

# A holomorphic and background independent partition function for matrix models and topological strings

Bertrand Eynard<sup>a</sup> and Marcos Mariño<sup>b</sup>

<sup>a</sup> Institut de Physique Théorique, CEA, IPhT  
F-91191 Gif-sur-Yvette, France  
CNRS, URA 2306, F-91191 Gif-sur-Yvette, France  
eynard@spht.saclay.cea.fr

<sup>b</sup>Section de Mathématiques et Département de Physique Théorique  
Université de Genève, CH-1211 Genève, Switzerland  
marcos.marino@unige.ch

## Abstract

We study various properties of a nonperturbative partition function which can be associated to any spectral curve. When the spectral curve arises from a matrix model, this nonperturbative partition function is given by a sum of matrix integrals over all possible filling fractions, and includes all the multi-instanton corrections to the perturbative  $1/N$  expansion. We show that the nonperturbative partition function, which is manifestly holomorphic, is also modular and background independent: it transforms as the partition function of a twisted fermion on the spectral curve. Therefore, modularity is restored by nonperturbative corrections. We also show that this nonperturbative partition function obeys the Hirota equation and provides a natural nonperturbative completion for topological string theory on local Calabi–Yau threefolds.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>The nonperturbative partition function</b>	<b>4</b>
<b>3</b>	<b>Modular properties of the nonperturbative partition function</b>	<b>6</b>
3.1	Modular transformations of the perturbative amplitudes . . . . .	6
3.2	Transformation properties of the theta function . . . . .	9
3.3	Generalization to higher rank . . . . .	11
3.4	Transformation properties of the nonperturbative partition function . . . . .	15
3.5	Background independence . . . . .	20
<b>4</b>	<b>Matrix models and sums over filling fractions</b>	<b>20</b>
4.1	Matrix integrals and domains of integration . . . . .	21
4.2	Topological perturbative expansion . . . . .	22
4.3	Summation over filling fractions . . . . .	22
4.4	Remarks on background independence . . . . .	24
<b>5</b>	<b>The nonperturbative partition function for topological strings</b>	<b>25</b>
5.1	Topological strings, matrix models and spectral curves . . . . .	25
5.2	Topological strings and the nonperturbative partition function . . . . .	26
<b>6</b>	<b>Integrability</b>	<b>30</b>
<b>7</b>	<b>Discussion</b>	<b>32</b>
<b>A</b>	<b>Symplectic invariants</b>	<b>34</b>
A.1	Spectral curve . . . . .	34
A.2	Correlators and symplectic invariants . . . . .	35
A.3	Derivatives . . . . .	35

## 1 Introduction

The perturbative partition function of a matrix model with fixed filling fractions  $\epsilon$  has the form

$$Z = \exp\left\{\sum_{g=0}^{\infty} N^{2-2g} F_g(\epsilon)\right\}, \quad (1.1)$$

where  $F_g(\epsilon)$  is the generating function for fatgraphs of genus  $g$  [19]. The same structure appears in the partition function of closed topological string theory, where  $1/N$  becomes the string coupling constant  $g_s$  and the  $\epsilon$  become closed string moduli. In both cases, the partition function depends on a choice of background  $\epsilon$ .

The background dependence of  $Z$  is closely related to its behavior under the modular group of the theory. In the context of matrix models, the large  $N$  limit is described by an algebraic curve called the spectral curve, and the modular group is simply the symplectic group  $\mathrm{Sp}(2\bar{g}, \mathbb{Z})$ , where  $\bar{g}$  is the genus of the spectral curve (not to be confused with the genera  $g$  appearing in the topological expansion). In topological string theory on a Calabi–Yau threefold  $X$ , the modular group is the symplectic group  $\mathrm{Sp}(2n, \mathbb{Z})$  of symmetries which preserve the symplectic form, where  $n = b_3(X)/2$ . The mathematical manifestation of background dependence is that, as emphasized in [1], the partition function  $Z$  does not have good transformation properties under the modular group of the theory. In fact, as shown in [1] in topological string theory and in [25] in the context of matrix models, the  $F_g$  transform as quasi-modular forms, with shifts (the prototype for this behavior is the second Eisenstein series).

It is possible to restore modularity of  $F_g$  (hence of  $Z$ ) at the price of introducing a non-holomorphic dependence on  $\bar{\epsilon}$ . When this is done, the resulting partition function  $Z(\epsilon, \bar{\epsilon})$  satisfies the holomorphic anomaly equations of [8]. In the context of matrix models this was shown in [24]. Indeed, the holomorphic anomaly was interpreted in [47] as an obstruction to background independence. Therefore, the lack of background independence seems to face us with a choice between modularity and holomorphicity.

In the case of matrix models, it is however clear that the original matrix integral which leads to the above  $1/N$  expansion can not depend on the choice of filling fractions. It should only depend on the coupling constants of the potential, the rank of the matrix  $N$ , and a choice of integration path for the eigenvalues. Therefore, for matrix models, background dependence is an artifact of the  $1/N$  expansion, coming from the fact that one has chosen a particular saddle configuration at large  $N$ . Background independence should be restored by including the rest of the saddle points in an appropriate way, i.e. by including the instanton configurations of the matrix model. In this sense, matrix integrals provide a simple framework in which we might understand both the breakdown of background independence and the appropriate mechanism to restore it.

In [22], one of us proposed an asymptotic formula for the partition function of convergent matrix integrals, generalizing the results of [11]. This formula includes, together with the perturbative expansion (1.1), a series of nonperturbative corrections which can be interpreted in terms of instantons of the matrix model. Since this nonperturbative partition function is obtained by summing over all possible filling fractions, it was suggested in [22] that it gives a natural proposal for a background independent partition function.

As in the case of the  $F_g$  [25], the nonperturbative partition function introduced in [22] can be defined for *any* spectral curve  $\Sigma$ . On top of the spectral curve data, one also needs a choice of characteristics  $(\mu, \nu)$ , just as for theta functions on a Riemann surface. These characteristics encode nonperturbative information; for example, in a matrix model they encode the choice of integration contour for the eigenvalues. In this paper we study in detail the transformation properties of the nonperturbative partition function  $Z_\Sigma(\mu, \nu; \epsilon)$  under the modular group. It turns out that they have good modular properties. More precisely, they transform in a matrix representation of the modular group, i.e.

$$\tilde{Z}_\Sigma(\tilde{\mu}, \tilde{\nu}; \tilde{\epsilon}) = \zeta_{[\nu]}^{[\mu]}(\Gamma) Z_\Sigma(\mu, \nu; \epsilon), \quad (1.2)$$

where  $\zeta_{[\nu]}^{[\mu]}(\Gamma)$  is a phase depending on the characteristics  $\mu, \nu$  and the modular transformation  $\Gamma$ , and  $\tilde{\mu}, \tilde{\nu}$  are new, transformed characteristics. Both  $\zeta_{[\nu]}^{[\mu]}(\Gamma)$  and  $\tilde{\mu}, \tilde{\nu}$  are the same quantities which appear in the transformation properties of higher rank theta functions. Since  $Z_\Sigma(\mu, \nu; \epsilon)$  is manifestly holomorphic, we obtain a partition function which is holomorphic, modular, and background independent. In other words, modularity can be restored in a holomorphic way by including nonperturbative effects. Notice that, according to (1.2),  $Z_\Sigma(\mu, \nu; \epsilon)$  transforms like a twisted fermion on the Riemann surface, with twists given by the characteristics  $\mu, \nu$ . Therefore, it seems to be the most natural object from the point of view of the free fermion theory on  $\Sigma$  advocated in [2, 18] and many other papers. As in CFT, one can also regard the  $Z_\Sigma(\mu, \nu; \epsilon)$  as chiral conformal blocks which can be used to construct invariants under suitable subgroups of the modular group.

We also show that the nonperturbative partition function is a tau function, in the sense that it satisfies a Hirota-type equation. This was observed in [25] in the case of genus zero spectral curve, where the nonperturbative instanton corrections are absent. Here we show that the inclusion of these corrections makes possible to generalize the construction of [25] to any spectral curve.

Since the construction of  $Z_\Sigma(\mu, \nu; \epsilon)$  only needs data coming from  $\Sigma$ , the nonperturbative partition function introduced in [22] can be also defined for topological strings on a variety of local Calabi–Yau manifolds, including mirrors of toric Calabi–Yau’s. In fact, it has been advocated in [41, 40] that the nonperturbative topological string partition function should include matrix model-like instanton effects, and in [40] it was pointed out that the nonperturbative partition function of [22] appears naturally in topological string models with large  $N$  Chern–Simons theory duals. Therefore, these nonperturbative partition functions provide holomorphic, modular and background independent partition functions for a wide class of topological string theory models.

The organization of this paper is as follows. In section 2 we recall the definition of the nonperturbative partition function introduced in [22], and we stress the fact that it

can be associated to any spectral curve. In section 3, which is the core of the paper, we show in detail that this partition function has good transformation properties under the modular group. In section 4 we review how the nonperturbative partition function appears in the context of matrix models, by summing over filling fractions. In section 5 we discuss the applications to topological string theory on local Calabi–Yau manifolds, and we propose that the nonperturbative partition function gives a natural nonperturbative object for topological strings. In section 6 we analyze the integrability properties and we show that  $Z_\Sigma(\mu, \nu; \epsilon)$  is a tau function. Finally, in section 7 we list some conclusions and avenues for further research.

## 2 The nonperturbative partition function

Consider an arbitrary “spectral curve”  $\Sigma = (\mathcal{C}, x, y)$ , i.e. the data of a compact Riemann surface  $\mathcal{C}$  of genus  $\bar{g}$ , together with two analytical functions  $x, y$  on some open domain of  $\mathcal{C}$ . Its symplectic invariants  $F_g$ ’s were defined in [25] (and we recall the definition in appendix A). They are such that (if  $g \geq 2$ ):

$$F_g(\mathcal{C}, x, \lambda y) = \lambda^{2-2g} F_g(\mathcal{C}, x, y), \quad (2.1)$$

and if two spectral curves have the same symplectic form  $d\tilde{x} \wedge d\tilde{y} = dx \wedge dy$ , we have

$$F_g(\mathcal{C}, \tilde{x}, \tilde{y}) = F_g(\mathcal{C}, x, y). \quad (2.2)$$

For the spectral curve  $\Sigma = (\mathcal{C}, x, y)$ , the *nonperturbative partition function* introduced in [22] in the context of matrix models, by summing over filling fractions (see section 4 for more details on the origin of this definition), is defined by

$$\begin{aligned} Z_\Sigma(\mu, \nu; \epsilon) &= e^{\sum_{g \geq 0} N^{2-2g} F_g(\epsilon)} \sum_k \sum_{l_i > 0} \sum_{h_i > 1 - \frac{l_i}{2}} \frac{N^{\sum_i (2-2h_i - l_i)}}{k! l_1! \dots l_k!} F_{h_1}^{(l_1)} \dots F_{h_k}^{(l_k)} \Theta_{\mu, \nu}^{(\sum_i l_i)}(NF'_0, \tau) \\ &= e^{N^2 F_0} e^{F_1} e^{(N^{-2} F_2 + N^{-4} F_3 + \dots)} \left\{ \Theta_{\mu, \nu} + \frac{1}{N} \left( \Theta'_{\mu, \nu} F'_1 + \frac{1}{6} \Theta'''_{\mu, \nu} F_0''' \right) \right. \\ &\quad \left. + \frac{1}{N^2} \left( \frac{1}{2} \Theta''_{\mu, \nu} F_1'' + \frac{1}{2} \Theta''_{\mu, \nu} F_1'^2 + \frac{1}{24} \Theta_{\mu, \nu}^{(4)} F_0'''' + \frac{1}{6} \Theta_{\mu, \nu}^{(4)} F_0''' F'_1 + \frac{1}{72} \Theta_{\nu, \mu}^{(6)} F_0''''^2 \right) + \dots \right\}. \end{aligned} \quad (2.3)$$

In this partition function, the  $F_g$ ’s are the symplectic invariants [25] of the spectral curve  $\Sigma$ , their derivatives are with respect to the background filling fraction  $\epsilon$  and computed at:

$$\epsilon = \frac{1}{2\pi i} \oint_{\mathcal{A}} y dx, \quad (2.4)$$

and the  $F_g(\mathcal{C}, x, y)$ 's and their derivatives depend on a choice of symplectic basis of  $2\bar{g}$  one-cycles  $\mathcal{A}_i, \mathcal{B}_j$  on  $\mathcal{C}$ . Finally, the theta function  $\Theta_{\mu, \nu}$  of characteristics  $(\mu, \nu)$  is defined by

$$\Theta_{\mu, \nu}(u, \tau) = \sum_{n \in \mathbb{Z}^{\bar{g}}} e^{(n+\mu-N\epsilon)u} e^{\pi i(n+\mu-N\epsilon)\tau(n+\mu-N\epsilon)} e^{2i\pi n\nu} \quad (2.5)$$

and is evaluated at

$$u = NF'_0, \quad F'_0 = \oint_{\mathcal{B}} y(x)dx, \quad \tau = \frac{1}{2\pi i} F''_0. \quad (2.6)$$

In (2.3), the derivatives of the theta function (2.5) are w.r.t.  $u$ , therefore each derivative introduces a factor of  $n + \mu - N\epsilon$  in the sum (2.5). The derivatives of  $\Theta$  and the derivatives of  $F_g$ , are written with tensorial notations. For instance,  $\frac{1}{6}\Theta_{\mu, \nu}^{(4)}F''_0F'_1$  actually means:

$$\frac{1}{6}\Theta_{\mu, \nu}^{(4)}F''_0F'_1 \equiv \frac{1}{2!3!1!} \sum_{i_1, i_2, i_3, i_4} \frac{\partial^4 \Theta_{\mu, \nu}}{\partial u_{i_1} \partial u_{i_2} \partial u_{i_3} \partial u_{i_4}} \frac{\partial^3 F_0}{\partial \epsilon_{i_1} \partial \epsilon_{i_2} \partial \epsilon_{i_3}} \frac{\partial F_1}{\partial \epsilon_{i_4}} \quad (2.7)$$

and the symmetry factor (here  $\frac{1}{6} = \frac{2}{2!3!1!}$ ) is the number of relabellings of the indices, giving the same pairings, and divided by the order of the group of relabellings, i.e.  $k!l_1! \dots l_k!$ , as usual in Feynmann graphs.

