J. Eur. Math. Soc. 27, 5093–5131 (2025) DOI 10.4171/JEMS/1469

Vincent Delecroix · Mingkun Liu

Length partition of random multicurves on large genus hyperbolic surfaces

Received 4 March 2022; revised 15 January 2024

Abstract. We study the length statistics of the components of a random multicurve on a closed surface of genus at least 2. This investigation was initiated by Mirzakhani in a paper published in 2016 where she studied the case of random pants decompositions. We prove that as the genus tends to infinity these statistics converge in law to the Poisson–Dirichlet distribution with parameter 1/2. In particular, the mean lengths of the three longest components converge to 75.8%, 17.1% and 4.9% of the total length, respectively.

Keywords: asymptotic geometry, surfaces, multicurves, Poisson–Dirichlet process.

1. Introduction

1.1. Lengths statistics of random multicurves in large genus

Let X be a closed Riemann surface of genus $g \geq 2$ endowed with its conformal hyperbolic metric of constant curvature -1. A *simple closed curve* is a homotopically non-trivial connected closed curve which has no self-intersection. In the free homotopy class of a simple closed curve γ on X, there exists a unique geodesic representative with respect to the hyperbolic metric X. We denote by $\ell_X(\gamma)$ the length of this geodesic representative.

A *multicurve* on X is a multiset of disjoint simple closed curves on X. Given a multicurve γ, a *component* of γ is a maximal family of freely homotopic curves in γ. The cardinality of a component is called its *multiplicity* and the *length of a component* is the sum of the lengths of the simple curves belonging to the component (or equivalently its multiplicity multiplied by the length of any simple closed curve in the component). A multicurve is called *primitive* if all its components have multiplicity one. We denote by $\boldsymbol{\ell}_X^{\downarrow}(\gamma)$ the vector of the lengths of each component sorted in descending order, by $\mathbf{mult}(\gamma)$ the

Vincent Delecroix: Laboratoire Bordelais de Recherche en Informatique, CNRS, Université de Bordeaux, 351 cours de la Libération, 33400 Talence, France; vincent.delecroix@u-bordeaux.fr

Mingkun Liu: Institut de Mathématiques de Jussieu – Paris Rive Gauche, Université Paris Cité, Bâtiment Sophie Germain, Boite Courrier 7012, 8 Place Aurélie Nemours, 75205 Paris; Sorbonne Université, Boite Courrier 247, 4 Place Jussieu, 75252 Paris, France; mingkun.liu@imj-prg.fr

Mathematics Subject Classification 2020: 30F60 (primary); 60C05, 60B12 (secondary).

multiset of the multiplicities of the components of γ, and by $\operatorname{mult}(\gamma)$ the maximum of $\mathbf{mult}(\gamma)$. Neither $\mathbf{mult}(\gamma)$ nor $\operatorname{mult}(\gamma)$ depends on the hyperbolic structure of X. We write $\ell_X(\gamma)$ for the sum of the entries of $\boldsymbol{\ell}_X^{\downarrow}(\gamma)$ and define its *normalized length vector* to be

$$\widehat{\boldsymbol{\ell}}_X^{\downarrow}(\gamma) := \frac{\boldsymbol{\ell}_X^{\downarrow}(\gamma)}{\ell_X(\gamma)}.$$

We denote by $\mathcal{ML}_X(\mathbb{Z})$ the set of free homotopy classes of multicurves on X. This notation is explained by the fact that multicurves are the integral points of the space of measured laminations usually denoted by $\mathcal{ML}_X$.

In order to make sense of convergence, we need all normalized vectors to belong to the same space. For an integer $k \geq 1$ and a real number $r > 0$, let us define

$$\Delta_{\leq r}^{k} := \{(x_1, x_2, \dots, x_k) \in [0, \infty)^k : x_1 + x_2 + \cdots + x_k \leq r\},$$
$$\Delta_{\leq r}^{\infty} := \{(x_1, x_2, \dots) \in [0, \infty)^{\mathbb{N}} : x_1 + x_2 + \cdots \leq r\},$$

where $\mathbb{N} := \{1, 2, \dots\}$. For $k_1 \leq k_2$, an injection $\Delta_{\leq r}^{k_1} \to \Delta_{\leq r}^{k_2}$ is naturally defined by completing vectors with zeros. The infinite simplex $\Delta_{\leq r}^{\infty}$ is the inductive limit of these injections, and we always identify $\Delta_{\leq r}^{k}$ as a subspace of $\Delta_{\leq r}^{\infty}$. In particular, each vector $\widehat{\boldsymbol{\ell}}_X^{\downarrow}(\gamma)$ can be seen as an element of $\Delta_{\leq 1}^{\infty}$ by completing its coordinates with infinitely many zeros.

As our aim is to study convergence of random infinite vectors, let us mention that $\Delta_{\leq 1}^{\infty}$ is a closed subset of $[0, 1]^{\mathbb{N}}$ endowed with the product topology. This topology coincides with the topology of the inductive limit. When we consider convergence in distribution on $\Delta_{\leq 1}^{\infty}$, we mean convergence in the space of Borel probability measures on $\Delta_{\leq 1}^{\infty}$ which is a compact set.

The following result is a consequence of the works [4, 19].

Theorem 1.1. *Let $g \geq 2$ and $m \in \mathbb{N} \cup \{\infty\}$. There exists a random variable $L^{(g,m)\downarrow} = (L^{(g,m)\downarrow}{}_1, \dots, L^{(g,m)\downarrow}{}_{3g-3})$ on $\Delta_{\leq 1}^{3g-3}$ with the following property. For any Riemann surface X of genus g, as $R \to \infty$, we have the following convergence in distribution:*

$$\frac{1}{s_X(R,m)} \sum_{\substack{\gamma \in \mathcal{ML}_X(\mathbb{Z}) \\ \ell_X(\gamma) \leq R \\ \operatorname{mult}(\gamma) \leq m}} \delta_{\widehat{\boldsymbol{\ell}}_X^{\downarrow}(\gamma)} \xrightarrow[R \to \infty]{} L^{(g,m)\downarrow},$$

where $\delta_{\widehat{\boldsymbol{\ell}}_X^{\downarrow}(\gamma)}$ is the Dirac mass at the vector $\widehat{\boldsymbol{\ell}}_X^{\downarrow}(\gamma)$ and $s_X(R,m) := \#\{\gamma \in \mathcal{ML}_X(\mathbb{Z}) : \ell_X(\gamma) \leq R \text{ and } \operatorname{mult}(\gamma) \leq m\}$ is the number of multicurves on X of length at most R and multiplicity at most m.

We actually prove a more precise version of the above statement, Theorem 3.2, in which the law of $L^{(g,m)\downarrow}$ is made explicit. Remark that the limit depends only on the genus of X and not on its hyperbolic metric.

The Poisson–Dirichlet distribution is a probability measure on $\Delta_{\leq 1}^{\infty}$. The simplest way to introduce it is via the *stick-breaking process*. Let $\theta > 0$ and let $U_1, U_2, \dots$ be independent and identically distributed random variables with law Beta$(1, \theta)$ (i.e., they are supported on $(0, 1]$ with density $\theta(1-x)^{\theta-1}$). Consider the random sequence

$$V := (U_1, (1-U_1)U_2, (1-U_1)(1-U_2)U_3, \dots).$$

Informally, the components of V are obtained by starting from a stick of length 1 identified with $[0, 1]$. At the first stage, U_1 determines where we break the first piece, and we are left with a stick of size $1-U_1$. We then repeat the process ad libitum. The law of V is the *Griffiths–Engen–McCloskey distribution with parameter* θ, which we shall denote by GEM(θ). The *Poisson–Dirichlet distribution with parameter* θ, denoted by PD(θ), is the distribution of $V^{\downarrow}$, the vector V whose entries are sorted in decreasing order. For more details, we refer the reader to Section 5.2. The distribution PD(1) is the limit distribution of the orbit length of uniform random permutations. The distribution PD(θ) appears when considering the Ewens distribution with parameter θ on the symmetric group. See Section 1.2.2 below for a more detailed discussion on permutations.

Our main result is the following.

Theorem 1.2. *For any $m \in \mathbb{N} \cup \{\infty\}$, the sequence $(L^{(g,m)\downarrow})_{g \geq 2}$ converges in distribution to* PD$(1/2)$ *as $g \to \infty$.*

The most interesting cases of this convergence are $m = 1$ (primitive multicurves) and $m = \infty$ (all multicurves). Let us insist that both $L^{(g,1)\downarrow}$ and $L^{(g,\infty)\downarrow}$ converge to the same limit as $g \to \infty$.

All marginals of the Poisson–Dirichlet law can be computed; see, for example, [5, Section 4.11]. In particular, if $V = (V_1, V_2, \dots) \sim$ PD(θ), then

$$\mathbb{E}((V_j)^n) = \frac{\Gamma(\theta+1)}{\Gamma(\theta+n)} \int_0^\infty \frac{(\theta E_1(x))^{j-1}}{(j-1)!} x^{n-1} e^{-x-\theta E_1(x)}\, \mathrm{d}x,$$

where $E_1(x) := \int_x^\infty e^{-y} y^{-1}\, \mathrm{d}y$. The formulas can be turned into a computer program, and values were tabulated in [12, 13]. For $\theta = 1/2$, we have

$$\mathbb{E}(V_1) \approx 0.758, \quad \mathbb{E}(V_2) \approx 0.171, \quad \text{and} \quad \mathbb{E}(V_3) \approx 0.049.$$

Therefore, Theorem 1.2 implies that, informally, the average lengths of the first three largest components of a random multigeodesic on a closed hyperbolic surface of high genus are approximately 75.8%, 17.1%, and 4.9% of the total length, respectively.

Moreover, we prove that, on a large genus surface, all macroscopic components of a random multicurve are primitive with high probability.

For a multicurve $\gamma = m_1\gamma_1 + \cdots + m_k\gamma_k$, define

$$p_X(\gamma) := \frac{1}{\ell_X(\gamma)} \sum_{\substack{1 \leq i \leq k \\ m_i = 1}} m_i \ell_X(\gamma_i).$$

In other words, $p_X(\gamma)$ stands for the proportion of the total length contributed by primitive components.

Theorem 1.3. *For any $g \geq 2$ and $m \in \mathbb{Z}_{\geq 1} \cup \{\infty\}$, there exists a random variable $P^{(g,m)}$ on $[0, 1]$ such that for any Riemann surface X of genus g, when $R \to \infty$, we have convergence in distribution*

$$\frac{1}{s_X(R,m)} \sum_{\substack{\gamma \in \mathcal{ML}_X(\mathbb{Z}) \\ \ell_X(\gamma) \leq R \\ \mathrm{mult}(\gamma) \leq m}} \delta_{p_X(\gamma)} \xrightarrow[R \to \infty]{} P^{(g,m)},$$

and for any $\varepsilon > 0$, we have

$$\lim_{g \to \infty} \mathbb{P}(P^{(g,m)} \geq 1 - \varepsilon) = 1.$$

In particular, a long component is more likely to be "visually" long than to have a huge multiplicity.

1.2. Further remarks

1.2.1. Square-tiled surfaces. In this section, we give an alternative statement of Theorem 1.2 in terms of square-tiled surfaces. The correspondence between statistics of multicurves and statistics of square-tiled surfaces is developed in [3, 7], and we refer the reader to these two references.

A *square-tiled surface* is a connected surface obtained from gluing finitely many unit squares $[0, 1] \times [0, 1]$ along their edges by translation $z \mapsto z + u$ or "half-translation" $z \mapsto -z + u$. Combinatorially, one can label the squares by $\{1, \ldots, N\}$ and then a square-tiled surface is encoded by two involutions (σ, τ) of $\{\pm 1, \pm 2, \ldots, \pm N\}$ without fixed points. More precisely, σ encodes the horizontal gluings: $+i$ and $-i$ are respectively the right and left sides of the i-th square. The orbits of σ with different signs are glued by translations and the ones with same signs are glued by half-translations. And τ encodes the vertical gluings: $+i$ and $-i$ are the top and bottom sides of the i-th squares, respectively. The labelling is irrelevant in our definition and two pairs (σ_1, τ_1) and (σ_2, τ_2) encode the same square-tiled surface if there exists a permutation α of $\{\pm 1, \pm 2, \ldots, \pm N\}$ so that $\alpha(-i) = -\alpha(+i)$, $\sigma_2 = \alpha \circ \sigma_1 \circ \alpha^{-1}$ and $\tau_2 = \alpha \circ \tau_1 \circ \alpha^{-1}$.

A square-tiled surface comes with a conformal structure and a quadratic form coming from the conformal structure of the unit square and the quadratic form $\mathrm{d}z^2$ (both are being preserved by translations and half-translations). This quadratic form might have simple poles, and we denote by $\mathcal{Q}_g(\mathbb{Z})$ the set of holomorphic square-tiled surfaces of genus g.

A square-tiled surface comes equipped with a filling pair of multicurves (γ_h, γ_v) coming from the gluings of the horizontal segments $[0, 1] \times \{1/2\}$ and vertical segments $\{1/2\} \times [0, 1]$ of each square, respectively. Conversely, the dual graph of a filling pair of multicurves in a surface of genus g defines a square-tiled surface in $\mathcal{Q}_g(\mathbb{Z})$. Our notation comes from the fact that holomorphic square-tiled surfaces can be seen as integral points in the moduli space of quadratic differentials $\mathcal{Q}_g$. A component of the multicurve γ_h corresponds geometrically to a horizontal cylinder. For a square-tiled surface M, we denote by $A^{\downarrow}(M)$ the normalized vector of areas of these horizontal cylinders sorted in decreasing order and by height(M) the maximum of their heights. Here as in the introduction, normalized means that we divide by the sum of entries of a vector which coincides with area(M). The following is a particular case of [7, Theorem 1.29] using the explicit formulas for $L^{(g,m)}$ given in Theorem 3.2.

Theorem 1.4 ([7]). *Let $g \geq 2$ and $m \in \mathbb{N} \cup \{\infty\}$. Let $L^{(g,m)}$ be the random variable from Theorem 1.1. Then, as $N \to \infty$, we have the following convergence in distribution:*

$$\frac{1}{\#\{M \in \mathcal{Q}_g(\mathbb{Z}) : \mathrm{height}(M) \leq m,\ \mathrm{area}(M) \leq N\}} \sum_{\substack{M \in \mathcal{Q}_g(\mathbb{Z}) \\ \mathrm{height}(M) \leq m \\ \mathrm{area}(M) \leq N}} \delta_{A^{\downarrow}(M)} \to L^{(g,m)\downarrow}.$$

An important difference to notice between Theorems 1.1 and 1.4 is that in the former (hyperbolic) metric, X is fixed and we sum over the multicurves γ while in the latter we sum over the discrete set of holomorphic square-tiled surfaces M.

Using Theorem 1.4, our Theorem 1.2 admits the following reformulation.

Corollary 1.5. *The vector of normalized areas of horizontal cylinders of a random square-tiled surface of genus g converges in distribution to* PD$(1/2)$ *as $g \to \infty$.*

1.2.2. Permutations and multicurves. Given a permutation σ in S_n, we denote by $K_n(\sigma)$ the number of orbits it has on $\{1, 2, \dots, n\}$ or equivalently the number of cycles in its disjoint cycle decomposition. The *Ewens measure with parameter θ* on S_n is the probability measure defined by

$$\mathbb{P}_{n,\theta}(\sigma) := \frac{\theta^{K_n(\sigma)}}{Z_{n,\theta}}, \quad \text{where } Z_{n,\theta} := \sum_{\sigma \in S_n} \theta^{K_n(\sigma)}.$$

Then under $\mathbb{P}_{n,\theta}$, as $n \to \infty$, we have that

- the random variable K_n behaves as a Poisson distribution Poi$(\theta \log(n))$ of parameter $\theta \log(n)$ (e.g., by means of a local limit theorem),
- the normalized sorted vector of cycle lengths of σ tends to PD(θ),
- the number of cycles of length k of σ converges to Poi(θ/k).

See, for example, [5].

By analogy, let us denote by $K^{(g,m)}$ the number of non-zero components of $L^{(g,m)}$. In [8], it is proven that $K^{(g,m)}$ behaves as a Poisson distribution with parameter $\log(g)/2$ (by means of a local limit theorem) independently of m. In other words, it behaves as the number of cycles $K_g(\sigma)$ for a random permutation σ under $\mathbb{P}_{g,1/2}$.

Theorem 1.2 provides another connection between $L^{(g,m)}$ and $\mathbb{P}_{g,1/2}$. Namely, $L^{(g,m)\downarrow}$ is asymptotically close to the normalized sorted vector of the cycle lengths of σ under $\mathbb{P}_{g,1/2}$.

