

# MASUR–VEECH VOLUMES, FREQUENCIES OF SIMPLE CLOSED GEODESICS, AND INTERSECTION NUMBERS OF MODULI SPACES OF CURVES

---

VINCENT DELECROIX, ÉLISE GOUJARD, PETER ZOGRAF, and ANTON ZORICH

## Abstract

We express the Masur–Veech volume and the area Siegel–Veech constant of the moduli space  $\mathcal{Q}_{g,n}$  of genus  $g$  meromorphic quadratic differentials with at most  $n$  simple poles and no other poles as polynomials in the intersection numbers  $\int_{\overline{\mathcal{M}}_{g',n'}} \psi_1^{d_1} \cdots \psi_{n'}^{d_{n'}}$  with explicit rational coefficients, where  $g' < g$  and  $n' < 2g + n$ . The formulas obtained in this article are derived from lattice point counts involving the Kontsevich volume polynomials  $N_{g',n'}(b_1, \dots, b_{n'})$  that also appear in Mirzakhani’s recursion for the Weil–Petersson volumes of the moduli spaces  $\mathcal{M}_{g',n'}(b_1, \dots, b_{n'})$  of bordered hyperbolic surfaces with geodesic boundaries of lengths  $b_1, \dots, b_{n'}$ . A similar formula for the Masur–Veech volume (but without explicit evaluation) was obtained earlier by Mirzakhani through a completely different approach. We prove a further result: the density of the mapping class group orbit  $\text{Mod}_{g,n} \cdot \gamma$  of any simple closed multicurve  $\gamma$  inside the ambient set  $\mathcal{ML}_{g,n}(\mathbb{Z})$  of integral measured laminations, computed by Mirzakhani, coincides with the density of square-tiled surfaces having horizontal cylinder decomposition associated to  $\gamma$  among all square-tiled surfaces in  $\mathcal{Q}_{g,n}$ . We study the resulting densities (or, equivalently, volume contributions) in more detail in the special case when  $n = 0$ . In particular, we compute explicitly the asymptotic frequencies of separating and nonseparating simple closed geodesics on a closed hyperbolic surface of genus  $g$  for all small genera  $g$ , and we show that in large genera the separating closed geodesics are  $\sqrt{\frac{2}{3\pi g}} \cdot \frac{1}{4^g}$  times less frequent.

## Contents

1. Introduction and statements of main theorems . . . . .	2634
---	------

DUKE MATHEMATICAL JOURNAL

Vol. 170, No. 12, © 2021 DOI [10.1215/00127094-2021-0054](https://doi.org/10.1215/00127094-2021-0054)

Received 13 February 2020. Revision received 1 October 2020.

First published online 9 August 2021.

2020 *Mathematics Subject Classification*. Primary 32G15; Secondary 57M50.

- 2. Proofs of the formulas for the Masur–Veech volumes and for the Siegel–Veech constants . . . . . 2662
- 3. Comparison with Mirzakhani’s formula for  $b_{g,n}$  . . . . . 2677
- 4. Large genus asymptotics for frequencies of simple closed curves and 1-cylinder square-tiled surfaces . . . . . 2684
- Appendix A. Proof of the asymptotic formula for 2-correlators . . . . . 2695
- Appendix B. Stable graphs: Formal definition . . . . . 2705
- Appendix C. Examples of explicit calculations . . . . . 2707
- Appendix D. Tables of volumes and Siegel–Veech constants . . . . . 2711
- References . . . . . 2714

**1. Introduction and statements of main theorems**

*1.1. Masur–Veech volume of the moduli space of quadratic differentials*

Let  $g, n$  be nonnegative integers with  $2g + n > 2$ . Consider the moduli space  $\mathcal{M}_{g,n}$  of complex curves of genus  $g$  with  $n$  distinct labeled marked points. The total space  $\mathcal{Q}_{g,n}$  of the cotangent bundle over  $\mathcal{M}_{g,n}$  can be identified with the moduli space of pairs  $(C, q)$ , where  $C \in \mathcal{M}_{g,n}$  is a smooth complex curve with  $n$  (labeled) marked points and  $q$  is a meromorphic quadratic differential on  $C$  with at most simple poles at the marked points and no other poles. In the case  $n = 0$  the quadratic differential  $q$  is holomorphic. Thus, as any total space of the cotangent bundle, the *moduli space of quadratic differentials*  $\mathcal{Q}_{g,n}$  is endowed with the canonical (real) symplectic structure. The induced volume element on  $\mathcal{Q}_{g,n}$  is called the *Masur–Veech volume element*. In Section 2.1 we provide an alternative, more common definition of the Masur–Veech volume element and explain why the two definitions are equivalent up to a global normalization constant.

A nonzero differential  $q$  in  $\mathcal{Q}_{g,n}$  defines a flat metric  $|q|$  on the complex curve  $C$ . The resulting metric has conical singularities at zeros and simple poles of  $q$ . The total area of  $(C, q)$ ,

$$\text{Area}(C, q) = \int_C |q|,$$

is positive and finite. For any real  $a > 0$ , consider the following subset in  $\mathcal{Q}_{g,n}$ :

$$\mathcal{Q}_{g,n}^{\text{Area} \leq a} := \{(C, q) \in \mathcal{Q}_{g,n} \mid \text{Area}(C, q) \leq a\}. \tag{1.1}$$

Since  $\text{Area}(C, q)$  is a norm in each fiber of the bundle  $\mathcal{Q}_{g,n} \rightarrow \mathcal{M}_{g,n}$ , the set  $\mathcal{Q}_{g,n}^{\text{Area} \leq a}$  is a ball bundle over  $\mathcal{M}_{g,n}$ . In particular, it is noncompact. However, by independent results of Masur [42] and Veech [55], the total mass of  $\mathcal{Q}_{g,n}^{\text{Area} \leq a}$  with respect to the Masur–Veech volume element is finite. One of the objectives of this article is to

provide a formula for this total mass. Our construction relies on square-tiled surface counting that we introduce next.

1.2. *Square-tiled surfaces, simple closed multicurves, and stable graphs*

We have already mentioned that a nonzero meromorphic quadratic differential  $q$  on a complex curve  $C$  defines a flat metric with conical singularities. One can construct a discrete collection of quadratic differentials of this kind by assembling identical flat squares in the following way. Take a finite set of copies of the oriented  $(1/2 \times 1/2)$ -square for which two opposite sides are chosen to be horizontal and the remaining two sides are declared to be vertical. Identify pairs of sides of the squares by isometries in such a way that horizontal sides are glued to horizontal sides and vertical sides to vertical. We get a topological surface  $S$  without boundary. We consider only those surfaces obtained in this way which are connected and oriented. The form  $dz^2$  on each square is compatible with the gluing and endows  $S$  with a complex structure and with a nonzero quadratic differential  $q$  with at most simple poles. The total area  $\text{Area}(S, q)$  is  $\frac{1}{4}$  times the number of squares. We call such a surface a *square-tiled surface*, see also [62].

Given a sequence of integers  $\mu = [\mu_1 \dots \mu_m, \mu_{m+1} \dots \mu_{m+n}]$ , where  $[\mu_1 \dots \mu_m]$  is a partition of  $4g - 4 + n$  and  $\mu_{m+1} = \dots = \mu_{m+n} = -1$ , the corresponding *stratum of quadratic differentials*  $\mathcal{Q}(\mu)$  is the space of equivalence classes of pairs consisting of a complex curve  $C$  with  $m + n$  distinct marked points  $z_1, \dots, z_m, p_1, \dots, p_n$  and a quadratic differential  $q$  with the divisor  $\sum_{i=1}^m \mu_i z_i - \sum_{j=1}^n p_j$  (both zeros and poles of  $q$  are considered to be labeled).

For any pair of nonnegative integers  $(g, n)$  satisfying  $2g + n > 3$ , the *principal stratum* is  $\mathcal{Q}(1^{4g-4+n}, -1^n)$  (that is,  $\mu = [1^{4g-4+n}, -1^n]$ ). The natural morphism  $\mathcal{Q}(1^{4g-4+n}, -1^n) \rightarrow \mathcal{Q}_{g,n}$  that forgets the labeling of zeros of  $q$  is a  $(4g - 4 + n)!$ -sheeted ramified cover of its image in  $\mathcal{Q}_{g,n}$ . Moreover, this image is open and dense in  $\mathcal{Q}_{g,n}$ .

*Remark 1.1*

The special cases  $(g, n) = (0, 3)$  and  $(g, n) = (1, 1)$  which correspond to  $2g + n = 3$  are discussed in Appendix C.1. In these two cases,  $\mathcal{Q}_{g,n}$  does not admit any natural interpretation in terms of meromorphic quadratic differentials with simple zeros and simple poles.

We denote by  $\mathcal{ST}(\mathcal{Q}(\mu), N)$  the set of square-tiled surfaces in the stratum  $\mathcal{Q}(\mu)$  made of at most  $N$  squares. For example, the square-tiled surface in Figure 1 has genus  $g = 0$ . It has three simple zeros and  $n = 7$  conical singularities with angle  $\pi$ . Hence, it belongs to the principal stratum  $\mathcal{Q}(1^3, -1^7)$ .

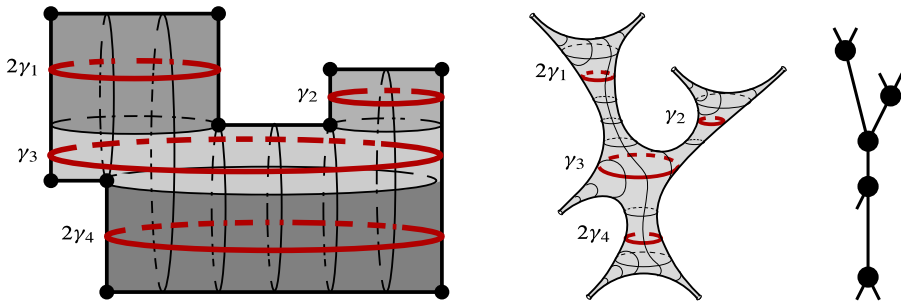


Figure 1. (Color online) A square-tiled surface in  $\mathcal{Q}(1^3, -1^7)$ , and its associated multicurve and stable graph.

We will see in Section 2.1 that the principal strata have a natural linear structure and that the square-tiled surfaces form a covolume-1 lattice in associated period coordinates. This justifies the following definition of the Masur–Veech volume of  $\mathcal{Q}_{g,n}$  (for  $(g, n)$  different from  $(0, 3)$  and  $(1, 1)$ ):

$$\begin{aligned} \text{Vol } \mathcal{Q}_{g,n} &:= \text{Vol } \mathcal{Q}(1^{4g-4+n}, -1^n) \\ &= 2d \cdot \lim_{N \rightarrow +\infty} \frac{\text{card}(\mathcal{ST}(\mathcal{Q}(1^{4g-4+n}, -1^n), 2N))}{N^d}, \end{aligned} \tag{1.2}$$

where  $d = \dim_{\mathbb{C}} \mathcal{Q}(1^{4g-4+n}, -1^n) = \dim_{\mathbb{C}} \mathcal{Q}_{g,n} = 6g - 6 + 2n$ . We emphasize that, in the above formula, we assume that all conical singularities of square-tiled surfaces are labeled. Formula (1.2) is the starting point of our expression for  $\text{Vol } \mathcal{Q}_{g,n}$ .

*Cylinder decomposition, multicurve, and stable graph.* A square-tiled surface admits a decomposition into maximal horizontal cylinders filled with isometric closed regular flat geodesics. Every such maximal horizontal cylinder has at least one conical singularity on each of the two boundary components. The square-tiled surface in Figure 1 has four maximal horizontal cylinders which are represented in the picture by different shades.

Let  $S$  be a square-tiled surface, and let  $S = \text{cyl}_1 \cup \dots \cup \text{cyl}_k$  be its decomposition into the set of maximal horizontal cylinders. To each cylinder  $\text{cyl}_i$  we associate the corresponding waist curve  $\gamma_i$  considered up to a free homotopy. The curves  $\gamma_i$  are nonperipheral (i.e., none of them bounds a disk containing a single pole) and pairwise nonhomotopic. We encode the number of circular horizontal bands of squares contained in the corresponding maximal horizontal cylinder by the integer weight  $H_i$  associated to the curve  $\gamma_i$ . The formal linear combination  $\gamma = \sum H_i \gamma_i$  is a simple closed integral multicurve in the space  $\mathcal{ML}_{g,n}(\mathbb{Z})$  of measured laminations. For

example, the simple closed multicurve associated to the square-tiled surface as in Figure 1 has the form  $2\gamma_1 + \gamma_2 + \gamma_3 + 2\gamma_4$ .

The multicurve  $\gamma = \sum H_i \gamma_i$  as above defines the associated *reduced* multicurve  $\gamma_{\text{reduced}} = \sum \gamma_i$ . Here we assume that  $\gamma_i$  and  $\gamma_j$  are pairwise nonisotopic for  $i \neq j$ . We associate to  $\gamma_{\text{reduced}}$  its *stable graph*  $\Gamma(\gamma_{\text{reduced}})$  which should be thought as the dual graph to  $\gamma_{\text{reduced}}$ . More precisely,  $\Gamma(\gamma)$  is the decorated graph whose vertices represent the components of  $S \setminus \gamma_{\text{reduced}}$  and are decorated with the genus of the corresponding component. By convention, when this number is not explicitly indicated, it is zero. The edges of  $\Gamma(\gamma)$  represent the components  $\gamma_i$  of  $\gamma_{\text{reduced}}$ , where the end-points of the edge associated to  $\gamma_i$  are the two vertices corresponding to the two components of  $S \setminus \gamma_{\text{reduced}}$  adjacent to  $\gamma_i$ . When  $\gamma_i$  has the same component of  $S \setminus \gamma_{\text{reduced}}$  on both sides, the corresponding edge of  $\Gamma(\gamma)$  is a loop. Finally,  $\Gamma(\gamma)$  is endowed with  $n$  “legs” (or half-edges) labeled from 1 to  $n$ . The leg with label  $i$  is attached to the vertex that represents the component that contains the  $i$ th marked point of  $S$ . The right picture in Figure 1 shows the stable graph associated to the multicurve  $\gamma$ . A formal combinatorial definition of a stable graph is provided in Appendix B.

The total number of stable graphs (considered up to isomorphism) is finite and is equal to the number of  $\text{Mod}_{g,n}$ -orbits of reduced multicurves in  $\mathcal{ML}_{g,n}(\mathbb{Z})$ . For  $2g + n > 2$ , we denote by  $\mathcal{G}_{g,n}$  the set of stable graphs. Table 1 in Section 1.4 and Table 5 in Appendix C.4 list all stable graphs in  $\mathcal{G}_{2,0}$  and  $\mathcal{G}_{1,2}$  respectively. The special cases  $\mathcal{G}_{0,3}$  and  $\mathcal{G}_{1,1}$  are considered in Appendix C.1.

Given a pair of nonnegative integers  $(g, n)$  satisfying  $2g + n > 3$  and a stable graph  $\Gamma$  in  $\mathcal{G}_{g,n}$ , let us consider the subset  $\mathcal{ST}_{\Gamma, \mathbf{H}}(\mathcal{Q}(1^{4g-4+n}, -1^n))$  of those square-tiled surfaces, for which the associated stable graph is  $\Gamma$  and the heights of the cylinders are  $\mathbf{H} = (H_1, \dots, H_k)$ . Let us denote by  $\mathcal{ST}_{\Gamma}(\mathcal{Q}(1^{4g-4+n}, -1^n))$  the analogous subset without restriction on the heights. Let us define  $\text{Vol}(\Gamma, \mathbf{H})$  and  $\text{Vol}(\Gamma)$  to be, respectively, the contributions to  $\text{Vol} \mathcal{Q}_{g,n}$  of square-tiled surfaces from the subsets  $\mathcal{ST}_{\Gamma, \mathbf{H}}(\mathcal{Q}(1^{4g-4+n}, -1^n))$  and  $\mathcal{ST}_{\Gamma}(\mathcal{Q}(1^{4g-4+n}, -1^n))$ :

$$\text{Vol}(\Gamma, \mathbf{H}) := 2d \cdot \lim_{N \rightarrow +\infty} \frac{\text{card}(\mathcal{ST}_{\Gamma, \mathbf{H}}(\mathcal{Q}(1^{4g-4+n}, -1^n), 2N))}{N^d}, \tag{1.3}$$

$$\text{Vol}(\Gamma) := 2d \cdot \lim_{N \rightarrow +\infty} \frac{\text{card}(\mathcal{ST}_{\Gamma}(\mathcal{Q}(1^{4g-4+n}, -1^n), 2N))}{N^d}, \tag{1.4}$$

where  $d = 6g - 6 + 2n$ . The results in [21] imply that, for any  $\Gamma$  in  $\mathcal{G}_{g,n}$ , the above limits exist, are strictly positive, and that

$$\text{Vol} \mathcal{Q}_{g,n} = \sum_{\Gamma \in \mathcal{G}_{g,n}} \text{Vol}(\Gamma) = \sum_{\Gamma \in \mathcal{G}_{g,n}} \sum_{\mathbf{H} \in \mathbb{N}^{E(\Gamma)}} \text{Vol}(\Gamma, \mathbf{H}). \tag{1.5}$$

Dividing both sides of (1.5) by  $\text{Vol } \mathcal{Q}_{g,n}$  we see that the ratio  $\text{Vol}(\Gamma)/\text{Vol } \mathcal{Q}_{g,n}$  can be interpreted as the “asymptotic probability” that a square-tiled surface taken at random has  $\Gamma$  as the stable graph associated to its horizontal cylinder decomposition.

*Remark 1.2*

A stable graph is commonly used to encode the boundary classes of the Deligne–Mumford compactification  $\overline{\mathcal{M}}_{g,n}$  of  $\mathcal{M}_{g,n}$ . Informally speaking, the stable curves in the boundary of  $\overline{\mathcal{M}}_{g,n}$  are obtained by pinching along appropriate multicurves  $\gamma_{\text{reduced}}$ . In our situation, this can be done algebraically in the following way. Let  $(C, q)$  be a quadratic differential whose horizontal cylinders fill the associated flat surface  $S$ . Consider the sequence of quadratic differentials  $(C_t, q_t)$  obtained as follows. Define  $q_t := \text{Re}(q) + i e^t \text{Im}(q)$ . The differential  $q_t$  is meromorphic for a unique complex structure  $C_t$ . The horizontal cylinder decompositions of all  $(C_t, q_t)$ ,  $t \in \mathbb{R}$ , are topologically identical. Metrically, cylinders of  $(C_t, q_t)$  are  $e^t$  times higher than cylinders of  $(C, q)$ . The path  $(C_t, q_t)$  in  $\overline{\mathcal{Q}}_{g,n}$  converges toward a stable quadratic differential  $(C_\infty, q_\infty)$  in  $\overline{\mathcal{Q}}_{g,n}$  with exactly double poles at the nodes of  $C_\infty$ . Each double pole corresponds to a half-infinite cylinder associated to the corresponding boundary component of  $S \setminus \gamma_{\text{reduced}}$ .

*1.3. Ribbon graphs, intersection numbers, and volume polynomials*

In this section we introduce multivariate polynomials  $N_{g,n}(b_1, \dots, b_n)$  that appear in different contexts. They are an essential ingredient to our formula for the Masur–Veech volume.

Let  $g, n$  be nonnegative integers with  $2g + n > 2$ . Let  $b_1, \dots, b_n$  be variables. For a multi-index  $\mathbf{d} = (d_1, \dots, d_n)$ , we denote by  $b^{2\mathbf{d}}$  the product  $b_1^{2d_1} \dots b_n^{2d_n}$ , by  $|\mathbf{d}|$  the sum  $d_1 + \dots + d_n$ , and by  $\mathbf{d}!$  the product  $d_1! \dots d_n!$ .

Define the homogeneous polynomial  $N_{g,n}(b_1, \dots, b_n)$  of degree  $6g - 6 + 2n$  in the variables  $b_1, \dots, b_n$  as

$$N_{g,n}(b_1, \dots, b_n) = \sum_{|\mathbf{d}|=3g-3+n} c_{\mathbf{d}} b^{2\mathbf{d}}, \tag{1.6}$$

where

$$c_{\mathbf{d}} = \frac{1}{2^{5g-6+2n} \mathbf{d}!} \langle \psi_1^{d_1} \dots \psi_n^{d_n} \rangle, \tag{1.7}$$

$$\langle \psi_1^{d_1} \dots \psi_n^{d_n} \rangle = \int_{\overline{\mathcal{M}}_{g,n}} \psi_1^{d_1} \dots \psi_n^{d_n}, \tag{1.8}$$

where  $\psi_1, \dots, \psi_n$  are the  $\psi$ -classes on the Deligne–Mumford compactification  $\overline{\mathcal{M}}_{g,n}$ . That is,  $\psi_i$  is the first Chern class of the  $i$ th tautological bundle over  $\overline{\mathcal{M}}_{g,n}$ . Informally,

the fiber of this bundle over  $(C, p_1, \dots, p_n)$  is the cotangent line  $T_{p_i}^* C$  to  $C$  at the  $i$ th marked point. Note that  $N_{g,n}(b_1, \dots, b_n)$  contains only even powers of  $b_i$ , where  $i = 1, \dots, n$ . For small  $g$  and  $n$ , we obtain

$$\begin{aligned} N_{0,3}(b_1, b_2, b_3) &= 1, \\ N_{0,4}(b_1, b_2, b_3, b_4) &= \frac{1}{4}(b_1^2 + b_2^2 + b_3^2 + b_4^2), \\ N_{1,1}(b_1) &= \frac{1}{48}(b_1^2), \\ N_{1,2}(b_1, b_2) &= \frac{1}{384}(b_1^2 + b_2^2)(b_1^2 + b_2^2). \end{aligned}$$

A ribbon graph  $G$  is a connected graph endowed at each vertex with a cyclic ordering of incident half-edges. The cyclic ordering determines faces, so one can consider a tubular neighborhood of  $G$  as a surface with boundary, where boundary components correspond to faces of  $G$ . We denote by  $\mathcal{R}_{g,n}$  the set of isomorphism classes of trivalent ribbon graphs of genus  $g$  with  $n$  faces labeled from 1 to  $n$ . For each trivalent ribbon graph  $G$  in  $\mathcal{R}_{g,n}$  and any collection of integers  $b_1, \dots, b_n$ , denote by  $\mathcal{N}_G(b_1, \dots, b_n)$  the number of integral metrics on  $G$  (assigning a positive integral length to each edge) such that the perimeters of the faces get lengths  $b_1, \dots, b_n$ .

THEOREM (see [39], [52])

Consider a collection of positive integers  $b_1, \dots, b_n$  such that  $\sum_{i=1}^n b_i$  is even. The weighted count of genus  $g$  connected trivalent metric ribbon graphs  $G$  with integer edges and with  $n$  labeled boundary components of lengths  $b_1, \dots, b_n$  is equal to  $N_{g,n}(b_1, \dots, b_n)$  up to the lower-order terms,

$$\sum_{G \in \mathcal{R}_{g,n}} \frac{1}{|\text{Aut}(G)|} \mathcal{N}_G(b_1, \dots, b_n) = N_{g,n}(b_1, \dots, b_n) + \text{lower-order terms},$$

where  $\mathcal{R}_{g,n}$  denotes the set of (nonisomorphic) trivalent ribbon graphs  $G$  of genus  $g$  and with  $n$  boundary components.

A version of this theorem (see the “main identity” in the framed formula in [39, section 3.1]) is a part of Kontsevich’s proof of Witten’s conjecture from [58, Section 2a].

*Remark 1.3*

Norbury in [52, Theorems 1, 3] (see also Chapman, Mulase, and Safnuk in [15]) refined the count of Kontsevich, proving that the function counting lattice points in the moduli space  $\mathcal{M}_{g,n}$  corresponding to covers of the sphere ramified over three points

(the so-called *Grothendieck’s dessins d’enfants*) is a quasipolynomial in variables  $b_i$ . In other terms, when considering all ribbon graphs (and not only trivalent ones) the lower-order terms in Kontsevich’s theorem form a quasipolynomial. This quasipolynomiality of the expression on the left-hand side endows the notion of “lower-order terms” with a natural formal sense.

We also use the following common notation for the intersection numbers (1.8). Given an ordered partition  $d_1 + \dots + d_n = 3g - 3 + n$  of  $3g - 3 + n$  into a sum of nonnegative integers, we define

$$\langle \tau_{d_1} \cdots \tau_{d_n} \rangle = \int_{\mathcal{M}_{g,n}} \psi_1^{d_1} \cdots \psi_n^{d_n}. \tag{1.9}$$

1.4. *Formula for the Masur–Veech volumes*

Following [11], we consider the linear operators  $\mathcal{Y}(\mathbf{H})$  and  $\mathcal{Z}$  on the spaces of polynomials in variables  $b_1, b_2, \dots$ , where  $H_1, H_2, \dots$  are positive integers. The operator  $\mathcal{Y}(\mathbf{H})$  is defined on monomials as

$$\mathcal{Y}(\mathbf{H}) : \prod_{i=1}^k b_i^{m_i} \mapsto \prod_{i=1}^k \frac{m_i!}{H_i^{m_i+1}}, \tag{1.10}$$

and it is extended to arbitrary polynomials by linearity. The operator  $\mathcal{Z}$  is defined on monomials as

$$\mathcal{Z} : \prod_{i=1}^k b_i^{m_i} \mapsto \prod_{i=1}^k (m_i! \cdot \zeta(m_i + 1)), \tag{1.11}$$

and extended to arbitrary polynomials by linearity. In the above formula  $\zeta$  is the Riemann zeta function

$$\zeta(s) = \sum_{n \geq 1} \frac{1}{n^s},$$

so for any collection of strictly positive integers  $(m_1, \dots, m_k)$  one has

$$\mathcal{Z}\left(\prod_{i=1}^k b_i^{m_i}\right) = \sum_{\mathbf{H} \in \mathbb{N}^k} \mathcal{Y}(\mathbf{H})\left(\prod_{i=1}^k b_i^{m_i}\right).$$



*Remark 1.4*

For even integers  $2m$ , we have

$$\zeta(2m) = (-1)^{m+1} \frac{B_{2m}(2\pi)^{2m}}{2(2m)!},$$

where  $B_{2m}$  are the Bernoulli numbers. Consider a homogeneous polynomial in  $k$  variables of degree  $2m - k$  with rational coefficients such that all powers of all variables in each monomial are odd. The observation above implies that the value of  $Z$  on such a polynomial is a rational number multiplied by  $\pi^{2m}$ .

Given a stable graph  $\Gamma$ , denote by  $V(\Gamma)$  the set of its vertices and denote by  $E(\Gamma)$  the set of its edges. To each stable graph  $\Gamma \in \mathcal{G}_{g,n}$  we associate the following homogeneous polynomial  $P_\Gamma$  of degree  $6g - 6 + 2n$ . To every edge  $e \in E(\Gamma)$  we assign a formal variable  $b_e$ . Given a vertex  $v \in V(\Gamma)$ , denote by  $g_v$  the integer number decorating  $v$  and denote by  $n_v$  the valency of  $v$ , where the legs adjacent to  $v$  are counted toward the valency of  $v$ . Take a small neighborhood of  $v$  in  $\Gamma$ . We associate to each half-edge (“germ” of edge)  $e$  incident to  $v$  the monomial  $b_e$ ; we associate 0 to each leg. We denote by  $\mathbf{b}_v$  the resulting collection of size  $n_v$ . If some edge  $e$  is a loop joining  $v$  to itself, then  $b_e$  is present in  $\mathbf{b}_v$  twice; if an edge  $e$  joins  $v$  to a distinct vertex, then  $b_e$  is present in  $\mathbf{b}_v$  once; all the other entries of  $\mathbf{b}_v$  correspond to legs and are represented by zeros. To each vertex  $v \in E(\Gamma)$  we associate the polynomial  $N_{g_v, n_v}(\mathbf{b}_v)$ , where  $N_{g,v}$  is defined in (1.6). We associate to the stable graph  $\Gamma$  the polynomial obtained as the product  $\prod b_e$  over all edges  $e \in E(\Gamma)$  multiplied by the product  $\prod N_{g_v, n_v}(\mathbf{b}_v)$  over all  $v \in V(\Gamma)$ . We define  $P_\Gamma$  as follows:

$$P_\Gamma(\mathbf{b}) = \frac{2^{6g-5+2n} \cdot (4g - 4 + n)!}{(6g - 7 + 2n)!} \cdot \frac{1}{2^{|V(\Gamma)|-1}} \cdot \frac{1}{|\text{Aut}(\Gamma)|} \cdot \prod_{e \in E(\Gamma)} b_e \cdot \prod_{v \in V(\Gamma)} N_{g_v, n_v}(\mathbf{b}_v). \tag{1.12}$$

Table 1 in Section 1.4 and Table 5 in Appendix C.4 list the polynomials associated to all stable graphs in  $\mathcal{G}_{2,0}$  and in  $\mathcal{G}_{1,2}$ .

**THEOREM 1.5**

*The Masur–Veech volume of the stratum of quadratic differentials with  $4g - 4 + n$  simple zeros and  $n$  simple poles has the following value:*

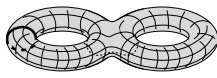

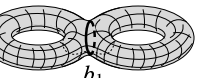

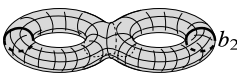

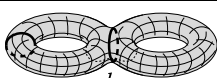
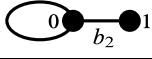
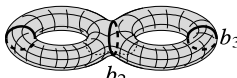
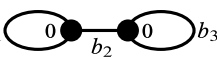
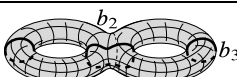
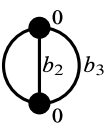
$$\begin{aligned} \text{Vol } \mathcal{Q}_{g,n} &= \text{Vol } \mathcal{Q}(1^{4g-4+n}, -1^n) \\ &= \sum_{\Gamma \in \mathcal{G}_{g,n}} \text{Vol}(\Gamma) = \sum_{\Gamma \in \mathcal{G}_{g,n}} \sum_{\mathbf{H} \in \mathbb{N}^{E(\Gamma)}} \text{Vol}(\Gamma, \mathbf{H}), \end{aligned} \tag{1.13}$$

where the contributions of individual stable graphs  $\Gamma$  defined by (1.3) (resp., the contributions of pairs  $(\Gamma, \mathbf{H})$  defined by (1.4)) are equal to

$$\text{Vol}(\Gamma) = \mathcal{Z}(P_\Gamma) \quad \text{and} \quad \text{Vol}(\Gamma, \mathbf{H}) = \mathcal{Y}(\mathbf{H})(P_\Gamma). \quad (1.14)$$

Table 1 illustrates the computation of the polynomials  $P_\Gamma$  and of the contributions  $\text{Vol}(\Gamma)$  to the Masur–Veech volume  $\text{Vol} \mathcal{Q}_{g,n}$  in the particular case of  $(g, n) = (2, 0)$ .

Table 1. Computation of  $\text{Vol} \mathcal{Q}_{2,0}$ . The left column represents the stable graphs  $\Gamma$  and their associated multicurves; the middle column gives the polynomials  $P_\Gamma$ ; the right column provides  $\text{Vol}(\Gamma)$ .

 <p><math>b_1</math></p>	$\frac{128}{5} \cdot 1 \cdot \frac{1}{2} \cdot b_1 \cdot N_{1,2}(b_1, b_1)$ $= \frac{64}{5} \cdot b_1 \left( \frac{1}{384} (2b_1^2) (2b_1^2) \right) = \frac{2}{15} \cdot b_1^5 \xrightarrow{\mathcal{Z}}$	$\frac{2}{15} \cdot (5! \cdot \zeta(6))$ $= \frac{16}{945} \cdot \pi^6$
 <p><math>b_1</math></p>		
 <p><math>b_1</math></p>	$\frac{128}{5} \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot b_1 \cdot N_{1,1}(b_1) \cdot N_{1,1}(b_1)$ $= \frac{32}{5} \cdot b_1 \cdot \left( \frac{1}{48} b_1^2 \right) \left( \frac{1}{48} b_1^2 \right) = \frac{1}{360} \cdot b_1^5 \xrightarrow{\mathcal{Z}}$	$\frac{1}{360} \cdot (5! \cdot \zeta(6))$ $= \frac{1}{2835} \cdot \pi^6$
 <p><math>b_1</math></p>		
 <p><math>b_1</math></p>	$\frac{128}{5} \cdot 1 \cdot \frac{1}{8} \cdot b_1 b_2 \cdot N_{0,4}(b_1, b_1, b_2, b_2)$ $= \frac{16}{5} \cdot b_1 b_2 \cdot \left( \frac{1}{4} (2b_1^2 + 2b_2^2) \right)$ $= \frac{8}{5} (b_1^3 b_2 + b_1 b_2^3) \xrightarrow{\mathcal{Z}}$	$\frac{8}{5} \cdot 2 \cdot 3! \zeta(4) \cdot 1! \zeta(2)$ $= \frac{8}{225} \cdot \pi^6$
 <p><math>b_1</math></p>		
 <p><math>b_1</math></p>	$\frac{128}{5} \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot b_1 b_2 \cdot N_{0,3}(b_1, b_1, b_2) \cdot N_{1,1}(b_2)$ $= \frac{32}{5} \cdot b_1 b_2 \cdot (1) \cdot \left( \frac{1}{48} b_2^2 \right) = \frac{2}{15} \cdot b_1 b_2^3 \xrightarrow{\mathcal{Z}}$	$\frac{2}{15} \cdot 1! \zeta(2) \cdot 3! \zeta(4)$ $= \frac{1}{675} \cdot \pi^6$
 <p><math>b_1</math></p>		
 <p><math>b_1</math></p>	$\frac{128}{5} \cdot \frac{1}{2} \cdot \frac{1}{8} \cdot b_1 b_2 b_3 \cdot N_{0,3}(b_1, b_1, b_2)$ $\cdot N_{0,3}(b_2, b_3, b_3) = \frac{8}{5} \cdot b_1 b_2 b_3 \cdot (1) \cdot (1) \xrightarrow{\mathcal{Z}}$	$\frac{8}{5} \cdot (1! \zeta(2))^3$ $= \frac{1}{135} \cdot \pi^6$
 <p><math>b_1</math></p>		
 <p><math>b_1</math></p>	$\frac{128}{5} \cdot \frac{1}{2} \cdot \frac{1}{12} \cdot b_1 b_2 b_3 \cdot N_{0,3}(b_1, b_1, b_2)$ $\cdot N_{0,3}(b_2, b_3, b_3) = \frac{16}{15} \cdot b_1 b_2 b_3 \cdot (1) \cdot (1) \xrightarrow{\mathcal{Z}}$	$\frac{16}{15} \cdot (1! \zeta(2))^3$ $= \frac{2}{405} \cdot \pi^6$
 <p><math>b_1</math></p>		

To make the calculation tractable, we follow the structure of formula (1.12). The first numerical factor  $\frac{128}{5}$  represents the factor  $\frac{2^{6g-5+2n} \cdot (4g-4+n)!}{(6g-7+2n)!}$ . It is common for all stable graphs in  $\mathcal{G}_{2,0}$ . The second numerical factor in the first line of each calculation in Table 1 is  $\frac{1}{2^{|V(\Gamma)|-1}}$ , where  $|V(\Gamma)|$  is the number of vertices of the corresponding stable graph  $\Gamma$  (equivalently—the number of connected components of the complement to the associated reduced multicurve). The third numerical factor is  $\frac{1}{|\text{Aut}(\Gamma)|}$ . Recall that neither vertices nor edges of  $\Gamma$  are labeled. We first evaluate the order of the corresponding automorphism group  $|\text{Aut}(\Gamma)|$  (this group respects the decoration of the graph), and only then we associate variables  $b_1, \dots, b_k$  to the edges of  $\Gamma$  in an arbitrary way.

