VOLUMES AND RANDOM MATRICES

by EDWARD WITTEN

(*Institute for Advanced Study, Einstein Drive, Princeton, NJ 08540 USA*)

Abstract

This article is an introduction to newly discovered relations between volumes of moduli spaces of Riemann surfaces or super Riemann surfaces, simple models of gravity or supergravity in two dimensions, and random matrix ensembles. (The article is based on a lecture at the conference on the Mathematics of Gauge Theory and String Theory, University of Auckland, January 2020)

1. Introduction

In this article, I will sketch some recent developments involving the volume of the moduli space $\mathcal{M}_g$ of Riemann surfaces of genus g, and also the volume of the corresponding moduli space $\mathfrak{M}_g$ of super Riemann surfaces. The discussion also encompasses Riemann surfaces with punctures and/or boundaries. The main goal is to explain how these volumes are related to random matrix ensembles. This is an old story with a very contemporary twist.

What is meant by the volume of $\mathcal{M}_g$? One answer is that $\mathcal{M}_g$ for g > 1 has a natural Weil-Petersson symplectic form. A Riemann surface Σ of genus g > 1 can be regarded as the quotient of the upper half plane $\mathcal{H} \cong \mathrm{PSL}(2,\mathbb{R})/\mathrm{U}(1)$ by a discrete group. Accordingly Σ carries a natural Riemannian metric g pulled back from $\mathcal{H}$; this metric has constant scalar curvature $R = -2$ and is called a hyperbolic metric. Similarly, Σ is endowed with a natural flat $\mathrm{PSL}(2,\mathbb{R})$ bundle[1] with connection A pulled back from $\mathcal{H}$. A point in $\mathcal{M}_g$ determines a Riemann surface together with such a flat connection, and the symplectic form ω of $\mathcal{M}_g$ can be defined by

$$\omega = \frac{1}{4\pi} \int_\Sigma \mathrm{Tr}\, \delta A \wedge \delta A, \tag{1}$$

in close analogy with the definition used by Atiyah and Bott [**1**] for a symplectic form on the moduli space of flat bundles with compact structure group G. The volume of $\mathcal{M}_g$ is then

$$V_g = \int_{\mathcal{M}_g} \mathrm{Pf}(\omega) = \int_{\mathcal{M}_g} e^\omega, \tag{2}$$

where Pf is the Pfaffian.

This approach generalizes perfectly well for super Riemann surfaces. One replaces $\mathrm{SL}(2,\mathbb{R})$, which is the group of linear transformations of $\mathbb{R}^2$ that preserve the symplectic form $\mathrm{d}u\,\mathrm{d}v$, with $\mathrm{OSp}(1|2)$, which is the supergroup of linear transformations of $\mathbb{R}^{2|1}$ that preserve the symplectic form $\mathrm{d}u\,\mathrm{d}v - \mathrm{d}\theta^2$. $\mathrm{OSp}(1|2)$ is a Lie supergroup of dimension 3|2. Its Lie algebra carries a nondegenerate bilinear form that I will denote as Tr. A point in $\mathfrak{M}_g$, the moduli space of super Riemann surfaces

---

Footnotes can be found in the appendix.
Email: witten@ias.edu

of genus g, determines an ordinary Riemann surface Σ together with a flat $\mathrm{OSp}(1|2)$ connection A. The symplectic form of $\mathfrak{M}_g$ can be defined by the same formula as before, only for flat $\mathrm{OSp}(1|2)$ connections rather than flat $\mathrm{PSL}(2,\mathbb{R})$ connections:

$$\widehat{\omega} = \frac{1}{4\pi} \int_\Sigma \mathrm{Tr}\, \delta A \wedge \delta A. \tag{3}$$

The volume $\widehat{V}_g$ of $\mathfrak{M}_g$ can be defined

$$\widehat{V}_g = \int_{\mathfrak{M}_g} \sqrt{\mathrm{Ber}\,\widehat{\omega}}, \tag{4}$$

where Ber is the Berezinian, the superanalog of the determinant. This is analogous to the first of the two formulas in Equation (2). There is no good superanalog of the integral of e^ω, but the formula $\mathrm{Pf}(\omega)$ for a measure does have a good superanalog, namely $\sqrt{\mathrm{Ber}(\widehat{\omega})}$.

I will now make a slight digression to explain how an ordinary Riemann surface Σ with a flat $\mathrm{OSp}(1|2)$ connection of the appropriate topological type (see the explanation of footnote 1 in the appendix) determines a super Riemann surface $\widehat{\Sigma}$. The superanalog of the upper half plane $\mathcal{H}$ is $\widehat{\mathcal{H}} = \mathrm{OSp}(1|2)/\mathrm{U}(1)$. $\widehat{\mathcal{H}}$ is a smooth supermanifold of real dimension $2|2$; it carries a complex structure in which it has complex dimension $1|1$. $\widehat{\mathcal{H}}$ also carries a canonical 'completely unintegrable distribution' making it a super Riemann surface. (There is no natural splitting of the Lie superalgebra $\mathrm{osp}(1|2)$ as the direct sum of even and odd parts, but the choice of a point in $\widehat{\mathcal{H}}$ determines such a splitting, and the odd part defines a subbundle of the tangent bundle to $\widehat{\mathcal{H}}$. This is the unintegrable distribution.) By taking the monodromies of the flat connection $A \to \Sigma$, we get a homomorphism $\rho : \pi_1(\Sigma) \to \mathrm{OSp}(1|2)$. Let $\Gamma = \rho(\pi_1(\Sigma))$. The quotient $\widehat{\Sigma} = \widehat{\mathcal{H}}/\Gamma$ is a smooth supermanifold of dimension $2|2$ that inherits from $\widehat{\mathcal{H}}$ the structure of a super Riemann surface. $\widehat{\Sigma}$ is the super Riemann surface associated to the pair Σ, A. For our purposes here, however, it is convenient to study the pair Σ, A rather than the super Riemann surface $\widehat{\Sigma}$.

It is possible to describe the super volumes $\widehat{V}_g$ in purely bosonic terms, that is in terms of ordinary geometry. The 'reduced space' of $\mathfrak{M}_g$ is the moduli space $\mathcal{M}'_g$ that parametrizes an ordinary Riemann surface Σ with a spin structure, which we can think of as a square root $K^{1/2}$ of the canonical bundle $K \to \Sigma$. $\mathcal{M}'_g$ is a finite cover of $\mathcal{M}_g$. The normal bundle to $\mathcal{M}'_g$ in $\mathfrak{M}_g$ is the vector bundle $U \to \mathcal{M}'_g$ whose fiber is $H^1(\Sigma, K^{-1/2})$. Viewing U as a real vector bundle (of twice its complex dimension), we denote its Euler class as $\chi(U)$. The symplectic form $\widehat{\omega}$ of $\mathfrak{M}_g$ restricts along $\mathcal{M}'_g$ to the ordinary symplectic form ω of $\mathcal{M}'_g$. By general arguments about symplectic supermanifolds, one can show that

$$\widehat{V}_g = \int_{\mathcal{M}'_g} \chi(U) e^\omega. \tag{5}$$

Thus what I will say about supervolumes can be interpreted as a purely classical statement about $\mathcal{M}'_g$. The quantities on the right-hand side of Equation (5) were studied by Norbury [2] from a different point of view (related however to the same spectral curve that appears in our discussion in Section 5).