Finally, let us mention that components of $L^{(g,m)}$ of the order of o(1) are invisible in the convergence towards PD(1/2). It is a consequence of Theorem 1.2 that the macroscopic components of the order of a constant carry the total mass. Building on the intuition that in the large genus asymptotic regime random multicurves on a surface X of genus g behave like the cycles of a random permutation in the symmetric group S_g, one should expect to have a Poisson limit for components of order g^{-1} and that there is no component of order $g^{-1-\varepsilon}$. In a work in progress, we provide an affirmative answer to this intuition. However, because lengths are continuous parameters, the limit is a continuous Poisson process and not a discrete one supported on $\mathbb{N}$ as in the permutation case.

1.3. Proof overview and structure of the paper

The first step of the proof consists in writing an explicit expression for the random variable $L^{(g,m)\downarrow}$ that appears in Theorem 1.1. See Theorem 3.2 in Section 3. The formula follows from the work of Mirzakhani on pants decompositions [20] and the result of Arana-Herrera [4] and Liu [19] on length distribution for each fixed topological type of multicurves. The expression of $L^{(g,m)\downarrow}$ can be seen as a refinement of the formula for the Masur–Veech volume of the moduli space of quadratic differentials from [7].

The formula for $L^{(g,m)\downarrow}$ involves a super-exponential number of terms in g (one term for each topological type of multicurve on a surface of genus g). However, in the large genus limit only O(log(g)) terms contribute. This allows us to consider a simpler random variable $\widetilde{L}^{(g,m,\kappa)\downarrow}$ which, asymptotically, coincides with $L^{(g,m)\downarrow}$. See Theorem 4.1 in Section 4. This reduction is very similar to the one used for the large genus asymptotics of Masur–Veech volumes in [1, 8].

The core of our proof consists in proving the convergence of moments of the simpler variable $\widetilde{L}^{(g,m,\kappa)\downarrow}$. We do not use directly $\widetilde{L}^{(g,m,\kappa)\downarrow}$ but its size-biased version $\widetilde{L}^{(g,m,\kappa)*}$. The definition of size bias and the link with the Poisson–Dirichlet distribution are explained in Section 5. In Section 6, we show that the moments of $\widetilde{L}^{(g,m,\kappa)*}$ converge to the moments of GEM(1/2) which is the size-biased version of the Poisson–Dirichlet process PD(1/2); see Theorem 6.1.

2. Background material

In this section, we introduce notations and state results from the literature that are used in our proof.

2.1. Multicurves and stable graphs

Recall from the introduction that a multicurve on a closed hyperbolic surface X of genus g is a finite multiset of free homotopy classes of disjoint simple closed curves. We denote by $\mathcal{ML}_X(\mathbb{Z})$ the set of multicurves on X. The homotopy classes that appear in a multicurve γ are called components. A multicurve on a closed surface of genus g has at most

$3g - 3$ components. The multiplicity of a component is the number of times it is repeated in γ, and γ is primitive if multiplicities of all components are 1.

Let us also recall our notations:

- $\ell_X(\gamma)$ is the total length of γ,
- $\boldsymbol{\ell}_X^{\downarrow}(\gamma)$ is the vector of the lengths of the components of γ sorted in descending order,
- $\widehat{\ell}_X^{\downarrow}(\gamma)$ is the (unit) vector of the normalized lengths of the components of γ sorted in descending order,
- $\mathrm{mult}(\gamma)$ is the maximum multiplicity of the components in γ,
- $\mathbf{mult}(\gamma)$ is the multiset of multiplicities of the components in γ.

The mapping class group $\mathrm{Mod}(X)$ of X acts on multicurves. The *topological type* of a multicurve γ is its equivalence class $[\gamma]$ under the $\mathrm{Mod}(X)$-action. For each fixed genus g, there are finitely many topological types of primitive multicurves and countably many topological types of multicurves. They are conveniently encoded by respectively stable graphs and weighted stable graphs that we define next. Informally, given a multicurve γ with components $\gamma_1, \dots, \gamma_k$ and multiplicities $m_1, \dots, m_k$, we build a dual graph Γ as follows:

- we add a vertex for each connected component of the complement $X \smallsetminus (\gamma_1 \cup \cdots \cup \gamma_k)$; the vertex v carries an integral weight the genus $g(v)$ of the corresponding component,
- we add an edge for each component γ_i of the multicurve between the two vertices corresponding to the connected components bounded by γ_i; this edge carries a weight m_i.

More formally, a *stable graph* Γ is a 5-tuple (V, H, ι, σ, g), where

- V is a finite set called *vertices*,
- H is a finite set called *half-edges*,
- $\iota\colon H \to H$ is an involution without fixed points on H; each pair $\{h, \iota(h)\}$ is called an *edge*, and we denote by $E(\Gamma)$ the set of edges,
- $\sigma\colon H \to V$ is a surjective map ($\sigma(h)$ is the vertex at which h is rooted),
- g is a map from V to $\mathbb{Z}_{\geq 0}$,

such that

- (*connectedness*) for each pair of vertices $u, v \in V$, there exists a sequence of edges, $\{x_1, y_1\}, \{x_2, y_2\}, \dots, \{x_n, y_n\}$ such that $\sigma(x_1) = u$, $\sigma(y_n) = v$ and $\sigma(y_i) = \sigma(x_{i+1})$, for $i \in \{1, \dots, n-1\}$,
- (*stability*) for each vertex $v \in V$, we have

$$2g(v) - 2 + \deg(v) > 0,$$

where $\deg(v) := |\sigma^{-1}(v)|$ is the *degree* of the vertex v.

Given a stable graph Γ, its *genus* $g(\Gamma)$ is defined by

$$g(\Gamma) := |E| - |V| + 1 + \sum_{v \in V} g(v).$$

An *isomorphism* between stable graphs

$$\Gamma_1 = (V_1, H_1, \iota_1, \sigma_1, g_1) \quad \text{and} \quad \Gamma_2 = (V_2, H_2, \iota_2, \sigma_2, g_2)$$

is a pair of bijections $\varphi\colon V_1 \to V_2$ and $\psi\colon H_1 \to H_2$ such that

- $\psi \circ \iota_1 = \iota_2 \circ \psi$ (in other words, ψ maps an edge to an edge),
- $\varphi \circ \sigma_1 = \sigma_2 \circ \psi$,
- for each $v \in V_1$, we have $g_2(\varphi(v)) = g_1(v)$.

Note that ψ determines φ but it is convenient to record the automorphism as a pair (φ, ψ). We denote by $\mathrm{Aut}(\Gamma)$ the set of automorphisms of Γ and by $\mathcal{G}_g$ the finite set of isomorphism classes of stable graphs of genus g.

A *weighted stable graph* is a pair $(\Gamma, \boldsymbol{m})$, where Γ is a stable graph and $\boldsymbol{m} \in \mathbb{N}^{E(\Gamma)}$ is a map which assigns a positive integral weight $\boldsymbol{m}(e)$, sometimes written as $\boldsymbol{m}_e$, to each edge e of Γ. An *isomorphism* between two weighted stable graphs $(\Gamma_1, \boldsymbol{m}_1)$ and $(\Gamma_2, \boldsymbol{m}_2)$ is an isomorphism (φ, ψ) between Γ_1 and Γ_2 such that for each edge e of Γ_1, we have $\boldsymbol{m}_1(e) = \boldsymbol{m}_2(\psi(e))$ (where we use $\psi(e)$ to denote $\{\psi(x), \psi(y)\}$ for the edge $e = \{x, y\} \subset H_1$). We denote by $\mathrm{Aut}(\Gamma, \boldsymbol{m})$ the set of automorphisms of the weighted graph $(\Gamma, \boldsymbol{m})$. There is a one-to-one correspondence between topological types of multicurves and weighted stable graphs. Primitive multicurves correspond to the case where all edges carry weight 1.

2.2. ψ-classes and Kontsevich polynomial

The formula for the random variable $L^{(g,m)\downarrow}$ that appears in Theorem 1.1 involves intersection numbers of ψ-classes that we introduce now. These rational numbers are famously related to the Witten conjecture [22] proven by Kontsevich [18].

We denote by $\overline{\mathcal{M}}_{g,n}$ the Deligne–Mumford compactification of the moduli space of smooth complex curves of genus g with n marked points. There exist n so-called *tautological line bundles* $\mathcal{L}_1, \ldots, \mathcal{L}_n \to \overline{\mathcal{M}}_{g,n}$ over $\overline{\mathcal{M}}_{g,n}$ such that the fiber of $\mathcal{L}_i$ at $(C; x_1, \ldots, x_n) \in \overline{\mathcal{M}}_{g,n}$ is the cotangent space of C at the i-th marked point x_i. The i-th *psi-class* ψ_i is defined as the first Chern class of the i-th tautological line bundle $c_1(\mathcal{L}_i) \in H^2(\overline{\mathcal{M}}_{g,n}, \mathbb{Q})$. We use the following standard notation

$$\langle \tau_{d_1} \cdots \tau_{d_n} \rangle_{g,n} := \int_{\overline{\mathcal{M}}_{g,n}} \psi_1^{d_1} \cdots \psi_n^{d_n},$$

where

$$d_1 + \cdots + d_n = \dim_{\mathbb{C}} \overline{\mathcal{M}}_{g,n} = 3g - 3 + n.$$

All these intersection numbers are positive rational numbers and can be computed by recursive equations from

$$\langle \tau_0^3 \rangle_{0,3} = 1 \quad \text{and} \quad \langle \tau_1 \rangle_{1,1} = \frac{1}{24};$$

see, for example, [16].

For our purposes, it will be convenient to consider the *Kontsevich polynomial* $V_{g,n} \in \mathbb{Q}[x_1, \dots, x_n]$ that gathers the intersection number into a symmetric polynomial on n variables. More precisely,

$$V_{g,n}(x_1, \dots, x_n) := \frac{1}{2^{3g-3+n}} \sum_{\substack{(d_1,\dots,d_n)\in\mathbb{Z}_{\geq 0}^n \\ d_1+\dots+d_n=3g-3+n}} \frac{\langle \tau_{d_1} \cdots \tau_{d_n} \rangle_{g,n}}{d_1! \cdots d_n!} \cdot x_1^{2d_1} \cdots x_n^{2d_n}$$
$$= \int_{\bar{\mathcal{M}}_{g,n}} \exp\left(\sum_{i=1}^{n} \frac{\psi_i}{2} x_i^2 \right).$$

For later use, we gather the list of small Kontsevich polynomials below:

$$V_{0,3}(x_1, x_2, x_3) = 1,$$
$$V_{0,4}(x_1, x_2, x_3, x_4) = \frac{1}{2}(x_1^2 + x_2^2 + x_3^2 + x_4^2),$$
$$V_{1,1}(x_1) = \frac{1}{48} x_1^2,$$
$$V_{1,2}(x_1, x_2) = \frac{1}{192}(x_1^2 + x_2^2)^2.$$

2.3. *Random multicurves*

Mirzakhani proved the polynomial growth of the number of multicurves on hyperbolic surfaces with respect to its length. This result and some extensions of it are nicely presented in the book of Erlandsson and Souto [10].

Let X be a hyperbolic surface of genus g. We define

$$s_X(R) := \#\{\eta \in \mathcal{ML}_X(\mathbb{Z}) : \ell_X(\eta) \leq R\},$$
$$s_X(R, \gamma) := \#\{\eta \in \mathrm{Mod}(X) \cdot \gamma : \ell_X(\eta) \leq R\}.$$

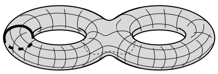
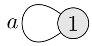
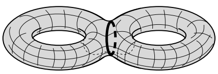
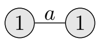
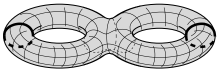
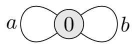
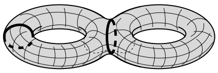
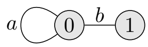
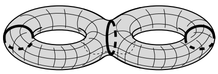
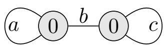
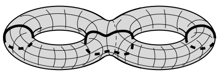
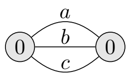
Theorem 2.1 ([20, Theorems 1.1, 1.2, 3.1 and 5.3]). *Let X be a hyperbolic surface. For any multicurve $\gamma \in \mathcal{ML}_X(\mathbb{Z})$, there exists a positive rational constant $c(\gamma)$ such that we have, as $R \to \infty$,*

$$s_X(R) \sim B(X) \cdot R^{6g-6}, \quad s_X(R, \gamma) \sim B(X) \cdot \frac{c(\gamma)}{b_g} \cdot R^{6g-6},$$

where $B(X)$ is the Thurston volume of the unit ball in the space of measured laminations $\mathcal{ML}_X$ with respect to the length function ℓ_X, and

$$b_g := \int_{\mathcal{M}_g} B(X)\, \mathrm{d}X = \sum_{[\gamma]\in\mathcal{ML}_X(\mathbb{Z})/\mathrm{Mod}(X)} c(\gamma).$$

The above theorem allows us to give sense to the notion of a random multicurve. Namely, we endow the set of topological types of multicurves $\mathcal{ML}_X(\mathbb{Z})/\mathrm{Mod}(X)$ with

Multicurve	Stable graph	Polynomial F_Γ
		$x_a V_{1,2}(x_a, x_a) = \frac{x_a^5}{48}$
		$x_a V_{1,1}(x_a) V_{1,1}(x_a) = \frac{x_a^5}{2304}$
		$x_a x_b V_{0,4}(x_a, x_a, x_b, x_b) = x_a^3 x_b + x_a x_b^3$
		$x_a x_b V_{0,3}(x_a, x_a, x_b) V_{1,1}(x_b) = \frac{x_a x_b^3}{48}$
		$x_a x_b x_c V_{0,3}(x_a, x_a, x_b) V_{0,3}(x_b, x_c, x_c) = x_a x_b x_c$
		$x_a x_b x_c V_{0,3}(x_a, x_b, x_c) V_{0,3}(x_a, x_b, x_c) = x_a x_b x_c$

Tab. 1. The list of topological types of primitive multicurves of genus 2, their associated stable graphs, and their corresponding polynomial F_Γ. The labels on edges are used as variable indices in F_Γ.

the probability measure which assigns $c(\gamma)/b_g$ to $[\gamma]$. We now provide the explicit expression for this probability. For a stable graph $\Gamma \in \mathcal{G}_g$, we define the polynomial F_Γ on the variables $\boldsymbol{x}_E := \{x_e : e \in E(\Gamma)\}$ by

$$F_\Gamma(\boldsymbol{x}_E) := \prod_{e \in E(\Gamma)} x_e \cdot \prod_{v \in V(\Gamma)} V_{g_v, n_v}(\boldsymbol{x}_v). \tag{2.1}$$

Here $\boldsymbol{x}_v$ is the multiset of variables $x_{e(h)}$, where $h \in H(\Gamma)$ runs over the set of half-edges incident to v and $e(h) \in E(\Gamma)$ stands for the edge that contains h (so x_e appears twice if e is a self-loop), $g_v := g(v)$ (see the definition of a stable graph), n_v is the cardinality of $\boldsymbol{x}_v$, and V_{g_v,n_v} is the Kontsevich polynomial defined in Section 2.2. In Table 1, we list the stable graphs Γ of genus 2 together with their associated polynomial F_Γ.

Remark 2.2. The polynomial F_Γ appeared first in Mirzakhani's work [20]; see in particular Theorem 5.3. It was related to square-tiled surfaces and Masur–Veech volumes in [7] though with a different normalization; see equation (1.12) and Remark 2.5. Namely, the polynomial P_Γ from [7] is related to F_Γ by

$$P_\Gamma = 2^{4g-2} \frac{(4g-4)!}{(6g-7)!} \frac{1}{|\mathrm{Aut}(\Gamma)|} F_\Gamma.$$

The normalization of F_Γ is identical to the conventions used in [2] and simplifies the computations of the present article.

Following [7], for a weighted stable graph $(\Gamma, \boldsymbol{m})$, we denote by

$$\mathcal{Y}_{\boldsymbol{m}} : \mathbb{Q}[\{x_e : e \in E(\Gamma)\}] \to \mathbb{Q}$$

the linear operator defined on monomials

$$\mathcal{Y}_{\boldsymbol{m}}\Big(\prod_{e \in E(\Gamma)} x_e^{n_e}\Big) := \prod_{e \in E(\Gamma)} \frac{n_e!}{m(e)^{n_e+1}}, \tag{2.2}$$

and for $m \in \mathbb{N} \cup \{\infty\}$, set

$$\mathcal{Z}_m := \sum_{\substack{\boldsymbol{m} \in \mathbb{N}^{E(\Gamma)} \\ \boldsymbol{m}_e \leq m \\ e \in E(\Gamma)}} \mathcal{Y}_{\boldsymbol{m}}.$$

We derive the following directly from [7].