Taking the sum of the six contributions we obtain the answer, matching the value found in [36] by implementing the method of Eskin and Okounkov in [32]:

$$\text{Vol } \mathcal{Q}_{2,0} = \left( \left( \frac{16}{945} + \frac{1}{2835} \right) + \left( \frac{8}{225} + \frac{1}{675} \right) + \left( \frac{1}{135} + \frac{2}{405} \right) \right) \pi^6 = \frac{1}{15} \pi^6.$$

*Remark 1.6*

In genus 0, the formula simplifies considerably. It was conjectured by Kontsevich and proved by Athreya, Eskin, and Zorich in [11] that, for all  $n \geq 4$ ,

$$\text{Vol } \mathcal{Q}_{0,n} = \text{Vol } \mathcal{Q}(1^{n-4}, -1^n) = \frac{\pi^{2n-6}}{2^{n-5}}. \tag{1.15}$$

Note that, in genus 0, all correlators of  $\psi$ -classes admit a closed explicit expression as multinomial coefficients. Rewriting all polynomials  $N_{0,n_v}(\mathbf{b}_v)$  for all stable graphs in  $\mathcal{G}_{0,n}$  in formula (1.13) for  $\text{Vol } \mathcal{Q}_{0,n}$  in terms of the corresponding multinomial coefficients, we get the formula originally obtained in [10]. A lot of techniques in this article are borrowed from [10]. Note, however, that the proof of (1.15) is indirect and is based on the analytic Riemann–Roch–Hirzebruch formula and on fine comparison of asymptotics of determinants of flat and hyperbolic Laplacians as  $(X, q)$  in  $\mathcal{Q}_{0,n}$  approaches the boundary.

*Remark 1.7*

In their recent article [17], Chen, Möller, and Sauvaget proved an alternative formula for  $\text{Vol } \mathcal{Q}(1^{4g-4+n}, -1^n)$ , expressing it as a weighted sum of certain linear Hodge integrals. Their approach is based on intersection theory. Combined with the recursive formula for the linear Hodge integrals obtained by Kazarian in [37], it allows us to compute the exact values of  $\text{Vol } \mathcal{Q}(1^{4g-4+n}, -1^n)$  for  $g$  up to 250 and, basically, for any  $n$ . In particular, it provides an alternative proof of (1.15).

Note the following important feature of (1.13) which distinguishes it from the approach of Eskin, Okounkov and Pandharipande in [31], [32] and [33] based on the

quasimodularity of certain generating functions or from the approach of Chen, Möller, and Sauvaget in [17] based on intersection theory. Formula (1.13) allows us to analyze the contribution of individual stable graphs to  $\text{Vol } \mathcal{Q}_{g,n}$ . In particular, it allows us to study the statistical geometry of random square-tiled surfaces as in Section 1.9 and the statistical properties of random simple closed hyperbolic geodesics as in Section 4.

Formula (1.13) also implies the following asymptotic lower bound for the Masur–Veech volume  $\text{Vol } \mathcal{Q}_{g,0}$  and a conjectural asymptotic value.

**THEOREM 1.8**

*The following inequality holds for any  $g \geq 2$ :*

$$\text{Vol } \mathcal{Q}_{g,0} > \sqrt{\frac{2}{3\pi g}} \cdot \left(\frac{8}{3}\right)^{4g-4} \cdot \left(1 - \frac{2}{6g-7}\right). \tag{1.16}$$

Theorem 1.8 is proved in Section 4.2.

**CONJECTURE 1.9**

*The Masur–Veech volume of the moduli space of holomorphic quadratic differentials has the following large genus asymptotics:<sup>1</sup>*

$$\text{Vol } \mathcal{Q}_{g,0} \stackrel{?}{\sim} \frac{4}{\pi} \cdot \left(\frac{8}{3}\right)^{4g-4} \cdot \left(1 + O\left(\frac{1}{g}\right)\right) \quad \text{as } g \rightarrow +\infty. \tag{1.17}$$

The exact values of  $\text{Vol } \mathcal{Q}_{g,0}$  obtained by combining results in [17] and [37] corroborate the conjectural formula (1.17). The work [60] also corroborates and develops this conjecture, suggesting several first terms of the asymptotic expansion in the negative powers of  $g$ .

A conjectural generalization of formula (1.17) to other strata of meromorphic quadratic differentials and numerical evidence beyond this conjecture are presented in [4]. The statistical geometry of random square-tiled surfaces of large genus and of random simple closed multicurves on surfaces of large genus is discussed in the separate article [22].

*Remark 1.10*

In the case of Abelian differentials, the large genus asymptotics of Masur–Veech volumes and of Siegel–Veech constants are obtained for all strata by Aggarwal [1] [2] and by Chen, Möller, Sauvaget and Zagier in [18] and [19] conforming conjectures stated by Eskin and Zorich in [34]. The large genus asymptotics of the Weil–Peterson

<sup>1</sup>After submission of the current paper, a stronger form of this conjecture was proved by Aggarwal in [3].

volumes of the moduli space of complex curves is obtained by Mirzakhani and Zograf in [50] up to a global normalization factor which remains conjectural.

*Remark 1.11*

By the result of Eskin and Okounkov in [31], the Masur–Veech volume of any stratum of Abelian differentials is a rational multiple of  $\pi^{2g}$ . For the strata of meromorphic quadratic differentials the results of [32] directly imply a slightly weaker property: the Masur–Veech volume is a polynomial in powers of  $\pi^2$  with rational coefficients (see the proof of Proposition 5.10 in [36]). Actually, one can derive from the results of [32] a stronger statement, namely, that the Masur–Veech volume of any stratum of meromorphic quadratic differentials is a rational multiple of  $\pi^{2g_{\text{eff}}}$ , where the integer number  $g_{\text{eff}}$  denotes the *effective genus* of the stratum. However, this implication is already nontrivial and was never written down. Alternative proofs of the latter statement were recently obtained by Chen, Möller, and Sauvaget in [17] for the principal stratum and by Koziarz and Nguyen in [40] for certain more general arithmetic affine  $\text{GL}(2, \mathbb{R})$ -invariant submanifolds of the strata.

By construction, the polynomial  $P_{\Gamma}(\mathbf{b})$  associated to a stable graph  $\Gamma \in \mathcal{G}_{g,n}$  by expression (1.12) is a homogeneous polynomial of degree  $6g - 6 + 2n - |E(\Gamma)|$ . Moreover, each variable  $b_e$  in each monomial appears with an odd power. Thus, (1.13) implies that the contribution  $\text{Vol}(\Gamma)$  of each stable graph  $\Gamma \in \mathcal{G}_{g,n}$  to  $\text{Vol} \mathcal{Q}_{g,n}$  is already a rational multiple of  $\pi^{6g-6+2n}$ . Moreover, using the refined version of  $N_{g,n}(b_1, \dots, b_n)$  due to Norbury [52] expressing the counting functions of ribbon graphs as quasipolynomials in  $b_i$ , one can even show that the generating series of square-tiled surfaces corresponding to a given stable graph is a quasimodular form. This result develops the results of Eskin and Okounkov in [32] that say that in each stratum, the generating series for the count of pillowcase covers (in the sense of Eskin and Okounkov) is a quasimodular form and the analogous result of Engel in [25] and [26] for the count of square-tiled surfaces.

*Remark 1.12*

Up to a normalization constant given by the explicit formula (3.17) (depending only on  $g$  and  $n$ ), the polynomial  $N_{g,n}(b_1, \dots, b_n)$  coincides with the top homogeneous part of Mirzakhani’s volume polynomial  $V_{g,n}(b_1, \dots, b_n)$  providing the Weil–Petersson volume of the moduli space of bordered Riemann surfaces (see [44]). The classical Weil–Petersson volume of  $\mathcal{M}_{g,n}$  corresponds to the constant term of  $V_{g,n}(b_1, \dots, b_n)$  when the lengths of all boundary components are equal to zero. This relation between correlators  $\langle \psi_1^{d_1} \cdots \psi_n^{d_n} \rangle$  and Weil–Petersson volumes allowed Mirzakhani to provide an alternative proof of Witten’s conjecture. In [5], a *Masur–*

*Veech* polynomial is defined whose top degree term is also a (rescaled) version of  $N_{g,n}(b_1, \dots, b_n)$  and whose constant term is the Masur–Veech volume.

Both the Weil–Petersson polynomial and the Masur–Veech polynomial satisfy a topological recursion that allows direct computations without relying on formulas such as (1.13) which involve the (huge) list of stable graphs. Note that, contrarily to the volume polynomial  $V_{g,n}$ , the topological recursion for the Masur–Veech polynomial does not admit a geometric interpretation yet.

*Remark 1.13*

Formulas (1.12)–(1.14) admit a generalization allowing one to express Masur–Veech volumes of those strata  $\mathcal{Q}(d_1, \dots, d_n)$ , for which all zeros have odd degrees  $d_i$ , in terms of intersection numbers of  $\psi$ -classes with the combinatorial cycles in  $\overline{\mathcal{M}}_{g,n}$  associated to the strata (denoted by  $W_{m_*,n}$  in [9], where  $m_*$  is the sequence of multiplicities of the zeros). In this more general case the formula requires additional correction subtracting the contribution of those quadratic differentials which degenerate to squares of globally defined Abelian differentials. This generalization is a work in progress.

*Remark 1.14*

Note that the contribution of square-tiled surfaces having a fixed number of cylinders to the Masur–Veech volume of more general strata of quadratic differentials might have a much more sophisticated arithmetic nature. In [20], we describe the contribution of square-tiled surfaces having a single maximal horizontal cylinder to the Masur–Veech volume of any stratum of Abelian or quadratic differentials. We conjecture that the contribution of  $k$ -cylinder square-tiled surfaces to the Masur–Veech volume of the ambient stratum is expressed as a polynomial with rational coefficients in multiple zeta values.

### 1.5. Siegel–Veech constants

We now turn to a formula for the Siegel–Veech constants of  $\mathcal{Q}_{g,n}$ . We first recall the definition of Siegel–Veech constant that involves the flat geometry of quadratic differentials.

Let  $(C, q)$  be a nonzero quadratic differential in  $\mathcal{Q}_{g,n}$ . It naturally defines a Riemannian metric  $|q|$  which is flat with conical singularities exactly at the zeros and poles of  $q$ . This metric allows us to define geodesics, and we say that a geodesic is *regular* if it does not pass through the singularities of  $q$ . Closed regular flat geodesics appear in families composed of parallel closed geodesics of the same length. Each such family fills a maximal flat cylinder  $\text{cyl}$  having a conical singularity (possibly the same) at each of the two boundary components. The length of any regular geodesic

in this family is called the *width* (or *circumference*) of the cylinder. The *height* of the cylinder is the distance between its boundary components measured inside the cylinder. In particular, the flat area of the cylinder is the product of its width and its height.

For  $S = (C, q)$  in  $\mathcal{Q}_{g,n} \setminus \{0\}$  and any  $L \in \mathbb{R}$ , the number of maximal cylinders in  $S$  filled with regular closed geodesics of bounded length  $w(\text{cyl}) \leq L$  is finite. Thus, the following quantity is well defined:

$$N_{\text{area}}(S, L) := \frac{1}{\text{Area}(S)} \sum_{\substack{\text{cyl} \subset S \\ w(\text{cyl}) \leq L}} \text{Area}(\text{cyl}). \tag{1.18}$$

For any pair of nonnegative integers  $(g, n)$  satisfying  $2g + n > 3$ , choose the Masur–Veech volume element  $d \text{Vol}$  in  $\mathcal{Q}_{g,n}$  which coincides with the Masur–Veech volume element on the principal stratum  $\mathcal{Q}(1^{4g-4+n}, -1^n)$ . This volume element induces a canonical volume element  $d \text{Vol}_1$  on any level hypersurface  $\mathcal{Q}_{g,n}^{\text{Area}=a}$ . The following theorem is a special case of the fundamental result of Veech in [56, theorem 0.5] developed by Vorobets in [57, theorem 1.6].

**THEOREM** (see [56], [57])

*Let  $(g, n)$  be a pair of nonnegative integers such that  $2g + n > 3$ . There exists a strictly positive constant  $c_{\text{area}}(\mathcal{Q}_{g,n})$  such that, for any strictly positive numbers  $a$  and  $L$ , the following holds:*

$$\frac{a}{\pi L^2} \int_{\mathcal{Q}_{g,n}^{\text{Area}=a}} N_{\text{area}}(S, L) d \text{Vol}_1(S) = \text{Vol}_1 \mathcal{Q}_{g,n}^{\text{Area}=a} \cdot c_{\text{area}}(\mathcal{Q}_{g,n}). \tag{1.19}$$

This formula is called the *Siegel–Veech formula*, and the corresponding constant  $c_{\text{area}}(\mathcal{Q}_{g,n})$  is called the *Siegel–Veech constant*. Note that  $c_{\text{area}}(\mathcal{Q}_{g,n})$  does not actually depend on the choice of the normalization of the Masur–Veech volume.

Eskin and Masur [29] proved that, for almost all  $S = (C, q)$  in  $\mathcal{Q}_{g,n}$  (with respect to the Masur–Veech measure),

$$\lim_{L \rightarrow +\infty} \text{Area}(S) \cdot \frac{N_{\text{area}}(S, L)}{\pi L^2} = c_{\text{area}}(\mathcal{Q}_{g,n}). \tag{1.20}$$

*Remark 1.15*

Beyond its geometrical relevance, let us mention that the area Siegel–Veech constant is the most important ingredient in the Eskin–Kontsevich–Zorich formula for the sum of the Lyapunov exponents of the Hodge bundle along the Teichmüller geodesic flow (see [28]).

An edge of a connected graph is called a *bridge* if the operation of removing this edge breaks the graph into two connected components. We define the following function  $\chi_\Gamma : E(\Gamma) \rightarrow \{\frac{1}{2}, 1\}$  on the set of edges of any connected graph  $\Gamma$ :

$$\chi_\Gamma(e) = \begin{cases} \frac{1}{2} & \text{if the edge } e \text{ is a bridge,} \\ 1 & \text{otherwise.} \end{cases}$$

We define the following operator  $\partial_\Gamma$  on polynomials  $P$  in variables  $b_e$  associated to the edges of stable graphs  $\Gamma \in \mathcal{G}_{g,n}$ . For every  $e \in E(\Gamma)$ , let

$$\partial_\Gamma^e P := \chi_\Gamma(e) b_e \frac{\partial P}{\partial b_e} \Big|_{b_e=0}, \tag{1.21}$$

and let

$$\partial_\Gamma P := \sum_{e \in E(\Gamma)} \partial_\Gamma^e P. \tag{1.22}$$

**THEOREM 1.16**

Let  $g, n$  be nonnegative integers satisfying  $2g + n > 3$ . The Siegel–Veech constant  $c_{\text{area}}(\mathcal{Q}_{g,n})$  satisfies the following relation:

$$\text{Vol } \mathcal{Q}_{g,n} \cdot c_{\text{area}}(\mathcal{Q}_{g,n}) = \frac{3}{\pi^2} \cdot \sum_{\Gamma \in \mathcal{G}_{g,n}} \mathcal{Z}(\partial_\Gamma P_\Gamma). \tag{1.23}$$

As an illustration of the above theorem we compute  $c_{\text{area}}(\mathcal{Q}_{2,0})$  and  $c_{\text{area}}(\mathcal{Q}_{1,2})$  in Appendices C.2 and C.4, respectively.

*1.6. Masur–Veech volumes and Siegel–Veech constants*

The Siegel–Veech constant can be expressed in terms of the Masur–Veech volumes of certain boundary strata. The formula for strata of Abelian differentials was obtained in [30] and for strata of quadratic differentials in [35]. Before presenting a reformulation of Corollary 1 in [35], we introduce the following conventions for  $(g, n)$  being  $(0, 3)$  or  $(1, 1)$  (see Appendix C.1 for a discussion):

$$\text{Vol } \mathcal{Q}_{0,3} := 4, \tag{1.24}$$

$$\text{Vol } \mathcal{Q}_{1,1} := \frac{2\pi^2}{3}. \tag{1.25}$$

**THEOREM ([35, Corollary 1])**

Let  $g, n$  be nonnegative integers with  $2g + n > 3$ . Under conventions (1.24)–(1.25), the following formula is valid:



$$\begin{aligned}
 c_{\text{area}}(\mathcal{Q}_{g,n}) \cdot \text{Vol } \mathcal{Q}_{g,n} &= \frac{1}{8} \sum_{\substack{g_1+g_2=g \\ n_1+n_2=n+2 \\ g_i \geq 0, n_i \geq 1, d_i \geq 1}} \frac{\ell!}{\ell_1! \ell_2!} \frac{n!}{(n_1-1)!(n_2-1)!} \\
 &\cdot \frac{(d_1-1)!(d_2-1)!}{(d-1)!} \text{Vol } \mathcal{Q}_{g_1, n_1} \times \text{Vol } \mathcal{Q}_{g_2, n_2} \\
 &+ \frac{1}{16} \cdot \frac{(4g-4+n)n(n-1)}{(6g-7+2n)(6g-8+2n)} \text{Vol } \mathcal{Q}_{0,3} \times \text{Vol } \mathcal{Q}_{g, n-1} \\
 &+ \frac{\ell!}{(\ell-2)!} \frac{(d-3)!}{(d-1)!} \text{Vol } \mathcal{Q}_{g-1, n+2}. \tag{1.26}
 \end{aligned}$$

Here  $d = \dim_{\mathbb{C}} \mathcal{Q}_{g,n} = 6g - 6 + 2n$ ,  $d_i = 6g_i - 6 + 2n_i$ ,  $\ell = 4g - 4 + n$ ,  $\ell_i = 4g_i - 4 + n_i$ .

For  $g = 0$  and any integer  $n$  satisfying  $n \geq 4$ , the following formula is valid:

$$\begin{aligned}
 c_{\text{area}}(\mathcal{Q}_{0,n}) \cdot \text{Vol } \mathcal{Q}_{0,n} &= \frac{1}{8} \sum_{\substack{n_1+n_2=n+2 \\ n_i \geq 4}} \frac{(n-4)!}{(n_1-4)!(n_2-4)!} \cdot \frac{n!}{(n_1-1)!(n_2-1)!} \\
 &\cdot \frac{(2n_1-7)!(2n_2-7)!}{(2n-7)!} \text{Vol } \mathcal{Q}_{0, n_1} \times \text{Vol } \mathcal{Q}_{0, n_2} \\
 &+ \frac{1}{16} \cdot \frac{(n-4)n(n-1)}{(2n-7)(2n-8)} \text{Vol } \mathcal{Q}_{0,3} \times \text{Vol } \mathcal{Q}_{0, n-1}. \tag{1.27}
 \end{aligned}$$

The terms which involve  $(g, n) = (0, 3)$  in the formulas above can be interpreted as particular cases of the corresponding general terms under the following convention. In the context of these formulas it is natural to define

$$\left. \frac{(d_i-1)!}{\ell_i!} \right|_{\substack{g=0 \\ n_i=3}} := \left. \frac{(2n_i-7)!}{(n_i-4)!} \right|_{n_i=3} := \lim_{n \rightarrow 3} \frac{\Gamma(2n-6)}{\Gamma(n-3)} = \frac{1}{2}. \tag{1.28}$$

Note that the expressions appearing on the right-hand sides of (1.23), (1.26), and (1.27) can be seen as polynomials in correlators. More precisely, in the definition of  $N_{g_v, n_v}(\mathbf{b}_v)$  one can keep the correlators  $\langle \psi_1^{d_1} \cdots \psi_k^{d_k} \rangle = \langle \tau_{d_1} \cdots \tau_{d_k} \rangle$  in (1.7) without evaluation. We extend the operators  $\mathcal{Z}$  and  $\partial_{\Gamma}$  to polynomials in the variables  $b_e$  and in “unevaluated” correlators by linearity. For example, under such a convention one gets

$$\begin{aligned}
 \text{Vol } \mathcal{Q}_{0,5} &= \frac{\pi^2}{9} (5\langle \tau_0^3 \tau_1 \rangle \langle \tau_0^3 \rangle + 4\langle \tau_0^3 \rangle^3) \quad \text{and} \\
 \frac{\pi^2}{3} c_{\text{area}}(\mathcal{Q}_{0,5}) \text{Vol } \mathcal{Q}_{0,5} &= \frac{5}{9} \langle \tau_0^3 \rangle^3.
 \end{aligned}$$

Numerical values of volumes  $\text{Vol } \mathcal{Q}_{g,n}$  and of Siegel–Veech constants  $c_{\text{area}}(\mathcal{Q}_{g,n})$  for small  $g$  and  $n$  are presented in Table 7 of Appendix D. The corresponding expressions in terms of the intersection numbers are available in [24].

Viewed in this way, the right-hand sides of (1.23) and (1.26) in the case  $g \geq 1$  (resp., the right-hand sides of (1.23) and (1.27) in the case  $g = 0$ ) provide identities between polynomials in intersection numbers. We show that these identities are, actually, trivial.

**THEOREM 1.17**

*The right-hand sides of (1.23) and (1.26) for  $g \geq 1$  (resp., the right-hand sides of (1.23) and (1.27) for  $g = 0$ ) considered as polynomials in intersection numbers of  $\psi$ -classes coincide.*

Theorem 1.17 is proved in Section 2.8.

*1.7. Frequencies of multicurves (after Mirzakhani)*

Let  $g, n$  be nonnegative integers with  $2g + n > 2$ . We say that two integral multicurves on the same smooth surface of genus  $g$  with  $n$  punctures have the *same topological type* if they belong to the same orbit of the mapping class group  $\text{Mod}_{g,n}$ . As we have already seen, topological types of primitive multicurves (resp., multicurves) are in bijection with stable graphs in  $\mathcal{G}_{g,n}$  (resp., stable graphs in  $\mathcal{G}_{g,n}$  labeled with a height  $H_i$  at each edge).

Let  $C$  be a complex curve,  $C \in \mathcal{M}_{g,n}$ . We denote by  $X$  the underlying Riemann surface endowed with its hyperbolic metric of constant curvature  $-1$ . Following Mirzakhani, given an integral multicurve  $\gamma$  in  $\mathcal{ML}_{g,n}(\mathbb{Z})$ , we define  $s_X(L, \gamma)$  as the number of simple closed geodesic multicurves on  $X$  of length at most  $L$  having the same topological type as  $\gamma$ . Mirzakhani analyzed the asymptotic behavior of  $s_X(L, \gamma)$  which involves several quantities that we define now.

The hyperbolic length function  $\ell_X$  defined on multicurves admits a continuous extension to  $\mathcal{ML}_{g,n}$  (see [38]). Hence, we can consider the unit ball  $B_X$  defined as

$$B_X := \{ \gamma \in \mathcal{ML} : \ell_X(\gamma) \leq 1 \} \tag{1.29}$$

in  $\mathcal{ML}_{g,n}$  and the associated volume with respect to Thurston’s measure as<sup>2</sup>

$$B(X) = \mu_{\text{Th}}(B_X).$$

<sup>2</sup>Thurston’s measure on  $\mathcal{ML}_{g,n}$  admits two natural normalizations which differ by a constant factor. Following Thurston and Mirzakhani, we use the normalization under which the set  $\mathcal{ML}_{g,n}(\mathbb{Z})$  of integral multicurves, playing the role of an integer lattice, has covolume 1 in the ambient piecewise linear space  $\mathcal{ML}_{g,n}$ . The alternative normalization is induced from the symplectic structure (see [27], [51], and Corollary 1.28 below for further details and for the value of the constant factor relating the two normalizations).

Next, we define the number  $b_{g,n}$  as the mean value of  $B(X)$

$$b_{g,n} := \int_{\mathcal{M}_{g,n}} B(X) dX. \tag{1.30}$$

Here we integrate with respect to the Weil–Petersson volume form  $dX$  on  $\mathcal{M}_{g,n}$ .

We can now state one of the main results of Mirzakhani from [47, theorem 6.1].

**THEOREM** ([47, Theorem 6.1])

*Let  $(g, n)$  be nonnegative integers with  $2g + n > 2$ . Let  $\gamma$  be a multicurve in  $\mathcal{ML}_{g,n}(\mathbb{Z})$ . Then there exists a positive constant  $c(\gamma)$  such that for any Riemann surface  $X$  of genus  $g$  with  $n$  punctures we have*

$$s_X(L, \gamma) \sim B(X) \cdot \frac{c(\gamma)}{b_{g,n}} \cdot L^{6g-6+2n},$$

as  $L \rightarrow +\infty$ .

We note that in this beautiful asymptotic formula, all information about the hyperbolic metric  $X$  is carried by the factor  $B(X)$  (which does not depend on  $\gamma$ ) and the topological information about  $\gamma$  is carried by the constant  $c(\gamma)$  (which does not depend on  $X$ ). Mirzakhani showed, furthermore, that

$$b_{g,n} = \sum_{[\gamma] \in \mathcal{P}} c(\gamma), \tag{1.31}$$

where the sum of  $c(\gamma)$  is taken with respect to representatives  $[\gamma]$  of all orbits  $\mathcal{P}$  of the mapping class group  $\text{Mod}_{g,n}$  in  $\mathcal{ML}_{g,n}(\mathbb{Z})$  as in the sum (1.5).

This allows us to interpret the ratio  $\frac{c(\gamma)}{b_{g,n}}$  as the probability of getting a multicurve of type  $\gamma$  taking a “long random” multicurve (in the same sense as the probability that the coordinates of a “random” point in  $\mathbb{Z}^2$  are coprime equals  $\frac{6}{\pi^2}$ ). More precisely, Mirzakhani showed that the asymptotic frequency  $\frac{c(\gamma)}{b_{g,n}}$  represents the density of the orbit  $\text{Mod}_{g,n} \cdot \gamma$  inside the set of all integral simple closed multicurves  $\mathcal{ML}_{g,n}(\mathbb{Z})$ . This density is analogous to the density  $\frac{6}{\pi^2}$  of integral points with coprime coordinates in  $\mathbb{Z}^2$  represented by the  $\text{SL}(2, \mathbb{Z})$ -orbit of the vector  $(1, 0)$ .

We now present a consequence of the bridge between square-tiled surfaces and multicurves that will appear in the next section. Let  $g \geq 2$ . There is a single topological type of a nonseparating simple closed curve in  $\mathcal{ML}_{g,0}$  and  $\lfloor g/2 \rfloor$  classes of separating simple closed curves. We define

$$c_{g,\text{nonsep}} := c(\gamma_{\text{nonsep}}), \quad c_{g,\text{sep}} := \sum_{[\gamma_{\text{sep}}]} c(\gamma_{\text{sep}}),$$

where  $\gamma_{\text{nonsep}}$  is the nonseparating simple closed curve and the sum in the second term is over the  $\lfloor g/2 \rfloor$  classes of separating simple closed curves.

**THEOREM 1.18**

*The frequency of separating simple closed geodesics on a closed hyperbolic surface of large genus  $g$  is exponentially small with respect to the frequency of nonseparating simple closed geodesics:*

$$\frac{c_{g,\text{sep}}}{c_{g,\text{nonsep}}} \sim \sqrt{\frac{2}{3\pi g}} \cdot \frac{1}{4g} \quad \text{as } g \rightarrow +\infty. \tag{1.32}$$

Theorem 1.18 follows from analyzing some individual contributions of particularly simple stable graphs to the Masur–Veech volume of  $\mathcal{Q}_{g,0}$  and is proved in Section 4.3. The proof is based on the large genus asymptotic formula for 2-correlators  $\langle \psi_1^{d_1} \psi_2^{d_2} \rangle$  uniform for all partitions  $d_1 + d_2 = 3g - 1$ . This formula is obtained in Section 4 using results of [61].

*Remark 1.19*

In order to go beyond the case of simple closed curves, one has to carry a much more involved asymptotic analysis of correlators that we perform in full generality in [22].

Using hyperbolic geometry and the recursion relations for Weil–Petersson volumes, Mirzakhani found an explicit expression for the coefficient  $c(\gamma)$  and for the global normalization constant  $b_{g,n}$  in terms of the intersection numbers of  $\psi$ -classes.

*Example 1.20*

For any hyperbolic metric  $X$  on a sphere with six cusps as in Figure 2, a long simple

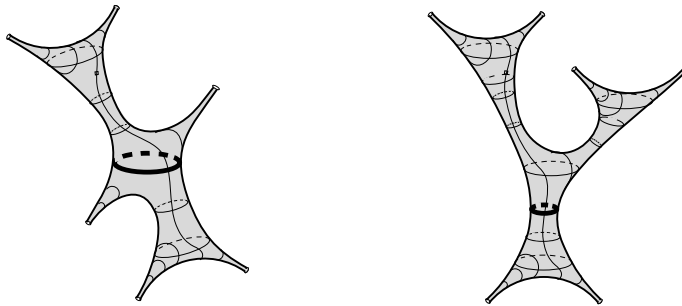


Figure 2. Simple closed curves on a six-punctured sphere.

closed geodesic separates the cusps into groups of 3 + 3 cusps with probability  $\frac{4}{7}$  and into 2 + 4 cusps with probability  $\frac{3}{7}$  (see (2) on page 123 in [47] for the calculation).

*Remark 1.21*

These values were confirmed experimentally in 2017 by Bell (see [13] for a description of the computations). They were also confirmed by a more implicit independent computer experiment by Delecroix.

*1.8. Frequencies of square-tiled surfaces of fixed combinatorial type*

The following theorem bridges both flat and hyperbolic counts. Recall from Section 1.1 that weighted stable graphs are in bijection with topological classes of multicurves.

**THEOREM 1.22**

*Let  $(g, n)$  be a pair of nonnegative integers satisfying  $2g + n > 3$  and different from  $(2, 0)$ . Let  $\gamma \in \mathcal{ML}_{g,n}(\mathbb{Z})$  be a multicurve, and let  $(\Gamma, \mathbf{H})$  be the associated stable graph and weights. Then the volume contribution  $\text{Vol}(\Gamma, \mathbf{H})$  to the Masur–Veech volume  $\text{Vol } \mathcal{Q}_{g,n}$  coincides with Mirzakhani’s asymptotic frequency  $c(\gamma)$  of closed geodesic multicurves of topological type  $\gamma$  up to an explicit factor depending only on  $g$  and  $n$ :*

$$\text{Vol}(\Gamma, \mathbf{H}) = 2 \cdot (6g - 6 + 2n) \cdot (4g - 4 + n)! \cdot 2^{4g-3+n} \cdot c(\gamma). \tag{1.33}$$

Theorem 1.22 is proved in Section 3.

*Example 1.23*

A 1-cylinder square-tiled surface in the moduli space  $\mathcal{Q}_{0,6}$  can have three simple poles on each of the two boundary components of the maximal horizontal cylinder, or it can have two simple poles on one boundary component and four simple poles on the other boundary component. The asymptotic frequency of square-tiled surfaces of the first type is  $\frac{4}{7}$  and the asymptotic frequency of the square-tiled surfaces of second type is  $\frac{3}{7}$  (cf. Example 1.20).

Combining (1.5), (1.31), and (1.33), we get the following immediate corollary.

**COROLLARY 1.24**

*For any pair of nonnegative integers satisfying  $2g + n > 3$  and different from  $(2, 0)$ , the Masur–Veech volume  $\text{Vol } \mathcal{Q}_{g,n}$  and the average Thurston measure of a unit ball  $b_{g,n}$  are related as follows:*

$$\text{Vol } \mathcal{Q}_{g,n} = 2 \cdot (6g - 6 + 2n) \cdot (4g - 4 + n)! \cdot 2^{4g-3+n} \cdot b_{g,n}. \tag{1.34}$$

*Remark 1.25*

In the case  $n = 0$ , Mirzakhani established in Theorem 1.4 of [46] a relation similar to (1.34) between  $b_g = b_{g,0}$  (computed in Theorem 5.3 of [47]) and the volume of  $\mathcal{Q}_g$ . However, Mirzakhani does not give any formula for the value of the normalization constant presented in (1.34). This constant was recently computed by Arana-Herrera [6, corollary 1.4] and Monin and Telpukhovskiy [51] simultaneously and independently of us by different methods. The same value of the constant in (1.34) is obtained by Erlandsson and Souto in [27] through an approach different from all the ones mentioned above.

Despite the fact that our main formula in Theorem 1.5 is obtained by a completely different method, it has, basically, the same structure as Mirzakhani's formula for  $b_{g,n}$ . We provide a detailed comparison of these two formulas in Section 3.

Our announcement of the explicit relation between flat and hyperbolic counts described here inspired Arana-Herrera to suggest in [6] an alternative geometric proof of these results in the spirit of Mirzakhani.

*Remark 1.26*

The pairs  $(0, 3)$ ,  $(1, 1)$ , and  $(2, 0)$  of  $(g, n)$  are exceptional for several reasons which affect, in particular, expression (1.30) for  $b_{g,n}$ . The cases of  $(0, 3)$  and  $(1, 1)$  are discussed in Appendix C.1 and Remark 3.3 (see also footnotes 1 and 2 on pages 9 and 12–13, respectively, in [59]).

The following observation distinguished the case  $(g, n) = (2, 0)$ . Any complex curve  $C$  in  $\mathcal{M}_2$  admits a hyperelliptic involution  $\tau$ . Any holomorphic quadratic differential  $q$  is invariant under this involution,  $\tau^*q = q$ . Suppose that  $q$  belongs to the principal stratum  $\mathcal{Q}(1^4)$ ; that is, suppose that  $q$  has four distinct simple zeros. Then the zeros of  $q$  are arranged into two groups of two zeros in each group, and the hyperelliptic involution interchanges the zeros in each of the two groups. Thus, there are  $4!/2$  ways to label these four zeros and not  $4!$  ways as suggests the factor  $(4g - 4 + n)!$  in the general formulas (1.33) and (1.34).

Formula (1.13) from Theorem 1.5 allows us to compute  $\text{Vol } \mathcal{Q}_{g,n}$  for all sufficiently small values of  $(g, n)$ . Remark 1.7 describes very recent methods which allow to compute  $\text{Vol } \mathcal{Q}_{g,n}$  for larger values of  $g$  and  $n$ . Corollary 1.24 thus provides explicit values of  $b_{g,n}$  for all such pairs.

When  $g = 0$ , the value  $\text{Vol } \mathcal{Q}_{0,n}$  admits closed formula (1.15) obtained in [11]. Corollary 1.24 translates this formula into the following explicit expression for  $b_{0,n}$ .