The bosonic volumes V_g can also be expressed in terms of the intersection theory of $\mathcal{M}_g$. Indeed, the Weil-Petersson form is one of the tautological classes on $\mathcal{M}_g$ that were introduced by Mumford, Morita and Miller. Maryam Mirzakhani described [3] a sort of converse of this statement: from a

knowledge of the volumes (for Riemann surfaces possibly with geodesic boundary, as discussed momentarily) one can deduce the intersection numbers of tautological classes. These facts do not generalize directly to $\mathfrak{M}_g$, as there is not a natural intersection theory on a supermanifold.

Volumes for surfaces with boundary are introduced as follows. Let Σ be a hyperbolic Riemann surface of genus g with n boundaries. We require the boundaries to be geodesics of prescribed lengths $b_1, b_2, \ldots, b_n$. Let $\mathcal{M}_{g,\vec{b}}$ be the moduli space of such objects. It has a symplectic form and volume that can be defined by precisely the same formulas (1) and (2) as before. Mirzakhani showed that $V_{g,\vec{b}}$ is a polynomial in $b_1, b_2, \ldots, b_n$, and that the canonical intersection numbers are the coefficients of the top degree terms in this polynomial.

One can similarly define a moduli space $\mathfrak{M}_{g,\vec{b}}$ of super Riemann surfaces with geodesic boundaries of specified lengths $b_1, b_2, \ldots, b_n$, and its volume $\widehat{V}_{g,\vec{b}}$.

The relation of volumes to intersection numbers gives one way to compute them [4, 5]. Mirzakhani, however, discovered a new direct way to compute the volumes [6]. Regardless of how volumes are to be computed, the relation between volumes and intersection numbers shows that volumes are related to random matrix ensembles, since intersection numbers are known to be related to random matrix ensembles in multiple ways. Indeed, my conjecture [7] on intersection theory on $\mathcal{M}_g$ was inspired directly by discoveries made in that period [8–10] relating two-dimensional gravity to random matrix ensembles. Moreover, Kontsevich's proof [11] involved a relation of the intersection numbers to a different random matrix ensemble. However, at least in my opinion, the role of the random matrices in all of these constructions was somewhat obscure. The new developments that I will describe give a much clearer picture, since the random matrix will have a simple physical meaning.

In this article, I will sketch how Saad, Shenker and Stanford [12], following Eynard and Orantin [13], reinterpreted Mirzakhani's results in terms of a random matrix ensemble. They were motivated by considerations of quantum gravity, and I will try to give at least some idea about their motivation and why their results are of physical interest. Then I will explain how Stanford and I [14] generalized the story to super Riemann surfaces and quantum supergravity.

2. Universal Teichmüller space

Let S^1 be a circle and diffS^1 its group of orientation-preserving diffeomorphisms. The homogeneous spaces diff$S^1/\mathrm{PSL}(2, \mathbb{R})$ or diff$S^1/\mathrm{U}(1)$ can be viewed as coadjoint orbits of a central extension of diffS^1, so they carry natural symplectic structures. Let $\mathcal{X}$ be one of these spaces, and denote its symplectic structure by ω. We will see that $\mathcal{X}$ with its symplectic structure is a sort of infinite-dimensional analog of $\mathcal{M}_g$. Actually, diff$S^1/\mathrm{PSL}(2, \mathbb{R})$ is sometimes called universal Teichmüller space, and the developments that will be reviewed here perhaps give a perspective on the sense in which this name is justified.

It is believed that there is no reasonable definition of the 'volume' $\int_{\mathcal{X}} e^\omega$. However, we can do the following. Consider a subgroup $\mathrm{U}(1) \cong S^1 \subset$ diffS^1, consisting of rigid rotations of S^1. In other words, for some parametrization of S^1 by an angle θ, $\mathrm{U}(1)$ acts by $\theta \to \theta +$ constant. Then there is a moment map K for this action of $\mathrm{U}(1)$; in other words, if V is the vector field on $\mathcal{X}$ that generates $\mathrm{U}(1)$ and i_V is contraction with V, then

$$\mathrm{d}\mathsf{K} = -\mathsf{i}_\mathsf{V}\omega. \tag{6}$$

Then introducing a real constant β, the integral

$$Z(\beta) = \int_{\mathcal{X}} \exp(\mathsf{K}/\beta + \omega) \tag{7}$$

is better-behaved.

In fact, this integral can be viewed as an infinite-dimensional example of a situation that was studied by Duistermaat and Heckman [18], and then reinterpreted by Atiyah and Bott [19] in terms of equivariant cohomology. Let $\mathcal{Y}$ be a symplectic manifold (compact, or with some suitable conditions at infinity) with symplectic form ω and action of $U(1)$. Let $p_1, \ldots, p_s$ be the fixed points of the $U(1)$ action. For simplicity I assume that there are finitely many. Let K be the moment map for the $U(1)$ action. The Duistermaat–Heckman/Atiyah–Bott (DH/AB) formula gives

$$\int_{\mathcal{Y}} \exp(\mathsf{K}/\beta + \omega) = \sum_i \frac{\exp(\mathsf{K}(p_i)/\beta)}{\prod_\alpha (e_{i,\alpha}/2\pi\beta)},$$

where the $e_{i,\alpha}$ are integers that represent the eigenvalues of the $U(1)$ action on the tangent space to $\mathcal{Y}$ at p_i.

In the present example, there is only one fixed point in the $U(1)$ action on $\mathrm{diff}S^1/\mathrm{PSL}(2,\mathbb{R})$ or $\mathrm{diff}S^1/U(1)$. The product over eigenvalues at this fixed point becomes formally $\prod_{n=2}^\infty n/2\pi\beta$ in the example of $\mathrm{diff}S^1/\mathrm{PSL}(2,\mathbb{R})$ (or $\prod_{n=1}^\infty n/2\pi\beta$ in the other example). This infinite product is treated with (for example) ζ-function regularization. For $\mathrm{diff}S^1/\mathrm{PSL}(2,\mathbb{R})$, the result is

$$Z(\beta) = \frac{C}{4\pi^{1/2}\beta^{3/2}} \exp(\pi^2/\beta), \tag{8}$$

where the constant C, which has been normalized for later convenience, depends on the regularization and so is considered inessential, but the rest is 'universal'. (This problem was first studied by Kitaev [15] and subsequently analyzed in [16]. There are many derivations of the formula (8) in the physics literature. The explanation that I have sketched is from [17].) There is a similar formula for the other example $\mathrm{diff}S^1/U(1)$.

To use the DH/AB formula, we did not need to know what is the moment map K (only its value at the fixed point). But in fact it is a function of interest. To pick the $U(1)$ subgroup of $\mathrm{diff}\, S^1$ that was used in this 'localization', we had to pick an angular parameter θ on the circle; an element of $\mathrm{diff}\, S^1$ maps this to another parameter u, and K is the integral of the Schwarzian derivative $\{u, \theta\}$.

By an inverse Laplace transform, we can write

$$Z(\beta) = \int_0^\infty \mathrm{d}E \rho(E) \exp(-\beta E), \tag{9}$$

with

$$\rho(E) = \frac{C}{4\pi^2} \sinh(2\pi\sqrt{E}). \tag{10}$$

3. Quantum gravity in two dimensions

I will next explain why this result was considered problematical and how it has been interpreted [12]. This will require explaining some physics.