Theorem 2.3. *Let γ be a multicurve of genus g and $(\Gamma, \boldsymbol{m})$ the dual weighted stable graph. Then*

$$c(\gamma) = \frac{1}{(6g-6)!} \frac{\mathcal{Y}_{\boldsymbol{m}}(F_\Gamma)}{|\mathrm{Aut}(\Gamma, \boldsymbol{m})|}. \tag{2.3}$$

Furthermore, for any $m \in \mathbb{N} \cup \{\infty\}$,

$$b_{g,m} := \sum_{\substack{[\gamma] \in \mathcal{ML}_X(\mathbb{Z})/\mathrm{Mod}(X) \\ \mathrm{mult}(\gamma) \leq m}} c(\gamma) = \frac{1}{(6g-6)!} \sum_{\Gamma \in \mathcal{G}_g} \frac{\mathcal{Z}_m(F_\Gamma)}{|\mathrm{Aut}(\Gamma)|}. \tag{2.4}$$

Note that $b_{g,\infty} = b_g$ defined in Theorem 2.1. The list of values $\mathcal{Y}_{\boldsymbol{m}}(F_\Gamma)$ for the stable graphs of genus 2 is shown in Table 2.

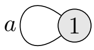
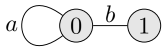
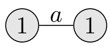
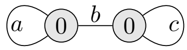
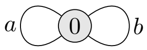
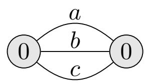
Remark 2.4. We warn the reader that the constant denoted by $b_{g,m}$ in this article has nothing to do with the analogue of b_g in the context of surfaces of genus g with n boundaries which is denoted by $b_{g,n}$ in [7, 21].

Remark 2.5. In Theorem 2.3, we fix a misconception in [7] about automorphisms of multicurves (or equivalently weighted stable graphs). Indeed, the way we defined automorphisms of stable graphs and weighted stable graphs in Section 2.1 makes it so that the following formula is valid

$$\sum_{\Gamma} \frac{\mathcal{Z}_\infty(F_\Gamma)}{|\mathrm{Aut}(\Gamma)|} = \sum_{(\Gamma, \boldsymbol{m})} \frac{\mathcal{Y}_{\boldsymbol{m}}(F_\Gamma)}{|\mathrm{Aut}(\Gamma, \boldsymbol{m})|},$$

where the sums are taken over isomorphism classes of stable graphs of genus g and weighted stable graphs of genus g, respectively.

Proof of Theorem 2.3. Taking into account the correction of Remark 2.5, this is exactly [7, Theorem 1.22] (see Remark 2.2 for the difference between P_Γ and F_Γ). ∎

Stable graph Γ	Value of $\mathcal{Y}_{\boldsymbol{m}}(\Gamma)$	Stable graph Γ	Value of $\mathcal{Y}_{\boldsymbol{m}}(\Gamma)$
	$\frac{5}{2}\cdot\frac{1}{m_a^6}$		$\frac{1}{2}\cdot\frac{1}{m_a^2 m_b^4}$
	$\frac{5}{96}\cdot\frac{1}{m_a^6}$		$\frac{1}{m_a^2 m_b^2 m_c^2}$
	$24\Big(\frac{1}{m_a^4 m_b^2}+\frac{1}{m_a^2 m_b^4}\Big)$		$\frac{1}{m_a^2 m_b^2 m_c^2}$

Tab. 2. The list of topological types of primitive multicurves of genus 2 and the associated values $\mathcal{Y}_{\boldsymbol{m}}(F_\Gamma)$ that is proportional to $c(\gamma)$ (see Theorem 2.3).

2.4. Asymptotics of ψ-correlators and b_g

Our proof of Theorem 1.2 uses crucially the asymptotics of ψ-intersections and Masur–Veech volumes from [1] that were further developed in [8].

Theorem 2.6 ([1, Theorem 1.5]). *For $g, n \in \mathbb{N}$ and $\boldsymbol{d} = (d_1, \ldots, d_n) \in \mathbb{Z}^n_{\geq 0}$ with $d_1 + \cdots + d_n = 3g - 3 + n$, let $\varepsilon(\boldsymbol{d})$ be defined by the equation*

$$\langle \tau_{d_1} \cdots \tau_{d_n} \rangle_{g,n} = \frac{(6g-5+2n)!!}{(2d_1+1)!! \cdots (2d_n+1)!!} \frac{1}{g! \cdot 24^g} \cdot (1 + \varepsilon(\boldsymbol{d})).$$

For $\delta > 0$, write

$$D(g,\delta) := \{\boldsymbol{d} = (d_1, \ldots, d_n) \in \mathbb{Z}^n_{\geq 0} : d_1 + \cdots + d_n = 3g + n - 3 \text{ and } n < \delta\sqrt{g}\}.$$

Then

$$\lim_{\delta \to 0} \lim_{g \to \infty} \max_{\boldsymbol{d} \in D(g,\delta)} \varepsilon(\boldsymbol{d}) = 0.$$

For $m \in \mathbb{N} \cup \{\infty\}$ and $\kappa \in \mathbb{R}_{>0}$, we define

$$\widetilde{b}_{g,m,\kappa} := \frac{1}{(6g-6)!} \sum_{\substack{\Gamma \in \mathcal{G}_g \\ |V(\Gamma)|=1 \\ |E(\Gamma)| \leq \kappa(\log(6g-6))/2}} \frac{\mathcal{Z}_m(F_\Gamma)}{|\mathrm{Aut}(\Gamma)|}. \tag{2.5}$$

As we have fewer terms in this definition, $\widetilde{b}_{g,m,\kappa} \leq b_{g,m}$.

We will use the asymptotic results of [1, 8] in the following form.

Theorem 2.7. *Let $m \in \mathbb{N} \cup \{\infty\}$ and $\kappa > 1$. Then, as $g \to \infty$, we have*

$$b_{g,m} \sim \widetilde{b}_{g,m,\kappa} \sim \frac{1}{\pi} \frac{1}{(6g-6)\cdot(4g-4)!} \sqrt{\frac{m}{m+1}} \left(\frac{4}{3}\right)^{4g-4}. \tag{2.6}$$

Proof. By [8, Theorem 3.1] (see also [1, Theorem 1.7] for the $m = \infty$ case), we have, as $g \to \infty$,

$$\sum_{\substack{\Gamma \in \mathcal{G}_g \\ |V(\Gamma)|=1}} \frac{\mathcal{Z}_m(F_\Gamma)}{|\mathrm{Aut}(\Gamma)|} \sim \frac{1}{\pi} \frac{(6g-7)!}{(4g-4)!} \sqrt{\frac{m}{m+1}} \left(\frac{4}{3}\right)^{4g-4}.$$

Next, by [8, Proposition 4.2] and [8, Proposition 4.3] ([1, Proposition 8.4] and [1, Proposition 8.5] for the case $m = \infty$), the contribution of the stable graphs with at least two vertices in (2.4) is negligible. More precisely,

$$\sum_{\substack{\Gamma \in \mathcal{G}_g \\ |V(\Gamma)|\geq 2}} \frac{\mathcal{Z}_m(F_\Gamma)}{|\mathrm{Aut}(\Gamma)|} \leq C \frac{(\log g)^{24}}{\sqrt{g}} \sum_{\substack{\Gamma \in \mathcal{G}_g \\ |V(\Gamma)|=1}} \frac{\mathcal{Z}_m(F_\Gamma)}{|\mathrm{Aut}(\Gamma)|} \tag{2.7}$$

for some constant $C > 0$. This proves the asymptotic formula for $b_{g,m}$.

Now we turn to $\widetilde{b}_{g,m,\kappa}$. We denote by $\Gamma_{g,k}$ the stable graph of genus g with a vertex of genus $g-k$ and k self-loops. Next, as in [8, (3.52)], for $1 \leq k \leq g$ let $p^{(1)}_{g,m}(k)$ denote the relative contribution of the one vertex graph with k self-loops

$$p^{(1)}_{g,m}(k) := \frac{\mathcal{Z}_m(F_{\Gamma_{g,k}})}{\sum_{k=1}^{g} \mathcal{Z}_m(F_{\Gamma_{g,k}})}.$$

By [8, Corollary 3.28], the sequence of probability distributions $(p^{(1)}_{g,m})_{g\geq 2}$ converges mod-Poisson with parameter

$$\lambda_g = \frac{1}{2} \log\left(\frac{2m}{m+1} \cdot (3g-3)\right),$$

radius $R = 2$, and limiting function $\Gamma(1/2)/\Gamma(x/2)$. As a consequence, by [8, Theorem 3.13], based on Hwang's works [14, 15], we obtain that for any $\kappa > 1$ and g large enough, we have

$$\sum_{k > (\kappa/2)\log((3g-3)(2m)/(m+1))} p^{(1)}_{g,m}(k) < \exp(-\varepsilon(\kappa) g) \tag{2.8}$$

for some constant $\varepsilon(\kappa) > 0$.

Now, it follows from (2.7) and (2.8) that $b_{g,m}$ and $\widetilde{b}_{g,m,\kappa}$ are asymptotically equivalent (using the fact that $2m/(m+1) \leq 2$ for all $m \in \mathbb{N} \cup \{\infty\}$). ∎

3. Length vectors of random multicurves

The aim of this section is to state and prove a refinement of Theorem 1.1 that provides an explicit description of the random variable $L^{(g,m)\downarrow}$. The latter involves two layers of randomness. The first layer is the choice of the topological type of the multicurve which is a discrete random variable with values in the set of weighted stable graphs $(\Gamma, \boldsymbol{m})$. Next, for each fixed $(\Gamma, \boldsymbol{m})$ one has an explicit length distribution of the components of a multicurve of type $(\Gamma, \boldsymbol{m})$.

The description of $L^{(g,m)\downarrow}$ involves a slight modification of the polynomials F_Γ introduced in (2.1) from Section 2.3. Let $(\Gamma, \boldsymbol{m})$ be a weighted stable graph. We define a polynomial in $\mathbb{Q}[\{x_e : e \in E(\Gamma)\}]$ by

$$f_{\Gamma,\boldsymbol{m}}((x_e)_{e\in E(\Gamma)}) := \frac{(6g-7)!}{\mathcal{Y}_{\boldsymbol{m}}(F_\Gamma)} \cdot \prod_{e\in E(\Gamma)} \frac{1}{m_e} \cdot F_\Gamma\Big(\Big(\frac{x_e}{m_e}\Big)_{e\in E(\Gamma)}\Big), \tag{3.1}$$

where $\mathcal{Y}_{\boldsymbol{m}}$ is defined in (2.2).

We endow $\Delta^k_{\le r}$ with the restriction of the Lebesgue measure on $\mathbb{R}^k$ that we denote by $\lambda^k_{\le r}$. We define the slice $\Delta^k_{=r}$ inside $\Delta^k_{\le r}$ as

$$\Delta^k_{=r} := \{(x_1, \dots, x_k) \in [0,\infty)^k : x_1 + \cdots + x_k = r\}.$$

On $\Delta^k_{=r}$ which is contained in a hyperplane in $\mathbb{R}^k$, we consider the Lebesgue measure induced by any choice of $k-1$ coordinates among $x_1, \dots, x_k$. The latter measure is well defined (since the change of variables between different choices has determinant ± 1), and we shall denote it by $\lambda^k_{=r}$. We denote similarly by $\lambda^{E(\Gamma)}_{\le r}$ and $\lambda^{E(\Gamma)}_{=r}$ the Lebesgue measures on the associated simplices in $\mathbb{R}^{E(\Gamma)}$. We have the following elementary result whose proof is postponed to Section 5.2.

Lemma 3.1. *The polynomial $f_{\Gamma,\boldsymbol{m}}$ defined by formula (3.1) is the density of a probability measure on $\Delta^{E(\Gamma)}_{=1}$.*

For each $g \ge 2$, we consider a family of independent random variables $\{U^{(\Gamma,\boldsymbol{m})}\}_{(\Gamma,\boldsymbol{m})}$ indexed by (isomorphism classes of) weighted stable graphs of genus g, where $U^{(\Gamma,\boldsymbol{m})}$ has density given by (3.1). We denote by $U^{(\Gamma,\boldsymbol{m})\downarrow}$ the random variable on $\Delta^{|E(\Gamma)|}_{=1}$ obtained from $U^{(\Gamma,\boldsymbol{m})}$ by sorting its entries in descending order. We slightly abuse notation and still denote by $U^{(\Gamma,\boldsymbol{m})\downarrow}$ the image in any $\Delta^k_{=1}$ for $k \ge |E(\Gamma)|$ by completing the last $k - |E(\Gamma)|$ variables with zeros.

In order to define $L^{(g,m)\downarrow}$ on $\Delta^{3g-3}_{=1}$, we need an auxiliary discrete variable $G_{g,m}$ independent of the family $\{U^{(\Gamma,\boldsymbol{m})}\}$ with values in the weighted stable graphs with multiplicities at most m defined as

$$\mathbb{P}(G_{g,m} = (\Gamma, \boldsymbol{m})) := \frac{1}{(6g-6)!\, b_{g,m}} \frac{\mathcal{Y}_{\boldsymbol{m}}(F_\Gamma)}{|\mathrm{Aut}(\Gamma, \boldsymbol{m})|}, \tag{3.2}$$

where $b_{g,m}$ is given by (2.4) and $\mathcal{Y}_{\boldsymbol{m}}$ by (2.2). By Theorem 2.3, this is a well-defined probability measure. Moreover, by Theorems 2.1 and 2.3, the distribution of $G_{g,m}$ is the asymptotic distribution of topological types of random multicurves with multiplicities bounded by m on any hyperbolic surface of genus g.

We define

$$L^{(g,m)\downarrow} := U^{G_{g,m}\downarrow}. \tag{3.3}$$

Note that $L^{(g,m)\downarrow}$ does not admit a density on $\Delta_{=1}^{3g-3}$: it is a convex combination of probability measures supported on subsimplices $\Delta_{=1}^{k} \subset \Delta_{=1}^{3g-3}$. Indeed, for each weighted stable graph $(\Gamma, \boldsymbol{m})$ the random variable $U^{(\Gamma,\boldsymbol{m})\downarrow}$ supported on $\Delta_{=1}^{|E(\Gamma)|}$ has a density.

The refinement of Theorem 1.1 is the following.

Theorem 3.2. *For any $g \geq 2$ and any $m \in \mathbb{N} \cup \{\infty\}$, we have convergence in distribution*

$$\frac{1}{s_X(R,m)} \sum_{\substack{\gamma \in \mathcal{ML}_X(\mathbb{Z}) \\ \ell_X(\gamma) \leq R \\ \mathrm{mult}\,\gamma \leq m}} \delta_{\hat{\boldsymbol{\ell}}_X^{\downarrow}(\gamma)} \xrightarrow[R\to\infty]{} L^{(g,m)\downarrow}, \tag{3.4}$$

where $L^{(g,m)\downarrow}$ is defined in (3.3) and

$$s_X(R,m) := \#\{\gamma \in \mathcal{ML}_X(\mathbb{Z}) : \ell_X(\gamma) \leq R,\ \mathrm{mult}(\gamma) \leq m\}$$

is the number of multicurves on X of length at most R and multiplicity at most m.

The study of the length vector of (ordered) multicurves of a given topological type was initiated by Mirzakhani in [21]. She studied the special case of a maximal multicurve corresponding to a pants decomposition. The general case that we present now was proved independently in [4, 19].

Let X be a closed hyperbolic surface of genus g and $(\Gamma, \boldsymbol{m})$ be a weighted stable graph of genus g. A $(\Gamma, \boldsymbol{m})$*-labelled multicurve* is a multicurve γ on X of topological type $(\Gamma, \boldsymbol{m})$ together with two bijections: one between the complements of γ in X and the vertices of Γ, the other between the components of γ and the edges of Γ, such that under these bijections $(\Gamma, \boldsymbol{m})$ is the dual of γ. More precisely, for each component of γ associated to an edge e, its multiplicity is m_e, and the vertices corresponding to the two components bounding this component are the endpoints of e. Given a $(\Gamma, \boldsymbol{m})$-labelled multicurve γ, we denote by γ_e the component of γ corresponding to the edge $e \in E(\Gamma)$. We also denote by $\hat{\boldsymbol{\ell}}_X(\gamma)$ the vector $(m_e \ell_X(\gamma_e)/\ell_X(\gamma))_{e \in E(\Gamma)}$ (where we do not make $(\Gamma, \boldsymbol{m})$ appear in the notation $\hat{\boldsymbol{\ell}}_X$ even though it depends on it).

Theorem 3.3 ([4,19]). *Let $g \geq 2$, and let X be a closed hyperbolic surface of genus g and $(\Gamma, \boldsymbol{m})$ be a weighted stable graph of genus g. Then we have the following convergence of probability measures on the simplex $\Delta_{=1}^{E(\Gamma)}$:*

$$\frac{1}{|\mathrm{Aut}(\Gamma, \boldsymbol{m})| \cdot s_X(R, (\Gamma, \boldsymbol{m}))} \sum_{\substack{\gamma\ (\Gamma,\boldsymbol{m})\text{-labelled} \\ \ell_X(\gamma) \leq R}} \delta_{\hat{\boldsymbol{\ell}}_X(\gamma)} \xrightarrow[R\to\infty]{} U^{(\Gamma,\boldsymbol{m})},$$

where $s_X(R, (\Gamma, \boldsymbol{m}))$ stands for the number of $(\Gamma, \boldsymbol{m})$-labelled multicurves of length at most R on X and $U^{(\Gamma,\boldsymbol{m})}$ has density given in (3.1), and where the length $\hat{\boldsymbol{\ell}}_X(\gamma)$ of a $(\Gamma, \boldsymbol{m})$-labelled multicurve γ is considered as an element of $\mathbb{R}^{E(\Gamma)}$.