The  $\Theta$  function is closely related to the standard theta function, which is defined by

$$\vartheta_{[\nu]}^{[\mu]}(\xi|\tau) = \sum_{\mathbf{n} \in \mathbb{Z}^{\bar{g}}} \exp[i\pi(n + \mu)\tau(n + \mu) + 2\pi i(n + \mu)(\xi + \nu)]. \quad (2.8)$$

It is easy to see that these two functions are related as follows

$$\Theta_{\mu, \nu}(u, \tau) = \exp\left[-N^2\left(\epsilon F'_0 + \frac{1}{2}\epsilon^2 F''_0\right)\right] \vartheta_{[\nu]}^{[\mu]}(\xi|\tau) \quad (2.9)$$

where

$$\xi = \frac{N}{2\pi i} \oint_{\mathcal{B}-\tau\mathcal{A}} y(x)dx = N\left(\frac{F'_0}{2\pi i} - \tau\epsilon\right). \quad (2.10)$$

Finally, we point out that, as shown in [22], the  $1/N$  corrections in (2.3) can be resummed in terms of a single theta function, and  $Z_{\Sigma}(\mu, \nu; \epsilon)$  can be written as

$$Z_{\Sigma}(\mu, \nu; \epsilon) = e^{\sum_{g \geq 0} N^{2-2g} F_g(\epsilon)} \Theta_{\mu, \nu}(u^N, \tau^N), \quad (2.11)$$

where

$$u^N = NF'_0(1 + \mathcal{O}(1/N)), \quad \tau^N = \tau + \mathcal{O}(1/N), \quad (2.12)$$

and the corrections (which depend on the  $F_g$ ) can be easily determined order by order in  $1/N$ .

### 3 Modular properties of the nonperturbative partition function

In this section we discuss the transformation properties of the nonperturbative partition function (2.3) under the modular group. For that purpose, we shall first remind the transformation properties of the  $F_g$ 's and their derivatives, and then the transformation properties of the  $\Theta$ -function. For simplicity, we will first discuss the case in which  $\bar{g} = 1$  and the theta functions involved are of rank one. There are little changes when we go to the general case, but we will indicate these in section 3.3 below.

For spectral curves of  $\bar{g} = 1$ , a general modular transformation is an element of  $\text{SL}(2, \mathbb{Z})$ ,

$$\Gamma = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in \text{SL}(2, \mathbb{Z}), \quad (3.1)$$

where  $\alpha, \beta, \gamma, \delta \in \mathbb{Z}$  and  $\alpha\delta - \gamma\beta = 1$ . Under this transformation the integrals of  $ydx$  over the  $\mathcal{B}$  and the  $\mathcal{A}$  cycles transform as

$$\begin{pmatrix} F'_0 \\ 2\pi i\epsilon \end{pmatrix} \rightarrow \begin{pmatrix} \tilde{F}'_0 \\ 2\pi i\tilde{\epsilon} \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \begin{pmatrix} F'_0 \\ 2\pi i\epsilon \end{pmatrix}. \quad (3.2)$$

It follows that

$$\tilde{\tau} = \frac{\alpha\tau + \beta}{\gamma\tau + \delta}, \quad \tilde{\xi} = \frac{\xi}{\gamma\tau + \delta}, \quad (3.3)$$

so in particular  $\xi$  is a modular form of weight  $-1$ .

#### 3.1 Modular transformations of the perturbative amplitudes

The transformation properties of the genus zero free energy  $F_0$  can be derived from those of  $F'_0$  by integration, and they read

$$\tilde{F}_0 = F_0 + \frac{1}{2}\delta\beta(2\pi i\epsilon)^2 + \frac{1}{2}\gamma\alpha(F'_0)^2 + \beta\gamma(2\pi i\epsilon)F'_0. \quad (3.4)$$

This implies that the combination

$$\tilde{F}_0 - \frac{1}{2}\tilde{\epsilon}\tilde{F}'_0 = F_0 - \frac{1}{2}\epsilon F'_0 \quad (3.5)$$

stays invariant<sup>1</sup>.

---

<sup>1</sup>The invariance of (3.5) is a well-known fact in the context of Seiberg–Witten theory [44], where this combination turns out to be the modulus of the Seiberg–Witten curve. This invariance can also be seen from the explicit expression of  $F_0$  in [25].

Another transformation property which will be useful in the following is that

$$2\pi i \epsilon (F'_0 - 2\pi i \epsilon \tau) \rightarrow 2\pi i \epsilon (F'_0 - 2\pi i \epsilon \tau) - \kappa (F'_0 - 2\pi i \epsilon \tau)^2, \quad (3.6)$$

where

$$\kappa = -(\gamma\tau + \delta)^{-1} \gamma. \quad (3.7)$$

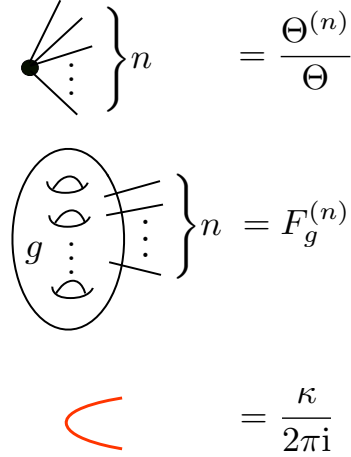


Figure 1: We will represent the different ingredients appearing in the modular transformation of the nonperturbative partition function by the graphic symbols depicted above.

The transformation properties for the  $F_g$ ,  $g \geq 1$  were studied in [1] from the point of view of topological string theory. [1] also found an elegant diagrammatic formalism to express these transformations. The formalism can be depicted as follows: after a modular transformation,  $F_g$  gets contributions corresponding to all possible stable degenerations of a genus  $g$  Riemann surface. Degenerations are obtained by pinching a non-trivial cycle. A stable degenerate surface is a nodal surface with marked points whose components have strictly negative Euler characteristics. Each degenerate cycle becomes a nodal point, and carries a factor  $(2i\pi)^{-1} \kappa$ . For example, the modular transformation of  $F_2$  gets contributions corresponding to pinching either 1, 2 or 3 cycles. This leads to the following transformation properties for the  $F_g$ :

$$\begin{aligned} \tilde{F}_1 &= F_1 - \frac{1}{2} \ln(\gamma\tau + \delta), \\ \tilde{F}_2 &= F_2 + \frac{\kappa}{2\pi i} \frac{1}{2} (F_1'' + F_1'^2) + \frac{1}{8} \left( \frac{\kappa}{2\pi i} \right)^2 (F_0'''' + 4F_0''' F_1') + \frac{1}{48} \left( \frac{\kappa}{2\pi i} \right)^3 (6F_0'^3 + 4F_0'^3) \end{aligned} \quad (3.8)$$



and so on. For  $F_2$ , in the higher rank  $\bar{g} > 1$  case, the last two terms have a different index structure, as it can be seen from the graphical representation, and this is why we have written them in separate form, anticipating our general analysis. In general for  $g \geq 2$ ,  $\tilde{F}_g$  is a polynomial in  $\kappa$  of degree  $3g - 3$ . This method can be extended (see for example [25]) to include all derivatives of the  $F_g$ , and the derivatives correspond simply to insertions of marked points. One should also note that the transformation of  $F_g^{(n)}$  includes an overall factor of

$$(\gamma\tau + \delta)^{-n}. \quad (3.9)$$

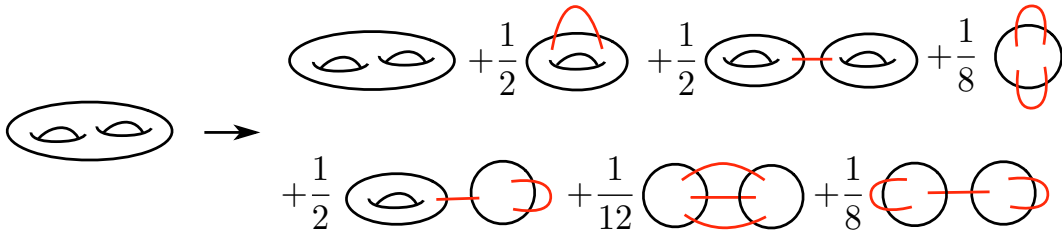


Figure 2: A graphical depiction of the modular transformation of  $F_2$ .

In the following, we will rely very heavily on diagrammatic methods, and we will use the graphical representation shown in Fig. 1 for the different ingredients appearing in the calculations. In this diagrammatic language, the modular transformation of  $F_2$  can be represented as in Fig. 2, while in Fig. 3 we show the transformation properties of  $F'_1$ . The numerical factors in front of each diagram are the symmetry factors of the corresponding diagram, i.e. the number of ways of pinching giving the same diagram, and divided by the order of the automorphism group. This is the usual symmetry factor of Feynman graphs.

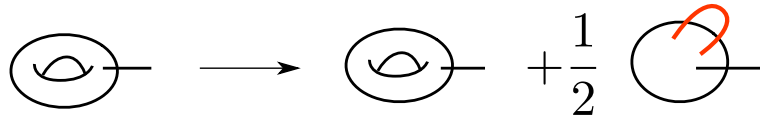


Figure 3: A graphical depiction of the modular transformation of  $F'_1$ .

In the context of matrix models, the transformation properties of the  $F_g$  can be derived from the formalism of [25]. The basic ingredient in this formalism is the transformation property of the Bergmann kernel under a modular transformation,

$$\tilde{B}(z_1, z_2) = B(z_1, z_2) + 2i\pi\omega(z_1)\kappa\omega(z_2), \quad (3.10)$$

where  $\omega(z)$  is a basis of Abelian holomorphic differentials on  $\Sigma$ . The  $F_g$ 's of [25] are made of residues of products of Bergmann kernels (see appendix A), namely  $F_g^{(n)}$  is a sum of residues of products of  $3g - 3 + n$  Bergmann kernels, and multiplied by terms independent of a choice of cycles. Each term of  $F_g^{(n)}$  is thus represented in [25] by a trivalent diagram, with  $3g - 3 + n$  edges. A modular transformation amounts to cutting (or not cutting) edges of each diagram in all possible ways (such that all subdiagrams have at least one vertex), and each cut edge is replaced by a factor of  $\kappa$ . We illustrate these rules in Fig. 4. One has for example

$$\tilde{F}_0^{(4)} = F_0^{(4)} + 3\frac{\kappa}{2i\pi}(F_0''')^2, \quad (3.11)$$

Thus, the modular transformation of  $F_g^{(n)}$  can be written as a polynomial of degree  $3g - 3 + n$  of  $\kappa$ . One can show [24], using this property, that one obtains the same diagrammatic calculus as in [1].

### 3.2 Transformation properties of the theta function

We now study the transformation properties of the  $\Theta$ -function, and first, we remind the transformation properties of the standard theta function,

$$\tilde{\vartheta} \left[ \begin{smallmatrix} \tilde{\mu} \\ \tilde{\nu} \end{smallmatrix} \right] (\tilde{\xi} | \tilde{\tau}) = \zeta \left[ \begin{smallmatrix} \mu \\ \nu \end{smallmatrix} \right] (\Gamma) (\gamma\tau + \delta)^{\frac{1}{2}} \exp \left[ -\pi i \kappa \xi^2 \right] \vartheta \left[ \begin{smallmatrix} \mu \\ \nu \end{smallmatrix} \right] (\xi | \tau), \quad (3.12)$$

where

$$\begin{aligned} \tilde{\mu} &= \delta\mu - \gamma\nu + \frac{1}{2}\gamma\delta, \\ \tilde{\nu} &= -\beta\mu + \alpha\nu + \frac{1}{2}\alpha\beta, \end{aligned} \quad (3.13)$$

and

$$\zeta \left[ \begin{smallmatrix} \mu \\ \nu \end{smallmatrix} \right] (\Gamma) = \exp \left\{ -\pi i \left( \delta\beta\mu^2 + \gamma\alpha\nu^2 - 2\beta\gamma\mu\nu + \alpha\beta(\delta\mu - \gamma\nu) \right) \right\} \zeta \left[ \begin{smallmatrix} 0 \\ 0 \end{smallmatrix} \right] (\Gamma). \quad (3.14)$$

Here,  $\zeta \left[ \begin{smallmatrix} 0 \\ 0 \end{smallmatrix} \right] (\Gamma)$  is a root of unity which does not depend on the characteristics.