General Relativity is difficult to understand as a quantum theory. Searching for understanding, physicists have looked for a simpler model in a lower dimension. Two dimensions is a good place to look. An obvious idea might be to start with the Einstein–Hilbert action in two dimensions, $I_{EH} = \int_\Sigma d^2x\sqrt{g}R/2\pi$, with R the Ricci scalar of a Riemannian metric g. This does not work well, as in two dimensions this action is a topological invariant, the Euler characteristic χ, according to the Gauss–Bonnet theorem. Instead it turns out to be better to add a scalar (real-valued) field ϕ. For many purposes, a simple and illuminating model of two-dimensional gravity is 'Jackiw–Teitelboim (JT) gravity', with action

$$I_{\text{JT}} = -\frac{1}{2}\int_\Sigma d^2x\sqrt{g}\,\phi(R+2). \tag{11}$$

Actually, even though the Einstein–Hilbert action is a topological invariant, it turns out that it is important to include a multiple of this term in the action. The combined action is then

$$I = I_{\text{JT}} - S_0 I_{\text{EH}} = -\frac{1}{2}\int_\Sigma d^2x\sqrt{g}\,\phi(R+2) - S_0\int d^2s\sqrt{g}\frac{R}{2\pi}. \tag{12}$$

In conventional General Relativity, the coefficient of the Einstein–Hilbert term is $1/16\pi G_N$, where G_N is Newton's constant. In the real world, G_N is extremely small in natural units set by the values of Planck's constant, the speed of light, and atomic masses. So $1/16\pi G_N$ is very large in natural units. Similarly, here we are going to think of S_0 as being large. The same goes for the renormalized parameter S introduced below.

Because I_{EH} is a topological invariant, the Euler–Lagrange equations come entirely from I_{JT}. The Euler–Lagrange equation for ϕ is simply $R+2=0$, so a classical solution is a hyperbolic Riemann surface. The Feynman path integral for compact Σ without boundary (or with geodesic boundary of prescribed length) is very simple. The path integral is

$$Z_\Sigma = \frac{1}{\text{vol}}\int \mathcal{D}\phi\,\mathcal{D}g\,e^{-I}, \tag{13}$$

where formally vol is the volume of the diffeomorphism group and the factor $1/\text{vol}$ is a way to indicate that we have to divide by the diffeomorphism group (that is, we integrate over pairs ϕ, g up to diffeomorphism). Since $I = I_{\text{JT}} - S_0 I_{\text{EH}} = I_{\text{JT}} - S_0\chi$, we have in more detail

$$Z_\Sigma = e^{S_0\chi} \cdot \frac{1}{\text{vol}}\int \mathcal{D}\phi\,\mathcal{D}g\exp\left(\frac{1}{2}\int d^2x\sqrt{g}\phi(R+2)\right). \tag{14}$$

The integral over ϕ and g is studied by integrating first over ϕ (after rotating the integration contour $\phi \to i\phi$) and gives a delta function setting $R+2=0$. Since we have to divide by the diffeomorphism group, the integral 'localizes' on the moduli space of two-manifolds with hyperbolic structure, modulo diffeomorphism.

If Σ is orientable and of genus g, the moduli space of two-manifolds with metric of constant curvature $R=-2$ is the usual moduli space $\mathcal{M}_g$ of Riemann surfaces of genus g, and one can show

that the integral over $\mathcal{M}_{\mathrm{g}}$ gives its usual volume, up to an elementary factor:

$$\frac{1}{\mathrm{vol}} \int \mathcal{D}\phi \, \mathcal{D}g \exp\left(\frac{1}{2}\int \mathrm{d}^2x\sqrt{g}\phi(R+2)\right) = C^\chi \int_{\mathcal{M}_{\mathrm{g}}} e^\omega. \tag{15}$$

C is a constant, independent of g, that depends on the regularization used in defining the integral. There is no natural choice of regularization, so there is no preferred value of the constant C. However, in Equation (13), we see that Z_Σ has an additional factor $e^{S_0\chi}$. We simply eliminate the arbitrary constant C by setting $e^S = e^{S_0}C$. So finally

$$Z_\Sigma = e^{S\chi} \int_{\mathcal{M}_{\mathrm{g}}} e^\omega. \tag{16}$$

What has just been explained is a typical, though elementary, example of renormalization theory. The theory does not depend on the choice of regularization, as long as it is expressed in terms of the 'renormalized' parameter S rather than the 'bare' parameter S_0. If Σ is unorientable, the moduli space of hyperbolic structures is not a symplectic manifold, and the formula analogous to Equation (16) involves the Reidemeister or Ray-Singer torsion [14].

    So far, I have assumed that Σ is a compact surface without boundary. The case that Σ has geodesic boundaries of specified length is similar. What really led to progress in the last few years was applying JT gravity to, roughly speaking, the whole upper half-plane $\mathcal{H}$. The Euler–Lagrange field equations of JT gravity tell us that $R + 2 = 0$, and also give a certain equation for the real-valued field ϕ. Since the natural metric on $\mathcal{H}$ has $R + 2 = 0$, $\mathcal{H}$, endowed with a suitable ϕ field, can be regarded as a classical solution of JT gravity. The Euler–Lagrange equations for ϕ have a natural interpretation if one views $\mathcal{H}$ as a Kahler manifold, and thus in particular as a Riemannian manifold that also has a symplectic structure. These equations say that ϕ is the moment map for a vector field V on $\mathcal{H}$ that generates an automorphism of $\mathcal{H}$ as a Kahler manifold: thus V is a Hamiltonian vector field, with Hamiltonian function ϕ, for $\mathcal{H}$ viewed as a symplectic manifold, and also is a Killing vector field for $\mathcal{H}$ viewed as a Riemannian manifold. In short, V generates a one-parameter subgroup of the automorphism group $\mathrm{PSL}(2, \mathbb{R})$ of $\mathcal{H}$. The case that one wants (because it makes possible a limiting procedure that is described shortly) is that this is a compact subgroup $\mathrm{U}(1) \subset \mathrm{PSL}(2, \mathbb{R})$. Any two such subgroups are conjugate so it does not matter which one we choose. Concretely, if one describes $\mathcal{H}$ as the sheet $x > 0$ in a hyperboloid $x^2 - y^2 - z^2 = 1$, then we can choose the $\mathrm{U}(1)$ that rotates y and z. The moment map for a generator of this $\mathrm{U}(1)$ is simply the function x, and we take ϕ to be a positive multiple of this:

$$\phi = cx, \quad c > 0. \tag{17}$$

In this description, the conformal boundary of $\mathcal{H}$ is at $x \to \infty$, and we see that ϕ blows up everywhere near this conformal boundary.

    But it turns out that literally studying JT gravity on all of $\mathcal{H}$ is not the right thing to do. This would be rather like trying to calculate the naive integral $\int_{\mathcal{X}} e^\omega$, with $\mathcal{X} = \mathrm{diff}S^1/\mathrm{PSL}(2, \mathbb{R})$, rather than the improved version (7). A better thing to do is to study JT gravity not on all of H but on a very large region $U \subset \mathcal{H}$ [23].

    Such a large region is sketched on the left of Fig. 1. (The right of the figure will be discussed later.) We are going to study JT gravity on a two-manifold U that topologically is a disc. To get a good variational problem for JT gravity on a manifold with boundary, one needs to include in the
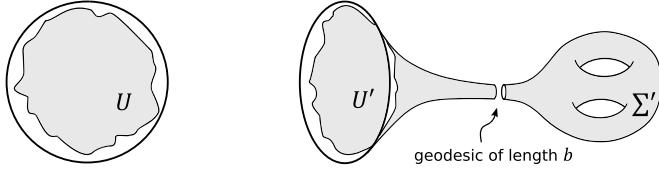
Figure 1. On the left, the upper half plane $\mathcal{H}$ is represented as a disc, and the shaded portion of the disc is a large region $U \subset \mathcal{H}$. On the right, a Riemann surface Σ is built by gluing together two pieces U' and Σ' (shaded) along an embedded geodesic of circumference b. Σ' is a surface of $g > 0$ with one geodesic boundary. U' is an annulus with a geodesic 'inner' boundary and a large 'outer' boundary. Near its outer boundary, U' resembles locally a large region in $\mathcal{H}$.

action a boundary contribution, somewhat like the boundary correction to the Einstein–Hilbert action of General Relativity [20, 21]. The JT action, with the boundary term, is

$$I_{\text{JT}} = -\frac{1}{2} \int_\Sigma \mathrm{d}^2 x \sqrt{g} \phi (R+2) - \frac{1}{2} \int_{\partial\Sigma} \mathrm{d}x \sqrt{h} \phi (K-1), \tag{18}$$

where K is the extrinsic curvature of the boundary $\partial\Sigma$, and h is the induced metric of the boundary. The Einstein–Hilbert action I_{EH} similarly needs a boundary term, so that it still equals the Euler characteristic χ. On ∂U, we impose a version of[2] Dirichlet boundary conditions: we specify the induced metric h of the boundary ∂U, along with the boundary value of[3] ϕ. Moreover, we specify them so that the circumference of ∂U is very large, as is $\phi|_{\partial U}$.