In the formula above, $|\mathrm{Aut}(\Gamma, \boldsymbol{m})|$ accounts for the number of possible labellings of a multicurve γ of type $(\Gamma, \boldsymbol{m})$. In other words, $|\mathrm{Aut}(\Gamma, \boldsymbol{m})| \cdot s_X(R, (\Gamma, \boldsymbol{m}))$ is the number of $(\Gamma, \boldsymbol{m})$-labelled multicurves of length at most R.

We start with a measure-theoretic lemma.

Lemma 3.4. *Let $(\mu_{n,i})_{n,i\in\mathbb{N}}$ and $(\mu_i)_{i\in\mathbb{N}}$ be sequences of finite measures on $\Delta^{\infty}_{\leq 1}$ such that for any $i \in \mathbb{N}$, $\mu_{n,i}$ converges to μ_i as $n \to \infty$. If*

$$\lim_{n\to\infty} \sum_{i\in\mathbb{N}} \mu_{n,i}(\Delta^{\infty}_{\leq 1}) = \sum_{i\in\mathbb{N}} \mu_i(\Delta^{\infty}_{\leq 1}) < \infty,$$

then $\sum_{i\in\mathbb{N}} \mu_{n,i}$ converges to $\sum_{i\in\mathbb{N}} \mu_i$ as $n \to \infty$.

Proof. This can be seen as an application of the fact that vague and weak convergences of measures are the same under the assumption of tightness (working on $\Delta^{\infty}_{\leq 1} \times \mathbb{N}$ rather than $\Delta^{\infty}_{\leq 1}$). It could also be deduced from Fatou lemma in the following way. Suppose that $f : \Delta^{\infty}_{\leq 1} \to \mathbb{R}$ is a continuous function and let $M := \|f\|_{\infty}$. Write $f_n[i]$ for $\int_{\Delta^{\infty}_{\leq 1}} f \, d\mu_{n,i}$ and $f[i]$ for $\lim_{n\to\infty} f_n[i]$. It follows from Fatou's lemma that

$$\sum_{i\in\mathbb{N}} f[i] = \sum_{i\in\mathbb{N}} \liminf_{n\to\infty} f_n[i] \leq \liminf_{n\to\infty} \sum_{i\in\mathbb{N}} f_n[i].$$

On the other hand, applying again Fatou's lemma to $M\mu_{n,i}(\Delta^{\infty}_{\leq 1}) - f_n[i]$, we find that

$$\begin{aligned} M \sum_{i\in\mathbb{N}} \mu_i(\Delta^{\infty}_{\leq 1}) - \sum_{i\in\mathbb{N}} f[i] &= \sum_{i\in\mathbb{N}} \liminf_{n\to\infty} (M\mu_{n,i}(\Delta^{\infty}_{\leq 1}) - f_n[i]) \\ &\leq M \sum_{i\in\mathbb{N}} \mu_i(\Delta^{\infty}_{\leq 1}) + \liminf_{n\to\infty} \Big(- \sum_{i\in\mathbb{N}} f_n[i] \Big), \end{aligned}$$

and therefore,

$$\limsup_{n\to\infty} \sum_{i\in\mathbb{N}} f_n[i] \leq \sum_{i\in\mathbb{N}} f[i].$$

This completes the proof. ∎

Proof of Theorem 3.2. The left-hand side of (3.4) can be decomposed as

$$\sum_{\substack{(\Gamma,\boldsymbol{m}) \\ m_e \leq m}} \frac{s_X(R, (\Gamma, \boldsymbol{m}))}{s_X(R, m)} \frac{1}{s_X(R, (\Gamma, \boldsymbol{m}))} \sum_{\substack{\gamma \text{ of type } (\Gamma,\boldsymbol{m}) \\ \ell_X(\gamma) \leq R}} \delta_{\hat{\boldsymbol{\ell}}^{\downarrow}_X(\gamma)}. \tag{3.5}$$

A multicurve γ with topological type $(\Gamma, \boldsymbol{m})$ has exactly $|\mathrm{Aut}(\Gamma, \boldsymbol{m})|$ ways of being labelled. We thus obtain that the inner sum in (3.5) satisfies

$$\begin{aligned} &\frac{1}{s_X(R, (\Gamma, \boldsymbol{m}))} \sum_{\substack{\gamma \text{ of type } (\Gamma,\boldsymbol{m}) \\ \ell_X(\gamma) \leq R}} \delta_{\hat{\boldsymbol{\ell}}^{\downarrow}_X(\gamma)} \\ &= \frac{1}{|\mathrm{Aut}(\Gamma, \boldsymbol{m})| \cdot s_X(R, (\Gamma, \boldsymbol{m}))} \sum_{\substack{\gamma \ (\Gamma,\boldsymbol{m})\text{-labelled} \\ \ell_X(\gamma) \leq R}} \delta_{\hat{\boldsymbol{\ell}}^{\downarrow}_X(\gamma)} \xrightarrow[R\to\infty]{} U^{(\Gamma,\boldsymbol{m})\downarrow}, \end{aligned}$$

where we applied Theorem 3.3 and the fact that convergence of measures implies the convergence of their reordering (see [9, Theorem 3]).

Now, recall that it follows from Theorems 2.1 and 2.3 that the ratio $s_X(R, (\Gamma, \boldsymbol{m}))/s_X(R, m)$ converges to $c(\gamma)/b_{g,m}$, where γ denotes any multicurve of type $(\Gamma, \boldsymbol{m})$, and $c(\gamma)$ and $b_{g,m}$ are defined in (2.3) and (2.4), respectively.

Now if $m \neq \infty$, there is only a finite number of topological types $(\Gamma, \boldsymbol{m})$ with multiplicities at most m, and we can conclude directly by applying (2.3) and Theorem 2.1. For the case $m = \infty$, we apply Lemma 3.4 using the following consequence of Theorem 2.1:

$$\sum_{(\Gamma,\boldsymbol{m})} \lim_{R\to\infty} \frac{s_X(R,(\Gamma,\boldsymbol{m}))}{s_X(R)} = 1,$$

where the sum is over all weighted stable graphs $(\Gamma, \boldsymbol{m})$ of genus g. ∎

4. Reduction in the asymptotic regime

The random variable $L^{(g,m)}$ appearing in Theorem 3.2 is delicate to study because it involves a huge number of terms. Using Theorem 2.7, we show that we can restrict to a sum involving only $\mathrm{O}(\log(g))$ terms associated to non-separating multicurves.

We denote by $\Gamma_{g,k}$ the stable graph of genus g with a vertex of genus $g-k$ and k self-loops. To simplify the notation, we fix a bijection between the edges of $\Gamma_{g,k}$ and $\{1, 2, \ldots, k\}$ so that $F_{\Gamma_{g,k}}$ can be seen as a polynomial in $\mathbb{Q}[x_1, \ldots, x_k]$. Note that because the edges in $\Gamma_{g,k}$ are not distinguishable, the polynomial $F_{\Gamma_{g,k}}$ is symmetric.

We now define a simplification of $L^{(g,m)}$ as follows. Let $g \geq 2$, $m \in \mathbb{N} \cup \{\infty\}$, and $\kappa > 1$. Let $\widetilde{G}_{g,m,\kappa}$ be a random variable independent from the $U^{(\Gamma,\boldsymbol{m})}$ with distribution

$$\mathbb{P}(\widetilde{G}_{g,m,\kappa} = (\Gamma_{g,k}, \boldsymbol{m})) := \frac{1}{(6g-6)! \cdot \widetilde{b}_{g,m,\kappa}} \frac{\mathcal{Y}_{\boldsymbol{m}}(F_{\Gamma_{g,k}})}{|\mathrm{Aut}(\Gamma_{g,k}, \boldsymbol{m})|}, \tag{4.1}$$

where $k \leq \kappa \log(6g-6)/2$ and $\widetilde{b}_{g,m,\kappa}$ is defined in (2.5). Note that $\widetilde{G}_{g,m,\kappa}$ is a conditional expectation of $G_{g,m}$ defined by (3.2). Now define $\widetilde{L}^{(g,m,\kappa)}$ as

$$\widetilde{L}^{(g,m,\kappa)} := U^{\widetilde{G}_{g,m,\kappa}}, \tag{4.2}$$

where $U^{(\Gamma,\boldsymbol{m})}$ is as in Section 3. Contrarily to $L^{(g,m)\downarrow}$, we are allowed to consider a non-sorted version $\widetilde{L}^{(g,m,\kappa)}$. Indeed, $U^{\Gamma_{g,k}}$ has a density which is symmetric and hence can be considered as an element of $\Delta^k_{=1}$.

We have the following result.

Theorem 4.1. *For any $m \geq 1$, any $\kappa > 1$ and any function $h \in L^\infty(\Delta^\infty_{\leq 1})$, we have, as $g \to \infty$,*

$$\mathbb{E}(h(L^{(g,m)\downarrow})) \sim \mathbb{E}(h(\widetilde{L}^{(g,m,\kappa)\downarrow})).$$

Proof. By Theorem 2.7, a random multicurve of high genus is almost surely non-separating with less than $\kappa(\log(6g-6))/2$ edges. As h is bounded, we obtain the result. ∎

5. Size-biased sampling and Poisson–Dirichlet distribution

All our convergences are happening on the infinite simplex

$$\Delta^{\infty}_{\leq r} := \{(x_1, x_2, \ldots) \in [0, \infty)^{\mathbb{N}} : x_1 + x_2 + \cdots = r\},$$

where the topology is inherited from the product topology on $[0, r]^{\mathbb{N}}$ and hence compact. By analogy with the finite-dimensional case, one could also consider $\Delta^{\infty}_{=r}$. The latter is however dense in $\Delta^{\infty}_{\leq r}$. For this reason, it is more convenient to work with measures on $\Delta^{\infty}_{\leq r}$ even though they are ultimately supported on $\Delta^{\infty}_{=r}$.

5.1. Size-biased reordering

The components of a multicurve are not ordered in any natural way. In Theorem 1.1, we solve this issue by considering a sorted variable $L^{(g,m)\downarrow}$ on $\Delta^{3g-3}_{\leq 1}$. In this section, we introduce another natural way of ordering the entries: the size-biased ordering. Contrarily to the decreasing order, it is a random ordering. The size-biased ordering turns out to be convenient in the proof of Theorem 1.2.

We work with vectors $x = (x_1, \ldots, x_k)$ in $\Delta^k_{\leq 1}$. A *reordering* of x is a random variable of the form $(x_{\sigma(1)}, \ldots, x_{\sigma(k)})$, where σ is a random permutation in S_k. We aim to define the size-biased reordering $x^* = (x_{\sigma(1)}, \ldots, x_{\sigma(k)})$ of x.

First, let us illustrate the idea with an example. Consider $k = 4$ and

$$x = (x_1, x_2, x_3, x_4) = (0.2, 0.1, 0.5, 0.1).$$

The first coordinate x^*_1 of x's size-biased reordering $x^* = (x^*_1, x^*_2, x^*_3, x^*_4)$ can be any of its four coordinates, each with probability

$$\mathbb{P}(x^*_1 = x_i) = \frac{x_i}{x_1 + x_2 + x_3 + x_4},$$

for $i = 1, 2, 3, 4$. In other words, the probability of picking x_i in the first round is proportional to its size. Then for the second entry x^*_2, define

$$\mathbb{P}(x^*_2 = x_i \mid x^*_1 = x_i) = 0 \quad \text{and} \quad \mathbb{P}(x^*_2 = x_j \mid x^*_1 = x_i) = \frac{x_j}{x_1 + x_2 + x_3 + x_4 - x_i},$$

for $i, j \in \{1, 2, 3, 4\}$ and $j \neq i$. That is to say, if x_i has been picked as x^*_1, then x^*_2 should be chosen between x_j's where $j \neq i$, and again, the probability of picking each is proportional to its size. The other probabilities $\mathbb{P}(x^*_3 = x_k \mid x^*_1 = x_i, x^*_2 = x_j)$ are defined similarly, and so on so forth.

Formally, each $x \in \Delta^k_{\leq s}$ defines a random permutation $\sigma \in S_k$, namely a probability measure $\mathbb{P}_x$ on S_k, in the following recursive way. If $x_i = 0$ for all i, then $\sigma \in S_k$ is the identity with probability 1. If not, write $s := x_1 + \cdots + x_k$; for any $1 \leq i_1 \leq k$, set

$$\mathbb{P}_x(\sigma(1) = i_1) = \frac{x_{i_1}}{s},$$

and for any $n \geq 1$, set

$$\mathbb{P}_x(\sigma(n+1) = i_{n+1} \mid \sigma(1) = i_1, \ldots, \sigma(n) = i_n)$$

to be

- $x_{i_{n+1}}/(s - x_{i_1} - \cdots - x_{i_n})$ if $i_1, \ldots, i_{n+1}$ are distinct and $s - x_{i_1} - \cdots - x_{i_n} > 0$,
- 1 if $s = x_{i_1} + \cdots + x_{i_n}$ and $i_{n+1} = n+1$,
- 0 otherwise.

Note that if $i_1, \ldots, i_n \in \{1, \ldots, k\}$ are distinct and $s - x_{i_1} - \cdots - x_{i_n} > 0$, then

$$\mathbb{P}_x(\sigma(1) = i_1, \ldots, \sigma(n) = i_n) = \frac{x_{i_1} \cdots x_{i_n}}{s(s - x_{i_1}) \cdots (s - x_{i_1} - \cdots - x_{i_{n-1}})}. \tag{5.1}$$

Now let $X \colon \Omega \to \Delta^k_{\leq 1}$ be a random variable. In order to define its size-biased reordering $X^* \colon \Omega \to \Delta^k_{\leq 1}$, we consider for each $x \in \Delta^k_{\leq 1}$ independent random variables σ_x distributed according to $\mathbb{P}_x$ as defined above which are furthermore independent from X. We then define for each $\omega \in \Omega$,

$$X^*(\omega) := \sigma_{X(\omega)} \cdot X(\omega),$$

where $\sigma \cdot x = (x_{\sigma(1)}, \ldots, x_{\sigma(k)})$.

Lemma 5.1. *Let X be a random variable on $\Delta^k_{=1}$ with density $f_X \colon \Delta^k_{=1} \to \mathbb{R}$. Let $1 \leq r \leq k$. Then the r-th marginal of the size-biased reordering of X, which is to say the density of the vector $(X^*_1, \ldots, X^*_r)$, is*

$$\begin{aligned} &f_{(X^*_1,\ldots,X^*_r)}(x_1, \ldots, x_r) \\ &\quad = \frac{1}{(k-r)!} \frac{x_1 \cdots x_r}{(1 - x_1) \cdots (1 - x_1 - \cdots - x_{r-1})} \\ &\qquad \times \int_{\Delta^{k-r}_{=1-x_1-\cdots-x_r}} \sum_{\sigma \in S_k} f_X(x_{\sigma(1)}, \ldots, x_{\sigma(k)}) \, d\lambda^{k-r}_{=1-x_1-\cdots-x_r}(x_{r+1}, \ldots, x_k). \end{aligned}$$

Proof. Let us define $g(x_1, \ldots, x_k) := \sum_{\sigma \in S_k} f(x_{\sigma(1)}, \ldots, x_{\sigma(k)})$.

We first consider the case $r = k$. Since X admits a density, almost surely all components are positive and distinct. Hence, one can use (5.1) to write its density as

$$f_{X^*}(x_1, \ldots, x_k) = \frac{x_1 \cdots x_k}{(1 - x_1) \cdots (1 - x_1 - \cdots - x_{k-1})} g(x_1, \ldots, x_k).$$

In the above formula, we used the fact that the sum of X is $s = 1$ almost surely.

