \geq 1$, for all $T > 0$, $\tilde{\tau}_T$ is a state on $(\mathbb{C}[\mathbb{F}_{2g}], 1, *)$ with*

$$\lim_{T \to 0} \tilde{\tau}_T(w) = \tau_{reg_{\Gamma_g}} \circ \pi(w) \text{ and } \lim_{T \to \infty} \tilde{\tau}_T(w) = \tau_{reg_{\mathbb{F}_{2g}}} \circ \pi(w), \ \forall w \in \mathbb{F}_{2g}.$$

https://doi.org/10.1017/fms.2024.152 Published online by Cambridge University Press

Remark 6.4.

1. Mind that under $\tau_{reg_{\mathbb{F}_{2g}}}$, $a_1, b_1, \ldots, a_g, b_g$ are freely independent, whereas when $g = 1$, under $\tau_{reg_{\mathbb{Z}^2}}$, a_1, b_1 are classically independent. Hence, when $g = 1$, $(\tau_T)_{T>0}$ gives an interpolation between freely and classically independent Haar unitaries.
2. Recall that when $(\mu_T)_{T>0}$ is given by a free unitary Brownian motion, according to Lemma (3.10), τ_T can be identified with the restriction of the master field on $\tilde{\mathbb{G}}_g$ where all polygon faces have area T.

Proof. It is enough to prove that $(\tau_T)_{T>0}$ satisfies the assumptions of Lemma 6.2. We shall only prove (63) and leave the proof of the other conditions to the reader. According to Lemma 6.5, there is a surjective group morphism $p : P \to \mathbb{F}_{2g}$, a state η_T on $(\mathbb{C}[P], 1, *)$ and a subgroup L such that $p : L \to K$ is surjective with

$$\eta_T(\ell) = \tau_T \circ \pi(\ell), \ \forall \ell \in L.$$

Hence, for any $w \in \mathbb{F}_{2g}$ and $n \in K$, there are $\gamma \in P, \ell \in L$ with $p(\gamma) = w$, $p(\ell) = n$, and since η_T is a trace,

$$\tau_T(wnw^{-1}) = \tau_T(p(\gamma\ell\gamma^{-1})) = \eta_T(\gamma\ell\gamma^{-1}) = \eta_T(\ell) = \tau_T(n).$$

□

Following the same convention as in Section 2.3, consider the covering map $\tilde{\mathbb{G}}_g = (\tilde{V}, \tilde{E}, \tilde{F})$ of the 2g-bouquet map, $\mathbb{G}_g = (V, E, F)$, its $2g$ distinct edges $a_1, b_1, \ldots, a_g, b_g$, a vertex $r \in \tilde{V}$, an orientation $\tilde{E}_+$ of the edges of $\tilde{\mathbb{G}}_g$, and the free group $P = \mathbb{F}(\tilde{E}_+)$ over the $\tilde{E}_+$. When $e \in \tilde{E}_+$, let us identify the inverse of e in H with the edge e^{-1} of $\tilde{\mathbb{G}}$ with reverse orientation. Denote by $p : P \to \mathbb{F}_{2g}$ the group morphism mapping any edge $\tilde{e} \in \tilde{E}$ to its projection $p(e) \in E$. Note that we can identify any nontrivial reduced path of $(\tilde{V}, \tilde{E})$ with a (strict) subset of P, and through this identification, the group of reduced loops of $(\tilde{V}, \tilde{E})$ based at r is identified with a subgroup L_r of P such that

$$p : L_r \to \ker(\pi) = K$$

is an isomorphism. Let us fix a spanning tree $\mathcal{T}$ of $(\tilde{V}, \tilde{E})$. As in Lemma 2.16, consider the associated basis $(\omega^r_\gamma)_{\gamma\in\Gamma_g}$ of K and denote by $(\ell^r_{\gamma.r})_{\gamma\in\Gamma_g}$ its pre-image in L_r. Let us recall another basis of L_r. Denote by $E_+(\mathcal{T})$ the subset of edges of $\mathcal{T}$ in $\tilde{E}_+$. For any vertex $v \in \tilde{V}$, there is a unique reduced path in $\mathcal{T}$ from r to v, and we identify it with an element $[r, v]_\mathcal{T} \in P$. Then, setting for any $e \in \tilde{E}_+ \setminus E_+(\mathcal{T})$,

$$\ell_{r,e} = [r, v]_\mathcal{T} e [r, v]^{-1}_\mathcal{T}$$

defines a free basis of L_r indexed by $\tilde{E}_+ \setminus \mathcal{T}$. It is easy to check that the family $(\ell_{r,e})_{e\in E_+\setminus\mathcal{T}}, (e)_{e\in E_+(\mathcal{T})}$ forms a free basis of P. In particular, P is isomorphic to the free product

$$\mathbb{F}(\tilde{E}_+) = \mathbb{F}(E_+(T)) * L_r.$$

Consider now freely independent unitary noncommutative random variables indexed by $\tilde{E}_+$, such that a random variable of this family is Haar unitary if it is indexed by $E_+(\mathcal{T})$ and is distributed according to μ_T otherwise. Denote by $\eta_T : \mathbb{C}[\mathbb{F}(\tilde{E}_+)] \to \mathbb{C}$ its noncommutative distribution. Since the distribution of $(w_\gamma)_{\gamma\in\Gamma_g}$ under τ_T is identical to the one of $(\ell_{\gamma.r})_{\gamma\in\Gamma_g}$, the next lemma follows.

Lemma 6.5. *For any $T > 0$, $\eta_T : \mathbb{C}[\mathbb{F}(\tilde{E})] \to \mathbb{C}$ is a state, the morphism $p : \mathbb{F}(\tilde{E}_+) \to \mathbb{F}_{2g}$ is surjective, such that $p : L_r \to K$ is an isomorphism with*

$$\eta_T(\ell) = \tau_T(p(\ell)), \ \forall \ell \in L_r.$$

https://doi.org/10.1017/fms.2024.152 Published online by Cambridge University Press

Extension to the group of reduced loops: Consider a compact surface Σ and r a point of Σ. The set $\mathrm{L}_r(\Sigma)$ of Lipschitz[49] loop of Σ based at r is a monoid with multiplication given by concatenation, whose unit element is the constant loop at r. It can be turned into a group through the following quotient [34, 43]. Following [43, Sect. 6.7], let us say that a loop $\ell \in \mathrm{L}_r(\Sigma)$ is a *thin loop* if it is homotopic a the constant loop at r within its own range. For any pair $\ell, \ell' \in \mathrm{L}_r(\Sigma)$, let us define a binary relation setting $\ell \sim \ell'$ whenever $\ell\ell^{-1}$ is a thin loop. Let us recall the following.

Theorem 6.6 [43].