In a limit as the circumference of ∂U goes to infinity and $\phi|_{\partial U}$ is taken to be a large constant, the Feynman integral $Z_U(\beta)$ of JT gravity on U can be evaluated in terms of an integral that we already studied in Section 2: $Z_U(\beta) = e^{S_0} \int_{\mathcal{X}} \exp(\mathsf{K}/\beta + \omega)$, where $\mathcal{X} = \mathrm{diff}\, S^1/\mathrm{PSL}(2,\mathbb{R})$. The prefactor e^{S_0} comes from the Einstein–Hilbert part of the action; that is, it is $e^{S_0 \chi}$, where the disc U has $\chi = 1$. More subtle is to explain how the rest of the path integral gives the integral $\int_{\mathcal{X}} \exp(\mathsf{K}/\beta + \omega)$. Note first that $S^1 = \partial\mathcal{H}$ comes with an angular parameter that is uniquely determined up to the action of $\mathrm{PSL}(2,\mathbb{R})$. Moreover, since its induced metric is specified, ∂U has a natural arclength parameter, up to an additive constant. In the limit of interest in which U is very large, it makes sense to compare these two parameters, and the comparison defines an element of $\mathcal{X} = \mathrm{diff}\, S^1/\mathrm{PSL}(2,\mathbb{R})$. So we can think of $\mathcal{X}$ as parametrizing a family of large regions $U \subset \mathcal{H}$, up to $\mathrm{PSL}(2,\mathbb{R})$. It turns out that in the limit that $\phi|_{\partial U}$ is large, forcing ∂U closer to $\partial\mathcal{H}$, $\int_{\partial\Sigma}(K-1)$ becomes a multiple of the moment map K for the $U(1)$ action on $\mathcal{X}$. Hence in this limit, apart from the factor e^{S_0}, the Feynman integral Z_U of JT gravity on the disc U becomes our friend $\int_{\mathcal{X}} \exp(\mathsf{K}/\beta + \omega)$. To achieve convergence to this limit, both the circumference of ∂U and the constant value of $\phi|_{\partial U}$ are taken to infinity, keeping fixed their ratio, β. From Equations. (9) and (10), and setting $e^{S_0} C = e^S$, we get

$$Z_U(\beta) = \int_0^\infty \mathrm{d}E\, \rho(E) \exp(-\beta E), \quad \rho(E) = \frac{e^S}{4\pi^2} \sinh(2\pi\sqrt{E}). \tag{19}$$

Here β is regarded as the renormalized circumference of ∂U, and S is the renormalized coefficient of the Einstein–Hilbert term.

This is a deeply problematic answer for the Feynman integral on the disc U. To understand why, one should be familiar with holographic duality between gravity in the bulk of spacetime and an ordinary quantum system on the boundary [24]. If the bulk were four dimensional, as in ordinary physics,

the boundary would be three dimensional and the 'ordinary quantum system' on the boundary would be a quantum field theory—not a very easy concept to understand mathematically. But here the bulk is two dimensional, so the boundary is one dimensional and matters are simpler. An ordinary quantum system in one dimension is just described by giving a Hilbert space H and a Hamiltonian operator H acting on H. The basic recipe of holographic duality predicts that $Z_U(\beta) = \mathrm{Tr}_{\mathsf{H}} \exp(-\beta H)$.

In a moment, we will check that that prediction is false, but before doing so, I want to explain that this actually did not come as a complete surprise. Analogous calculations going back to the 1970s [21, 25] have always given the same problem: Euclidean path integrals give results that lack the expected Hilbert space interpretation. The calculations were traditionally done in models (like four-dimensional General Relativity) that were too complicated for a complete calculation, and there was always a possibility that a more complete calculation would make the issue go away. The novelty is that holographic duality and a variety of other related developments have made it possible to ask the question in a model—JT gravity—that is so simple that one can do a complete calculation, demonstrating the problem.

To see that the prediction of the duality is false, we just note the following. If we do have a Hilbert space H and a Hamiltonian H acting on it such that the operator $e^{-\beta H}$ has a trace, then H must have a discrete spectrum with eigenvalues $E_1, E_2, \cdots$ (which moreover must tend to infinity fast enough) and

$$\mathrm{Tr}_{\mathsf{H}} \exp(-\beta H) = \sum_i e^{-\beta E_i} = \int_0^\infty \mathrm{d}E \sum_i \delta(E - E_i)\, e^{-\beta E}. \tag{20}$$

But the integral over diff $S^1/\mathrm{PSL}(2,\mathbb{R})$ gave

$$Z_U(\beta) = \int_0^\infty \mathrm{d}E\, \frac{e^S}{4\pi^2} \sinh(2\pi\sqrt{E}) e^{-\beta E}. \tag{21}$$

The function $\frac{e^S}{4\pi^2}\sinh(2\pi\sqrt{E})$ is not a sum of delta functions, so the prediction of the duality is false.

However, the interpretation via JT gravity gives us a key insight that we did not have when we were just abstractly integrating over $\mathcal{X} = \mathrm{diff}S^1/\mathrm{PSL}(2,\mathbb{R})$. In Equation (10), C was an arbitrary constant, but now C has been replaced with e^S where S is should be large to get a model more similar to the real world. This was explained following Equation (12). To amplify on the point a bit, S is analogous to black hole entropy in four dimensions; it is very large in any situation in which the usual puzzles of quantum black hole physics arise. If S is just moderately large, say of order 100, then e^S is huge. So we think of e^S as a huge number. With this in mind, the function $\frac{e^S}{4\pi^2}\sinh(2\pi\sqrt{E})$ actually can be well-approximated for many purposes as $\sum_i \delta(E - E_i)$, for suitable E_i. One must look very closely to see the difference. One could not reasonably approximate $2\sinh(2\pi\sqrt{E})$ (for example) by a such a sum, but $e^{100}\sinh(2\pi\sqrt{E})$ is another matter.

4. A random matrix ensemble

The novel idea of Saad, Shenker and Stanford [12] was to interpret the function $\frac{e^S}{4\pi^2}\sinh 2\pi\sqrt{E}$ not as the density of energy levels of a particular Hamiltonian but as the average level density of an ensemble of Hamiltonians—a random matrix. In terms of the physics involved, this interpretation is rather provocative, though it may be a challenge to convey this fully.