We now deduce an important transformation property of the derivatives of the theta function which will be useful in the following. The derivatives appearing in (2.3) can be computed in terms of the standard theta function by considering

$$\frac{1}{(2\pi i)^\ell} \left( \partial_\xi - N2\pi i \epsilon \right)^\ell \vartheta \left[ \begin{smallmatrix} \mu \\ \nu \end{smallmatrix} \right] (\xi | \tau), \quad (3.15)$$

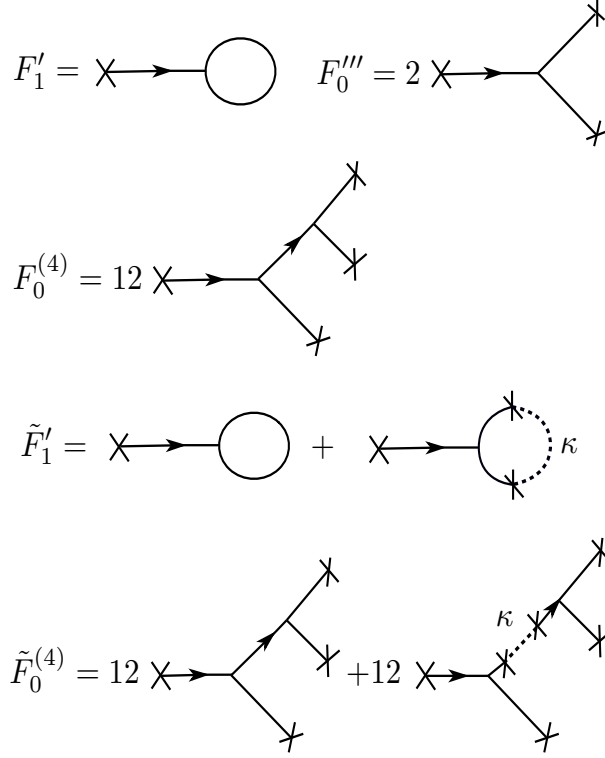


Figure 4: A graphical depiction of the modular transformation of  $F'_1$  and  $F_0^{(4)}$  using the diagrammatic representation of [21, 25]. Each arrowed edge means a propagator  $K$  in [25], and each non-arrowed edge means a Bergmann kernel. Each cross at the end of edges, means that we take the  $\mathcal{B}_i$  cycle integral corresponding to  $\partial/\partial\epsilon_i$ . For example  $\partial F_1/\partial\epsilon_i = \oint_{z \in \mathcal{B}_i} \sum_j \text{Res}_{z' \rightarrow a_j} K(z, z') B(z', \bar{z}')$ . Modular transformations amount to cutting edges in all possible ways such that each subdiagram contains at least one vertex.

and we are interested on the transformation properties of such quantities under a modular transformation. In evaluating these quantities, the variable  $\xi$  is regarded as an independent variable, unrelated to  $\epsilon$  and transforming as a modular form of weight  $-1$ . Only at the very end we set it equal to its true value. Under a modular transformation,

$$\left(\partial_{\xi} - N2\pi i\epsilon\right)^{\ell} \tilde{\vartheta}_{[\nu]}^{[\tilde{\mu}]}(\tilde{\xi}|\tilde{\tau}) = \mathcal{N}\left(\partial_{\xi} - N2\pi i\epsilon\right)^{\ell} \exp\left[-\pi i\kappa\xi^2\right] \vartheta_{[\nu]}^{[\mu]}(\xi|\tau) \quad (3.16)$$

where  $\mathcal{N}$  does not depend on  $\xi$ . We now evaluate

$$A_\ell(f) = \exp\left[\pi i \kappa \xi^2\right] \left(\partial_{\tilde{\xi}} - N 2\pi i \tilde{\epsilon}\right)^\ell \exp\left[-\pi i \kappa \xi^2\right] f(\xi) \quad (3.17)$$

where  $f(\xi)$  is an arbitrary function of  $\xi$ . We find,

$$A_\ell(f) = \left(\partial_{\tilde{\xi}} + 2\pi i \gamma \xi - N 2\pi i \tilde{\epsilon}\right)^\ell f(\xi) \quad (3.18)$$

The operators appearing in the r.h.s. do not commute, and it is easier to consider a generating functional

$$S(x) = \sum_{\ell=0}^{\infty} \frac{A_\ell(f)}{\ell!} x^\ell = e^{x(\partial_{\tilde{\xi}} + 2\pi i \gamma \xi - N 2\pi i \tilde{\epsilon})} f(\xi). \quad (3.19)$$

Using the Baker–Campbell–Hausdorff formula, we get

$$S(x) = e^{x(2\pi i \gamma \xi - N 2\pi i \tilde{\epsilon})} e^{-x^2 \pi i \kappa} e^{x \partial_{\tilde{\xi}}} f(\xi). \quad (3.20)$$

After extracting the  $\ell$ -th power of  $x$  in this generating functional, we can already set  $\xi$  to its value (2.10). Since

$$2\pi i \gamma \xi - N 2\pi i \tilde{\epsilon} = -(\gamma \tau + \delta) N 2\pi i \epsilon \quad (3.21)$$

we finally obtain

$$A_\ell(f) = (c\tau + \delta)^\ell \sum_{j=0}^{[\ell/2]} (2j-1)!! \binom{\ell}{2j} (-2\pi i \kappa)^j (\partial_{\tilde{\xi}} - N 2\pi i \epsilon)^{\ell-2j} f(\xi), \quad (3.22)$$

and this implies the following transformation law for the derivatives of the theta function,

$$\begin{aligned} & \frac{1}{\tilde{\vartheta}\left[\begin{smallmatrix} \tilde{\mu} \\ \tilde{\nu} \end{smallmatrix}\right]}(\tilde{\xi}|\tilde{\tau}) \left(\frac{1}{2\pi i} \partial_{\tilde{\xi}} - N \epsilon\right)^\ell \tilde{\vartheta}\left[\begin{smallmatrix} \tilde{\mu} \\ \tilde{\nu} \end{smallmatrix}\right]}(\tilde{\xi}|\tilde{\tau}) = \\ & \frac{1}{\vartheta\left[\begin{smallmatrix} \mu \\ \nu \end{smallmatrix}\right]}(\xi|\tau) (\gamma \tau + \delta)^\ell \sum_{j=0}^{[\ell/2]} (2j-1)!! \binom{\ell}{2j} \left(-\frac{\kappa}{2\pi i}\right)^j \left(\frac{1}{2\pi i} \partial_{\xi} - N \epsilon\right)^{\ell-2j} \vartheta\left[\begin{smallmatrix} \mu \\ \nu \end{smallmatrix}\right]}(\xi|\tau). \end{aligned} \quad (3.23)$$

### 3.3 Generalization to higher rank

In the higher rank case, the modular group is  $\text{Sp}(2\bar{g}, \mathbb{Z})$ . A modular transformation  $\Gamma$  satisfies

$$\Gamma^T \Omega \Gamma = \Omega, \quad \Omega = \begin{pmatrix} 0 & \mathbf{1} \\ -\mathbf{1} & 0 \end{pmatrix}, \quad (3.24)$$

and it can be written as:

$$\Gamma = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \quad (3.25)$$

where the  $\bar{g} \times \bar{g}$  matrices  $A, B, C, D$ , with integer-valued entries, satisfy

$$A^T D - C^T B = \mathbf{1}_{\bar{g}}, \quad A^T C = C^T A, \quad B^T D = D^T B. \quad (3.26)$$

All previous quantities are promoted to vectors and matrices, with obvious generalizations. For example,

$$\begin{aligned} \tilde{F}'_{0,j} &= A_j^k F'_{0,k} + B_{jk} 2\pi i \epsilon^k, \\ 2\pi i \tilde{\epsilon}^j &= C^{jk} F'_{0,k} + D^j_k 2\pi i \epsilon^k, \end{aligned} \quad (3.27)$$

where summation over repeated indices is understood, and

$$\tilde{\tau} = (A\tau + B)(C\tau + D)^{-1}, \quad \tilde{\xi}_i = \left[ (C\tau + D)^{-1} \right]_i^j \xi_j, \quad i = 1, \dots, \bar{g}. \quad (3.28)$$

The genus zero free energy transforms now as

$$\tilde{F}_0 = F_0 + \frac{1}{2} 2\pi i \epsilon^k (D^T B)_{kj} 2\pi i \epsilon^j + \frac{1}{2} F'_{0,k} (C^T A)^{kj} F'_{0,j} + 2\pi i \epsilon^k (B^T C)_k^j F'_{0,j}. \quad (3.29)$$

The quantity

$$F_0 - \frac{1}{2} \epsilon F'_0 \equiv F_0 - \frac{1}{2} \epsilon^k F'_{0,k} \quad (3.30)$$

is still invariant, as it can be easily checked. In the following, for general  $\bar{g}$ , the expressions we used for  $\bar{g} = 1$  will denote the obvious contraction of indices, as in (3.30) above. We also have the obvious generalization of (3.6),

$$2\pi i \epsilon (F'_0 - 2\pi i \epsilon \tau) \rightarrow 2\pi i \epsilon (F'_0 - 2\pi i \epsilon \tau) - (F'_0 - 2\pi i \tau \epsilon) \kappa (F'_0 - 2\pi i \tau \epsilon), \quad (3.31)$$

where

$$\kappa^{ij} = - \left[ (C\tau + D)^{-1} C \right]^{ij}. \quad (3.32)$$

This matrix is *symmetric*, as a consequence of the symplectic properties of  $\Gamma$ .

The transformation properties of the  $F_g$  are given by the diagrammatic method explained above. We note that, for  $\bar{g} > 1$ ,

$$\tilde{F}_1 = F_1 - \frac{1}{2} \log \det(C\tau + D). \quad (3.33)$$

Let us now consider the transformation properties of the theta function and its derivatives, in the higher genus case. We have (see for example [5])

$$\tilde{\vartheta}_{[\tilde{\nu}]}^{[\tilde{\mu}]}(\tilde{\xi}|\tilde{\tau}) = \zeta_{[\nu]}^{[\mu]}(\Gamma) \left( \det(C\tau + D) \right)^{\frac{1}{2}} \exp \left[ -\pi i \xi \kappa \xi \right] \vartheta_{[\nu]}^{[\mu]}(\xi|\tau). \quad (3.34)$$

In this transformation, the characteristics are given by

$$\begin{aligned} \tilde{\mu} &= D\mu - C\nu + \frac{1}{2}(CD^T)_d, \\ \tilde{\nu} &= -B\mu + A\nu + \frac{1}{2}(AB^T)_d. \end{aligned} \quad (3.35)$$

where  $(M)_d$  denotes a column vector whose entries are the diagonal entries of the matrix  $M$ . The phase is now given by

$$\zeta_{[\nu]}^{[\mu]}(\Gamma) = \exp \left\{ -\pi i \left( \mu D^T B \mu + \nu C^T A \nu - 2\mu B^T C \nu + (\mu D^T - \nu C^T)(AB^T)_d \right) \right\} \zeta_{[0]}^{[0]}(\Gamma). \quad (3.36)$$

We now consider the derivatives of the theta function. We want to generalize (3.23) and to compute the modular transformation of

$$\left( \partial^{j_1} - N2\pi i \epsilon^{j_1} \right) \cdots \left( \partial^{j_\ell} - N2\pi i \epsilon^{j_\ell} \right) \vartheta_{[\nu]}^{[\mu]}(\xi|\tau), \quad (3.37)$$

where

$$\partial^j = \frac{\partial}{\partial \xi_j}. \quad (3.38)$$

Under a modular transformation,

$$\begin{aligned} &\left( \tilde{\partial}^{j_1} - N2\pi i \tilde{\epsilon}^{j_1} \right) \cdots \left( \tilde{\partial}^{j_\ell} - N2\pi i \tilde{\epsilon}^{j_\ell} \right) \vartheta_{[\tilde{\nu}]}^{[\tilde{\mu}]}(\tilde{\xi}|\tilde{\tau}) = \\ &\mathcal{N} \left( \tilde{\partial}^{j_1} - N2\pi i \tilde{\epsilon}^{j_1} \right) \cdots \left( \tilde{\partial}^{j_\ell} - N2\pi i \tilde{\epsilon}^{j_\ell} \right) \exp \left[ -\pi i \xi \kappa \xi \right] \vartheta_{[\nu]}^{[\mu]}(\xi|\tau) \end{aligned} \quad (3.39)$$

where  $\mathcal{N}$  is again independent of  $\xi$ . Since

$$\left( \tilde{\partial}^j - N2\pi i \tilde{\epsilon}^j \right) \exp \left[ -\pi i \xi \kappa \xi \right] = \exp \left[ -\pi i \xi \kappa \xi \right] \left( \tilde{\partial}^j + 2\pi i C^{jk} \xi_k - N2\pi i \tilde{\epsilon}^j \right) \quad (3.40)$$

we just have to compute

$$A^{j_1} \cdots A^{j_\ell} \vartheta_{[\nu]}^{[\mu]}(\xi|\tau). \quad (3.41)$$

where the operator  $A^j$  is given by

$$A^j = (C\tau + D)_k^j \partial^k + 2\pi i C^{jk} \xi_k - N2\pi i \tilde{\epsilon}^j. \quad (3.42)$$

As it happened in the  $\bar{g} = 1$  case, the operators  $A^{j_1}, \dots, A^{j_\ell}$  do not commute, since we cannot set  $\xi_k$  to its value until we have not commuted all the derivatives to the right. Of course, this is precisely the type of computation that Wick's theorem does. In this context, we define a normal-ordered operator as an operator in which we set

$$\xi_k = N \left( \frac{F'_{0,k}}{2\pi i} - \tau_{kl} \epsilon^l \right), \quad (3.43)$$

i.e. we have

$$: A^{j_1} \dots A^{j_\ell} := (C\tau + D)_{k_1}^{j_1} \dots (C\tau + D)_{k_\ell}^{j_\ell} (\partial^{k_1} - N2\pi i \epsilon^{k_1}) \dots (\partial^{k_\ell} - N2\pi i \epsilon^{k_\ell}). \quad (3.44)$$

The contraction is given by

$$\langle A^j A^k \rangle = 2\pi i C^{jk} = (C\tau + D)_l^j (-2\pi i \kappa)^{lk}. \quad (3.45)$$

We can now apply Wick's theorem to write the operator in (3.41) as a sum of normal-ordered operators. In the end we obtain

$$\begin{aligned} & \frac{1}{\tilde{\vartheta}[\tilde{\mu}]{\tilde{\nu}}(\tilde{\xi}|\tilde{\tau})} \left( \frac{1}{2\pi i} \tilde{\partial}^{p_1} - N\tilde{\epsilon}^{p_1} \right) \dots \left( \frac{1}{2\pi i} \tilde{\partial}^{p_\ell} - N\tilde{\epsilon}^{p_\ell} \right) \vartheta[\tilde{\mu}]{\tilde{\nu}}(\tilde{\xi}|\tilde{\tau}) = \\ & \frac{1}{\vartheta[\mu]{\nu}(\xi|\tau)} (C\tau + D)_{q_1}^{p_1} \dots (C\tau + D)_{q_\ell}^{p_\ell} \sum_{j=0}^{[\ell/2]} \sum_{\sigma} \left( -\frac{\kappa}{2\pi i} \right)^{q_{\sigma(1)} q_{\sigma(2)}} \dots \left( -\frac{\kappa}{2\pi i} \right)^{q_{\sigma(2j-1)} q_{\sigma(2j)}} \quad (3.46) \\ & \left( \frac{1}{2\pi i} \partial^{q_{\sigma(2j+1)}} - N\epsilon^{q_{\sigma(2j+1)}} \right) \dots \left( \frac{1}{2\pi i} \partial^{q_{\sigma(\ell)}} - N\epsilon^{q_{\sigma(\ell)}} \right) \vartheta[\mu]{\nu}(\xi|\tau). \end{aligned}$$

In the r.h.s., the first sum is over the number  $j$  of contractions. The second sum is over all possible ways of performing the contractions, and  $\sigma$  denotes the permutation of the indices  $1, \dots, \ell$  associated to a given contraction (note that not all possible permutations appear). (3.46) is the generalization of (3.23) to the multi-index case. Indeed, the combinatorial factor in the r.h.s. of (3.23) is simply the number of possible ways of performing  $j$  contractions in  $\ell$  terms: the combinatorial number accounts for the choices of  $2j$  legs to be contracted among  $\ell$  legs in total, and  $(2j-1)!!$  is the number of possible pairings of the  $2j$  legs.