1. *The relation* $\sim$ *is an equivalence relation and* $\mathrm{RL}_r(\Sigma) = \mathrm{L}_r(\Sigma)/\sim$ *is a group.*
2. *When* $\Sigma = \mathbb{R}^2$ *or* $\mathbb{D}_{\mathfrak{h}}$, *the master field* Φ_Σ *on* Σ *satisfies*
   (a) *for any pair* $\ell, \ell' \in \mathrm{L}_r(\Sigma)$,

$$\ell \sim \ell' \Rightarrow \Phi_\Sigma(\ell) = \Phi_\Sigma(\ell') \tag{65}$$

and

$$\Phi_\Sigma(\ell) = 1 \Rightarrow \ell \sim 1. \tag{66}$$

Setting $\Phi_\Sigma(l) = \Phi_\Sigma(\ell)$ *for any* $\ell \in \mathrm{L}_r$ *with quotient image* $l \in \mathrm{RL}_r(\Sigma)$ *defines by linear extension a state* Φ_Σ *on the group algebra* $(\mathbb{C}[\mathrm{RL}_r(\Sigma)], 1, *)$.

   (b) *For any path* $a, b \in \mathrm{P}(\Sigma)$ *with* $\overline{a} = \underline{b}$ *and* $\overline{b} = \underline{a}$,

$$\Phi_\Sigma(ab) = \Phi_\Sigma(ba). \tag{67}$$

Consider now a compact orientable Riemannian manifold Σ, its foundamental cover $p : \tilde{\Sigma} \to \Sigma$, a point $\tilde{r}$ of $\tilde{\Sigma}$ and $r = p(\tilde{r})$. It is elementary to check that the map

$$\pi : \mathrm{RL}_r(\Sigma) \to \pi_1(\Sigma)$$

sending a based loop to its based-homotopy class is a group morphism and that its kernel is given by

$$K = p(\mathrm{RL}_{\tilde{r}}(\tilde{\Sigma})).$$

For any $l \in K$, let $\tilde{l} \in \mathrm{RL}_{\tilde{r}}(\tilde{\Sigma})$ be its unique lift starting at $\tilde{r}$.

Lemma 6.7. *Setting*

$$\Phi_\Sigma(l) = \begin{cases} \Phi_{\tilde{\Sigma}}(\tilde{l}) & \text{if } \pi(l) = 1, \\ 0 & \text{otherwise,} \end{cases} \tag{68}$$

and extending Φ_Σ *linearly defines a unital state on* $(\mathbb{C}[\mathrm{RL}_r(\Sigma)], 1, *)$.

Proof. Since $\Phi_{\tilde{\Sigma}} \circ p^{-1}$ defines a state on $(\mathbb{C}[K], 1, *)$, thanks to Lemma 6.1, it is enough to check (63). The latter follows from (67) applied to $\Phi_{\tilde{\Sigma}}$. □

6.2. Master field on the torus and t-freeness

Let us give here a proof of Corollary 1.12. For $T > 0$, let us consider the two-dimensional torus $\mathbb{T}^2_T$ obtained as the quotient $\mathbb{R}^2/\sqrt{T}\mathbb{Z}^2$ endowed with the push-forward of the Euclidean metric, so that it has total volume T. Denote by α and β the loop of $\mathbb{T}^2_T$ obtained by projecting the segments from $(0, 0)$ to, respectively, $(\sqrt{T}, 0)$ and $(0, \sqrt{T})$. Then, under YM_Σ, the law (a, b) on G^2 is given by (13). Therefore, for any word w in $\alpha, \beta, \alpha^{-1}, \beta^{-1}$ denoting by $[w] \in \mathbb{Z}^2$ the signed number of occurrences of α and β

[49] Recall the notation introduced in page 40.

https://doi.org/10.1017/fms.2024.152 Published online by Cambridge University Press

and by $\tilde{\gamma}_w$ the path of $\mathbb{R}^2$ starting from $(0,0)$ obtained by lifting the loop Σ formed by w, under YM_Σ, the following converge holds in probability as $N \to \infty$:

$$\tau_{\rho_N}(w) \to \begin{cases} \Phi_{\mathbb{R}^2}(\tilde{\gamma}_w) & \text{if } [\gamma_w] = 0 \\ 0 & \text{if } [\gamma_w] \neq 0. \end{cases}$$

The first statement of Corollary 1.12 follows considering the noncommutative distribution Φ_T of αt and β under the limit of τ_{ρ_N} as $N \to \infty$.

On the one hand, for any word w with $[w] = 0$, γ_w is a loop, and by continuity of the master field (Point 1 of Theorem 3.9), $\Phi_T(w) = \Phi_{\mathbb{R}^2}(\gamma_w) \to 1$ as $T \to 0$. On the other hand, for any word w with $[w] \neq 0$, $\tilde{\gamma}_w$ is not a loop, $[\gamma_w] \neq 0$, and for all $T > 0$, $\Phi_T(w) = 0$. Therefore, for any word in $\alpha, \beta, \alpha^{-1}, \beta^{-1}$, $\lim_{T\to 0} \Phi_T(w) = \tau_u \star_c \tau_u(w)$, since

$$\tau_u \star_c \tau_u(w) = \begin{cases} 1 & \text{if } [w] = 0, \\ 0 & \text{otherwise.} \end{cases}$$

Consider now the second limit of Corollary 1.12. When $(\mathbb{G}, a)$ is an area weighted map embedded in $\mathbb{R}^2$ with v a vertex of $\mathbb{G}$ sent to 0 by the embedding, consider the state $\hat{\tau}_T$ on $(\mathrm{RL}_v(\mathbb{G}), *)$ such that $\hat{\tau}_T(\ell) = \Phi_{\mathbb{R}^2}(\ell_T)$, where ℓ is the drawing of ℓ, while $\ell_T = \sqrt{T}\ell$. Consider a free basis of lassos $\ell_1, \ldots \ell_r$ of $\mathrm{RL}_v(\mathbb{G})$, with meanders given by distinct faces of area $a_1, \ldots, a_r$. Under $\hat{\tau}_T$, $\ell_1, \ldots, \ell_r$ are r independent free unitary Brownian motion marginals at time $\sqrt{T}a_1, \ldots, \sqrt{T}a_r$. It follows easily from its definition in moments that the free unitary Brownian motion at time s converges weakly towards a Haar unitary as $s \to \infty$. Since $a \in \Delta^o(T)$, $(\ell_1, \ldots, \ell_r)$ converges weakly toward r freely independent Haar unitary variables as $T \to \infty$. Therefore, for any reduced loop ℓ, $\lim_{T\to\infty} \hat{\tau}_T(\ell) = 1$ if ℓ is the constant loop and 0 otherwise. Now for any word w in $\alpha, \beta, \alpha^{-1}, \beta^{-1}$, with $[w] = 0$, it follows that

$$\lim_{T\to\infty} \Phi_T(w) = \begin{cases} 1 & \text{if } \gamma_w \sim_r c \text{ with } c \text{ constant,} \\ 0 & \text{otherwise.} \end{cases}$$

Since $\gamma_w \sim_r c$, where c is a constant loop if and only if w can be reduced to the empty word, it follows that $\lim_{T\to\infty} \Phi_T(w) = \tau_u \star \tau_u(w)$.