What motivated this interpretation? One clue came from the work of Kitaev [15], who discovered a simple model of holographic duality based on a random ensemble that was more complicated than the one of Saad et al. Unfortunately, to explain this would take us rather far afield. Another clue came from the prior history of relations between random matrices and two-dimensional gravity [8–11]. A final clue, more directly related to the topic of the present article, involved the volumes of the moduli spaces of Riemann surfaces. As noted in the introduction, Mirzakhani [6] had found a powerful new way to compute the volumes of these moduli spaces. And Eynard and Orantin [13] had interpreted her formulas in terms of their notion of topological recursion [22], which is related to a random matrix ensemble. The key technical observation of Saad et al. was that the eigenvalue density of a random matrix ensemble related to the spectral curve of Eynard and Orantin is precisely the function $\frac{e^S}{4\pi^2}\sinh 2\pi\sqrt{E}$ that arises in JT gravity and integration over $\mathrm{diff}\,S^1/\mathrm{PSL}(2,\mathbb{R})$.

The sort of random matrix ensemble used in [12] is the following. M will be an $N \times N$ hermitian matrix for some N. We are really interested in N very large and ultimately in a limit with $N \to \infty$. Picking some suitable real-valued function $T(M)$, we consider the integral

$$Z(T;N) = \frac{1}{\mathrm{vol}(\mathrm{U}(N))} \int \mathrm{d}M \exp(-N\mathrm{Tr}\,T(M)). \tag{22}$$

(The space of $N \times N$ Hermitian matrices is a copy of $\mathbb{R}^{N^2}$, and the measure $\mathrm{d}M$ is the standard Euclidean measure on $\mathbb{R}^{N^2}$.) If the function $T(M)$ is quadratic, this is a Gaussian random matrix ensemble, as studied originally in the 1960s by Wigner, Dyson, Mehta and many others. We are interested in the case that $T(M)$ is not quadratic. In this case, $Z(T;N)$ or more precisely its logarithm has an asymptotic expansion for large N:

$$\log Z(T;N) = N^2 F_0(T) + F_1(T) + \frac{1}{N^2}F_2(T) + \cdots = \sum_{g=0}^{\infty} N^{2-2g} F_g(T). \tag{23}$$

This expansion can be constructed by standard Feynman diagram methods. As shown by 't Hooft [26], the Feynman diagrams for a matrix model are in a natural way ribbon graphs, each of which can be naturally drawn on a certain Riemann surface. $F_g(T)$ is the contribution of connected ribbon graphs that can be drawn on a surface of genus g.

However, instead of making a Feynman diagram expansion, we can try to just evaluate the integral [27]. We do this by first diagonalizing M, writing $M = U\Lambda U^{-1}$, with $U \in \mathrm{U}(N)$, and Λ a diagonal matrix $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_N)$, $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_N$. The measure becomes $\mathrm{d}M = \mathrm{d}U \prod_i \mathrm{d}\lambda_i \prod_{j<k}(\lambda_j - \lambda_k)^2$. Here $\mathrm{d}U$ is Haar measure on the group $\mathrm{U}(N)$, and the integral over U just cancels the factor $1/\mathrm{vol}(\mathrm{U}(N))$ in the definition of $Z(T;N)$. So we reduce to

$$Z(T;N) = \int \mathrm{d}\lambda_1 \ldots \mathrm{d}\lambda_N \prod_{i<j}(\lambda_i - \lambda_j)^2 \exp\left(-N\sum_k T(\lambda_k)\right). \tag{24}$$

For large N, the integrand $\prod_{i<j}(\lambda_i - \lambda_j)^2 \exp(-N\sum_k T(\lambda_k))$ has a sharp maximum as a function of the λ_i. To describe this maximum, assuming N to be very large, we think in terms of a continuous distribution of eigenvalues, $\rho(\lambda) = N\rho_0(\lambda)$ for some function $\rho_0(\lambda)$ (normalized to $\int \mathrm{d}\lambda\rho_0(\lambda) = 1$). In terms of such a distribution, the integrand in (24) becomes
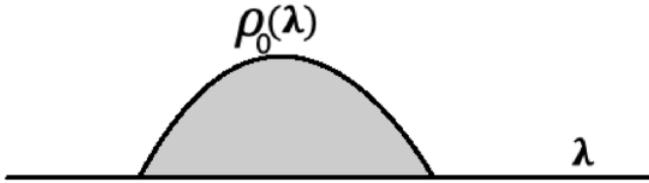
Figure 2. For a nice class of functions $T(\lambda)$, the support of the eigenvalue density $\rho_0(\lambda)$ is an interval $[a, b]$

$$\exp\left(N^2\left(-\int d\lambda \rho_0(\lambda)T(\lambda) + \int d\lambda d\lambda' \rho_0(\lambda)\rho_0(\lambda')\log|\lambda - \lambda'|\right)\right). \tag{25}$$

For a nice class of functions $T(\lambda)$, the exponent has a unique maximum at some function $\rho_0(\lambda)$. The most famous case is that T is quadratic; then (with suitable normalization) $\rho_0(\lambda) = \frac{2}{\pi}\sqrt{1 - \lambda^2}$ in the interval $[-1, 1]$ and $\rho_0(\lambda) = 0$ outside the interval. This example illustrates a general behavior; for a nice class of functions T, the support of the function $\rho_0(\lambda)$ is an interval $[a, b]$; moreover, $\rho_0(\lambda)$ vanishes near the endpoints of the interval as the square root of the distance to the endpoint, and can be continued to an algebraic function of λ with only those singularities.[4]

If $F_0(T)$ is the value of the exponent at its maximum, then the leading approximation to the logarithm of the integral is

$$\log Z(N; T) = N^2 F_0(T) + \mathcal{O}(1). \tag{26}$$

How can we calculate the further terms in this expansion? Something quite remarkable happens, though a detailed explanation would go beyond the scope of this article. (There is a brief introduction in Section 4 of [14], for example.) One should define the 'spectral curve' in the $y - \lambda$ plane:[5]

$$y^2 = -\pi^2 \rho_0^2(\lambda). \tag{27}$$

This curve is a double cover of the λ plane, branched at the endpoints of the eigenvalue distribution, where the function $\rho_0^2(\lambda)$ has simple zeroes. Once one knows this spectral curve, one can forget about doing integrals and one can forget the original function $T(M)$. The whole expansion

$$\log Z(T; N) = N^2 F_0(T) + F_1(T) + \frac{1}{N^2}F_2(T) + \cdots = \sum_{g=0}^{\infty} N^{2-2g}F_g(T) \tag{28}$$

(and a wide variety of other things about this ensemble) can be worked out just from a knowledge of the spectral curve. A very useful version of this process is the 'topological recursion' of Eynard and Orantin [22].

Now let us go back to the volumes of moduli spaces. As I have explained, Saad, Shenker and Stanford interpreted the function $\frac{e^S}{4\pi^2}\sinh 2\pi\sqrt{E}$ as the average density of eigenvalues for a random matrix drawn from a non-Gaussian hermitian matrix ensemble. In principle, the procedure is to start with a function T and then compute the density of eigenvalues $N\rho_{0,T}(E)$, where we now denote the eigenvalue as E, and we also make explicit the dependence of ρ_0 on T. Then we take $N \to \infty$ while adjusting T so that $N\rho_{0,T}(E)$ converges to the desired $\frac{e^S}{4\pi^2}\sinh 2\pi\sqrt{E}$. This process is called double-scaling.

But we can skip all that work, because everything we want to compute depends only on the spectral curve, which we know is going to be

$$y^2 = -\frac{1}{16\pi^2}\sinh^2(2\pi\sqrt{E}).\tag{29}$$

This curve is a double cover of the complex E-plane, with branch points at $E = 0$ and $E = \infty$. (In the process of double-scaling to get this curve, the interval $[a, b]$ converges to $[0, \infty)$.) In short, all we have to do is to start with that spectral curve, and apply topological recursion to get the expansion $\log Z(T) = \sum_{g=0}^{\infty} e^{(2-2g)S}F_g$, as well as other quantities of interest that are introduced momentarily. Here, after double-scaling, the expansion parameter is e^{-S} rather than $1/N$.

