Abstract. We study the length statistics of the components of a random multicurve on a surface of genus \( g \geq 2 \). For each fixed genus, the existence of such statistics follows from the work of M. Mirzakhani [Mir], F. Arana-Herrera [A-H] and M. Liu [Liu]. We prove that as the genus \( g \) tends to infinity the statistics converge in law to the Poisson–Dirichlet distribution of parameter \( \theta = 1/2 \). In particular, as the genus tends to infinity the mean length of the three longest components converge respectively to 75.8%, 17.1% and 4.9% of the total length.

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 on \( X \) is a connected closed curve on \( X \), non-homotopic to a point and without self-intersection. In the free homotopy class of a simple closed curve \( \gamma \), there exists a unique geodesic representative with respect to \( X \). 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 \( \gamma \), a component of \( \gamma \) is a maximal family of freely homotopic curves in \( \gamma \). The cardinal 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 if the component). A multicurve is called primitive if all its components have multiplicity one. We denote by \( \ell_X^\downarrow(\gamma) \) the vector of the lengths of each component sorted in decreasing order and by \( \text{mult}(\gamma) \) the multiset of the multiplicities of the components of \( \gamma \), and by \( \text{mult}(\gamma) \) the maximum of \( \text{mult}(\gamma) \). Neither \( \text{mult}(\gamma) \) nor \( \text{mult}(\gamma) \) depend on the hyperbolic structure \( X \). We define \( \ell_X(\gamma) \) as the sum of the entries of \( \ell_X^\downarrow(\gamma) \) and the normalized length vector to be

\[
\hat{\ell}_X^\downarrow(\gamma) := \frac{\ell_X^\downarrow(\gamma)}{\ell_X(\gamma)}.
\]

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

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

\[
\Delta_{k,r} := \{(x_1, x_2, \ldots, x_k) \in [0, \infty)^k : x_1 + x_2 + \cdots + x_k \leq r \}.
\]
Let also define
\[ \Delta_{\leq r}^\infty := \{(x_1, x_2, \ldots) \in [0, \infty)^N : x_1 + x_2 + \cdots \leq r \}. \]

For \( k \leq k' \) we have an injection \( \Delta_{\leq r}^k \to \Delta_{\leq r}^{k'} \) 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 \( \hat{\ell}_X(\gamma) \) is naturally 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 \([0, 1]^N\) endowed with the product topology. This topology coincides with the topology of the inductive limit. When we consider a 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 [A-H] and [Liu].

**Theorem 1.** Let \( g \geq 2 \) and \( m \in \mathbb{N} \cup \{+\infty\} \). There exists a random variable \( L^{g,m} := (L^{g,m}_1, L^{g,m}_2, \ldots) \) on \( \Delta_{\leq 1}^{g-3} \) with the following properties. For any \( g \)-Riemann surface \( X \) of genus \( g \), as \( R \to \infty \) we have the following convergence in distribution
\[ \frac{1}{s_X(R,m)} \sum_{\gamma \in \mathcal{ML}_X(Z)} \delta_{\hat{\ell}_X(\gamma)} \xrightarrow{R \to \infty} L^{g,m}_X \]
where \( \delta_{\hat{\ell}_X(\gamma)} \) is the Dirac mass at the vector \( \hat{\ell}_X(\gamma) \) and \( s_X(R,m) := \#\{\gamma \in \mathcal{ML}_X(Z) : \ell_X(\gamma) \leq R \text{ and } \text{mult}(\gamma) \leq m\} \) is the number of multicurves on \( X \) of length at most \( R \) and multiplicities at most \( m \).

We actually prove a more precise version of the above statement, Theorem 12, in which the law of \( L^{g,m}_X \) 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 \( U_1, U_2, \ldots, \) be i.i.d. random variables with law Beta(1, \( \theta \)) (i.e. they are supported on \([0, 1]\) with density \( \theta(1-x)^{\theta-1} \)). Define the vector
\[ V := (U_1, (1-U_1)U_2, (1-U_1)(1-U_2)U_3, \ldots). \]
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 of parameter \( \theta \) that we denote GEM(\( \theta \)). The Poisson–Dirichlet distribution of parameter \( \theta \), denoted PD(\( \theta \)), 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 1.2.2. The distribution PD(\( \theta \)) appears when considering the Ewens distribution with parameter \( \theta \) on the symmetric group. See Section 1.2.2 below for a more detailed discussion on permutations.

Our main result is the following.

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

The most interesting cases of this convergence are for \( m = 1 \) (primitive multicurves) and \( m = \infty \) (all multicurves). Let us insist that \( L^{g,1} \) and \( L^{g,\infty} \) converge to the same limit as \( g \to \infty \).
All marginals of the Poisson–Dirichlet law can be computed, see for example [ABT03, Section 4.11]. In particular if \( V = (V_1, V_2, \ldots) \sim \text{PD}(\theta) \) then

\[
\mathbb{E}(V_i^n) = \frac{\Gamma(\theta + 1)}{\Gamma(\theta + n)} \int_0^\infty \theta E_1(x)^{(j-1)} \exp(-x) x^{n-1} \, dx
\]

where \( E_1(x) := \int_x^\infty \frac{e^{-y}}{y} \, dy \). The formulas can be turned into a computer program and values were tabulated in [Gri79] [Gri88]. For \( \theta = 1/2 \) we have

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

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 [DGZZ21] and [A-H20] and we refer the readers 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 from 1 to \( N \) and then a square-tiled surface is encoded by two involutions without fixed points \( (\sigma, \tau) \) of \( \{ \pm 1, \pm 2, \ldots, \pm N \} \). More precisely, \( \sigma \) encodes the horizontal gluings: \( +i \) and \( -i \) are respectively the right and left sides of the \( i \)-th squares. The orbits of \( \sigma \) with different signs are glued by translations and the ones with same signs are glued by half-translations. And \( \tau \) encodes the vertical gluings: \( +i \) and \( -i \) are respectively the top and bottom sides of the \( i \)-th squares. The labelling is irrelevant in our definition and two pairs \( (\sigma, \tau) \) and \( (\sigma', \tau') \) encode the same square-tiled surface if there exists a permutation \( \alpha \) of \( \{ \pm 1, \pm 2, \ldots, \pm N \} \) so that \( \alpha(-i) = -\alpha(+i) \), \( \sigma' = \alpha \circ \sigma \circ \alpha^{-1} \) and \( \tau' = \alpha \circ \tau \circ \alpha^{-1} \).

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 \( dz^2 \) (both are being preserved by translations and half-translations). This quadratic form might have simple poles and we denote by \( Q_g(\mathbb{Z}) \) the set of holomorphic square-tiled surfaces of genus \( g \).

A square-tiled surface come equipped with a filling pair of multicurves \( (\gamma_h, \gamma_v) \) coming respectively from the gluings of the horizontal segments \([0,1] \times \{1/2\}\) and vertical segments \( \{1/2\} \times [0,1] \) of each square. Conversely, the dual graph of a filling pair of multicurves in a surface of genus \( g \) defines a square-tiled surface in \( Q_g(\mathbb{Z}) \). Our notation comes from the fact that holomorphic square-tiled surfaces can be seen as integral points in the moduli space of quadratic differentials \( Q_g \). A component of the multicurve \( \gamma_h \) corresponds geometrically to a horizontal cylinder. For a square-tiled surface \( M \) we denote by \( A^i(M) \) the normalized vector of areas of these horizontal cylinders sorted in decreasing order and by \( \text{height}(M) \) the maximum of their heights. Here as in the introduction, normalized mean that we divide by the sum of entries of a vector which coincides with \( \text{area}(M) \). The following is a particular case of [DGZZ21, Theorem 1.29] using the explicit formulas for \( L^{(g,m)} \) given in Theorem 12.

**Theorem 3** ([DGZZ21]). Let \( g \geq 2 \) and \( m \in \mathbb{N} \cup \{+\infty\} \). Let \( L^{(g,m)} \) be the random variable from Theorem 7. Then as \( N \to \infty \) we have the following convergence in distribution

\[
\sum_{\substack{M \in Q_g(\mathbb{Z}) : \\
\text{area}(M) \leq N \text{ \ and} \ \text{height}(M) \leq m}} \delta_{A^i(M)} \to L^{(g,m)}.
\]
An important difference to notice between Theorem 1 and Theorem 3 is that in the former the (hyperbolic) metric $X$ is fixed and we sum over the multicurves $\gamma$ while in the latter we sum over the discrete set of holomorphic square-tiled surfaces $M$.

Using Theorem 3, our Theorem 2 admits the following reformulation.

**Corollary 4.** 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 $\sigma$ in $S_n$ we denote by $K_n(\sigma)$ the number of orbits it has on $\{1, 2, \ldots, n\}$ or equivalently the number of cycles in its disjoint cycle decomposition. The Ewens measure with parameter $\theta$ on $S_n$ is the probability measure defined by

$$P_{n,\theta}(\sigma) := \frac{\theta^{K_n(\sigma)}}{Z_{n,\theta}}$$

where

$$Z_{n,\theta} := \sum_{\sigma \in S_n} \theta^{K_n(\sigma)}.$$

Then under $P_{n,\theta}$, as $n \to \infty$ we have that

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

See for example [ABT03].

By analogy let us denote by $K(g,m)$ the number of non-zero components of $L(g,m)\downarrow$. In [DGZZ], it is proven that $K(g,m)$ behaves as a Poisson distribution with parameter $\frac{\log(g)}{2}$ (by mean 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 $\sigma$ under $P_{g,1/2}$.

Our Theorem 2 provides another connection between $L(g,m)$ and $P_{g,1/2}$. Namely, $L(g,m)\downarrow$ is asymptotically close to the normalized sorted vector of the cycle length of $\sigma$ under $P_{g,1/2}$.

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 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 $g^{-1-\epsilon}$. 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; see Theorem 12 in Section 3. The formula follows from the work of M. Mirzakhani on pants decompositions [Mir08] and the result of F. Arana-Herrera [A-H] and M. Liu [Liu] 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 [DGZZ21].

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 $\tilde{L}(g,m,\kappa)\downarrow$ which, asymptotically, coincides with $L(g,m)\downarrow$. 
See Theorem 17 in Section 4. This reduction is very similar to the one used for the large genus asymptotics of Masur–Veech volumes in [Agg21] and [DGZZ].