Of course, the easiest way of keeping track of combinatorial formulae like the above is by means of a graphical representation. Using the building blocks shown in Fig. 1, the equation (3.46) might be represented as in Fig. 5, where some simple examples are also shown.

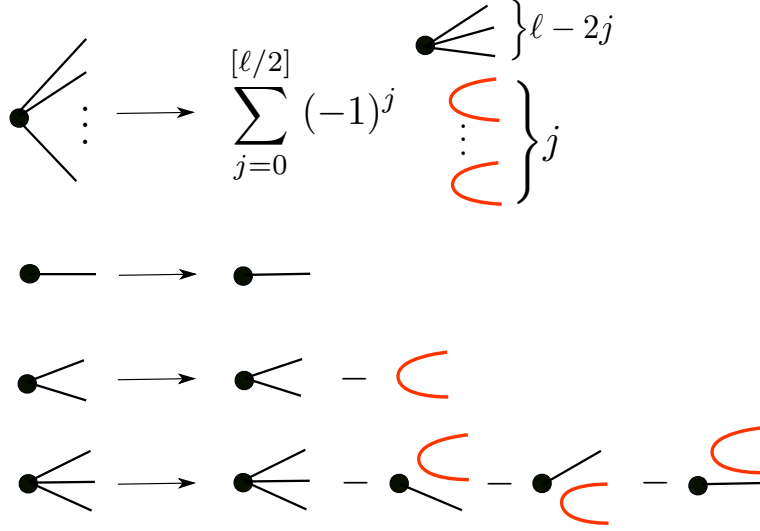


Figure 5: The first line gives a graphical representation of (3.46), while the next lines exemplify it for  $\ell = 1, 2, 3$ . Notice that the edges are labeled, and when summing over all possible contractions we have to take this labeling into account, as shown in the last example.

### 3.4 Transformation properties of the nonperturbative partition function

We will now study in detail the transformation properties of the nonperturbative partition function (2.3). We rewrite (2.3) as:

$$\begin{aligned}
Z_{\Sigma}(\mu, \nu; \epsilon) &= e^{N^2 F_0 + F_1} \Theta_{\mu, \nu} \left\{ 1 + \sum_{j=1}^{\infty} N^{-j} Z_j(\mu, \nu; \epsilon) \right\} \\
&= e^{N^2 F_0 + F_1} \Theta_{\mu, \nu} \left\{ 1 + \frac{1}{N} \left( \frac{\Theta'_{\mu, \nu}}{\Theta_{\mu, \nu}} F'_1 + \frac{1}{6} \frac{\Theta'''_{\mu, \nu}}{\Theta_{\mu, \nu}} F_0''' \right) \right. \\
&\quad + \frac{1}{N^2} \left( F_2 + \frac{1}{2} \frac{\Theta''_{\mu, \nu}}{\Theta_{\mu, \nu}} F_1'' + \frac{1}{2} \frac{\Theta''_{\mu, \nu}}{\Theta_{\mu, \nu}} F_1'^2 + \frac{1}{24} \frac{\Theta_{\mu, \nu}^{(4)}}{\Theta_{\mu, \nu}} F_0'''' + \frac{1}{6} \frac{\Theta_{\mu, \nu}^{(4)}}{\Theta_{\mu, \nu}} F_0''' F_1' + \frac{1}{72} \frac{\Theta_{\mu, \nu}^{(6)}}{\Theta_{\mu, \nu}} (F_0''')^2 \right) \\
&\quad \left. + \dots \right\}
\end{aligned} \tag{3.47}$$

i.e.  $Z_j$  is the sum of all terms contributing to order  $N^{-j}$ .

We first study the transformation properties of the leading term  $e^{N^2 F_0 + F_1} \Theta_{\mu, \nu}$ . To do



this, we use the relation (2.9) to write

$$e^{N^2 F_0} \Theta_{\mu, \nu} = \exp \left[ N^2 \left( F_0 - \epsilon F'_0 - \frac{1}{2} \epsilon^2 F''_0 \right) \right] \vartheta_{[\nu]}^{[\mu]}(\xi | \tau). \quad (3.48)$$

In order to study the transformation properties of this quantity, we have to study the transformation properties of

$$\Xi = F_0 - \epsilon F'_0 - \frac{1}{2} \epsilon^2 F''_0 = F_0 - \frac{1}{2} \epsilon F'_0 - \frac{1}{2} \epsilon \left( F'_0 - 2\pi i \epsilon \tau \right). \quad (3.49)$$

Using (3.31) and the invariance of (3.30), one finds that

$$\tilde{\Xi} = \Xi + \pi i \left( \frac{F'_0}{2\pi i} - \tau \epsilon \right) \kappa \left( \frac{F'_0}{2\pi i} - \tau \epsilon \right), \quad (3.50)$$

therefore the shift in  $\Xi$  exactly compensates the  $\xi$  dependent exponent in the transformation (3.34) and we find,

$$e^{N^2 \tilde{F}_0} \tilde{\Theta}_{\tilde{\mu}, \tilde{\nu}}(\tilde{u}, \tilde{\tau}) = \zeta_{[\nu]}^{[\mu]}(\Gamma) (\det(C\tau + D))^{\frac{1}{2}} e^{N^2 F_0} \Theta_{\mu, \nu}(u, \tau). \quad (3.51)$$

It follows immediately from (3.33) that

$$e^{N^2 \tilde{F}_0 + \tilde{F}_1} \tilde{\Theta}_{\tilde{\mu}, \tilde{\nu}}(\tilde{u}, \tilde{\tau}) = \zeta_{[\nu]}^{[\mu]}(\Gamma) e^{N^2 F_0 + F_1} \Theta_{\mu, \nu}(u, \tau) \quad (3.52)$$

under a general  $\text{Sp}(2\bar{g}, \mathbb{Z})$  transformation. This proves (1.2) to leading order in  $N$ .

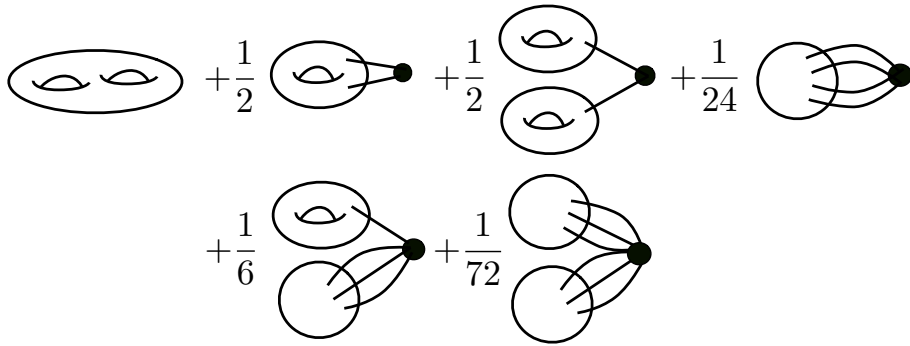


Figure 6: Graphic representation of  $Z_2$ , the term in  $1/N^2$  in the expansion of  $Z$ .

Our goal is now to prove that each  $Z_j$  is itself modular invariant, up to the change of characteristics (3.35).

$Z_j$  is given by a sum, and each term in this sum is the product of a derivative of  $\Theta$ , with a product of a certain number of derivatives of  $F_h$ 's corresponding to the stable degeneracy of (possibly non-connected) surfaces with Euler characteristics  $\chi = -j$ . Notice that, in each term, the total number of derivatives acting on the  $F_h$  equals the number of derivatives acting on  $\Theta$ . As shown in Fig. 1,  $\Theta^{(n)}/\Theta$ , the  $n^{\text{th}}$  derivative of  $\Theta$ , is depicted as a black dot with  $n$  legs, while  $F_h^{(n)}$  is depicted as a Riemann surface of genus  $h$  with  $n$  marked points. Each term in  $Z_j$  can thus be depicted as a set of stable Riemann surfaces with marked points linked to the legs of the  $\Theta^{(n)}/\Theta$  term. Notice that there can be disconnected terms corresponding to  $F_h^{(n)}$  with  $n = 0$ . Also, each term has a symmetry factor, which is, as usual in Feynmann graphs, the number of possible pairings of indices, divided by the group order  $k! l_1! \dots l_k!$ . As an example, we depict  $Z_2$  in Fig. 6.

Figure 7: Graphic proof of the modular invariance of the term  $Z_1$  of order  $1/N$  in the expansion of  $Z$ .

Under a modular transformation, each  $F_h^{(n)}$  transforms into a sum of stable degenerated surfaces, with a factor  $\kappa/(2\pi i)$  for each degenerated cycle. As shown in (3.46) (and as illustrated in Fig. 5) the  $\Theta^{(n)}/\Theta$  term transforms into a sum of  $\Theta^{(k)}/\Theta$  where  $(n - k)/2$  legs are replaced by a  $-\kappa/(2\pi i)$  in all possible ways (with a change of characteristics as (3.35)). The factors of  $C\tau + D$  appearing in (3.46) cancel against the factors of  $(C\tau + D)^{-1}$  coming from the derivatives of the  $F_h$ 's.

In the end,  $\tilde{Z}_j$  is a sum of terms, each of them is the product of a  $\Theta^{(k)}/\Theta$ , with a certain number of  $F_{h_i}^{(n_i)}$ 's, and a certain number of  $\kappa$  propagators. Each propagator can be obtained in two ways, either from a degeneracy of a  $F_h^{(n)}$ , with a factor  $+(2i\pi)^{-1}$ , or from a term  $\Theta^{(k)}/\Theta$ , with a factor  $-(2i\pi)^{-1}$ . Also, the symmetry factors are counting the number of possible pairings corresponding to a given diagramm, and are the same whether obtained

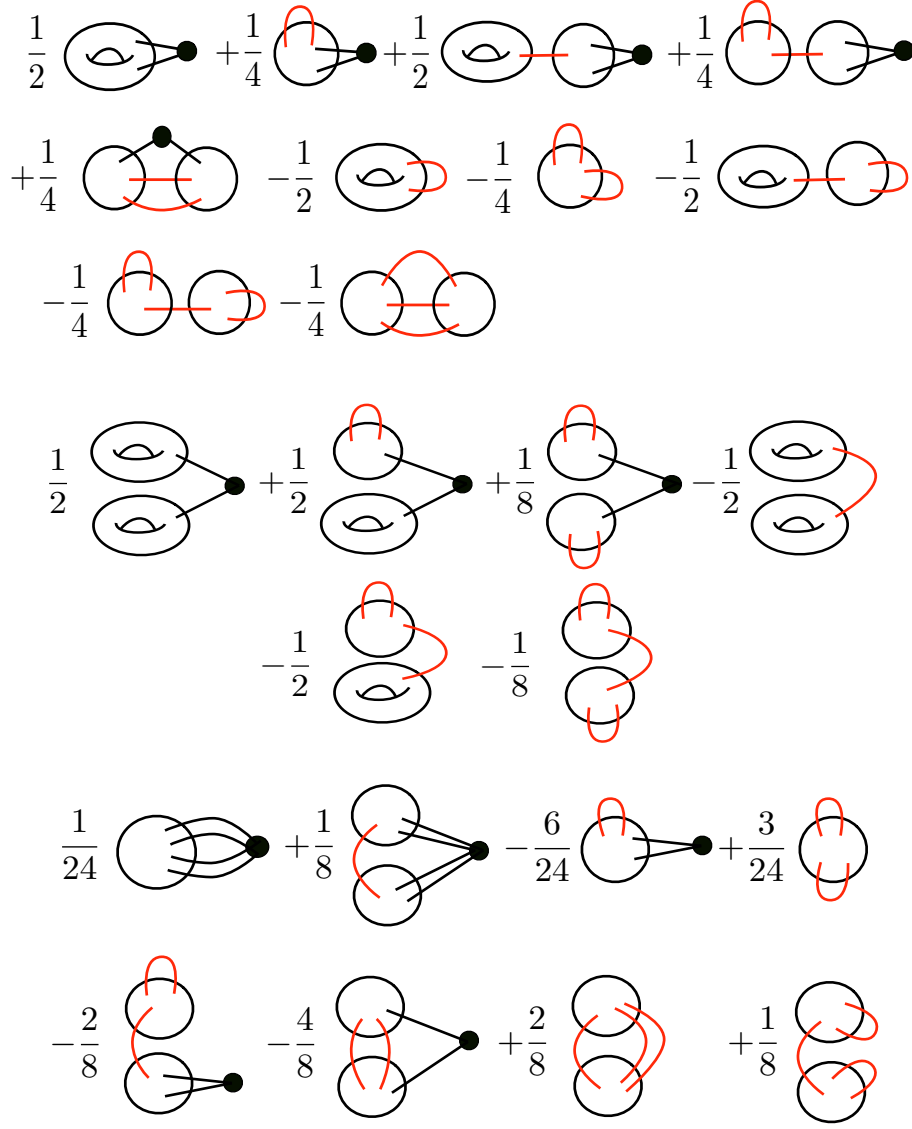


Figure 8: Modular transformation of the second, third and fourth graphs in Fig. 6 .

from the modular transformation of  $\Theta^{(k)}/\Theta$ , or from the modular transformation of  $F_{h_i}^{(n_i)'}_S$ . Therefore, the total contribution of each  $\kappa$  propagator is 0. This shows that

$$\tilde{Z}_j(\tilde{\mu}, \tilde{\nu}) = Z_j(\mu, \nu) \quad (3.53)$$

therefore we have proved that

$$\tilde{Z}_\Sigma(\tilde{\mu}, \tilde{\nu}; \tilde{\epsilon}) = \zeta \left[ \begin{smallmatrix} \mu \\ \nu \end{smallmatrix} \right] (\Gamma) Z_\Sigma(\mu, \nu; \epsilon) \quad (3.54)$$

We illustrate the proof by the examples of  $Z_1$  and  $Z_2$ . In the case of  $Z_1$ , we know from Fig. 3 the transformation rule of  $F'_1$ , and from Fig. 5 we know the transformation properties of  $\Theta'/\Theta$  and  $\Theta'''/\Theta$ . Since  $F_0'''$  transforms as a modular form of weight  $-3$  (i.e., there are no shifts involved) and it is completely symmetric in its indices, we immediately obtain the graphic proof of modular invariance of  $Z_1$  depicted in Fig. 7.

