On Geometry and Matrix Models

Robbert Dijkgraaf
Institute for Theoretical Physics &
Korteweg-de Vries Institute for Mathematics
University of Amsterdam
1018 TV Amsterdam, The Netherlands

and

Cumrun Vafa
Jefferson Physical Laboratory
Harvard University
Cambridge, MA 02138, USA

Abstract
We point out two extensions of the relation between matrix models, topological strings and $\mathcal{N} = 1$ supersymmetric gauge theories. First, we note that by considering double scaling limits of unitary matrix models one can obtain large $N$ duals of the local Calabi-Yau geometries that engineer $\mathcal{N} = 2$ gauge theories. In particular, a double scaling limit of the Gross-Witten one-plaquette lattice model gives the $SU(2)$ Seiberg-Witten solution, including its induced gravitational corrections. Secondly, we point out that the effective superpotential terms for $\mathcal{N} = 1$ $ADE$ quiver gauge theories is similarly computed by large $N$ multi-matrix models, that have been considered in the context of $ADE$ minimal models on random surfaces.
The associated spectral curves are multiple branched covers obtained as Virasoro and $W$-constraints of the partition function.
1. Introduction

In [1] we have shown how the effective superpotential in $\mathcal{N} = 1$ supersymmetric gauge theories, that are obtained by breaking an $\mathcal{N} = 2$ super Yang-Mills theory by adding a tree-level superpotential $W(\Phi)$ for the adjoint scalar $\Phi$, can be computed by a large $N$ hermitian matrix models. More precisely, the effective superpotential of the $\mathcal{N} = 1$ theory considered as a function of the gluino condensates $S_i$ is—apart from the universal $S_i \log(S_i/\Lambda)$ terms coming from the pure $\mathcal{N} = 1$ Yang-Mills theory—given exactly by a perturbative series that is computed by the planar diagrams of the matrix model with potential $W(\Phi)$. Furthermore, the contributions of this gauge theory to the induced supergravity corrections $R^2 F^{2g-2}$ (with $R$ the Riemann curvature and $F$ the graviphoton field strength) are similarly computed exactly by the genus $g > 0$ matrix diagrams.

This gauge theory/matrix model correspondence was a consequence of the large $N$ dualities of [2,3,4] that relate the computation of holomorphic F-terms in the world-volume theories of D-branes to partition functions of topological strings in local Calabi-Yau geometries—a relation that was further explored in [5,6,7,8,9,10,11]. In the simplest case these local non-compact Calabi-Yau manifolds take the form

$$ vv' + y^2 - W'(x)^2 + f(x) = 0. $$

One finds that in the B-model topological string the tree-level free energy can be computed in terms of the periods of the meromorphic differential $y\,dx$ on the associated Riemann surface

$$ y^2 - W'(x)^2 + f(x) = 0. $$

As we argued in [1] this curve and the associated special geometry arises naturally from the large $N$ dynamics of the matrix integral with action $W(\Phi)$. But we should stress again that the relation with matrix models goes beyond the planar limit. The higher genus string partition functions $\mathcal{F}_g$ and the related gravitational couplings of the gauge theories are exactly computed in the $1/N$ expansion of the matrix models.

Let us briefly summarize these connections, for more details see [1]. We start with the hermitian matrix integral

$$ \int d\Phi \cdot e^{-S(\Phi)} $$

with action

$$ S(\Phi) = \frac{1}{g_s} \text{Tr} W(\Phi), $$

$$ 2 $$
and $W(x)$ is a polynomial of degree $n + 1$. The matrix integral can be reduced to an integral over the eigenvalues $x_1, \ldots, x_N$ of $\Phi$ in the potential $W(x)$. In the classical limit $g_s \to 0$, where one ignores the interactions among the eigenvalues, the equation of motion is given by

$$y(x) = g_s \frac{\partial S}{\partial x} = W'(x) = 0. \quad (1.1)$$

The associated classical spectral curve is

$$y^2 - W'(x)^2 = 0, \quad (1.2)$$

where $x, y$ can be considered as complex variables. Writing $W'(x) = \prod_i (x - a_i)$ we see that this singular genus zero planar curve has $n$ double points at the critical points $x = a_i$.