The core of our proof consists in proving the convergence of moments of the simpler variable $\tilde{L}(g,m,\kappa)$. We do not use directly $\tilde{L}(g,m,\kappa)$ but its size-biased version $\tilde{L}(g,m,\kappa)^\ast$. The definition of size bias and the link with the Poisson–Dirichlet distribution is explained in Section 5. In Section 6, we show that the moments $\tilde{L}(g,m,\kappa)^\ast$ converge to the moments of GEM(1/2) which is the size-biased version of the Poisson–Dirichlet process PD(1/2), see Theorem 23.

1.4. Acknowledgement. We warmly thank Anton Zorich who encouraged us to join our forces and knowledge from [Liu] and [DGZZ] 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.

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

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 hyperbolic surface $X$ of genus $g$ is a finite multiset of free homotopy classes of disjoint simple closed curves. We denote by $ML_X(Z)$ the set of multicurves on $X$. The homotopy classes that appear in a multicurve $\gamma$ are called components. There are at most $3g-3$ of them. The multiplicity of a component is the number of times it is repeated in $\gamma$, and $\gamma$ is primitive if all the multiplicities are 1.

Let us also recall our notations:
- $\ell_X(\gamma)$: total length of $\gamma$,
- $\ell_X^k(\gamma)$: length vector of the components of $\gamma$,
- $\text{mult}(\gamma)$: maximum multiplicity of component in $\gamma$,
- $\text{mult}(\gamma)$: multiset of multiplicities of components in $\gamma$.

The mapping class group Mod($X$) of $X$ acts on multicurves. We call topological type of a multicurve its equivalence class under the 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 $\gamma$ with components $\gamma_1, \ldots, \gamma_k$ and multiplicities $m_1, \ldots, m_k$ we build a dual graph $\Gamma$ as follows:

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

More formally, a stable graph $\Gamma$ is a 5-tuple $(V,H,\iota,\sigma,\{g_v\}_{v \in V})$ where

- $V$ is a finite set called vertices,
- $H$ is a finite set called half-edges,
- $\iota : H \rightarrow H$ 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,
- $\sigma : H \rightarrow V$ is a surjective map ($\sigma(h)$ is the vertex at which $h$ is rooted),
- $g_v \in \mathbb{Z}_{\geq 0}$,
such that

- (connectedness) for each pair of vertices \( u, v \in V \) there exists a sequence of edges, \([h_1, h'_1], [h_2, h'_2], \ldots, [h_n, h'_n]\) such that \( \sigma(h_1) = u, \sigma(h'_n) = v \) and for \( i \in \{1, \ldots, n - 1\} \) we have \( \sigma(h'_i) = \sigma(h_{i+1}) \),
- (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 \( \Gamma \), its genus is

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

An isomorphism between two stable graphs \( \Gamma = (V, H, \iota, \sigma, g) \) and \( \Gamma' = (V', H', \iota', \sigma', g') \) is a pair of bijections \( \phi : V \to V' \) and \( \psi : H \to H' \) such that

- \( \psi \circ \iota = \iota' \circ \psi \) (in other words, \( \psi \) maps an edge to an edge)
- \( \phi \circ \sigma = \sigma' \circ \psi \)
- for each \( v \in V \), we have \( g'_{\phi(v)} = g_v \).

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

A weighted stable graph is a pair \((\Gamma, m)\) where \( \Gamma \) is a stable graph and \( m \in \mathbb{N}^{E(\Gamma)} \).

An isomorphism between two weighted stable graphs \((\Gamma, m)\) and \((\Gamma', m')\) is an isomorphism \((\phi, \psi)\) between \( \Gamma \) and \( \Gamma' \) such that for each edge \( e \) of \( \Gamma' \) we have \( m_e = m'_{\psi(e)} \) (where we use \( \psi(e) \) to denote \( \{\psi(h), \psi(h')\} \) for the edge \( e = \{h, h'\} \subset H' \)).

We denote by \( \text{Aut}(\Gamma, m) \) the set of automorphisms of the weighted graph \((\Gamma, m)\). 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. \( \psi \)-classes and Kontsevich polynomial

The formula for the random variable \( L^{(g,m)} \) that appears in Theorem 3 involves intersection numbers of \( \psi \)-classes that we introduce now. These rational numbers are famously related to the Witten conjecture [Wit91] proven by Kontsevich [Kon92].

Let \( \overline{\mathcal{M}}_{g,n} \) denote the Deligne–Mumford compactification of 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 \( \psi_i \) 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

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

when \( d_1 + \cdots + d_n = \dim_x \overline{\mathcal{M}}_{g,n} = 3g - 3 + n \). All these intersection numbers are positive rational numbers and can be computed by recursive equations from \( \langle \tau_0^3 \rangle_{0,3} = 1 \) and \( \langle \tau_1 \rangle_{1,1} = \frac{1}{32} \), see for example [ItzZub92].

For our purpose, it is convenient to consider the Kontsevich polynomial \( V_{g,n} \in \mathbb{Q}[x_1, \ldots, x_n] \) that gathers the intersection number into a symmetric polynomial on
\( n \) variables. More precisely,
\[
V_{g,n}(x_1, \ldots, x_n) := \frac{1}{2^{3g-3+n}} \sum_{(d_1, \ldots, d_n) \in \mathbb{Z}^n_+ \atop d_1 + \cdots + d_n = 3g-3+n} \frac{\tau_{d_1} \cdots \tau_{d_n} \gamma, n}{d_1! \cdots d_n!} x_1^{2d_1} \cdots x_n^{2d_n}.
\]

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. M. 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 V. Erlandsson and J. Souto \cite{ES}.

Let \( X \) be a hyperbolic surface of genus \( g \). We define
\[
s_X(R, \gamma) := \# \{ \eta \in \Mod(X) : \gamma \cdot \ell_X(\eta) \leq N \}.
\]

**Theorem 5** (\cite[Theorem 1.1, 1.2 and 5.3]{Mir08}). 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 \( R \to \infty \),
\[
|s_X(R, \gamma)| \sim B(X) \cdot \frac{c(\gamma)}{b_g} \cdot R^{6g-6}
\]
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 \( \ell_X \), and
\[
b_g = \sum_{|\gamma| \in \mathcal{ML}_X(\mathbb{Z})/\Mod(X)} c(\gamma) = \int_{\mathcal{M}_g} B(X) dX.
\]

The above theorem allows to give sense to the notion of a random multicurve. Namely we endow the set of topological types of multicurves \( \mathcal{ML}_X(\mathbb{Z})/\Mod(X) \) with the probability measure which assigns \( c(\gamma)/b_g \) to \( |\gamma| \). We now provide the explicit expression for this probability. For \( \Gamma \in \mathcal{G}_g \) a stable graph we define the polynomial \( F_\Gamma \) on the variables \( \{ x_e \}_{e \in E(\Gamma)} \) by
\[
F_\Gamma \left( \{ x_e \}_{e \in E(\Gamma)} \right) = \prod_{e \in E(\Gamma)} x_e \cdot \prod_{v \in V(\Gamma)} V_{g_{e,v}}(x_v),
\]
where \( x_v \) is the multiset of variables \( x_e \) where \( e \) is an edge adjacent to \( v \) and \( V_{g_{e,v}} \) are the Kontsevich polynomial defined in Section 2.2. In the case \( e \) is a loop based at \( v \), the variable \( x_e \) is repeated twice in \( x_v \).

**Remark 6.** The polynomial \( F_\Gamma \) appeared first in Mirzakhani’s work \cite{Mir08}, see in particular Theorem 5.3. They were related to square-tiled surfaces and Masur–Veech volumes in \cite{DGZZ21} though with a different normalization. Namely, the polynomial \( P_\Gamma \) from \cite{DGZZ21} is related to \( F_\Gamma \) by
\[
P_\Gamma = 2^{4g-2} \cdot \frac{(4g-4)!}{(6g-7)!} \cdot \frac{1}{|\Aut(\Gamma)|} \cdot F_\Gamma.
\]
| multicurve | stable graph | polynomial $F_\Gamma$ |
|------------|--------------|------------------------|
| ![Multicurve 1](image1) | $a \circ 1$ | $x_a V_{1,2}(x_a, x_a)$ = $\frac{1}{18} x_a^5$ |
| ![Multicurve 2](image2) | $1 \circ a \circ 1$ | $x_a V_{1,1}(x_a) V_{1,1}(x_a)$ = $\frac{1}{2} x_a^5$ |
| ![Multicurve 3](image3) | $a \circ 0 \circ b$ | $x_a x_b V_{0,4}(x_a, x_a, x_b, x_b)$ = $x_a^3 x_b + x_a x_b^3$ |
| ![Multicurve 4](image4) | $a \circ 0 \circ b \circ 1$ | $x_a x_b V_{0,3}(x_a, x_a, x_b) V_{1,1}(x_b)$ = $\frac{1}{18} x_a x_b^3$ |
| ![Multicurve 5](image5) | $a \circ 0 \circ b \circ 0 \circ c$ | $x_a x_b x_c V_{0,3}(x_a, x_b, x_c) V_{0,3}(x_a, x_b, x_c)$ = $x_a x_b x_c$ |
| ![Multicurve 6](image6) | $0 \circ a \circ b \circ c$ | $x_a x_b x_c V_{0,3}(x_a, x_b, x_c) V_{0,3}(x_a, x_b, x_c)$ = $x_a x_b x_c$ |

**Figure 1.** The list of topological types of primitive multicurves in genus 2, their associated stable graphs and their corresponding polynomial $F_\Gamma$. The labels on edges are used as variable indices in $F_\Gamma$.

The normalization of $F_\Gamma$ is identical to the conventions used in [ABCDGLW] and simplifies the computations of the present article.

Following [DGZZ21], for a weighted stable graph $(\Gamma, m)$ and we denote by $Y_m: \mathbb{Q}[\{x_e\}_{e \in E(\Gamma)}] \to \mathbb{Q}$ the linear operator defined on monomials

$$Y_m(x_1^{n_1} x_2^{n_2} \ldots x_k^{n_k}) := \frac{n_1! n_2! \ldots n_k!}{m_1^{n_1+1} m_2^{n_2+1} \ldots m_k^{n_k+1}}$$

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

$$Z_m := \sum_{m \in \mathbb{N}^{\infty}(\Gamma) \atop \forall e \in E(\Gamma), m_e \leq m} Y_m.$$

We derive the following directly from [DGZZ21]:

**Theorem 7.** Let $\gamma$ be a multicurve in genus $g$ and $(\Gamma, m)$ the dual weighted stable graph. Then

$$c(\gamma) = \frac{1}{(6g-6)!} \frac{1}{|\text{Aut}(\Gamma, m)|} Y_m(F_\Gamma).$$

Furthermore

$$b_g = \frac{1}{(6g-6)!} \sum_{F \in \mathcal{G}_g} \frac{1}{|\text{Aut}(F)|} Z(F_\Gamma).$$

**Remark 8.** In Theorem 7 we fix a misconception in [DGZZ21] about automorphisms of multicurves (or equivalently weighted stable graph). Indeed, the way we defined
automorphisms of stable graphs and weighted stable graphs in Section 2.1 make it 
so that the following formula is valid
\[ \sum_{\Gamma} \frac{1}{|\text{Aut}(\Gamma)|} \mathcal{Z}(F_{\Gamma}) = \sum_{(\Gamma, m)} \frac{1}{|\text{Aut}(\Gamma, m)|} \mathcal{Y}_m(F_{\Gamma}) \]
where the sums are taken over isomorphism classes of respectively stable graphs of 
genus g and weighted stable graphs of genus g.