Now, for $1 \leq r \leq k-1$, the r-th marginal is obtained by integrating the free variables

$$f_{(X^*_1,\ldots,X^*_r)}(x_1, \ldots, x_r) = \int_{\Delta^{k-r}_{=1-s}} f_{X^*}(x_1, \ldots, x_k) \, d\lambda^{k-r}_{=1-s}(x_{r+1}, \ldots, x_k), \tag{5.2}$$

where $s = x_1 + \cdots + x_r$. For a permutation $\tau \in S(\{r+1, \ldots, k\})$ of the set $\{r+1, \ldots, k\}$, we define the subsimplex

$$\Delta^{k-r}_{=1-s;\tau} := \{(x_{r+1}, \ldots, x_k) \in \Delta^{k-r}_{=1-s} : x_{\tau(r+1)} > x_{\tau(r+2)} > \cdots > x_{\tau(k)}\}.$$

We can decompose integral (5.2) as a sum over these subsimplices

$$f_{(X_1^*,\dots,X_r^*)}(x_1,\dots,x_r) = \frac{x_1\cdots x_r}{(1-x_1)\cdots(1-x_1-\cdots-x_{r-1})(1-s)} \times \sum_{\tau\in S(\{r+1,\dots,k\})} \int_{\Delta^{k-r}_{=1-s;\tau}} \frac{x_{r+1}\cdots x_k\, g(x_1,\dots,x_k)\,\mathrm{d}\lambda^{k-r}_{=1-s}(x_{r+1},\dots,x_k)}{(1-s-x_{r+1})\cdots(1-s-x_{r+1}-\cdots-x_{k-1})}.$$

Using the fact that g is symmetric, we can rewrite it by means of a change of variables on the standard simplex $\Delta^{k-r}_{=1-s;\mathrm{id}}$

$$f_{(X_1^*,\dots,X_r^*)}(x_1,\dots,x_r) = \frac{x_1\cdots x_r}{(1-x_1)\cdots(1-x_1-\cdots-x_{r-1})(1-s)} \times \int_{\Delta^{k-r}_{=1-s;\mathrm{id}}} \sum_{\tau\in S(\{r+1,\dots,k\})} \frac{x_{\tau(r+1)}\cdots x_{\tau(k)}\, g(x_1,\dots,x_k)\,\mathrm{d}\lambda^{k-r}_{=1-s}(x_{r+1},\dots,x_k)}{(1-s-x_{\tau(r+1)})\cdots(1-s-x_{\tau(r+1)}-\cdots-x_{\tau(k-1)})}.$$

Using the facts that

$$\sum_{\tau\in S(\{r+1,\dots,k\})} \frac{x_{\tau(r+1)}\cdots x_{\tau(k)}}{(1-s-x_{\tau(r+1)})\cdots(1-s-x_{\tau(r+1)}-\cdots-x_{\tau(k-1)})} = 1-s$$

and

$$\int_{\Delta^{k-r}_{=1-s;\mathrm{id}}} \sum_{\sigma\in S_k} f(x_{\sigma(1)},\dots,x_{\sigma(k)})\,\mathrm{d}\lambda^{k-r}_{=1-s}(x_{r+1},\dots,x_k) = \frac{1}{(k-r)!}\int_{\Delta^{k-r}_{=1-s}} \sum_{\sigma\in S_k} f(x_{\sigma(1)},\dots,x_{\sigma(k)})\,\mathrm{d}\lambda^{k-r}_{=1-s}(x_{r+1},\dots,x_k),$$

we obtain the result. ∎

We finish this section by mentioning that the size-biased reordering extends to infinite vectors that are elements of $\Delta^\infty_{\le 1}$.

5.2. *Poisson–Dirichlet and Griffiths–Engen–McCloskey distributions*

Recall that the GEM(θ) distribution was defined in the introduction via the stick-breaking process. We also defined the PD(θ) as the sorted reordering of GEM(θ). The Poisson–Dirichlet distribution admits an intrinsic definition in terms of the Poisson process first introduced by Kingman [17]. We refer to [5, Section 4.11] for this definition. Instead, we concentrate on the simpler Griffiths–Engen–McCloskey distribution.

In the introduction, we passed from GEM(θ) to PD(θ). The following result formalizes the equivalence between these two distributions.

Theorem 5.2 ([9, Theorem 5]). *Let $\theta > 0$. If a random variable $X \in \Delta^{\infty}_{=1}$ is distributed according to* PD(θ)*, then its size-biased permutation X^* follows* GEM(θ).

The following result is a direct consequence of [9, Theorems 1, 2 and 5].

Corollary 5.3 ([9]). *Let $X^{(n)}$ be a sequence of random variables with values in $\Delta^{\infty}_{=1}$ and let $\theta > 0$. Then the sorted sequence $X^{(n)\downarrow}$ converges in distribution to* PD(θ) *if and only if the size-biased sequence $X^{(n)*}$ converges in distribution to* GEM(θ).

In order to prove convergence towards GEM, we will need the explicit description of its marginals.

Proposition 5.4 ([9]). *Let $X = (X_1, X_2, \dots)$ be a random variable with distribution* GEM(θ). *Then the distribution of the r-first components $(X_1, X_2, \dots, X_r)$ of X supported on $\Delta^r_{\leq 1}$ admits a density given by*

$$\frac{\theta^r (1 - x_1 - \cdots - x_r)^{\theta - 1}}{(1 - x_1)(1 - x_1 - x_2) \cdots (1 - x_1 - \cdots - x_{r-1})}. \tag{5.3}$$

In order to simplify computations, we consider moments of the GEM distribution that get rid of the denominator in density (5.3). Namely, for a random variable $X = (X_1, X_2, \dots)$ on $\Delta^{\infty}_{\leq 1}$ and an r-tuple of non-negative integers $p = (p_1, \dots, p_r)$, we define

$$M_p(X) := \mathbb{E}((1 - X_1) \cdots (1 - X_1 - \cdots - X_{r-1}) \cdot X_1^{p_1} \cdots X_r^{p_r}). \tag{5.4}$$

These moments of the GEM(θ) can be computed using the following elementary lemma.

Lemma 5.5. *Let $d_1, \dots, d_k \in \mathbb{R}_{\geq 0}$. Then*

$$\int_{\Delta^k_{\leq r}} x_1^{d_1} \cdots x_k^{d_k} \, \mathrm{d}\lambda^k_{\leq r} = \frac{d_1! \cdots d_k!}{(d_1 + \cdots + d_k + k)!} \cdot r^{d_1 + \cdots + d_k + k}$$

and

$$\int_{\Delta^k_{=r}} x_1^{d_1} \cdots x_k^{d_k} \, \mathrm{d}\lambda^k_{=r} = \frac{d_1! \cdots d_k!}{(d_1 + \cdots + d_k + k - 1)!} \cdot r^{d_1 + \cdots + d_k + k - 1}.$$

Here the factorial of a real number has to be considered by means of the analytic continuation given by the gamma function: $x! = \Gamma(x + 1)$.

Proof. For $x \in \mathbb{R}_{>0}$ and $\alpha, \beta \in \mathbb{R}_{>1}$, we have the following scaling of the beta function:

$$\int_0^x t^{\alpha - 1}(x - t)^{\beta - 1} \, \mathrm{d}t = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha + \beta)} x^{\alpha + \beta - 1}.$$

This implies that

$$\int_{\Delta^k_{\leq r}} \frac{x_1^{d_1}}{d_1!} \cdots \frac{x_k^{d_k}}{d_k!} \, \mathrm{d}\lambda^k_{\leq r} = \int_{\Delta^{k-1}_{\leq r}} \frac{x_1^{d_1}}{d_1!} \cdots \frac{x_{k-1}^{d_{k-1} + d_k + 1}}{(d_{k-1} + d_k + 1)!} \, \mathrm{d}\lambda^{k-1}_{\leq r}$$

and

$$\int_{\Delta^k_{=r}} \frac{x_1^{d_1}}{d_1!} \cdots \frac{x_k^{d_k}}{d_k!} \, \mathrm{d}\lambda^k_{=r} = \int_{\Delta^{k-1}_{=r}} \frac{x_1^{d_1}}{d_1!} \cdots \frac{x_{k-1}^{d_{k-1}+d_k+1}}{(d_{k-1}+d_k+1)!} \, \mathrm{d}\lambda^{k-1}_{=r}.$$

The two equations in the statement then follow by induction. ∎

We now show how to deduce Lemma 3.1 using Lemma 5.5.

Proof of Lemma 3.1. From the second equation in the statement of Lemma 5.5, it follows that

$$\mathcal{Y}_{\boldsymbol{m}}(F_\Gamma) = (6g-7)! \int_{\Delta^k_{=1}} \frac{1}{m_1 \cdots m_k} F_\Gamma\Big(\frac{x_1}{m_1}, \ldots, \frac{x_k}{m_k}\Big) \, \mathrm{d}\lambda^k_{=1}(x_1, \ldots, x_k).$$

Indeed, each monomial that appears in F_Γ has k variables and total degree $6g-6-k$. Hence, the denominator coming from the formula of Lemma 5.5 compensates the $(6g-7)!$ term from (3.1). The numerator in the formula of Lemma 5.5 matches the definition of $\mathcal{Y}_{\boldsymbol{m}}$. ∎

Lemma 5.6. *If* $X = (X_1, X_2, \ldots) \sim \mathrm{GEM}(\theta)$ *and* $(p_1, \ldots, p_r)$ *is a non-negative integral vector, then the moment* $M_p(X)$ *defined in* (5.4) *has the following value:*

$$M_p(X) = \frac{\theta^r \cdot \Gamma(\theta) \cdot p_1! \cdots p_r!}{\Gamma(p_1 + \cdots + p_r + \theta + r)}.$$

Proof. By Proposition 5.4, we have

$$\begin{aligned} M_p(X) &= \int_{\Delta^r_{\le 1}} \theta^r x_1^{p_1} \cdots x_r^{p_r} (1 - x_1 - \cdots - x_r)^{\theta-1} \, \mathrm{d}\lambda^r_{\le 1} \\ &= \theta^r \int_{\Delta^{r+1}_{=1}} x_1^{p_1} \cdots x_r^{p_r} x_{r+1}^{\theta-1} \, \mathrm{d}\lambda^{r+1}_{=1}. \end{aligned}$$

The last term is an instance of Lemma 5.5 on the simplex $\Delta^{r+1}_{=1}$. This completes the proof. ∎

6. Proof of the main theorem

The aim of this section is to prove the following result.

Theorem 6.1. *For* $g \ge 2$ *integral,* $m \in \mathbb{N} \cup \{\infty\}$ *and* $\kappa > 1$ *real, let* $\widetilde{L}^{(g,m,\kappa)\downarrow}$ *be as in Theorem* 4.1. *Then, as* g *tends to infinity,* $\widetilde{L}^{(g,m,\kappa)*}$ *converges in distribution to* GEM(1/2).

Let us first show how to derive our main Theorem 1.2 from Theorem 6.1.

Proof of Theorem 1.2. By Theorem 4.1, the random variables $L^{(g,m,\kappa)\downarrow}$ and $L^{(g,m)\downarrow}$ have the same limiting distribution as $g \to \infty$. Hence, by Theorem 6.1 and Corollary 5.3, the random variable $L^{(g,m)*}$ converges in distribution towards GEM(1/2).

Finally, Corollary 5.3 shows that the convergence in distribution of $L^{(g,m)*}$ towards GEM(1/2) is equivalent to the convergence of $L^{(g,m)\downarrow}$ towards PD(1/2). This concludes the proof of Theorem 1.2. ∎

6.1. Moments method

Let us recall from equation (5.4) that we defined some specific moments $M_{(p_1,\dots,p_r)}(X)$ for a random variable X on $\Delta^{\infty}_{=1}$. In this section, we show that the convergence of a sequence of random variables $X^{(n)}$ is equivalent to the convergence of all the moments $M_p(X^{(n)})$. This strategy called the *method of moments* is a standard tool in probability; see, for example, [6, Section 30] for the case of real variables.

Lemma 6.2. *A sequence of random variables $X^{(n)} = (X_1^{(n)}, X_2^{(n)}, \dots) \in \Delta^{\infty}_{=1}$ converges in distribution to a random variable $X^{(\infty)}$ in $\Delta^{\infty}_{=1}$ if and only if for all $p = (p_1, \dots, p_r)$ vector of non-negative integers, we have $\lim_{n\to\infty} M_p(X^{(n)}) = M_p(X^{(\infty)})$.*

Proof. The infinite-dimensional cube $[0,1]^{\mathbb{N}}$ is compact with respect to the product topology by Tychonoff's theorem. The set $\Delta^{\infty}_{\leq 1}$ is a closed subset of $[0,1]^{\mathbb{N}}$, and is therefore compact. The signed measures on $\Delta^{\infty}_{\leq 1}$ are identified with the dual of real continuous functions $C(\Delta^{\infty}_{\leq 1}, \mathbb{R})$. In particular, we have the convergence of $X^{(n)}$ towards $X^{(\infty)}$ in distribution if and only if for any continuous function $f \in C(\Delta^{\infty}_{\leq 1}, \mathbb{R})$, we have the convergence of $\mathbb{E}(f(X^{(n)}))$ towards $\mathbb{E}(f(X^{(\infty)}))$.

Now let S be the set of functions in $C(\Delta^{\infty}_{\leq 1}, \mathbb{R})$ of the form

$$(1 - x_1)(1 - x_1 - x_2) \cdots (1 - x_1 - \cdots - x_{r-1}) \cdot x_1^{p_1} \cdots x_r^{p_r}$$

with $r \geq 0$, $p_1, \dots, p_r \geq 0$. We claim that the span of S (that is, finite linear combinations of elements of S) is dense in $C(\Delta^{\infty}_{\leq 1}, \mathbb{R})$.

Indeed, S contains the constant function 1, and the algebra generated by S is equal to its span.

Now, the set S is a separating subset of $C(\Delta^{\infty}_{\leq 1}, \mathbb{R})$, and density follows from the Stone–Weierstrass theorem. ■

We will use the following asymptotic simplification of the moments.

Theorem 6.3. *For $g \geq 2$ integral, $m \in \mathbb{N} \cup \{\infty\}$ and $\kappa > 1$ real, let $\widetilde{L}^{(g,m,\kappa)*}$ be the size-biased reordering of the random variable $\widetilde{L}^{(g,m,\kappa)}$ from Theorem 4.1. Let $r \geq 1$ and $p_1, \dots, p_r \in \mathbb{N}$. Then, as $g \to \infty$, the moment $M_p(\widetilde{L}^{(g,m,\kappa)*})$ is asymptotically equivalent to*

$$\frac{\sqrt{(m+1)/m} \cdot \sqrt{\pi}}{2 \cdot (6g-6)^{p_1 + \cdots + p_r + r - 1/2}} \sum_{k=r}^{\kappa(\log(6g-6))/2} \frac{1}{(k-r)!} \times \sum_{\substack{(j_1,\dots,j_k) \in \mathbb{N}^k \\ j_1 + \cdots + j_k = 3g-3}} \prod_{i=1}^{k} \frac{\zeta_m(2j_i)}{2j_i} \prod_{i=1}^{r} \frac{(2j_i + p_i)!}{(2j_i - 1)!},$$

where

$$\zeta_m(s) := \sum_{n=1}^{m} \frac{1}{n^s}$$

is the partial Riemann zeta function.

Following [8, (14)], we define

$$c_{g,k}(d_1,\ldots,d_k) := \frac{g!(3g-3+2k)!}{(6g-5+4k)!}\frac{3^g}{2^{3g-6+5k}}(2d_1+2)!\cdots(2d_k+2)! \times \sum_{\substack{d_i^-+d_i^+=d_i\\ d_i^-,d_i^+\geq 0\\ 1\leq i\leq k}} \frac{\langle \tau_{d_1^-}\tau_{d_1^+}\cdots\tau_{d_k^-}\tau_{d_k^+}\rangle_{g,2k}}{d_1^-!\,d_1^+!\cdots d_k^-!\,d_k^+!}.$$

The above coefficients were introduced in [8, Lemma 3.5], and we have for any $C>0$

$$\lim_{g\to\infty}\sup_{k\leq C\log(g)}|c_{g,k}(d_1,\ldots,d_k)-1|=0. \tag{6.1}$$

This asymptotic result is a direct consequence of Theorem 2.6 of Aggarwal that we stated in the introduction. We define

$$\widetilde{c}_{g,k}(j_1,\ldots,j_k) := c_{g-k,k}(j_1-1,\ldots,j_k-1). \tag{6.2}$$

Lemma 6.4. *For each $k=1,\ldots,3g-3$, let $U^{(g,m,k)}$ be the random variable on $\Delta^k_{=1}$ with density*

$$\frac{(6g-7)!}{\mathcal{Z}_m(F_{g,k})}\sum_{\substack{(m_1,\ldots,m_k)\in\mathbb{N}^k\\ m_i\leq m\\ i\in\{1,\ldots,k\}}} F_{g,k}\left(\frac{x_1}{m_1},\ldots,\frac{x_k}{m_k}\right)\frac{1}{m_1\cdots m_k}, \tag{6.3}$$

where for simplicity we write $F_{g,k}$ for the graph polynomial $F_{\Gamma_{g,k}}$ defined in (2.1). Then for any integer $r\in\mathbb{N}$ and $p=(p_1,\ldots,p_r)\in\mathbb{N}^r$, we have

$$M_p(U^{(g,m,k)*})=0$$

if $r>k$, and

$$M_p(U^{(g,m,k)*}) = \frac{w_{g,k}\cdot k!}{\mathcal{Z}_m(F_{g,k})\cdot(k-r)!\cdot(6g-7+p_1+\cdots+p_r+r)!} \times \sum_{\substack{(j_1,\ldots,j_k)\in\mathbb{N}^k\\ j_1+\cdots+j_k=3g-3}} \widetilde{c}_{g,k}(j_1,\ldots,j_k)\prod_{i=1}^k\frac{\zeta_m(2j_i)}{2j_i}\prod_{i=1}^r\frac{(2j_i+p_i)!}{(2j_i-1)!},$$

$$w_{g,k} := \frac{(6g-5-2k)!\cdot(6g-7)!}{(g-k)!\cdot(3g-3-k)!}\cdot\frac{2^{3k-3}}{3^{g-k}} \tag{6.4}$$

if $1\leq r\leq k$.