Sometimes it can be helpful to think of the $(x, y)$-plane as a phase space, with $y$ the momentum conjugate to $x$, as given by the Hamilton-Jacobi equation (1.1). Then (1.2) has an interpretation as the zero-energy level set of the (bosonic part of the) supersymmetric quantum mechanics Hamiltonian associated to the superpotential $W(x)$, and $S(x)$ can be thought of as the semi-classical WKB action of the associated quantum mechanical ground state $\Psi(x) \sim e^{-S(x)}$.

Classically, the $N$ eigenvalues will cluster in groups of $N_i$ in the critical points $a_i$ where they will form some meta-stable state. The relative number of eigenvalues or filling fraction of the critical point $a_i$ we will denote as

$$\nu_i = \frac{N_i}{N}.$$ 

If $g_s$ is not zero, we have to take into account the Coulomb interaction that results from integrating out the angular, off-diagonal components of the matrix $\Phi$. The equation of motion of a single eigenvalue $x$ in the presence of the Dyson gas of eigenvalues $x_1, \ldots, x_N$ is now modified to

$$y = W'(x) - 2g_s \sum_{I=1}^N \frac{1}{x - x_I}.$$ 

We will now take the large $N$ 't Hooft limit keeping both $\mu = g_s N$ and the filling fractions $\nu_i$ fixed. In this case each critical point has its own 't Hooft coupling

$$\mu_i = g_s N_i = \mu \nu_i.$$ 

The collective dynamics of these eigenvalues in the large $N$ limit can be summarized geometrically as follows. Each of the $n$ double points $x = a_i$ gets resolved into two
branch points $a_i^+, a_i^-$. The resulting branch cuts $A_i = [a_i^-, a_i^+]$ are filled by a continuous density of eigenvalues that behave as fermions and spread out due to the Pauli exclusion principle. This process of splitting up of double points is very analogous to transition from the classical to the quantum moduli space in the Seiberg-Witten solution of $\mathcal{N} = 2$ supersymmetric gauge theories [12]—a relation that was explained in [11]. The resolution of double points is captured by deforming the classical spectral curve (1.2) into the quantum curve

$$y^2 - W'(x)^2 + \mu f(x) = 0, \quad (1.3)$$

where the quantum deformation $f(x)$ is a polynomial of degree $n - 1$. The filling fraction $\nu_i$ is related to the size of the branch cut $A_i$. Roughly, the higher the proportion of eigenvalue at the critical point $a_i$, the larger the cut. More precisely, we have

$$\mu_i = \mu \nu_i = \frac{1}{2\pi i} \oint_{A_i} y(x) dx$$

(This equation can be considered as analogous to the Bohr-Sommerfeld quantization condition.) Finally the tree-level free energy $\mathcal{F}(\mu_i)$ can be computed in terms of the dual $B$-periods by the special geometry relations

$$\frac{\partial \mathcal{F}}{\partial \mu_i} = \oint_{B_i} y(x) dx,$$

where the cycles $B_i$ run from the branch cuts to some cut-off point at infinity.

### 2. Unitary matrix models and the Seiberg-Witten solution

We have reviewed how the effective superpotential in $\mathcal{N} = 1$ supersymmetric gauge theories obtained by breaking an $\mathcal{N} = 2$ super Yang-Mills theory can be computed by a large $N$ hermitian matrix model. This raises the question whether there exists a matrix model that computes directly holomorphic F-terms in the underlying undeformed $\mathcal{N} = 2$ theory as described by the rigid special geometry of the Seiberg-Witten solution [12]. As explained in [11] it is indeed possible to extract the geometry of the $\mathcal{N} = 2$ solution from the effective superpotential of the $\mathcal{N} = 1$ deformation of $\mathcal{N} = 2$ theory and thereby, indirectly, from the associated matrix model. But here we will pursue a more direct relation—we will show that indeed there is a large $N$ matrix model that computes the SW solution directly. The matrix model for $\mathcal{N} = 2$ supersymmetric $SU(2)$ gauge theory turns out to be a double scaling limit of the most simple unitary matrix model—the so-called one-plaquette or Gross-Witten model [13].
2.1. Geometrical engineering of $\mathcal{N} = 2$ theories