**Proof.** Up to the correction of Remark 8 this is exactly [DGZZ21, Theorem 1.22] (see Remark 6 for the difference between \( P_{\Gamma} \) and \( F_{\Gamma} \)).

| stable graph \( \Gamma' \) | value of \( \mathcal{Y}_m(\Gamma) \) |
|--------------------------|------------------|
| \( a(1) \)               | \( 5 \cdot \frac{1}{2} \cdot m_a^5 \) |
| \( 1 \cdot a(1) \)       | \( \frac{5}{96} \cdot \frac{1}{m_a^6} \) |
| \( a(0b) \)              | \( 24 \left( \frac{1}{m_a^4 \cdot m_b^2} + \frac{1}{m_b^2 \cdot m_c^2} \right) \) |
| \( a(0b1) \)             | \( \frac{1}{2} \cdot \frac{1}{m_a^2 \cdot m_b^2} \) |
| \( a(0b0) \)             | \( \frac{1}{m_a^2 \cdot m_b^2 \cdot m_c^2} \) |
| \( 0(0b) \)              | \( \frac{1}{m_a^2 \cdot m_b^2 \cdot m_c^2} \) |

**Figure 2.** The list of topological types of primitive multicurves in genus 2 and the associated values \( \mathcal{Y}_m(\Gamma) \) that is proportional to \( c(\gamma) \) (see Theorem 7).

2.4. Asymptotics of \( \psi \)-correlators and \( b_g \). Our proof of Theorem 2 uses crucially the asymptotics of \( \psi \)-intersections and Masur–Veech volumes from [Agg21] and further developed in [DGZZZ].

**Theorem 9 (Asg21).** For \( g, n \in \mathbb{N} \) and \( d = (d_1, \ldots, d_n) \in \mathbb{Z}_{\geq 0}^n \) with \( d_1 + \cdots + d_n = 3g - 3 + n \), let \( \epsilon(d) \) be defined by
\[
\langle \tau_{d_1}, \ldots, \tau_{d_n} \rangle_{g,n} = \frac{(6g - 5 + 2n)!!}{(2d_1 + 1)!! \cdots (2d_n + 1)!!} \cdot \frac{1}{g! \cdot 24^g} \cdot (1 + \epsilon(d)).
\]
Then
\[
\lim_{g \to \infty} \sup_{n < \sqrt{g/800}} \sup_{d=(d_1,\ldots,d_n)} \sup_{d_1+\cdots+d_n=3g-3+n} \epsilon(d) = 0.
\]

For \( m \in \mathbb{N} \cup \{+\infty\} \) we define
\[
\mathcal{Z}_m(\Gamma) := \frac{1}{(6g - 6)!} \sum_{\Gamma \in \mathcal{G}_g} \frac{\mathcal{Z}_m(\Gamma)}{|\text{Aut}(\Gamma)|}.
\]
Note that \( b_g = b_{g, +\infty} \).

Remark 10. We warn the reader that the constant denoted \( b_{g, m} \) 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 \( b_{g, n} \) in [Mir] and [DGZZ21].

For \( m \in \mathbb{N} \cup \{+\infty\} \) and a real number \( \kappa > 1 \) we also define

\[
\hat{b}_{g, m, \kappa} := \frac{1}{(6g - 6)!} \sum_{\Gamma \in G_g} \frac{1}{|\text{Aut}(\Gamma)|} \sum_{x \in [\mathbb{N} \cap E(\Gamma)]} \mathcal{Y}_m(F_{\Gamma}).
\]

As we have less terms in its definition, \( \hat{b}_{g, m, \kappa} \) is nothing to do with the analogue of \( b_{g, m} \).

We will use the asymptotic results of [Agg21] and [DGZZ] in the following form.

**Theorem 11 (Agg21, DGZZ).** Let \( \Gamma \) be a stable graph and let \( \kappa > 1 \). Then as \( g \to \infty \) we have

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

3. LENGTH VECTORS OF RANDOM MULTICURVES

The aim of this section is to state and prove a refinement of Theorem 1 that provides an explicit description of the random variable \( L_{(g^\gamma, m)} \). For each weighted stable graph \((\Gamma, m)\) we define a random variable \( L_{(\Gamma, m)} \). We then explain how \( L_{(g^\gamma, m)} \) is obtained from them.

Let \( \Gamma \) be a stable graph and let \( k \geq |E(\Gamma)| \). For each injection \( \iota: E(\Gamma) \to \{1, 2, \ldots, k\} \), we define an injection \( g_{\Gamma, \iota}: \mathbb{R}^{|E(\Gamma)|} \to \mathbb{R}^k \) by

\[
g_{\Gamma, \iota}(\{x_e\})_{e \in E(\Gamma)} = (y_1, y_2, \ldots, y_k), \quad \text{where } y_i = \begin{cases} x_{\iota^{-1}(i)} & \text{if } i \in \iota(E(\Gamma)) \\ 0 & \text{otherwise} \end{cases}
\]

Given a measure \( \mu \) on \( \mathbb{R}^{|E(\Gamma)|} \) we define its \( k \)-th symmetrization to be the measure on \( \mathbb{R}^k \) given by

\[
G_{\Gamma, k}(\mu) = \frac{(k - |E(\Gamma)|)!}{k!} \sum_{\iota} (g_{\Gamma, \iota})^*(\mu).
\]

The \( k \)-th symmetrization is supported on the subspaces of dimension \( |E(\Gamma)| \) generated by basis vectors. Because of the coefficient \( (k - |E(\Gamma)|)!)/k! \), the total weights of the measures \( \mu \) and \( (g_{\Gamma, \iota})^*(\mu) \) are the same.

We prove the following refinement of Theorem 4.

**Theorem 12.** Let \( L_{(g, m)} \) be the random variable on \( \Delta_{\leq 1}^{3g - 3} \) with density

\[
\frac{1}{(6g - 6)! \cdot b_{g, m}} \sum_{\Gamma \in G_g} \frac{1}{|\text{Aut}(\Gamma)|} \sum_{x \in [\mathbb{N} \cap E(\Gamma)]} G_{\Gamma, 3g - 3}(\mu_{\Gamma, m})
\]

where \( b_{g, m} \) is defined in (3) and \( \mu_{\Gamma, m} \) is the measure on \( \Delta_{\leq 1}^{E(\Gamma)} \) with density

\[
F_{\Gamma} \left( \left\{ \frac{x_e}{m_e} \right\}_{e \in E(\Gamma)} \right) \cdot \prod_{e \in E(\Gamma)} \frac{1}{m_e}
\]

where \( F_{\Gamma} \) is the polynomial defined in (2). Then

\[
\frac{1}{s_X(R, m)} \sum_{\gamma \in \mathcal{L}_{X}^{m} \cap (2)} \delta_{\ell_X}(\gamma) \xrightarrow{R \to \infty} L_{(g, m)}
\]
where \( s_X(R,m) \) := \#\{ \gamma \in \mathcal{M}_X : \ell_X(\gamma) \leq R \text{ and } \text{mult}(\gamma) \leq m \} \) is the number of multicurves on \( R \) of length at most \( R \) and multiplicity at most \( m \) and \( L^{(g,m)} \) is the vector \( L^{(g,m)} \) sorted in decreasing order.

The study of the length vector of multicurves of a given topological type was initiated by M. Mirzakhani in [Mir]. She studied the special case of maximal multicurve corresponding to a pants decomposition. The general case that we present now was proved independently in [A-H] and [Liu].

**Theorem 13 ([A-H], [Liu]).** Let \( X \) be a hyperbolic surface and \( \gamma \) a multicurve on \( X \) with \( k \) components. Let \( (\Gamma, m) \) be a weighted stable graph dual to \( \gamma \). Let \( U^{(\Gamma,m)} \) be the random variable on \( \Delta_{\leq 1}^k \) with density

\[
(8) \quad \frac{(6g-7)!}{\mathcal{Y}_m(\Gamma) \ m_1 \cdots m_k} \cdot F_{\Gamma} \left( \frac{x_1}{m_1}, \ldots, \frac{x_k}{m_k} \right)
\]

where \( E(\Gamma) \) and \( V(\Gamma) \) are the set of edges and the set of vertices of \( \Gamma \), respectively. Then we have the convergence in distribution

\[
\frac{1}{s_X(R, \gamma)} \sum_{\ell \in \text{Mod}(X \gamma)} \delta_{\ell_X(\eta)} \xrightarrow{R \to \infty} U^{(\Gamma,m)}_{\leq 1}
\]

where \( U^{(\Gamma,m)}_{\leq 1} \) is the sorted version of \( U^{(\Gamma,m)} \) and \( s_X(R, \gamma) \) is defined in [1].

We endow \( \Delta_{\leq r}^{k} \) with the restriction of the Lebesgue measure on \( \mathbb{R}^k \) that we denote by \( \lambda_{\leq r}^{k} \). We define the slice \( \Delta_{\leq r}^{k} \) inside \( \Delta_{\leq r}^{k} \) as

\[
\Delta_{\leq r}^{k} := \{(x_1, x_2, \ldots, x_k) \in [0, \infty)^k : x_1 + x_2 + \cdots + x_k = r\}
\]

and its infinite counterpart

\[
\Delta_{\leq r}^{\infty} := \{(x_1, x_2, \ldots) \in [0, \infty)^{\mathbb{N}} : x_1 + x_2 + \cdots = r\}.
\]

Let us mention that \( \Delta_{\leq r}^{\infty} \) is closed for the product topology in \([0, r)^{\mathbb{N}}\) and hence compact. However \( \Delta_{\leq r}^{\infty} \) is dense in \( \Delta_{\leq r}^{\infty} \). For this reason, it is more convenient to work with measures on \( \Delta_{\leq r}^{\infty} \) even though they are ultimately supported on \( \Delta_{\leq r}^{\infty} \).