Note that expression (6.3) is the density of the conditional expectation $\mathbb{E}(\widetilde{L}^{(g,m,\kappa)} \mid \widetilde{G}_{g,m,\kappa}=\Gamma_{g,k})$, where $\widetilde{L}^{(g,m,\kappa)}$ was defined in Section 4. It is more precisely the density of the asymptotic normalized vector of lengths of random multicurves restricted to multicurves of type $\Gamma_{g,k}$.

Proof of Lemma 6.4. It follows from the definition of M_p that

$$M_p(U^{(g,m,k)*}) = 0$$

if $r > k$. We assume from now on that $r \le k$. By definition of the stable graph polynomial, we have

$$\begin{aligned} F_{g,k}(x_1, x_2, \ldots, x_k) &= x_1 \cdots x_k \cdot V_{g-k,2k}(x_1, x_1, x_2, x_2, \ldots, x_k, x_k) \\ &= \frac{1}{2^{3g-3-k}} \sum_{\substack{(d_1^-, d_1^+, \ldots, d_k^-, d_k^+) \in \mathbb{Z}_{\ge 0}^{2k} \\ d_1^- + d_1^+ + \cdots + d_k^- + d_k^+ = 3g-3-k}} \frac{\langle \tau_{d_1^-} \tau_{d_1^+} \cdots \tau_{d_k^-} \tau_{d_k^+} \rangle_{g-k,2k}}{d_1^-! d_1^+! \cdots d_k^-! d_k^+!} \\ &\quad \times x_1^{2(d_1^+ + d_1^-)+1} \cdots x_k^{2(d_k^+ + d_k^-)+1}. \end{aligned}$$

Using the coefficients $\widetilde{c}_{g,k}$ defined just above the statement of the lemma, we rewrite the polynomial $F_{g,k}$ as

$$F_{g,k} = \frac{(6g-5-2k)!}{(g-k)! \cdot (3g-3-k)!} \frac{2^{3k-3}}{3^{g-k}} \sum_{\substack{(j_1, \ldots, j_k) \in \mathbb{N}^k \\ j_1 + \cdots + j_k = 3g-3}} \widetilde{c}_{g,k}(j_1, \ldots, j_k) \prod_{i=1}^{k} \frac{x_i^{2j_i - 1}}{(2j_i)!}.$$

Hence, the density of $U^{(g,m,k)}$ in (6.3) can be rewritten as

$$\frac{w_{g,k}}{\mathcal{Z}_m(F_{g,k})} \sum_{\substack{(j_1, \ldots, j_k) \in \mathbb{N}^k \\ j_1 + \cdots + j_k = 3g-3}} \widetilde{c}_{g,k}(j_1, \ldots, j_k) \prod_{i=1}^{k} \zeta_m(2j_i) \frac{x_i^{2j_i - 1}}{(2j_i)!}.$$

Now, according to Lemma 5.1, the r-th marginal of the sized-biased version $U^{(g,m,k)*}$ of $U^{(g,m,k)}$ is

$$\begin{aligned} &\frac{w_{g,k} \cdot k!}{\mathcal{Z}_m(F_{g,k}) \cdot (k-r)!} \cdot \frac{1}{(1-x_1) \cdots (1 - x_1 - \cdots - x_{r-1})} \\ &\quad \times \sum_{\substack{(j_1, \ldots, j_k) \in \mathbb{N}^k \\ j_1 + \cdots + j_k = 3g-3}} \widetilde{c}_{g,k}(j_1, \ldots, j_k) \prod_{i=1}^{k} \frac{\zeta_m(2j_i)}{(2j_i)!} \prod_{i=1}^{r} x_i^{2j_i} \\ &\quad \times \int_{\Delta^k_{=1-x_1-\cdots-x_r}} x_{r+1}^{2j_{r+1}-1} \cdots x_k^{2j_k - 1} \, d\lambda^{k-r}_{=1-x_1-\cdots-x_r}(x_{r+1}, \ldots, x_k), \end{aligned}$$

where we used the fact that the density of $U^{(g,m,k)}$ is a symmetric function. Hence, the sum over all permutations of k elements only pops out a $k!$ coefficient. The value of the integral in the above sum follows from Lemma 5.5 and is equal to

$$\frac{(2j_{r+1} - 1)! \cdots (2j_k - 1)!}{(2j_{r+1} + \cdots + 2j_k - 1)!} (1 - x_1 - \cdots - x_r)^{2j_{r+1} + \cdots + 2j_k - 1}.$$

Now we end up with the following formula for the distribution of the r-th marginal of $U^{(g,m,k)*}$:

$$\frac{w_{g,k}\cdot k!}{\mathcal{Z}_m(F_{g,k})\cdot(k-r)!}\sum_{\substack{(j_1,\dots,j_k)\in\mathbb{N}^k\\ j_1+\dots+j_k=3g-3}}\frac{\widetilde{c}_{g,k}(j_1,\dots,j_k)}{(2j_{r+1}+\dots+2j_k-1)!}\cdot\prod_{i=1}^{k}\frac{\zeta_m(2j_i)}{2j_i}$$
$$\times\prod_{i=1}^{r}\frac{x_i^{2j_i}}{(2j_i-1)!}\frac{(1-x_1-\dots-x_r)^{2j_{r+1}+\dots+2j_k-1}}{(1-x_1)\cdots(1-x_1-\dots-x_{r-1})}.$$

From the above formula and the definition of the moment M_p in (5.4), $M_p(U^{(g,m,k)*})$ equals

$$\frac{w_{g,k}\cdot k!}{\mathcal{Z}_m(F_{g,k})\cdot(k-r)!}\sum_{\substack{(j_1,\dots,j_k)\in\mathbb{N}^k\\ j_1+\dots+j_k=3g-3}}\frac{\widetilde{c}_{g,k}(j_1,\dots,j_k)}{(2j_{r+1}+\dots+2j_k-1)!}$$
$$\times\prod_{i=1}^{k}\frac{\zeta_m(2j_i)}{2j_i}\prod_{i=1}^{r}\frac{1}{(2j_i-1)!}$$
$$\times\int_{\Delta^r_{\le 1}}x_1^{2j_1+p_1}\cdots x_r^{2j_r+p_r}(1-x_1-\dots-x_r)^{2j_{r+1}+\dots+2j_k-1}\,d\lambda^r_{\le 1}.$$

Lemma 5.5 gives the value of the above integral

$$\int_{\Delta^r_{\le 1}}x_1^{2j_1+p_1}\cdots x_r^{2j_r+p_r}(1-x_1-\dots-x_r)^{2j_{r+1}+\dots+2j_k-1}\,d\lambda^r_{\le 1}$$
$$=\int_{\Delta^{r+1}_{=1}}x_1^{2j_1+p_1}\cdots x_r^{2j_r+p_r}x_{r+1}^{2j_{r+1}+\dots+2j_k-1}\,d\lambda^{r+1}_{=1}$$
$$=\frac{(2j_1+p_1)!\cdots(2j_r+p_r)!\cdot(2j_{r+1}+\dots+2j_k-1)!}{(6g-6+p_1+\dots+p_r+r-1)!}.$$

Substituting the above value in our last expression for $M_p(U^{(g,m,k)*})$ gives the announced formula. ∎

Proof of Theorem 6.3. By definition (4.2) of $\widetilde{L}^{(g,m,\kappa)}$ and the fact that

$$|\mathrm{Aut}(\Gamma_{g,k})|=2^k k!,$$

we have

$$M_p(\widetilde{L}^{(g,m,\kappa)*})=\frac{1}{(6g-6)!\,\widetilde{b}_{g,m,\kappa}}\sum_{k=1}^{\kappa(\log(6g-6))/2}\frac{\mathcal{Z}_m(F_{g,k})}{2^k k!}\cdot M_p(U^{(g,m,k)*}),\qquad(6.5)$$

where $\widetilde{b}_{g,m,\kappa}$ was defined in (2.5).

Now substituting the formula for $M_p(U^{(g,m,k)*})$ from Lemma 6.4 and the asymptotic value of $\widetilde{b}_{g,m,\kappa}$ from Theorem 2.7 in sum (6.5), we have, as $g \to \infty$, the asymptotic equivalence

$$M_p(L^{(g,m,\kappa)*}) \sim \frac{(4g-4)! \cdot \pi}{(6g-7)! \cdot (6g-7+p_1+\cdots+p_r+r)!} \cdot \sqrt{\frac{m+1}{m}} \cdot \left(\frac{3}{4}\right)^{4g-4} \times \sum_{k=r}^{\kappa(\log(6g-6))/2} \left(\frac{1}{2^k \cdot (k-r)!} \cdot \frac{(6g-5-2k)! \cdot (6g-7)! \cdot 2^{3k-3}}{(g-k)! \cdot (3g-3-k)! \cdot 3^{g-k}} \times \sum_{\substack{(j_1,\ldots,j_k)\in\mathbb{N}^k \\ j_1+\cdots+j_k=3g-3}} \prod_{i=1}^{k} \frac{\zeta_m(2j_i)}{2j_i} \prod_{i=1}^{r} \frac{(2j_i+p_i)!}{(2j_i-1)!}\right), \quad (6.6)$$

where we have used that $\widetilde{c}_{g,k}(j_1,\ldots,j_k) \sim 1$ uniformly in $k \in [1, \kappa \log(6g-6)/2]$, which follows from (6.1). By [8, proof of Theorem 3.4, (3.13)], we have

$$\frac{(4g-4)! \cdot (6g-5-2k)!}{(6g-7)! \cdot (g-k)! \cdot (3g-3-k)!} \sim (6g-6)^{1/2} \frac{1}{\sqrt{\pi}} \frac{2^{8g-6-2k}}{3^{3g-4+k}}. \quad (6.7)$$

On the other hand,

$$\frac{(6g-7)!}{(6g-7+p_1+\cdots+p_r+r)!} \sim \frac{1}{(6g-6)^{p_1+\cdots+p_r+r}} \quad (6.8)$$

as $g \to \infty$. Replacing (6.7) and (6.8) in (6.6), we obtain

$$M_p(\widetilde{L}^{(g,m,\kappa)*}) \sim \frac{1}{2(6g-6)^{p_1+\cdots+p_r+r-1/2}} \cdot \sqrt{\frac{m+1}{m}} \cdot \sqrt{\pi} \cdot \sum_{k=r}^{\kappa(\log(6g-6))/2} \frac{1}{(k-r)!} \times \sum_{\substack{(j_1,\ldots,j_k)\in\mathbb{N}^k \\ j_1+\cdots+j_k=3g-3}} \prod_{i=1}^{k} \frac{\zeta_m(2j_i)}{2j_i} \prod_{i=1}^{r} \frac{(2j_i+p_i)!}{(2j_i-1)!}$$

which is the announced formula. ∎

6.2. *Asymptotic expansion of a related sum*

Let $\theta = (\theta_i)_{i\geq 1}$ be a sequence of non-negative real numbers and let $p = (p_1,\ldots,p_r)$ be a non-negative integral vector. This section is dedicated to the asymptotics in n of the numbers

$$S_{\theta,p,n} := \sum_{k=r}^{\infty} \frac{1}{(k-r)!} \sum_{\substack{(j_1,\ldots,j_k)\in\mathbb{N}^k \\ j_1+\cdots+j_k=n}} \left(\prod_{i=1}^{k} \frac{\theta_i}{2j_i}\right)\left(\prod_{i=1}^{r} \frac{(2j_i+p_i)!}{(2j_i-1)!}\right), \quad (6.9)$$

which should be reminiscent of the formula from Theorem 6.3.

Definition 6.5. Let $\theta = (\theta_j)_{j\geq 1}$ be non-negative real numbers and let $g_\theta(z)$ be the formal series

$$g_\theta(z) := \sum_{j\geq 1} \theta_j \frac{z^j}{j}.$$

We say that θ is *admissible* if the function $g_\theta(z)$

- converges in the open disc $D(0,1) \subset \mathbb{C}$ centred at 0 of radius 1,
- $g_\theta(z) + \log(1-z)$ extends to a holomorphic function on $D(0,R)$ with $R > 1$.

Theorem 6.6. *Let $\theta = (\theta_k)_{k\geq 1}$ be admissible, then, as $n \to \infty$, we have*

$$S_{\theta,p,n} \sim \sqrt{\frac{e^\beta}{2}} \cdot \frac{p_1!\cdots p_r!}{2^{r-1}} \frac{(2n)^{p_1+\cdots+p_r+r-1/2}}{\Gamma(p_1+\cdots+p_r+r+1/2)},$$

where β is the value at $z=1$ of $g_\theta(z) + \log(1-z)$.

The following is essentially [8, Lemma 3.8] which we reproduce for completeness.

Lemma 6.7. *For $m \in \mathbb{N} \cup \{\infty\}$, let*

$$g_m(z) := \sum_{j\geq 1} \zeta_m(2j) \frac{z^j}{j}. \tag{6.10}$$

Then $g_m(z)$ is summable in $D(0,1)$ and $g_m(z) + \log(1-z)$ extends to a holomorphic function on $D(0,4)$. In particular, the sequence $(\zeta(2j))_{j\geq 1}$ is admissible. Moreover,

$$(g_m(z) + \log(1-z))|_{z=1} = \log\Big(\frac{2m}{(m+1)}\Big).$$

Proof. Since $\zeta_m(2j)$ is bounded uniformly in j, the series converges in $D(0,1)$. Now, expanding the definition of the partial zeta function ζ_m and changing the order of summation, we have for $z \in D(0,1)$

$$g_m(z) = -\sum_{n=1}^{m} \log\Big(1 - \frac{z}{n^2}\Big),$$

and hence

$$g_m(z) + \log(1-z) = -\sum_{n=2}^{m} \log\Big(1 - \frac{z}{n^2}\Big).$$

The term $\log(1 - z/n^2)$ defines a holomorphic function on $D(0,n^2)$. Since

$$\Big|\log\Big(1 - \frac{z}{n^2}\Big)\Big| \leq \frac{4}{n^2}\Big|\log\Big(1 - \frac{z}{4}\Big)\Big|,$$

we have absolute convergence even for $m = \infty$ and $g_m(z) + \log(1-z)$ defines a holomorphic function in $D(0,4)$.

Now for the value at $z = 1$, we obtain

$$\begin{aligned}(g_m(z) + \log(1-z))|_{z=1} &= -\sum_{n=2}^{m} \log\Big(1 - \frac{1}{n^2}\Big) \\ &= \sum_{n=2}^{m} (2\log(n) - \log(n-1) - \log(n+1)) \\ &= \log\Big(\frac{2m}{m+1}\Big).\end{aligned}$$

This completes the proof. ■

Combining Theorem 6.6 and Lemma 6.7, we obtain the following corollary.

Corollary 6.8. *Let $m \in \mathbb{N} \cup \{\infty\}$. For $\theta = (\zeta_m(2i))_{i\geq 1}$, we have*

$$S_{\theta,p,n} \sim \sqrt{\frac{m}{m+1}} \cdot \frac{p_1! \cdots p_r!}{2^{r-1}} \frac{(2n)^{p_1+\cdots+p_r+r-1/2}}{\Gamma(p_1+\cdots+p_r+r+1/2)}.$$

Now let us prove Theorem 6.6. For a non-negative integer p, we define the differential operator on $\mathbb{C}[\![z]\!]$ by

$$D_p(f) := z\frac{\mathrm{d}^{p+1}}{\mathrm{d}z^{p+1}}(z^p f).$$

We start with some preliminary lemmas.

Lemma 6.9. *Let $\theta = (\theta_i)_i$ and let $g_\theta(z)$ be as in Theorem 6.6. Let $p = (p_1,\ldots,p_r)$ be a tuple of non-negative integers and let*

$$G_{\theta,p}(z) := \exp\Big(\frac{1}{2}g_\theta(z^2)\Big)\prod_{i=1}^{r} D_{p_i}\Big(\frac{1}{2}g_\theta(z^2)\Big).$$

Then, for any $n \geq 0$ we have

$$[z^{2n}]G_{\theta,p}(z) = S_{\theta,p,n},$$

where $[z^{2n}]$ is the coefficient extraction operator and $S_{\theta,p,n}$ is the sum in (6.9).