Let us now recall a way introduced in [5] to compute the evaluation of $\tau_{\mathcal{A}} \star_t \tau_{\mathcal{B}}$ given $\tau_{\mathcal{A}}$ and $\tau_{\mathcal{B}}$, solving systems of ODEs in the parameter t and present an argument for (14). Let us say that a noncommutative monomial P in $(X_{1,i})_{i\in I}, (X_{2,j})_{j\in I}$ is alternated if it is of the form $X_{\varepsilon_1,i_1} X_{\varepsilon_2,i_2} \ldots X_{\varepsilon_n,i_n}$ with $\varepsilon_k \neq \varepsilon_{k+1}$ for all $1 \leq k < n$. Denote by d_{X_2} is degree in the variables $(X_{2,j})_{i\in I}$. For such a monomial, let us set

$$\begin{aligned} \Delta_{ad}.P = &-\frac{d_{X_2}(P)}{2}(P \otimes 1 + 1 \otimes P) + \sum_{Q_1,Q_2,i} X_{2,i} \otimes Q_1 Q_2, \\ &- \sum_{P_{1,1},P_{1,2},P_2,i,j} \big[X_{2,i}P_2 \otimes (P_{1,1}X_{2,j}P_{1,2}) + (P_{1,1}X_{2,i}P_{1,2}) \otimes P_2 X_{2,j} \\ &\qquad - (P_{1,1}P_{1,2}) \otimes (X_{2,i}P_2X_{2,j}) - (P_{1,1}X_{2,i}X_{2,j}P_{1,2}) \otimes P_2\big], \end{aligned}$$

where the first sum is over all monomials Q_1, Q_2 and $i \in I$ such that $P = Q_1 X_{2,i} Q_2$, while the second is over all monomials $P_{1,1}, P_{1,2}, P_2$ and $i, j \in I$ such that $P = P_{1,1}X_{2,i}P_2X_{2,j}P_{1,2}$. With these notations, Theorem 3.4 of [5] states that for all alternated noncommutative monomial P in $(X_{1,i})_{i\in I}, (X_{2,j})_{j\in I}$, $\tau_{\mathcal{A}} \star_t \tau_{\mathcal{B}}(P)$ is differentiable with

$$\partial_t \tau_{\mathcal{A}} \star_t \tau_{\mathcal{B}}(P) = (\tau_{\mathcal{A}} \star_t \tau_{\mathcal{B}})^{\otimes 2}(\Delta_{ad}.P), \ \forall t \geq 0.$$

https://doi.org/10.1017/fms.2024.152 Published online by Cambridge University Press

For instance, assume that for all $t \geq 0$, (a, b) is a t-free couple within a noncommutative probability space $(\mathcal{C}, \tau_t)$, such that a and b are Haar unitaries for all $t > 0$. Then for any $n \geq 1$,

$$\partial_t \tau_t(ab^n) = -\tau_t(ab^n) + \tau(a)\tau_t(b^n) = -\tau_t(ab^n), \ \forall t \geq 0,$$

and since $\tau_0(ab^n) = \tau_0(a)\tau_0(b^n) = 0$,

$$\tau_t(ab^n) = 0.$$

Likewise, up to algebraic manipulations,

$$\partial_t \tau_t(ab^n a^*(b^*)^n) = -2\tau_t(ab^n a^*(b^*)^n).$$

Since $\tau_0(ab^n a^*(b^*)^n) = \tau_0(aa^*)\tau_0(b^n(b^*)^n) = 1$, this implies

$$\tau_t(ab^n a^*(b^*)^n) = e^{-2t}. \tag{69}$$

A similar argument together with (12) implies the following lemma.

Lemma 6.8.

1. *For any word* w *in* a, b, a^{-1}, b^{-1}, *if* $[w] \neq 0$,

$$\tau_t(w) = 0.$$

2. *For any* $n \geq 1$,

$$\partial_t \tau_t([a, b]^n) = -2n\tau_t([a, b]^n) - 2n \sum_{k=1}^{n-1} \tau_t([a, b]^k)\tau([a, b]^{n-k}).$$

3. *For any* $n \in \mathbb{Z}$ *and* $t \geq 0$,

$$\tau_t([a, b]^n) = v_{4t}(|n|).$$

The last equality of Corollary 1.12 follows from the last point of the above Lemma. Besides, for any $t > 0, T > 0$

$$\tau_u \star_t \tau_u(XYX^*Y^*) = e^{-2t} \text{ and } \Phi_T(XYX^*Y^*) = e^{-\frac{T}{2}},$$

so that if $\tau_u \star_t \tau_u = \Phi_T$, then $T = 4t$. But (69) implies $\tau_u \star_t \tau_u(XY^2X^*Y^{-2}) = e^{-2t} > e^{-4t} = \Phi_{4t}(XY^2X^*Y^{-2})$. Therefore, for all $t, T > 0$, $\Phi_T \neq \tau_u \star_t \tau_u$.

7. Appendix

7.1. Casimir element and trace formulas

Let us recall some tensor identities instrumental to prove Makeenko–Migdal relations.

Definition 7.1. Consider a Lie algebra $\mathfrak{g}$ endowed with an inner product $\langle \cdot, \cdot \rangle$. The *Casimir element* of $(\mathfrak{g}, \langle \cdot, \cdot \rangle)$ is the tensor $C_{\mathfrak{g}} \in \mathcal{M}_d(\mathbb{C}) \otimes_{\mathbb{R}} \mathcal{M}_d(\mathbb{C})$ defined by

$$C_{\mathfrak{g}} = \sum_{X \in \mathcal{B}} X \otimes X, \tag{70}$$

where $\mathcal{B}$ is an orthonormal basis of $\mathfrak{g}$ for the inner product $\langle \cdot, \cdot \rangle$.

It is simple to check that the definition of the Casimir element does not depend on the choice of the basis but only on the inner product $\langle \cdot, \cdot \rangle$. We focus on the setting recalled in Section 3.1; we consider the Lie algebra $\mathfrak{g}_N$ of a compact classical group G_N with the inner product (1) considered in [20, Section 2.1.]. We set the value β to be, respectively, 1 and 4 when G_N is $O(N)$ and $Sp(N)$ and 2 otherwise – that is, when G_N is $SU(N)$ or $U(N)$. We set $\gamma = 1$ when $G_N = SU(N)$ and 0 otherwise.

Most of the following results can be proved by a direct computation using an arbitrary chosen basis. For any $(a, b) \in \{1, \dots, N\}^2$, the elementary matrix $E_{ab} \in \mathcal{M}_N(\mathbb{R})$ is defined by $(E_{ab})_{ij} = \delta_{ai}\delta_{bj}$.

We shall need the following standard result on the Casimir element in this setting, which gives computation rules for traces of products and product of traces involving elements of $\mathcal{B}$.