On \( \Delta_{\leq r}^{k} \) which is contained in a hyperplane in \( \mathbb{R}^k \) we consider the Lebesgue measure induced by any choice of \( k-1 \) coordinates among \( x_1, \ldots, x_k \). The latter measure is well defined since the change of variables between different choices has determinant \( \pm 1 \). We first start with an elementary integration lemma.

**Lemma 14.** Let \( d_1, \ldots, d_k \in \mathbb{R}_{\geq 0} \). Then

\[
\int_{\Delta_{\leq r}^{k}} x_1^{d_1} x_2^{d_2} \cdots x_k^{d_k} d\lambda_{\leq r}^{k} = \frac{d_1! \cdots d_k!}{(d_1 + \cdots + d_k + k)!} \cdot x_1^{d_1+\cdots+d_k+k}
\]

and

\[
\int_{\Delta_{\leq r}^{k}} x_1^{d_1} x_2^{d_2} \cdots x_k^{d_k} d\lambda_{=1}^{k} = \frac{d_1! \cdots d_k!}{(d_1 + \cdots + d_k + k + 1)!} \cdot x_1^{d_1+\cdots+d_k+k+1}.
\]

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

**Remark 15.** Using Lemma 14 let us check that (6) and (8) are indeed densities of probability measures. From the second equation in the statement of Lemma 14 it follows that the total mass of (6) is \( \mathcal{Y}_m(\Gamma) \), namely

\[
\mathcal{Y}_m(\Gamma) = (6g-7)! \int_{\Delta_{=1}^{k}} \frac{1}{m_1 \cdots m_k} F_{\Gamma} \left( \frac{x_1}{m_1}, \ldots, \frac{x_k}{m_k} \right) d\lambda_{=1}^{k}(x_1, \ldots, x_k).
\]

Indeed, each monomial that appears in \( F_{\Gamma} \) has \( k \) variables and total degree \( 6g-6-k \). Hence the denominator coming from the formula of Lemma 14 compensates the
This implies that as \( g \to \infty \) loops. To simplify the notation we fix a bijection between the edges of \( \Gamma \) and \( \gamma \). Using Theorem 17.

\[ Y(10) = \frac{1}{2} \sum_{\text{non-separating multicurves}} \phi \]

we show that we can restrict to a sum involving only \( O \) terms. Using Theorem 11 from \([Agg21, DGZZ]\) Remark 16.

\[ F \text{ in terms of the stable graph polynomials } ML \]

density in \( \gamma \). Theorem 5 of M. Mirzakhani, for any multicurve \( \gamma \) coming from Theorem 13 that F. Arana-Herrera and M. Liu gave. From Proof of Theorem 12.

We just have to gather the different contributions of each \( \ell \) term. The two equations in the statement then follow by induction.

Proof of Lemma 14.

For \( \alpha, \beta \geq 1 \) integral, we have the following scaling of the beta function

\[ \int_0^\infty t^{\alpha-1} (x-t)^{\beta-1} \, dt = \frac{\Gamma(\alpha) \Gamma(\beta)}{\Gamma(\alpha + \beta)} x^{\alpha + \beta - 1}. \]

This implies that

\[ \int_{\Delta^g_{\leq r}} \frac{x_{1}}{d_1!} \cdots \frac{x_{k}}{d_k!} \, d\lambda^{k}_{\leq r} = \int_{\Delta^g_{\leq r}} \frac{x_{1}}{d_1!} \cdots \frac{x_{k}}{d_k!} \, d\lambda^{k}_{\leq r} \]

and

\[ \int_{\Delta^g_{\leq r}} \frac{x_{1}}{d_1!} \cdots \frac{x_{k}}{d_k!} \, d\lambda^{k}_{\leq r} = \int_{\Delta^g_{\leq r}} \frac{x_{1}}{d_1!} \cdots \frac{x_{k}}{d_k!} \, d\lambda^{k}_{\leq r}. \]

The two equations in the statement then follow by induction.

Proof of Theorem 11.

We just have to gather the different contributions of each multicurve coming from Theorem 13 that F. Arana-Herrera and M. Liu gave. From Theorem 5 of M. Mirzakhani, for any multicurve \( \gamma \) in \( \mathcal{ML}_X(\mathbb{Z}) \), its asymptotic density in \( \mathcal{ML}_X(\mathbb{Z}) \) is \( \frac{c(\gamma)}{b_{g,m}} \). Now Theorem 17 provides the values of \( c(\gamma) \) and \( b_{g,m} \) in terms of the stable graph polynomials \( F_T \).

Remark 16. In [Lim], the length vector \( \ell_X(\gamma) \) for a multicurve \( \gamma = m_1 \gamma_1 + \cdots + m_k \gamma_k \) is defined as \( (\ell_X(\gamma_1), \ldots, \ell_X(\gamma_k)) \) instead of \( (\ell_X(m_1 \gamma_1), \ldots, \ell_X(m_k \gamma_k)) \) as in this paper, and the limit distribution is thus supported on the (non-standard) simplex

\[ \{(x_1, \ldots, x_k) \in \mathbb{R}^k : x_i \geq 0, m_1 x_1 + \cdots + m_k x_k = 1\}. \]

As a consequence, our distribution is the push-forward of the distribution in [Lim] under the map \((x_1, \ldots, x_k) \mapsto (m_1 x_1, \ldots, m_k x_k)\).

4. Reduction in the asymptotic regime

The random variable \( L(g,m) \) appearing in Theorem 12 is delicate to study because it involves a huge number of terms. Using Theorem 11 from [Agg21] and [DGZZ] we show that we can restrict to a sum involving only \( O(\log(g)) \) terms associated to non-separating multicurves.

We denote by \( \Gamma_{g,k} \) the stable graph of genus \( g \) with a vertex of genus \( g-k \) loops. To simplify the notation we fix a bijection between the edges of \( \Gamma_{g,k} \) and \( \{1, 2, \ldots, k\} \) so that \( F_{\Gamma_{g,k}} \) is 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.

Using the same notation as in Theorem 12 we have the following result.

Theorem 17. For \( m \in \mathbb{N} \cup \{+\infty\} \) and \( \kappa > 1 \), let \( \tilde{L}(g,m,\kappa) \) be the random variable on \( \Delta^g_{\leq 3} \) with density

\[ \frac{1}{(6g-6) \cdot b_{g,m,\kappa}} \sum_{\Gamma_{g,k} \in \text{Aut} \Gamma_{g,k}} \frac{1}{|\Gamma_{g,k}|} \sum_{m \in \mathbb{N}^{+k}} G_{\Gamma_{g,k}, 3g-3}(\mu_{\Gamma_{g,k}, m}) \]

where \( b_{g,m,\kappa} \) is defined in [1]. Then for any function \( h \in L^\infty(\Delta^g_{\leq 1}) \) we have

\[ \mathbb{E}(h(L(g,m))) \sim \mathbb{E}(h(\tilde{L}(g,m,\kappa))) \]

as \( g \to \infty \).
Note that terms appearing in the sum \( \sum_{\mathcal{I}} \) in Theorem 10 form a subset of the terms in the sum \( \sum_{\mathcal{J}} \) in Theorem 12.

Proof. By Theorem 11 a random multicurve of high genus is almost surely non-separating with less than \( \kappa \frac{\log(6g-6)}{2} \) edges. As \( h \) is bounded, we obtain the result. \( \square \)

5. Size-biased sampling and Poisson–Dirichlet distribution

5.1. Size-biased reordering. The components of a multicurve are not ordered in any natural way. In Theorem 1 we solve this issue by defining a symmetric random variable \( L^{(g,m)} \) on \( \Delta_{g,1}^{k-5} \) and making the convergence happen towards \( L^{(g,m)} \).whose entries are sorted in decreasing order. In this section we introduce another natural way of ordering the entries: the size-biased ordering. Contrarily to the symmetrization or the decreasing order, it is a random ordering. The size-biased ordering turns out to be convenient in the proof of Theorem 1.

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 \( \sigma \) is a random permutation in \( S_k \). We aim to define the size-biased reordering \( x^{*} = (x_{\sigma(1)}, x_{\sigma(2)}, \ldots, x_{\sigma(k)}) \) of \( x \).

The idea under the size-biased reordering is to pick components according to their values. One can define the random permutation \( \sigma \) inductively as follows. If \( x \) is the zero vector, then \( x^{*} = x_{\sigma} \) where \( \sigma \) is taken uniformly at random. Otherwise, we set \( \sigma(1) \) according to

\[
P_x(\sigma(1) = i) := \frac{x_i}{x_1 + \cdots + x_k}
\]

and define a new vector \( y = (x_1, \ldots, \hat{x}_{\sigma(1)}, \ldots, x_k) \) on \( \Delta_{k}^{\leq 1} \) which is the vector \( x \) with the component \( x_{\sigma(1)} \) removed. In order to keep track of the components we denote \( \phi : \{1, 2, \ldots, k - 1\} \rightarrow \{1, 2, \ldots, k\} \) the unique increasing injection such that its image avoids \( \sigma(1) \). In other words

\[
\phi(i) := \begin{cases} 
  i & \text{if } 1 \leq i < \sigma(1) \\
  i + 1 & \text{if } \sigma(1) \leq i \leq k - 1.
\end{cases}
\]

Assuming that by induction \( y \) has a size-biased reordering \( \sigma_y \) we define for \( i \in \{1, 2, \ldots, k - 1\} \) the other values by \( \sigma(\phi(i)) := \phi(\sigma_y(i)) \). This defines inductively the size-biased reordering.

A more direct definition can be given as follows. Given \( r \) such that at least \( r \) components of \( x \) are positive, for \( 1 \leq i_1, \ldots, i_r \leq k \) distinct integers, we have

\[
P_x(\sigma(1) = i_1, \ldots, \sigma(r) = i_r) = \frac{x_{i_1} \cdots x_{i_r}}{s(s - x_{i_1})(s - x_{i_1} - x_{i_2}) \cdots (s - x_{i_1} - \cdots - x_{i_{r-1}})}
\]

where \( s = x_1 + x_2 + \cdots + x_k \). Note that for \( r = k \) we have \( x_r = 1 - x_1 - \cdots - x_{r-1} \) and one can perform a simplification of the last terms in the numerator and denominator.