COROLLARY 1.27

The quantity  $b_{0,n}$  defined in (1.30) is equal to

$$b_{0,n} = \frac{1}{(n-3)!} \cdot \left(\frac{\pi}{2}\right)^{2(n-3)}. \tag{1.35}$$

By Stirling’s formula, we get the following asymptotics for large  $n$ :

$$b_{0,n} \sim \frac{1}{\sqrt{2\pi n}} \cdot \left(\frac{\pi^2 e}{4n}\right)^{n-3} \text{ as } n \rightarrow +\infty. \tag{1.36}$$

Mirzakhani notes in [47] (with a reference to [49]) that the frequency of simple closed curves of any fixed topological type  $\gamma$  can be described in a purely topological way as

$$\lim_{N \rightarrow +\infty} \frac{\text{card}(\{\lambda \text{ an integral multicurve} \mid \iota(\lambda, \gamma) \leq N\} / \text{Stab}(\gamma))}{N^{6g-6+2n}} = \tilde{c}(\gamma).$$

J. Souto drew our attention to the fact that Theorem 1.22 implies the following corollary from his forthcoming book with Erlandsson [27].

COROLLARY 1.28

The frequencies  $c(\gamma)$  and  $\tilde{c}(\gamma)$  are related by the following scaling factor:

$$c(\gamma) = 2^{2g-3+n} \cdot \tilde{c}(\gamma). \tag{1.37}$$

*Proof*

The subset of the set  $\{\text{an integral multicurve } \lambda \mid \iota(\lambda, \gamma) \leq N\} / \text{Stab}(\gamma)$  for which the pair  $(\lambda, \gamma)$  is not filling has cardinality  $o(N^{6g-6+2n})$ . The complementary subset for which the pair  $(\lambda, \gamma)$  is filling are images under the morphism that forgets the labeling of square-tiled surfaces in  $\mathcal{ST}_\Gamma(\mathcal{Q}(1^{4g-4+n}, -1^n))$  defined in Section 1.2. Hence (1.4) implies that

$$\text{Vol}(\Gamma, \mathbf{H}) = 2(6g - 6 + 2n) \cdot (4g - 4 + n)! \cdot 2^{6g-6+2n} \cdot \tilde{c}(\gamma),$$

where  $(\Gamma, \mathbf{H})$  is the weighted stable graph corresponding to  $\gamma$ . The factor  $(4g - 4 + n)!$  comes from possible ways to label the  $(4g - 4 + n)$  zeros, and the factor  $2^{6g-6+2n}$  comes from

$$\begin{aligned} &\text{card}(\mathcal{ST}_\gamma(\mathcal{Q}(1^{4g-4+n}, -1^n), 2N)) \\ &\sim 2^{6g-6+2n} \cdot \text{card}(\mathcal{ST}_\gamma(\mathcal{Q}(1^{4g-4+n}, -1^n), N)). \end{aligned}$$

It remains to apply (1.33). □

An alternative proof based on the result of Monin and Telpukhovskiy in [51] is suggested in [27].

1.9. *Statistical geometry of square-tiled surfaces*

Theorem 1.5 provides a detailed description of the statistical geometric properties of square-tiled surfaces in  $\mathcal{Q}_{g,n}$  tiled with a large number of squares. It has the same spirit as Mirzakhani’s result in [48, Theorem 1.2] describing statistics of lengths of simple closed geodesics in a random pants decomposition. More precisely, she fixes a reduced multicurve  $\gamma = \gamma_1 + \dots + \gamma_{3g-3}$  decomposing the surface of genus  $g$  into pairs of pants and considers its  $\text{Mod}_g$ -orbit. For any hyperbolic metric, she describes the asymptotic distribution of (normalized) lengths of simple closed geodesics represented by the components  $\gamma_i, i = 1, \dots, 3g - 3$ , of the multicurve  $\gamma$ .

Our result concerns, in particular, the asymptotic statistics of (normalized) perimeters of a random square-tiled surface corresponding to a given stable graph  $\Gamma$  tiled with a large number of squares. The resulting statistics disclose the geometric meaning of the coefficients of the polynomials  $P_\Gamma$  associated to a stable graph  $\Gamma$  appearing in our formulas for the Masur–Veech volumes as in Theorem 1.5 and for the Siegel–Veech constants as in Theorem 1.16.

As we have seen in Section 1.8, asymptotic statistical properties of random square-tiled surfaces can be translated into asymptotic statistical properties of geodesic multicurves on random hyperbolic surfaces and vice versa. This general correspondence translates the results mentioned above into analogues of a mean version of Theorem 1.2 in [48], in the sense that we obtain the average of her statistics over all hyperbolic surfaces in  $\mathcal{M}_g$ , where the average is computed using the Weil–Petersson measure on  $\mathcal{M}_g$ .

Let us define the operator  $\mathcal{X}(\mathbf{x}, \mathbf{H})$  on polynomials  $\mathbb{Q}[b_1, \dots, b_k]$  as follows. We define it on monomials

$$\mathcal{X}(\mathbf{x}, \mathbf{H})\left(\prod_{i=1}^k b_i^{m_i}\right) = (m_1 + \dots + m_k + k)! \cdot \prod_{i=1}^k \frac{x_i^{m_i}}{H_i^{m_i+1}}$$

and extend it by  $\mathbb{Q}$ -linearity. We denote by

$$\Delta^k := \{(x_1, \dots, x_k) \mid x_i \geq 0; x_1 + \dots + x_k \leq 1\}$$

the standard  $k$ -dimensional simplex. The operator  $\mathcal{X}$  generalizes both  $\mathcal{Y}(\mathbf{H})$  and  $\mathcal{Z}$  (defined by (1.10) and (1.11), respectively) in the sense that we can recover  $\mathcal{Y}(\mathbf{H})$  and  $\mathcal{Z}$  by integration:

$$\mathcal{Y}(\mathbf{H})(P) = \int_{\Delta_k} \mathcal{X}(\mathbf{x}, \mathbf{H})(P) dx_1 \cdots dx_k, \tag{1.38}$$

$$\mathcal{Z}(P) = \sum_{H_i \geq 1} \int_{\Delta_k} \mathcal{X}(\mathbf{x}, \mathbf{H})(P) dx_1 \cdots dx_k. \tag{1.39}$$



To any square-tiled surface  $S$  in  $\mathcal{ST}(\mathcal{Q}(1^{4g-4+n}, -1^n))$  we associate the following data:

$$\left(\Gamma, \mathbf{H}, \frac{\mathbf{b}}{2N}\right) \in \mathcal{G}_{g,n} \times \mathbb{N}^k \times \Delta^k,$$

where  $\Gamma$  is the stable graph associated to the horizontal cylinder decomposition,  $k$  is the number of maximal horizontal cylinders (i.e., number of edges of  $\Gamma$ ), and  $\mathbf{H} = (H_1, \dots, H_k)$  and  $\mathbf{b} = (b_1, \dots, b_k)$  are, respectively, the heights and perimeters of the maximal horizontal cylinders measured in those units, in which the square of the tiling has unit sides. Consider the set  $\mathcal{ST}(\mathcal{Q}(1^{4g-4+n}, -1^n), 2N)$  of all square-tiled surfaces in  $\mathcal{Q}(1^{4g-4+n}, -1^n)$  tiled with at most  $2N$  squares as in (1.2). Recall that  $\mathcal{ST}_{\Gamma, \mathbf{H}}(\mathcal{Q}(1^{4g-4+n}, -1^n), 2N)$  denotes the set of square-tiled surfaces associated to the stable graph  $\Gamma$  and having the vector of heights  $\mathbf{H}$  with at most  $2N$  squares.

For each stable graph  $\Gamma \in \mathcal{G}_{g,n}$ , each  $\mathbf{H} \in \mathbb{N}^k$ , where  $k = |E(\Gamma)|$ , and each  $N \in \mathbb{N}$ , we can define the following measure  $\mu_{g,n,N}^{\gamma(\Gamma, \mathbf{H})}$  on the simplex  $\Delta^k$ :

$$\mu_{g,n,N}^{\gamma(\Gamma, \mathbf{H})} := 2(6g - 6 + 2n) \cdot \frac{1}{N^d} \sum_{S \in \mathcal{ST}_{\Gamma, \mathbf{H}}(\mathcal{Q}(1^{4g-4+n}, -1^n), 2N)} \frac{1}{|\text{Aut}(S)|} \delta_{\mathbf{b}(S)/(2N)},$$

where  $\delta_{\mathbf{b}(S)/(2N)}$  is the Dirac measure concentrated at the point  $\frac{\mathbf{b}(S)}{2N} \in \Delta^k$ . We can disintegrate the discrete part  $(\Gamma, \mathbf{H})$  defining the following measure on  $\Delta^k$ :

$$\mu_{g,n,N}^\Gamma = \sum_{\mathbf{H}} \mu_{g,n,N}^{\gamma(\Gamma, \mathbf{H})}.$$

Let  $g, n$  be nonnegative integers with  $2g + n > 3$ . Let  $\mu_{g,n,N}^{\gamma(\Gamma, \mathbf{H})}$  and  $\mu_{g,n,N}^\Gamma$  be the measures defined above.

**THEOREM 1.29**

For each stable graph  $\Gamma \in \mathcal{G}_{g,n}$  and each  $\mathbf{H} \in \mathbb{N}^k$ , we have the weak convergence of measures:

$$(\mu_{g,n,N}^{\gamma(\Gamma, \mathbf{H})})_N \rightarrow \mathcal{X}(\mathbf{x}, \mathbf{H})(P_\Gamma) dx \quad \text{as } N \rightarrow +\infty. \tag{1.40}$$

Here  $P_\Gamma = P_\Gamma(b_1, \dots, b_k)$  is the global polynomial associated to the stable graph  $\Gamma$  by formula (1.12) and  $dx = dx_1 \cdots dx_k$  is the Lebesgue measure on the simplex  $\Delta^k$ .

Similarly, we have the weak convergence of measures:

$$(\mu_{g,n,N}^\Gamma)_N \rightarrow \sum_{\mathbf{H}} \mathcal{X}(\mathbf{x}, \mathbf{H})(P_\Gamma) dx. \tag{1.41}$$

Theorem 1.29 is proved in Section 2.5.

Comparing (1.38) and (1.39) with (1.14), we conclude that the total masses of the limiting measures are finite and have the following geometric meaning:

$$\int_{\Delta^k} \mathcal{X}(x, \mathbf{H})(P_\Gamma) dx = \mathcal{Y}(\mathbf{H})(P_\Gamma) = \text{Vol}(\Gamma, \mathbf{H}),$$

$$\int_{\Delta^k} \sum_{\mathbf{H}} \mathcal{X}(x, \mathbf{H})(P_\Gamma) dx = \sum_{\mathbf{H}} \int_{\Delta^k} \mathcal{X}(x, \mathbf{H})(P_\Gamma) dx = \mathcal{Z}(P_\Gamma) = \text{Vol}(\Gamma).$$

The above theorem allows us to describe some statistical properties of random square-tiled surfaces. For example, it allows us to compute the asymptotic probability that a random square-tiled surface tiled with a large number of squares corresponds to a given stable graph  $\Gamma$ . Considering only square-tiled surfaces associated to a given stable graph  $\Gamma$ , we can compute asymptotic distributions of the heights  $\mathbf{H}$  of the maximal horizontal cylinders and asymptotic distribution of their areas normalized by the area of the surface. We can also compute asymptotic statistics of perimeters of the cylinders under appropriate normalization; for example, statistics of the ratios of any two perimeters. Note that for the ratios of length variables, the unit of measurement becomes irrelevant; in particular,

$$\frac{H_i}{H_j} = \frac{h_i}{h_j} \quad \text{and} \quad \frac{b_i}{b_j} = \frac{w_i}{w_j}.$$

We will use the notation  $\mathbb{E}_\Gamma$  (resp.,  $\mathbb{E}_{\Gamma, \mathbf{H}}$ ) to denote the asymptotic expectation values of quantities evaluated on square-tiled surfaces with given cylinder decomposition associated to  $\Gamma$  (resp., associated to  $\Gamma$  and given heights  $\mathbf{H}$ ). Let us consider several simple examples.

*Example 1.30*

Consider the following stable graph  $\Phi_1$  in  $\mathcal{S}_{2,0}$  and the associated reduced multicurve:



From Table 1 we see that  $\text{Vol}(\Phi_1) = \frac{1}{675} \cdot \pi^6$  and  $\text{Vol} \mathcal{Q}_{2,0} = \frac{1}{15} \cdot \pi^6$ . Thus, a random square-tiled surface in  $\mathcal{Q}_{2,0}$  (tiled with a very large number of squares) corresponds to the stable graph  $\Phi_1$  with (asymptotic) probability  $\frac{\text{Vol}(\Phi_1)}{\text{Vol} \mathcal{Q}_{2,0}} = \frac{1}{45}$ .

We have also computed in Table 1 the polynomial  $P_{\Phi_1} = \frac{2}{15} b_1 b_2^3$ . For any given  $\mathbf{H} = (H_1, H_2)$ , we obtain

$$\mathbb{E}_{\Phi_1, \mathbf{H}} \left( \frac{b_1}{b_2} \right) = \frac{\int_{\Delta^2} \mathcal{X}(x, \mathbf{H}) \left( \frac{b_1}{b_2} P_{\Phi_1}(b_1, b_2) \right) dx_1 dx_2}{\int_{\Delta^2} \mathcal{X}(x, \mathbf{H}) (P_{\Phi_1}(b_1, b_2)) dx_1 dx_2}$$

$$\begin{aligned} &= \frac{\int_{\Delta^2} (\frac{x_1^2 x_2^2}{H_1^3 H_2^3}) dx_1 dx_2}{\int_{\Delta^2} (\frac{x_1 x_2^3}{H_1^2 H_2^4}) dx_1 dx_2} \\ &= \frac{2! \cdot 2!}{1! \cdot 3!} \frac{H_2}{H_1} = \frac{2}{3} \cdot \frac{H_2}{H_1}. \end{aligned}$$

In other words, if we force a square-tiled surface “of type  $\Phi_1$ ” to have cylinders of the same height, then the perimeter  $b_2$  of the “separating” cylinder is on average slightly longer than the perimeter  $b_1$  of the “nonseparating” cylinder. However, if we impose a large height  $H_2$  on the “separating” cylinder, then its perimeter becomes proportionally shorter on average, which is quite natural.

What might seem somehow counterintuitive is that if we do not fix  $\mathbf{H}$ , we obtain

$$\begin{aligned} \mathbb{E}_{\Phi_1} \left( \frac{b_1}{b_2} \right) &= \frac{\int_{\Delta^2} \sum_{\mathbf{H}} \mathcal{X}(\mathbf{x}, \mathbf{H}) (\frac{b_1}{b_2} P_{\Phi_1}(b_1, b_2)) dx_1 dx_2}{\int_{\Delta^2} \sum_{\mathbf{H}} \mathcal{X}(\mathbf{x}, \mathbf{H}) (P_{\Phi_1}(b_1, b_2)) dx_1 dx_2} \\ &= \frac{\int_{\Delta^2} \sum_{\mathbf{H}} (\frac{x_1^2 x_2^2}{H_1^3 H_2^3}) dx_1 dx_2}{\int_{\Delta^2} \sum_{\mathbf{H}} (\frac{x_1 x_2^3}{H_1^2 H_2^4}) dx_1 dx_2} \\ &= \frac{(\sum_{\mathbf{H} \in \mathbb{N}^2} \frac{1}{H_1^3 H_2^3}) \cdot \int_{\Delta^2} x_1^2 x_2^2 dx_1 dx_2}{(\sum_{\mathbf{H} \in \mathbb{N}^2} \frac{1}{H_1^2 H_2^4}) \cdot \int_{\Delta^2} x_1 x_2^3 dx_1 dx_2} \\ &= \frac{\zeta(3)\zeta(3)}{\zeta(2)\zeta(4)} \cdot \frac{2! \cdot 2!}{1! \cdot 3!} \approx 0.811605 \cdot \frac{2}{3} \approx 0.54107, \end{aligned}$$

while

$$\begin{aligned} \mathbb{E}_{\Phi_1} \left( \frac{b_2}{b_1} \right) &= \frac{\int_{\Delta^2} \sum_{\mathbf{H}} \mathcal{X}(\mathbf{x}, \mathbf{H}) (\frac{b_2}{b_1} P_{\Phi_1}(b_1, b_2)) dx_1 dx_2}{\int_{\Delta^2} \sum_{\mathbf{H}} \mathcal{X}(\mathbf{x}, \mathbf{H}) (P_{\Phi_1}(b_1, b_2)) dx_1 dx_2} \\ &= \frac{\int_{\Delta^2} \sum_{\mathbf{H}} (\frac{x_2^4}{H_1 H_2^5}) dx_1 dx_2}{\int_{\Delta^2} \sum_{\mathbf{H}} (\frac{x_1 x_2^3}{H_1^2 H_2^4}) dx_1 dx_2} \\ &= \frac{(\sum_{\mathbf{H} \in \mathbb{N}^2} \frac{1}{H_1 H_2^5}) \cdot \int_{\Delta^2} x_2^4 dx_1 dx_2}{(\sum_{\mathbf{H} \in \mathbb{N}^2} \frac{1}{H_1^2 H_2^4}) \cdot \int_{\Delta^2} x_1 x_2^3 dx_1 dx_2} \\ &= \frac{\zeta(1)\zeta(5)}{\zeta(2)\zeta(4)} \cdot \frac{0! \cdot 4!}{1! \cdot 3!} = +\infty. \end{aligned}$$

*Example 1.31*

Now consider the following graph  $\Phi_2$  with two edges:



It was computed in Table 1 that  $P_{\Phi_2} = \frac{8}{5}(b_1^3 b_2 + b_1 b_2^3)$ . Then for any given  $\mathbf{H} = (H_1, H_2)$ , we have

$$\mathbb{E}_{\Phi_2, \mathbf{H}} \left( \frac{b_1}{b_2} \right) = \frac{\frac{4! \cdot 0!}{H_1^5 H_2} + \frac{2! \cdot 2!}{H_1^3 H_2^3}}{\frac{3! \cdot 1!}{H_1^4 H_2^2} + \frac{1! \cdot 3!}{H_1^2 H_2^4}} = \frac{2H_2(H_1^2 + 6H_2^2)}{3H_1(H_1^2 + H_2^2)}.$$

In particular, in the symmetric case, when  $H_1 = H_2$ , we have

$$\mathbb{E}_{\Phi_2(2), \mathbf{H}} \left( \frac{b_1}{b_2} \right) = \mathbb{E}_{\Gamma, \mathbf{H}} \left( \frac{b_2}{b_1} \right) = \frac{7}{3}.$$

*Example 1.32*

Note also that we get for free the averaged version of [48, Theorem 1.2]. Namely, when the stable graph  $\Gamma \in \mathcal{G}_{g,0}$  has the maximal possible number  $3g - 3$  of vertices (i.e., when the corresponding multicurve provides a pants decomposition of the surface), equation (1.12) for  $P_\Gamma$  takes the following form:

$$P_\Gamma(\mathbf{b}) = (\text{numerical factor}) \cdot b_1 \cdots b_{3g-3} \cdot \prod_{v \in V(\Gamma)} N_{0,3}(\mathbf{b}_v).$$

Since  $N_{0,3} = 1$  identically, we conclude that for  $H_1 = H_2 = \cdots = H_{3g-3}$  the density function of the statistics of the normalized lengths is the product  $x_1 \cdots x_{3g-3}$  up to a constant normalization factor depending only on the genus  $g$ :

$$\begin{aligned} \mu_{g,n,N}^{\gamma(\Gamma, (1, \dots, 1))} &\rightarrow \mathcal{X}(\mathbf{x}, (1, \dots, 1))(C_1(g) \cdot b_1 \cdots b_{3g-3}) dx \\ &= C_2(g) \cdot x_1 \cdots x_{3g-3} dx_1 \cdots dx_{3g-3} \quad \text{as } N \rightarrow +\infty. \end{aligned}$$

Mirzakhani [48, Theorem 1.2] proved that the same asymptotic length statistics are valid for any individual hyperbolic surface in  $\mathcal{M}_g$  (and not only on average, as we do). Arana-Herrera and Liu independently proved (see [7], [8], [41]) a generalization of this result to arbitrary multicurves: namely, for any stable graph  $\Gamma \in \mathcal{G}_{g,n}$ , any associated collection of positive integer weights  $\mathbf{H}$ , and any hyperbolic surface  $X \in \mathcal{M}_{g,n}$ , the asymptotic statistics of normalized lengths of components of hyperbolic geodesic multicurves in  $\text{Mod}_{g,n} \cdot \gamma(\Gamma, \mathbf{H})$  coincide (up to a global normalization constant depending only on  $g$  and  $n$ ) with  $\mathcal{X}(\mathbf{x}, \mathbf{H})(P_\Gamma) dx$ . In particular, they do not depend on the hyperbolic metric  $X$ .

We complete this section by considering two examples describing statistics of heights of the maximal horizontal cylinders of a random square-tiled surface. We start with the following elementary lemma.

LEMMA 1.33

Consider a random square-tiled surface in  $\mathcal{Q}_{g,n}$  having a single maximal horizontal cylinder. The asymptotic probability that this cylinder is represented by a single band of squares (i.e., that  $H_1 = 1$ ) equals  $\frac{1}{\zeta(6g-6+2n)}$ .

*Proof*

When a stable graph  $\Gamma \in \mathcal{G}_{g,n}$  has a single edge  $b_1$ , formula (1.12) gives  $P_\Gamma =$  (numerical factor)  $\cdot b_1^{6g-7+2n}$ , and the lemma follows.  $\square$

In terms of multicurves this means that a random single-component integral multicurve  $n\gamma$  in  $\mathcal{ML}(\mathbb{Z})$  (where  $n \in \mathbb{N}$  and  $\gamma$  is a simple closed curve) is reduced (i.e.,  $n = 1$ ) with asymptotic probability  $\frac{1}{\zeta(6g-6+2n)}$ .

Note that  $\zeta(x)$  tends to 1 exponentially rapidly as the real-valued argument  $x$  grows. Thus, our result implies that when at least one of  $g$  or  $n$  is large enough, a random 1-cylinder square-tiled surface is tiled with a single horizontal band of squares with a very large probability, and a random single-component integral multicurve is just a simple closed curve with a very large probability.

The polynomial  $P_\Gamma$  enables us to compute analogous probabilities for any given stable graph  $\Gamma$ . For example, a random square-tiled surface in  $\mathcal{Q}_2$  associated to the stable graph  $\Phi_1$ , considered earlier in this section, has both cylinders of height  $H_1 = H_2 = 1$  with probability

$$\frac{\mathcal{Y}(1, 1)(P_{\Phi_1})}{\mathcal{Z}(P_{\Phi_1})} = \frac{1}{\zeta(2)\zeta(4)} = \frac{540}{\pi^6} \approx 0.561687.$$

A random square-tiled surface in  $\mathcal{Q}_2$  associated to the stable graph  $\Phi_2$ , considered earlier in this section, has heights  $H_i$  of both horizontal cylinders bounded by 2 with probability

$$\begin{aligned} & \frac{\mathcal{Y}(1, 1)(P_{\Phi_2}) + \mathcal{Y}(1, 2)(P_{\Phi_2}) + \mathcal{Y}(2, 1)(P_{\Phi_2}) + \mathcal{Y}(2, 2)(P_{\Phi_2})}{\mathcal{Z}(P_{\Phi_2})} \\ &= \frac{2 + \frac{2}{16} + \frac{2}{4} + \frac{2}{64}}{2\zeta(2)\zeta(4)} = \frac{\frac{85}{64}}{\zeta(2)\zeta(4)} \approx 0.745991. \end{aligned}$$

1.10. Structure of this article

The sections here are organized as follows. In Section 2, we prove Theorem 1.5 as stated in Section 1.4, providing the formula for the Masur–Veech volume  $\text{Vol } \mathcal{Q}_{g,n}$ , and we prove Theorem 1.16 as stated in Section 1.5, providing the formula for the area Siegel–Veech constant  $c_{\text{area}}(\mathcal{Q}_{g,n})$ . In Section 3, we compare our formula (1.13) for  $\text{Vol } \mathcal{Q}_{g,n}$  with Mirzakhani’s formula for  $b_{g,n}$ , and we compare our formula (1.14) for  $\text{Vol}(\Gamma)$  for a stable graph  $\Gamma$  with Mirzakhani’s formula for the associated  $c(\gamma)$  for

the associated multicurve  $\gamma$ . We elaborate the translation between the two languages and prove Theorem 1.22 as stated in Section 1.8, evaluating the normalization constant (1.33) between the corresponding quantities. In Section 4, we state a uniform asymptotic formula (4.3) for correlators  $\langle \tau_{d_1} \tau_{d_2} \rangle$ , which has independent interest. We apply it to the computation of the asymptotic frequencies  $c_{g,\text{sep}}$  and  $c_{g,\text{nonsep}}$  of separating and of nonseparating simple closed hyperbolic geodesics on a hyperbolic surface of large genus  $g$ , thus proving Theorem 1.18 stated in Section 1.7.

As for the appendices, in Appendix A we present the proof of the asymptotic formula (4.3) used in Section 4. This proof consists of a sequence of combinatorial manipulations with expressions in binomial coefficients. For the sake of completeness, we present a detailed formal definition of a stable graph in Appendix B. Appendix C provides examples of explicit calculations of the Masur–Veech volume  $\text{Vol } \mathcal{Q}_{g,n}$  and of the Siegel–Veech constant  $c_{\text{area}}(\mathcal{Q}_{g,n})$  for small  $g$  and  $n$ . Appendix D presents tables of  $\text{Vol } \mathcal{Q}_{g,n}$  and of  $c_{\text{area}}(\mathcal{Q}_{g,n})$  for small  $g$  and  $n$ .

**2. Proofs of the formulas for the Masur–Veech volumes and for the Siegel–Veech constants**

We start this section by recalling the necessary background and normalization conventions that are used in the subsequent sections of the article.

*2.1. The principal stratum and the Masur–Veech measure*

In this section, we recall the canonical construction of the Masur–Veech measure on  $\mathcal{Q}(1^{4g-4+n}, -1^n)$  and its link with the integral structure given by square-tiled surfaces.

Consider a compact nonsingular complex curve  $C$  of genus  $g$  endowed with a meromorphic quadratic differential  $q$  with  $\ell = 4g - 4 + n$  simple zeros and with  $n$  simple poles. Any such pair  $(C, q)$  defines a canonical ramified double cover  $\pi : \hat{C} \rightarrow C$  such that  $\pi^*q = \hat{\omega}^2$ , where  $\hat{\omega}$  is an Abelian differential  $\hat{\omega}$  on the double cover  $\hat{C}$ . The ramification points of  $\pi$  are exactly the zeros and poles of  $q$ . The double cover  $\hat{C}$  is endowed with the canonical involution  $\iota$  interchanging the two preimages of every regular point of the cover. The stratum  $\mathcal{Q}(1^\ell, -1^n)$  of such differentials is modeled on the subspace of the relative cohomology of the double cover  $\hat{C}$ , anti-invariant with respect to the involution  $\iota$ . This anti-invariant subspace is denoted by  $H_-^1(\hat{C}, \{\hat{P}_1, \dots, \hat{P}_\ell\}; \mathbb{C})$ , where  $\{\hat{P}_1, \dots, \hat{P}_\ell\}$  are zeros of the induced Abelian differential  $\hat{\omega}$ . For  $(g, n)$  with  $2g + n > 3$ , the image of the stratum  $\mathcal{Q}(1^\ell, -1^n)$  is open and dense in  $\mathcal{Q}_{g,n}$  (its complement in  $\mathcal{Q}_{g,n}$  is closed and of positive codimension). In what follows, we always work with the principal stratum.

We define a lattice in  $H_-^1(\hat{S}, \{\hat{P}_1, \dots, \hat{P}_\ell\}; \mathbb{C})$  as the subset of those linear forms which take values in  $\mathbb{Z} \oplus i\mathbb{Z}$  on  $H_1^-(\hat{S}, \{\hat{P}_1, \dots, \hat{P}_\ell\}; \mathbb{Z})$ . The integer points in  $\mathcal{Q}_{g,n}$

are exactly those quadratic differentials for which the associated flat surface with the metric  $|q|$  can be tiled with  $1/2 \times 1/2$  squares. In this way, the integer points in  $\mathcal{Q}(1^\ell, -1^n)$  are represented by a *square-tiled surface* as defined in Section 1.2.

We define the Masur–Veech volume element  $d \text{Vol}$  on  $\mathcal{Q}(1^\ell, -1^n)$  as the linear volume element in the vector space  $H^1(S, \{\widehat{P}_1, \dots, \widehat{P}_\ell\}; \mathbb{C})$  normalized in such a way that the fundamental domain of the above lattice has unit volume. The Masur–Veech volume element  $d \text{Vol}$  in  $\mathcal{Q}(1^\ell, -1^n)$  induces a volume element on the level sets of the Area function. In particular, on the level hypersurface  $\mathcal{Q}_{g,n}^{\text{Area}=\frac{1}{2}}$  we get

$$\begin{aligned} \text{Vol } \mathcal{Q}(1^\ell, -1^n) &:= \text{Vol}_1 \mathcal{Q}^{\text{Area}=\frac{1}{2}}(1^\ell, -1^n) \\ &= 2 \cdot d \cdot \text{Vol } \mathcal{Q}^{\text{Area}\leq\frac{1}{2}}(1^\ell, -1^n). \end{aligned} \tag{2.1}$$

Here  $d = 6g - 6 + 2n = \dim_{\mathbb{C}} \mathcal{Q}_{g,n}$ . The right-hand side term in the above formula can be considered as the definition of the middle term and the left-hand side term.

By construction, the volume element  $d \text{Vol}^{\text{symplectic}}$  in  $\mathcal{Q}_{g,n}$  induced by the canonical symplectic structure considered in Section 1.1 and the linear volume element  $d \text{Vol}^{\text{period}}$  in period coordinates defined in this section belong to the same Lebesgue measure class. It was proved by Masur in [43] that the Teichmüller flow is Hamiltonian, and in particular, that  $d \text{Vol}^{\text{symplectic}}$  is preserved by the Teichmüller flow. By the results of Masur [42] and Veech [55], the volume element  $d \text{Vol}^{\text{period}}$  is also preserved by the Teichmüller flow. Ergodicity of the Teichmüller flow now implies that the two volume forms are pointwise proportional with constant coefficient.

We postpone evaluation of this constant factor to a future paper. Throughout this work we consider the normalization of the Masur–Veech volume element  $d \text{Vol} = d \text{Vol}^{\text{period}}$  as defined in the current section and then define  $\text{Vol } \mathcal{Q}_{g,n} = \text{Vol } \mathcal{Q}(1^\ell, -1^n)$  by means of (2.1). This definition incorporates the conventions on the choice of the lattice, on the choice of the level of the area function, and the convention on the dimensional constant. We follow [10], [11], [21], [36], and [20] in the choice of these conventions (see Section 4.1 in [11] and Appendix A in [21] for the arguments in favor of this normalization).

### 2.2. Jenkins–Strebel differentials and stable graphs

A quadratic differential  $q$  in  $\mathcal{Q}_{g,n}$  is called *Jenkins–Strebel* if its horizontal foliation contains only closed leaves. Any Jenkins–Strebel differential can be decomposed into maximal horizontal cylinders with zeros and simple poles located on the boundaries of these cylinders. We call these boundaries *singular layers*. Each singular layer defines a metric ribbon graph representing an oriented surface with boundary. When the quadratic differential belongs to the principal stratum  $\mathcal{Q}(1^{4g-4+n}, -1^n)$ , the ribbon graph has vertices of valence 3 at simple zeros of  $q$ , vertices of valence 1 at simple

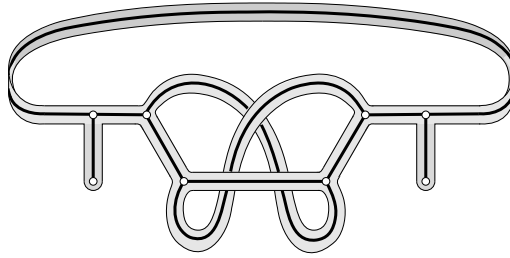


Figure 3. This ribbon graph  $G$  has genus  $g(G) = 1$ ; it has  $n(G) = 2$  boundary components and  $p(G) = 2$  univalent vertices.

poles of  $q$ , and no vertices of any other valency. Throughout the present article, we always assume that the quadratic differential  $q$  belongs to the principal stratum.

Every ribbon graph  $G$  considered as an oriented surface with boundary has a certain genus  $g(G)$ , number  $n(G)$  of boundary components, and number  $p(G)$  of univalent vertices often called *leaves* of the graph (see an example in Figure 3). The number  $m(G)$  of trivalent vertices can be expressed through these quantities as

$$m(G) = 4g(G) - 4 + 2n(G) + p(G). \quad (2.2)$$

To every Jenkins–Strebel differential  $q$  as above we associate a *stable graph*  $\Gamma = (V, H, \alpha, \iota, \mathbf{g}, L)$  in the same way as we did it in Section 1.2 in the particular case when the Jenkins–Strebel differential  $q$  represents a square-tiled surface. (A formal definition of a stable graph can be found, for example, in [53]; we reproduce it in Appendix B for completeness.) We briefly recall the construction of  $\Gamma$ .

The vertices of  $\Gamma$  encode the singular layers. The set of all vertices (singular layers) is denoted by  $V$ . A tubular neighborhood of a singular layer  $v$  is a surface with boundary, which has some genus  $g_v$ . The *genus decoration*  $\mathbf{g} = (g_v)_v$  associates to each  $v \in V$  the nonnegative integer  $g_v$ . Any maximal horizontal cylinder of  $q$  has two boundary components which are canonically identified with appropriate boundary components of tubular neighborhoods of appropriate singular layers  $v_i, v_j$  (where  $v_i$  and  $v_j$  coincide when the cylinder goes from the singular layer to itself). In this way, each maximal horizontal cylinder defines an edge of  $\Gamma$  joining the boundary layers  $v_i, v_j$ . Finally, simple poles of  $q$  are encoded by the *legs* of  $\Gamma$ . By convention the  $n$  simple poles are labeled, so the legs of  $\Gamma$  inherit the labeling  $L$ . Relation (2.2) implies the stability condition  $2g_v - 2 + n_v > 0$  for every vertex  $v$  of  $\Gamma$ .

### Remark 2.1

Consider the following trivial stable graph: it has a unique vertex decorated by the integer  $g$ ; it has  $n$  legs; it has no edges. Such a graph does not correspond to any



(nonzero) Strebel differential. For  $(g, n) \neq (0, 3)$ , it provides zero contribution in formulas (1.13) for  $\text{Vol } \mathcal{Q}_{g,n}$  and (1.23) for  $c_{\text{area}}(\mathcal{Q}_{g,n})$ .

2.3. *Conditions on the lengths of the waist curves of the cylinders*

Having a square-tiled surface and its associated stable graph  $\Gamma$ , we denote by  $k = |E(\Gamma)|$  the number of maximal cylinders filled with closed horizontal trajectories. Denote by  $w_1, \dots, w_k$  the lengths of the waist curves of these cylinders. Since every edge of any singular layer  $v$  is followed by the boundary of the corresponding ribbon graph twice, the sum of the lengths of all boundary components of each singular layer  $V$  is integral (and not only half-integral).

Let  $\Gamma$  be a stable graph, and let us consider the collection of linear forms  $l_v = \sum_{e \in E_v(\Gamma)} w_e$  in variables  $w_1, \dots, w_k$ , where  $k = |E(\Gamma)|$ ,  $v$  runs over the vertices  $V(\Gamma)$ , and  $E_v(\Gamma)$  is the set of edges adjacent to the vertex  $v$  (ignoring legs). It is immediate to see that the  $(\mathbb{Z}/2\mathbb{Z})$ -vector space spanned by all such linear forms has dimension  $|V(\Gamma)| - 1$ .