Now we can compute volumes, with a few steps, as follows. First we compute the average of $\mathrm{Tr}\exp(-\beta H)$ in this matrix ensemble, where now I refer to the random matrix M as a Hamiltonian H. This can be done explicitly, applying topological recursion to the spectral curve (29). The result is an expansion of the ensemble average $\langle\mathrm{Tr}\exp(-\beta H)\rangle$ in powers of e^{-2S}.

To interpret the result in terms of volumes, one proceeds as follows. The Feynman diagram expansion of $\langle\mathrm{Tr}\exp(-\beta H)\rangle$ involves Feynman diagrams drawn on a Riemann surface Σ with one boundary component, like those drawn in Fig. 1. This happens because when we make a Feynman diagram expansion, the trace $\mathrm{Tr}\exp(-\beta H)$ turns into a boundary. The picture on the left of Fig. 1 actually corresponds to the leading contribution, the special case that Σ is a disc. In Section 3, we analyzed the JT path integral for the case of a disc of regularized circumference β, and interpreted the answer as $\int_0^{\infty}\mathrm{d}E\rho(E)\exp(-\beta E)$ with $\rho(E) = \frac{e^S}{4\pi^2}\sinh 2\pi\sqrt{E}$. We have chosen a matrix ensemble that reproduces this answer for the leading contribution. We are now interested in the higher topologies shown on the right of Fig. 1. They contribute the higher order terms in the expansion in powers of e^{-S}. The contribution from Feynman diagrams drawn on a genus g surface Σ is of relative order $\exp(-2gS)$.

In the right-hand panel of Fig. 1, the surface Σ has been cut in two along a geodesic of length b. To the right of the cut is a hyperbolic Riemann surface Σ' of genus g, with one geodesic boundary of length b. Let $\mathcal{M}_{g,b}$ be the moduli space of such hyperbolic surfaces, and $V_{g,b}$ its volume. To the left of the cut is a hyperbolic two-manifold U' that is topologically an annulus. U' has one geodesic boundary of length b and another 'large' boundary near which U' looks locally like a large portion of $\mathcal{H}$. In the same sense that the disc U depicted on the left of Fig. 1 represents diffS^1/PSL$(2, \mathbb{R})$, U' represents diffS^1/U(1). Apart from an overall multiplicative constant, the symplectic structure of diffS^1/U(1) depends on one parameter[6] b, which corresponds to the length of the geodesic boundary of U'. Let us just write $\Theta(b; \beta)$ for the JT integral on this orbit (β being again the regularized circumference of the 'large' boundary of U'). The function $\Theta(b; \beta)$ can be computed from the DH/AB formula and is given by a formula similar to Equation (9). Roughly because of locality properties of quantum field theory, or because the length of a geodesic on a hyperbolic two-manifold is the Hamiltonian that generates a corresponding Dehn twist, the JT path integral Z_Σ on Σ is obtained by integrating over b the product of the JT path integral on U' times the JT path integral on Σ':

$$Z_\Sigma = e^{S(1-2g)}\int_0^{\infty} b\,\mathrm{d}b\,\Theta(b; \beta)V_{g,b}.\tag{30}$$

One has to include here a factor of b that comes from integrating over a 'twist' parameter in the gluing of U' to Σ'. We have used the fact that the JT path integral on Σ' is $e^{S\chi(\Sigma')}V_{g,b} = e^{S(1-2g)}V_{g,b}$.

On the other hand, the Feynman diagram expansion of the matrix model makes us expect that Z_Σ will be the term of order $e^{S(1-2g)}$ in the expansion of $\langle\mathrm{Tr}\exp(-\beta H)\rangle$. This can be computed by

applying topological recursion to the spectral curve (29). Using also the explicit result for $\Theta(b; \beta)$, one can make an inverse Laplace transform of the result for Z_Σ to obtain an explicit formula for $V_{g,b}$.

This procedure can be generalized to give the volumes $V_{g,\vec{b}}$ of moduli spaces $\mathcal{M}_{g,\vec{b}}$ of hyperbolic surfaces of genus g with boundary geodesics of specified lengths $\vec{b} = (b_1, b_2, \ldots, b_n)$. The resulting formulas are correct, in view of the relation demonstrated by Eynard and Orantin [13] between topological recursion for the spectral curve $y^2 = -\frac{1}{16\pi^2} \sinh^2 2\pi\sqrt{E}$ and the recursion relation used by Mirzakhani [6] to compute $V_{g,\vec{b}}$.

Matching with Mirzakhani's recursion relation was how Eynard and Orantin determined which spectral curve to use. Another way is to use the relation between volumes and intersection numbers and the general relation of intersection numbers to spectral curves. (This approach is sketched in [28]; see Section 2.4 and Equation (4.46).) As I have explained, Saad, Shenker and Stanford instead arrived at the same spectral curve from the path integral of JT gravity on a disc.

The approach involving JT gravity is very interesting for physicists, but if one only cares about volumes of moduli spaces, one might ask why it is important. One answer is that this derivation sheds a new light on the relationship between diff S^1/PSL$(2, \mathbb{R})$ and the moduli spaces of Riemann surfaces. Another answer is that this approach possibly gives a better understanding of why random matrix ensembles are related to volumes and intersection numbers. A third answer is given by my work with Stanford [14]. We ran the entire story for super Riemann surfaces. Every step has a direct analog for that case.

5. Matrix ensembles and volumes of super moduli spaces

The superanalog of JT gravity is JT supergravity, which computes the volumes of moduli spaces of super Riemann surfaces, in general with geodesic boundaries of specified lengths. As before, it is important to consider the special case of a super Riemann surface which is the super upper half plane $\widehat{\mathcal{H}}$, or more precisely a very large region $\widehat{U} \subset \widehat{\mathcal{H}}$, as in the left-hand side of the familiar Fig. 1. The boundary of $\widehat{U}$, or simply the boundary of $\widehat{\mathcal{H}}$, is a super circle that we may call $S^{1|1}$. I also write Sdiff S^1 for its group of orientation-preserving diffeomorphisms. Let $\widehat{\mathcal{X}} = $ Sdiff S^1/OSp$(1|2)$, which we might think of as universal super Teichmüller space. By similar arguments to those that were described in Section 3, the super JT path integral on a large region $\widehat{U} \subset \widehat{\mathcal{H}}$ computes

$$Z_{\widehat{U}}(\beta) = e^{S_0} \int_{\widehat{\mathcal{X}}} \exp(\widehat{\mathsf{K}}/\beta + \widehat{\omega}), \tag{31}$$

where $\widehat{\mathsf{K}}$ is the moment map for a subgroup U$(1) \subset$ Sdiff S^1, $\widehat{\omega}$ is the symplectic form of $\widehat{\mathcal{X}}$, and S_0 is the coefficient of the Einstein–Hilbert term in the action. This integral can be computed using DH/AB localization, with the result [14]

$$Z_{\widehat{U}}(\beta) = \int_0^\infty \mathrm{d}E\, e^{-\beta E} \widehat{\rho}(E), \quad \widehat{\rho}(E) = \frac{e^S \sqrt{2}}{\pi} \frac{\cosh(2\pi\sqrt{E})}{\sqrt{E}}. \tag{32}$$

As before, S is the renormalized coefficient of the Einstein–Hilbert term. Again, this is not $\mathrm{Tr}_{\mathsf{H}} \exp(-\beta H)$ for a Hamiltonian H acting on a Hilbert space H, but now we know what to do: we have to consider a random quantum ensemble.