Now let \( X : \Omega \rightarrow \Delta_{k}^{\leq 1} \) be a random variable. In order to define its size-biased reordering \( X^{*} : \Omega \rightarrow \Delta_{k}^{\leq 1} \), we consider for each \( x \in \Delta_{k}^{\leq 1} \) independent random variables \( \sigma_x \) distributed according to \( P_x \) 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 18. Let $X$ a random variable on $\Delta_{\geq 1}^k$ with density $f_X: \Delta_{\geq 1}^k \to \mathbb{R}$. Let $1 \leq r \leq k$. Then the $r$-th marginal of the size-biased reordering of $X$, that is to say the density of the vector $(X_1^r, \ldots, X_r^r)$ is

$$f(X_1^r, \ldots, X_r^r)(x_1, \ldots, x_r) = \frac{1}{(k-r)!} \frac{x_1 \cdots x_r}{(1-x_1) \cdots (1-x_1 \cdots x_{r-1})},$$

$$\cdot \int_{\Delta_{\geq 1-x_1}^{k-r} \cdot \cdots \cdot \Delta_{\geq 1-x_{r-1}}^{k-r}} f(x_{\tau(1)}, \ldots, x_{\tau(k)}) \, d\lambda_{x_1}^{k-r}(x_{r+1}, \ldots, x_k),$$

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 (11) 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

$$f(X_1^r, \ldots, X_r^r)(x_1, \ldots, x_r) = \int_{\Delta_{\geq 1-x}^{k-r}} f_X(x_1, \ldots, x_k) \, d\lambda_{x_1}^{k-r}(x_{r+1}, \ldots, x_k)$$

where $s = x_1 + \cdots + x_r$. For a permutation $\tau \in S_{k-r}$ we define the subsimplex

$$\Delta_{s=1-r, \tau}^{k-r} := \{(x_1, \ldots, x_{k-r}) \in \Delta_{\geq 1-s}^{k-r} : x_{\tau(1)} > x_{\tau(2)} > \cdots > x_{\tau(k-r)}\}.$$

We can decompose the integral (12) as a sum over these subsimplices

$$f(X_1^r, \ldots, X_r^r)(x_1, \ldots, x_r) = \sum_{\tau \in S_{k-r}} \int_{\Delta_{\geq 1-x_1}^{k-r} \cdot \cdots \cdot \Delta_{\geq 1-x_{r-1}}^{k-r}} x_{r+1} \cdots x_k \, d\lambda_{x_1}^{k-r}(x_{r+1}, \ldots, x_k) g(x_1, \ldots, x_k) \frac{x_1 \cdots x_r}{(1-x_1) \cdots (1-x_1 \cdots x_{r-1})}.$$
5.2. Poisson–Dirichlet and GEM 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 [K.1975]. We refer to [ABT2013 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 the two distributions.

Theorem 19 ([DJ89]). Let \( X = (X_1, X_2, \ldots) \) be a random variable on \( \Delta_{\leq 1}^\infty \). Let \( \theta > 0 \). Then the sorted reordering \( X^* \) has distribution PD(θ) if and only if the size-biased reordering \( X^* \) has distribution GEM(θ).

We will use the above result in the following form.

Corollary 20 ([DJ89]). Let \( X^{(n)} \) be a sequence of random variables on \( \Delta_{\leq 1}^\infty \). Let \( \theta > 0 \). Then the sorted sequence \( X^{(n)}_{\geq 1} \) converges in distribution to PD(θ) if and only if the size-biased sequence \( X^{(n)*}_{\geq 1} \) converges in distribution to GEM(θ).

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

Proposition 21 ([DJ89]). Let \( X = (X_1, X_2, \ldots) \) be a random variable with distribution GEM(θ). Then the distribution the \( r \)-first components \( (X_1, X_2, \ldots, X_r) \) of \( X \) supported on \( \Delta_{\leq 1}^r \) admit a distribution with density given by

\[
\frac{\theta^r (1 - x_1 - \cdots - x_r)^{\theta - 1}}{(1 - x_1)(1 - x_1 - x_2) \cdots (1 - x_1 - \cdots - x_{r-1})}.
\]

In order to simplify computations, we consider moments of the GEM distribution that get rid of the denominator in the density (13). Namely, for a random variable \( g \) tends to \( \infty \), Lemma 22. These moments of GEM(θ) are as follows.

Lemma 22. If \( X = (X_1, X_2, \ldots) \sim \text{GEM}(\theta) \) and \( (p_1, \ldots, p_r) \) is a non-negative \( r \)-tuple of non-negative integers we define

\[
M_p(X) := \mathbb{E}((1 - X_1) \cdots (1 - X_1 - \cdots - X_{r-1}) \cdot X_1^{p_1} \cdots X_r^{p_r})
\]

These moments of GEM(θ) are as follows.

\[
M_p(X) = \frac{\theta^r \cdot (\theta - 1)! \cdot p_1! \cdots p_r!}{\Gamma(p_1 + \cdots + p_r + \theta + r)!}
\]

Proof. By Proposition 21 we have

\[
M_p(X) = \int_{\Delta_{\leq 1}^r} \theta^r x_1^{p_1} \cdots x_r^{p_r} (1 - x_1 - \cdots - x_r)^{\theta - 1} dx_{\leq 1}
\]

\[
= \theta^r \int_{\Delta_{\leq 1}^{r+1}} x_1^{p_1} \cdots x_r^{p_r} x_r^{\theta - 1} d\lambda_{=1}^{r+1}.
\]

The last term is an instance of Lemma 14 on the simplex \( \Delta_{=1}^{r+1} \). Replacing the value obtained from the integration lemma gives the result. \( \square \)

6. PROOF OF THE MAIN THEOREM

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

Theorem 23. For \( g \geq 2 \) integral, \( m \in \mathbb{N} \cup \{+\infty\} \) and \( \kappa > 1 \) real, let \( \tilde{L}^{(g,m,\kappa)*}_{\leq 1} \) be the size-biased version of the random variable \( \tilde{L}^{(g,m,\kappa)}_{\leq 1} \) from Theorem 17. Then as \( g \) tends to \( \infty \), \( \tilde{L}^{(g,m,\kappa)*}_{\leq 1} \) converges in distribution to GEM(1/2).
Let us first show how to derive our main Theorem 2 from Theorem 23.

Proof of Theorem 2. By Theorem 17, the random variables \( L(g, m, \kappa)^\ast \) and \( L(g, m)^\ast \) have the same limit distribution as \( g \to +\infty \). Hence by Theorem 23, the random variable \( L(g, m)^\ast \) converges in distribution towards \( \text{GEM}(1/2) \).

Finally Corollary 20 shows that the convergence in distribution of \( L(g, m)^\ast \) towards \( \text{GEM}(1/2) \) is equivalent to the convergence of \( L(g, m)^\ast \) towards \( \text{PD}(1/2) \). This concludes the proof of Theorem 2. \( \square \)

6.1. Moment’s method. Let us recall from Section 5.2 Equation (14) that we defined some specific moments \( M_{(p_1, \ldots, p_r)}(X) \) for a random variable \( X \) on \( \Delta_{\kappa}^\infty \). 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 [Bill95, Section 30] for the case of real variables.

Lemma 24. A sequence of random variables \( X^{(n)} = (X_1^{(n)}, X_2^{(n)}, \ldots) \) in \( \Delta_{\kappa}^\infty \) converges in distribution to a random variable \( X^{(\infty)} \) in \( \Delta_{\kappa}^\infty \) if and only if for all \( p = (p_1, \ldots, p_r) \) vector of non-negative integers we have \( \lim_{n \to \infty} M_p(X^{(n)}) = M_p(X^{(\infty)}) \).

Proof. The infinite-dimensional cube \([0, 1]^N\) is compact with respect to the product topology by Tychnoff’s theorem. The set \( \Delta_{\kappa}^\infty \) is a closed subset of \([0, 1]^N\), and is therefore compact. The signed measures on \( \Delta_{\kappa}^\infty \) are identified with the dual of the real continuous function \( C(\Delta_{\kappa}^\infty, \mathbb{R}) \). In particular, we have the convergence of \( X^{(n)} \) towards \( X^{(\infty)} \) in distribution if and only if for any continuous function \( f \in C(\Delta_{\kappa}^\infty, \mathbb{R}) \) 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 \( C(\Delta_{\kappa}^\infty, \mathbb{R}) \) 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 \( r \geq 0, p_1, \ldots, p_r \geq 0 \). We claim that the span of \( S \) (that is finite linear combinations of elements of \( S \)) is dense in \( C(\Delta_{\kappa}^\infty, \mathbb{R}) \).

Indeed, \( S \) contains \( 1 \) and is stable under multiplication. Therefore, the algebra generated by \( S \) is equal to its span.

Now, the set \( S \) is a separating subset of \( C(\Delta_{\kappa}^\infty, \mathbb{R}) \) and density follows from the Stone–Weierstrass theorem. \( \square \)

We will use the following asymptotic simplification of the moments.

Theorem 25. For \( g \geq 2 \) integral, \( m \in \mathbb{N} \cup \{+\infty\} \) and \( \kappa > 1 \) real, let \( \tilde{L}(g, m, \kappa)^\ast \) be the size-biased reordering of the random variable \( L(g, m, \kappa)^\ast \) from Theorem 17. Let \( r \geq 1 \) and \( p_1, \ldots, p_r \in \mathbb{N} \). Then, as \( g \to \infty \), the moment \( M_p(\tilde{L}(g, m, \kappa)^\ast) \) is asymptotically equivalent to
\[
\sqrt{\frac{m+1}{m}} \sqrt{\pi} \sum_{k=r}^{\kappa \log(4g-6)} \frac{1}{(k - r)!} \sum_{(j_1, \ldots, j_k) \in \mathbb{N}^k} \prod_{i=1}^{k} \frac{\zeta_m(2j_i)}{2j_i} \prod_{i=1}^{r} \frac{(2j_i + p_i)}{(2j_i - 1)!}
\]

where
\[
\zeta_m(s) := \sum_{n=1}^{m} \frac{1}{n^s}
\]
is the partial Riemann zeta function.
Following [DGZZ] Equation (14), we define
\[
c_{g,k}(d_1, \ldots, d_k) := \frac{g! \cdot (3g - 3 + 2k)!}{(6g - 5 + 4k)!} \cdot \frac{3^p}{2^{g-6+5k}} \cdot (2d_1 + 2) \cdots (2d_k + 2)!
\]
\[
\cdot \sum_{d_i^+ + d_i^- = d_i, \ \text{w} \ \	ext{w}} \frac{\prod (\tau_{d_i^+} \tau_{d_i^-})^{g-2k}}{d_1^+! \cdots d_k^+! \cdot d_1^-! \cdots d_k^-!}.
\]

The above coefficients were introduced because by [DGZZ] Lemma 3.5, and we have
\[
\lim_{g \to \infty} \sup_{k \leq \sqrt{g/200}} |c_{g,k}(d_1, \ldots, d_k) - 1| = 0.
\]

This asymptotic result is a direct consequence of Theorem 3 of A. Aggarwal that we stated in the introduction. We define \(\tilde{c}_{g,k}(j_1, \ldots, j_k) := c_{g-k,k}(j_1-1, \ldots, j_k-1)\).