Proof. Let us first note that $g_\theta(z^2)/2 = \sum_i \theta_i z^{2i}/(2i)$. We aim to compute the expansion of $D_p(g_\theta(z^2)/2)$. By linearity, it is enough to compute a single term, and we have

$$D_p(z^{2j}) = z\frac{d^{p+1}}{dz^{p+1}}(z^{2j+p}) = \frac{(2j+p)!}{(2j-1)!}z^{2j}.$$

Hence,

$$D_p\Big(\frac{1}{2}g_\theta(z^2)\Big) = \sum_{j=1}^{\infty}\frac{(2j+p)!}{(2j-1)!}\theta_j\frac{z^{2j}}{2j}.$$

The lemma follows by expanding the exponential. ■

Lemma 6.10. *For any $p \in \mathbb{Z}_{\geq 0}$, we have*

$$\frac{1}{p!} D_p(-\log(1 \pm z)) = \frac{1}{(1 \pm z)^{p+1}} - 1.$$

Proof. By Leibniz's rule,

$$\begin{aligned}\frac{z}{p!}\frac{\mathrm{d}^{p+1}}{\mathrm{d}z^{p+1}}\Big(z^p \log\frac{1}{1-z}\Big) &= \frac{z}{p!}\sum_{i=0}^{p}\binom{p+1}{i} p(p-1)\cdots(p-i+1)z^{p-i}\frac{(p-i)!}{(1-z)^{p+1-i}} \\ &= -1 + \sum_{i=0}^{p+1}\binom{p+1}{i}\Big(\frac{z}{1-z}\Big)^{p-i+1} \\ &= -1 + \Big(1 + \frac{z}{1-z}\Big)^{p+1} = -1 + \frac{1}{(1-z)^{p+1}}.\end{aligned}$$

The proof for $-\log(1+z)$ is similar. ■

Proof of Theorem 6.6. By Lemma 6.9, the sum $S_{\theta,p,n}$ is the coefficient of z^{2n} in $G_{\theta,p}$. By the conditions in the statement, we can write $g_\theta(z^2) = -\log(1-z^2) + \beta + r_\theta(z)$, where $r_\theta(z)$ is holomorphic on $D(0,\sqrt{R})$ and $r_\theta(1) = 0$. Using Lemma 6.10, we deduce that, for any $p \geq 0$,

$$D_p(g_\theta(z^2)) = \frac{p!}{(1-z)^{p+1}} + \frac{p!}{(1+z)^{p+1}} + r_{\theta,p}(z),$$

where $r_{\theta,p}$ is holomorphic in $D(0,\sqrt{R})$. Thus, $G_{\theta,p}(z)$ is holomorphic in $D(0,\sqrt{R}) \setminus ([1,\sqrt{R}) \cup (-\sqrt{R},-1])$ and satisfies, as $z \to 1$,

$$\begin{aligned}G_{\theta,p}(z) &= \exp\Big(-\frac{1}{2}(\log(1-z) + \log(2) -\Big) + \mathrm{O}(1-z)\Big) \\ &\quad \times \prod_{i=1}^{r}\frac{p_i!}{2}\Big(\frac{1}{(1-z)^{p_i+1}} + \mathrm{O}(1)\Big) \\ &= \sqrt{\frac{e^\beta}{2}}\cdot\frac{p_1!\cdots p_r!}{2^r}\frac{1}{(1-z)^{p_1+\cdots+p_r+r+1/2}}(1+\mathrm{o}(1)).\end{aligned}$$

Similarly, as $z \to -1$, we have

$$G_{\theta,p}(z) = \sqrt{\frac{e^\beta}{2}}\cdot\frac{p_1!\cdots p_r!}{2^r}\frac{1}{(1+z)^{p_1+\cdots+p_r+r+1/2}}(1+\mathrm{o}(1)).$$

Now using [11, Theorem VI.5], we obtain

$$[z^{2n}]G_{\theta,p}(z) \sim 2\cdot\sqrt{\frac{e^\beta}{2}}\cdot\frac{p_1!\cdots p_r!}{2^r}\cdot\frac{(2n)^{p_1+\cdots+p_r+r-1/2}}{\Gamma(p_1+\cdots+p_r+r+1/2)}.$$

This completes the proof. ■

6.3. Truncation estimates

Recall that Theorem 6.3 provided an expression for the moment $M_p(\widetilde{L}^{(g,m,\kappa)*})$ which involves a sum which is a truncated version of $S_{\theta,p,n}$ from (6.9). In this section, we show that the difference between $S_{\theta,p,n}$ and its truncation is negligible compared to the asymptotics of Theorem 6.6.

Theorem 6.11. *Let θ and $g_\theta(z)$ be as in Theorem 6.6. Then for any real $\kappa > 1$, we have, as $n \to \infty$,*

$$S_{\theta,p,n} \sim \sum_{k=r}^{\kappa(\log(2n))/2} \frac{1}{(k-r)!} \sum_{\substack{(j_1,\dots,j_k)\in\mathbb{N}^k \\ j_1+\dots+j_k=n}} \prod_{i=1}^{k} \frac{\theta_i}{2j_i} \prod_{i=1}^{r} \frac{(2j_i+p_i)!}{(2j_i-1)!}. \tag{6.11}$$

Bounding the coefficient in a Taylor expansion is a standard tool in asymptotic analysis known as the "Big-O transfer" [11, Theorem VI.3]. However, in our situation we need to bound the n-th Taylor coefficient of a function f_n that depends on n. To do so, we track down the dependencies on the functions inside the transfer theorem.

Lemma 6.12 ([8, Lemma 4.4]). *Let λ and x be positive real numbers. We have*

$$\sum_{k=\lceil x\lambda\rceil}^{\infty} \frac{\lambda^k}{k!} \le \exp(-\lambda(x\log x - x)).$$

Lemma 6.13. *Let $h(z)$ be a holomorphic function on $D(0,R) \setminus ([1,R) \cup (-R,-1])$ such that, as $z \to \pm 1$, we have*

$$h(z) = -\frac{1}{2}\log(1-z^2) + \mathrm{O}(1).$$

Fix a real $\kappa > 1$ and non-negative integers p and q. For $n \ge 1$, let

$$f_n(z) = \frac{1}{(1-z)^p(1+z)^q} \sum_{k=\lfloor \kappa(\log n)/2\rfloor}^{\infty} \frac{h(z)^k}{k!}.$$

Then, we have, as $n \to \infty$,

$$[z^n] f_n(z) = \mathrm{O}(n^{\max\{p,q\}-1-(\kappa\log\kappa-\kappa)/2}).$$

Proof. Let $0 < \eta < R-1$ and $0 < \phi < \pi/2$, and define the contour γ as the union $\sigma_+ \cup \sigma_- \cup \lambda_\nearrow \cup \lambda_\nwarrow \cup \lambda_\searrow \cup \lambda_\swarrow \cup \Sigma_+ \cup \Sigma_-$ with

$$\begin{aligned}
\sigma_+ &= \{z : |z-1| = 1/n,\ |\arg(z-1)| \ge \phi\},\\
\sigma_- &= \{z : |z+1| = 1/n,\ |\arg(z-1)| \le \pi-\phi\},\\
\lambda_\nearrow &= \{z : |z-1| \ge 1/n,\ |z| \le 1+\eta,\ \arg(z-1) = \phi\},\\
\lambda_\nwarrow &= \{z : |z-1| \ge 1/n,\ |z| \le 1+\eta,\ \arg(z-1) = -\phi\},\\
\lambda_\searrow &= \{z : |z+1| \ge 1/n,\ |z| \le 1+\eta,\ \arg(z+1) = \pi-\phi\},
\end{aligned}$$

$$\lambda_{\swarrow} = \{z : |z+1| \geq 1/n,\ |z| \leq 1+\eta,\ \arg(z+1) = -\pi+\phi\},$$
$$\Sigma_{+} = \{z : |z| = 1+\eta,\ \arg(z-1) \geq \phi,\ \arg(z+1) \leq \pi-\phi\},$$
$$\Sigma_{-} = \{z : |z| = 1+\eta,\ \arg(z-1) \leq -\phi,\ \arg(z+1) \geq -\pi+\phi\}.$$

See Figure 1 for a picture of γ. Since f_n is holomorphic on $D(0,R) \smallsetminus ([1,R) \cup [-R,-1))$, we have Cauchy's residue theorem for its coefficients

$$[z^n] f_n(z) = \frac{1}{2\pi i} \int_\gamma \frac{f_n(z)}{z^{n+1}}\, \mathrm{d}z. \tag{6.12}$$

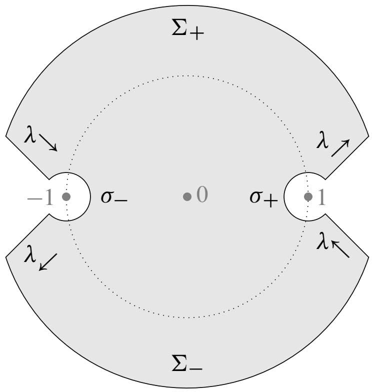


Fig. 1. The contour γ.

Taking absolute values in (6.12), we obtain

$$|[z^n] f_n(z)| \leq \frac{1}{2\pi} \int_\gamma \frac{|\mathrm{d}z|}{|z|^{n+1}} \frac{1}{|1-z|^p |1+z|^q} \sum_{k=\lfloor \kappa(\log n)/2 \rfloor}^{\infty} \frac{|h(z)|^k}{k!}. \tag{6.13}$$

The proof proceeds by analyzing the right-hand side in (6.13) for each piece of the contour γ.

Let us start with the small arc of the circle σ_+. The change of variables $z = 1 - e^{i\theta}/n$ yields

$$\left| \frac{1}{2\pi i} \int_{\sigma_+} \frac{f_n(z)\, \mathrm{d}z}{z^{n+1}} \right| \leq \frac{n^{p-1}}{2\pi} \int_{-\pi+\phi}^{\pi-\phi} \frac{\mathrm{d}\theta}{|1 - e^{i\theta}/n|^{n+1}} \sum_{k=\lfloor \kappa(\log n)/2 \rfloor}^{\infty} \frac{|h(1-e^{i\theta}/n)|^k}{k!}.$$

First $h(1 - e^{i\theta}/n) = \log(n)/2 + \mathrm{O}(1)$ uniformly in θ. Hence, by Lemma 6.12, uniformly in θ, as $n \to \infty$, we have

$$\sum_{k=\lceil \kappa(\log n)/2 \rceil}^{\infty} \frac{|h(z)|^k}{k!} \leq \exp\Big(-(\kappa \log \kappa - \kappa) \cdot \frac{\log n + \mathrm{O}(1)}{2}\Big) = \mathrm{O}(n^{-(\kappa \log \kappa - \kappa)/2}).$$

Since $1/|1-e^{i\theta}/n|^{n+1}$ is uniformly bounded in n,

$$\left|\frac{1}{2\pi i}\int_{\sigma_+}\frac{f_n(z)\,\mathrm{d}z}{z^{n+1}}\right| = \mathrm{O}(n^{p-1-(\kappa\log\kappa-\kappa)/2}).$$

Similarly,

$$\left|\frac{1}{2\pi i}\int_{\sigma_-}\frac{f_n(z)\,\mathrm{d}z}{z^{n+1}}\right| = \mathrm{O}(n^{q-1-(\kappa\log\kappa-\kappa)/2}).$$

Let us now consider the case of $\lambda_{\nearrow}$. Let r be the positive solution of the equation $|1+re^{i\phi}| = 1+\eta$. Perform the change of variable $z = 1+e^{i\phi}\cdot t/n$, we have

$$\left|\frac{1}{2\pi i}\int_{\lambda_{\nearrow}}\frac{f_n(z)}{z^{n+1}}\,\mathrm{d}z\right| \leq \frac{n^{p-1}}{2\pi\cdot(R+1)^q}\int_1^{nr}\mathrm{d}t\cdot t^{-p}\left|1+e^{i\phi}\frac{t}{n}\right|^{-n-1} \times \sum_{k=\lfloor\kappa(\log n)/2\rfloor}^{\infty}\frac{|h(1+e^{i\phi}t/n)|^k}{k!}.$$

For n large enough and uniformly in t, $|h(1+e^{i\phi}t/n)| = \log(n)/2+\mathrm{O}(1)$. Lemma 6.12 gives

$$\sum_{k=\lfloor\kappa(\log n)/2\rfloor}^{\infty}\frac{|h(1+e^{i/\phi}t/n)|^k}{k!} = \mathrm{O}(n^{-(\kappa\log\kappa-\kappa)/2}).$$

From the boundedness of $|1+e^{i\phi}t/n|^{-n-1}$, it follows that

$$\int_1^{nr} t^{-p}\left|1+e^{i\phi}\frac{t}{n}\right|^{-n-1}\mathrm{d}t = \mathrm{O}(n^{-p+1}),$$

and therefore

$$\left|\frac{1}{2\pi i}\int_{\lambda_{\nearrow}}\frac{f_n(z)\,\mathrm{d}z}{z^{n+1}}\right| = \mathrm{O}(n^{-(\kappa\log\kappa-\kappa)/2}).$$

The same estimate is valid for the integral along the other three segments $\lambda_{\nwarrow}$, $\lambda_{\searrow}$, and $\lambda_{\swarrow}$.

For the large demi-circle Σ_+, we have

$$\left|\frac{1}{2\pi i}\int_{\Sigma_+}\frac{f_n(z)\,\mathrm{d}z}{z^{n+1}}\right| \leq \frac{1}{2\pi}\cdot\frac{\max_{z\in\Sigma_+}|f_n(z)|}{(1+\eta)^{n+1}}\cdot 2\pi(1+\eta) = \mathrm{O}((1+\eta)^{-n})$$

which decreases exponentially fast. The integral along Σ_- can be bounded similarly.

We conclude the proof by combining the above estimates. ∎

Proof of Theorem 6.11. Similar to Lemma 6.9, if we write

$$G_{n,\theta,p}(z) := \sum_{k=\lfloor\kappa\log(2n)/2\rfloor}^{\infty}\frac{(g_\theta(z^2)/2)^k}{k!}\prod_{i=1}^{r}D_{p_i}\Big(\frac{1}{2}g_\theta(z^2)\Big),$$

then $[z^{2n}]G_{n,\theta,p}$ is the complement of the partial sum in the right-hand side of (6.11). Following the proof of Theorem 6.6, we obtain, as $z \to 1$,

$$G_{n,\theta,p}(z) = \sum_{k=\lfloor \kappa \log(2n)/2 \rfloor}^{\infty} \frac{(g_\theta(z^2)/2)^k}{k!} \prod_{i=1}^{r} \frac{p_i!}{2}\Big(\frac{1}{(1-z)^{p_i+1}} + \mathrm{O}(1)\Big),$$

where the O(1) is uniform in n (it only depends on $g_\theta(z)$). Applying Lemma 6.13, we obtain

$$[z^{2n}]G_{n,\theta,p}(z) = \mathrm{O}((2n)^{p_1+\cdots+p_r+r-1-(\kappa \log \kappa - \kappa)/2}).$$

For $\kappa > 1$, we have $-1-(\kappa\log\kappa-\kappa)/2 < -1/2$, and the above sum is negligible compared to the asymptotics of the full sum $S_{\theta,p,n}$ from Theorem 6.6. ∎

6.4. *Proof of Theorem* 6.1

By Lemma 6.2, it suffices to prove the convergence of the moments $M_p(\widetilde{L}^{(g,m,\kappa)*})$ for all $p = (p_1, \dots, p_r)$ towards the moments of the GEM(1/2) distribution that were computed in Lemma 5.6.

Now, Theorem 6.3, provides an asymptotic equivalence of $M_p(\widetilde{L}^{(g,m,\kappa)*})$ involving the sum

$$\sum_{k=r}^{\kappa(\log(6g-6))/2} \frac{1}{(k-r)!} \sum_{\substack{(j_1,\dots,j_k)\in\mathbb{N}^k \\ j_1+\cdots+j_k=3g-3}} \prod_{i=1}^{k} \frac{\zeta_m(2j_i)}{2j_i} \prod_{i=1}^{r} \frac{(2j_i+p_i)!}{(2j_i-1)!}.$$

The asymptotics of the above sum were then obtained from Corollary 6.8 and Theorem 6.11. Namely, the above is asymptotically equivalent to

$$\sqrt{\frac{m}{m+1}} \cdot \frac{p_1!\cdots p_r!}{2^{r-1}} \cdot \frac{(6g-6)^{p_1+\cdots+p_r+r-1/2}}{\Gamma(p_1+\cdots+p_r+r+1/2)}.$$

Substituting this value in the formula of Theorem 6.3, we obtain, as $g \to \infty$,

$$M_p(\widetilde{L}^{(g,m,\kappa)*}) \sim \frac{\sqrt{\pi}}{2^r} \cdot \frac{p_1!\cdots p_r!}{\Gamma(p_1+\cdots+p_r+r+1/2)}.$$

The above is the value of the moments M_p of the distribution GEM(1/2) from Lemma 5.6 as $\theta = 1/2$ and $(\theta-1)! = (-1/2)! = \Gamma(1/2) = \sqrt{\pi}$.