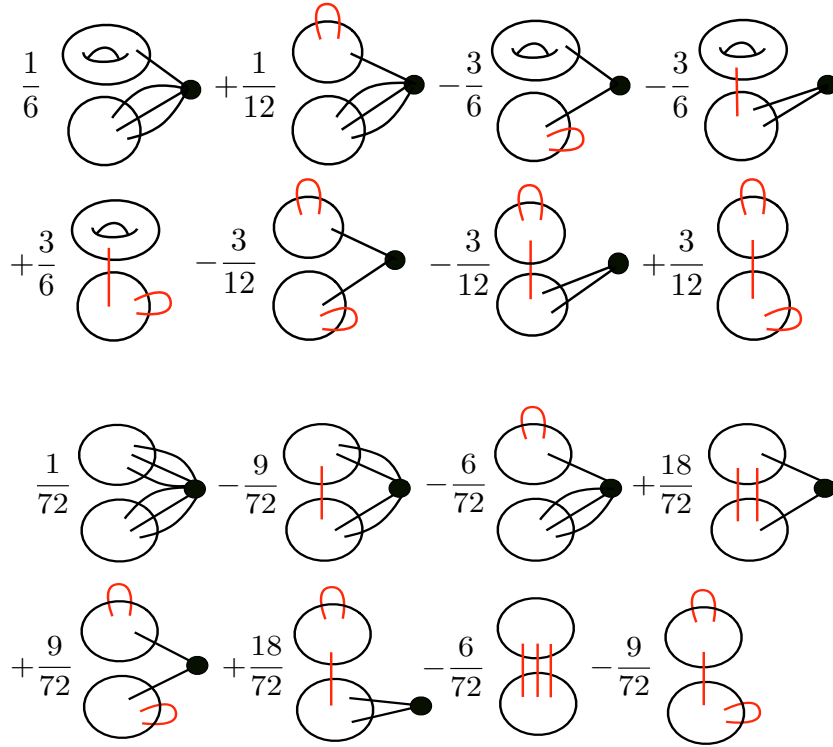


Figure 9: Modular transformation of the fifth and sixth graphs in Fig. 6 .

The terms appearing in  $Z_2$  were depicted graphically in Fig. 6. The modular transformation of  $F_2$ , which is the first term, was shown in Fig. 2. The modular transformation of the rest of the graphs in Fig. 6 is shown in Fig. 8 and Fig. 9. One can check that all the graphs generated on top of the original ones (i.e., all the shifts generated by the modular transformation, i.e. all graphs containing propagators  $\kappa$ ) eventually cancel, and  $Z_2$  remains

invariant up to the change of characteristics. For example, the second graph in the first line of Fig. 8 cancels against the third graph in the sixth line of the same figure, while the third graph in the first line of Fig. 8 cancels against the fourth graph in the first line of Fig. 9. In a similar manner, all graphs in Fig. 2, Fig. 8 and Fig. 9 containing propagators eventually cancel.

### 3.5 Background independence

We now show that the non-perturbative partition function is background independent, i.e. it does not depend on the background filling fraction  $\epsilon$ . This is not surprising, because this is how it was first constructed in [22].

Let

$$\hat{Z}(\eta) = e^{\sum_{g \geq 0} N^{2-2g} F_g(\eta)} \quad (3.55)$$

be the perturbative partition function.

We may Taylor expand it near an arbitrary background filling fraction  $\epsilon$ :

$$\begin{aligned} \hat{Z}(\eta) &= \sum_j \frac{1}{j!} (\eta - \epsilon)^j \hat{Z}^{(j)}(\epsilon) \\ &= e^{N^2(F_0 + (\eta - \epsilon)F'_0 + \frac{1}{2}(\eta - \epsilon)^2 F''_0)} e^{F_1} e^{\sum_{g \geq 2} N^{2-2g} F_g} \\ &\quad \cdot \sum_k \sum_{l_i > 0} \sum_{h_i > 1 - \frac{l_i}{2}} \frac{N^{\sum_i (2-2h_i)}}{k! l_1! \dots l_k!} F_{h_1}^{(l_1)} \dots F_{h_k}^{(l_k)} (\eta - \epsilon)^{\sum_i l_i} \end{aligned} \quad (3.56)$$

Therefore we recognize that the nonperturbative partition function (2.3), is the Taylor expansion of:

$$Z_\Sigma(\mu, \nu; \epsilon) = \sum_{\mathbf{n} \in \mathbb{Z}^g} \hat{Z}\left(\frac{n + \mu}{N}\right) e^{2i\pi n\nu} \quad (3.57)$$

The right hand side of this expression is clearly independent of  $\epsilon$ :

$$Z_\Sigma(\mu, \nu; \epsilon) = Z_\Sigma(\mu, \nu; \epsilon') \quad (3.58)$$

and thus  $Z_\Sigma(\mu, \nu; \epsilon) = Z_\Sigma(\mu, \nu)$  is background independent.

## 4 Matrix models and sums over filling fractions

Let us briefly summarize [22], which is just an improved version of [11], to explain the origin of the definition (2.3) for the non-perturbative partition function.

## 4.1 Matrix integrals and domains of integration

A matrix integral is typically an integral of the form:

$$Z = \frac{1}{N!} \int dx_1 \dots dx_N \prod_{i>j} (x_i - x_j)^2 \prod_i e^{-NV(x_i)} \quad (4.59)$$

where we have not written the integration domain. For any choice of integration domain,  $Z$  satisfies the same loop equations (i.e. Schwinger-Dyson equations), as long as there is no boundary term when one integrates by parts.

For example, one may consider the matrix integral over  $H_N(\gamma)$  (the set of normal matrices with eigenvalues constrained on a path  $\gamma$ ):

$$Z(\gamma) = \int_{H_N(\gamma)} dM e^{-N \text{Tr} V(M)} = \frac{1}{N!} \int_{\gamma^N} dx_1 \dots dx_N \prod_{i>j} (x_i - x_j)^2 \prod_i e^{-NV(x_i)} \quad (4.60)$$

This integral depends only on a choice of path  $\gamma$ , and it has a  $U(N)$  invariance.

One may also consider the following integral where  $N = \sum_i n_i$ :

$$\hat{Z}_{n_1, \dots, n_k} = \frac{1}{\prod_i n_i!} \int_{\gamma_1^{n_1} \times \dots \times \gamma_k^{n_k}} dx_1 \dots dx_N \prod_{i>j} (x_i - x_j)^2 \prod_i e^{-NV(x_i)} \quad (4.61)$$

The parameters  $\epsilon_i = n_i/N$  are called the filling fractions, because they represent the fraction of the eigenvalues on each path  $\gamma_i$ . This integral depends on a choice of filling fractions, called “background filling fraction”, and it only has a  $U(n_1) \times U(n_2) \times \dots \times U(n_k)$  invariance.

One may also write a path in different ways, and for instance if  $\gamma_1, \dots, \gamma_k$  form a homological basis of paths on which integrals of type (4.59) can be performed [10], then any path  $\gamma$  can be decomposed on such a basis:

$$\gamma = \sum_i c_i \gamma_i \quad (4.62)$$

and one may relate  $Z(\gamma)$  and  $\hat{Z}_{n_1, \dots, n_k}$ :

$$Z(\gamma) = \sum_{\sum_i n_i = N} \prod_i c_i^{n_i} \hat{Z}_{n_1, \dots, n_k} \quad (4.63)$$

Since not all  $n_i$  are independent, we can fix the overall normalization of  $Z(\gamma)$  by setting  $c_1 = 1$ . One may also study changes of paths in  $\hat{Z}_{n_1, \dots, n_k}$ .

## 4.2 Topological perturbative expansion

For some special choices of  $\gamma$ , or alternatively for some special choices of the basis  $\gamma_1, \dots, \gamma_k$  and some special choices of filling fractions  $n_1, \dots, n_k$ , it may happen that  $\ln Z$ , or  $\ln \hat{Z}$  have a perturbative large  $N$  expansion:

$$\ln \hat{Z}_{n_1, \dots, n_k} = \sum_{g=0}^{\infty} N^{2-2g} F_g \quad (4.64)$$

If such  $F_g$ 's exist, then Schwinger-Dyson equations imply [21, 14] that they coincide with the symplectic invariants defined in [25], for some spectral curve. This matrix model spectral curve is the so called “equilibrium density” of eigenvalues.

However, for most paths  $\gamma$ , or for arbitrary choices of the basis  $\gamma_1, \dots, \gamma_k$  and arbitrary choices of filling fractions  $n_1, \dots, n_k$ , matrix integrals have no perturbative large  $N$  expansion.

We conjecture (which is proved for special cases, for instance for the one-matrix model [9]), that given a potential  $V$ , and given  $n_1, \dots, n_k$ , there always exists a “good” basis of paths  $\gamma_1, \dots, \gamma_k$ , such that  $\hat{Z}_{n_1, \dots, n_k}$  has a topological expansion of the form (4.64). In some sense, this “good” basis of paths is made of the steepest descent contours for the integral (4.59).

## 4.3 Summation over filling fractions

Using (4.63):

$$Z(\gamma) = \sum_{\sum_i n_i = N} \prod_i c_i^{n_i} \hat{Z}_{n_1, \dots, n_k} \quad (4.65)$$

and using the fact that each  $\hat{Z}_{n_1, \dots, n_k}$  has a perturbative expansion of type (4.64), we can find the asymptotic expansion of  $Z(\gamma)$  as a combination of  $\hat{Z}_{n_1, \dots, n_k}$  after summation over the filling fractions  $n_i$ . This goes as follows.

For various matrix models, including the 1-matrix model, the 2-matrix model, the chain of matrices, and matrix models with external fields, the coefficients  $F_g$  in (4.64) have been computed in [14, 15, 26, 25], and are the symplectic invariants of [25]. They happen to be analytic in the filling fraction variables  $n_i/N$ , and thus they can be Taylor expanded near an arbitrary background value  $\epsilon_i$ :

$$N^{2-2g} F_g(n_i/N) = \sum_k \frac{N^{2-2g-k} (n_i - N\epsilon_i)^k}{k!} F_g^{(k)}(\epsilon_i) \quad (4.66)$$

The only terms with non-negative powers of  $N$  correspond to  $N^2 F_0 + F_1 + (n - N\epsilon)NF'_0 + \frac{1}{2}(n - N\epsilon)^2 F''_0$ . All the other terms have negative power of  $N$ , and thus can be expanded at large  $N$ , so that:

$$\hat{Z}_{n_1, \dots, n_k} \sim e^{N^2 F_0 + F_1 + \sum_{g \geq 2} N^{2-2g} F_g} \sum_k \sum_{l_i > 0} \sum_{h_i > 1 - \frac{l_i}{2}} \frac{N^{\sum_i (2-2h_i - l_i)}}{k! l_1! \dots l_k!} F_{h_1}^{(l_1)} \dots F_{h_k}^{(l_k)} (n - N\epsilon)^{\sum_i l_i} e^{(n - N\epsilon)NF'_0 + \frac{1}{2}(n - N\epsilon)^2 F''_0} \quad (4.67)$$

The sum over filling fractions (4.65) generates the non-perturbative partition function of [22], i.e. (2.3), under the identification:

$$\mu_i = 0, \quad c_i = e^{2i\pi\nu_i} \quad (4.68)$$

This shows that the choice of a characteristics  $(\mu, \nu)$  is in relationship with a choice of integration contour  $\gamma$  for the matrix integral (see Fig. 10).

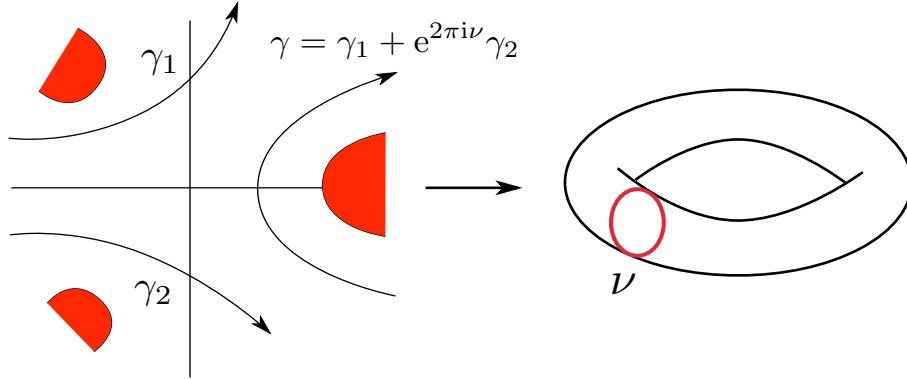


Figure 10: A choice of integration path for the matrix integral determines a choice of characteristic in the spectral curve, corresponding to A-cycles. We illustrate this correspondence in this figure for the case of the cubic matrix model, with a spectral curve of genus one. A typical path of integration for the eigenvalues is  $\gamma = \gamma_1 + e^{2i\pi\nu} \gamma_2$  and corresponds to a characteristic  $(0, \nu)$ .

This simplified description of the origin of the non-perturbative partition function for matrix models seems to involve only characteristics of type  $(0, \nu)$ , and not all characteristics  $(\mu, \nu)$ .