**Lemma 26.** For each \(k = 1, \ldots, 3g - 3\), let \(U^{(g,m,k)}\) be the random variable on \(\Delta_k\) with density
\[
\frac{(6g - 7)!}{Z_m(F_{g,k})} \sum_{m \in \mathbb{N}^k} \frac{F_{g,k}(x_1, \ldots, x_k)}{m_1 \cdots m_k} \cdot \frac{1}{m_1 \cdots m_k}
\]
where we use the notation \(F_{g,k}\) for \(F_{g,k}\). Then for any \(p = (p_1, \ldots, p_r) \in \mathbb{N}^k\) we have
\[
M_p(U^{(g,k)*}) = \frac{Z_m(F_{g,k}) \cdot (k-r)! \cdot (6g - 7 + p_1 + \cdots + p_r + r)!}{\prod_{i=1}^k \zeta_m(2j_i)} \prod_{i=1}^r \frac{(2j_i + p_i)!}{(2j_i - 1)!}
\]
where
\[
w_{g,k} := \frac{(6g - 5 - 2k)! \cdot (6g - 7)!}{(g - k)! \cdot (3g - 3 - k)!} \cdot \frac{2^{3k-3}}{3^{3g-k}}.
\]

Note that Formula (15) is the density of a probability measure by Remark 15. It is more precisely the density of the asymptotic normalized vector of length of random multicurves restricted to multicurves of the type \(F_{g,k}\).

**Proof.** By definition of the stable graph polynomial we have
\[
F_{g,k}(x_1, x_2, \ldots, x_k) = x_1 \cdots x_k \cdot V_{g-k,2k}(x_1, x_1, x_2, x_2, x_3, \ldots, x_k, x_k) =
\frac{1}{2^{3g-3-k}} \sum_{(d_1^+, \ldots, d_k^+, d_1^-, \ldots, d_k^-) \in \mathbb{N}^{2k}} \frac{\prod (\tau_{d_i^+} \tau_{d_i^-})^{g-k-2k}}{d_1^+! \cdots d_k^+! \cdot d_1^-! \cdots d_k^-!} \cdot x_1^{2(d_1^+ + d_1^-) + \cdots + x_k^{2(d_k^+ + d_k^-) + 1}}.
\]

Using the coefficients \(\tilde{c}_{g,k}\) defined just above the statement of the lemma, we rewrite the polynomial \(F_{g,k}\) as
\[
F_{g,k} = \frac{(6g - 5 - 2k)!}{(g - k)! \cdot (3g - 3 - k)!} \cdot \frac{2^{3k-3}}{3^{3g-k}} \sum_{(j_1, \ldots, j_k) \in \mathbb{N}^k, j_1 + \cdots + j_k = 3g-3} \tilde{c}_{g,k}(j_1, \ldots, j_k) \prod_{i=1}^k \zeta_m(2j_i) \frac{x_1^{2j_i-1}}{(2j_i)!}.
\]

Hence the density of \(U^{(g,m,k)}\) in (15) can be rewritten as
\[
\frac{w_{g,k}}{Z_m(F_{g,k})} \sum_{(j_1, \ldots, j_k) \in \mathbb{N}^k, j_1 + \cdots + j_k = 3g-3} \tilde{c}_{g,k}(j_1, \ldots, j_k) \prod_{i=1}^k \zeta_m(2j_i) \frac{x_1^{2j_i-1}}{(2j_i)!}.
\]
Now, by Lemma 18, the r-th marginal of the sized-biased version $U^{(g,m,κ)^*}$ of $U^{(g,m,κ)}$ is
\[
\frac{w_{g,k} \cdot k!}{Z_m(F_{g,k}) \cdot (k - r)!} \cdot \frac{1}{(1 - x_1) \cdots (1 - x_1 - \cdots - x_{r-1})} \cdot \sum_{(j_1, \ldots, j_k) \in \mathbb{N}^k \atop j_1 + \cdots + j_k = 3g - 3} \tilde{c}_{g,k}(j_1, \ldots, j_k) \prod_{i=1}^k \frac{ζ_m(2j_i)}{(2j_i)!} \prod_{i=1}^r x_i^{2j_i} \cdot \int_{\Delta_{r+1}^\circ} x_r^{2j_{r+1} - 1} \cdots x_k^{2j_k - 1} \, dλ_1^k(x_{r+1}, \ldots, x_k).
\]
In the above, 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 14 and is equal to
\[
\frac{(2j_{r+1} - 1)! \cdots (2j_k - 1)!}{(2j_{r+1} + \cdots + 2j_k - 1)!} \frac{1}{(1 - x_1 - \cdots - x_r)^{2j_{r+1} + \cdots + 2j_k - 1}}.
\]
We end up with the following formula for the distribution of the r-th marginal of $U^{(g,m,κ)^*}$
\[
\frac{w_{g,k} \cdot k!}{Z_m(F_{g,k}) \cdot (k - r)!} \cdot \frac{1}{(1 - x_1) \cdots (1 - x_1 - \cdots - x_{r-1})} \cdot \sum_{(j_1, \ldots, j_k) \in \mathbb{N}^k \atop j_1 + \cdots + j_k = 3g - 3} \tilde{c}_{g,k}(j_1, \ldots, j_k) \prod_{i=1}^k \frac{ζ_m(2j_i)}{(2j_i)!} \prod_{i=1}^r x_i^{2j_i} \cdot \int_{\Delta_{r+1}^\circ} x_r^{2j_{r+1} - 1} \cdots x_k^{2j_k - 1} \, dλ_1^k(x_{r+1}, \ldots, x_k).
\]
From the above formula and the definition of the moment $M_p$ in [14], the moment $M_p(U^{(g,m,κ)^*})$ equals
\[
\frac{w_{g,k} \cdot k!}{Z_m(F_{g,k}) \cdot (k - r)!} \cdot \frac{1}{(1 - x_1) \cdots (1 - x_1 - \cdots - x_{r-1})} \cdot \sum_{(j_1, \ldots, j_k) \in \mathbb{N}^k \atop j_1 + \cdots + j_k = 3g - 3} \tilde{c}_{g,k}(j_1, \ldots, j_k) \prod_{i=1}^k \frac{ζ_m(2j_i)}{(2j_i)!} \prod_{i=1}^r x_i^{2j_i} \cdot \int_{\Delta_{r+1}^\circ} x_r^{2j_{r+1} - 1} \cdots x_k^{2j_k - 1} \, dλ_1^k(x_{r+1}, \ldots, x_k).
\]
Lemma 14 gives the value of the above integral
\[
\int_{\Delta_{r+1}^\circ} x_r^{2j_{r+1} - 1} \cdots x_k^{2j_k - 1} \, dλ_{r+1}^k = \frac{(2j_1 + p_1)! \cdots (2j_r + p_r)! \cdot (2j_{r+1} + \cdots + 2j_k - 1)!}{(6g - 6 + p_1 + \cdots + p_r + r - 1)!}.
\]
Substituting the above value in our last expression for $M_p(U^{(g,m,κ)^*})$ gives the announced formula.

Proof of Theorem 23. Because the distribution of $\tilde{L}^{(g,m,κ)}$ is a weighted sum of distributions, we can perform the computation of the moments for each term in the sum and gather the result in the end. More precisely, we have
\[
M_p(\tilde{L}^{(g,m,κ)^*}) = \frac{1}{(6g - 6)!} \sum_k \frac{\hat{h}_{g,m,κ}}{2^k \cdot k!} \cdot M_p(U^{(g,m,κ)})
\]
where $\hat{b}_{g,m,κ}$ was defined in [4].
Now substituting the formula for $M_p(U^{(g,m,k)})$ from Lemma 26 and the asymptotic value of $\tilde{b}_{g,m,k}$ from Theorem 11 in the sum 16, we have as $g \to \infty$ the asymptotic equivalence
\begin{equation}
M_p(L^{(a,m,k)}) \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{4g - 4}{4}\right)^{4g - 4}
\end{equation}
\begin{equation*}
\quad \quad \cdot \sum_{k=1}^{\left\lfloor \frac{\log((6g - 6))}{2}\right\rfloor} \frac{1}{(2k \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}}
\end{equation*}
\begin{equation*}
\quad \quad \cdot \sum_{(j_1, \ldots, j_k) \in \mathbb{N}^k} \prod_{j_i=1}^k \zeta_m(2j_i) \prod_{i=1}^r \frac{(2j_i + p_i)!}{(2j_i - 1)!}
\end{equation*}
where we have used that $\tilde{c}_{g,k}(j_1, \ldots, j_k) \sim 1$ uniformly in $k \in [1, \kappa \log(6g - 6)/2]$. On the one hand, by [DGZZ, Equation (3.13)] (in the proof of Theorem 3.4) we have
\begin{equation}
\frac{(4g - 4)! \cdot (6g - 5 - 2k)!}{(6g - 7)! \cdot (g - k)! \cdot (3g - 3 - k)!} \sim (6g - 6)^{1/2} \sqrt{\pi} \frac{2^{8g - 2k}}{3^{g - 4 + k}}.
\end{equation}
On the other hand
\begin{equation}
\frac{(6g - 7)!}{(6g - 7 + p_1 + \cdots + p_r + r)!} \sim \frac{1}{(6g - 6)^p_{1 + \cdots + p_r}}.
\end{equation}
Replacing (18) and (19) in (17) we obtain
\begin{equation}
M_p(L^{(a,m,k)}) \sim \frac{1}{2 \cdot (6g - 6)^{p_{1 + \cdots + p_r - 1/2}}} \cdot \sqrt{\frac{m + 1}{m}} \cdot \sqrt{\pi}
\end{equation}
\begin{equation*}
\quad \quad \cdot \sum_{k=1}^{\left\lfloor \frac{\log((6g - 6))}{2}\right\rfloor} \frac{1}{(k - r)!} \cdot \sum_{(j_1, \ldots, j_k) \in \mathbb{N}^k \atop j_1 + \cdots + j_k = 3g - 3} \prod_{j_i=1}^k \frac{(2j_i + p_i)!}{(2j_i - 1)!}
\end{equation*}
which is the announced formula. \hfill \Box

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
\begin{equation}
S_{\theta, p, n} := \sum_{k=r}^{\infty} \frac{1}{(k - r)!} \sum_{(j_1, \ldots, j_k) \in \mathbb{N}^k \atop j_1 + \cdots + j_k = n} \prod_{j_i=1}^k \frac{\theta_i}{2j_i} \prod_{i=1}^r \frac{(2j_i + p_i)!}{(2j_i - 1)!},
\end{equation}
which should be reminiscent of the formula from Theorem 25.