Lemma 7.2. *For any $A, B \in G_N$ we have*

$$\sum_{X \in \mathcal{B}} \mathrm{tr}(AXBX) = -\mathrm{tr}(A)\mathrm{tr}(B) - \frac{\beta - 2}{\beta N}\mathrm{tr}(AB^{-1}) + \frac{\gamma}{N^2}\mathrm{tr}(AB) \tag{71}$$

and

$$\sum_{X \in \mathcal{B}} \mathrm{tr}(AX)\mathrm{tr}(BX) = -\mathrm{tr}(AB) - \frac{\beta - 2}{\beta N}\mathrm{tr}(AB^{-1}) + \gamma\mathrm{tr}(A)\mathrm{tr}(B). \tag{72}$$

Proof. We only sketch the proof in order to show where the expressions come from. First of all, remark that by linearity they only need to be proved for $A = E_{ij}$ and $B = E_{k\ell}$. We have, for instance,

$$\sum_{X \in \mathcal{B}} \mathrm{tr}(AXBX) = \frac{1}{N} \sum_{X \in \mathcal{B}} \sum_{a,b,c,d} A_{ab}X_{bc}B_{cd}X_{da} = \frac{1}{N}(C_{\mathfrak{g}})_{jk\ell i},$$

where we have set

$$\left(\sum_i X^i \otimes Y^i\right)_{abcd} = \sum_i X^i_{ab} Y^i_{cd}.$$

Using the expression of $C_{\mathfrak{g}}$ for each value of $\mathfrak{g}$ leads to Equation (71). By similar computations, we also obtain Equation (72). □

In the unitary case, the formulas in Lemma 7.2 are known as the 'magic formulas', as stated in [25] for instance, and appeared already in [59]; they are crucial to the derivation of Makeenko–Migdal equations for Wilson loops that we briefly recall in the next section. Although we do not detail it, there exists a beautiful interpretation of Lemma 7.2 in terms of Schur–Weyl duality; the interested reader can refer to [41] or [19] for an explanation and discussion of this fact and to [43, Chap. I, Section 1.2] about the above Lemma.

7.2. Makeenko–Migdal equations

Given a topological map $\mathbb{G}$ of genus g with m edges, a vertex of $\mathbb{G}$ will be said to be an *admissible crossing* if it possesses four outgoing edges labelled e_1, e_2, e_3, e_4 counterclockwise.

Definition 7.3. Let $\mathbb{G}$ be map of genus g with m edges, and v be an admissible crossing. A function $f : G^m \to \mathbb{C}$ has an *extended gauge invariance at v* if for any $x \in G$,

$$f(a_1, a_2, a_3, a_4, \mathbf{b}) = f(a_1x, a_2, a_3x, a_4, \mathbf{b}) = f(a_1, a_2x, a_3, a_4x, \mathbf{b}), \tag{73}$$

where a_i denotes the variable associated to the edge e_i and $\mathbf{b}$ denotes the tuple of other edge variables than e_1, e_2, e_3, e_4.

https://doi.org/10.1017/fms.2024.152 Published online by Cambridge University Press

The extended gauge-invariance was first introduced by Lévy in [42] to prove Makeenko–Migdal equations in the plane, then used in [26] to give alternative, local proofs of these equations, which allowed in [25] to prove their validity on any surface; these last equations were then applied in [21, 33].

Theorem 7.4 (Abstract Makeenko–Migdal equations). *Let* $(\mathbb{G}, a)$ *be an area weighted map of area* T *and genus* g *with* m *edges, and* $f : G^m \to \mathbb{C}$ *be a function with extended gauge invariance at an admissible crossing* v*. Denote by* f_1 *(resp.* f_2, f_3, f_4*) the face of* $\mathbb{G}$ *whose boundary contains* (e_1, e_2) *(resp.* (e_2, e_3), (e_3, e_4), (e_4, e_1)*). Denote by* t_i *the area of the face* f_i*, choose an orthonormal basis* $\mathcal{B}$ *of* $\mathfrak{g}$ *with respect to the chosen inner product, and set*

$$(\nabla^{a_1} \cdot \nabla^{a_2} f)(a_1, a_2, a_3, a_4, \mathbf{b}) = \sum_{X \in \mathcal{B}} \frac{\partial^2}{\partial s \partial t} f(a_1 e^{sX}, a_2 e^{tX}, a_3, a_4, \mathbf{b})\big|_{s=t=0}.$$

We have

$$\left(\frac{\partial}{\partial t_1} - \frac{\partial}{\partial t_2} + \frac{\partial}{\partial t_3} - \frac{\partial}{\partial t_4}\right) \int_{G^m} f d\mu = - \int_{G^m} \nabla^{a_1} \cdot \nabla^{a_2} f d\mu. \tag{74}$$

Equation (74) might be confusing, as it involves partial derivatives with respect to variables that do not appear explicitly in the function $\int_{G^m} f d\mu$; it becomes, in fact, clearer after being translated in terms of the area simplex. We define the differential operator μ_v on functions $\Delta_{\mathbb{G}}(T) \to \mathbb{C}$ by

$$\mu_v = \frac{\partial}{\partial a_1} - \frac{\partial}{\partial a_2} + \frac{\partial}{\partial a_3} - \frac{\partial}{\partial a_4},$$

using the labelling of $a = (a_1, \ldots, a_p) \in \Delta_{\mathbb{G}}(T)$ such that a_i corresponds to the face f_i. Equation (74) becomes then

$$\mu_v \mathbb{E}(f) = -\mathbb{E}(\nabla^{a_1} \cdot \nabla^{a_2} f),$$

and now everything only depends on the areas of the faces. We want to apply these abstract Makeenko–Migdal equations to functionals of Wilson loops in order to obtain the convergence to the master field. We define, for k unrooted loops $\ell_1, \ldots, \ell_k \in \mathrm{L}_c(\mathbb{G})$, the k-point function $\phi^G_{\ell_1,\ldots,\ell_k} : \Delta_{\mathbb{G}}(T) \to \mathbb{C}$ by

$$\phi^G_{\ell_1 \otimes \ldots \otimes \ell_k} = \mathbb{E}(W_{\ell_1} \cdots W_{\ell_k}).$$

and extend it linearly to $\mathbb{C}[\mathrm{L}_c(\mathbb{G})]^{\otimes k}$. The following proposition offers an estimate of the face-area variation of the functions $\phi^G_{\ell_1 \otimes \ldots \otimes \ell_k}$.