However, one should keep in mind that a “good basis of paths”  $\gamma_1, \dots, \gamma_k$  is not unique. Changes of basis correspond to modular transformations, and basis can be more complicated than what naive intuition seems to show. In the “naive” picture, it is often thought that, in the large  $N$  limit of multicut matrix integrals, eigenvalues tend to localize along disconnected segments called “cuts”. In this case, there is a natural choice of  $A$ -cycles, as contours surrounding the cuts (see Fig. 11, left).

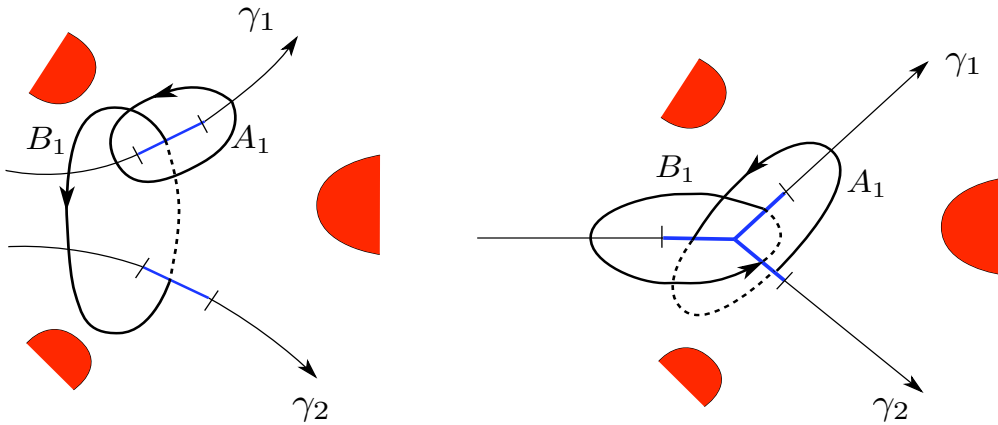


Figure 11: On the left we show a situation in which there is a natural choice of  $A$ -cycle surrounding the cut along the path  $\gamma_1$ . However, in some situations the eigenvalues might localize on tree structures, as shown in the figure on the right.

But in general, it is well known [17, 9] that, in the large  $N$  limit of multicut matrix integrals, eigenvalues tend to localize along forests of 3-valent trees rather than union of segments (see Fig. 11, right). In that case, there is no natural distinction between  $A$ -cycles and  $B$ -cycles, and maybe in such tree structures one could probably interpret the general characteristics  $(\mu, \nu)$  with  $\mu \neq 0$ .

#### 4.4 Remarks on background independence

First, notice that we have chosen to perform a Taylor expansion near an arbitrary background filling fraction  $\epsilon$ , although the partition function  $Z(\gamma)$  is independent of  $\epsilon$ . We clearly have background independence.

Also, we have used the existence of a good basis  $\gamma_1, \dots, \gamma_k$ . However, such a good basis is not unique, and one may perform changes of basis. Such changes of basis are equivalent to symplectic changes of cycles on the spectral curve. The fact that  $Z(\gamma)$  depends intrinsically

on  $\gamma$ , and not on a choice of basis on which  $\gamma$  is decomposed as in (4.62), is a hint towards modular invariance.

Finally, we notice that a choice of background leads to a breaking of the original, unitary  $U(N)$  symmetry of the matrix integral down to a subgroup  $U(n_1) \times \cdots \times U(n_k)$ . Therefore, the restoration of modular invariance and background independence by non-perturbative corrections seems to be deeply related to the restoration of the  $U(N)$  symmetry.

## 5 The nonperturbative partition function for topological strings

### 5.1 Topological strings, matrix models and spectral curves

In [20], Dijkgraaf and Vafa showed that the type B topological string, on certain Calabi-Yau (CY) geometries, is equivalent to a Hermitian matrix model with polynomial potential  $V(x)$  (see [38] for a review). These target geometries are deformed singularities of the form

$$uv = H(x, y) \tag{5.1}$$

where  $x, y, u, v \in \mathbb{C}$  and

$$H(x, y) = y^2 - V^{\prime 2}(x) - f(x). \tag{5.2}$$

The nontrivial information about this geometry turns out to be encoded in the Riemann surface  $\Sigma$  described by  $H(x, y) = 0$ . An important insight of the analysis of Dijkgraaf and Vafa is that  $\Sigma$  is precisely the spectral curve of the corresponding Hermitian matrix model, providing in this way a beautiful example in which the master field of the  $1/N$  expansion generates the target geometry of a string theory.

More recently, the correspondence between matrix models/spectral curves and topological strings was extended to toric CY threefolds [39, 12]. This class of examples is very interesting since (in contrast to the geometries considered in [20]) they have mirror geometries. The mirrors are described by an equation of the form (5.1), but where the variables  $x, y$  now belong to  $\mathbb{C}^*$ , and again the nontrivial part of the geometry reduces to a Riemann surface  $\Sigma$  described by  $H(x, y) = 0$ . In [25] it was shown that, given any algebraic curve, not necessarily coming from a matrix model, one can compute a set of “open” and “closed” amplitudes. When the curve is the spectral curve of a matrix model, these amplitudes reproduce the  $1/N$  expansion of correlators and free energies, respectively. However, the construction is valid for any spectral curve. The proposal of [39, 12] is to regard the Riemann surface  $H(x, y) = 0$  appearing in (5.1) as a spectral curve in the general sense of [25], and to identify the open and closed string amplitudes of the B–model with the open and closed amplitudes constructed

in [25] purely in terms of data of the spectral curves. Therefore, for this more general class of examples, we rather have a correspondence between the type B topological string amplitudes on the mirror of a toric manifold, and the amplitudes associated to a spectral curve defined in [25].

For certain backgrounds it is however still possible to provide a matrix integral representation. In the case of local curves, i.e. manifolds of the form

$$X_p = \mathcal{O}(p) \oplus \mathcal{O}(-p-2) \rightarrow \mathbb{P}^1 \quad (5.3)$$

this representation can be obtained directly from the A-model answer for the amplitudes, written as a sum over partitions [23]. Another family of Calabi–Yau’s where a matrix integral representation is available is given by

$$Y_p = S_{A_{p-1}} \rightarrow \mathbb{P}^1, \quad (5.4)$$

where  $S_{A_{p-1}}$  is a four-dimensional  $A_{p-1}$  resolved singularity. For example, for  $p = 2$ ,  $S_{A_1}$  is a resolved  $\mathbb{C}^2/\mathbb{Z}_2$  singularity and we can represent it also as

$$Y_2 = K^{-1} \rightarrow \mathbb{P}^1 \times \mathbb{P}^1, \quad (5.5)$$

i.e. as the total space of the anticanonical bundle of  $\mathbb{P}^1 \times \mathbb{P}^1$ . These backgrounds have a large  $N$  dual [3] given by Chern–Simons theory with gauge group  $U(N)$  on the lens space

$$L(p, 1) = \mathbb{S}^3/\mathbb{Z}_p. \quad (5.6)$$

For  $p = 1$  this is just the Gopakumar–Vafa duality of [29] (extensions to  $L(p, q)$  have been recently discussed in [13]). It turns out that Chern–Simons theory on  $L(p, 1)$  has a matrix integral realization [37], and it can be shown that the spectral curve of this matrix integral reproduces the mirror geometry [31]. Therefore, for those backgrounds, one can indeed give a nonperturbative definition in terms of a matrix model.

## 5.2 Topological strings and the nonperturbative partition function

It has been argued in [41, 40] that the above correspondence between matrix models/spectral curves and topological strings could also be used to obtain nonperturbative information on the topological string side. In particular, instanton effects on the matrix model side should lead to spacetime instantons in topological string theory, generalizing in this way the connection between eigenvalue tunneling and ZZ branes in noncritical string theory [43, 4].

This has been verified in some simple models in [41, 40] by checking that these instanton effects control the large order behavior of string perturbation theory.

Since the nonperturbative partition function of [22] studied in this paper incorporates the full multi-instanton series of the matrix model and has good modular properties, it is natural to generalize [41, 40] and use  $Z_\Sigma(\mu, \nu)$  to define topological string theory nonperturbatively. Notice that the ingredients to construct  $Z_\Sigma(\mu, \nu)$  are simply the spectral curve and a choice of characteristics. All the nonperturbative information is encoded in the latter, since they correspond to (and generalize) the choice of contour in the matrix integral.

We then make the following proposal. Let  $X_\Sigma$  be a local Calabi–Yau manifold whose geometry is encoded in a spectral curve  $\Sigma$ . For each choice of characteristics  $\mu, \nu$  on  $\Sigma$ , we define the nonperturbative topological string partition function  $Z_{X_\Sigma}(\mu, \nu)$  associated to the characteristics  $\mu, \nu$  as

$$Z_{X_\Sigma}(\mu, \nu) = Z_\Sigma(\mu, \nu). \quad (5.7)$$

In this equality, the filling fractions are identified with the moduli of the Calabi–Yau as in [20, 3], and the string coupling constant is identified with  $1/N$ :

$$t^i = \epsilon^i, \quad g_s = \frac{1}{N}. \quad (5.8)$$

In principle, each choice of characteristic provides a nonperturbative completion of topological string theory, as it was argued in [40] in some simple examples. This situation is of course reminiscent of the nonperturbative ambiguity in 2d gravity, where different choices of integration contour provide different nonperturbative completions.

Our proposal (5.7) incorporates previous suggestions concerning the nonperturbative structure of topological string theory. In [35] it was pointed out, based on results for non-critical strings, that the exact topological string partition function should be a sum over all vacua, i.e. over all instantons obtained by filling all critical points of the potential. This is precisely what was done in [22] to obtain  $Z_\Sigma(\mu, \nu)$ , and in fact it is at the origin of its background independence.  $Z_{X_\Sigma}(\nu, \mu)$  is also closely related to the generalized partition functions introduced in section 4.7 of [2] and in section 3.4 of [18], in the context of a fermion formulation of topological string theory on the spectral curve  $\Sigma$ . In both cases, one considers a partition function where one sums over all possible fermion numbers with fixed twists around the A cycles, i.e. one goes to a grand canonical ensemble for the fermions where the twists correspond to chemical potentials. The relation to a fermion system is now further clarified, since what we have seen in this paper is that  $Z_\Sigma(\mu, \nu)$  has the same modular properties as the partition function of a twisted fermion on  $\Sigma$ , with twists  $\mu, \nu$ . We will extract more consequences of this below.

Although our proposal does not completely solve the problem of providing a nonperturbative definition of topological string theory, even in the local case, it cuts down the freedom

in making such a definition to a well-defined choice: a characteristic on the Riemann surface. Moreover, there are some nonperturbative effects that do depend on the multi-instanton amplitudes encoded in (5.7) but not on the choice of characteristic. One example is the large order behaviour of string perturbation theory, as discussed in [41, 40].

How reasonable is our proposal (5.7)? As it was pointed out in [40], for the Calabi–Yau’s of the form (5.4), the topological string theory has a large  $N$  Chern–Simons dual, and the total partition function is indeed of the form  $Z_\Sigma(\mu, \nu)$ . Moreover the characteristics, i.e. the nonperturbative data, are fixed. To see this, one notices that at finite  $N$  one has [37, 3, 40]

$$Z_{\text{CS}}(N, g_s) = \sum_{N_1 + \dots + N_p = N} c_1^{N_1} \dots c_p^{N_p} Z(N_1, \dots, N_p), \quad (5.9)$$

where

$$c_j = \exp\left(\frac{\pi i \hat{k}}{p}(j-1)^2\right), \quad j = 1, \dots, p \quad (5.10)$$

and  $\hat{k} = k + N$  is the shifted coupling constant of Chern–Simons theory, which is related to the string coupling constant by

$$g_s = \frac{2\pi i}{p \hat{k}}. \quad (5.11)$$

The r.h.s of (5.9) is of the form (4.65), therefore at large  $N$  is nothing but  $Z_{\Sigma_p}(0, \nu)$ , where the spectral curve  $\Sigma_p$  is the Riemann surface appearing in the mirror of  $Y_p$  and it has genus  $\bar{g} = p - 1$ . Since  $c_1$  is already normalized to one, we read that the characteristic  $\nu$  is given by

$$\nu_j = \frac{\hat{k}}{2p} j^2, \quad j = 1, \dots, \bar{g} \quad (5.12)$$

while  $\mu = 0$ . Notice that, in Chern–Simons theory,  $\hat{k}$  is an integer, and the  $c_j$  are indeed phases which depend on truly nonperturbative data, i.e. the value of  $k + N$  as a quantized integer (the quantization of  $k$  in Chern–Simons theory is not visible in perturbation theory, nor in the large  $N$  expansion). Therefore, for these backgrounds, the partition function of the holographic Chern–Simons dual is precisely the nonperturbative partition function  $Z_{\Sigma_p}(\mu, \nu)$  with a definite choice of characteristics. The proposal (5.7) is then completely natural in this case.

If we now consider a general local Calabi–Yau background, is there any way of fixing the characteristics, i.e. the nonperturbative content of the theory? A possible guiding principle is the following. Note that, if  $\mu_i, \nu_i \in \mathbb{R}$ , the partition function  $Z_\Sigma(\mu, \nu)$  transforms with phases. Let us suppose that there is a discrete set of real characteristics,  $\mathcal{S}$ , which is invariant

under a suitable subgroup  $\mathcal{G} \subset \mathrm{Sp}(2\bar{g}, \mathbb{Z})$  of the full modular group. It follows that

$$Z_{\Sigma, S} = \sum_{\mu, \nu \in S} \left| Z_{\Sigma}(\mu, \nu) \right|^2 \quad (5.13)$$

is invariant under  $\mathcal{G}$ . In a Calabi–Yau threefold  $X$ , the invariance subgroup of the modular group is the so-called *monodromy group*  $\mathcal{M}_X$ , generated by the monodromies of the periods around the singularities in the moduli space of the spectral curve. It is natural to postulate that the nonperturbative structure of topological string theory on a local Calabi–Yau  $X$  is given by a set  $S$  of characteristics  $\mu, \nu$  such that the sum (5.13) is invariant under the monodromy group  $\mathcal{M}_X$ . We remark that the sum (5.13) is a “diagonal” invariant, and it is conceivable that one can find non-diagonal invariants, as in the ADE classification of modular invariants for minimal models.