Definition 27. Let $\theta = (\theta_j)_{j \geq 1}$ be non-negative real numbers and let $g_\theta(z)$ be the formal series
\begin{equation}
g_\theta(z) := \sum_{j \geq 1} \theta_j \frac{z^j}{j}.
\end{equation}
We say that $\theta$ is admissible if the function $g_\theta(z)$
- converges in the open disk $D(0, 1) \subset \mathbb{C}$ centered at 0 of radius 1,
- $g_\theta(z) + \log(1 - z)$ extends to a holomorphic function on $D(0, R)$ with $R > 1$. 

**Theorem 28.** Let $\theta = (\theta_k)_{k \geq 1}$ be admissible, then as $n \to \infty$ we have

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

where $\beta$ is the value at $z = 1$ of $g_0(z) + \log(1-z)$.

The following is essentially [DGZZ, Lemma 3.8] that we reproduce for completeness.

**Lemma 29.** For $m \in \mathbb{N} \cup \{+\infty\}$, let

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

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 $(\zeta_m(2j))_{j \geq 1}$ is admissible. Moreover $(g_m(z) + \log(1-z))|_{z=1} = \log \left( \frac{2m}{m+1} \right)$.

**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 $\zeta_m$ and changing the order of summation we have for $z \in D(0,1)$

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

and hence

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

The term $\log \left( 1 - \frac{z}{n^2} \right)$ defines a holomorphic function on $D(0,n^2)$. Since $|\log \left( 1 - \frac{z}{n^2} \right)| \leq \frac{4}{n^2} |\log \left( 1 - \frac{z}{n^2} \right)|$ 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

$$(g_m(z) + \log(1-z))|_{z=1} = -\sum_{n=2}^{m} \log \left( 1 - \frac{1}{n^2} \right)$$

$$= \sum_{n=2}^{m} (2 \log(n) - \log(n-1) + \log(n+1))$$

$$= \log \left( \frac{2m}{m+1} \right).$$

This completes the proof.

**Corollary 30.** Let $m \in \mathbb{N} \cup \{+\infty\}$. For $\theta = (\zeta_m(2i))_{i \geq 1}$ we have

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

For a non-negative integer $p$ we define the differential operator on $\mathbb{C}[[z]]$ by

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

We start with some preliminary lemmas.

**Lemma 31.** Let $\theta = (\theta_i)_i$ and $g_0(z)$ as in Theorem 28. Let $p = (p_1, \ldots, p_r)$ be a tuple of non-negative integers and let

$$G_{\theta,p}(z) := \exp \left( \frac{1}{2} g_0(z^2) \right) \prod_{i=1}^{r} D_{p_i} \left( \frac{1}{2} g_0(z^2) \right).$$
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 \([20]\).

**Proof.** Let us first note that \( \frac{1}{2} g_0(z^2) = \sum \theta_i \frac{z}{2i}. \) We aim to compute the expansion of \( D_p \left( \frac{1}{2} g_0(z^2) \right). \) By linearity, it is enough to compute a single term and we have
\[
D_p \left( \frac{1}{2} g_0(z^2) \right) = \frac{d^{p+1}}{dz^{p+1}} \left( \frac{1}{2} g_0(z^2) \right) = \frac{(2j + p)!}{(2j)!} z^{2j}.
\]
Hence
\[
D_p \left( \frac{1}{2} g_0(z^2) \right) = \frac{\sum (2j + p)! \theta_j z^{2j}}{(2j - 1)!}.
\]
The lemma follows by expanding the exponential. \( \square \)

**Lemma 32.** For any \( p \geq -1 \) we have
\[
\frac{1}{p!} D_p \left( \log(1 + z) \right) = \frac{1}{(1 + z)^{p+1}} - 1.
\]

**Proof.** By Leibniz’s rule,
\[
z \frac{d^{p+1}}{dz^{p+1}} \left( z^p \log \frac{1}{1 - z} \right) = \frac{z}{p!} \sum_{i=0}^{p} \binom{p+1}{i} p(p-1) \cdots (p-i+1) z^{p-i} (p-i)! \frac{1}{(1 - z)^{p+1-i}}
\]
\[
= -1 + \frac{z}{p!} \sum_{i=0}^{p+1} \binom{p+1}{i} \frac{z^i}{1 - z} \frac{p+1}{p+1-i} \frac{1}{(1 - z)^{p+1-i}}
\]
\[
= -1 + \frac{1}{1 - z} \frac{z^{p+1}}{(1 - z)^{p+1}}.
\]
The proof for \( - \log(1 + z) \) is similar. \( \square \)

**Proof of Theorem 28.** By Lemma \([31]\) the sum \( S_{\theta,p,n}\) is the coefficient in front of \( z^{2n} \) of \( G_{\theta,p}\). By the conditions in the statement, \( g_0(z^2) = - \log(1 - z^2) + \beta + r_\theta(z) \) where \( r_\theta(z) \) is holomorphic on \( D(0, \sqrt{R}) \) and \( r_\theta(1) = 0 \). Using Lemma \([32]\) we deduce that for any \( p \geq 0 \) we have
\[
D_p \left( g_0(z^2) \right) = \frac{1}{(1 - z)^{p+1}} + \frac{1}{(1 + z)^{p+1}} + r_{\theta,p}(z)
\]
where \( r_{\theta,p}\) is holomorphic in \( D(0, \sqrt{R}) \). We deduce that \( G_{\theta,p}(z) \) is meromorphic in \( D(0, \sqrt{R}) \setminus \{1, \sqrt{R}\} \) and satisfies as \( z \to 1 \)
\[
G_{\theta,p}(z) = \exp \left( - \frac{1}{2} (\log(1 - z) + \log(2) - \beta) + O(1 - z) \right) \prod_{i=1}^{\rho} \frac{2^i}{2^r} \left( \frac{1}{(1 - z)^{p_i+1}} + O(1) \right)
\]
\[
= \sqrt{\frac{e^\beta}{2}} \cdot \frac{p_1! \cdots p_r!}{2^r} \frac{1}{(1 - z)^{p_1 + \cdots + r + 1/2}} (1 + o(1)).
\]
Similarly, as \( z \to -1 \) we have
\[
G_{\theta,p}(z) = \sqrt{\frac{e^\beta}{2}} \cdot \frac{p_1! \cdots p_r!}{2^r} \frac{1}{(1 + z)^{p_1 + \cdots + p_r + r + 1/2}} (1 + o(1)).
\]
Now using \([20, \text{Theorem VI.5}]\), 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. \( \square \)
6.3. Truncation estimates. Recall that Theorem 28 provided an expression for the moment $M_p(L^{(0,0)})$ which involves a sum which is a truncated version of $S_{θ,p,n}$ from (20). In this section, we show that the difference between $S_{θ,p,n}$ and its truncation is negligible compared to the asymptotics of Theorem 28.

Theorem 33. Let $θ$ and $g_0(z)$ be as in Theorem 28. Then for any real $κ > 1$ we have as $n → ∞$

$$S_{θ,p,n} ∼ \sum_{k=r}^{κ \log(2n)} \frac{1}{(k-r)!} \frac{θ}{j_1} \prod_{i=1}^{r} (2j_i + p_i)!$$

Bounding the coefficient in a Taylor expansion is a standard tool in asymptotic analysis as the “Big-Oh transfer” [FS09, 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 34 ([DCZZ Lemma 4.4]). Let $λ$ and $x$ be positive real numbers. We have,

$$\sum_{k=[xλ]}^{∞} \frac{λ^k}{k!} ≤ \exp(-λ(x \log x - x)).$$

Lemma 35. Let $g(z)$ be a holomorphic function on $D(0,R) \setminus ([1,R) \cup [-1,-R])$ such that, as $z → ±1$ we have

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

Fix a real $κ > 1$, and non-negative integers $p$ and $q$. For $n ≥ 1$ let

$$f_n(z) = \frac{1}{(1-z)^p(1+z)^q} \sum_{k=[κ \log n]}^{∞} \frac{h(z)^k}{k!}.$$

Then, we have as $n → ∞$

$$\lfloor z^n \rfloor f_n(z) = O \left( n^{\max\{p,q\} - 1 - \frac{1}{2}(κ \log κ - κ)} \right).$$

Proof. Let $0 < η < R - 1$ and $0 < φ < π/2$ and define the contour $γ$ as the union $σ_+ \cup σ_- \cup λ_+ \cup λ_- \cup λ_+ \cup λ_- \cup Σ_+ \cup Σ_-$ with

$$σ_+ = \{ z : |z - 1| = 1/n, \arg(z - 1) ≥ φ \},$$

$$σ_- = \{ z : |z + 1| = 1/n, \arg(z - 1) ≥ φ \},$$

$$λ_+ = \{ z : |z - 1| ≥ 1/n, |z| ≤ 1 + η, \arg(z - 1) = φ \},$$

$$λ_- = \{ z : |z - 1| ≥ 1/n, |z| ≤ 1 + η, \arg(z - 1) = -φ \},$$

$$Σ_+ = \{ z : |z| = 1 + η, \arg(z - 1) ≥ φ, \arg(z + 1) ≤ -π - φ \},$$

$$Σ_- = \{ z : |z| = 1 + η, \arg(z - 1) ≤ -φ, \arg(z + 1) ≤ -π + φ \}.$$

See Figure 3 for a picture of $γ$. Since $f_n$ is holomorphic on $D(0,R) \setminus ([1,R) \cup [-1,-R])$, we have the Cauchy’s residue theorem for its coefficients

$$\lfloor z^n \rfloor f_n(z) = \frac{1}{2πi} \int_γ \frac{f_n(z)}{z^n + 1} dz.$$
Taking absolute values in (25) we obtain

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

The proof proceeds by analyzing the right-hand side in (26) for each piece of the contour \(\gamma\).