Proposition 7.5 (Makeenko–Migdal equations for Wilson loops). *Assume that* G_N *is a compact classical group and* $\langle \cdot, \cdot \rangle$ *is fixed as in Section 3.1. Let* $(\mathbb{G}, a)$ *be a weighted map of area* T *and genus* g *with* m *edges, and* v *be an admissible crossing in* $\mathbb{G}$*.*

1. *If* v *is a self-intersection of a single loop* ℓ_1 *such that the edges* $(e_j^{\pm 1}, 1 \leq j \leq 4)$ *are visited in the following order:* $e_1, e_4^{-1}, e_2, e_3^{-1}$*, then define* ℓ_{11} *the sub-loop of* ℓ_1 *starting at* e_1 *and finishing at* e_4^{-1}*,* ℓ_{12} *the subloop starting at* e_2 *and finishing at* e_3^{-1}*. We have, for any loops* $\ell_2, \ldots, \ell_k$ *that do not cross* v*,*

$$\mu_v \phi^G_{\ell_1 \otimes \ldots \otimes \ell_k} = \phi^G_{\ell_{11} \otimes \ell_{12} \otimes \ell_2 \otimes \ldots \otimes \ell_k} + \frac{\beta - 2}{\beta N} \phi^G_{\ell_{11} \ell_{12}^{-1} \otimes \ell_2 \otimes \ldots \otimes \ell_k} - \frac{\gamma}{N^2} \phi_{\ell_1 \otimes \ldots \otimes \ell_k}, \tag{75}$$

$$\mu_v \phi^G_{\ell_1 \otimes \ell_1^{-1}} = \phi^G_{\ell_{11} \otimes \ell_{12} \otimes \ell_1^{-1}} + \phi^G_{\ell_1 \otimes \ell_{11}^{-1} \otimes \ell_{12}^{-1}} + \frac{R_{\ell_1}}{N}, \tag{76}$$

where the $|R_{\ell_1}| \leq 10$ *uniformly on* $\Delta_{\mathbb{G}}(T)$*.*

https://doi.org/10.1017/fms.2024.152 Published online by Cambridge University Press

2. *If v is the intersection between two loops ℓ_1 and ℓ_2 such that ℓ_1 starts at e_1 and finishes at e_3^{-1}, and ℓ_2 starts at e_2 and finishes at e_4^{-1}, then define ℓ the loop obtained by concatenation of ℓ_1 and ℓ_2. We have, for any loops $\ell_3, \ldots, \ell_k$ that do not cross v,*

$$\mu_v \phi^G_{\ell_1 \otimes \ell_2 \otimes \ldots \otimes \ell_k} = \frac{R_{\ell_1 \otimes \ell_2 \otimes \ldots \otimes \ell_k}}{N^2} \tag{77}$$

with $|R_{\ell_1 \otimes \ell_2 \otimes \ldots \otimes \ell_k}| \leq 3$ uniformly on $\Delta_{\mathbb{G}}(T)$.

It was proved for all classical Lie algebras if $\mathbb{G}$ is a planar combinatorial graph by Lévy in [43, Prop. 6.16] when the loops form what he called a skein. If $\mathbb{G}$ is a map of genus 0 and $\mathfrak{g}$ is the Lie algebra of $\mathrm{U}(N)$, this result was proved by the first author with Norris in [21, Prop. 4.3]. See also [26, Thm. 1.1]

Proof of Proposition 7.5. Let us start with the first case, which is when v is a self-intersection of a loop ℓ_1. We take $E = \{e_1, e_2, e_3, e_4, e'_1, \ldots, e'_{m-4}\}$ as an orientation of E, with e_1, e_2, e_3, e_4 the four outgoing edges from v. We identify any multiplicative function $h \in \mathcal{M}(P(\mathbb{G}), G)$ to a tuple $(a_1, a_2, a_3, a_4, \mathbf{b})$ by setting $a_i = h_{e_i}$ and $\mathbf{b} = (h_{e'_i})_{1 \leq i \leq m-4}$ the tuple of all other images of edges by h. There are words $\alpha, \beta, w_2, \ldots, w_k$ in the elements of $\mathbf{b}$ such that

$$h_{\ell_1} = a_3^{-1} \alpha a_2 a_4^{-1} \beta a_1, h_{\ell_i} = w_i \ \forall 2 \leq i \leq k.$$

It appears that $\phi^G_{\ell_1, \ldots, \ell_k} = E(f)$, where f is the extended gauge-invariant function

$$f : \begin{cases} G^m & \to \mathbb{C} \\ (a_1, a_2, a_3, a_4, \mathbf{b}) & \mapsto \mathrm{tr}(a_3^{-1} \alpha a_2 a_4^{-1} \beta a_1) \mathrm{tr}(w_2) \cdots \mathrm{tr}(w_k). \end{cases}$$

Then, by the abstract Makeenko–Migdal equation (74), we get

$$\mu_v \mathbb{E}(f) = -\mathbb{E}(\nabla^{a_1} \cdot \nabla^{a_2} f),$$

and by definition,

$$\nabla^{a_1} \cdot \nabla^{a_2} f = \left(\sum_X \mathrm{tr}(a_3^{-1} \alpha a_2 X a_4^{-1} \beta a_1 X) \right) \mathrm{tr}(w_2) \cdots \mathrm{tr}(w_k),$$

where X runs through an orthonormal basis of $\mathfrak{g}$. A straightforward application of (71) from Lemma 7.2 yields (75), by noticing that $h_{\ell_{11}} = a_4^{-1} \beta a_1$ and $h_{\ell_{12}} = a_3^{-1} \alpha a_2$.

Similarly, we have $\phi^G_{\ell, \ell^{-1}} = E(f')$, where

$$f' : \begin{cases} G^m & \to \mathbb{C} \\ (a_1, a_2, a_3, a_4, \mathbf{b}) & \mapsto \mathrm{tr}(a_3^{-1} \alpha a_2 a_4^{-1} \beta a_1) \mathrm{tr}(a_1^{-1} \beta^{-1} a_4 a_2^{-1} \alpha^{-1} a_3). \end{cases}$$

We have

$$\begin{aligned} \nabla^{a_1} \cdot \nabla^{a_2} f' = \sum_X \big\{ & \mathrm{tr}(a_3^{-1} \alpha a_2 X a_4^{-1} \beta a_1 X) \mathrm{tr}(a_1^{-1} \beta^{-1} a_4 a_2^{-1} \alpha^{-1} a_3) \\ & - \mathrm{tr}(a_3^{-1} \alpha a_2 a_4^{-1} \beta a_1 X) \mathrm{tr}(a_1^{-1} \beta^{-1} a_4 X a_2^{-1} \alpha^{-1} a_3) \\ & - \mathrm{tr}(a_3^{-1} \alpha a_2 X a_4^{-1} \beta a_1) \mathrm{tr}(X a_1^{-1} \beta^{-1} a_4 a_2^{-1} \alpha^{-1} a_3) \\ & + \mathrm{tr}(a_3^{-1} \alpha a_2 a_4^{-1} \beta a_1) \mathrm{tr}(X a_1^{-1} \beta^{-1} a_4 X a_2^{-1} \alpha^{-1} a_3) \big\}, \end{aligned}$$

and a simultaneous application of (71) and (72) leads to the result. We detail the case of $\mathrm{SU}(N)$ and leave the others as an exercise: if we set $A = h_{\ell_{11}}$ and $B = h_{\ell_{12}}$, then

https://doi.org/10.1017/fms.2024.152 Published online by Cambridge University Press