We can rerun the previous story with a few changes. The formula for $Z_{\widehat{U}}(\beta)$ tells us the spectral curve:

$$y^2 = -\frac{2}{E}\cosh^2(2\pi\sqrt{E}). \tag{33}$$

However, the matrix ensemble is of a different type than we encountered before. One way to see that it must be different is that the eigenvalue density $\widehat{\rho}(E)$ in Equation (32) behaves as $E^{-1/2}$ near the endpoint of the energy spectrum at $E=0$, in contrast to the typical $E^{+1/2}$ behavior of a matrix ensemble of the sort that we studied in Section 4.

The ensemble must be different because as we are now studying super Riemann surfaces rather than ordinary ones, the dual quantum mechanical system is now supposed to be supersymmetric. So we need to do random supersymmetric quantum mechanics, not just random quantum mechanics.

In supersymmetric quantum mechanics, the Hilbert space H is $\mathbb{Z}_2$-graded by an operator

$$(-1)^F = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}, \tag{34}$$

where I is the identity operator. The hamiltonian H commutes with the $\mathbb{Z}_2$-grading, but it is supposed to be the square of an odd self-adjoint operator Q, that is, a self-adjoint operator that anticommutes with $(-1)^F$:

$$Q = \begin{pmatrix} 0 & P \\ P^\dagger & 0 \end{pmatrix}, \quad H = Q^2 = \begin{pmatrix} PP^\dagger & 0 \\ 0 & P^\dagger P \end{pmatrix}. \tag{35}$$

We then consider a random ensemble for Q defined by the measure $\exp(-N\mathrm{Tr}\,T(Q^2))$, for some suitable function T. If the function T is linear, this is a Gaussian ensemble; here we are interested in a non-Gaussian ensemble. Ensembles of this type have been introduced (originally in the Gaussian case) in [29–31].

Random supersymmetric quantum mechanics leads easily to the $E^{-1/2}$ behavior of the density of eigenvalues near $E=0$ that we see in Equation (32). Let μ be an eigenvalue of Q, and let $f(\mu)\mathrm{d}\mu$ be the corresponding normalized density of eigenvalues. Since the ensemble for Q is invariant under $Q \to -Q$, we have $f(\mu) = f(-\mu)$. The condition $f(0) \neq 0$ is generic. Now since $H = Q^2$, the relation between an eigenvalue E of H and an eigenvalue μ of Q is $E = \mu^2$. Since $f(\mu)\mathrm{d}\mu = f(\sqrt{E})\mathrm{d}E/2\sqrt{E}$, the density of eigenvalues of H behaves as $E^{-1/2}$ near $E=0$. So an ensemble of this kind is a good candidate for interpreting the formula (32) for $Z_{\widehat{U}}(\beta)$.

For such a 'supersymmetric' matrix ensemble, there is again a version of topological recursion. Applying this to the spectral curve of Equation (33), we get an expansion of $\langle \mathrm{Tr}_{\mathsf{H}}\exp(-\beta H)\rangle$ in powers of e^{-2S}. More generally, we can study the expectation value of a product of traces $\langle \prod_{i=1}^n \mathrm{Tr}_{\mathsf{H}}\exp(-\beta_i H)\rangle$. The terms in the expansion are related to the volumes of supermoduli spaces in the same manner as in the bosonic case.

In this way, Stanford and I deduced a recursion relation that determines the volumes of the supermoduli spaces $\mathfrak{M}_{g,\vec{b}}$. Moreover, we were able to prove this formula by finding analogs of results of Mirzakhani and of Eynard and Orantin. By imitating Mirzakhani's derivation, we obtained a Mirzakhani-style recursion relation for the volumes of supermoduli spaces. (See also [32] for a related super McShane identity.) And similarly to the arguments of Eynard and Orantin, we

showed that the recursion relation that comes from the matrix ensemble agrees with the Mirzakhani-style recursion relation. Thus the relation of random matrices, volumes and gravity has a perfect counterpart in the supersymmetric case.

Actually, the supersymmetric random matrix ensemble that was just described potentially involves an integer invariant, the index of the operator P, or equivalently the difference in dimension between the even and odd subspaces of H. It turns out that to compute volumes of supermoduli spaces $\mathfrak{M}_{g,\vec{b}}$ of super Riemann surfaces with geodesic boundaries, one should take the index to vanish. But what happens if the index is nonzero? At least in low orders of the expansion in powers of e^{-S}, the answer is known: the same type of ensemble based on the same spectral curve but with a nonzero index can be used to compute volumes of moduli spaces of super Riemann surfaces with Ramond punctures as well as geodesic boundaries. (A Ramond puncture represents a certain type of singularity in the superconformal structure of a super Riemann surface; roughly, the spin structure is branched over a Ramond puncture.) For this, one takes the size of the matrices to infinity, keeping the index fixed. A complete proof to all orders in e^{-S} is not yet available in this case.

Ramond punctures are the only punctures that add something essentially new to the computation of volumes, once one has already analyzed surfaces with geodesic boundaries. For purposes of computing volumes, a puncture of an ordinary Riemann surface, or a Neveu–Schwarz puncture of a super Riemann surface (which is a similar notion), is equivalent to the small b limit of a geodesic boundary of length b. Nothing like that is true for Ramond punctures.

Another important detail concerns the spin structures on a Riemann surface, which may be 'even' or 'odd'. The two cases are distinguished by a topological invariant, interpreted by Atiyah as the mod 2 index of the Dirac operator [**33**]. Accordingly, $\mathfrak{M}_g$ has two components, parametrizing super Riemann surfaces with even or odd spin structure. In the case of a super Riemann surface with boundary, there is another topological invariant (also interpretable as a mod 2 index) that specifies whether the spin structure on a given boundary component is even or odd. To get a full answer for the volumes in all cases, one has to consider also a somewhat different matrix ensemble in which it is still true that $H = Q^2$ (leading again to $\widehat{\rho}(E) \sim E^{-1/2}$ near $E = 0$), but a $\mathbb{Z}_2$-grading by $(-1)^F$ is not assumed.

Finally, one can make a similar analysis for unorientable two-manifolds. Even if Σ is unorientable, we can still define a moduli space of hyperbolic structures on Σ. These moduli spaces are not symplectic, but JT gravity or supergravity determines volume forms on them. It turns out that these volume forms can be computed using the Reidemeister or Ray-Singer torsion of a flat connection (whose structure group is the full symmetry group of $\mathcal{H}$ or $\widehat{\mathcal{H}}$, including symmetries that reverse the orientation). In the bosonic case, the same volume form was first defined by Norbury [**34**] in another way. The moduli space volumes are in most cases divergent if Σ is unorientable, but the volume forms can be compared to random matrix theory. In a detailed analysis [**14**], all 10 standard types of random matrix ensemble [**31**] make an appearance.

Acknowledgement

Appendix

1. The group $\mathrm{PSL}(2, \mathbb{R})$ is contractible onto $\mathrm{U}(1)$, so a $\mathrm{PSL}(2, \mathbb{R})$ bundle over a surface Σ has an integer invariant, the first Chern class. The flat bundle related to a hyperbolic metric has first

Chern class $2 - 2g$. A similar remark applies for the supergroup $\mathrm{OSp}(1|2)$ introduced shortly. The maximal bosonic subgroup of $\mathrm{OSp}(1|2)$ is the spin double cover $\mathrm{SL}(2, \mathbb{R})$ of $\mathrm{PSL}(2, \mathbb{R})$, so in that case the relevant value of the first Chern class is $1 - g$.