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

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

First \(h(1-e^{i\theta}/n) = \frac{\log(n)}{2} + O(1)\) uniformly in \(\theta\) as \(n \to \infty\) we have

\[
\sum_{\kappa=\lfloor \kappa \log n/2 \rfloor}^{\infty} \frac{|h(z)|^k}{k!} \leq \exp \left(-\left(\kappa \log \kappa - \kappa\right) \cdot \frac{\log n + O(1)}{2}\right) = O\left(n^{-\frac{1}{2}(\kappa \log \kappa - \kappa)}\right).
\]

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

\[
\left| \frac{1}{2\pi i} \int_{\sigma_+} \frac{f_n(z)dz}{z^{n+1}} \right| = O\left(n^{p-1-\frac{1}{2}(\kappa \log \kappa - \kappa)}\right).
\]

Similarly,

\[
\left| \frac{1}{2\pi i} \int_{\sigma_-} \frac{f_n(z)dz}{z^{n+1}} \right| = O\left(n^{q-1-\frac{1}{2}(\kappa \log \kappa - \kappa)}\right).
\]
Let us now consider the case of $\lambda_\gamma$. 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

$$\left| \frac{1}{2\pi i} \int_{\lambda_\gamma} \frac{f(z)}{z^{n+1}} dz \right| \leq \frac{\eta^{p-1}}{2\pi \cdot (R + 1)^n} \int_1^{nr} dt \cdot t^{-p} |1 + e^{i\phi} t/n|^{-n-1} \sum_{k=\lceil \kappa \log n \rceil} |h(1 + e^{i\phi} t/n)|^k k!.$$

For $n$ large enough and uniformly in $t$, $|h(1 + e^{i\phi} t/n)| = \frac{\log(n)}{2} + O(1)$. Lemma 34 gives

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

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

$$\int_1^{nr} t^{-p} |1 + e^{i\phi} t/n|^{-n-1} dt = O(1),$$

and therefore

$$\left| \frac{1}{2\pi i} \int_{\lambda_\gamma} \frac{f(z)dz}{z^{n+1}} \right| = O\left(n^{-\frac{1}{2} (\kappa \log \kappa - \kappa)}\right).$$

The same estimate is valid for the integral along the other three segments $\lambda_{\kappa}$, $\lambda_{\kappa'}$, and $\lambda_{r'}$.

For the two large demi-circles $\Sigma_+$ and $\Sigma_-$, we have

$$\left| \frac{1}{2\pi i} \int_{\Sigma_+} \frac{f(z)dz}{z^{n+1}} \right| \leq \frac{1}{2\pi} \cdot \eta^{p+1/2} \cdot (1 + \eta)^{-n-1} \cdot 2\pi (1 + \eta) = \frac{\eta^{p+1/2}}{(1 + \eta)^{n+1}}$$

which decreases exponentially fast.

We conclude the proof by combining the above estimates.

**Proof of Theorem 33.** Similarly to Lemma 31 if we write

$$G_{n,\theta,p}(z) := \sum_{k=\lceil \log(n) \rceil} (\frac{1}{2} g_{\theta}(z^2))^k \frac{r!}{k!} \prod_{i=1}^r \left( \frac{1}{2} g_{\theta}(z^2) \right),$$

then $[z^{2n}] G_{n,\theta,p}$ is the complement of the partial sum in the right hand side of (24).

Following the proof of Theorem 28 we obtain as $z \to 1$

$$G_{n,\theta,p}(z) = \sum_{k=\lceil \log(2n) \rceil} (\frac{1}{2} g_{\theta}(z^2))^k \frac{r!}{k!} \prod_{i=1}^r \left( \frac{1}{2} g_{\theta}(z^2) \right) + O(1)$$

where the $O(1)$ is uniform in $n$ (it only depends on $g_{\theta}(z)$). Applying Lemma 35 we obtain

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

For $\kappa > 1$ we have $-1 - \frac{1}{2} (\kappa \log \kappa - \kappa) < -1/2$ and the above sum is negligible compared to the asymptotics of the full sum $S_{\theta,p,n}$ from Theorem 28. \qed
6.4. Proof of Theorem 23

Proof of Theorem 23. By Lemma 24 it suffices to prove the convergence of the moments $M_p(\tilde{L}(g,m,\kappa)^*)$ for all $p = (p_1, \ldots, p_r)$ towards the moments of the GEM(1/2) distribution that were computed in Lemma 22.

Now, Theorem 25 provides an asymptotic equivalence of $M_p(\tilde{L}(g,m,\kappa)^*)$ involving the sum

$$
\sum_{k=r}^{\infty} \frac{1}{(k-r)!} \prod_{i=1}^{k} \frac{\zeta_m(2j_i)}{2j_i} \prod_{i=1}^{r} \frac{(2j_i + p_i)!}{(2j_i - 1)!}.
$$

The asymptotics of the above sum was then obtained from Corollary 30 and Theorem 33. Namely, the above is asymptotically equivalent to

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

Substituting this value in the formula of Theorem 25 we obtain as $g \to \infty$

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

The above is the value of the moments $M_p$ of the distribution GEM(1/2) from Lemma 22 as $\theta = 1/2$ and $(\theta - 1)! = (-1/2)! = \Gamma(1/2) = \sqrt{\pi}$.

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

References

[ABCDGLW] J. E. Andersen, G. Borot, S. Charbonnier, V. Delecroix, A. Giacchetto, D. Lewanski, C. Wheeler Topological recursion for Masur-Veech volumes preprint [arXiv:1905.10352] [math.GT].

[ABT03] R. Arratia, A. D. Barbour and S. Tavaré Logarithmic combinatorial structures: A probabilistic approach EMS Monographs in Mathematics. Zürich (2003) 363 p.

[Agg21] A. Aggarwal Large genus asymptotics for intersection numbers and principal strata volumes of quadratic differentials Invent. Math. 226 (2021) 897–1010.

[A-H20a] F. Arana-Herrera Equidistribution of families of expanding horospheres on moduli spaces of hyperbolic surfaces Geometriae Dedicata 210 (2020) 65–102.

[A-H20b] F. Arana-Herrera 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 (2020) 81–107.

[A-I] F. Arana-Herrera Counting hyperbolic multi-geodesics with respect to the lengths of individual components and asymptotics of Weil-Petersson volumes preprint [arXiv:2002.10906] [math.DS].

[Bil95] P. Billingsley Probability and measure. 3rd ed John Wiley & Sons Ltd., (1995) 593p.

[Bus10] P. Buser Geometry and spectra of compact Riemann surfaces Modern Birkhäuser Classics. Birkhäuser Boston, Inc., Boston, MA, 2010. Reprint of the 1992 edition.

[CMS] D. Chen, M. Möller, A. Sauvaget, with an appendix by G. Borot, A. Giacchetto, D. Lewanski Masur-Veech volumes and intersection theory: the principal strata of quadratic differentials preprint [arXiv:1912.02867] [math.AG].

[DGZZ] V. Delecroix, É. Goujard, P. Zograf and A. Zorich Large genus asymptotic geometry of random square-tiled surfaces and of random multicurves preprint [arXiv:2007.04740] [math.GT].

[DGZZ21] V. Delecroix, É. Goujard, P. Zograf and A. Zorich Masur-Veech volumes, frequencies of simple closed geodesics and intersection numbers of moduli spaces of curves Duke Math. J. 170:12 (2021) 2633–2718.

[DJ89] P. Donnelly and P. Joyce Continuity and weak convergence of ranked and size-biased permutations on the infinite simplex Stochastic Processes Appl. 31:1 (1989) 89–103.

[ES] V. Erlandsson and J. Souto Mirzakhani’s Curve Counting and geodesic currents to appear

[Fo90] P. Flajolet and A. Odlyzko Singularity analysis of generating functions SIAM J. Discrete Math. 3:2 (1990) 216–240.
[FS09] P. Flajolet and R. Sedgewick Analytic Combinatorics Cambridge University Press, 2009.

[Griff97] R. C. Griffiths On the distribution of allele frequencies in a diffusion model Theor. Popul. Biol. 15 (1979) 140–158.

[Griff88] R. C. Griffiths On the distribution of points in a Poisson Dirichlet process J. Appl. Probab. 25:2 (1988) 336–345.

[ItzyZu92] C. Itzykson and J.-B. Zuber Combinatorics of the modular group II: The Kontsevich integrals, Int. J. Mod. Phys. A 7:23 (1992) 5661–5705.

[Kin75] J. F. C. Kingman Random Discrete Distributions Journal of the Royal Statistical Society. Series B (Methodological) 37:1 (1975) 1–22.

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

[Liu] M. Liu Length statistics of random multicurves on closed hyperbolic surfaces preprint arXiv:1912:11155.

[Mas82] H. Masur Interval exchange transformations and measured foliations Ann. Math. (2) 115 (1982) 169–200.

[Mir07a] M. Mirzakhani. Random hyperbolic surfaces and measured laminations In the tradition of Ahlfors–Bers. IV, Contemp. Math., vol. 432, Amer. Math. Soc., Providence, RI, (2007), no. 1, 179–222.

[Mir07b] M. Mirzakhani Simple geodesics and Weil–Petersson volumes of moduli spaces of bordered Riemann surfaces Invent. Math. 167:1 (2007) 179–222.

[Mir07c] M. Mirzakhani. Weil–Petersson volumes and intersection theory on the moduli space of curves J. Amer. Math. Soc. 20:1 (2007) 1–23.

[Mir08] M. Mirzakhani. Growth of the number of simple closed geodesics on hyperbolic surfaces Ann. Math. (2) 168:1 (2008) 97–125.

[MP19] M. Mirzakhani and B. Petri Lengths of closed geodesics on random surfaces of large genus Comment. Math. Helv. 94:4 (2019) 869–889.

[Mir] M. Mirzakhani. Counting Mapping Class group orbits on hyperbolic surfaces preprint arXiv:1601.03342 [math.GT].

[Vee82] W. A. Veech Gauss measures for transformations on the space of interval exchange maps Ann. Math. (2) 115 (1982) 201–242.

[Wit91] E. Witten Two-dimensional gravity and intersection theory on moduli space Proc. Conf., Cambridge/MA (USA) 1990, Surv. Differ. Geom., Suppl. J. Diff. Geom. 1 (1991) 243–310.

[Wol83] S. Wolpert. On the homology of the moduli of stable curves Annals of Mathematics Second Series, 118:3 (1983) 491–523.

[YZZ] D. Yang, D. Zagier, Y. Zhang Masur–Veech volumes of quadratic differentials and their asymptotics preprint arXiv:2005.02275 [math-ph].