$$\sum \mathrm{tr}(AXBX)\mathrm{tr}(B^{-1}A^{-1}) = -\mathrm{tr}(A)\mathrm{tr}(B)\mathrm{tr}((AB)^{-1}) + \frac{1}{N}\mathrm{tr}(AB)\mathrm{tr}((AB)^{-1})$$
$$\sum \mathrm{tr}(ABX)\mathrm{tr}(B^{-1}XA^{-1}) = -\frac{1}{N^2}\mathrm{tr}([A,B]) + \frac{1}{N}\mathrm{tr}(AB)\mathrm{tr}(A^{-1}B^{-1})$$
$$\sum \mathrm{tr}(AXB)\mathrm{tr}(XB^{-1}A^{-1}) = -\frac{1}{N^2}\mathrm{tr}([A,B]^{-1}) + \frac{1}{N}\mathrm{tr}(BA)\mathrm{tr}(B^{-1}A^{-1})$$
$$\sum \mathrm{tr}(AB)\mathrm{tr}(XB^{-1}XA^{-1}) = -\mathrm{tr}(A^{-1})\mathrm{tr}(B^{-1})\mathrm{tr}((AB)) + \frac{1}{N}\mathrm{tr}(AB)\mathrm{tr}(A^{-1}B^{-1}),$$

where all sums are over X in the orthonormal basis. We can then take the expectation of the alternated sum of these expressions, and as all traces are bounded by 1 because they apply to special unitary matrices, we find that all terms with a coefficient $\frac{1}{N}$ or $\frac{1}{N^2}$ fall into $O\big(\frac{1}{N}\big)$ which does not depend on any loop,[50] so that

$$\phi^{\mathrm{SU}(N)}_{\ell_1\otimes\ell_1^{-1}} = \phi^{\mathrm{SU}(N)}_{\ell_{11}\otimes\ell_{12}\otimes(\ell_{11}\ell_{12})^{-1}} + \phi^{\mathrm{SU}(N)}_{\ell_{11}^{-1}\otimes\ell_{12}^{-1}\otimes(\ell_{11}\ell_{12})} + O\Big(\frac{1}{N}\Big).$$

Let us now turn to the second case, when v is the intersection of ℓ_1 and ℓ_2. We take $E = \{e_1, e_2, e_3, e_4, e'_1, \ldots, e'_{m-4}\}$ as an orientation of E, with e_1, e_2, e_3, e_4 the four outgoing edges from v. There are words $\alpha, \beta, w_2, \ldots, w_k$ in the elements of $\mathbf{b}$ such that

$$h_{\ell_1} = a_3^{-1}\alpha a_1, h_{\ell_2} = a_4^{-1}\alpha a_2, h_{\ell_i} = w_i \ \forall 3 \le i \le k.$$

We have $\phi^G_{\ell_1,\ldots,\ell_k} = E(f)$, where f is the extended gauge-invariant function

$$f : \begin{cases} G^m & \to & \mathbb{C} \\ (a_1, a_2, a_3, a_4, \mathbf{b}) & \mapsto & \mathrm{tr}(a_3^{-1}\alpha a_1)\mathrm{tr}(a_4^{-1}\beta a_2)\mathrm{tr}(w_2)\cdots\mathrm{tr}(w_k), \end{cases}$$

then

$$\mu_v E(f) = -E(\nabla^{a_1}\cdot\nabla^{a_2} f),$$

where

$$\nabla^{a_1}\cdot\nabla^{a_2} f = \left(\sum_X \mathrm{tr}(a_3^{-1}\alpha a_1)\mathrm{tr}(Xa_4^{-1}\beta a_2 X)\right)\mathrm{tr}(w_2)\cdots\mathrm{tr}(w_k).$$

The result follows then from (72). □

By letting $N \to \infty$ in Proposition 7.5, one immediately gets the following.

Corollary 7.6 (Makeenko–Migdal equations for a master field). *Assume for some some sequence $(G_N)_N$ of compact classical groups, we have for all maps $\mathbb{G}$ of genus $g \ge 1$ and $\ell \in \mathrm{L}(\mathbb{G})$, $\lim_{N\to\infty}\Phi^{G_N}_\ell$ and $\lim_{N\to\infty}\Phi^{G_N}_{\ell\otimes\ell^{-1}} = |\Phi_\ell|^2$ uniformly on $\Delta_{\mathbb{G}}(T)$. Then Φ defines an exact solution of the Makeenko–Migdal solution as defined in Section 3.5.*

To address uniqueness questions, it is convenient to work with *centered Wilson loops*. Define, for any $\ell_1, \ldots, \ell_k$ in an area-weighted graph $(\mathbb{G}, a)$,

$$\psi^G_{\ell_1\otimes\ldots\otimes\ell_k} = \mathbb{E}\left[\prod_{i=1}^k (W_{\ell_i} - \Phi_{\ell_i})\right].$$

[50] We add up a finite number of terms, 6 to be precise, which are bounded by $\frac{1}{N}$, so their sum is bounded by $\frac{6}{N}$ which is indeed independent from the loops or the face-area vector.

https://doi.org/10.1017/fms.2024.152 Published online by Cambridge University Press

Proposition 7.7 (Makeenko–Migdal equations for centered Wilson loops). *Assume $g \geq 0$, $T > 0$, $\ell \in \mathfrak{L}_g$, $v \in V_\ell$ with $\delta_v \ell = \ell_1 \otimes \ell_2$. Then for any compact classical group G_N,*

$$\begin{aligned}\mu_v \psi^G_{\ell\otimes\ell^{-1}} = \psi^G_{\ell_1\otimes\ell_2\otimes\ell^{-1}} + \psi^G_{\ell_1^{-1}\otimes\ell_2^{-1},\ell} + \psi^G_{\ell_1\otimes\ell^{-1}}\Phi_{\ell_2} + \psi^G_{\ell_1^{-1}\otimes\ell}\Phi_{\ell_2^{-1}} \\ + \psi^G_{\ell_2\otimes\ell^{-1}}\Phi_{\ell_1} + \psi_{\ell_2^{-1}\otimes\ell}\Phi_{\ell_1^{-1}} + \frac{R_\ell}{N},\end{aligned} \tag{78}$$

where the $|R_\ell| \leq 10$ uniformly on $\Delta_{\mathbb{G}}(T)$. There is a constant C_ℓ independent of G, such that for all $X \in \mathfrak{m}_\ell$,

$$\begin{aligned}\mu_v \psi^G_{\ell\otimes\ell^{-1}} = \psi_{\delta_X(\ell)\otimes\ell^{-1}} + \psi^G_{\ell,\delta_X(\ell^{-1})} + \psi_{\ell_1\otimes\ell^{-1}}\Phi_{\ell_2} + \psi_{\ell_1^{-1}\otimes\ell}\Phi_{\ell_2^{-1}} \\ + \psi^G_{\ell_2\otimes\ell^{-1}}\Phi_{\ell_1} + \psi_{\ell_2^{-1}\otimes\ell}\Phi_{\ell_1^{-1}} + \frac{R_\ell}{N},\end{aligned}$$

with $|R_\ell| \leq 10$ uniformly on $\Delta_{\mathbb{G}}(T)$.

Acknowledgements. A. D. acknowledges partial support from ‘The Dr Perry James (Jim) Browne Research Centre on Mathematics and its Applications’. T. L. acknowledges support from ANR AI chair BACCARAT (ANR-20-CHIA-0002). Many thanks are due to the reviewers for their careful reading and useful comments.