Let us make a change of variables passing from half-integer to integer parameters  $b_i := 2w_i$ , where  $i = 1, \dots, k$ . Consider the integer sublattice  $\mathbb{L}_\Gamma \subset \mathbb{Z}^k$  defined by the linear relations

$$l_v(b_1, \dots, b_k) = \sum_{e \in E_v(\Gamma)} b_e = 0 \pmod{2} \tag{2.3}$$

for all vertices  $v \in V(\Gamma)$ . By the above remark, the sublattice  $\mathbb{L}_\Gamma$  has index  $2^{|V(\Gamma)|-1}$  in  $\mathbb{Z}^k$ . We summarize the observations of this section in the following criterion.

COROLLARY 2.2

*A collection of strictly positive numbers  $w_1, \dots, w_k$ , where  $w_i \in \frac{1}{2}\mathbb{N}$  for  $i = 1, \dots, k$ , corresponds to a square-tiled surface realized by a stable graph  $\Gamma \in \mathcal{G}_{g,n}$  if and only if  $k = |E(\Gamma)|$  and the corresponding vector  $\mathbf{b} = 2\mathbf{w}$  belongs to the sublattice  $\mathbb{L}_\Gamma$ . This sublattice has index  $|\mathbb{Z}^k : \mathbb{L}_\Gamma| = 2^{|V(\Gamma)|-1}$  in the integer lattice  $\mathbb{Z}^k$ .*

We complete this section with a generalization of Lemma 3.7 in [10] which would be used in the proof of our main formula (1.14) for the Masur–Veech volume  $\text{Vol } \mathcal{Q}_{g,n}$ .

LEMMA 2.3

*Let  $\mathbb{L}$  be a sublattice of finite index  $|\mathbb{Z}^k : \mathbb{L}|$  in the integer lattice  $\mathbb{Z}^k$ , and let  $m_1, \dots, m_k \in \mathbb{N}$  be any positive integers. The following formula holds:*

$$\begin{aligned}
 & \lim_{N \rightarrow +\infty} \frac{1}{N^{|m|+k}} \sum_{\substack{\mathbf{b} \cdot \mathbf{H} \leq N \\ b_i, H_i \in \mathbb{N} \\ \mathbf{b} \in \mathbb{L}}} b_1^{m_1} \dots b_k^{m_k} \\
 &= \frac{1}{|\mathbb{Z}^k : \mathbb{L}|} \cdot \frac{1}{(|m| + k)!} \cdot \prod_{i=1}^k (m_i! \cdot \zeta(m_i + 1)) \\
 &= \frac{1}{|\mathbb{Z}^k : \mathbb{L}|} \cdot \frac{1}{(|m| + k)!} \cdot \mathcal{Z}(b_1^{m_1} \dots b_k^{m_k}). \tag{2.4}
 \end{aligned}$$

Moreover, the sum in  $\mathbf{H}$  and the limit commute,

$$\begin{aligned}
 & \lim_{N \rightarrow +\infty} \frac{1}{N^{|m|+k}} \sum_{\substack{\mathbf{b} \cdot \mathbf{H} \leq N \\ b_i, H_i \in \mathbb{N} \\ \mathbf{b} \in \mathbb{L}}} b_1^{m_1} \dots b_k^{m_k} \\
 &= \sum_{\substack{\mathbf{H} \\ H_i \in \mathbb{N}}} \lim_{N \rightarrow +\infty} \frac{1}{N^{|m|+k}} \sum_{\substack{\mathbf{b} \cdot \mathbf{H} \leq N \\ b_i \in \mathbb{N} \\ \mathbf{b} \in \mathbb{L}}} b_1^{m_1} \dots b_k^{m_k},
 \end{aligned}$$

and we have

$$\begin{aligned}
 & \lim_{N \rightarrow +\infty} \frac{1}{N^{|m|+k}} \sum_{\substack{\mathbf{b} \cdot \mathbf{H} \leq N \\ b_i \in \mathbb{N} \\ \mathbf{b} \in \mathbb{L}}} b_1^{m_1} \dots b_k^{m_k} \\
 &= \frac{1}{|\mathbb{Z}^k : \mathbb{L}|} \cdot \frac{1}{(|m| + k)!} \cdot \mathcal{Y}(\mathbf{H})(b_1^{m_1} \dots b_k^{m_k}).
 \end{aligned}$$

*Proof*

The limit with omitted restriction  $\mathbf{b} \in \mathbb{L}$  is computed in Lemma 3.7 of [10]. Note that the inversion of sum and limits here is valid by virtue of the dominated convergence theorem. More precisely, under the substitution  $x_i = \frac{b_i H_i}{N}$ , the sum approximates the integral from below.

The restriction  $\mathbf{b} \in \mathbb{L}$  rescales the volume element in the corresponding integral sum by the index  $|\mathbb{Z}^k : \mathbb{L}|$  of the sublattice  $\mathbb{L}$  in  $\mathbb{Z}^k$  which produces the extra factor  $|\mathbb{Z}^k : \mathbb{L}|^{-1}$ .

Finally, the last equality in (2.4) and the last equality in the last line of the assertion of Lemma 2.3 are just the definitions (1.11) and (1.10) of  $\mathcal{Z}$  and of  $\mathcal{Y}$ , respectively. □

2.4. Counting trivalent metric ribbon graphs with leaves

We need the following elementary generalization of Kontsevich’s theorem stated in Section 1.3, allowing our metric ribbon graph to have univalent vertices (*leaves*) in addition to trivalent vertices.

We use the letter  $p$  to denote the number of leaves. Consider a collection of positive integers  $b_1, \dots, b_n$  such that  $\sum_{i=1}^n b_i$  is even. Similarly to  $\mathcal{N}_{g,n}(b_1, \dots, b_n)$  defined in Section 1.3, let us denote by  $\mathcal{N}_{g,n,p}(b_1, \dots, b_n)$  the weighted count of connected metric ribbon graphs  $G$  of genus  $g$  with  $n$  labeled boundary components of integer lengths  $b_1, \dots, b_n$  and  $p$  univalent vertices. In other words,

$$\mathcal{N}_{g,n,p}(b_1, \dots, b_n) := \sum_{G \in \mathcal{R}_{g,n,p}} \frac{1}{|\text{Aut}(G)|} N_G(b_1, \dots, b_n),$$

where  $\mathcal{R}_{g,n,p}$  is the set of equivalence classes of ribbon graphs of genus  $g$ , with  $n$  boundary components and  $p$  univalent vertices. The counting function  $\mathcal{N}_{g,n,p}$  generalizes Kontsevich’s polynomials.

PROPOSITION 2.4

Consider  $n$ -tuples of large positive integers  $b_1, \dots, b_n$  such that  $\sum_{i=1}^n b_i$  is even. The following relation holds:

$$\mathcal{N}_{g,n,p}(b_1, \dots, b_n) = \mathcal{N}_{g,n+p,0}(b_1, \dots, b_n, \underbrace{0, \dots, 0}_p) \tag{2.5}$$

$$= \mathcal{N}_{g,n+p}(b_1, \dots, b_n, \underbrace{0, \dots, 0}_p) + \text{lower-order terms}, \tag{2.6}$$

where the Kontsevich polynomials  $N_{g,n}$  are defined by formula (1.6).

The proof of Proposition 2.4 is the combination of the following two lemmas.

LEMMA

Suppose that, for some  $g, n, p$ , the leading term of  $\mathcal{N}_{g,n,p}(b_1, \dots, b_n)$  is a homogeneous polynomial  $N_{g,n,p}^{\text{top}}(b_1, \dots, b_n)$  in  $b_1, \dots, b_n$ . Then the leading term of  $\mathcal{N}_{g,n,p+1}(b_1, \dots, b_n)$  is also a homogeneous polynomial  $N_{g,n,p+1}^{\text{top}}(b_1, \dots, b_n)$  in  $b_1, \dots, b_n$ . Moreover, it satisfies the relation

$$N_{g,n,p+1}^{\text{top}} = \frac{1}{2} \cdot \mathcal{J}(N_{g,n,p}^{\text{top}}), \tag{2.7}$$

where  $\mathcal{J} = \sum_{i=1}^n \mathcal{J}_{b_i}$ , and the operators  $\mathcal{J}_{b_i}$  are defined on monomials by

$$\mathcal{J}_{b_i}(b_1^{j_1} \dots b_{i-1}^{j_{i-1}} \cdot b_i^{j_i} \cdot b_{i+1}^{j_{i+1}} \dots b_n^{j_n}) = b_1^{j_1} \dots b_{i-1}^{j_{i-1}} \cdot \frac{b_i^{j_i+2}}{j_i + 2} \cdot b_{i+1}^{j_{i+1}} \dots b_n^{j_n}$$

and are extended to arbitrary polynomials by linearity.

*Proof*

This lemma mimics Lemma 3.5 in [10]. Formally speaking, in [10] the corresponding statement is formulated only for  $g = 0$ , but it is immediate to see that the inductive proof is applicable without any changes to any genus as soon as the base of induction, corresponding to  $p = 0$ , is valid.  $\square$

LEMMA

The polynomials  $N_{g,n}$  defined by equation (1.6) satisfy the relations:

$$N_{g,n+p+1}(b_1, \dots, b_n, 0, \dots, 0) = \frac{1}{2} \cdot \mathcal{J}(N_{g,n+p}(b_1, \dots, b_n, 0, \dots, 0)). \tag{2.8}$$

*Proof*

Using the explicit expressions  $N_{g,n+p+1}$  and  $N_{g,n+p}$  in terms of  $\psi$ -classes, we see that the above relation is equivalent to

$$\begin{aligned} \langle \psi_1^{d_1} \psi_2^{d_2} \dots \psi_n^{d_n} \rangle_{g,n+p+1} &= \langle \psi_1^{d_1-1} \psi_2^{d_2} \dots \psi_n^{d_n} \rangle_{g,n+p} \\ &\quad + \langle \psi_1^{d_1} \psi_2^{d_2-1} \dots \psi_n^{d_n} \rangle_{g,n+p} + \dots \\ &\quad + \langle \psi_1^{d_1} \psi_2^{d_2} \dots \psi_n^{d_n-1} \rangle_{g,n+p}, \end{aligned}$$

which is the well-known identity for the intersection numbers known as *string equation* (see [58]).  $\square$

*Proof of Proposition 2.4*

For  $p = 0$  (when there are no poles at all), the statement corresponds to the original theorem of Kontsevich stated in Section 1.3. We use this as the base of induction in  $p$  for any fixed pair  $(g, n)$ . It remains to notice that (2.7) and (2.8) recursively define the corresponding polynomials for any  $(g, n, p)$  starting from  $(g, n, 0)$ .  $\square$

2.5. *Proof of the volume formula*

A square-tiled surface corresponding to a fixed stable graph  $\Gamma$  can be described by three groups of parameters. The parameters in different groups can be varied independently. The parameters in the first group are responsible for the lengths of horizontal saddle connections. In this group we fix only the lengths  $w_1, \dots, w_k$  of the waist curves of the cylinders filled with closed horizontal trajectories, where  $k$  is the number of edges in  $\Gamma$ . This leaves a certain freedom for the choice of the lengths of horizontal saddle connections. The criterion of admissibility of a given collection  $\mathbf{w} = (w_1, \dots, w_k)$  is given by Corollary 2.2. The count for the number of choices of

the lengths of all individual saddle connections for a fixed choice of  $\mathbf{w}$  is given in Proposition 2.4.

There are no restrictions on the choice of strictly positive integer or half-integer heights  $h_1, \dots, h_k$  of the cylinders.

Having chosen the widths  $w_1, \dots, w_k$  of all maximal cylinders (i.e., the lengths of the closed horizontal trajectories) and the heights  $h_1, \dots, h_k$  of the cylinders, the flat area of the entire surface is already uniquely determined as the sum  $\mathbf{w} \cdot \mathbf{h} = w_1h_1 + \dots + w_kh_k$  of flat areas of individual cylinders.

However, when the lengths of all horizontal saddle connections and the heights  $h_i$  of all cylinders are fixed, there is still some freedom in the third independent group of parameters. Namely, we can twist each cylinder by some twist  $\phi_i \in \frac{1}{2}\mathbb{N}$  before attaching it to the layer. So, applying if necessary the appropriate Dehn twist, we can assume that  $0 \leq \phi_i < w_i$ , where  $w_i$  is the perimeter (length of the waist curve) of the corresponding cylinder. Thus, for any choice of lengths of horizontal saddle connections realizing some square-tiled surface with the stable graph  $\Gamma$  and for any choice  $h_1, \dots, h_k$  of heights of the cylinders, we obtain  $(2w_1) \cdots (2w_k)$  square-tiled surfaces sharing the same lengths of the horizontal saddle connections and the same heights of the cylinders.

In Proposition 2.4, we assume that the lengths of the edges of the metric ribbon graph are integer. Clearly, if we allow these lengths to also be half-integer, then we get  $N_{g,n+p}(2w_1, \dots, 2w_n, 0, \dots, 0)$  as the leading term of the new count. The realizability condition  $2\mathbf{w} \in \mathbb{L}_\Gamma$  from Corollary 2.2 translates as the compatibility condition of the parity of the sum of the lengths of the boundary components of each individual connected ribbon graph as in Proposition 2.4.

We are ready to write a formula for the leading term in the number of all square-tiled surfaces tiled with at most  $2N$  squares represented by the stable graph  $\Gamma$  when the integer bound  $N$  becomes sufficiently large:

$$\begin{aligned} &\text{card}(\mathcal{ST}_\Gamma(\mathcal{Q}(1^{4g-4+n}, 2N))) \\ &\sim (4g - 4 + n)! \cdot \frac{1}{|\text{Aut}(\Gamma)|} \cdot \sum_{\substack{\mathbf{w} \cdot \mathbf{h} \leq N/2 \\ w_j, h_i \in \frac{1}{2}\mathbb{N} \\ 2\mathbf{w} \in \mathbb{L}_\Gamma}} (2w_1) \cdots (2w_k) \cdot \prod_{v \in V(\Gamma)} N_{g_v, n_v}(2\mathbf{w}_v). \end{aligned}$$

The notation in the above expression mimics notation in (1.13): namely,  $k = |E(\Gamma)|$ , and  $\mathbf{w}_v$  is defined analogously to  $\mathbf{b}_v$  in (1.13). The factor  $(4g - 4 + n)!$  represents the number of ways to label the  $4g - 4 + n$  trivalent vertices of the ribbon graphs, which correspond to  $\ell = 4g - 4 + n$  simple zeros of the corresponding Strebel quadratic differential  $q$ . Note that by convention the univalent vertices (leaves) (corresponding to simple poles of  $q$  and also to  $n$  marked points) are already labeled.

Making a change of variables  $b_h := (2w_h) \in \mathbb{N}$  and  $H_i := (2h_i) \in \mathbb{N}$ , we can rewrite the above expression as

$$\begin{aligned} &\text{card}(\mathcal{ST}_\Gamma(\mathcal{Q}(1^{4g-4+n}, -1^n), 2N)) \\ &\sim (4g - 4 + n)! \cdot \frac{1}{|\text{Aut}(\Gamma)|} \cdot \sum_{\substack{\mathbf{b} \cdot \mathbf{H} \leq 2N \\ b_i, H_i \in \mathbb{N} \\ \mathbf{b} \in \mathbb{L}^k}} b_1 \cdots b_k \cdot \prod_{v \in V(\Gamma)} N_{g_v, n_v}(\mathbf{b}_v). \end{aligned} \tag{2.9}$$

The expression above is a homogeneous polynomial of degree  $6g - 6 + 2n - k$ . For any individual monomial the corresponding sum was evaluated in Lemma 2.3. It remains to adapt formula (2.4) to our specific context.

By Corollary 2.2, the sublattice  $\mathbb{L}_\Gamma$  in the above formula has index  $2^{|V(\Gamma)|-1}$  in  $\mathbb{Z}^k$ . The corresponding factor  $1/2^{|V(\Gamma)|-1}$  appears as the first factor in the second line of definition (1.12) of  $P_\Gamma(\mathbf{b})$ .

The degree of the homogeneous polynomial denoted by  $|m|$  in (2.4) equals in our case  $6g - 6 + 2n - k$ , so

$$|m| + k = 6g - 6 + 2n = \dim_{\mathbb{C}} \mathcal{Q}(1^{4g-4+n}, -1^n) = d.$$

Note also that in (2.4) we perform the summation under the condition  $\mathbf{b} \cdot \mathbf{H} \leq N$ , while in the above formula we sum over the region  $\mathbf{b} \cdot \mathbf{H} \leq 2N$ . This provides an extra factor  $2^d$ . Finally, passing from  $\text{card}(\mathcal{ST}_\Gamma(\mathcal{Q}(1^{4g-4+n}, -1^n), 2N))$  in (2.9) to  $\text{Vol } \mathcal{Q}_{g,n}$  by (1.2) we introduce the extra factor  $2(6g - 6 + 2n)$ . The resulting product factor

$$2(6g - 6 + 2n) \cdot \frac{2^d}{d!} \cdot (4g - 4 + n)! = \frac{2^{6g-5+2n} \cdot (4g - 4 + n)!}{(6g - 7 + 2n)!}$$

is the factor in the first line of definition (1.12) of  $P_\Gamma(\mathbf{b})$ .

Theorem 1.5 is proved. □

The lemma below is a straightforward combination of Lemma 2.3 stated in terms of densities and of Theorem 1.5. This lemma corresponds to relation (1.41) from Theorem 1.29. The other statements of Theorem 1.29 are proved analogously.

LEMMA 2.5

Let  $F : \Delta^k \times \mathbb{N}^k \rightarrow \mathbb{R}$  be a continuous function integrable with respect to the density  $\mathcal{X}(b_1^{m_1} \cdots b_k^{m_k}) \delta_{\mathbf{H}} d\mathbf{x}$  defined in (1.41). Let  $\mathbb{L}$  be a lattice of finite index in  $\mathbb{Z}^k$ . Then

$$\lim_{N \rightarrow +\infty} \frac{1}{N^{|m|+k}} \sum_{\substack{\mathbf{b} \cdot \mathbf{H} \leq 2N \\ b_i, H_i \in \mathbb{N} \\ \mathbf{b} \in \mathbb{L}}} F\left(\left(\frac{b_1 H_1}{2N}, \dots, \frac{b_k H_k}{2N}\right), \mathbf{H}\right) b_1^{m_1} \cdots b_k^{m_k}$$

$$= \frac{1}{|\mathbb{Z}^k : \mathbb{L}|} \cdot \sum_{\mathbf{H}} \int_{\Delta^k} F(\mathbf{x}, \mathbf{H}) \mathcal{X}(b_1^{m_1} \dots b_k^{m_k}) d\mathbf{x}.$$

2.6. *Yet another expression for the Siegel–Veech constant*

For any square-tiled surface  $S$ , define the following quantity. Suppose that  $S$  has  $k$  maximal cylinders filled with closed horizontal trajectories. Denote as usual by  $w_i$  the length of the closed horizontal trajectory (length of the waist curve) of the  $i$ th cylinder, and denote by  $h_i$  its height. Define

$$M(S) := \sum_{i=1}^k \frac{h_i}{w_i}.$$

The  $GL(2, \mathbb{R})$ -orbit of any square-tiled surface  $S$  is a closed invariant submanifold  $\mathcal{L}(S)$  in the ambient stratum of quadratic (or Abelian) differentials. In the same way in which we defined in Section 1.5 the area Siegel–Veech constant  $c_{\text{area}}(\mathcal{Q}_{g,n})$  for  $\mathcal{Q}_{g,n}$ , we can define the area Siegel–Veech constant  $c_{\text{area}}(\mathcal{L}(S))$  for  $\mathcal{L}(S)$ . It satisfies, in particular, analogues of (1.19) and (1.20). Theorem 4 in [28] proves the following assertion.

THEOREM ([28, Theorem 4])

For any square-tiled surface  $S$ , its Siegel–Veech constant satisfies

$$c_{\text{area}}(\mathcal{L}(S)) = \frac{3}{\pi^2} \cdot \frac{1}{\text{card}(\text{SL}(2, \mathbb{Z}) \cdot S)} \sum_{S_i \in \text{SL}(2, \mathbb{Z}) \cdot S} M(S). \tag{2.10}$$

Recall that we denote by  $\mathcal{ST}(\mathcal{Q}(1^{4g-4+n}, -1^n), N)$  the set of square-tiled surfaces in the principal stratum  $\mathcal{Q}(1^{4g-4+n}, -1^n)$  tiled with at most  $N$  squares. Define now the quantity:

$$\mathcal{DST}(\mathcal{Q}(1^{4g-4+n}, -1^n), N) := \sum_{S \in \mathcal{ST}(\mathcal{Q}(1^{4g-4+n}, -1^n), N)} M(S).$$

The latter quantity has the following geometric interpretation. Consider all square-tiled surfaces obtained from square-tiled surfaces as above by cutting exactly one cylinder  $\text{cyl}_i$  along the closed horizontal trajectory at the level  $h$ , where  $0 \leq h < h_i$  and  $h$  is integer in the case of Abelian differentials and half-integer in the case of quadratic differentials. The above quantity enumerates bordered square-tiled surfaces obtained in this way (and “ $\mathcal{D}$ ” in the notation “ $\mathcal{DST}$ ” is an allusion to the boundary). Indeed, we lose the twist parameter  $w_i$  (correspondingly  $2w_i$ ) along the cylinder which is now cut open, but we gain the new height parameter  $h_i$  (correspondingly  $2h_i$ ) responsible for the level of the cut.

As a corollary of the above theorem, Chen and Eskin proved the following result.

THEOREM ([16, Appendix A])

The area Siegel–Veech constant  $c_{\text{area}}(\mathcal{Q}_{g,n}) = c_{\text{area}}(\mathcal{Q}(1^{4g-4+n}, -1^n))$  has the following value:

$$c_{\text{area}}(\mathcal{Q}(1^{4g-4+n}, -1^n)) = \frac{3}{\pi^2} \cdot \lim_{N \rightarrow \infty} \frac{\mathcal{D}\mathcal{S}\mathcal{T}(\mathcal{Q}(1^{4g-4+n}, -1^n), 2N)}{\text{card}(\mathcal{S}\mathcal{T}(\mathcal{Q}(1^{4g-4+n}, -1^n), 2N))}. \tag{2.11}$$

Formally speaking, the original theorem is proved only for the components of the strata of Abelian differentials. However, all the arguments are applicable to any arithmetic  $\text{GL}(2, \mathbb{R})$ -invariant submanifold, in particular to the loci induced from strata of quadratic differentials by the ramified double covering construction.

An alternative way to derive (2.11) is to use the following more elaborate technique. Neglecting exceptional orbits of square-tiled surfaces in  $\mathcal{Q}(1^{4g-4+n}, -1^n)$  containing negligibly small number of square-tiled surfaces, we can arrange other orbits in a sequence of affine invariant manifolds for which the natural  $\text{SL}(2, \mathbb{R})$ -invariant measures supported on the orbits converge to the invariant measure of the ambient stratum. By Theorem 2.8 from [14], all individual Lyapunov exponents of affine invariant manifolds converge to the Lyapunov exponents of the ambient stratum. Results from [28] now imply that the area Siegel–Veech constants of the corresponding arithmetic Teichmüller disks converge to the area Siegel–Veech constant of the ambient stratum, which implies (2.11).

2.7. Proof of the formula for the area Siegel–Veech constant

We have already evaluated the denominator in (2.11). Evaluating the numerator following the lines of the initial computation, we reduce the problem to the evaluation of the sum (2.9) counted with the weight  $\frac{h_i}{w_i} = \frac{H_i}{b_i}$ , for each  $i = 1, \dots, k$ , where we use the notation of formula (2.9):

$$\sum_{\substack{\mathbf{b} \cdot \mathbf{H} \leq 2N \\ b_j, H_j \in \mathbb{N} \\ \mathbf{b} \in \mathbb{L}_\Gamma}} \left( b_1 \cdots b_k \cdot \prod_{v \in V(\Gamma)} N_{g_v, n_v}(\mathbf{b}_v) \right) \cdot \frac{H_i}{b_i}. \tag{2.12}$$

The numerator in (2.11) is the sum of the above expressions with respect to the summation index  $i$  varying from 1 to  $k$ .

Denote by  $P(b_1, \dots, b_k)$  the homogeneous polynomial  $\prod N_{g_v, n_v}(\mathbf{b}_v)$  in the formula above. It is easy to see that the condition  $b_i H_i < \mathbf{b} \cdot \mathbf{H} \leq 2N$  implies that the contribution of any monomial of  $P(b_1, \dots, b_k)$  containing the variable  $b_i$  to the above sum is of order  $o(N^d)$ , so it does not contribute to the limit (2.11). Thus, up to lower-order terms, the sum (2.12) coincides with the sum



$$\sum_{\substack{\mathbf{b} \cdot \mathbf{H} \leq 2N \\ b_j, H_j \in \mathbb{N} \\ \mathbf{b} \in \mathbb{L}_\Gamma}} b_1 \cdots b_{i-1} \cdot H_i \cdot b_{i+1} \cdots b_k \cdot P(b_1, \dots, b_{i-1}, 0, b_{i+1}, \dots, b_k). \tag{2.13}$$

It is sufficient to interchange the notation  $b_i$  and  $H_i$  to see that

$$\begin{aligned} &\sum_{\substack{\mathbf{b} \cdot \mathbf{H} \leq 2N \\ b_i, H_i \in \mathbb{N}}} b_1 \cdots b_{i-1} \cdot H_i \cdot b_{i+1} \cdots b_k \cdot P(b_1, \dots, b_{i-1}, 0, b_{i+1}, \dots, b_k) \\ &= \sum_{\substack{\mathbf{b} \cdot \mathbf{H} \leq 2N \\ b_i, H_i \in \mathbb{N}}} b_1 \cdots b_k \cdot P(b_1, \dots, b_{i-1}, 0, b_{i+1}, \dots, b_k). \end{aligned}$$

We already know how to evaluate the latter sum, so it remains to study the impact of the extra condition  $\mathbf{b} \in \mathbb{L}_\Gamma$  present in the sum (2.13).

Recall the strategy of evaluation of the sum (2.13) (see the proof of analogous Lemma 3.7 in [10] for reduction to integral sums and the proof of Lemma 2.3 for the impact of the sublattice condition). The variables  $H_1, \dots, H_{i-1}, b_i, H_{i+1}, \dots, H_k$  are considered as parameters. For each collection of such parameters we evaluate the corresponding integral sum over a simplex in the  $k$ -dimensional space with coordinates  $b_1, \dots, b_{i-1}, H_i, b_{i+1}, \dots, b_k$ . After that, we perform summation with respect to the parameters  $H_1, \dots, H_{i-1}, b_i, H_{i+1}, \dots, H_k$ .

When the edge of the graph  $\Gamma$  corresponding to the variable  $b_i$  is a *bridge* (i.e., when this edge is separating), the *parameter*  $b_i$  is always even. The space of integration now has coordinates  $b_1, \dots, b_{i-1}, H_i, b_{i+1}, \dots, b_k$ ; the sublattice  $\mathbb{L}_\Gamma$  in it is defined by the system of equations (2.3) where we let  $b_i = 0$ . Such a sublattice has index  $2^{|V(\Gamma)|-2}$  and not  $2^{|V(\Gamma)|-1}$  as before. Thus, on the level of integration, we gain the factor 2 with respect to the initial count. However, since the *parameter*  $b_i$  is now always even, evaluating the corresponding sum with respect to possible values of this parameter we obtain the sum

$$\frac{1}{2^2} + \frac{1}{4^2} + \frac{1}{6^2} \cdots = \frac{1}{4} \cdot \zeta(2)$$

instead of the original sum

$$\frac{1}{1^2} + \frac{1}{2^2} + \frac{1}{3^2} \cdots = \zeta(2).$$

Thus, when  $b_i$  corresponds to a bridge (i.e., to a separating edge), we obtain

$$\sum_{\substack{\mathbf{b} \cdot \mathbf{H} \leq 2N \\ b_j, H_j \in \mathbb{N} \\ \mathbf{b} \in \mathbb{L}_\Gamma}} b_1 \cdots b_{i-1} \cdot H_i \cdot b_{i+1} \cdots b_k \cdot P(b_1, \dots, b_{i-1}, 0, b_{i+1}, \dots, b_k)$$

$$= \frac{1}{2} \cdot \sum_{\substack{\mathbf{b} \cdot \mathbf{H} \leq 2N \\ b_j, H_j \in \mathbb{N} \\ \mathbf{b} \in \mathbb{L}_\Gamma}} b_1 \cdots b_k \cdot P(b_1, \dots, b_{i-1}, 0, b_{i+1}, \dots, b_k).$$

When  $b_i$  corresponds to a nonseparating edge, the *parameter*  $b_i$  in the sum (2.13) can take even and odd values. The new space of integration has coordinates  $b_1, \dots, b_{i-1}, H_i, b_{i+1}, \dots, b_k$ , where the sublattice in it is defined by the system of equations (2.3) in which we substitute  $b_i = 0$  or  $b_i = 1$  depending on the parity of the value of the parameter  $b_i$ . The sublattice is linear in the first case and affine in the second case. Such a sublattice has index  $2^{|V(\Gamma)|-1}$  as before. Thus, when  $b_i$  corresponds to a nonseparating edge, we get

$$\begin{aligned} & \sum_{\substack{\mathbf{b} \cdot \mathbf{H} \leq 2N \\ b_j, H_j \in \mathbb{N} \\ \mathbf{b} \in \mathbb{L}_\Gamma}} b_1 \cdots b_{i-1} \cdot H_i \cdot b_{i+1} b_k \cdot P(b_1, \dots, b_{i-1}, 0, b_{i+1}, \dots, b_k) \\ &= \sum_{\substack{\mathbf{b} \cdot \mathbf{H} \leq 2N \\ b_j, H_j \in \mathbb{N} \\ \mathbf{b} \in \mathbb{L}_\Gamma}} b_1 \cdots b_k \cdot P(b_1, \dots, b_{i-1}, 0, b_{i+1}, \dots, b_k). \end{aligned}$$

We have proved that

$$\begin{aligned} & \sum_{\substack{\mathbf{b} \cdot \mathbf{H} \leq 2N \\ b_j, H_j \in \mathbb{N} \\ \mathbf{b} \in \mathbb{L}_\Gamma}} \left( b_1 \cdots b_k \cdot \prod_{v \in V(\Gamma)} N_{g_v, n_v}(\mathbf{b}_v) \right) \cdot \left( \sum_{i=1}^k \frac{H_i}{b_i} \right) \\ &= \sum_{\substack{\mathbf{b} \cdot \mathbf{H} \leq 2N \\ b_j, H_j \in \mathbb{N} \\ \mathbf{b} \in \mathbb{L}_\Gamma}} b_1 \cdots b_k \cdot \mathcal{D}_\Gamma \left( \prod_{v \in V(\Gamma)} N_{g_v, n_v}(\mathbf{b}_v) \right) + \text{lower-order terms,} \end{aligned}$$

where the operator  $\mathcal{D}_\Gamma$  is defined in (1.22). Applying to the latter sum the technique used in the proof of Theorem 1.5, we complete the proof of Theorem 1.16. □

### 2.8. Equivalence of two expressions for the Siegel–Veech constant

In this section we prove Theorem 1.17. We start the proof by establishing a natural correspondence between the summands of the two expressions. For any stable graph  $\Gamma \in \mathcal{G}_{g,n}$  and any edge  $e$  of  $\Gamma$ , we define a combinatorial surgery  $\text{Cut}_e \Gamma$  of  $\Gamma$ . We describe it separately in the case when  $e$  is a bridge (i.e., a separating edge), and when it is not.

We start with the case when  $e$  is a bridge. Cut the edge  $e$  transforming it into two legs. Assign index 1 to one of the resulting graphs, and index 2 to the remaining

one. We do not modify the genus decoration of the vertices. The set of vertices  $V(\Gamma)$  gets naturally partitioned into two complementary subsets  $V = V_1 \sqcup V_2$ . Define  $g_i = \sum_{v \in V_i} g_v$  for  $i = 1, 2$ . Similarly, the  $n$  original legs are partitioned into  $n_1$  legs which go to  $\Gamma_1$  and into  $n_2$  legs which go to  $\Gamma_2$ . For  $i = 1, 2$ , relabel the  $n_i$  legs of  $\Gamma_i$  to the consecutive labels  $1, 2, \dots, n_i$ , preserving the order of labels. Assign the label  $n_i + 1$  to the new leg of  $\Gamma_i$  created during the surgery. The stability condition  $2g_v - 2 + n_v > 0$ , which is valid for every vertex  $v$  of  $\Gamma$ , implies that we get two stable graphs  $\Gamma_i \in \mathcal{G}_{g_i, n_i + 1}$ .

The only ambiguity in this construction is the choice of the label (1 or 2) for one of the components  $\Gamma_i$  of the graph  $\Gamma$  with removed bridge  $e$ . In general, there are two choices, except the case when there is a symmetry of  $\Gamma$  acting on the edge  $e$  as a flip (i.e., a symmetry which sends  $e$  to itself exchanging its two ends).

Note that the surgery is reversible in the following sense. Given two stable graphs  $\Gamma_i \in \mathcal{G}_{g_i, n_i + 1}$ , we can glue the endpoint of the leg with index  $n_1 + 1$  of  $\Gamma_1$  to the endpoint of the leg with index  $n_2 + 1$  of  $\Gamma_2$  creating a connected graph with  $n = n_1 + n_2$  legs and with an extra bridge joining  $\Gamma_1$  to  $\Gamma_2$ . The only ambiguity in this construction is in relabeling the  $n = n_1 + n_2$  legs to a consecutive list  $(1, 2, \dots, n)$ ; there are  $\binom{n}{n_1}$  ways to do it.

We now describe the surgery  $\text{Cut}_e \Gamma$  in the remaining case when the edge  $e$  of  $\Gamma$  is not a bridge (i.e., is not separating). Cutting such an edge, we transform it into two legs. We keep the same labels for the preexisting legs, and we associate labels  $n + 1$  and  $n + 2$  to the two created legs. In general, there are two ways to do that, except when there is a symmetry of  $\Gamma$  acting on the edge  $e$  as a flip (i.e., a symmetry which sends  $e$  to itself exchanging its two ends). We obtain a stable graph  $\Gamma' \in \mathcal{G}_{g-1, n+2}$ .

The inverse operation (applicable to stable graphs with at least two legs) consists in gluing the two legs of higher index together, transforming them into an edge and keeping the same labeling for the other legs. Once again we do not modify the genus decoration of the vertices.

Note that the operator  $\partial_\Gamma$  is defined in (1.22) as a sum  $\sum \partial_\Gamma^e$  over edges  $e \in E(\Gamma)$  of a stable graph  $\Gamma$ . Thus, our key sum  $\sum_{\Gamma \in \mathcal{G}_{g,n}} \mathcal{Z}(\partial_\Gamma P_\Gamma)$  on the right-hand side of (1.23) in Theorem 1.16 can be seen as the sum over all pairs  $(\Gamma, e)$ , where  $\Gamma \in \mathcal{G}_{g,n}$  and  $e \in E(\Gamma)$ . We show below that for every such pair  $(\Gamma, e)$ , the corresponding term of the resulting sum has simple expression in terms of the product  $\mathcal{Z}(P_{\Gamma_1})\mathcal{Z}(P_{\Gamma_2})$  when  $e$  is a bridge and in terms of  $\mathcal{Z}(P_{\Gamma'})$  when  $e$  is not a bridge, where  $\Gamma_1 \sqcup \Gamma_2$  (resp.,  $\Gamma'$ ) are the stable graphs obtained under applying the surgery  $\text{Cut}_e \Gamma$ .