2. Note in particular that we do *not* constrain ∂U to be a geodesic. Classically, it would be impossible to impose such a constraint, since a disc does not admit a hyperbolic metric with geodesic boundary. But more to the point, we want a boundary condition that leads U to be, in some sense, a good approximation to $\mathcal{H}$.

3. With this boundary condition, Equation (18) leads to a sensible classical variational problem (and quantum path integral) regardless of whether we use K or $K - 1$ in the boundary term. We use $K - 1$ because $K - 1$ vanishes in the limit of a large disc in $\mathcal{H}$. This is important for the existence of the limit discussed momentarily in which U becomes large.

4. Because of the square root singularities at $\lambda = a, b$, the analytically continued function $\rho_0(\lambda)$ is imaginary on the real axis outside the interval $[a, b]$, so in terms of the continued function, the eigenvalue density is $\mathrm{Re}\,\rho_0(\lambda)$ (Fig. 2). For example, note that the function $\frac{2}{\pi}\sqrt{1 - \lambda^2}$ which arises in the Gaussian case is imaginary on the real axis outside the interval $[-1, 1]$.

5. The factor of π^2 here is conventional. Its purpose is to avoid a factor of π in the relationship between y and the expectation value of the matrix resolvent $\mathrm{Tr}\,\frac{1}{\lambda - M}$.

6. As explained for example in section 3.1 of [**14**], $\mathrm{diff}S^1/\mathrm{U}(1)$ can be viewed as a coadjoint orbit of the Virasoro group (that is, the central extension of $\mathrm{diff}S^1$). But actually, there is a one-parameter family of coadjoint orbits that are all isomorphic to $\mathrm{diff}S^1/\mathrm{U}(1)$, so the symplectic structure of $\mathrm{diff}S^1/\mathrm{U}(1)$ depends on a parameter b, in addition to the possibility of scaling it by an overall constant.

References

1. M. F. Atiyah and R. Bott, The Yang-Mills equations over Riemann surfaces, *Phil. Trans. Roy. Soc. Longon*, **A308** (1982), 52–3.
2. P. Norbury, A new cohomology class on the moduli space of curves, arXiv:1717.03662.
3. M. Mirzakhani, Weil-Petersson Volumes and intersection theory on the moduli space Of curves, *Journal of the American Mathematical Society*, **20** no. 1 (2006), 1–23.
4. Y. I. Manin and P. Zograf, Invertible cohomological field theories and Weil-Petersson volumes, *Annales de l'institut Fourier*, **50** no. 2 (2000), 519–535 arXiv: math/9902051.
5. P. Zograf, On the large genus asymptotics of Weil-Petersson volumes, arXiv:0812.0544.
6. M. Mirzakhani, Simple geodesics and Weil-Petersson volumes of moduli spaces of bordered Riemann surfaces, *Invent. Math*, **167** no. 1 (2006), 179–222.
7. E. Witten, Two-dimensional gravity and intersection theory on moduli space, *Surveys Diff. Geom*, **1** (1991), 243–310.
8. E. Brezin and V. Kazakov, Exactly solvable field theories of closed strings, *Physics Letters* B, **236** no. 2 (1990), 14–4.
9. M. Douglas and S. Shenker, Strings in less than one dimension, *Nuclear Physics* B, **335** no. 3 (1990), 63–5 **B335**.
10. D. J. Gross and A. Migdal, Nonperturbative Two-dimensional quantum gravity, *Phys. Rev. Lett.*, **64** no. 2 (1990), 12–7.
11. M. Kontsevich, Intersection theory on the moduli space of curves and the matrix airy function, *Commun. Math. Phys.*, **147** no. 1 (1992), 1–23.

12. P. Saad, S. Shenker and D. Stanford, *JT gravity as a matrix integral*, arXiv:1903.11115.

13. B. Eynard and N. Orantin, Invariants of algebraic curves and topological expansion, *Communications in Number Theory and Physics*, **1** no. 2 (2007), 347–452.

14. D. Stanford and E. Witten, *JT gravity and the ensembles of random matrix theory*, arXiv: 1907.03363.

15. A. Kitaev, *A simple model of quantum holography*, http://online.kitp.ucsb.edu/online/entangled15/kitaev/, http://online.kitp.ucsb.edu/online/entangled15/kitaev2/, talks at KITP, April 7 2015, May 27 2015.

16. J. Maldacena and D. Stanford, Comments on the Sachdev-Ye-Kitaev Model, *Phys. Rev.*, **D94** (2016), 106002. arXiv:1604.07818.

17. D. Stanford and E. Witten, Fermionic localization of the Schwarzian theory, *Journal of High Energy Physics*, **2017** no. 10 (2017), 00–8 arXiv:170-3.04612.

18. J. J. Duistermaat and G. J. Heckman, On the variation in the Cohomology of the Symplectic form of the reduced phase space, *Inventiones Mathematicae*, **69** no. 2 (1982), 259–268.

19. M. F. Atiyah and R. Bott, The moment map and equivariant Cohomology, *Topology*, **23** no. 1 (1984), 1–28.

20. J. W. York, Role of conformal Three-Geometry in the dynamics of gravitation, *Phys. Rev. Lett.*, **28** no. 16 (1972), 108–2.

21. G. W. Gibbons and S. W. Hawking, Action integrals and partition functions in quantum gravity, *Phys. Rev.*, **D15** (1977), 275–2.

22. B. Eynard and N. Orantin, *Invariants of algebraic curves and topological expansion*, arXiv:math-ph/0702045.

23. J. Maldacena, D. Stanford and Z. Yang, Conformal symmetry and its breaking in two-dimensional nearly anti-de Sitter space, *Prog. Theor. Exp. Phys.*, **2016** no. 12 (2016), 12C104 arXiv:1606.01857.

24. J. Maldacena, The large N limit of superconformal field Theories and Supergravity, *International Journal of Theoretical Physics*, **38** no. 4 (1999), 1113–1133.

25. S. W. Hawking, The path integral approach to quantum gravity, in *General Relativity - An Einstein Centenary Survey* (Cambridge University Press, 1977), eds. S. W. Hawking and W. Israel.

26. G. 't Hooft, A planar diagram theory for strong interactions, *Nucl. Phys.*, (1974), 461–73 **B72**.

27. E. Brézin, C. Itzykson, G. Parisi and J. B. Zuber, Planar Diagrams, *Communications in Mathematical Physics*, **59** no. 1 (1978), 35–51.

28. R. Dijkgraaf and E. Witten, Developments in topological gravity, in Topology and Physics, eds. C. N. Yang, M. L. Ge, and Y. H. He (World Scientific, 2018), arxiv:1804.03275.

29. J. Wishart, The generalised product moment distribution in samples from a normal multivariate population, *Biometrika*, **20A** no. 1-2 (1928), 32–52.

30. J. J. M. Verbaarschot and I. Zahed, Spectral density of the QCD Dirac operator near zero virtuality, *Physical Review Letters*, **70** no. 25 (1993), 3852–3855 arXiv:hep-th/9303012.

31. A. Altland and M. R. Zirnbauer, Nonstandard symmetry classes in Mesoscopic Normal-Superconducting Hybrid Structures, *Physical Review B*, **55** no. 2 (1997), 114–2.

32. Y. Huang, R. C. Penner and A. Zeitlin, srcSuper McShane Identity, arXiv:1907.09978.

33. M. F. Atiyah, Riemann surfaces and Spin structures, *Ann. Sci. de l'É.N.S.*, **4** (1971), 47–62.

34. P. Norbury, Lengths of geodesics on Non-Orientable Hyperbolic Surfaces, *Geometriae Dedicata*, **134** no. 1 (2008), 153–76 arXiv:math/0612128.