Competing interest. The authors have no competing interest to declare.

Data availability statement. No data was produced in connection with this research.

Ethical standards. The research meets all ethical guidelines, including adherence to the legal requirements of the study country.

Author contributions. Both authors contributed equally to this article.

References

[1] S. Albeverio, R. Høegh-Krohn and H. Holden, ‘Stochastic Lie group-valued measures and their relations to stochastic curve integrals, gauge fields and Markov cosurfaces’, in *Stochastic Processes—Mathematics and Physics (Bielefeld, 1984)* (Lecture Notes in Math.) vol. 1158 (Springer, Berlin, 1986), 1–24.
[2] S. Albeverio, R. Høegh-Krohn and H. Holden, ‘Stochastic multiplicative measures, generalized Markov semigroups, and group-valued stochastic processes and fields’, *J. Funct. Anal.* **78**(1) (1988), 154–184.
[3] M. Anshelevich and A. N. Sengupta, ‘Quantum free Yang–Mills on the plane’, *J. Geom. Phys.* **62**(2) (2012), 330–343.
[4] M. F. Atiyah and R. Bott, ‘The Yang-Mills equations over Riemann surfaces’, *Philos. Trans. Roy. Soc. London Ser. A* **308**(1505) (1983), 523–615.
[5] F. Benaych-Georges and T. Lévy, ‘A continuous semigroup of notions of independence between the classical and the free one’, *Ann. Probab.* **39**(3) (2011), 904–938.
[6] P. Biane, ‘Free Brownian motion, free stochastic calculus and random matrices’, in *Free Probability Theory (Waterloo, ON, 1995)* (Fields Inst. Commun.) vol. 12 (Amer. Math. Soc., Providence, RI, 1997), 1–19.
[7] J. S. Birman and C. Series, ‘Dehn’s algorithm revisited, with applications to simple curves on surfaces’, in *Combinatorial Group Theory and Topology (Alta, Utah, 1984)* (Ann. of Math. Stud.) vol. 111 (Princeton Univ. Press, Princeton, NJ, 1987), 451–478.
[8] J.-M. Bismut and F. Labourie, ‘Symplectic geometry and the Verlinde formulas’, in *Surveys in Differential Geometry: Differential Geometry Inspired by String Theory* (Surv. Differ. Geom.) vol. 5 (Int. Press, Boston, MA, 1999), 97–311.
[9] G. Cébron, A. Dahlqvist and F. Gabriel, ‘The generalized master fields’, *J. Geom. Phys.* **119** (2017), 34–53.
[10] A. Chandra, I. Chevyrev, M. Hairer and H. Shen, ‘Langevin dynamic for the 2D Yang-Mills measure’, *Publ. Math. Inst. Hautes Études Sci.* **136** (2022), 1–147.
[11] A. Chandra, I. Chevyrev, M. Hairer and H. Shen, ‘Stochastic quantisation of yang–mills–higgs in 3d’, *Invent. Math.* **237**(2) (2024), 541–696.
[12] S. Chatterjee, ‘Rigorous solution of strongly coupled so(n) lattice gauge theory in the large n limit’, *Comm. Math. Phys.* **366**(1) (2019), 203–268.
[13] I. Chevyrev, ‘Yang-Mills measure on the two-dimensional torus as a random distribution’, *Comm. Math. Phys.* **372**(3) (2019), 1027–1058.
[14] I. Chevyrev, ‘Stochastic quantization of Yang-Mills’, *J. Math. Phys.* **63**(9) (2022), Paper No. 091101, 19.
[15] I. Chevyrev and H. Shen, ‘ *Invariant measure and universality of the 2d yang-mills langevin dynamic*’, 2023.