As an example of this discussion, let us consider Seiberg–Witten theory for pure  $\mathcal{N} = 2$  Yang–Mills theory with gauge group  $SU(2)$ , which can be regarded as a special limit of topological string theory (see [33, 32] for reviews of this model and its stringy origin). The spectral curve of this theory is the Seiberg–Witten curve  $\Sigma_{\mathrm{SW}}$

$$y^2 = x \left( x^2 - ux + \frac{1}{4} \right), \quad (5.14)$$

which is an elliptic curve with genus  $\bar{g} = 1$ . The moduli space of this curve is the  $u$ -plane, with three singular points  $u = \infty, 1, -1$ . The monodromy group is  $\Gamma^0(4)$ , and it is generated by the monodromies around the singular points at  $\infty$  and  $u = 1$ ,

$$M_{\infty} = \begin{pmatrix} -1 & 4 \\ 0 & -1 \end{pmatrix}, \quad M_1 = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}. \quad (5.15)$$

Given the Seiberg–Witten curve, we can define nonperturbative partition functions for any characteristic  $(\mu, \nu)$ . It turns out that the characteristic  $(0, 1/2)$  is invariant under  $\Gamma^0(4)$ , therefore

$$\left| Z_{\Sigma_{\mathrm{SW}}}(0, 1/2) \right|^2, \quad (5.16)$$

is also invariant under the monodromy group. The theta function with this characteristic,  $\vartheta_4$ , appears naturally in Seiberg–Witten theory and in Donaldson–Witten theory in relation to the so-called blow-up function, see [36] for a review and a list of references.

Another example is the manifold  $Y_2$  in (5.5). Its monodromy group is  $\Gamma(2)$  [1]. In this case, one has from (5.10) that  $\mu = 0$  and

$$\nu = \frac{\pi i \hat{k}}{2} \quad (5.17)$$

and it depends on the value of the integer  $\hat{k} = k + N \bmod 4$ . For example, if  $\hat{k} = 4n + 2$ ,  $n \in \mathbb{Z}$ , one has  $\mu = 0, \nu = 1/2$  which is also invariant under the monodromy group.

The requirement of monodromy invariance of (5.13) might not be enough to completely fix the nonperturbative structure of the theory. It rather suggests that we should think of the nonperturbative partition functions  $Z_\Sigma(\mu, \nu)$  as a set of conformal blocks associated to the mirror curve  $\Sigma$ . Finally, we note the natural appearance in this formalism of the modulus square of the nonperturbative partition function, which features in the OSV conjecture [45]. In our context, this is simply due to the requirement of modular invariance.

## 6 Integrability

Here, we show that the non-perturbative partition function (2.3), is a tau-function in the sense that it obeys Hirota equations. We follow the same ideas as in [25], which were applied there only for spectral curves  $\Sigma = (\mathcal{C}, x, y)$  of genus  $\bar{g} = 0$ , precisely because the  $\Theta$ -terms were missing in [25]. So, here we complete the proof of [25] in the general case.

For a spectral curve  $\Sigma = (\mathcal{C}, x, y)$ , let us write:

$$\tau(\Sigma) = \tau(\mathcal{C}, x, y) = Z_\Sigma(\mu, \nu; \epsilon) \quad (6.1)$$

We shall define the Baker-Akhiezer kernel through a Sato formula [46]:

$$K(z_1, z_2) = \frac{1}{x(z_1) - x(z_2)} \frac{\tau(\mathcal{C}, x, y + \frac{1}{N} \frac{dS_{z_1, z_2}}{dx})}{\tau(\mathcal{C}, x, y)} \quad (6.2)$$

where  $dS_{z_1, z_2}(z)$  is the 3rd kind differential having a simple pole at  $z = z_1$  with residue  $+1$ , and a pole at  $z = z_2$  with residue  $-1$ , and no other pole, and normalized on  $A$ -cycles, i.e. it is the integral of the Bergmann kernel:

$$dS_{z_1, z_2}(z) = \int_{z_2}^{z_1} B(z, z') \quad (6.3)$$

In [25], it is explained how to compute derivatives of  $F_g$  with respect to any parameter of the spectral curve (see appendix A), and in particular, if we consider the spectral curve  $(\mathcal{C}, x, y + r \frac{dS_{z_1, z_2}}{dx})$ , we have:

$$\frac{\partial}{\partial r} W_n^{(g)}(p_1, \dots, p_n) = \int_{z_2}^{z_1} W_{n+1}^{(g)}(p_1, \dots, p_n, z) \quad (6.4)$$

and we may use this to compute the value of  $\tau$  at  $r = 1/N$  by Taylor expansion:

$$\tau(\mathcal{C}, x, y + \frac{1}{N} \frac{dS_{z_1, z_2}}{dx}) = \sum_{k=0}^{\infty} \frac{N^{-k}}{k!} \partial_r^k \tau(\mathcal{C}, x, y) \quad (6.5)$$

This Taylor expansion can be performed more explicitly

$$\begin{aligned} K(z_1, z_2) &= \frac{e^{N^2 F_0 + F_1}}{Z_{\Sigma}(\mu, \nu)} \frac{e^{-N \int_{z_2}^{z_1} y dx}}{E(z_1, z_2) \sqrt{dx(z_1) dx(z_2)}} \\ &= \frac{\sum_k \sum_{n_i, l_i} \sum_{h_i > 1 - \frac{l_i + n_i}{2}} \frac{N^{\sum_i (2-2h_i - l_i - n_i)}}{k! l_1! \dots l_k! n_1! \dots n_k!} \prod_{i=1}^k \left( \int_{z_2}^{z_1} \dots \int_{z_2}^{z_1} \partial_{\epsilon}^{l_i} W_{n_i}^{(h_i)} \right) \Theta_{\mu, \nu}^{(\sum_i l_i)}(NF'_0, \tau)}{e^{-N \int_{z_2}^{z_1} y dx}} \\ &= \frac{1}{E(z_1, z_2) \sqrt{dx(z_1) dx(z_2)}} \\ &\quad \left\{ 1 + \frac{1}{N} \left( \int_{z_2}^{z_1} W_1^{(1)} + \frac{1}{6} \int_{z_2}^{z_1} \int_{z_2}^{z_1} \int_{z_2}^{z_1} W_3^{(0)} + \frac{1}{2} \int_{z_2}^{z_1} \int_{z_2}^{z_1} W_2^{(0)'} \frac{\Theta'}{\Theta} + \frac{1}{2} \int_{z_2}^{z_1} W_1^{(0)''} \frac{\Theta''}{\Theta} \right) + \dots \right\} \end{aligned} \quad (6.6)$$

where  $E(z_1, z_2)$  is the prime form. In this expression, each  $\partial_{\epsilon}^{l_i} W_{n_i}^{(h_i)}$  is an abelian meromorphic form on  $\mathcal{C}$ , with no residue, and thus its integral from  $z_2$  to  $z_1$  is an abelian integral and defines a function of  $z_1$  and  $z_2$  only on the universal covering of  $\mathcal{C}$ . If  $\mathcal{C}$  is not simply connected, it is not obvious at all that expression (6.6) actually defines a function on  $\mathcal{C}$ . However, thanks to background independence, this is the case:  $K(z_1, z_2)$  is a function of  $z_1, z_2 \in \mathcal{C}$ . Indeed, if we move  $z_1$  around a  $B$ -cycle,  $dS_{z_1, z_2}$  gets shifted by the holomorphic differential  $\omega_i$ :

$$dS_{z_1 + B_i, z_2} = dS_{z_1, z_2} + 2i\pi\omega_i \quad (6.7)$$

and thus moving  $z_1$  around the  $\mathcal{B}_i$  cycle, is equivalent to shifting  $y dx$  by the holomorphic differential  $2i\pi\omega_i$ , which is also equivalent to changing the filling fraction  $\epsilon_i \rightarrow \epsilon_i + \frac{1}{N}$ . Since  $\tau$  is independent of  $\epsilon$ , this means that

$$\tau(\mathcal{C}, x, y + \frac{1}{N} \frac{dS_{z_1 + B_i, z_2}}{dx}) = \tau(\mathcal{C}, x, y + \frac{1}{N} \frac{dS_{z_1, z_2} + 2i\pi\omega_i}{dx}) = \tau(\mathcal{C}, x, y + \frac{1}{N} \frac{dS_{z_1, z_2}}{dx}) \quad (6.8)$$

and thus  $K(z_1, z_2)$  is a well defined function of  $z_1, z_2 \in \mathcal{C}$ .

$K(z_1, z_2)$  has essential singularities when  $z_1$  (resp.  $z_2$ ) approaches a pole of  $y dx$ , and is such that:

$$d_{z_1} \ln K(z_1, z_2) \sim -N y(z_1) dx(z_1) \quad (6.9)$$



$K(z_1, z_2)$  also has a simple pole at  $z_1 = z_2$ :

$$K(z_1, z_2) \sim \frac{1}{x(z_1) - x(z_2)} \quad (6.10)$$

and those are the only singularities of  $K$ .

For any pair of spectral curves  $\Sigma$  and  $\tilde{\Sigma}$  over  $\mathcal{C}$  we have the bilinear relation:

$$\operatorname{Res}_{z_2 \rightarrow z_3} dx(z_2) K_{\Sigma}(z_1, z_2) K_{\tilde{\Sigma}}(z_2, z_3) = K_{\Sigma}(z_1, z_3) \quad (6.11)$$

When written in terms of the moduli of  $\Sigma$ , this relation can be interpreted as the Hirota equation for the multicomponent KP hierarchy [6].

## 7 Discussion

Given an arbitrary spectral curve  $\Sigma = (\mathcal{C}, x, y)$ , and arbitrary characteristics  $(\mu, \nu)$ , we have defined a (non-perturbative) partition function (2.3), and we have proved that it is both background independent and modular invariant, and it is perfectly holomorphic. In other words, we have restored modular invariance not by breaking holomorphicity, but by breaking the perturbative expansion, i.e. by taking into account all instantons.

One consequence of our analysis is that the holomorphic anomaly of the topological string might not be as fundamental as previously thought. If one thinks about the non-holomorphic dependence of the partition function as a way to restore modularity, one is tempted to say, in view of the results in this paper, that non-holomorphicity was the prize to pay in order to forget about nonperturbative corrections. This observation could also explain a long-standing puzzle for the large  $N$  dualities of the topological string: the absence of a natural mechanism in the gauge theory side that might lead to non-holomorphic dependence on the 't Hooft couplings. In other words, there is no known, natural gauge theory dual of the holomorphic anomaly. According to our results, rather than mimicking the non-holomorphic way in which perturbative string theory achieves modularity, the gauge theory side restores the modularity properties of the full partition function by using non-perturbative effects, i.e. spacetime instantons.

Another interesting outcome of our analysis is that it explains the similarity between the holomorphic anomaly equations and the heat equations for the theta functions that has been pointed out in [47] and recently elaborated in [30, 22]. The nonperturbative partition function is simply the product of the perturbative partition function times a complicated theta function, as shown in (2.11). Since the product is modular, it follows that the perturbative piece has to transform precisely in a compensating way, therefore mirroring the transformation properties of theta functions. Notice however that the theta function involves  $1/N$

corrections that depend on the perturbative free energies in a complicated way. It would be interesting to compare the nonperturbative partition function studied in this paper to the holomorphic wavefunction introduced in [30], which was conjectured to be modular.

The results presented in this paper can be improved in many aspects. The expression (2.3) for the nonperturbative partition function is not completely satisfactory, since it is still defined by a  $1/N$  expansion. It would be interesting to provide a more intrinsic formulation of  $Z_\Sigma(\mu, \nu)$ , maybe by using the integrability properties discussed in section 6. Each term in the expansion (2.3) contains however a sum over all the multi-instantons. In this sense, the organization of the series is similar to what was done in [16] in the analysis of trans-series expansions: it is perturbative in  $1/N$ , but non-perturbative in  $e^{\pm N}$ .

It would be clearly important to understand the detailed implications of the nonperturbative completion that we are proposing. Since this completion involves multi-instanton corrections, it would be very interesting to relate them more precisely to the simplest cases studied in [41, 40, 42]. These papers analyzed multi-instantons in the one-cut case, which should be regarded as a limiting case of the general situation considered in this paper (in fact, in [42] the one-cut case was already regarded as a limit of the multi-cut case). One could use the multi-instanton information encoded in the nonperturbative partition function to understand the large order behavior of string perturbation theory and of the  $1/N$  expansion of multi-cut matrix models, generalizing what was done in [41] in the one-cut case. Finally, it would be also interesting to relate the multi-instanton configurations encoded in  $Z_\Sigma(\mu, \nu)$  to D-brane effects in topological string theory.

On a more speculative note, we would like to point out that the nonperturbative partition function introduced in this paper incorporates a mechanism for recovering background independence which is similar in many respects to the mechanism suggested in [34] for restoring unitarity in some black hole backgrounds. In both cases one needs to sum over subleading saddle configurations which are invisible in the  $1/N$  expansion. It is also interesting to notice that, when including nonperturbative corrections in the form of theta functions, the correlation functions of the matrix model display quasi-periodicity properties [11], as required in theories without information loss (see the nice review in [7] and references therein). Therefore, the mechanism to recover background independence studied in this paper might provide a useful toy model for the information paradox in the AdS/CFT correspondence.

## Acknowledgments

We would like to thank Vincent Bouchard, Andrea Brini, Boris Dubrovin, Tamara Grava, Albrecht Klemm, Hiroshi Ooguri, Sara Pasquetti, Ricardo Schiappa, Cumrun Vafa and Marlene Weiss for useful and fruitful discussions on this subject. B.E. would like to thank the

Department of Mathematics of the University of Geneva for hospitality. The work of B.E. is partly supported by the Enigma European network MRT-CT-2004-5652, by the ANR project Géométrie et intégrabilité en physique mathématique ANR-05-BLAN-0029-01, by the European Science Foundation through the Misgam program, by the Quebec government with the FQRNT. The work of M.M. is partially supported by FNS under the projects 200021-121605 and 200020-121675.