Having a stable graph  $\Gamma$ , we associate to it the polynomial

$$\Pi_\Gamma(\mathbf{b}) := \prod_{v \in V(\Gamma)} N_{g_v, n_v}(\mathbf{b}_v).$$

By Definition (1.12) of  $P_\Gamma(\mathbf{b})$ , we have

$$P_\Gamma(\mathbf{b}) = (\text{combinatorial factor}) \cdot \left( \prod_{e \in E(\Gamma)} b_e \right) \cdot \Pi_\Gamma(\mathbf{b}).$$

A pair  $(\Gamma, e_0)$  provides a nonzero contribution to the sum on the right-hand side of (1.23) if and only if the term  $\partial_\Gamma^{e_0} = \chi_\Gamma(e_0) b_{e_0} \cdot \frac{\partial}{\partial b_{e_0}}|_{b_{e_0}=0}$  in the operator  $\partial_\Gamma$  applied to  $(\prod_{e \in E(\Gamma)} b_e) \cdot \Pi_\Gamma(\mathbf{b})$  does not identically vanish (see (1.21)). The latter is equivalent to the condition that the polynomial  $\Pi_\Gamma|_{b_{e_0}=0}$  does not identically vanish.

If the edge  $e_0$  is a bridge, consider the stable graphs  $\Gamma_1, \Gamma_2$  obtained under the surgery  $\text{Cut}_{e_0} \Gamma$ . The polynomial  $\Pi_\Gamma|_{b_{e_0}=0}$  splits naturally into the product  $\Pi_\Gamma|_{b_j=0} = \Pi_{\Gamma_1} \Pi_{\Gamma_2}$ , so when  $e_0$  is a bridge, and when  $\Pi_\Gamma|_{b_{e_0}=0}$  does not identically vanish, we get

$$\begin{aligned} & \mathcal{Z}\left(\partial_\Gamma^{e_0} \left(\prod_e b_e \cdot \Pi_\Gamma\right)\right) \\ &= \mathcal{Z}\left(\frac{1}{2} \cdot \prod_e b_e \cdot \Pi_\Gamma|_{b_{e_0}=0}\right) \\ &= \mathcal{Z}\left(\frac{1}{2} \cdot b_{e_0} \cdot \prod_{e \in E(\Gamma_1)} b_e \cdot \Pi_{\Gamma_1} \cdot \prod_{e \in E(\Gamma_2)} b_e \cdot \Pi_{\Gamma_2}\right) \\ &= \frac{1}{2} \cdot \frac{\pi^2}{6} \cdot \mathcal{Z}\left(\prod_{e \in E(\Gamma_1)} b_e \cdot \Pi_{\Gamma_1}\right) \cdot \mathcal{Z}\left(\prod_{e \in E(\Gamma_2)} b_e \cdot \Pi_{\Gamma_2}\right). \end{aligned} \tag{2.14}$$

If the edge  $e_0$  is not a bridge and  $\Pi_\Gamma|_{b_{e_0}=0}$  does not identically vanish, then we get  $\Pi_\Gamma|_{b_{e_0}=0} = \Pi_{\Gamma'}$ , so

$$\begin{aligned} & \mathcal{Z}\left(\partial_\Gamma^{e_0} \left(\prod_e b_e \cdot \Pi_\Gamma\right)\right) \\ &= \mathcal{Z}\left(b_{e_0} \cdot \prod_{e \in E(\Gamma')} b_e \cdot \Pi_{\Gamma'}\right) \\ &= \frac{\pi^2}{6} \cdot \mathcal{Z}\left(\prod_{e \in E(\Gamma')} b_e \cdot \Pi_{\Gamma'}\right). \end{aligned} \tag{2.15}$$

Rewrite the right-hand side of (1.23) as

$$\frac{3}{\pi^2} \cdot \sum_{\Gamma \in \mathcal{G}_{g,n}} \mathcal{Z}(\partial_\Gamma P_\Gamma) = \frac{3}{\pi^2} \cdot \sum_{\Gamma \in \mathcal{G}_{g,n}} \sum_{e \in E(\Gamma)} \mathcal{Z}(\partial_\Gamma^e P_\Gamma).$$

Apply (2.14) and (2.15) to the resulting sum, keeping the intersection numbers as formal expressions, and simplify the product of the factors  $\frac{3}{\pi^2}$  and  $\frac{\pi^2}{6}$ .

Suppose now that  $g \geq 1$  (the consideration in the case  $g = 0$  is completely analogous). Consider the sum on the right-hand side of (1.26), and replace  $\text{Vol } \mathcal{Q}_{g_i, n_i}$  for  $i = 1, 2$  and  $\text{Vol } \mathcal{Q}_{g-1, n+2}$  in it by the corresponding sums (1.13) over  $\mathcal{G}_{g_1, n_1} \times \mathcal{G}_{g_2, n_2}$  and  $\mathcal{G}_{g-1, n+2}$ , respectively (where we keep the intersection numbers as formal expressions). It is easy to see that, term by term, we obtain the same sum as above. Theorem 1.17 is proved.  $\square$

### 3. Comparison with Mirzakhani’s formula for $b_{g,n}$

#### 3.1. Mirzakhani’s expression for the volume of $\mathcal{Q}_g$

Consider a pair  $(X, \lambda)$ , where  $X$  is a hyperbolic surface of genus  $g$  without punctures and  $\lambda$  is a measured lamination on  $X$ . In [46], Mirzakhani associates to almost any such pair  $(X, \lambda)$  a unique holomorphic quadratic differential  $q = F(\lambda, X)$  on the complex curve  $C = C(X)$  corresponding to the hyperbolic metric  $X$ .

Consider the measure  $\mu_{\text{WP}}$  on the moduli space  $\mathcal{M}_g$  coming from the Weil–Petersson volume element, and consider Thurston’s measure  $\mu_{\text{Th}}$  on the space of measured laminations  $\mathcal{ML}_g$ . Using ergodicity arguments, Mirzakhani proves in [46] that the pushforward measure

$$\mu_g := F_*(\mu_{\text{WP}} \otimes \mu_{\text{Th}}) \tag{3.1}$$

on  $\mathcal{Q}_g$  under the map

$$F : \mathcal{M}_g \times \mathcal{ML}_g \rightarrow \mathcal{Q}_g$$

is proportional to the Masur–Veech measure. Equation (3.1) should be considered as the definition of Mirzakhani’s normalization of the Masur–Veech measure on  $\mathcal{Q}_g$ .

*Remark 3.1*

Note that under such an implicit definition of the Masur–Veech measure it is not clear at all why its density should be constant in period coordinates. Definition (3.1) does not provide any distinguished lattice in period coordinates either. Thus, though we know, by results of Mirzakhani, that the density of the Masur–Veech measure defined by (3.1) differs from the Masur–Veech volume element defined in Section 2.1 by a constant numerical factor which depends only on  $g$ , evaluation of this factor is not straightforward.

Mirzakhani proves in [46] that the map  $F$  identifies the length  $\ell_\lambda(X)$  of the measured lamination  $\lambda$  evaluated in the hyperbolic metric  $X$  with the norm  $\|q(\lambda, X)\| = \int_C |q|$  of the quadratic differential  $q$  (defined as the area of the flat surface associated to the pair  $(C, q)$ ):

$$\int_{C(X)} |q(\lambda, X)| = \ell_\lambda(X). \tag{3.2}$$

Relation (3.2) implies that restricting  $F$  to the total space of the bundle of “unit balls”  $B_X$  over  $\mathcal{M}_g$ , defined by equation (1.29), one gets as the image of the restricted map the total space  $\mathcal{Q}_g^{\text{Area}\leq 1}$  of the bundle of “unit balls” in  $\mathcal{Q}_g$ , where

$$\mathcal{Q}_g^{\text{Area}\leq 1} = \{(C, q) \in \mathcal{Q}_g \mid \text{Area}(C, q) \leq 1\}. \tag{3.3}$$

We have seen in Section 1.1 that the real hypersurface

$$\mathcal{Q}_g^{\text{Area}=1} = \{(C, q) \in \mathcal{Q}_{g,n} \mid \text{Area}(C, q) = 1\} \tag{3.4}$$

can be seen as the unit cotangent bundle to  $\mathcal{M}_g$  (denoted by  $\mathcal{Q}^1\mathcal{M}_g$  in [46]). Mirzakhani defines the Masur–Veech volume of  $\mathcal{Q}_g$  as

$$\text{Vol}_{\text{Mir}} \mathcal{Q}_g := \mu_g(\mathcal{Q}_g^{\text{Area}\leq 1}). \tag{3.5}$$

The preceding observations imply Mirzakhani’s following formula for  $\text{Vol}_{\text{Mir}} \mathcal{Q}_g$  from [46].

**THEOREM** ([46, Theorem 1.4])

*The Masur–Veech volume of the moduli space of holomorphic quadratic differentials on complex curves of genus  $g$  defined by (3.5) under normalizations (3.1) satisfies the relation*

$$\text{Vol}_{\text{Mir}} \mathcal{Q}_g = b_g, \tag{3.6}$$

where  $b_g = b_{g,0}$  is defined by (1.30).

Formula (1.34) from Corollary 1.24 relating  $\text{Vol} \mathcal{Q}_g = \text{Vol} \mathcal{Q}_{g,0}$  and  $b_g = b_{g,0}$  implies that the Masur–Veech volume  $\text{Vol}_{\text{Mir}} \mathcal{Q}_g$  in Mirzakhani’s normalization (3.6) and the Masur–Veech volume  $\text{Vol} \mathcal{Q}_g$  in the normalization of formula (1.13) are related by the following factor:

$$\text{Vol} \mathcal{Q}_g = ((12g - 12) \cdot (4g - 4)! \cdot 2^{4g-3}) \cdot \text{Vol}_{\text{Mir}} \mathcal{Q}_g. \tag{3.7}$$

### 3.2. Mirzakhani’s formulas for the Masur–Veech volume of $\mathcal{Q}_g$ and for asymptotic frequencies of simple closed geodesic multicurves

As we have seen,  $\text{Vol}_{\text{Mir}} \mathcal{Q}_g$  appearing in (3.6) is a particular case of the more general quantity  $b_{g,n}$  defined in (1.30). The quantity  $b_{g,n}$  is computed in [47, Theorem 5.3, p. 118]. To reproduce the corresponding formula and closely related formula for the asymptotic frequencies  $c(\gamma)$  of simple closed geodesic multicurves  $\gamma$  of fixed topological type, we need to recall the notation from [47], as follows.

*Simple closed multicurves.* Depending on the context, we denote by the same symbol  $\gamma$  a collection of disjoint, essential, nonperipheral simple closed curves, no two of which are in the same homotopy class; a disjoint union of such curves; the corresponding primitive multicurve; and the corresponding orbit in the space  $\mathcal{ML}_{g,n}$  under the action of the mapping class group  $\text{Mod}_{g,n}$ ,

$$\gamma = \gamma_1 \sqcup \cdots \sqcup \gamma_k = \gamma_1 + \cdots + \gamma_k = (\gamma_1, \dots, \gamma_k).$$

(Mirzakhani [47] uses symbols  $\gamma$ ,  $\hat{\gamma}$ , and  $\tilde{\gamma}$  for these objects, depending on the context.) To every such multicurve  $\gamma$  Mirzakhani [47] associates a collection of quantities  $N(\gamma)$ ,  $M(\gamma)$ ,  $\text{Sym}(\gamma)$ ,  $b_\gamma$ ,  $\text{Vol}_{\text{WP}}(\mathcal{M}_{g,n}(\gamma, x))$  involved in the formula for  $b_{g,n}$ . For the sake of completeness and to simplify comparison formulas, we reproduce the definitions of these quantities (see [47] for details).

*The collection  $\mathfrak{S}_{g,n}$  of all topological types of primitive multicurves.* Recall that by  $S_{g,n}$  we denote a smooth orientable topological surface of genus  $g$  with  $n$  punctures. Consider the finite set  $\mathfrak{S}_{g,n}$  defined as (see [47, (5.4), p. 118])

$$\mathfrak{S}_{g,n} := \{\gamma \mid \gamma \text{ is a union of simple closed curves on } S_{g,n}\} / \text{Mod}_{g,n}. \tag{3.8}$$

It is immediate to see that  $\mathfrak{S}_{g,n}$  is in canonical bijection with the set  $\mathfrak{G}_{g,n}$  of stable graphs defined in Section 1.4,

$$\mathfrak{S}_{g,n} \simeq \mathfrak{G}_{g,n}. \tag{3.9}$$

*Symmetries  $\text{Stab}(\gamma)$  and  $N(\gamma)$  of a primitive multicurve.* For any set  $A$  of homotopy classes of simple closed curves on  $S_{g,n}$ , Mirzakhani defines  $\text{Stab}(A)$  as

$$\text{Stab}(A) := \{g \in \text{Mod } g, n \mid g \cdot A = A\} \subset \text{Mod}_{g,n}.$$

Having a multicurve  $\gamma$  on  $S_{g,n}$  as above, Mirzakhani defines

$$\text{Sym}(\gamma) := \text{Stab}(\gamma) / \bigcap_{i=1}^k \text{Stab}(\gamma_i).$$

(See the beginning of Section 4 on page 112 of [47] for both definitions.) For any single connected simple closed curve  $\gamma_i$ , one has  $|\text{Sym}(\gamma_i)| = 1$ .

For each connected simple closed curve  $\gamma_i$ , define  $\text{Stab}_0(\gamma_i) \subset \text{Stab}(\gamma_i)$  as the subgroup consisting of elements which preserve the orientation of  $\gamma_i$ . Define  $N(\gamma)$  (see [47, p. 113]) as

$$N(\gamma) := \left| \bigcap_{i=1}^k \text{Stab}(\gamma_i) / \bigcap_{i=1}^k \text{Stab}_0(\gamma_i) \right|.$$

Consider a stable graph  $\Gamma(\gamma)$  associated to a primitive multicurve  $\gamma$ . It follows from the definitions of  $\text{Stab}(\gamma)$ ,  $N(\gamma)$ , and  $\text{Aut}(\gamma)$  that

$$|\text{Aut}(\Gamma(\gamma))| = |\text{Sym}(\gamma)| \cdot N(\gamma). \tag{3.10}$$

*Number  $M(\gamma)$  of 1-handles.* Consider now the closed surface

$$S_{g,n}(\gamma) = \bigsqcup_{j=1}^s S_{g_j,n_j} \tag{3.11}$$

obtained from  $S_{g,n}$  by cutting along all  $\gamma_1, \dots, \gamma_k$ . Here  $S_{g_1,n_1}, \dots, S_{g_s,n_s}$  are the connected components of the resulting surface  $S_{g,n}(\gamma)$ . Define (see [47, Theorem 4.1, p. 114])

$$M(\gamma) := |\{i \mid \gamma_i \text{ separates off a 1-handle from } S_{g,n}\}|.$$

By definition, a 1-handle is a surface of genus 1 with one boundary component—that is, a surface of type  $S_{1,1}$ .

*Remark 3.2*

There is one very particular case when the index  $i$  in the definition of  $M(\gamma)$  should be counted with multiplicity 2. Namely, when  $\gamma$  is a connected separating simple closed curve on a surface  $S_2 = S_{2,0}$ , it simultaneously separates off *two* 1-handles, and thus should be counted with multiplicity 2. In other words, the quantity  $M(\gamma)$  might be defined as

$$M(\gamma) := \text{number of surfaces of type } S_{1,1} \text{ in decomposition (3.11)} \tag{3.12}$$

without any exceptions and multiplicities.

Wright suggested an alternative way to fix this issue (see [59, pp. 12–13, footnote 2]).

*Volume polynomials  $V_{g,n}$ .* In [44] and [45], Mirzakhani proved the following statement from [47].

**THEOREM** ([47, Theorems 4.2, 4.3])

*The Weil–Petersson volume  $\text{Vol}_{\text{WP}} \mathcal{M}_{g,n}(b)$  of the moduli space of bordered hyperbolic surfaces of genus  $g$  with hyperbolic boundary components of lengths  $b_1, \dots, b_n$  is a polynomial  $V_{g,n}(b_1, \dots, b_n)$  in even powers of  $b_1, \dots, b_n$ ; that is,*

$$\text{Vol}_{\text{WP}} \mathcal{M}_{g,n}(b) = V_{g,n}(b) = \sum_{\substack{\alpha \\ |\alpha| \leq 3g-3+n}} C_\alpha \cdot b^{2\alpha},$$



where  $C_\alpha > 0$  lies in  $\pi^{6g-6+2n-2|\alpha|} \cdot \mathbb{Q}$ . The coefficient  $C_\alpha$  is given by

$$C_\alpha = \frac{1}{2^{|\alpha|} \cdot \alpha! \cdot (3g - 3 + n - |\alpha|)!} \int_{\overline{\mathcal{M}}_{g,n}} \psi_1^{\alpha_1} \dots \psi_n^{\alpha_n} \cdot \omega^{3g-3+n-|\alpha|},$$

where  $\omega$  is the Weil–Petersson symplectic form,  $\alpha! = \prod_{i=1}^n \alpha_i!$ , and  $|\alpha| = \sum_{i=1}^n \alpha_i$ .

*Remark 3.3 (see [37])*

For one particular pair  $(g, n)$ , namely, for  $(g, n) = (1, 1)$ , Mirzakhani’s normalization (see [47, (4.5), p. 116] or [45]),

$$\text{Vol}_{\text{WP}}^{\text{Mir}} \mathcal{M}_{1,1}(b) = \frac{1}{24}(b^2 + 4\pi^2) \tag{3.13}$$

is two times greater than the normalization in many other papers. Topologically,  $\overline{\mathcal{M}}_{1,1}$  is homeomorphic to  $\mathbb{C}P^1$ . However, since every elliptic curve with a marked point admits an involution, the fundamental class of the orbifold  $\overline{\mathcal{M}}_{1,1}$  used in the integration equals  $[\overline{\mathcal{M}}_{1,1}] = \frac{1}{2}[\mathbb{C}P^1] \in H_2(\mathbb{C}P^1)$ , which gives the value

$$\text{Vol}_{\text{WP}} \mathcal{M}_{1,1}(b) = \frac{1}{48}(b^2 + 4\pi^2). \tag{3.14}$$

Denote by  $V_{g,n}^{\text{top}}(b)$  the homogeneous part of the top degree  $6g - 6 + 2n$  of  $V_{g,n}(b)$ . It follows from the definition of the volume polynomial  $V_{g,n}(b)$  that

$$V_{g,n}^{\text{top}}(b) = \sum_{|\alpha|=3g-3+n} C_\alpha \cdot b^{2\alpha}, \tag{3.15}$$

where  $C_\alpha$  is given by

$$C_\alpha = \frac{1}{2^{3g-3+n} \cdot \alpha!} \int_{\overline{\mathcal{M}}_{g,n}} \psi_1^{\alpha_1} \dots \psi_n^{\alpha_n}. \tag{3.16}$$

Comparing the definition of  $V_{g,n}^{\text{top}}(b)$  with the definitions (1.6)–(1.8) of the polynomial  $N_{g,n}(b)$ , we get the following result.

LEMMA 3.4

The homogeneous parts of the top degree of Mirzakhani’s volume polynomial  $V_{g,n}$  and of Kontsevich’s polynomial  $N_{g,n}$  coincide up to the constant factor:

$$V_{g,n}^{\text{top}}(b) = \begin{cases} 2^{2g-3+n} \cdot N_{g,n}(b) & \text{for } (g, n) \neq (1, 1), \\ 2 \cdot 2^{2g-3+n} \cdot N_{g,n}(b) & \text{for } (g, n) = (1, 1). \end{cases} \tag{3.17}$$

In the exceptional case  $(g, n) = (1, 1)$  the polynomial  $V_{1,1}^{\text{top}}(b) = \frac{1}{24}b^2$  used by Mirzakhani is two times greater than  $N_{1,1}(b) = \frac{1}{48}b^2$ . The origin of this extra factor 2 is explained in Remark 3.3.

*Volume polynomial*  $\text{Vol}_{\text{WP}}(\mathcal{M}_{g,n}(\gamma, x))$  associated to a multicurve  $\gamma$ . Assuming that the initial surface  $S_{g,n}$  is endowed with a hyperbolic metric, and that the simple closed curves  $\gamma_1, \dots, \gamma_k$  are realized by simple closed hyperbolic geodesics of hyperbolic lengths  $x_1 = \ell_{\gamma_1}(X), \dots, x_k = \ell_{\gamma_k}(X)$ , the boundary  $\partial S_{g,n}(\gamma)$  gets  $k$  pairs of distinguished boundary components of lengths  $x_1, \dots, x_k$ . We assume that the  $n$  marked points of  $S_{g,n}$  are represented by hyperbolic cusps—that is, by hyperbolic boundary components of zero length. Denote by  $g_j$  the genus and by  $n_j$  the number of boundary components (including cusps) of each surface  $S_j$ , where  $j = 1, \dots, s$ . To each boundary component of each surface  $S_j$  we have assigned a length variable which is equal to  $x_i$  if the boundary component comes from the cut along  $\gamma_i$ , or is equal to 0 if the boundary component comes from a cusp (one of the  $n$  marked points). Consider the corresponding Mirzakhani–Weil–Petersson volume polynomial  $V_{g_j, n_j}(x)$ , and define (see [47, (4.1), p. 113])

$$\text{Vol}_{\text{WP}}(\mathcal{M}_{g,n}(\gamma, x)) := \frac{1}{N(\gamma)} \prod_{j=1}^s V_{g_j, n_j}(x). \tag{3.18}$$

Denote by  $(2d_1, \dots, 2d_k)_\gamma$  the coefficient of  $x^{2d_1} \dots x^{2d_k}$  in this polynomial, and let (see [47, (5.3), p. 118])

$$b_\gamma(2d_1, \dots, 2d_k) := (2d_1, \dots, 2d_k)_\gamma \frac{\prod_{i=1}^k (2d_i + 1)!}{(6g - 6 + 2n)!}. \tag{3.19}$$

Now everything is ready to state Mirzakhani’s result and prove Theorem 1.22.

*Mirzakhani’s results and proof of comparison theorems.* Let  $\gamma = \gamma_1 + \dots + \gamma_k$  be a primitive simple closed multicurve as above. Let  $\gamma_a = \sum_{i=1}^k a_i \gamma_i$ , where  $a_i \in \mathbb{N}$  for  $i = 1, \dots, k$ . (To follow Mirzakhani’s original notation, we denote integer weights of components of a multicurve by  $a_i$  and not by  $H_i$  as before.)

THEOREM ([47, Theorem 5.3])

The frequency  $c(\gamma_a)$  of a multicurve  $\gamma_a = \sum_{i=1}^k a_i \gamma_i$  is equal to

$$c(\gamma_a) = \frac{2^{-M(\gamma)}}{|\text{Sym}(\gamma)|} \cdot \sum_{\substack{d \\ |d|=3g-3+n-k}} \frac{b_\gamma(2d_1, \dots, 2d_k)}{a_1^{2d_1+2} \dots a_k^{2d_k+2}}. \tag{3.20}$$

The expression  $b_{g,n}$  given by the integral (1.30) can be represented as the sum

$$b_{g,n} = \sum_{\gamma \in \delta_{g,n}} B_\gamma, \tag{3.21}$$

where for  $\gamma = \bigsqcup_{i=1}^k \gamma_i$  one has

$$B_\gamma = \frac{2^{-M(\gamma)}}{|\text{Sym}(\gamma)|} \cdot \sum_{|d|=3g-3+n-k} b_\gamma(2d_1, \dots, 2d_k) \prod_{i=1}^k \zeta(2s_i + 2). \tag{3.22}$$

*Proof of Theorem 1.22*

Let  $\gamma = \gamma_1 + \dots + \gamma_k$  be a primitive simple closed multicurve in  $\mathfrak{S}_{g,n}$ . Let  $\Gamma$  be the associated decorated graph. The sets  $\mathfrak{S}_{g,n}$  and  $\mathfrak{G}_{g,n}$  are in canonical one-to-one correspondence.

Defining  $M(\gamma)$  as in (3.12), we see that  $M(\gamma)$  is the number of factors  $V_{1,1}$  in the product (3.18)—that is, the number of indices  $j$  in the range  $\{1, \dots, s\}$  such that  $(g_j, n_j) = (1, 1)$ . Thus, the factor  $2^{-M(\gamma)}$  in equations (3.20) and (3.22) compensates the difference in normalizations (3.13) and (3.14) (see Remark 3.3). Hence, applying (3.17) to the right-hand side of (3.18), we get

$$2^{-M(\gamma)} \prod_{j=1}^s V_{g_j, n_j}(x) = \left( \prod_{j=1}^s N_{g_j, n_j}(x) \right) \cdot \left( \prod_{j=1}^s 2^{2g_j - 3 + n_j} \right).$$

For a decorated graph  $\Gamma$ , one gets

$$\begin{aligned} g &= 1 - \chi(\Gamma) + \sum_{v_j \in V(\Gamma)} g(v_j) = 1 - |V(\Gamma)| + |E(\Gamma)| + \sum_{j=1}^s g_j \\ &= 1 + \sum_{j=1}^s \left( g_j - 1 + \frac{n_j}{2} \right) - \frac{(\text{total number of half-edges (legs)})}{2}. \end{aligned}$$

(By convention,  $E(\Gamma)$  denotes the set of “true” edges of  $\Gamma$  versus  $n$  half-edges (legs) corresponding to  $n$  marked points on  $S_{g,n}$ .) Hence,

$$\sum_{j=1}^s (2g_j - 3 + n_j) = 2g - 2 + n - |V(\Gamma)| = 2g - 3 + n - (|V(\Gamma)| - 1).$$

Note also that by definition

$$|\text{Aut}(\Gamma)| = |\text{Sym}(\gamma)| \cdot N(\gamma).$$

We have proved that

$$\begin{aligned} &\frac{2^{-M(\gamma)}}{|\text{Sym}(\gamma)| \cdot N(\gamma)} \prod_{j=1}^s V_{g_j, n_j}(x) \\ &= 2^{2g+n-3} \cdot \left( \frac{1}{2^{|V(\Gamma)|-1}} \cdot \frac{1}{|\text{Aut}(\Gamma)|} \cdot \prod_{v \in V(\Gamma)} N_{g_v, n_v}(x) \right). \tag{3.23} \end{aligned}$$

Let us now prove relation (1.33). Expressions (3.19) and (3.20) represent  $c(\gamma_{\mathbf{H}})$  as the sum of terms constructed using the polynomial on the left-hand side of (3.23). Expression (1.14) represents  $\text{Vol}(\Gamma, \mathbf{H})$  as the sum of the corresponding terms constructed using the proportional polynomial in the brackets on the right-hand side of (3.23). The only difference between the two sums comes from the global normalization factors shared by all terms of the sums. Namely, (3.19) has an extra factor  $(6g - 6 + 2n)!$  in the denominator, while the first line of (1.12) has the extra global factor  $\frac{2^{6g-5+2n} \cdot (4g-4+n)!}{(6g-7+2n)!}$ . Taking into consideration the coefficient of proportionality  $2^{2g+n-3}$  relating the two polynomials in (3.23), we obtain

$$(6g - 6 + 2n)! \cdot c(\gamma_a) = 2^{2g+n-3} \cdot \left( \frac{2^{6g-5+2n} \cdot (4g - 4 + n)!}{(6g - 7 + 2n)!} \right)^{-1} \cdot \text{Vol}(\Gamma, \mathbf{H}),$$

and (1.33) follows.

It remains to notice that, by definitions (3.22) and (3.20) of  $B_\gamma$  and  $c(\gamma_a)$ , respectively, one gets

$$B_\gamma = \sum_{\mathbf{H} \in \mathbb{N}^k} c(\gamma_{\mathbf{H}}).$$

Similarly, by (1.5), we get

$$\text{Vol}(\Gamma) = \sum_{\mathbf{H} \in \mathbb{N}^k} \text{Vol}(\Gamma, \mathbf{H}).$$

Since the coefficient of proportionality  $\text{const}_{g,n}$  between  $\text{Vol}(\Gamma, \mathbf{H})$  and  $c(\gamma_{\mathbf{H}})$  in (1.33) is common for all  $\mathbf{H} \in \mathbb{N}^k$ , relation (1.33) for the individual terms implies the analogous relation  $\text{Vol}(\Gamma) = \text{const}_{g,n} \cdot B_\gamma$  for the above sums of all terms over all  $\mathbf{H} \in \mathbb{N}^k$  and for the corresponding sums  $\text{Vol } \mathcal{Q}_{g,n} = \text{const}_{g,n} \cdot b_{g,n}$  (see (1.13) and (3.21)) over all stable graphs  $\Gamma \in \mathcal{G}_{g,n}$ . Theorem 1.22 and Corollary 1.24 are proved. □

#### 4. Large genus asymptotics for frequencies of simple closed curves and 1-cylinder square-tiled surfaces

##### 4.1. Universal bounds for 2-correlators

Following Witten [58], define

$$\langle \tau_{d_1} \cdots \tau_{d_n} \rangle = \int_{\mathcal{M}_{g,n}} \psi^{d_1} \cdots \psi^{d_n}, \tag{4.1}$$

where  $d_1 + \cdots + d_n = 3g - 3 + n$ .

Consider the following normalization of the 2-correlators  $\langle \tau_k \tau_{3g-1-k} \rangle_g$  introduced in [61]:

$$a_{g,k} = \frac{(2k + 1)!! \cdot (6g - 1 - 2k)!!}{(6g - 1)!!} \cdot 24^g \cdot g! \cdot \langle \tau_k \tau_{3g-1-k} \rangle_g. \tag{4.2}$$

By (7) in [61], under such normalization the differences of 2-correlators admit the following explicit expression:

$$a_{g,k+1} - a_{g,k} = \frac{(6g - 3 - 2k)!!}{(6g - 1)!!} \cdot \begin{cases} \frac{(6j-1)!!}{j!} \cdot \frac{(g-1)!}{(g-j)!} \cdot (g - 2j) & \text{for } k = 3j - 1, \\ -2 \cdot \frac{(6j+1)!!}{j!} \cdot \frac{(g-1)!}{(g-1-j)!} & \text{for } k = 3j, \\ 2 \cdot \frac{(6j+3)!!}{j!} \cdot \frac{(g-1)!}{(g-1-j)!} & \text{for } k = 3j + 1, \end{cases} \tag{4.3}$$

where  $k = 0, 1, \dots, \lfloor \frac{3g-1}{2} \rfloor - 1$ , and  $a_{g,0} = 1$ .

It follows from definition (4.2) that  $a_{g,k} = a_{g,3g-1-k}$ , so relations (4.3) determine  $a_{g,k}$  for all  $g \geq 1$  and for all  $k$  satisfying  $0 \leq k \leq 3g - 1$ .

PROPOSITION 4.1

For all  $g \in \mathbb{N}$  and for all integers  $k$  in the range  $\{2, 3, \dots, 3g - 3\}$ , the following bounds are valid:

$$1 - \frac{2}{6g - 1} = a_{g,1} = a_{g,3g-2} < a_{g,k} < a_{g,0} = a_{g,3g-1} = 1. \tag{4.4}$$

Proposition 4.1 is proved in Appendix A.

4.2. Asymptotic volume contribution of 1-cylinder square-tiled surfaces

In this section we compute the large genus asymptotics for the contributions of the stable graphs having a single edge to the Masur–Veech volume  $\text{Vol } \mathcal{Q}_g$ . In other words, we compute the asymptotic contributions of 1-cylinder square-tiled surfaces to the Masur–Veech volume of the principal stratum  $\mathcal{Q}(1^{4g-4})$ . These contributions obviously provide lower bounds for  $\text{Vol } \mathcal{Q}_g$ .

Denote by  $\Gamma_1(g)$  the stable graph having a single vertex and a single edge and such that the vertex is decorated with integer  $g - 1$  (see Figure 4). The single edge forms a loop, so this stable graph represents a surface of genus  $g$  without marked points. The stable graph  $\Gamma_1(g)$  encodes the orbit of a simple closed nonseparating curve on a surface of genus  $g$ .

THEOREM 4.2

Consider the stable graph  $\Gamma_1(g)$  having a single vertex decorated with label  $g - 1$  and a single edge (see Figure 4). The contribution  $\text{Vol } \Gamma_1(g)$  of  $\Gamma_1(g)$  to the Masur–Veech volume  $\text{Vol } \mathcal{Q}_g$  of the moduli space  $\mathcal{Q}_g$  of holomorphic quadratic differentials on complex curves of genus  $g$  has the following asymptotics:

$$\text{Vol } \Gamma_1(g) = \sqrt{\frac{2}{3\pi g}} \cdot \left(\frac{8}{3}\right)^{4g-4} \cdot \left(1 + O\left(\frac{1}{g}\right)\right) \quad \text{as } g \rightarrow +\infty. \quad (4.5)$$

*Proof*

It would be slightly more convenient to shift  $g$  by 1. Assign variable  $b_1$  to the only edge of the graph. Formula (1.14) from Theorem 1.5 applied to the graph  $\Gamma_1(g + 1)$  gives

$$\begin{aligned} & \left(\frac{(4g)!}{(6g)!}\right) \cdot 2^{6g} \cdot 12g \cdot \frac{1}{2} \cdot 1 \cdot b_1 \cdot N_{g,2}(b_1, b_1) \\ \xrightarrow{Z} \text{Vol } \Gamma_1(g + 1) &= (4g)! \cdot 2^{g+2} \cdot \zeta(6g) \cdot \sum_{d_1+d_2=3g-1} \frac{\langle \psi_1^{d_1} \psi_2^{d_2} \rangle}{d_1! \cdot d_2!} \\ &= (4g)! \cdot 2^{g+2} \cdot \zeta(6g) \cdot \sum_{k=0}^{3g-1} \frac{\langle \tau_k \tau_{3g-1-k} \rangle_g}{k! \cdot (3g-1-k)!}. \end{aligned} \quad (4.6)$$

Now pass to the normalization (4.2) of the correlators  $\langle \tau_k \tau_{3g-1-k} \rangle_g$ :

$$a_{g,k} = \frac{(2k + 1)!! \cdot (6g - 1 - 2k)!!}{(6g - 1)!!} \cdot 24^g \cdot g! \cdot \langle \tau_k \tau_{3g-1-k} \rangle_g.$$

By Proposition 4.1, the 2-correlators admit the following uniform bounds under such normalization:

$$1 - \frac{2}{6g - 1} \leq a_{g,k} \leq 1, \quad \text{for } k = 0, 1, \dots, 3g - 1.$$

Rewriting the expression (4.6) for  $\text{Vol } \Gamma_1(g + 1)$  in terms of  $a_{g,k}$ , we obtain

$$\begin{aligned} & \text{Vol } \Gamma_1(g + 1) / \zeta(6g) \\ &= (4g)! \cdot 2^{g+2} \cdot \frac{(6g - 1)!!}{24^g \cdot g!} \\ & \cdot \sum_{k=0}^{3g-1} \frac{a_{g,k}}{k! \cdot (2k + 1)!! \cdot (3g - 1 - k)! \cdot (6g - 1 - 2k)!!}. \end{aligned} \quad (4.7)$$

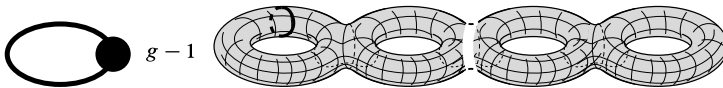


Figure 4. The graph  $\Gamma_1(g)$  on the left represents a nonseparating simple closed curve on a surface of genus  $g$ .