[16] B. Collins, A. Guionnet and E. Maurel-Segala, ‘Asymptotics of unitary and orthogonal matrix integrals’, *Adv. Math.* **222**(1) (2009), 172–215.
[17] B. Collins and P. Śniady, ‘Integration with respect to the Haar measure on unitary, orthogonal and symplectic group’, *Comm. Math. Phys.* **264**(3) (2006), 773–795.
[18] A. Dahlqvist, ‘Free energies and fluctuations for the unitary Brownian motion’, *Comm. Math. Phys.* **348**(2) (2016), 395–444.
[19] A. Dahlqvist, ‘Integration formulas for Brownian motion on classical compact Lie groups’, *Ann. Inst. Henri Poincaré Probab. Stat.* **53**(4) (2017), 1971–1990.
[20] A. Dahlqvist and T. Lemoine, ‘Large N limit of Yang–Mills partition function and Wilson loops on compact surfaces’, *Probab. Math. Phys.* **4**(4) (2023), 849–890.
[21] A. Dahlqvist and J. R. Norris, ‘Yang–mills measure and the master field on the sphere’, *Comm. Math. Phys.* **377**(2) (2020), 1163–1226.
[22] P. Diaconis and S. N. Evans, ‘Linear functionals of eigenvalues of random matrices’, *Trans. Amer. Math. Soc.* **353**(7) (2001), 2615–2633.
[23] B. K. Driver, ‘YM_2: continuum expectations, lattice convergence, and lassos’, *Comm. Math. Phys.* **123**(4) (1989), 575–616.
[24] B. K. Driver, ‘A functional integral approaches to the Makeenko-Migdal equations’, *Comm. Math. Phys.* **370**(1) (2019), 49–116.
[25] B. K. Driver, F. Gabriel, B. C. Hall and T. Kemp, ‘The Makeenko-Migdal equation for Yang-Mills theory on compact surfaces’, *Comm. Math. Phys.* **352**(3) (2017), 967–978.
[26] B. K. Driver, B. C. Hall and T. Kemp, ‘Three proofs of the Makeenko-Migdal equation for Yang-Mills theory on the plane’, *Comm. Math. Phys.* **351**(2) (2017), 741–774.
[27] W. M. Goldman, ‘The symplectic nature of fundamental groups of surfaces’, *Adv. Math.* **54**(2) (1984), 200–225.
[28] R. Gopakumar and D. J. Gross, ‘Mastering the master field’, *Nuclear Phys. B* **451**(1–2) (1995), 379–415.
[29] D. J. Gross and A. Matytsin, ‘Some properties of large-n two-dimensional yang-mills theory’, *Nuclear Phys. B* **437**(3) (1995), 541–584.
[30] D. J. Gross and W. Taylor, IV, ‘Two-dimensional QCD is a string theory’, *Nuclear Phys. B* **400**(1–3) (1993), 181–208.
[31] L. Gross, ‘The Maxwell equations for Yang-Mills theory’, in *Mathematical Quantum Field Theory and Related Topics (Montreal, PQ, 1987)* (Conf. Proc.) vol. 9 (Amer. Math. Soc., Providence, RI, 1988), 193–203.
[32] L. Gross, C. King and A. Sengupta, ‘Two-dimensional Yang-Mills theory via stochastic differential equations’, *Ann. Physics* **194**(1) (1989), 65–112.
[33] B. C. Hall, ‘The large-N limit for two-dimensional Yang–Mills theory’, *Comm. Math. Phys.* **363**(3) (2018), 789–828.
[34] B. Hambly and T. Lyons, ‘Uniqueness for the signature of a path of bounded variation and the reduced path group’, *Ann. of Math. (2)* **171**(1) (2010), 109–167.
[35] A. Kassel and T. Lévy, ‘Determinantal probability measures on Grassmannians’, *Ann. Inst. Henri Poincaré D* **9**(4) (2022), 659–732.
[36] [36] V. A. Kazakov, ‘Wilson loop average for an arbitrary contour in two-dimensional U(N) gauge theory’, *Nuclear Phys. B* **179**(2) (1981), 283–292.
[37] V. A. Kazakov and I. K. Kostov, ‘Nonlinear strings in two-dimensional U(∞) gauge theory’, *Nuclear Phys. B* **176**(1) (1980), 199–215.
[38] R. Kenyon, ‘Lectures on dimers’, in *Statistical Mechanics (IAS/Park City Math. Ser.)* vol. 16 (Amer. Math. Soc., Providence, RI, 2009), 191–230.
[39] T. Lévy, ‘Yang-Mills measure on compact surfaces’, *Mem. Amer. Math. Soc.* **166**(790) (2003), xiv+122.
[40] T. Lévy, ‘Discrete and continuous Yang-Mills measure for non-trivial bundles over compact surfaces’, *Probab. Theory Related Fields* **136**(2) (2006), 171–202.
[41] T. Lévy, ‘Schur-Weyl duality and the heat kernel measure on the unitary group’, *Adv. Math.* **218**(2) (2008), 537–575.
[42] T. Lévy, ‘Two-dimensional Markovian holonomy fields’, *Astérisque* **329** (2010), 172.
[43] T. Lévy, ‘The master field on the plane’, *Astérisque* **388** (2017), ix+201.
[44] T. Lévy, *Two-Dimensional Quantum Yang–Mills Theory and the Makeenko–Migdal Equations* (Springer International Publishing, Cham, 2020), 275–325.
[45] K. Liechty and D. Wang, ‘Nonintersecting brownian motions on the unit circle’, *Ann. Probab.* **44**(2) (2016), 1134–1211.
[46] K. Liu, ‘Heat kernels, symplectic geometry, moduli spaces and finite groups’, in *Surveys in Differential Geometry: Differential Geometry Inspired by String Theory* (Surv. Differ. Geom.) vol. 5 (Int. Press, Boston, MA, 1999), 527–542.
[47] M. Magee, ‘Random unitary representations of surface groups II: The large n limit’, Preprint, 2021, arXiv:2101.03224.
[48] M. Magee, ‘Random unitary representations of surface groups I: asymptotic expansions’, *Comm. Math. Phys.* **391**(1) (2022), 119–171.
[49] Y. Makeenko and A. A. Migdal, ‘Exact equation for the loop average in multicolor QCD’, *Phys. Lett. B* **88** (1979), 135–137.
[50] P. Mergny and M. Potters, ‘Rank one HCIZ at high temperature: interpolating between classical and free convolutions’, *SciPost Phys.* **12** (2022), 022.
[51] A. A. Migdal, ‘Recursion equations in gauge field theories’, *Sov. Phys. JETP* (1975) 413–418.
[52] J. A. Mingo and R. Speicher, *Free Probability and Random Matrices* (Fields Institute Monographs) vol. 35 (Springer, New York; Fields Institute for Research in Mathematical Sciences, Toronto, ON, 2017).
[53] W. Mlotkowski, ‘Λ-free probability’, *Infin. Dimens. Anal. Quantum Probab. Relat. Top.* **7**(1) (2004), 27–41.

[54] J. Moser, 'On the volume elements on a manifold', *Trans. Amer. Math. Soc.* **120** (1965), 286–294.
[55] M. Park, J. Pfeffer, S. Sheffield and P. Yu, 'Wilson loop expectations as sums over surfaces on the plane', 2023.
[56] A. Sengupta, 'Gauge theory on compact surfaces', *Mem. Amer. Math. Soc.* **126**(600) (1997), viii+85.
[57] A. Sengupta, 'Yang-Mills on surfaces with boundary: quantum theory and symplectic limit', *Comm. Math. Phys.* **183**(3) (1997), 661–705.
[58] A. N. Sengupta, 'The volume measure for flat connections as limit of the yang–mills measure', *J. Geom. Phys.* **47**(4) (2003), 398–426.
[59] [59] A. N. Sengupta, 'Traces in two-dimensional QCD: the large-N limit', in *Traces in Number Theory, Geometry and Quantum Fields* (Aspects Math.) E38 (Friedr. Vieweg, Wiesbaden, 2008), 193–212.
[60] H. Shen, 'A stochastic PDE approach to large N problems in quantum field theory: a survey', *J. Math. Phys.* **63**(8) (2022), Paper No. 081103, 21.
[61] [61] H. Shen, S. A. Smith and R. Zhu, 'A new derivation of the finite N master loop equation for lattice Yang-Mills', *Electron. J. Probab.* **29** (2024), Paper No. 29, 18.
[62] I. M. Singer, 'On the master field in two dimensions', in *Functional Analysis on the Eve of the 21st Century, Vol. 1 (New Brunswick, NJ, 1993)* (Progr. Math.) vol. 131 (Birkhäuser Boston, Boston, MA, 1995), 263–281.
[63] R. Speicher and J. Wysoczański, 'Mixtures of classical and free independence', *Arch. Math. (Basel)* **107**(4) (2016), 445–453.
[64] J. Stillwell, *Classical Topology and Combinatorial Group Theory* (Graduate Texts in Mathematics) vol. 72 (Springer-Verlag, New York, second edition, 1993).
[65] G. Hooft, 'A planar diagram theory for strong interactions', *Nuclear Phys. B* **72**(3) (1974), 461–473.
[66] D. Voiculescu, 'Limit laws for random matrices and free products', *Invent. Math.* **104**(1) (1991), 201–220.
[67] D. Voiculescu, 'The analogues of entropy and of Fisher's information measure in free probability theory. VI. Liberation and mutual free information', *Adv. Math.* **146**(2) (1999), 101–166.
[68] D. Voiculescu, 'Lectures on free probability theory', in *Lectures on Probability Theory and Statistics (Saint-Flour, 1998)* (Lecture Notes in Math.) vol. 1738 (Springer, Berlin, 2000), 279–349.
[69] K. G. Wilson, 'Confinement of quarks', *Phys. Rev.* **10** (1974).
[70] E. Witten, 'The $1/N$ expansion in atomic and particle physics', *NATO Sci. Ser. B* **59** (1980), 403–419.
[71] E. Witten, 'On quantum gauge theories in two dimensions', *Comm. Math. Phys.* **141**(1) (1991), 153–209.
[72] F. Xu, 'A random matrix model from two-dimensional Yang–Mills theory', *Comm. Math. Phys.* **190**(2) (1997), 287–307.