## A Symplectic invariants

Here, we recall a few basic ingredients of the symplectic invariants  $F_g$ 's introduced in [25].

### A.1 Spectral curve

Consider a spectral curve  $\Sigma = (\mathcal{C}, x, y)$ , where  $\mathcal{C}$  is a compact Riemann surface of genus  $\bar{g}$  (not to be confused with the topological index  $g$  of  $F_g$ ), and  $x$  and  $y$  are two analytical functions on some open domain of  $\mathcal{C}$ .

The compact Riemann surface  $\mathcal{C}$  is equipped with a symplectic basis of non-contractible cycles:

$$\mathcal{A}_i \cap \mathcal{B}_j = \delta_{i,j} \quad (\text{A.12})$$

and it comes with all classical tools of algebraic geometry (see [28,27]), in particular a prime form  $E(z_1, z_2)$ , which has a simple zero at  $z_1 = z_2$ , and a Bergmann kernel  $B(z_1, z_2)$ , which has a double pole at  $z_1 = z_2$ :

$$B(z_1, z_2) = \frac{df(z_1)df(z_2)}{(f(z_1) - f(z_2))^2} + \text{regular} \quad (\text{A.13})$$

in any local parametrization  $f(z)$ .  $B$  is normalized on  $\mathcal{A}$ -cycles:

$$\oint_{\mathcal{A}_i} B = 0 \quad (\text{A.14})$$

Explicitly we have ( $c$  being an odd characteristics):

$$B(z_1, z_2) = d_{z_1} d_{z_2} \ln \theta(z_1 - z_2 + c) \quad (\text{A.15})$$

For example if  $\mathcal{C}$  is a torus  $\mathbb{C}/\mathbb{Z} + \tau\mathbb{Z}$ ,  $B$  is the Weierstrass function  $B(z_1, z_2) = \wp(z_1 - z_2) dz_1 dz_2$ .

The branch points are the points  $a_i$  where  $dx = 0$ .

$$dx(a_i) = 0 \quad (\text{A.16})$$

We assume all branch points to be regular, i.e.  $a_i$  is a simple zero of  $dx$ , and  $dy(a_i) \neq 0$ . This means that, if  $z$  is in the vicinity of  $a_i$ , there exists a unique  $\bar{z} \neq z$  in the same vicinity of  $a_i$  such that

$$x(z) = x(\bar{z}) \quad (\text{A.17})$$

The recursion kernel  $K$  is then:

$$K(z_1, z) = \frac{\int_z^{\bar{z}} B(z_1, z')}{2(y(z) - y(\bar{z})) dx(z)} \quad (\text{A.18})$$

## A.2 Correlators and symplectic invariants

We define:

$$W_2^{(0)}(z_1, z_2) = B(z_1, z_2) \quad (\text{A.19})$$

and, if we denote  $J = \{z_1, z_2, \dots, z_n\}$ , we define recursively (on  $2g + n$ ):

$$\begin{aligned} W_{n+1}^{(g)}(J, z_{n+1}) &= \sum_i \operatorname{Res}_{z \rightarrow a_i} K(z_{n+1}, z) \left[ W_{n+2}^{(g-1)}(z, \bar{z}, J) \right. \\ &\quad \left. + \sum_{h=0}^g \sum_{I \subset J} W_{1+|I|}^{(h)}(z, I) W_{1+n-|I|}^{(g-h)}(\bar{z}, J/I) \right] \end{aligned} \quad (\text{A.20})$$

where  $\sum'$  means that we exclude the terms  $(h, I) = (0, \emptyset)$  and  $(h, I) = (g, J)$ .

The symplectic invariants  $F_g = W_0^{(g)}$  are defined by (if  $g \geq 2$ )

$$F_g = \frac{1}{2-2g} \sum_i \operatorname{Res}_{z \rightarrow a_i} W_1^{(g)}(z) \Phi(z) \quad (\text{A.21})$$

where  $d\Phi = ydx$ . For the definition of  $F_1$  and  $F_0$ , we refer the reader to [25].

## A.3 Derivatives

The filling fraction is:

$$\epsilon_i = \frac{1}{2i\pi} \oint_{\mathcal{A}_i} ydx \quad (\text{A.22})$$

Changing  $\epsilon_i$  is equivalent to changing  $x$  and  $y$  by:

$$\frac{\partial}{\partial \epsilon_i} ydx = 2i\pi\omega_i = \oint_{\mathcal{B}_i} B \quad (\text{A.23})$$

where  $\omega_j$  are the holomorphic abelian differentials on  $\mathcal{C}$ , normalized on  $\mathcal{A}$ -cycles:

$$\oint_{\mathcal{A}_i} \omega_j = \delta_{i,j} \quad (\text{A.24})$$

Then [25] tells that:

$$\frac{\partial}{\partial \epsilon_i} W_n^{(g)}(z_1, \dots, z_n) = \oint_{\mathcal{B}_i} W_{n+1}^{(g)}(z_1, \dots, z_n, z) \quad (\text{A.25})$$

Similarly, if we take a variation of  $ydx$  such that:

$$\frac{\partial}{\partial r} ydx = dS_{p_1, p_2} = \int_{p_2}^{p_1} B \quad (\text{A.26})$$

then [25] tells that:

$$\frac{\partial}{\partial r} W_n^{(g)}(z_1, \dots, z_n) = \int_{p_2}^{p_1} W_{n+1}^{(g)}(z_1, \dots, z_n, z) \quad (\text{A.27})$$

## References

- [1] M. Aganagic, V. Bouchard and A. Klemm, “Topological Strings and (Almost) Modular Forms,” *Commun. Math. Phys.* **277**, 771 (2008) [arXiv:hep-th/0607100].
- [2] M. Aganagic, R. Dijkgraaf, A. Klemm, M. Mariño and C. Vafa, “Topological strings and integrable hierarchies,” *Commun. Math. Phys.* **261**, 451 (2006) [arXiv:hep-th/0312085].
- [3] M. Aganagic, A. Klemm, M. Mariño and C. Vafa, “Matrix model as a mirror of Chern-Simons theory,” *JHEP* **0402**, 010 (2004) [arXiv:hep-th/0211098].
- [4] S. Y. Alexandrov, V. A. Kazakov and D. Kutasov, “Non-perturbative effects in matrix models and D-branes,” *JHEP* **0309**, 057 (2003) [arXiv:hep-th/0306177].
- [5] L. Álvarez-Gaumé, G. W. Moore and C. Vafa, “Theta functions, modular invariance, and strings,” *Commun. Math. Phys.* **106**, 1 (1986).
- [6] O. Babelon, D. Bernard, M. Talon, *Introduction to Classical Integrable System*, Cambridge University Press, 2003.

- [7] J. L. F. Barbón and E. Rabinovici, “Topology change and unitarity in quantum black hole dynamics,” arXiv:hep-th/0503144.
- [8] M. Bershadsky, S. Cecotti, H. Ooguri and C. Vafa, “Kodaira-Spencer theory of gravity and exact results for quantum string amplitudes,” *Commun. Math. Phys.* **165**, 311 (1994) [arXiv:hep-th/9309140].
- [9] M. Bertola, “Boutroux curves with external field: equilibrium measures without a minimization problem”, arXiv:0705.3062.
- [10] M. Bertola, “Bilinear semi-classical moment functionals and their integral representation”, arXiv:math/0205160.
- [11] G. Bonnet, F. David and B. Eynard, “Breakdown of universality in multi-cut matrix models,” *J. Phys. A* **33**, 6739 (2000) [arXiv:cond-mat/0003324].
- [12] V. Bouchard, A. Klemm, M. Mariño and S. Pasquetti, “Remodeling the B-model,” arXiv:0709.1453 [hep-th].
- [13] A. Brini, L. Griguolo, D. Seminara and A. Tanzini, “Chern-Simons theory on  $L(p,q)$  lens spaces and Gopakumar-Vafa duality,” arXiv:0809.1610 [math-ph].
- [14] L. Chekhov and B. Eynard, “Hermitean matrix model free energy: Feynman graph technique for all genera,” *JHEP* **0603**, 014 (2006) [arXiv:hep-th/0504116].
- [15] L. Chekhov, B. Eynard and N. Orantin, “Free energy topological expansion for the 2-matrix model,” *JHEP* **0612**, 053 (2006) [arXiv:math-ph/0603003].
- [16] O. Costin and R. Costin, “On the formation of singularities of solutions of nonlinear differential systems in antistokes directions,” *Invent. Math.* **145**, 425 (2001) [arXiv:math.CA/0202234].
- [17] F. David, “Nonperturbative effects in matrix models and vacua of two-dimensional gravity,” *Phys. Lett. B* **302**, 403 (1993) [arXiv:hep-th/9212106].
- [18] R. Dijkgraaf, L. Hollands, P. Sulkowski and C. Vafa, “Supersymmetric Gauge Theories, Intersecting Branes and Free Fermions,” *JHEP* **0802**, 106 (2008) [arXiv:0709.4446 [hep-th]].
- [19] P. Di Francesco, P. Ginsparg, J. Zinn-Justin, “2D Gravity and Random Matrices”, *Phys. Rep.* **254**, 1 (1995).

- [20] R. Dijkgraaf and C. Vafa, “Matrix models, topological strings, and supersymmetric gauge theories,” Nucl. Phys. B **644**, 3 (2002) [arXiv:hep-th/0206255].
- [21] B. Eynard, “Topological expansion for the 1-hermitian matrix model correlation functions,” JHEP **0411**, 031 (2004) [arXiv:hep-th/0407261].
- [22] B. Eynard, “Large N expansion of convergent matrix integrals, holomorphic anomalies, and background independence,” arXiv:0802.1788 [math-ph].
- [23] B. Eynard, “All orders asymptotic expansion of large partitions,” J. Stat. Mech. **0807**, P07023 (2008) [arXiv:0804.0381 [math-ph]].
- [24] B. Eynard, M. Mariño and N. Orantin, “Holomorphic anomaly and matrix models,” JHEP **0706**, 058 (2007) [arXiv:hep-th/0702110].
- [25] B. Eynard and N. Orantin, “Invariants of algebraic curves and topological expansion”, arXiv:math-ph/0702045.
- [26] B. Eynard and A. Prats Ferrer, “Topological expansion of the chain of matrices,” arXiv:0805.1368 [math-ph].
- [27] H. M. Farkas and I. Kra, *Riemann surfaces*, Springer.
- [28] J. D. Fay, *Theta functions on Riemann surfaces*, Springer.
- [29] R. Gopakumar and C. Vafa, “On the gauge theory/geometry correspondence,” Adv. Theor. Math. Phys. **3**, 1415 (1999) [arXiv:hep-th/9811131].
- [30] M. Gunaydin, A. Neitzke and B. Pioline, “Topological wave functions and heat equations,” JHEP **0612**, 070 (2006) [arXiv:hep-th/0607200].
- [31] N. Halmagyi and V. Yasnov, “The spectral curve of the lens space matrix model,” arXiv:hep-th/0311117. N. Halmagyi, T. Okuda and V. Yasnov, “Large N duality, lens spaces and the Chern-Simons matrix model,” JHEP **0404**, 014 (2004) [arXiv:hep-th/0312145].
- [32] A. Klemm, “On the geometry behind N = 2 supersymmetric effective actions in four dimensions,” arXiv:hep-th/9705131.
- [33] W. Lerche, “Introduction to Seiberg-Witten theory and its stringy origin,” Nucl. Phys. Proc. Suppl. **55B**, 83 (1997) [Fortsch. Phys. **45**, 293 (1997)] [arXiv:hep-th/9611190].

- [34] J. M. Maldacena, “Eternal black holes in Anti-de-Sitter,” *JHEP* **0304**, 021 (2003) [arXiv:hep-th/0106112].
- [35] J. M. Maldacena, G. W. Moore, N. Seiberg and D. Shih, “Exact vs. semiclassical target space of the minimal string,” *JHEP* **0410**, 020 (2004) [arXiv:hep-th/0408039].
- [36] M. Mariño, “The uses of Whitham hierarchies,” *Prog. Theor. Phys. Suppl.* **135**, 29 (1999) [arXiv:hep-th/9905053].
- [37] M. Mariño, “Chern-Simons theory, matrix integrals, and perturbative three-manifold invariants,” *Commun. Math. Phys.* **253**, 25 (2004) [arXiv:hep-th/0207096].
- [38] M. Mariño, “Les Houches lectures on matrix models and topological strings,” arXiv:hep-th/0410165.
- [39] M. Mariño, “Open string amplitudes and large order behavior in topological string theory,” *JHEP* **0803**, 060 (2008) [arXiv:hep-th/0612127].
- [40] M. Mariño, “Nonperturbative effects and nonperturbative definitions in matrix models and topological strings,” arXiv:0805.3033 [hep-th].
- [41] M. Mariño, R. Schiappa and M. Weiss, “Nonperturbative Effects and the Large-Order Behavior of Matrix Models and Topological Strings,” arXiv:0711.1954 [hep-th].
- [42] M. Mariño, R. Schiappa and M. Weiss, “Multi-Instantons and Multi-Cuts,” arXiv:0809.2619 [hep-th].
- [43] E. J. Martinec, “The annular report on non-critical string theory,” arXiv:hep-th/0305148.
- [44] M. Matone, “Instantons and recursion relations in N=2 SUSY gauge theory,” *Phys. Lett. B* **357**, 342 (1995) [arXiv:hep-th/9506102].
- [45] H. Ooguri, A. Strominger and C. Vafa, “Black hole attractors and the topological string,” *Phys. Rev. D* **70**, 106007 (2004) [arXiv:hep-th/0405146].
- [46] M. Sato and Y. Sato, “Soliton equations as dynamical systems on infinite-dimensional Grassmann manifolds,” *Lect. Notes in Num. Appl. An.* **5** (1982) 259.
- [47] E. Witten, “Quantum background independence in string theory,” arXiv:hep-th/9306122.