Passing from double factorials to factorials,

$$\begin{aligned} (6g - 1)!! &= \frac{(6g)!}{(3g)! \cdot 2^{3g}}, \\ (2k + 1)!! &= \frac{(2k + 1)!}{k! \cdot 2^k}, \\ (6g - 1 - 2k)!! &= \frac{(6g - 1 - 2k)!}{(3g - 1 - k)! \cdot 2^{3g - k - 1}}, \end{aligned}$$

we rewrite and simplify (4.7) as follows:

$$\begin{aligned} &\text{Vol } \Gamma_1(g + 1) / \zeta(6g) \\ &= (4g)! \cdot 2^{g+2} \cdot \frac{1}{3^g \cdot 2^{3g}} \cdot \frac{1}{g! \cdot (3g)!} \cdot \frac{1}{2} \sum_{k=0}^{3g-1} \binom{6g}{2k+1} \cdot a_{g,k} \\ &= \frac{(4g)!}{g! \cdot (3g)!} \cdot \frac{1}{3^g \cdot 2^{2g}} \cdot 2 \sum_{k=0}^{3g-1} \binom{6g}{2k+1} \cdot a_{g,k}. \end{aligned} \tag{4.8}$$

Taking the difference of binomial expansions of the left-hand sides of the identities  $(1 - 1)^{2n} = 0$  and  $(1 + 1)^{2n} = 2^{2n}$ , we derive the classical identity

$$\sum_{k=0}^{n-1} \binom{2n}{2k+1} = 2^{2n-1}.$$

Combining the latter identity evaluated for  $n = 3g$  with bounds (4.4), we get the following bounds for  $\text{Vol } \Gamma_1(g + 1)$ :

$$\binom{4g}{g} \cdot \left(\frac{2^4}{3}\right)^g \cdot \left(1 - \frac{2}{6g-1}\right) \leq \text{Vol } \Gamma_1(g + 1) / \zeta(6g) \leq \binom{4g}{g} \cdot \left(\frac{2^4}{3}\right)^g. \tag{4.9}$$

Note that

$$\zeta(6g) \rightarrow 1 \quad \text{as } g \rightarrow +\infty$$

and the convergence is exponentially fast.

By now applying Stirling’s formula to the factorials in the binomial coefficient  $\binom{4g}{g}$  in the latter expression, we obtain

$$\binom{4g}{g} = \sqrt{\frac{2}{3\pi g}} \cdot \left(\frac{2^8}{3^3}\right)^g \cdot \left(1 + O\left(\frac{1}{g}\right)\right) \quad \text{as } g \rightarrow +\infty. \tag{4.10}$$

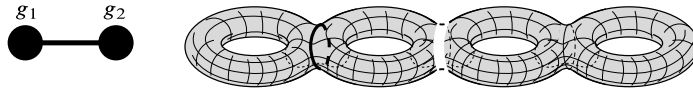


Figure 5. The graph  $\Delta(g_1, g_2)$  on the left represents a simple closed curve on a surface of genus  $g$  separating the surface into surfaces of genera  $g_1$  and  $g_2$ , where  $g_1 + g_2 = g$  (on the right).

Combining the latter equality with bounds (4.9), we obtain the desired formula (4.5) in genus  $g + 1$ :

$$\text{Vol } \Gamma_1(g + 1) = \sqrt{\frac{2}{3\pi g}} \cdot \left(\frac{8}{3}\right)^{4g} \cdot \left(1 + O\left(\frac{1}{g}\right)\right) \quad \text{as } g \rightarrow +\infty. \quad \square$$

*Remark 4.3*

Actually, we have a very good control on the asymptotic expansions of correlators  $a_{g,k}$  in powers of  $\frac{1}{g}$ , so it would not be difficult to specify several terms of the asymptotic expansion of  $O\left(\frac{1}{g}\right)$  in (4.5). We refrain from doing so only because we do not currently see any specific need for a more precise expression.

*Proof of Theorem 1.8*

Inequality (1.16) now follows from (4.9), where we use the following estimates for the factorials involved in the binomial coefficient. By Theorem 1.6 in [12], for all positive real numbers  $x \geq 1$  one has

$$x^x e^{-x} \sqrt{2\pi(x+a)} < \Gamma(x+1) < x^x e^{-x} \sqrt{2\pi(x+b)},$$

where  $a = 1/6 = 0.1666666\dots$  and  $b = \frac{e^2}{2\pi} - 1 = 0.176005\dots$  □

We proceed with the remaining graphs having a single edge. This time the graph has two vertices joined by the edge as in Figure 5. The two vertices are decorated with strictly positive integers  $g_1, g_2 \in \mathbb{N}$  such that  $g_1 + g_2 = g$  (see Figure 4). Without loss of generality, we may assume that  $g_1 \leq g_2$ . This stable graph encodes the orbit of a simple closed curve separating the compact surface of genus  $g$  without punctures into subsurfaces of genera  $g_1$  and  $g_2$ .

PROPOSITION 4.4

*Consider the stable graph  $\Delta(g_1, g - g_1)$  having a single edge joining two vertices decorated with labels  $g_1$  and  $g - g_1$ , respectively (see Figure 5). The contribution*



$\text{Vol}(\Delta(g_1, g - g_1))$  of  $\Delta(g_1, g - g_1)$  to the Masur–Veech volume  $\text{Vol} \mathcal{Q}_g$  is

$$\begin{aligned} &\text{Vol}(\Delta(g_1, g - g_1)) \\ &= \frac{4 \cdot \zeta(6g - 6)}{|\text{Aut}(\Delta(g_1, g - g_1))|} \cdot \binom{4g - 4}{g} \cdot \frac{1}{12^g} \cdot \binom{g}{g_1} \cdot \binom{3g - 4}{3g_1 - 2}. \end{aligned} \tag{4.11}$$

*Proof*

Let  $g_2 = g - g_1$ . The contribution of the graph  $\Delta(g_1, g_2)$  is given by formula (1.13) from Theorem 1.5:

$$\begin{aligned} \text{Vol}(\Delta(g_1, g_2)) &= \frac{2^{6g-5} \cdot (4g - 4)!}{(6g - 7)!} \cdot \frac{1}{2} \cdot \frac{1}{|\text{Aut}(\Gamma(g_1, g_2))|} \\ &\quad \cdot \mathcal{Z}(b \cdot N_{g_1,1}(b) \cdot N_{g_2,1}(b)), \end{aligned} \tag{4.12}$$

where

$$|\text{Aut}(\Gamma(g_1, g_2))| = \begin{cases} 2 & \text{when } g_1 = g_2, \\ 1 & \text{otherwise.} \end{cases} \tag{4.13}$$

By the result of Witten [58, lines following equation (2.26)], one has the following closed expression for 1-correlators:

$$\langle \psi_1^{3g-2} \rangle = \int_{\mathcal{M}_{g,1}} \psi_1^{3g-2} = \frac{1}{24^g \cdot g!}.$$

Applying definitions (1.6) and (1.7) to  $N_{g,1}(b)$  and using the above expression for  $\langle \psi_1^{3g-2} \rangle$ , we get the following closed-form expression for the polynomial  $N_{g,1}(b)$ :

$$\begin{aligned} N_{g,1}(b) &= c_{3g-2} b^{2(3g-2)} = \frac{1}{2^{5g-6+2} \cdot (3g - 2)!} \cdot \langle \psi_1^{3g-2} \rangle \cdot b^{6g-4} \\ &= \frac{1}{2^{8g-4} \cdot 3^g \cdot g! \cdot (3g - 2)!} \cdot b^{6g-4}. \end{aligned} \tag{4.14}$$

Using definition (1.11) of  $\mathcal{Z}$  and the assumption  $g_1 + g_2 = g$ , we can now develop the rightmost factor in (1.11) as follows:

$$\begin{aligned} &\mathcal{Z}(b \cdot N_{g_1,1}(b) \cdot N_{g_2,1}(b)) \\ &= \mathcal{Z}\left(\frac{1}{2^{8g_1-4} \cdot 2^{8g_2-4}} \cdot \frac{1}{3^{g_1} \cdot 3^{g_2}} \cdot \frac{1}{g_1!g_2!}\right. \\ &\quad \left. \cdot \frac{1}{(3g_1 - 2)!(3g_2 - 2)!} \cdot b \cdot b^{6g_1-4} \cdot b^{6g_2-4}\right) \end{aligned}$$

$$= \frac{1}{2^{8g-8}} \cdot \frac{1}{3^g} \cdot \frac{1}{g_1!(g-g_1)!} \cdot \frac{1}{(3g_1-2)!(3g-3g_1-2)!} \cdot (6g-7)! \cdot \zeta(6g-6).$$

Plugging this expression into (4.12), multiplying and dividing by  $g!$  and by  $(3g-4)!$  to pass to binomial coefficients and simplifying, we get the desired formula (4.11). □

We complete this section with the computation of the cumulative contribution of the graphs  $\Delta(g_1, g-g_1)$  to  $\text{Vol } \mathcal{Q}_g$  coming from all  $g_1 = 1, \dots, \lfloor \frac{g}{2} \rfloor$ .

PROPOSITION 4.5

As  $g \rightarrow \infty$ , we have the following asymptotic relation

$$\sum_{g_1=1}^{\lfloor \frac{g}{2} \rfloor} \text{Vol}(\Delta(g_1, g-g_1)) \sim \frac{2}{3\pi g} \cdot \frac{1}{4^g} \cdot \left(\frac{8}{3}\right)^{4g-4}. \tag{4.15}$$

*Proof*

Applying formula (4.12), making the summation index range from 1 to  $g-1$  (instead of up to  $\lfloor \frac{g}{2} \rfloor$ ), and taking into consideration that when  $g = 2g_1$ , the term  $\text{Vol } \Delta(g_1, g_1)$  has  $|\text{Aut}(\Delta(g_1, g_1))| = 2$  (see (4.13)), we obtain

$$\sum_{g_1=1}^{\lfloor \frac{g}{2} \rfloor} \text{Vol}(\Delta(g_1, g-g_1)) = 2 \cdot \zeta(6g-6) \cdot \binom{4g-4}{g} \cdot \frac{1}{12^g} \cdot \sum_{g_1=1}^{g-1} \binom{g}{g_1} \cdot \binom{3g-4}{3g_1-2}.$$

The zeta value  $\zeta(6g-6)$  tends to 1 exponentially fast when  $g \rightarrow +\infty$ . Stirling’s formula provides the following asymptotic value of the binomial coefficient:

$$\binom{4g-4}{g} \sim \frac{2}{\sqrt{6\pi g}} \cdot 3^g \cdot \left(\frac{4}{3}\right)^{4g-4}.$$

Thus, to complete the proof of formula (1.32) and, thus, of Theorem 1.18, it remains to prove the lemma below. □

LEMMA 4.6

*The asymptotic formula*

$$S(g) = \sum_{g_1=1}^{g-1} \binom{g}{g_1} \binom{3g-4}{3g_1-2} \sim \frac{1}{\sqrt{6\pi g}} \cdot 2^{4g-4} \tag{4.16}$$

holds as  $g \rightarrow +\infty$ .

*Proof*

The probability density of the normal distribution  $\mathcal{N}(\mu, \sigma^2)$  is given by the function

$$f(x | \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

Let  $g$  be a large positive integer. By the de Moivre–Laplace theorem, after normalization by  $2^g$  the distribution of the binomial coefficients  $\binom{g}{k}$ , where  $k = 0, 1, \dots, g$ , tends to the normal distribution  $\mathcal{N}(\frac{g}{2}, \frac{g}{4})$  as  $g \rightarrow +\infty$ .

Let  $m$  be any positive integer which we use as a fixed parameter. The normalized distribution of the binomial coefficients  $\binom{m \cdot g}{k}$ , where  $k = 0, 1, \dots, m \cdot g$ , tends to the normal distribution  $\mathcal{N}(\frac{m \cdot g}{2}, \frac{m \cdot g}{4})$  as  $g \rightarrow \infty$ . Hence, the normalized distribution of binomial coefficients  $\binom{m \cdot g}{m \cdot k}$ , where  $k = 0, 1, \dots, g$ , tends to the normal distribution  $\mathcal{N}(\frac{g}{2}, \frac{g}{4m})$  as  $g \rightarrow +\infty$ . In particular, letting  $m = 3$  we see that the normalized distribution of binomial coefficients  $\binom{3g-4}{3k-2}$ , where  $k = 1, \dots, g-1$ , tends to the normal distribution  $\mathcal{N}(\frac{g}{2}, \frac{g}{12} - \frac{1}{9})$ .

We have

$$f(x | \mu, \sigma_1^2) \cdot f(x | \mu, \sigma_2^2) = \frac{\sqrt{2\pi\sigma^2}}{\sqrt{2\pi\sigma_1^2} \cdot \sqrt{2\pi\sigma_2^2}} \cdot f(x | \mu, \sigma^2),$$

$$\text{where } \sigma^2 = \frac{\sigma_1^2 \sigma_2^2}{\sigma_1^2 + \sigma_2^2}.$$

In other words, the normalized distribution of the product of two independent normal distributions  $\mathcal{N}(\mu, \sigma_1^2)$  and  $\mathcal{N}(\mu, \sigma_2^2)$  sharing the same mean  $\mu$  is the normal distribution  $\mathcal{N}(\mu, \frac{\sigma_1^2 \sigma_2^2}{\sigma_1^2 + \sigma_2^2})$ . Hence, after normalization by  $S(g)$ , the product  $\binom{g}{g_1} \binom{3g-4}{3g_1-2}$  of the two binomial distributions, where  $g_1 = 1, \dots, g-1$ , tends to the normal distribution  $\mathcal{N}(g/2, \sigma^2)$  with

$$\sigma^2 = \frac{\frac{g}{4} (\frac{g}{12} - \frac{1}{9})}{\frac{g}{4} + (\frac{g}{12} - \frac{1}{9})} \sim \frac{g}{16}, \quad \text{as } g \rightarrow +\infty.$$

Thus, the asymptotic value of the sum  $S(g)$  can be computed as the value of the product distribution  $\binom{g}{g_1} \binom{3g-4}{3g_1-2}$  at  $\mu = g/2$  multiplied by  $\sqrt{2\pi\sigma^2}$ ; that is,

$$S(g) \sim \binom{g}{\lfloor \frac{g}{2} \rfloor} \cdot \binom{3g-4}{3\lfloor \frac{g}{2} \rfloor - 2} \cdot \sqrt{2\pi \frac{g}{16}}. \tag{4.17}$$

From Stirling’s formula we obtain

$$\binom{2m}{m} \sim \frac{2^{2m}}{\sqrt{\pi m}} \quad \text{and} \quad \binom{2m+1}{m} \sim \frac{2^{2m+1}}{\sqrt{\pi m}}.$$

Applying these asymptotic formulas to each of the binomial coefficients in (4.17), we obtain

$$S(g) \sim \frac{2^g}{\sqrt{\pi \frac{g}{2}}} \cdot \frac{2^{3g-4}}{\sqrt{\pi \frac{3g-4}{2}}} \cdot \frac{\sqrt{\pi g}}{\sqrt{8}} \sim \frac{1}{\sqrt{6\pi g}} \cdot 2^{4g-4}. \quad \square$$

*Remark 4.7*

The hypergeometric sum  $S(g)$  on the left-hand side of (4.16) satisfies the following recursive relation obtained by applying Zeilberger’s algorithm:

$$S(g + 2) = 2 \cdot \frac{(324g^4 + 432g^3 + 123g^2 - 49g - 8)}{(6g - 1)(3g + 4)(3g - 1)(g + 1)} \cdot S(g + 1) + 36 \cdot \frac{(6g + 5)(4g - 1)(4g - 3)}{(6g - 1)(3g + 4)(3g - 1)} \cdot S(g).$$

*4.3. Frequencies of simple closed geodesics*

In this section we use the setting and the notation as in Theorem 6.1 of Mirzakhani [47] reproduced at the end of Section 1.2.

Recall that equivalence classes of smooth simple closed curves on a compact oriented surface of genus  $g$  without boundary are classified as follows. The curve can be separating or nonseparating. All nonseparating curves as in Figure 4 belong to the same class; we denote the corresponding frequency by  $c_{g,\text{nonsep}}$ .

Separating simple closed curves are classified by the genera  $g_1, g_2$  of components in which the curve separates the surface (see Figure 5). Here  $g_1 + g_2 = g$ ,  $g_1, g_2 \geq 1$ , and pairs  $g_1, g_2$  and  $g_2, g_1$  correspond to the same equivalence class. Denote a simple closed curve of this type by  $\gamma_{g_1, g_2}$ . The stable graph corresponding to  $\gamma_{g_1, g_2}$  is  $\Delta(g_1, g_2)$  (see Figure 5).

Recall that the volume contribution  $\text{Vol}(\Delta(g_1, g_2))$  comes from all 1-cylinder square-tiled surfaces of genus  $g = g_1 + g_2$  such that the waist curve of the single cylinder separates the surface into two surfaces of genera  $g_1$  and  $g_2$ , respectively. This single horizontal cylinder can be composed of  $a = 1, 2, \dots$  horizontal bands of squares. The contribution  $\text{Vol}(\Delta(g_1, g_2))$  is the sum of contributions  $\text{Vol}(\Delta(g_1, g_2), H)$  of square-tiled surfaces having a fixed value  $H \in \mathbb{N}$ ,

$$\text{Vol}(\Delta(g_1, g_2)) = \sum_{H=1}^{+\infty} \text{Vol}(\Delta(g_1, g_2), H).$$

Recall also that

$$\text{Vol}(\Delta(g_1, g_2), H) = \frac{1}{H^d} \text{Vol}(\Delta(g_1, g_2)),$$

where  $d = 6g - 6 = \dim \mathcal{Q}_g$ , which implies that

$$\text{Vol}(\Delta(g_1, g_2)) = \zeta(6g - 6) \cdot \text{Vol}(\Delta(g_1, g_2), 1).$$

Combining the latter relation with formula (1.33) from Theorem 1.22, we obtain

$$\frac{\text{Vol}(\Delta(g_1, g_2))}{\zeta(6g - 6)} = 2 \cdot (6g - 6) \cdot (4g - 4)! \cdot 2^{4g-3} \cdot c(\gamma_{g_1, g_2}).$$

Applying (4.11) for  $\text{Vol } \Gamma(g_1, g_2)$ , we obtain the following formula:

$$\begin{aligned} &c(\gamma_{g_1, g-g_1}) \\ &= \frac{1}{|\text{Aut } \Delta(g_1, g-g_1)|} \\ &\quad \cdot \frac{1}{2^{3g-4} \cdot 24g \cdot g_1! \cdot (g-g_1)! \cdot (3g_1-2)! \cdot (3(g-g_1)-2)! \cdot (6g-6)}. \end{aligned} \tag{4.18}$$

In this way, we reproduce the formula for the frequency of simple closed separating geodesics first proved by Mirzakhani [47, p. 124].

*Remark 4.8*

The formula in [47, p. 124] contains two misprints: the power in the first factor in the denominator is indicated as  $2^{3g-2}$  while it should read  $2^{3g-4}$ , and the fifth factor is indicated as  $(3g-2)!$  while it should read  $(3i-2)!$ . Indeed, following Mirzakhani’s calculation, we have to use formula (5.5) from [47] for  $c(\gamma)$ . Using notation of this formula applied to our particular  $\gamma$ , we have  $n = 0$ ,  $k = 1$ ,  $a_1 = 1$ ,  $s_1 = 3g - 4$ . Mirzakhani assumes for simplicity that  $g > 2i > 2$ , which implies that  $M(\gamma) = 0$  and  $|\text{Sym}(\gamma)| = 1$  (and implies that  $|\text{Aut } \Delta(g_1, g - g_1)| = 1$  in notation of formula (4.18) above). Thus, applying formula (5.5) from [47], Mirzakhani obtains

$$c(\gamma) = b_\Gamma(2 \cdot (3g - 4)) = (2 \cdot (3g - 4))_\Gamma \cdot \frac{(2 \cdot (3g - 4) + 1)!}{(6g - 6)!},$$

where  $b_\Gamma(2s_1)$  and  $(2s_1)_\Gamma$  are introduced in [47] in formula (5.3) and in the line above it, respectively.

In order to evaluate  $(2 \cdot (3g - 4))_\Gamma$ , we compute—following Mirzakhani—the product of the coefficients of the leading terms of the polynomials  $V_{i,1}(x)$  and  $V_{g-i,1}(x)$  (see the bottom of page 123 in [47]). In this way, we obtain

$$\begin{aligned} &(2 \cdot (3g - 4))_\Gamma \\ &= \frac{1}{(3i - 2)! \cdot i! \cdot 24^i \cdot 2^{3i-2}} \cdot \frac{1}{(3(g - i) - 2)! \cdot (g - i)! \cdot 24^{g-i} \cdot 2^{3(g-i)-2}} \end{aligned}$$

$$= \frac{1}{2^{3g-4} \cdot 24^g \cdot i! \cdot (g-i)! \cdot (3i-2)! \cdot (3(g-i)-2)!}$$

(cf. (4.18) replacing  $g_1$  with  $i$ ). The remaining last factor  $(6g-6)$  in the denominator of Mirzakhani’s formula comes from  $\frac{(2 \cdot (3g-4)+1)!}{(6g-6)!} = \frac{1}{6g-6}$ .

*Remark 4.9*

The computation of  $c(\alpha_2)$  in [47, p. 123], where  $\alpha_2$  denotes a separating simple closed curve on a surface of genus 2, contains some misprints. It is correctly written that  $N(\alpha_2) = 2$ . However, the factor  $\frac{1}{N(\alpha_2)}$  involved in the definition (4.1) of  $\text{Vol}_{\text{WP}}(\mathcal{M}_{g,n}(\Gamma, \mathbf{x}))$  on page 113 of [47] is missing in the formula for  $\text{Vol}(\mathcal{M}(S(\alpha_2), \ell_{\alpha_2} = x))$  on page 123. We assume that this formula should read:

$$\text{Vol}(\mathcal{M}(S(\alpha_2), \ell_{\alpha_2} = x)) = \frac{1}{N(\alpha_2)} \cdot V_{1,1}(x) \times V_{1,1}(x) = \frac{1}{2} \cdot \left(\frac{x^2}{24} + \frac{\pi^2}{6}\right)^2.$$

Also, the factor  $2^{-M(\alpha_2)}$  present in the general formula (5.5) in [47, p. 118] is missing in the computation of  $c(\alpha_2)$ . Finally, in this particular case (and only in this case) either  $M(\alpha_2)$  should be readjusted as  $M(\alpha_2) = 2$ , as we suggest in Remark 3.2, or one has to redefine the symmetry group taking into consideration the hyperelliptic involutions, as suggested in [59, fn 2, pp. 12–13]. We get the following value for  $c(\alpha_2)$ :

$$c(\alpha_2) = c(\gamma_{1,1}) = \frac{1}{8 \times 24 \times 24 \times 6}.$$

In the above computation we followed the normalization conventions chosen by Mirzakhani in [47]. As it was pointed out in Remark 1.26, the case  $(g, n) = (2, 0)$  admits an alternative normalization. Nevertheless, such an alternative normalization changes  $c(\alpha_1)$  and  $c(\alpha_2)$  by a common scaling factor and, hence, does not affect the ratio

$$\frac{c(\alpha_1)}{c(\alpha_2)} = \frac{c(\gamma_{\text{nonsep},2})}{c(\gamma_{1,1})} = 48.$$

This value is independently confirmed in [13] experimentally and in [6] and [27] theoretically.

Denote by  $c(\gamma_{\text{sep},g})$  the sum of the frequencies  $c(\gamma_{g_1,g_2})$  over all equivalence classes of separating curves, that is, over all unordered pairs  $(g_1, g_2)$  satisfying  $g_1 + g_2 = g; g_1, g_2 \geq 1$ . We are now ready to prove Theorem 1.18.

Table 2. Numerical values of ratios of the frequency of separating over the frequency of nonseparating simple closed geodesics in small genera.

$g$	2	3	4	5	$11$
Exact	$\frac{1}{48}$	$\frac{5}{1776}$	$\frac{605}{790992}$	$\frac{4697}{27201408}$	$\frac{166833285883}{5360555755385245488}$
Approximate	$2.08 \cdot 10^{-2}$	$2.82 \cdot 10^{-3}$	$7.65 \cdot 10^{-4}$	$1.73 \cdot 10^{-4}$	$3.11 \cdot 10^{-8}$
Asymptotic formula	$2.03 \cdot 10^{-2}$	$4.16 \cdot 10^{-3}$	$9.00 \cdot 10^{-4}$	$2.01 \cdot 10^{-4}$	$3.31 \cdot 10^{-8}$

*Proof of Theorem 1.18*

By Theorem 1.22, we have

$$\frac{c_{g,\text{sep}}}{c_{g,\text{nonsep}}} = \frac{\sum_{g_1=1}^{\lfloor \frac{g}{2} \rfloor} \text{Vol}(\Delta(g_1, g - g_1))}{\text{Vol}(\Gamma_1(g))}.$$

Plugging the asymptotic values (4.5) and (4.15), respectively, into the denominator and numerator of the ratio on the right-hand side, we obtain the desired asymptotic value for the ratio on the left-hand side. □

In Table 2 we present the exact (first line) and approximate (second line) values of the ratio  $\frac{c_{g,\text{sep}}}{c_{g,\text{nonsep}}}$  of the two frequencies in small genera and the value given by the asymptotic formula (1.32) (third line).

**Appendix A. Proof of the asymptotic formula for 2-correlators**

In this appendix we prove Proposition 4.1.

*Structure of the proof.* The equality  $a_{g,1} = 1 - \frac{2}{6g-1}$  immediately follows from the fact that  $a_{g,0} = 1$  and the recursive relations (4.3). For  $g = 1$ , bounds (4.4) are trivial. The symmetry  $a_{g,3g-1-k} = a_{g,k}$  allows us to confine  $k$  to the range  $\{2, 3, \dots, \lfloor \frac{3g-1}{2} \rfloor\}$ .

Using recursive relations (4.3), we evaluate  $a_{g,k}$  explicitly for  $k = 2, \dots, 5$  and prove in Lemma A.1 bounds (4.4) for these small values of  $k$ . In genera  $g = 2, 3, 4$ , the expression  $\lfloor \frac{3g-1}{2} \rfloor$  is bounded by 5, which implies bounds (4.4) for any  $k$  when  $g = 2, 3, 4$ . From this point, on we always assume that  $g \geq 5$  and  $k$  is in the range  $\{6, \dots, \lfloor \frac{3g-1}{2} \rfloor\}$ .

We start the main part of the proof by rewriting the recurrence relations (4.3) in a form convenient for estimates. Namely, we introduce the function

$$R(g, j) = \frac{\binom{3g}{3j} \binom{g}{j}}{\binom{6g}{6j}} \tag{A.1}$$

and express the right-hand side of each of the recurrence relations (4.3) as a product  $R \cdot \frac{P_i}{Q}$ , where  $P_i$ ,  $i = 1, 2, 3$ , and  $Q$  are explicit polynomials in  $g$  and  $j$ . In

Lemma A.3 we show that for any  $g$ , the absolute value of each of the rational functions  $P_i/Q$  on the range of  $j$  corresponding to  $k \in \{6, \dots, \lfloor \frac{3g-1}{2} \rfloor\}$  is bounded from above by 1. In Lemma A.4 we show that for any fixed  $g$  and  $0 \leq j \leq \lfloor \frac{g-1}{2} \rfloor$ , the expression  $R(g, j)$  is monotonically decreasing as a function of  $j$ . Combining these two lemmas, we obtain in Lemma A.5 the estimate  $-R(g, 2) \cdot \frac{g-3}{2} \leq a_{g,k} - a_{g,5} \leq R(g, 2) \cdot \frac{3g-11}{3}$  valid for all  $g$  and  $k$  under consideration.

We use explicit expressions for rational functions  $\varepsilon_{\text{below}}(g)$  and  $\varepsilon_{\text{above}}(g)$  in  $a_{g,5} = 1 - \frac{2}{6g-1} + \varepsilon_{\text{below}}(g) = 1 - \varepsilon_{\text{above}}(g)$  obtained in Lemma A.1 to prove in Lemma A.6 that for  $g \geq 5$  the inequalities  $R(g, 2) \cdot \frac{g-3}{3} < \varepsilon_{\text{below}}(g)$  and  $R(g, 2) \cdot \frac{3g-11}{3} < \varepsilon_{\text{above}}(g)$  hold.

A.1. Small values of  $k$  and  $g$

Recall that  $a_{g,0} = 1$ . Recursive relations (4.3) provide the following first several terms  $a_{g,k}$  for  $k = 1, 2, 3, 4, 5$ , where for each  $k$  in this range we indicate the smallest value of  $g$  starting from which the corresponding equality holds:

$$a_{g,0} = 1 \qquad \text{for } g \geq 1, \text{ (A.2)}$$

$$a_{g,1} = 1 - \frac{2}{6g-1} \qquad \text{for } g \geq 1, \text{ (A.3)}$$

$$a_{g,2} = 1 - \frac{12(g-1)}{(6g-1)(6g-3)} \qquad \text{for } g \geq 2,$$

$$a_{g,3} = 1 - \frac{3(24g^2 - 49g + 30)}{(6g-1)(6g-3)(6g-5)} \qquad \text{for } g \geq 3,$$

$$a_{g,4} = 1 - \frac{2}{6g-1} + \frac{9(g-2)(34g-35)}{(6g-1)(6g-3)(6g-5)(6g-7)} \qquad \text{for } g \geq 3,$$

$$a_{g,5} = 1 - \frac{2}{6g-1} + \frac{27(68g^3 - 308g^2 + 519g - 280)}{(6g-1)(6g-3)(6g-5)(6g-7)(6g-9)} \qquad \text{(A.4)}$$

$$= 1 - \frac{9(g-2)(288g^3 - 780g^2 + 1012g - 525)}{(6g-1)(6g-3)(6g-5)(6g-7)(6g-9)} \qquad \text{for } g \geq 4. \text{ (A.5)}$$

LEMMA A.1

For all  $g \in \mathbb{N}$  and all  $k \in \mathbb{N}$  satisfying  $2 \leq k \leq \min(5, \lfloor \frac{3g-1}{2} \rfloor)$ , the relations (4.4) are valid:

$$1 - \frac{2}{6g-1} = a_{g,1} < a_{g,k} < a_{g,0} = 1.$$



*Proof*

By (A.2) and (A.3), the terms  $a_{g,0}$  and  $a_{g,1}$  indeed have values as claimed in the statement of Lemma A.1.

It follows from recurrence relations (4.3) that for  $k$  satisfying  $0 \leq k \leq \lfloor \frac{3g-1}{2} \rfloor - 1$  we have  $a_{g,k+1} < a_{g,k}$  if and only if  $k \equiv 0 \pmod{3}$ , and we have  $a_{g,k+1} > a_{g,k}$  for the remaining  $k$  in this range. In particular, for  $g \geq 2$  the difference  $a_{g,2} - a_{g,1}$  is strictly positive, which implies the desired lower bound (4.4) for  $a_{g,2}$  when  $g \geq 2$ . The explicit expression for  $a_{g,2}$  when  $g \geq 2$  implies the desired strict upper bound (4.4).

Recursive relations (4.3) imply that for  $g \geq 3$  the difference  $a_{g,3} - a_{g,2}$  is strictly positive. Since  $a_{g,2}$  satisfies the desired lower bound (4.4), the term  $a_{g,3}$  also does. The quadratic polynomial  $(24g^2 - 49g + 30)$  in the numerator of the explicit expression for  $a_{g,3}$  admits only strictly positive values, which implies the upper bound (4.4) for  $a_{g,3}$  when  $g \geq 3$ .

Recursive relations (4.3) imply that for  $g \geq 3$  the difference  $a_{g,4} - a_{g,3}$  is strictly negative. Since  $a_{g,3}$  satisfies the desired upper bound (4.4), the term  $a_{g,4}$  also does. The explicit expression for  $a_{g,4}$  implies the lower bound (4.4) for  $a_{g,4}$  when  $g \geq 3$ .