Since the convergence of $M_p(\widetilde{L}^{(g,m,\kappa)*})$ holds for all $p = (p_1, \dots, p_r)$, the sequence $\widetilde{L}^{(g,m,\kappa)*}$ converges in distribution towards GEM(1/2).

6.5. *Proof of Theorem* 1.3

The structure of the proof is similar to that of Theorem 1.2. Similar to $U^{(\Gamma,\boldsymbol{m})}$ defined in Section 3, we consider a family of independent variables $\{V^{(\Gamma,\boldsymbol{m})}\}_{(\Gamma,\boldsymbol{m})}$ on [0, 1] indexed by weighted stable graphs, where $V^{(\Gamma,\boldsymbol{m})} := s(U^{(\Gamma,\boldsymbol{m})})$ and $s\colon \Delta_{=1}^{E(\Gamma)} \to [0,1]$ is given by the formula

$$(x_e)_{e\in E(\Gamma)} \mapsto \sum_{\substack{e\in E(\Gamma)\\ m_e=1}} x_e.$$

Then we define

$$P^{(g,m)} := V^{G_{g,m}},$$

where $G_{g,m}$ is the random stable graph given in (3.2). Using an argument similar to one in the proof of Theorem 3.2, we can deduce that

$$\frac{1}{s_X(R,m)} \sum_{\substack{\gamma\in\mathcal{ML}_X(\mathbb{Z})\\ \ell_X(\gamma)\le R\\ \mathrm{mult}(\gamma)\le m}} \delta_{p_X(\gamma)} \xrightarrow[R\to\infty]{} P^{(g,m)},$$

where $p_X(\gamma)$ was defined just above Theorem 1.3. Now if we were able to show

$$\lim_{g\to\infty} \mathbb{E}(P^{(g,m)}) = 1,$$

then the result follows from Markov's inequality. To this end, we fix $\kappa > 1$ and consider the random variable

$$\widetilde{P}^{(g,m,\kappa)} := V^{\widetilde{G}_{g,m,\kappa}},$$

where $\widetilde{G}_{g,m,\kappa}$ is the random stable graph (with one vertex) defined by (4.1). Similar to Theorem 4.1, it follows from Theorem 2.7 and the boundedness of s defined above that

$$\mathbb{E}(P^{(g,m)}) \sim \mathbb{E}(\widetilde{P}^{(g,m,\kappa)}).$$

Thus, it is sufficient to prove

$$\lim_{g\to\infty} \mathbb{E}(\widetilde{P}^{(g,m,\kappa)}) = 1.$$

The computation for $\mathbb{E}(\widetilde{P}^{(g,m,\kappa)})$ is similar to the one carried out in Section 6.1,

$$\mathbb{E}(\widetilde{P}^{(g,m,\kappa)}) = \frac{1}{\widetilde{b}_{g,m,\kappa}} \sum_{k=1}^{\kappa\log(6g-6)/2} \sum_{(\Gamma_{g,k},\boldsymbol{m})} \frac{1}{(6g-6)!} \frac{\mathcal{Y}_{\boldsymbol{m}}(F_{\Gamma_{g,k}})}{|\mathrm{Aut}(\Gamma_{g,k},\boldsymbol{m})|}$$
$$\times \int_{\Delta_{=1}^{E(\Gamma_{g,k})}} \mathrm{d}\lambda_{=1}^{k}\, f_{\Gamma_{g,k},\boldsymbol{m}}((x_e)_{e\in E(\Gamma_{g,k})}) \sum_{\substack{e\in E(\Gamma_{g,k})\\ m_e=1}} x_e$$

which equals

$$\frac{1}{\widetilde{b}_{g,m,\kappa}} \sum_{k=1}^{\kappa\log(6g-6)/2} \frac{w_{g,k}}{(6g-6)!\,|\mathrm{Aut}(\Gamma_{g,k})|} \sum_{\substack{(m_1,\ldots,m_k)\in\mathbb{N}^k\\ m_i\le m\\ i\in\{1,\ldots,k\}}} \sum_{\substack{(j_1,\ldots,j_k)\in\mathbb{N}^k\\ j_1+\cdots+j_k=3g-3}} \widetilde{c}_{g,k}(j_1,\ldots,j_k)$$
$$\times \int_{\Delta_{=1}^{k}} \Big(\sum_{\substack{1\le i\le k\\ m_i=1}} x_i\Big)\Big(\prod_{i=1}^{k} \frac{x_i^{2j_i-1}}{m_i^{2j_i}(2j_i)!}\Big) \mathrm{d}\lambda_{=1}^{k}, \tag{6.14}$$

where $w_{g,k}$ and $\widetilde{c}_{g,k}$ are defined in (6.4) and (6.2), respectively. Now by partitioning $\{(m_1,\ldots,m_k)\in\mathbb{N}^k\}$ into $k+1$ subsets according to the number of i such that $m_i\neq 1$, the second line of (6.14) can be rewritten as

$$\sum_{r=0}^{k}\binom{k}{r}\sum_{\substack{(m_1,\ldots,m_r)\in\mathbb{N}^r\\ 2\le m_1,\ldots,m_r\le m}}\ \sum_{\substack{(j_1,\ldots,j_k)\in\mathbb{N}^k\\ j_1+\cdots+j_k=3g-3}}\widetilde{c}_{g,k}(j_1,\ldots,j_k)$$
$$\times\int_{\Delta^k_{=1}}(x_{r+1}+\cdots+x_k)\prod_{i=1}^{r}\frac{1}{m_i^{2j_i}}\prod_{i=1}^{k}\frac{x_i^{2j_i-1}}{(2j_i)!}\,\mathrm{d}\lambda^k_{=1}$$

which is equal to, by Lemma 5.5,

$$\sum_{r=0}^{k}\binom{k}{r}\frac{k-r}{(6g-6)!}\sum_{\substack{(j_1,\ldots,j_k)\in\mathbb{N}^k\\ j_1+\cdots+j_k=3g-3}}\widetilde{c}_{g,k}(j_1,\ldots,j_k)\prod_{i=1}^{r}(\zeta_m(2j_i)-1)\prod_{i=1}^{k-1}\frac{1}{2j_i}.$$

Hence,

$$\mathbb{E}(\widetilde{P}^{(g,m,\kappa)})=\frac{1}{\widetilde{b}_{g,m,\kappa}}\sum_{k=1}^{\kappa\log(6g-6)/2}\frac{w_{g,k}}{((6g-6)!)^2 2^k}\sum_{r=0}^{k}\frac{1}{r!\,(k-r-1)!}$$
$$\times\sum_{\substack{(j_1,\ldots,j_k)\in\mathbb{N}^k\\ j_1+\cdots+j_k=3g-3}}\widetilde{c}_{g,k}(j_1,\ldots,j_k)\prod_{i=1}^{r}\frac{\zeta_m(2j_i)-1}{2j_i}\prod_{i=r+1}^{k-1}\frac{1}{2j_i},$$

where we have used the fact

$$|\mathrm{Aut}(\Gamma_{g,k})|=2^k k!.$$

It follows from (2.6) and (6.7) that, as $g\to\infty$,

$$\frac{1}{\widetilde{b}_{g,m,\kappa}}\frac{w_{g,k}}{((6g-6)!)^2 2^k}\sim\frac{\sqrt{\pi}}{2}\sqrt{\frac{m+1}{m}}(6g-6)^{-1/2},$$

and hence it is enough to determine the asymptotic behaviour of

$$\sum_{k=1}^{\kappa\log(6g-6)/2}\sum_{r=0}^{k}\frac{1}{r!\,(k-r-1)!}\sum_{\substack{(j_1,\ldots,j_k)\in\mathbb{N}^k\\ j_1+\cdots+j_k=3g-3}}\prod_{i=1}^{r}\frac{\zeta_m(2j_i)-1}{2j_i}\prod_{i=r+1}^{k-1}\frac{1}{2j_i}$$

when $g\to\infty$. To achieve this, we follow a similar approach as outlined in Section 6.2. Define for any $n\ge k$,

$$T_m(n):=\sum_{k=1}^{\infty}\sum_{r=0}^{k}\frac{1}{r!\,(k-r-1)!}\sum_{\substack{(j_1,\ldots,j_k)\in\mathbb{N}^k\\ j_1+\cdots+j_k=n}}\prod_{i=1}^{r}\frac{\zeta_m(2j_i)-1}{2j_i}\prod_{i=r+1}^{k-1}\frac{1}{2j_i}.$$

The sum $T_m(n)$ can be written as

$$\begin{aligned} T_m(n) &= [z^n] \sum_{k=1}^{\infty} \sum_{r=0}^{k} \frac{1}{r!\,(k-r-1)!} \Big(\frac{1}{2} g_m(z) - \frac{1}{2} \log \frac{1}{1-z}\Big)^r \Big(\frac{1}{2} \log \frac{1}{1-z}\Big)^{k-r-1} \frac{z}{1-z} \\ &= [z^n] \frac{z}{1-z} \sum_{k=1}^{\infty} \frac{(g_m(z)/2)^{k-1}}{(k-1)!} = [z^n] \frac{z}{1-z} \exp\Big(\frac{g_m(z)}{2}\Big), \end{aligned}$$

where g_m is defined by (6.10). Now Lemma 6.7 implies that $\exp(g_m(z)/2)$ is holomorphic in $D(0,4) \smallsetminus [1,4)$, and when $z \to 1$ inside $D(0,4) \smallsetminus [1,4)$, we have

$$\exp\Big(\frac{g_m(z)}{2}\Big) = \sqrt{\frac{2m}{m+1}} \frac{1}{\sqrt{1-z}} (1 + \mathrm{O}(1-z)).$$

Therefore, using [11, Theorem VI.4], we obtain

$$T_m(n) = [z^n] \frac{z}{1-z} \exp\Big(\frac{g_m(z)}{2}\Big) \sim \sqrt{\frac{2m}{m+1}} \frac{\sqrt{n}}{\Gamma(3/2)}.$$

To conclude, it suffices to show that when $n \to \infty$,

$$\sum_{k=1}^{\kappa \log(2n)/2} \sum_{r=0}^{k} \frac{1}{r!\,(k-r-1)!} \sum_{\substack{(j_1,\dots,j_k)\in\mathbb{N}^k \\ j_1+\dots+j_k=n}} \prod_{i=1}^{r} \frac{\zeta_m(2j_i)-1}{2j_i} \prod_{i=r+1}^{k-1} \frac{1}{2j_i} \sim T_m(n).$$

This can be done by writing the complement sum

$$\sum_{k=\lfloor \kappa \log(2n)/2 \rfloor}^{\infty} \sum_{r=0}^{k} \frac{1}{r!\,(k-r-1)!} \sum_{\substack{(j_1,\dots,j_k)\in\mathbb{N}^k \\ j_1+\dots+j_k=n}} \prod_{i=1}^{r} \frac{\zeta_m(2j_i)-1}{2j_i} \prod_{i=r+1}^{k-1} \frac{1}{2j_i}$$

as

$$[z^{2n}] \frac{z^2}{1-z^2} \sum_{k=\lfloor \kappa \log(2n)/2 \rfloor}^{\infty} \frac{(g_m(z^2)/2)^{k-1}}{(k-1)!}$$

and then applying Lemma 6.13. This completes the proof of Theorem 1.3.

Acknowledgements. We warmly thank Anton Zorich who encouraged us to join our forces and knowledge from [8, 19] to study the lengths statistics of random multicurves. The second author would like to thank Grégoire Sergeant-Perthuis and Maud Szusterman for helpful conversations about probability theory. We owe a debt of gratitude to the referees who read the paper with great care. Their numerous valuable suggestions helped us to improve the presentation of the paper considerably.

Funding. The work of the first named author is partially supported by the ANR-19-CE40-0003 grant.

References

[1] Aggarwal, A.: Large genus asymptotics for intersection numbers and principal strata volumes of quadratic differentials. Invent. Math. **226**, 897–1010 (2021) Zbl 1480.14020 MR 4337974

[2] Andersen, J. E., Borot, G., Charbonnier, S., Delecroix, V., Giacchetto, A., Lewański, D., Wheeler, C.: Topological recursion for Masur–Veech volumes. J. Lond. Math. Soc. (2) **107**, 254–332 (2023) Zbl 07730992 MR 4535014

[3] Arana-Herrera, F.: Counting square-tiled surfaces with prescribed real and imaginary foliations and connections to Mirzakhani's asymptotics for simple closed hyperbolic geodesics. J. Mod. Dyn. **16**, 81–107 (2020) Zbl 1462.30078 MR 4097720

[4] Arana-Herrera, F.: Counting hyperbolic multigeodesics with respect to the lengths of individual components and asymptotics of Weil–Petersson volumes. Geom. Topol. **26**, 1291–1347 (2022) Zbl 1502.30131 MR 4466649

[5] Arratia, R., Barbour, A. D., Tavaré, S.: Logarithmic combinatorial structures: a probabilistic approach. EMS Monogr. Math., European Mathematical Society (EMS), Zürich (2003) Zbl 1040.60001 MR 2032426

[6] Billingsley, P.: Probability and measure. 3rd ed., Wiley Ser. Probab. Math. Stat., John Wiley & Sons, New York (1995) Zbl 0822.60002 MR 1324786

[7] Delecroix, V., Goujard, E., Zograf, P., Zorich, A.: Masur–Veech volumes, frequencies of simple closed geodesics, and intersection numbers of moduli spaces of curves. Duke Math. J. **170**, 2633–2718 (2021) Zbl 1471.14066 MR 4305379

[8] Delecroix, V., Goujard, E., Zograf, P., Zorich, A.: Large genus asymptotic geometry of random square-tiled surfaces and of random multicurves. Invent. Math. **230**, 123–224 (2022) Zbl 1498.14074 MR 4480147

[9] Donnelly, P., Joyce, P.: Continuity and weak convergence of ranked and size-biased permutations on the infinite simplex. Stochastic Process. Appl. **31**, 89–103 (1989) Zbl 0694.60009 MR 996613

[10] Erlandsson, V., Souto, J.: Mirzakhani's curve counting and geodesic currents. Progr. Math. 345, Birkhäuser, Cham (2022) Zbl 1514.32003 MR 4501245

[11] Flajolet, P., Sedgewick, R.: Analytic combinatorics. Cambridge University Press, Cambridge (2009) Zbl 1165.05001 MR 2483235

[12] Griffiths, R. C.: On the distribution of allele frequencies in a diffusion model. Theoret. Population Biol. **15**, 140–158 (1979) Zbl 0422.92015 MR 528914

[13] Griffiths, R. C.: On the distribution of points in a Poisson Dirichlet process. J. Appl. Probab. **25**, 336–345 (1988) Zbl 0691.92009 MR 938197

[14] Hwang, H.-K.: Théorèmes limites pour les structures combinatoires et les fonctions arithmétiques. Ph.D. thesis, École Polytechnique, http://www.theses.fr/1994EPXX0021, visited on 7 June 2024 (1994)

[15] Hwang, H.-K.: Asymptotics of Poisson approximation to random discrete distributions: an analytic approach. Adv. in Appl. Probab. **31**, 448–491 (1999) Zbl 0945.60001 MR 1724562

[16] Itzykson, C., Zuber, J.-B.: Combinatorics of the modular group. II. The Kontsevich integrals. Internat. J. Modern Phys. A **7**, 5661–5705 (1992) Zbl 0972.14500 MR 1180858

[17] Kingman, J. F. C., Taylor, S. J., Hawkes, A. G., Walker, A. M., Cox, D. R., Smith, A. F. M., Hill, B. M., Burville, P. J., Leonard, T.: Random discrete distribution. J. R. Stat. Soc. Ser. B. Stat. Methodol. **37**, 1–22 (1975) Zbl 0331.62019 MR 368264

[18] Kontsevich, M.: Intersection theory on the moduli space of curves and the matrix Airy function. Comm. Math. Phys. **147**, 1–23 (1992) Zbl 0756.35081 MR 1171758

[19] Liu, M.: Length statistics of random multicurves on closed hyperbolic surfaces. Groups Geom. Dyn. **16**, 437–459 (2022) Zbl 1509.57019 MR 4502612

[20] Mirzakhani, M.: Growth of the number of simple closed geodesics on hyperbolic surfaces. Ann. of Math. (2) **168**, 97–125 (2008) Zbl 1177.37036 MR 2415399

[21] Mirzakhani, M.: Counting mapping class group orbits on hyperbolic surfaces. arXiv:1601.03342 (2016)

[22] Witten, E.: Two-dimensional gravity and intersection theory on moduli space. In: Surveys in differential geometry (Cambridge, MA, 1990), Lehigh University, Bethlehem, PA, 243–310 (1991) Zbl 0757.53049 MR 1144529