Finally, recursive relations (4.3) imply that for  $g \geq 4$  the difference  $a_{g,5} - a_{g,4}$  is strictly positive, which implies the desired lower bound (4.4) for  $a_{g,5}$  when  $g \geq 4$ . It remains to verify that the polynomial  $(288g^3 - 780g^2 + 1012g - 525)$  in the explicit expression (A.5) for  $a_{g,5}$  attains only strictly positive values for  $g \geq 4$  to prove the desired upper bound (4.4) for  $a_{g,5}$ . For  $g \geq 4$ , we have

$$\begin{aligned} 288g^3 - 780g^2 + 1012g - 525 &> 288g^3 - 864g^2 + 864g - 576 \\ &= 288((g - 1)^3 - 1) > 0. \end{aligned} \quad \square$$

COROLLARY A.2

For any  $g$  in  $\{1, 2, 3, 4\}$  and for all  $k \in \mathbb{N}$  satisfying  $2 \leq k \leq 3g - 3$ , the desired bounds (4.4) are valid:

$$1 - \frac{2}{6g - 1} = a_{g,1} < a_{g,k} < a_{g,0} = 1.$$

*Proof*

Recall that the symmetry  $a_{g,k} = a_{g,3g-1-k}$  allows one to limit  $k$  to the range  $2 \leq k \leq \lfloor \frac{3g-1}{2} \rfloor$ . Thus, for  $g \leq 4$  the largest possible value of  $k$  satisfying  $k \leq \lfloor \frac{3g-1}{2} \rfloor$  equals  $\lfloor \frac{3 \cdot 4 - 1}{2} \rfloor = 5$ . The proof of bounds (4.4) for  $k \leq 5$  and any  $g$  is already completed in Lemma A.1. □

A.2. Alternative form of recurrence relations

We start by extracting the common factor in relations (4.3) and simplifying it. Define the following polynomial in  $g$  and  $j$ :

$$Q(g, j) = g(6g - 6j - 1)(6g - 6j - 3). \tag{A.6}$$

Rewriting double factorials in terms of factorials and rearranging, we obtain

$$\begin{aligned} & \frac{(6g - 6j - 5)!!}{(6g - 1)!!} \cdot \frac{(6j - 1)!! \cdot (g - 1)!}{j!(g - j)!} \\ &= \left( \frac{(6g - 6j - 5)!}{(3g - 3j - 3)!2^{3g-3j-3}} \right) \left( \frac{(3g - 1)!2^{3g-1}}{(6g - 1)!} \right) \left( \frac{(6j - 1)!}{(3j - 1)!2^{3j-1}} \right) \frac{(g - 1)!}{j!(g - j)!} \\ &= 8 \cdot \left( \frac{(6j - 1)!(6g - 6j - 5)!}{(6g - 1)!} \right) \left( \frac{(3g - 1)!}{(3j - 1)!(3g - 3j - 3)!} \right) \left( \frac{(g - 1)!}{j!(g - j)!} \right) \\ &= 8 \cdot \left( \frac{(6j)!(6g - 6j)!}{(6g)!} \right) \left( \frac{(3g)!}{(3j)!(3g - 3j)!} \right) \left( \frac{g!}{j!(g - j)!} \right) \\ &\quad \cdot \frac{6g}{(6j) \cdot (6g - 6j)(6g - 6j - 1)(6g - 6j - 2)(6g - 6j - 3)(6g - 6j - 4)} \\ &\quad \cdot \frac{(3j) \cdot (3g - 3j)(3g - 3j - 1)(3g - 3j - 2)}{3g} \cdot \frac{1}{g} \\ &= \frac{\binom{3g}{3j} \binom{g}{j}}{\binom{6g}{6j}} \cdot \frac{1}{g \cdot (6g - 6j - 1)(6g - 6j - 3)} = R(g, j) \cdot \frac{1}{Q(g, j)}. \end{aligned}$$

Defining the following polynomials in  $g$  and  $j$ ,

$$P_1(g, j) = (6g - 6j - 1)(6g - 6j - 3)(g - 2j), \tag{A.7}$$

$$P_2(g, j) = -2(6g - 6j - 3)(6j + 1)(g - j), \tag{A.8}$$

$$P_3(g, j) = 2(6j + 1)(6j + 3)(g - j), \tag{A.9}$$

we can express the recurrence relations (4.3) as

$$a_{g,3j} - a_{g,3j-1} = R(g, j) \cdot \frac{P_1(g, j)}{Q(g, j)}, \quad \text{where } 3 \leq 3j \leq \left\lceil \frac{3g - 1}{2} \right\rceil, \tag{A.10}$$

$$a_{g,3j+1} - a_{g,3j} = R(g, j) \cdot \frac{P_2(g, j)}{Q(g, j)}, \quad \text{where } 1 \leq 3j + 1 \leq \left\lceil \frac{3g - 1}{2} \right\rceil, \tag{A.11}$$

$$a_{g,3j+2} - a_{g,3j+1} = R(g, j) \cdot \frac{P_3(g, j)}{Q(g, j)}, \quad \text{where } 2 \leq 3j + 2 \leq \left\lceil \frac{3g - 1}{2} \right\rceil. \tag{A.12}$$

LEMMA A.3

For any  $g \in \mathbb{N}$ , the following bounds are valid:

$$0 < \frac{P_1(g, j)}{Q(g, j)} < 1, \quad \text{where } 3 \leq 3j \leq \left\lceil \frac{3g - 1}{2} \right\rceil, \tag{A.13}$$

$$-1 < \frac{P_2(g, j)}{Q(g, j)} < 0, \quad \text{where } 1 \leq 3j + 1 \leq \left\lceil \frac{3g-1}{2} \right\rceil, \tag{A.14}$$

$$0 < \frac{P_3(g, j)}{Q(g, j)} < 1, \quad \text{where } 2 \leq 3j + 2 \leq \left\lceil \frac{3g-1}{2} \right\rceil. \tag{A.15}$$

*Proof*

The bounds  $3 \leq 3j \leq \lceil \frac{3g-1}{2} \rceil$  in (A.13) imply that  $0 < j < \frac{g}{2}$ . Dividing expression (A.7) for  $P_1(g, j)$  by expression (A.6) for  $Q(g, j)$ , we obtain

$$\frac{P_1(g, j)}{Q(g, j)} = \frac{g - 2j}{g}.$$

Clearly,

$$0 < \frac{g - 2j}{g} < 1$$

for any  $g \in \mathbb{N}$  and for all  $j$  satisfying  $0 < j < \frac{g}{2}$ .

The bounds  $1 \leq 3j + 1 \leq \lceil \frac{3g-1}{2} \rceil$  in (A.14) imply that  $0 \leq j < \frac{g}{2}$ . Dividing expression (A.8) for  $P_2(g, j)$  by expression (A.6) for  $Q(g, j)$ , we obtain

$$\begin{aligned} \frac{P_2(g, j)}{Q(g, j)} &= -2 \cdot \frac{6j + 1}{6g - 6j - 1} \cdot \frac{g - j}{g} = -\frac{6g - 6j}{6g - 6j - 1} \cdot \frac{2j + \frac{1}{3}}{g} \\ &= -\left(1 + \frac{1}{6g - 6j - 1}\right) \cdot \left(\frac{2j + \frac{1}{3}}{g}\right). \end{aligned}$$

Since  $0 \leq 2j \leq g - 1$ , the latter expression is always strictly negative for this range of  $j$ . Both factors in the brackets in the latter expression are monotonically increasing on this range of  $j$ , so the maximum of the absolute value of the product is attained at  $2j = g - 1$ . We obtain

$$\begin{aligned} 0 < \left| \frac{P_2(g, j)}{Q(g, j)} \right| &\leq \left(1 + \frac{1}{6g - (3g - 3) - 1}\right) \cdot \left(\frac{g - \frac{2}{3}}{g}\right) \\ &= \left(1 + \frac{1}{3g + 2}\right) \cdot \left(1 - \frac{1}{\frac{3}{2}g}\right) < 1 \quad \text{for } g \in \mathbb{N} \text{ and } 0 \leq j < \frac{g}{2}. \end{aligned} \tag{A.16}$$

The bounds  $2 \leq 3j + 2 \leq \lceil \frac{3g-1}{2} \rceil$  in (A.15) imply that  $0 \leq j < \frac{g}{2}$  and that  $6j \leq 3g - 5$ . Dividing expression (A.9) for  $P_3(g, j)$  by expression (A.6) for  $Q(g, j)$ , we obtain

$$\frac{P_3(g, j)}{Q(g, j)} = 2 \cdot \frac{(6j + 1)(6j + 3)(g - j)}{g(6g - 6j - 1)(6g - 6j - 3)} = \left| \frac{P_2(g, j)}{Q(g, j)} \right| \cdot \frac{6j + 3}{6g - 6j - 3}.$$

The expression  $\frac{6j+3}{6g-6j-3}$  is strictly positive and monotonically increasing on the range of  $j$  under consideration, so it attains its maximum on the largest possible value of  $j$ . Since  $6j \leq 3g - 5$ , we obtain

$$0 < \frac{6j+3}{6g-6j-3} \leq \frac{3g-2}{3g+2} < 1 \quad \text{for } 2 \leq 3j+2 \leq \left\lfloor \frac{3g-1}{2} \right\rfloor \text{ and } g \in \mathbb{N}.$$

Combined with (A.16) this proves the desired bounds (A.15), which completes the proof of Lemma A.3.  $\square$

LEMMA A.4

For any fixed value of  $g \in \mathbb{N}$ , the expression  $R(g, j)$  considered as a function of  $j$  is strictly monotonically decreasing on the range  $\{0, 1, \dots, \lfloor \frac{g-1}{2} \rfloor\}$  of the argument  $j$ .

*Proof*

It is immediate to verify that

$$\begin{aligned} & R(g, j+1)/R(g, j) \\ &= \frac{(6j+5)(6j+3)(6j+1)}{(6g-6j-1)(6g-6j-3)(6g-6j-5)} \cdot \frac{g-j}{j+1} \\ &= \frac{j+5/6}{j+1} \cdot \frac{6j+1}{6g-6j-5} \cdot \frac{6j+3}{6g-6j-3} \cdot \frac{6g-6j}{6g-6j-1} \\ &= \left(1 - \frac{1}{6j+6}\right) \cdot \frac{6j+1}{6g-6j-5} \cdot \frac{6j+3}{6g-6j-3} \cdot \left(1 + \frac{1}{6g-6j-1}\right). \quad (\text{A.17}) \end{aligned}$$

For any fixed  $g \in \mathbb{N}$ , each of the four terms in the last line of the above expression is strictly monotonically increasing as a function of  $j$  on the range  $\{0, 1, \dots, \lfloor \frac{g-1}{2} \rfloor\}$ . It is immediate to verify that when  $2j = g - 1$ , the product of four terms in the last line of the above expression is identically equal to 1 for all  $g \in \mathbb{N}$ , and the lemma follows.  $\square$

LEMMA A.5

For any  $g \in \mathbb{N}$  and for any integer  $k$  in the range  $6 \leq k \leq \lfloor \frac{3g-1}{2} \rfloor$ , the following bounds hold:

$$-R(g, 2) \cdot \frac{g-3}{2} \leq a_{g,k} - a_{g,5} \leq R(g, 2) \cdot \frac{3g-11}{3}. \quad (\text{A.18})$$

*Proof*

We can assume that  $g \geq 5$ ; otherwise the range of  $k$  is empty and the statement is vacuous.

Consider the sequence  $\{a_{g,5}, a_{g,6}, \dots, a_{g,k}\}$ , and denote by  $n_+(k)$  the number of entries  $m$  in the set  $\{5, 6, \dots, k - 1\}$  for which the inequality  $a_{g,m+1} > a_{g,m}$  holds. Similarly, denote by  $n_-(k)$  the number of entries in the same set for which the inequality  $a_{g,m+1} < a_{g,m}$  holds.

Combining the recurrence relations in the form (A.10)–(A.12), bounds (A.13)–(A.15), and Lemma A.4, we conclude that

$$-R(g, 2) \cdot n_-(k) \leq a_{g,k} - a_{g,5} \leq R(g, 2) \cdot n_+(k).$$

It remains to translate the restriction  $k \leq \lfloor \frac{3g-1}{2} \rfloor$  into upper bounds for  $n_+(k)$  and  $n_-(k)$  as functions of  $g$ .

Recall that it follows from recurrence relations (4.3) that for  $m$  satisfying  $0 \leq m \leq \lfloor \frac{3g-1}{2} \rfloor - 1$  we have  $a_{g,m+1} < a_{g,m}$  if and only if  $m \equiv 0 \pmod{3}$  and we have  $a_{g,m+1} > a_{g,m}$  for the remaining  $m$  in this range. This implies that

$$\begin{aligned} n_+ &= 2(j - 2), & n_- &= j - 2, & \text{when } k &= 3j - 1, \\ n_+ &= 2(j - 2) + 1, & n_- &= j - 2, & \text{when } k &= 3j, \\ n_+ &= 2(j - 2) + 1, & n_- &= j - 1, & \text{when } k &= 3j + 1. \end{aligned}$$

In all these cases, we have

$$\begin{aligned} n_+(k) &\leq \frac{2k - 10}{3}, \\ n_-(k) &\leq \frac{k - 4}{3}. \end{aligned}$$

By assumption,  $k \leq \lfloor \frac{3g-1}{2} \rfloor$ , so the latter bounds imply that

$$\begin{aligned} n_+(k) &\leq \frac{3g - 11}{3}, \\ n_-(k) &\leq \frac{g - 3}{2}, \end{aligned}$$

and (A.18) follows. □

We assume that  $g \geq 5$  and  $k \in \{6, \dots, \lfloor \frac{3g-1}{2} \rfloor\}$ . Define

$$\varepsilon_{\text{below}}(g) = \frac{27(68g^3 - 308g^2 + 519g - 280)}{(6g - 1)(6g - 3)(6g - 5)(6g - 7)(6g - 9)}, \tag{A.19}$$

$$\varepsilon_{\text{above}}(g) = \frac{9(g - 2)(288g^3 - 780g^2 + 1012g - 525)}{(6g - 1)(6g - 3)(6g - 5)(6g - 7)(6g - 9)}. \tag{A.20}$$

In this notation, expressions (A.4) and (A.5) for  $a_{g,5}$  can be written as

$$a_{g,5} = 1 - \frac{2}{6g-1} + \varepsilon_{\text{below}}(g) = 1 - \varepsilon_{\text{above}}(g). \quad (\text{A.21})$$

LEMMA A.6

For any integer  $g$  satisfying  $g \geq 5$ , the following strict inequalities are valid:

$$R(g, 2) \cdot \frac{g-3}{2} < \varepsilon_{\text{below}}(g),$$

$$R(g, 2) \cdot \frac{3g-11}{3} < \varepsilon_{\text{above}}(g).$$

*Proof*

The proof is a straightforward calculation. First note that all the quantities  $R(g, 2)$ ,  $(g-3)$ ,  $(3g-11)$ ,  $\varepsilon_{\text{below}}(g)$ ,  $\varepsilon_{\text{above}}(g)$  are strictly positive for  $g \geq 5$ , where strict positivity of  $\varepsilon_{\text{below}}(g)$  and of  $\varepsilon_{\text{above}}(g)$  was proved in Lemma A.1. Thus, it is sufficient to prove that

$$\frac{2\varepsilon_{\text{below}}(g)}{R(g, 2) \cdot (g-3)} > 1 \quad \text{for } g \geq 5, \quad (\text{A.22})$$

$$\frac{3\varepsilon_{\text{above}}(g)}{R(g, 2) \cdot (3g-11)} > 1 \quad \text{for } g \geq 5. \quad (\text{A.23})$$

Applying definition (A.1) to evaluate  $R(g, 2)$  and canceling out common factors in the numerator and in the denominator of the resulting expression allows us to obtain

$$R(g, 2) = \frac{10395}{2} \cdot \frac{g(g-1)}{(6g-1)(6g-3)(6g-5)(6g-7)(6g-9)(6g-11)}. \quad (\text{A.24})$$

Plug expression (A.19) for  $\varepsilon_{\text{below}}(g)$  and the above expression (A.24) for  $R(g, 2)$  into the left-hand side of (A.22) and cancel out the common factors in the numerator and in the denominator of the resulting expression. Then applying polynomial division with remainder to the resulting numerator and denominator, we obtain

$$\begin{aligned} \frac{2\varepsilon_{\text{below}}(g)}{R(g, 2) \cdot (g-3)} &= 2 \cdot 27 \cdot \frac{2}{10395} \cdot \frac{(68g^3 - 308g^2 + 519g - 280)(6g-11)}{g(g-1)(g-3)} \\ &= \frac{4}{385} \cdot \left( 408g - 964 + \frac{1422g^2 - 4497g + 3080}{g(g-1)(g-3)} \right). \end{aligned}$$

It is immediate to check that  $(1422g^2 - 4497g + 3080)$  is positive for  $g \geq 5$ . It remains to note that, for  $g \geq 5$ , we have  $(408g - 964) \geq (408 \cdot 5 - 964) = 1076 > 385/4$ , which completes the proof of (A.22).

Performing analogous manipulations, we obtain

$$\begin{aligned} & \frac{3\varepsilon_{\text{above}}(g)}{R(g, 2) \cdot (3g - 11)} \\ &= 3 \cdot 9 \cdot \frac{2}{10395} \cdot \frac{(g - 2)(288g^3 - 780g^2 + 1012g - 525)(6g - 11)}{g(g - 1)(3g - 11)} \\ &= \frac{2}{385} \cdot \left( 576g^2 - 1080g + 2964 + \frac{9790g^2 + 1735g - 11550}{g(g - 1)(3g - 11)} \right). \end{aligned}$$

It is immediate to check that  $(9790g^2 + 1735g - 11550)$  is positive for  $g \geq 5$  as well as the denominator of the corresponding fraction. It remains to note that the function  $(576g^2 - 1080g + 2964)$  is monotonically increasing on the interval  $[5; +\infty[$ , so for  $g \geq 5$  we obtain

$$576g^2 - 1080g + 2964 \geq 576 \cdot 5^2 - 1080 \cdot 5 + 2964 = 11964 > 385/2,$$

which completes the proof of (A.23). □

*Proof of Proposition 4.1*

For small genera,  $g = 1, 2, 3, 4$ , Proposition 4.1 was proved in Corollary A.2.

For genera  $g \geq 5$  and  $k = 2, 3, 4, 5$ , Proposition 4.1 was proved in Lemma A.1. The symmetry  $a_{g,k} = a_{g,3g-1-k}$  implies Proposition 4.1 for symmetric values of  $k$ .

For  $g \geq 5$  and  $k$  in the range  $6 \leq k \leq \lfloor \frac{3g-1}{2} \rfloor$ , Proposition 4.1 immediately follows from the combination of Lemma A.5, expression (A.21) for  $a_{g,5}$ , and Lemma A.6. The symmetry  $a_{g,k} = a_{g,3g-1-k}$  implies Proposition 4.1 for symmetric values of  $k$ . □

*A.3. Asymptotic behavior of normalized 2-correlators in large genera*

In this section we briefly describe the behavior of  $a_{g,k}$  for  $g \gg 1$ . (A more detailed discussion will be given in a future and more general paper.)

When  $g \rightarrow +\infty$  and  $j$  remains bounded, expressions (A.7)–(A.9) and (A.6) for polynomials  $P_i(g, j), i = 1, \dots, 4$ , and  $Q(g, j)$ , respectively, imply that

$$\begin{aligned} \frac{P_1(g, j)}{Q(g, j)} &= 1 - (2j) \cdot \frac{1}{g}, \\ \frac{P_2(g, j)}{Q(g, j)} &= -\left(2j + \frac{1}{3}\right) \cdot \frac{1}{g} + o\left(\frac{1}{g}\right), \\ \frac{P_3(g, j)}{Q(g, j)} &= \left(2j + \frac{1}{3}\right)\left(j + \frac{1}{2}\right) \cdot \frac{1}{g^2} + o\left(\frac{1}{g^2}\right), \end{aligned}$$

as  $g \rightarrow +\infty$ . In particular, for  $g \gg 1$  we see that for small values of  $j$  the ratio  $\frac{P_1(g, j)}{Q(g, j)}$  is close to 1, while the ratio  $\frac{P_2(g, j)}{Q(g, j)}$  is of the order  $\frac{1}{g}$  and the ratio

$\frac{P_3(g,j)}{Q(g,j)}$  is of the order  $\frac{1}{g^2}$ . Thus, assuming that  $g \gg 1$ , and taking consecutive terms  $\{a_{g,3j-1}, a_{g,3j}, a_{g,3j+1}, a_{g,3j+1}\}$  with  $j \ll g$ , we observe a certain increment from  $a_{g,3j-1}$  to  $a_{g,3j}$ , a much smaller decrement from  $a_{g,3j}$  to  $a_{g,3j+1}$ , and a very small increment from  $a_{g,3j+1}$  to  $a_{g,3j+2}$ .

For any fixed  $g \gg 1$  and  $j \ll g$ , the quantity  $R(g, j)$  defined by (A.1) is very rapidly decreasing as  $j$  grows. We conclude from expression (A.17) that for bounded  $j$  and  $g \rightarrow +\infty$  one has

$$R(g, j + 1) = R(g, j) \cdot \frac{(j + \frac{5}{6})(j + \frac{3}{6})(j + \frac{1}{6})}{(j + 1)} \cdot \frac{1}{g^2} \cdot (1 + o(1)).$$

Since  $R(g, 0) = 1$ , we get the following expressions for  $j = 0, 1, 2, 3$ :

$$R(g, 0) = 1,$$

$$R(g, 1) = \frac{(0 + \frac{5}{6})(0 + \frac{3}{6})(0 + \frac{1}{6})}{(0 + 1)} \cdot \frac{1}{g^2} \cdot (1 + o(1)) = \frac{5}{72} \cdot \frac{1}{g^2} \cdot (1 + o(1)),$$

$$R(g, 2) = \frac{(1 + \frac{5}{6})(1 + \frac{3}{6})(1 + \frac{1}{6})}{(1 + 1)} \cdot \frac{5}{72} \cdot \frac{1}{g^4} \cdot (1 + o(1)) = \frac{385}{3456} \cdot \frac{1}{g^4} \cdot (1 + o(1)),$$

$$\begin{aligned} R(g, 3) &= \frac{(2 + \frac{5}{6})(2 + \frac{3}{6})(2 + \frac{1}{6})}{(2 + 1)} \cdot \frac{385}{3456} \cdot \frac{1}{g^6} \cdot (1 + o(1)) \\ &= \frac{425425}{746496} \cdot \frac{1}{g^6} \cdot (1 + o(1)). \end{aligned}$$

Lemma A.5 admits the following immediate generalization.

LEMMA A.7

For any  $g \in \mathbb{N}$  and for any integer  $k$  in the range  $3j \leq k \leq [\frac{3g-1}{2}]$ , the following bounds hold:

$$-R(g, j) \cdot \frac{g - j - 1}{2} \leq a_{g,k} - a_{g,3j-1} \leq R(g, j) \cdot \frac{3g - 2j - 7}{3}. \tag{A.25}$$

Thus, having found the asymptotic expansion (when  $g \rightarrow +\infty$ ) in  $\frac{1}{g}$  up to the term  $\frac{1}{g^{2j-2}}$  for some  $a_{g,3j-1}$ , we get the same asymptotic expansion up to the term  $\frac{1}{g^{2j-2}}$  for all  $a_{g,k}$  with  $k$  satisfying  $3j - 1 \leq k \leq 3g - 3j$ .

Finally, no matter whether  $g = 2j$  or  $g = 2j + 1$ , one easily derives from Stirling's formula that

$$R(g, j) \approx \frac{1}{2^{2g-1}} \cdot \frac{1}{\sqrt{\pi g}}.$$



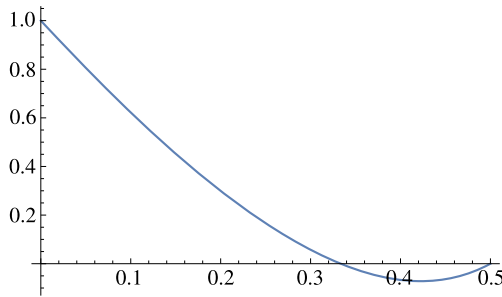


Figure 6. (Color online) Graph of  $f(x)$ .

Morally, when  $g \gg 1$  and the index  $k$  is located sufficiently far from the extremities of the range  $\{0, 1, \dots, 3g - 1\}$ , the values of  $a_{g,k}$  become, basically, indistinguishable.

It is curious to note that the sequence  $\{a_{g,2}, a_{g,5}, \dots, a_{g,3j_{\max}-1}\}$ , where  $j_{\max}$  is the maximum integer satisfying  $3j_{\max} - 1 \leq \lfloor \frac{3g-1}{2} \rfloor$ , is not monotonically increasing for  $g \geq 17$ . Let  $x = \frac{j}{g}$ . Our bounds on  $j$  imply that  $0 < x < \frac{1}{2}$ . For any fixed  $g \gg 1$ , define

$$f(x) = f\left(\frac{j}{g}\right) = \frac{(P_1(g, j) + P_2(g, j) + P_3(g, j))}{Q(g, j)}.$$

For large values of  $g$  the graph of  $f(x)$  has the form as in Figure 6, so up to some point the function  $f$  remains positive and the sequence  $a_{g,2}, a_{g,5}, \dots$  monotonically increases, but then it attains its maximum and very slowly monotonically decreases down to  $a_{g,3j_{\max}-1}$ .

**Appendix B. Stable graphs: Formal definition**

Following Kontsevich [39], we now introduce the definition of a stable graph. Let  $S$  be a closed oriented surface of genus  $g$  without boundary endowed with  $n$  labeled marked points. Let  $\gamma$  be a simple closed primitive multicurve on  $S$ . Consider the decomposition of  $S \setminus \gamma$  into a union of surfaces with boundaries endowed with marked points. The dual graph  $\Gamma$  to this decomposition is constructed as follows.

- Each connected component  $S_j$  of  $S \setminus \gamma$  gives rise to a vertex  $v_j$  of  $\Gamma$  decorated by the genus  $g(S_j)$ . The marked points on  $S_j$  are encoded by the *legs* attached to  $v_j$ , the boundary components of  $S_j$  correspond to half-edges incident to  $v_j$ .
- Each component  $\gamma_i$  of  $\gamma$  gives rise to an edge of  $\Gamma$ . When  $\gamma_i$  is the common boundary of two distinct connected components of  $S \setminus \gamma$ , the corresponding edge of  $\Gamma$  joins the two distinct vertices of  $\Gamma$  representing these two connected components. When both sides of  $\gamma_i$  are at the boundary of the same connected

component of  $S \setminus \gamma$ , the edge of  $\Gamma$  dual to  $\gamma_i$  is a loop joining the vertex associated to the corresponding connected component to itself.

Figures in the tables of Appendix C.2 illustrate the correspondence between multicurves and stable graphs in genus 2.

We now present the following formal definition.

*Definition B.1*

Consider a 6-tuple  $\Gamma = (V, H, \iota, \alpha, \mathbf{g}, L)$ , where

- $V$  is a finite set of *vertices*;
- $H$  is a finite set of *half-edges*;
- $\iota : H \rightarrow H$  is an involution; the fixed points of  $\iota$  are called the *legs* and the 2-cycles of  $\iota$  are called the *edges* of  $\Gamma$ ;
- $\alpha : H \rightarrow V$  is a map that attaches a half-edge to a vertex; the number of half-edges at a given vertex  $v$  is denoted by  $n_v := |\alpha^{-1}(v)|$ ;
- the graph is connected: for each pair of vertices  $(u, v)$  there exists a sequence of half-edges  $(h_1, h'_1, h_2, h'_2, \dots, h_k, h'_k)$  such that  $\iota(h_i) = h'_i, u = \alpha(h_1), v = \alpha(h'_k)$  and  $\alpha(h'_i) = \alpha(h_{i+1})$ ;
- $\mathbf{g} = \{g_v\}_{v \in V}$  is a set of nonnegative integers, one at each vertex, called the *genus decoration*;
- $L$  is a bijection from the set of legs to  $\overline{\{1, \dots, n\}}$ .

Such a 6-tuple  $\Gamma$  is called a *stable graph* for  $\overline{\mathcal{M}}_{g,n}$  if the genus decoration  $\mathbf{g}$  satisfies the following conditions:

- $g(\Gamma) = h_1(\Gamma) + \sum_{v \in V} g_v$ , where  $h_1(\Gamma)$  is the first Betti number of the graph,
- *stability condition*  $2g_v - 2 + n_v > 0$  at each vertex  $v$  of  $\Gamma$ .

To a stable graph  $\Gamma$ , we associate an underlying graph whose vertex set is  $V$  and each 2-cycle  $(h, h')$  of  $\iota$  gives an edge attached to  $\alpha(h)$  and  $\alpha(h')$ . We denote the set of these edges by  $E = E(\Gamma)$ . Such a graph can have multiple edges and loops. The additional information carried by a stable graph is the genus decoration  $\mathbf{g}$  and the  $n$  legs.

Two stable graphs  $\Gamma = (V, H, \iota, \alpha, \mathbf{g}, L)$  and  $\Gamma' = (V', H', \iota', \alpha', \mathbf{g}', L')$  are isomorphic if there exist two bijections  $\phi : V \rightarrow V'$  and  $\psi : H \rightarrow H'$  that preserve edges, legs, and genus decoration; that is,

$$\psi \circ \iota = \iota' \circ \psi, \quad L'(\psi(h)) = L(h), \quad g'_{\phi(v)} = g_v.$$

Note that automorphisms of stable graphs are allowed to interchange edges and vertices respecting the decoration, but not the legs which are numbered by  $L$ .

We denote by  $\mathcal{G}_{g,n}$  the set of isomorphism classes of stable graphs with given genus  $g$  and number of legs  $n$ .

As we have already mentioned, each stable graph in  $\mathcal{G}_{g,n}$  corresponds to a cycle of the Deligne–Mumford compactification  $\overline{\mathcal{M}}_{g,n}$ . More precisely, each vertex  $v$  of the graph corresponds to the component of a nodal curve of genus  $g_v$  and contains the marked points corresponding to the legs attached to this vertex. Each edge of  $E(\Gamma)$  represents a node. (See the survey [54] for an excellent introduction to this subject and for beautiful illustrations.) Hence, the unique stable graph with no edge and  $n$  legs in  $\mathcal{G}_{g,n}$  corresponds to the component  $\overline{\mathcal{M}}_{g,n}$  (the smooth curves).

**Appendix C. Examples of explicit calculations**

*C.1. The cases of  $\mathcal{Q}_{0,3}$  and  $\mathcal{Q}_{1,1}$*

We now consider the moduli spaces  $\mathcal{M}_{0,3}$  and  $\mathcal{M}_{1,1}$  and the associated cotangent bundles  $\mathcal{Q}_{0,3}$  and  $\mathcal{Q}_{1,1}$ .

There is a unique complex curve  $C$  in  $\mathcal{M}_{0,3}$  which is  $\mathbb{C}P^1 \setminus \{0, 1, \infty\}$ . This curve does not admit any nonzero quadratic differentials with at most simple poles at the marked points 0, 1 and  $\infty$  and with no other poles (which is coherent with the fact that the tangent space to a point is 0). There is a unique stable graph for  $\mathcal{M}_{0,3}$ :

$$\mathcal{G}(0, 3) = \left\{ \text{---} \bullet \begin{array}{l} \diagup \\ \diagdown \end{array} \right\}.$$

0

Let us denote this graph by  $\Phi_{0,3}$ . If we try to apply Theorem 1.5 to define the value of  $\text{Vol } \mathcal{Q}_{0,3}$ , then the polynomial  $P_{\Phi_{0,3}}$  defined by (1.12) is ill defined since  $4g - 4 + n = 6g - 7 + 2n = -1$  and  $(-1)!$  makes no sense. However, taking limits as in (1.28), we obtain the value

$$P_{\Phi_{0,3}} = 2 \cdot \frac{(-4 + n)!}{(-7 + 2n)!} \Big|_{n=3} = 4,$$

leading to  $\text{Vol } \mathcal{Q}_{0,3} = 4$  by means of (1.13) which is coherent with the value in (1.15) evaluated for  $n = 3$  and also coherent with (1.24).

For  $(g, n) = (1, 1)$ , there are two stable graphs as given below:

$$\mathcal{G}(1, 1) = \left\{ \text{---} \bullet_1 \text{---} \bullet_0 \bigcirc \right\}.$$

Only the second graph contributes to expression (1.13) from Theorem 1.5 for  $\text{Vol } \mathcal{Q}_{1,1}$  and to expression (1.23) from Theorem 1.16 for  $c_{\text{area}}(\mathcal{Q}_{1,1})$ . Let us denote this graph by  $\Phi_{1,1}$ . Applying (1.12), one finds  $P_{\Phi_{1,1}} = 4b_1$  from which we deduce that

$$\text{Vol } \mathcal{Q}_{1,1} = \mathcal{Z}(P_{\Phi_{1,1}}) = 4\zeta(2) = \frac{2\pi^2}{3} \quad \text{and}$$

$$c_{\text{area}}(\mathcal{Q}_{1,1}) = \frac{3}{\pi^2} \mathcal{Z}(\partial_{\Phi} P_{\Phi}) = 2\pi^2,$$

which coincides with our convention (1.25).

Note that any element of  $\mathcal{Q}_{1,1}$  is the square of a holomorphic 1-form (regular at the marked point). In other words, the principal stratum  $\mathcal{Q}(1, -1)$  is empty and definition (1.2) of  $\text{Vol } \mathcal{Q}_{g,n}$  does not apply. One can make geometric sense of the values  $\text{Vol } \mathcal{Q}_{1,1}$  and  $c_{\text{area}}(\mathcal{Q}_{1,1})$  by considering square-tiled surfaces for the stratum  $\mathcal{H}(0)$  of holomorphic Abelian differentials.

C.2. Holomorphic quadratic differentials in genus 2

We start by evaluation of the Siegel–Veech constant  $c_{\text{area}}(\mathcal{Q}_2)$ . Note that certain graphs do not contribute to  $c_{\text{area}}$  at all (see Table 3).

Taking the sum of the four contributions, we obtain

$$\left( \left( 1 \cdot \frac{8}{225} + 1 \cdot \frac{1}{675} \right) + \left( \frac{5}{2} \cdot \frac{1}{135} + 3 \cdot \frac{2}{405} \right) \right) \cdot \pi^6 = \left( \frac{1}{27} + \frac{1}{30} \right) \cdot \pi^6 = \frac{19}{270} \cdot \pi^6.$$

Dividing by  $\text{Vol } \mathcal{Q}_2 = \frac{1}{15} \cdot \pi^6$ , we get the answer which matches the value found in [35]:  $\frac{\pi^2}{3} \cdot c_{\text{area}}(\mathcal{Q}_2) = \left( \frac{19}{270} \pi^6 \right) : \left( \frac{1}{15} \pi^6 \right) = \frac{19}{18}$ .

The computation of the Masur–Veech volume  $\text{Vol } \mathcal{Q}_2$  was presented in Table 1 in Section 1.4. The first two graphs in Table 1 represent the contribution to the volume

Table 3. Computation of  $c_{\text{area}}(\mathcal{Q}_2)$ —Contributions to the sum in the RHS of equation (1.23).

	$\frac{128}{5} \cdot 1 \cdot \frac{1}{8} \cdot b_1 b_2 \cdot N_{0,4}(b_1, b_1, b_2, b_2)$ $= \frac{16}{5} \cdot b_1 b_2 \cdot \left( \frac{1}{4} (2b_1^2 + 2b_2^2) \right)$ $\stackrel{\partial \Gamma}{\mapsto} \frac{8}{5} (1 \cdot b_1 b_2^3 + 1 \cdot b_1^3 b_2) \stackrel{Z}{\mapsto}$	$(1 + 1)$ $\cdot \frac{8}{5} \cdot 3! \zeta(4) \cdot 1! \zeta(2)$ $= 1 \cdot \frac{8}{225} \cdot \pi^6$
	$\frac{128}{5} \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot b_1 b_2 \cdot N_{0,3}(b_1, b_1, b_2) \cdot N_{1,1}(b_2)$ $= \frac{32}{5} \cdot b_1 b_2 \cdot 1 \cdot \left( \frac{1}{48} b_2^2 \right) \stackrel{\partial \Gamma}{\mapsto} \frac{2}{15} \cdot 1 \cdot b_1 b_2^3 \stackrel{Z}{\mapsto}$	$1 \cdot \frac{2}{15} \cdot 1! \zeta(2) \cdot 3! \zeta(4)$ $= 1 \cdot \frac{1}{675} \cdot \pi^6$
	$\frac{128}{5} \cdot \frac{1}{2} \cdot \frac{1}{8} \cdot b_1 b_2 b_3 \cdot N_{0,3}(b_1, b_1, b_2)$ $\cdot N_{0,3}(b_2, b_3, b_3) = \frac{8}{5} \cdot b_1 b_2 b_3 \cdot (1) \cdot (1)$ $\stackrel{\partial \Gamma}{\mapsto} \frac{8}{5} \left( 1 \cdot b_1 b_2 b_3 + \frac{1}{2} b_1 b_2 b_3 + 1 \cdot b_1 b_2 b_3 \right) \stackrel{Z}{\mapsto}$	$(1 + \frac{1}{2} + 1)$ $\cdot \frac{8}{5} \cdot (1! \zeta(2))^3$ $= \frac{5}{2} \cdot \frac{1}{135} \cdot \pi^6$
	$\frac{128}{5} \cdot \frac{1}{2} \cdot \frac{1}{12} \cdot b_1 b_2 b_3 \cdot N_{0,3}(b_1, b_1, b_2)$ $\cdot N_{0,3}(b_2, b_3, b_3) = \frac{16}{15} \cdot b_1 b_2 b_3 \cdot (1) \cdot (1)$ $\stackrel{\partial \Gamma}{\mapsto} \frac{8}{5} \left( 1 \cdot b_1 b_2 b_3 + \frac{1}{2} b_1 b_2 b_3 + 1 \cdot b_1 b_2 b_3 \right) \stackrel{Z}{\mapsto}$	$(1 + 1 + 1)$ $\cdot \frac{16}{15} \cdot (1! \zeta(2))^3$ $= 3 \cdot \frac{2}{405} \cdot \pi^6$

Vol  $\mathcal{Q}_2$  of square-tiled surfaces having single maximal cylinder. The resulting contribution  $(\frac{16}{945} + \frac{1}{2835})\pi^6 = \frac{7}{405}\pi^6 = \frac{49}{3}\zeta(6)$  was found in Appendix C in [23] by a completely different technique.

The third and fourth graphs together represent the volume contribution  $(\frac{8}{225} + \frac{1}{675})\pi^6 = \frac{1}{27}\pi^6$  of square-tiled surfaces having two maximal cylinders. The last two graphs represent the contribution  $(\frac{1}{135} + \frac{2}{405})\pi^6 = \frac{1}{81}\pi^6$  of square-tiled surfaces having three maximal cylinders.

Normalizing the contribution of 1, 2, 3-cylinder square-tiled surfaces by the entire volume Vol  $\mathcal{Q}_2$  of the stratum, we get the quantity  $p_k(\mathcal{Q}_2)$  which can be interpreted as the “probability” for a “random” square-tiled surface in the stratum  $\mathcal{Q}_2$  to have exactly  $k$  horizontal cylinders. These same quantities  $p_k$  provide “probabilities” of getting a  $k$ -band generalized interval exchange transformation (linear involution) taking a “random” generalized interval exchange transformation in the Rauzy class representing the stratum  $\mathcal{Q}_2$  (see Section 3.2 in [23] for details). The latter quantities are particularly simple to evaluate in numerical experiments. The resulting proportions

$$(p_1(\mathcal{Q}_2), p_2(\mathcal{Q}_2), p_3(\mathcal{Q}_2)) = \left(\frac{7}{405}, \frac{1}{27}, \frac{1}{81}\right) : \frac{1}{15} = \left(\frac{7}{27}, \frac{15}{27}, \frac{5}{27}\right)$$

match the numerical experiments obtained earlier in Appendix C of [23].

### C.3. Holomorphic quadratic differentials in genus 3

In genus 3 there are already 41 different decorated ribbon graphs. We do not provide the graph-by-graph calculation as we did in genus 2, only the contributions of  $k$ -cylinder square-tiled surfaces for all possible values  $k = 1, \dots, 6$  of cylinders (see Table 4).

The resulting contribution of 1-cylinder surfaces was confirmed by the alternative combinatorial study of the Rauzy class of the stratum  $\mathcal{Q}(1^8)$  (see Section 3.2 in [23]). The approximate values  $p_k(\mathcal{Q}(1^8))$  were confirmed by numerical experiments with statistics of  $k$ -band generalized interval exchange transformations.

Table 4. Computation of Vol  $\mathcal{Q}_3$ .

Number of cylinders $k$	Number of graphs $\Gamma$	Contribution to the volume	Relative contribution $p_k$
1	2	$\frac{94667}{126299250} \cdot \pi^{12}$	$\frac{757336}{3493125}$
2	5	$\frac{150749}{108256500} \cdot \pi^{12}$	$\frac{4220972}{10479375}$
3	9	$\frac{84481}{86605200} \cdot \pi^{12}$	$\frac{591367}{2095875}$
4	12	$\frac{5989}{21651300} \cdot \pi^{12}$	$\frac{167692}{2095875}$
5	8	$\frac{1}{17820} \cdot \pi^{12}$	$\frac{28}{1725}$
6	5	$\frac{1}{144342} \cdot \pi^{12}$	$\frac{56}{27945}$

Taking the sum of all contributions, we obtain the volume of the moduli space  $\mathcal{Q}_3$  of holomorphic quadratic differentials in genus 3:  $\text{Vol } \mathcal{Q}_3 = \frac{115}{33264} \cdot \pi^{12}$ .

*C.4. Meromorphic quadratic differentials in genus 1*

In this section we apply formula (1.13) to compute the Masur–Veech volume  $\text{Vol } \mathcal{Q}(1^2, -1^2)$  of the moduli space  $\mathcal{Q}_{1,2}$  of meromorphic quadratic differentials in genus  $g = 1$  with two simple poles  $p = 2$ .

We use the same convention on the order of numerical factors in every first line of the middle column as in Section C.2. Namely, for  $(g, p) = (1, 2)$ , we have  $\ell! \cdot 2d \cdot \frac{2^d}{d!} = \frac{32}{5}$ , which is the first factor. The second factor is  $1/2^{|\mathcal{V}(\Gamma)|-1}$ .

The third factor is  $|\text{Aut}(\Gamma)|^{-1}$ . We note that the vertices and edges of  $\Gamma$  are not labeled, while the two legs are. An automorphism of  $\Gamma$  preserves the decoration of vertices and the labeling of the legs. For example, the graph  $\Gamma$  in the second line does not have any nontrivial automorphisms.

Table 5. Computation of  $\text{Vol } \mathcal{Q}_{1,2}$ . The left column represents stable graphs  $\Gamma \in \mathcal{G}_{1,2}$  and associated multicurves; the middle column gives associated polynomials  $P_\Gamma$ ; the right column provides volume contributions  $\text{Vol}(\Gamma)$ .

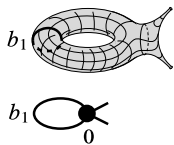
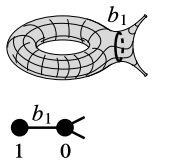
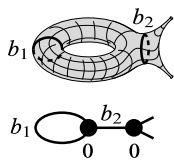
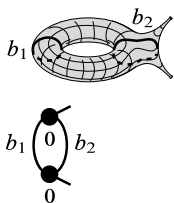
	$\frac{32}{3} \cdot 1 \cdot \frac{1}{2} \cdot b_1 \cdot N_{0,4}(b_1, b_1, 0, 0)$ $= \frac{16}{3} \cdot b_1 \cdot (\frac{1}{4}(2b_1^2)) = \frac{8}{3} \cdot b_1^3 \xrightarrow{\mathbb{Z}}$	$\frac{8}{3} \cdot 3! \cdot \zeta(4) = \frac{8}{45} \pi^4$
	$\frac{32}{3} \cdot \frac{1}{2} \cdot 1 \cdot b_1 \cdot N_{1,1}(b_1) \cdot N_{0,3}(0, 0, b_1)$ $= \frac{16}{3} \cdot b_1 \cdot (\frac{1}{48} b_1^2) \cdot (1) = \frac{1}{9} \cdot b_1^3 \xrightarrow{\mathbb{Z}}$	$\frac{1}{9} \cdot 3! \cdot \zeta(4) = \frac{1}{135} \cdot \pi^4$
	$\frac{32}{3} \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot b_1 b_2 \cdot N_{0,3}(b_1, b_1, b_2) \cdot N_{0,3}(b_1, 0, 0)$ $= \frac{8}{3} \cdot b_1 b_2 \cdot (1) \cdot (1) = \frac{8}{3} \cdot b_1 b_2 \xrightarrow{\mathbb{Z}}$	$\frac{8}{3} \cdot (\zeta(2))^2 = \frac{2}{27} \pi^4$
	$\frac{32}{3} \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot b_1 b_2 \cdot N_{0,3}(0, b_1, b_2) \cdot N_{0,3}(b_1, b_2, 0)$ $= \frac{8}{3} \cdot b_1 b_2 \cdot (1) \cdot (1) = \frac{8}{3} \cdot b_1 b_2 \xrightarrow{\mathbb{Z}}$	$\frac{8}{3} \cdot (\zeta(2))^2 = \frac{2}{27} \pi^4$

Table 6. Computation of  $c_{\text{area}}(\mathcal{Q}_{1,2})$ —Contributions to the sum in the RHS of equation (1.23).

	$\frac{32}{3} \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot b_1 b_2 \cdot N_{0,3}(b_1, b_1, b_2) \cdot N_{0,3}(b_1, 0, 0)$ $= \frac{8}{3} \cdot b_1 b_2 \cdot 1 \cdot 1 = \frac{8}{3} \cdot b_1 b_2 \xrightarrow{Z^{\circ\partial\Gamma}} (1 + \frac{1}{2}) \cdot \frac{8}{3} \cdot (\zeta(2))^2 = \frac{3}{2} \cdot \frac{2}{27} \pi^4$
	$\frac{32}{3} \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot b_1 b_2 \cdot N_{0,3}(0, b_1, b_2) \cdot N_{0,3}(b_1, b_2, 0)$ $= \frac{8}{3} \cdot b_1 b_2 \cdot 1 \cdot 1 = \frac{8}{3} \cdot b_1 b_2 \xrightarrow{Z^{\circ\partial\Gamma}} (1 + 1) \cdot \frac{8}{3} \cdot (\zeta(2))^2 = 2 \cdot \frac{2}{27} \pi^4$

The resulting value

$$\text{Vol } \mathcal{Q}_{1,2} = \left( \left( \frac{8}{45} + \frac{1}{135} \right) + \left( \frac{2}{27} + \frac{2}{27} \right) \right) \cdot \pi^4 = \left( \frac{5}{27} + \frac{4}{27} \right) \cdot \pi^4 = \frac{\pi^4}{3}$$

matches the one found in [36]. The contribution of 1-cylinder square-tiled surfaces and the 5:4 proportion between 1-cylinder and 2-cylinder contributions match the corresponding quantities found in Appendix C in [23].

We now evaluate the Siegel–Veech constant  $c_{\text{area}} \mathcal{Q}_{1,2}$  (see Table 6).

Taking the sum of the two contributions, we obtain

$$\left( \frac{3}{2} \cdot \frac{2}{27} + 2 \cdot \frac{2}{27} \right) \cdot \pi^4 = \frac{7}{27} \cdot \pi^4.$$

Dividing by  $\text{Vol } \mathcal{Q}(1^2, -1^2) = \frac{\pi^4}{3}$ , we obtain the answer which matches the value found in [35]:

$$\frac{\pi^2}{3} \cdot c_{\text{area}}(\mathcal{Q}(1^2, -1^2)) = \left( \frac{7}{27} \cdot \pi^4 \right) : \left( \frac{1}{3} \cdot \pi^4 \right) = \frac{7}{9}.$$

**Appendix D. Tables of volumes and Siegel–Veech constants**

In this appendix we present numerical data for  $\text{Vol}(\mathcal{Q}_{g,n})$  and  $c_{\text{area}}(\mathcal{Q}_{g,n})$  corresponding to small values of  $g$  and  $n$ . Table 7 gathers the numerical values of the volumes, of the Siegel–Veech constants, and of the sums  $\Lambda^+$  (resp.,  $\Lambda^-$ ) of the top  $g$  (resp.,  $g_{\text{eff}}$ ) Lyapunov exponents of the Kontsevich–Zorich cocycle over the principal stratum  $\mathcal{Q}(1^{4g-4+n}, -1^n)$ . By formulas (2.3) and (2.4) in the article by Eskin, Kontsevich, Zorich [28], these quantities are related as follows:

Table 7. Numerical values of volumes, Siegel–Veech constants, and sums of Lyapunov exponents for low-dimensional strata.

$g$	$n$	Stratum	Volume	$\pi^2/3 \cdot c_{\text{area}}$	$\Lambda^+$	$\Lambda^-$
0	5	$\mathcal{Q}(1, -1^5)$	$\pi^4$	5/9	0	4/3
0	6	$\mathcal{Q}(1^2, -1^6)$	$1/2 \cdot \pi^6$	11/18	0	5/3
0	7	$\mathcal{Q}(1^3, -1^7)$	$1/4 \cdot \pi^8$	2/3	0	2
1	2	$\mathcal{Q}(1^2, -1^2)$	$1/3 \cdot \pi^4$	7/9	2/3	4/3
1	3	$\mathcal{Q}(1^3, -1^3)$	$11/60 \cdot \pi^6$	47/66	6/11	17/11
1	4	$\mathcal{Q}(1^4, -1^4)$	$1/10 \cdot \pi^8$	44/63	10/21	38/21
1	5	$\mathcal{Q}(1^5, -1^5)$	$163/3024 \cdot \pi^{10}$	2075/2934	70/163	1025/489
2	0	$\mathcal{Q}(1^4)$	$1/15 \cdot \pi^6$	19/18	4/3	5/3
2	1	$\mathcal{Q}(1^5, -1)$	$29/840 \cdot \pi^8$	230/261	32/29	154/87
2	2	$\mathcal{Q}(1^6, -1^2)$	$337/18144 \cdot \pi^{10}$	8131/10110	1636/1685	3321/1685
3	0	$\mathcal{Q}(1^8)$	$115/33264 \cdot \pi^{12}$	24199/25875	4286/2875	18608/8625
4	0	$\mathcal{Q}(1^{12})$	$\pi^{18} \cdot 2106241/11548293120$	283794163/315936150	91179048/52656025	143835073/52656025

$$\Lambda^+ = \frac{(5g - 5 - n)}{18} + \frac{\pi^2}{3} c_{\text{area}}(\mathcal{Q}_{g,n}),$$

$$\Lambda^- = \Lambda^+ + \frac{(g - 1 + n)}{3}.$$

It is possible to approximate  $\Lambda^+$  and  $\Lambda^-$  numerically by computer simulations of the accelerated Rauzy induction. The values of  $\Lambda^+$  and  $\Lambda^-$  based on the values of  $c_{\text{area}}(\mathcal{Q}_{g,n})$  computed here match the approximate values obtained in these numerical experiments.

One has  $\Lambda^+ = 0$  in genus 0, so in genus 0 the Siegel–Veech constant admits a simple closed formula. By formula (1.1) in the paper of Athreya, Eskin, and Zorich [11], one has

$$\text{Vol } \mathcal{Q}(1^{n-4}, -1^n) = 2\pi^2 \left(\frac{\pi^2}{2}\right)^{n-4}.$$

We obtain the same expressions in the cells of Table 7 corresponding to genus 0 computed by formulas (1.13) and (1.23) of the current article.

Recent work of Chen, Möller, and Sauvaget [17] has suggested alternative formulas for the Masur–Veech volumes and the area Siegel–Veech constants of the principal strata  $\mathcal{Q}(1^{4g-4+n}, -1^n)$  as weighted sums of certain very special linear Hodge integrals. The subsequent paper of Kazarian [37] provided very efficient recursive formulas for these Hodge integrals, allowing one to compute  $\text{Vol } \mathcal{Q}_{g,n}$  and  $c_{\text{area}}(\mathcal{Q}_{g,n})$  for all sufficiently small values of  $g$  and  $n$  fast enough. In particular, following this alternative approach one obtains the same data as in Table 7.

Further numerical data can be found in [24], where we apply (1.13) and (1.23) to express, respectively, the Masur–Veech volumes  $\text{Vol}(\mathcal{Q}_{g,n})$  and the Siegel–Veech



constants  $c_{\text{area}}(\mathcal{Q}_{g,n})$  as polynomials in the intersection numbers of  $\psi$ -classes. Recall that applying formula (1.26) one can express the Siegel–Veech constant  $c_{\text{area}}$  in terms of the volumes of the principal boundary strata. One more table in [24] provides the corresponding explicit expressions for low-dimensional strata.

*Acknowledgments.* Numerous results in this paper were directly or indirectly inspired by the beautiful ideas of the late mathematician Maryam Mirzakhani. During our work, we could not avoid a constant feeling that we were retracing her path-breaking steps. This is particularly applicable to the relation between Masur–Veech volume  $\text{Vol } \mathcal{Q}_{g,n}$  and frequencies of hyperbolic multicurves, as well as to the relation between large genus asymptotics of the volumes of moduli spaces and intersection numbers of  $\psi$ -classes. Alex Eskin originally planned to use Kontsevich’s formula for the computation of Masur–Veech volumes before he and Okounkov invented their new approach in [31]; we thank him for useful conversations and for pointing out that our evaluation technique of the Masur–Veech volumes admits generalization to strata with only odd zeros (see Remark 1.13). We are grateful to Maxim Kazarian for his computer code for evaluating intersection numbers, which we used at an early stage of this project, and we are particularly grateful to him for the recursion for linear Hodge integrals obtained in [37]. A combination of recent results from [17] and [37] confirmed all numerical values of the Masur–Veech volumes  $\text{Vol } \mathcal{Q}_{g,n}$  and the Siegel–Veech constants  $c_{\text{area}}(\mathcal{Q}_{g,n})$ , which figure prominently in our work here and which provided very serious numerical evidence for Conjecture 1.9 (stated prior to these results). We highly appreciate the computer experiments of Mark Bell [13], which provided computer evidence—*independent of theoretical predictions*—of frequencies of multicurves on surfaces of low genera. We are extremely grateful to A. Aggarwal and J. Athreya who carefully read the draft manuscripts of this work and made numerous helpful remarks, as well as to F. Arana–Herrera, S. Barazer, A. Borodin, C. Ball, G. Borot, D. Chen, E. Duryev, M. Liu, M. Möller, L. Monin, B. Petri, K. Rafi, A. Sauvaget, J. Souto, I. Telpukhovskiy, S. Wolpert, A. Wright, and D. Zagier for useful discussions. We are grateful to the Max Planck Institute for Mathematics (MPIM) in Bonn, where a considerable part of this work was performed, and to the Mathematical Sciences Research Institute (MSRI) in Berkeley for providing a friendly, stimulating environment. Finally, we thank the anonymous referees for an attentive reading of the manuscript and for important suggestions which helped improve the presentation, and we thank Roy Pattishall for his very professional and responsible work which completed the transformation of the manuscript into the current form.

Goujard’s work was partially supported by a Projets Exploratoires Premier Soutien (PEPS) JCJC grant. The results of Section 4 were obtained at Saint Petersburg State University under the support of Russian Science Foundation grant 19-71-30002.

The work of Zograf was partially supported by the Government of Russian Federation megagrant 14.W03.31.0030. This material is based on work supported by National Science Foundation grant DMS-1440140 while Delecroix, Goujard, and Zorich were in residence at the MSRI during the Fall 2019 semester. It was also supported by Agence Nationale de la Recherche grant ANR-19-CE40-0021.

## References

- [1] A. AGGARWAL, *Large genus asymptotics for Siegel-Veech constants*, *Geom. Funct. Anal.* **29** (2019), no. 5, 1295–1324. [MR 4025514](#).  
[DOI 10.1007/s00039-019-00509-0](#). (2644)
- [2] ———, *Large genus asymptotics for volumes of strata of Abelian differentials*, with an appendix “Asymptotic values of Siegel-Veech constants” by A. Zorich, *J. Amer. Math. Soc.* **33** (2020), no. 4, 941–989. [MR 4155217](#).  
[DOI 10.1090/jams/947](#). (2644)
- [3] ———, *Large genus asymptotics for intersection numbers and principal strata volumes of quadratic differentials*, to appear in *Invent. Math.*, preprint, [arXiv:2004.05042v1](#) [math.AG]. (2644)
- [4] A. AGGARWAL, V. DELECROIX, É. GOUJARD, P. ZOGRAF, and A. ZORICH, *Conjectural large genus asymptotics of Masur-Veech volumes and of area Siegel-Veech constants of strata of quadratic differentials*, *Arnold Math. J.* **6** (2020), no. 2, 149–161. [MR 4134482](#). [DOI 10.1007/s40598-020-00139-7](#). (2644)
- [5] J. E. ANDERSEN, G. BOROT, S. CHARBONNIERY, V. DELECROIX, A. GIACCHETTO, D. LEWAŃSKI, and C. WHEELER, *Topological recursion for Masur-Veech volumes*, preprint, [arXiv:1905.10352v3](#) [math.GT]. (2645)
- [6] 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. [MR 4097720](#).  
[DOI 10.3934/jmd.2020004](#). (2654, 2694)
- [7] ———, *Equidistribution of families of expanding horospheres on moduli spaces of hyperbolic surfaces*, *Geom. Dedicata* **210** (2021), 65–102. [MR 4200910](#).  
[DOI 10.1007/s10711-020-00534-6](#). (2660)
- [8] ———, *Counting hyperbolic multi-geodesics with respect to the lengths of individual components*, preprint, [arXiv:2002.10906v1](#) [math.DS]. (2660)
- [9] E. ARBARELLO and M. CORNALBA, *Combinatorial and algebro-geometric cohomology classes on the moduli spaces of curves*, *J. Algebraic Geom.* **5** (1996), no. 4, 705–749. [MR 1486986](#). (2646)
- [10] J. S. ATHREYA, A. ESKIN, and A. ZORICH, *Counting generalized Jenkins-Strebel differentials*, *Geom. Dedicata* **170** (2014), no. 1, 195–217. [MR 3199485](#).  
[DOI 10.1007/s10711-013-9877-7](#). (2643, 2663, 2665, 2666, 2668, 2673)
- [11] ———, *Right-angled billiards and volumes of moduli spaces of quadratic differentials on  $\mathbb{C}P^1$* , with an appendix “Equidistribution of circle translates” by J. Chaika,

- Ann. Sci. Éc. Norm. Supér. (4) **49** (2016), no. 6, 1311–1386. MR 3592359.  
DOI 10.24033/asens.2310. (2640, 2643, 2654, 2663, 2712)
- [12] N. BATIR, *Inequalities for the gamma function*, Arch. Math. (Basel) **91** (2008), no. 6, 554–563. MR 2465874. DOI 10.1007/s00013-008-2856-9. (2688)
- [13] M. BELL, *Experimental statistics for Mirzakhani’s theorem*, to appear in Exp. Math., preprint, arXiv:1910.08155v1 [math.GT]. (2653, 2694, 2713)
- [14] C. BONATTI, A. ESKIN, and A. WILKINSON, *Projective cocycles over  $SL(2, \mathbb{R})$  actions: Measures invariant under the upper triangular group*, Astérisque **415** (2020), 157–180. MR 4142450. DOI 10.24033/ast. (2672)
- [15] K. M. CHAPMAN, M. MULASE, and B. SAFNUK, *The Kontsevich constants for the volume of the moduli of curves and topological recursion*, Commun. Number Theory Phys. **5** (2011), no. 3, 643–698. MR 2864463.  
DOI 10.4310/CNTP.2011.v5.n3.a3. (2639)
- [16] D. CHEN, *Square-tiled surfaces and rigid curves on moduli spaces*, Adv. Math. **228** (2011), no. 2, 1135–1162. MR 2822219. DOI 10.1016/j.aim.2011.06.002. (2672)
- [17] D. CHEN, M. MÖLLER, and A. SAUVAGET, *Masur-Veech volumes and intersection theory: The principal strata of quadratic differentials*, preprint, arXiv:1912.02267v1 [math.AG]. (2643, 2644, 2645, 2712, 2713)
- [18] D. CHEN, M. MÖLLER, A. SAUVAGET, and D. ZAGIER, *Masur-Veech volumes and intersection theory on moduli spaces of abelian differentials*, Invent. Math. **222** (2020), no. 1, 283–373. MR 4145791. DOI 10.1007/s00222-020-00969-4. (2644)
- [19] D. CHEN, M. MÖLLER, and D. ZAGIER, *Quasimodularity and large genus limits of Siegel-Veech constants*, J. Amer. Math. Soc. **31** (2018), no. 4, 1059–1163. MR 3836563. DOI 10.1090/jams/900. (2644)
- [20] V. DELECROIX, É. GOUJARD, P. ZOGRAF, and A. ZORICH, *Contribution of one-cylinder square-tiled surfaces to Masur-Veech volumes*, with an appendix “Square-tiled surfaces with one horizontal cylinder” by P. Engel, Astérisque **415** (2020), 223–274. MR 4142454. DOI 10.24033/ast. (2646, 2663)
- [21] ———, *Enumeration of meanders and Masur-Veech volumes*, Forum Math. Pi **8** (2020), no. e4. MR 4079755. DOI 10.1017/fmp.2020.2. (2637, 2663)
- [22] ———, *Large genus asymptotic geometry of random square-tiled surfaces and of random multicurves*, preprint, arXiv:2007.04740v1 [math.GT]. (2644, 2652)
- [23] ———, *Square-tiled surfaces of fixed combinatorial type: Equidistribution, counting, volumes of the ambient strata*, preprint, arXiv:1612.08374v2 [math.GT]. (2709, 2711)
- [24] ———, *Tables of Masur-Veech volumes and of area Siegel-Veech constants of principal strata of quadratic differentials*, preprint, 2020, <http://www.math.u-bordeaux.fr/~egoujard/tablevol2020.pdf>. (2650, 2712, 2713)
- [25] P. ENGEL, *Hurwitz theory of elliptic orbifolds, I*, Geom. Topol. **25** (2021), no. 1, 229–274. MR 4226230. DOI 10.2140/gt.2021.25.229. (2645)
- [26] ———, *Hurwitz theory of elliptic orbifolds, II*, preprint, arXiv:1809.07434v1 [math.AG]. (2645)

- [27] V. ERLANDSSON and J. SOUTO, *Geodesic currents and Mirzakhani's curve counting*, in preparation. (2650, 2654, 2655, 2694)
- [28] A. ESKIN, M. KONTSEVICH, and A. ZORICH, *Sum of Lyapunov exponents of the Hodge bundle with respect to the Teichmüller geodesic flow*, Publ. Math. Inst. Hautes Études Sci. **120** (2014), no. 1, 207–333. MR 3270590. DOI 10.1007/s10240-013-0060-3. (2647, 2671, 2672, 2711)
- [29] A. ESKIN and H. MASUR, *Asymptotic formulas on flat surfaces*, Ergodic Theory Dynam. Systems **21** (2001), no. 2, 443–478. MR 1827113. DOI 10.1017/S0143385701001225. (2647)
- [30] A. ESKIN, H. MASUR, and A. ZORICH, *Moduli spaces of abelian differentials: The principal boundary, counting problems, and the Siegel-Veech constants*, Publ. Math. Inst. Hautes Études Sci. **97** (2003), no. 1, 61–179. MR 2010740. DOI 10.1007/s10240-003-0015-1. (2648)
- [31] A. ESKIN and A. OKOUNKOV, *Asymptotics of numbers of branched coverings of a torus and volumes of moduli spaces of holomorphic differentials*, Invent. Math. **145** (2001), no. 1, 59–103. MR 1839286. DOI 10.1007/s002220100142. (2643, 2645, 2713)
- [32] ———, “Pillowcases and quasimodular forms” in *Algebraic Geometry and Number Theory*, Progr. Math. **253**, Birkhäuser Boston, Boston, 2006, 1–25. MR 2263191. DOI 10.1007/978-0-8176-4532-8\_1. (2643, 2645)
- [33] A. ESKIN, A. OKOUNKOV, and R. PANDHARIPANDE, *The theta characteristic of a branched covering*, Adv. Math. **217** (2008), no. 3, 873–888. MR 2383889. DOI 10.1016/j.aim.2006.08.001. (2643)
- [34] A. ESKIN and A. ZORICH, *Volumes of strata of Abelian differentials and Siegel-Veech constants in large genera*, Arnold Math. J. **1** (2015), no. 4, 481–488. MR 3434506. DOI 10.1007/s40598-015-0028-0. (2644)
- [35] E. GOUJARD, *Siegel-Veech constants for strata of moduli spaces of quadratic differentials*, Geom. Funct. Anal. **25** (2015), no. 5, 1440–1492. MR 3426059. DOI 10.1007/s00039-015-0345-4. (2648, 2708, 2711)
- [36] ———, *Volumes of strata of moduli spaces of quadratic differentials: Getting explicit values*, Ann. Inst. Fourier (Grenoble) **66** (2016), no. 6, 2203–2251. MR 3580171. (2643, 2645, 2663, 2711)
- [37] M. KAZARIAN, *Recursion for Masur-Veech volumes of moduli spaces of quadratic differentials*, to appear in J. Inst. Math. Jussieu, preprint, arXiv:1912.10422v1 [math.ph]. (2643, 2644, 2681, 2712, 2713)
- [38] S. P. KERCKHOFF, *Earthquakes are analytic*, Comment. Math. Helv. **60** (1985), no. 1, 17–30. MR 0787659. DOI 10.1007/BF02567397. (2650)
- [39] M. KONTSEVICH, *Intersection theory on the moduli space of curves and the matrix Airy function*, Comm. Math. Phys. **147** (1992), no. 1, 1–23. MR 1171758. DOI 10.1007/BF02099526. (2639, 2705)
- [40] V. KOZIARZ and D.-M. NGUYEN, *Variation of Hodge structure and enumerating tilings of surfaces by triangles and squares*, J. Éc. polytech. Math. **8** (2021), 831–857. MR 4237020. DOI 10.5802/jep.159. (2645)

- [41] M. LIU, *Length statistics of random multicurves on closed hyperbolic surfaces*, preprint, [arXiv:1912.11155v2](https://arxiv.org/abs/1912.11155v2) [math.GT]. (2660)
- [42] H. MASUR, *Interval exchange transformations and measured foliations*, *Ann. of Math.* (2) **115** (1982), no. 1, 169–200. [MR 0644018](#). [DOI 10.2307/1971341](#). (2634, 2663)
- [43] ———, *The Teichmüller flow is Hamiltonian*, *Proc. Amer. Math. Soc.* **123** (1995), no. 12, 3739–3747. [MR 1283559](#). [DOI 10.2307/2161902](#). (2663)
- [44] M. MIRZAKHANI, *Simple geodesics and Weil-Petersson volumes of moduli spaces of bordered Riemann surfaces*, *Invent. Math.* **167** (2007), no. 1, 179–222. [MR 2264808](#). [DOI 10.1007/s00222-006-0013-2](#). (2645, 2680)
- [45] ———, *Weil-Petersson volumes and intersection theory on the moduli space of curves*, *J. Amer. Math. Soc.* **20** (2007), no. 1, 1–23. [MR 2257394](#). [DOI 10.1090/S0894-0347-06-00526-1](#). (2680, 2681)
- [46] ———, *Ergodic theory of the earthquake flow*, *Int. Math. Res. Not. IMRN* **2008**, no. 3, art. ID. 116. [MR 2416997](#). [DOI 10.1093/imrn/rnm116](#). (2654, 2677, 2678)
- [47] ———, *Growth of the number of simple closed geodesics on hyperbolic surfaces*, *Ann. of Math.* (2) **168** (2008), no. 1, 97–125. [MR 2415399](#). [DOI 10.4007/annals.2008.168.97](#). (2651, 2653, 2654, 2655, 2678, 2679, 2680, 2681, 2682, 2692, 2693, 2694)
- [48] ———, *Counting mapping class group orbits on hyperbolic surfaces*, preprint, [arXiv:1601.03342v1](https://arxiv.org/abs/1601.03342v1) [math.GT]. (2656, 2660)
- [49] ———, *Simple geodesics on hyperbolic surfaces and the volume of the moduli space of curves*, Ph.D. dissertation, Harvard University, Cambridge, 2004. [MR 2705986](#). (2655)
- [50] M. MIRZAKHANI and P. ZOGRAF, *Towards large genus asymptotics of intersection numbers on moduli spaces of curves*, *Geom. Funct. Anal.* **25** (2015), no. 4, 1258–1289. [MR 3385633](#). [DOI 10.1007/s00039-015-0336-5](#). (2645)
- [51] L. MONIN and V. TELPUKHOVSKIY, *On normalizations of Thurston measure on the space of measured laminations*, *Topology Appl.* **267** (2019), art. ID 106878. [MR 4008568](#). [DOI 10.1016/j.topol.2019.106878](#). (2650, 2654, 2655)
- [52] P. NORBURY, *Counting lattice points in the moduli space of curves*, *Math. Res. Lett.* **17** (2010), no. 3, 467–481. [MR 2653682](#). [DOI 10.4310/MRL.2010.v17.n3.a7](#). (2639, 2645)
- [53] A. OKOUNKOV and R. PANDHARIPANDE, “Gromov-Witten theory, Hurwitz numbers, and matrix models” in *Algebraic Geometry (Seattle, 2005) Part 1*, *Proc. Sympos. Pure Math.* **80**, Amer. Math. Soc., Providence, 2009, 325–414. [MR 2483941](#). [DOI 10.1090/pspum/080.1/2483941](#). (2664)
- [54] R. VAKIL, *The moduli space of curves and its tautological ring*, *Notices Amer. Math. Soc.* **50** (2003), no. 6, 647–658. [MR 1988577](#). (2707)
- [55] W. A. VEECH, *Gauss measures for transformations on the space of interval exchange maps*, *Ann. of Math.* (2) **115** (1982), no. 1, 201–242. [MR 0644019](#). [DOI 10.2307/1971391](#). (2634, 2663)
- [56] ———, *Siegel measures*, *Ann. of Math.* (2) **148** (1998), no. 3, 895–944. [MR 1670061](#). [DOI 10.2307/121033](#). (2647)

- [57] Y. VOROBETS, “Periodic geodesics on generic translation surfaces” in *Algebraic and Topological Dynamics*, Contemp. Math. **385**, Amer. Math. Soc., Providence, 2005, 205–258. MR 2180238. DOI 10.1090/conm/385/07199. (2647)
- [58] E. WITTEN, “Two-dimensional gravity and intersection theory on moduli space” in *Surveys in Differential Geometry (Cambridge, 1990)*, Lehigh Univ., Bethlehem, 1991, 243–310. MR 1144529. (2639, 2668, 2684, 2689)
- [59] A. WRIGHT, *A tour through Mirzakhani’s work on moduli spaces of Riemann surfaces*, Bull. Amer. Math. Soc. (N.S.) **57** (2020), no. 3, 359–408. MR 4108090. DOI 10.1090/bull/1687. (2654, 2680, 2694)
- [60] D. YANG, D. ZAGIER, and Y. ZHANG, *Masur-Veech volumes of quadratic differentials and their asymptotics*, J. Geom. Phys. **158** (2020), art. ID 103870. MR 4150385. DOI 10.1016/j.geomphys.2020.103870. (2644)
- [61] P. ZOGRAF, *An explicit formula for Witten’s 2-correlators*, J. Math. Sci. **240** (2019), no. 5, 535–538. DOI 10.1007/s10958-019-04371-1. (2652, 2684, 2685)
- [62] A. ZORICH, “Square tiled surfaces and Teichmüller volumes of the moduli spaces of abelian differentials” in *Rigidity in Dynamics and Geometry (Cambridge, 2000)*, Springer, Berlin, 2002, 459–471. MR 1919417. (2635)

*Delecroix*

Laboratoire Bordelais de Recherche en Informatique, Université de Bordeaux, Talence, France;  
[vincent.delecroix@u-bordeaux.fr](mailto:vincent.delecroix@u-bordeaux.fr)

*Goujard*

Institut de Mathématiques de Bordeaux, Université de Bordeaux, Talence, France;  
[elise.goujard@gmail.com](mailto:elise.goujard@gmail.com)

*Zograf*

St. Petersburg Department of Steklov Mathematical Institute, St. Petersburg, Russia and  
 Chebyshev Laboratory, St. Petersburg State University, St. Petersburg, Russia;  
[zograf@pdmi.ras.ru](mailto:zograf@pdmi.ras.ru)

*Zorich*

Skoltech Center for Advanced Studies, Moscow, Russia and Institut de Mathématiques de  
 Jussieu–Paris Rive Gauche, Paris, France; [anton.zorich@gmail.com](mailto:anton.zorich@gmail.com)