Our strategy will be the following. The $\mathcal{N} = 2$ theory can be geometrically engineered by taking a suitable limit of type IIB string theory on a local Calabi-Yau [14,15,16,17]. This local CY produces directly the Seiberg-Witten curve that encodes the dynamics of the $\mathcal{N} = 2$ gauge theory. More precisely the genus zero topological B-model amplitudes on the local CY capture the Seiberg-Witten geometry. The higher genus amplitudes compute the contributions of the gauge theory to certain gravitational terms of the form $R^2 F^{2g-2}$ [18,19]. We will now engineer a matrix model that is large $N$ dual to this local CY geometry. In particular the planar limit gives the SW geometry and the $1/N$ corrections capture the generation of the corresponding gravitational terms.

To be specific, let us discuss here the simplest case of pure $SU(2)$ $\mathcal{N} = 2$ Yang-Mills theory. The SW solution is given in terms of the familiar elliptic curve

$$w^2 = (y^2 + u)^2 - \Lambda^4,$$

where $u$ is the coordinate on the moduli space, i.e. the vev of the adjoint $\frac{1}{2}\langle \text{tr}\Phi^2 \rangle$, and $\Lambda$ the gauge theory scale, that we will sometimes set conveniently to $\Lambda = 1$.

The local CY obtained in the geometric engineering is given by the algebraic variety [15,16,17]

$$vv' + \Lambda^2 (z + \frac{1}{z}) + 2(y^2 + u) = 0$$

with $z \in \mathbb{C}^*$, i.e. an invertible variable. After reducing over the $(v,v')$-plane the associated Riemann surface is

$$\Lambda^2 (z + \frac{1}{z}) + 2(y^2 + u) = 0.$$  

Since $z \neq 0$ we can multiply by $z$ and substituting $w = \Lambda^2 z + y^2 + u$ to bring the curve in the form (2.1).

Furthermore, the reduction of the holomorphic three-form on the CY gives directly the SW differential $ydz/z$. The prepotential $\mathcal{F}(u)$ is then obtained by computing the periods of this meromorphic one-form along the $A$-cycle and $B$-cycle of the elliptic curve (2.1). Note that there are four branch points at $y = \pm \sqrt{-u \pm \Lambda^2}$. In the classical limit $\Lambda \to 0$ they coalesce pairwise in two double points $y = \pm \sqrt{-u}$. The moduli space contains two singularities at $u = \pm 1$ where monopoles respectively dyons become massless.

Since $z$ is a $\mathbb{C}^*$ variable, it makes sense to write $z = e^{ix}$ with the variable $x$ periodic modulo $2\pi$, and reexpress the original local CY and the resulting curve (2.2) as

$$\Lambda^2 \cos x + y^2 + u = 0,$$

5
with SW differential $ydx$. This way of writing the equation suggests a relation to a matrix model with some suitable potential $W(x)$. In fact, since $x$ is now a periodic variable this suggests a unitary matrix model where $z = e^{ix}$ will get interpreted as the eigenvalue of a unitary matrix $U$.

2.2. Unitary matrix models

Unitary matrix models are defined as integrals over the group manifold $U(N)$ of the form

$$Z = \frac{1}{\text{Vol}(U(N))} \int_{U(N)} dU \cdot \exp\left(-\frac{1}{g_s} \text{Tr} W(U)\right),$$

(2.3)

where $dU$ is the Haar measure. As in the hermitian matrix models, one can diagonalize $U$ and express everything in integrals over its eigenvalues

$$U \sim \text{diag}(e^{i\alpha_1}, \ldots, e^{i\alpha_N}),$$

where the $\alpha_i$ are periodic variables, giving

$$Z = \int d\alpha_1 \cdots d\alpha_N \prod_{i<j} \sin^2\left(\frac{\alpha_i - \alpha_j}{2}\right) \exp\left(-\frac{1}{g_s} \sum_i W(\alpha_i)\right).$$

Note that such a unitary matrix model can be viewed as a special case of a hermitian model by writing $U = e^{i\Phi}$ with $\Phi$ a ‘compactified’ hermitian matrix, i.e. a matrix with a periodic spectrum $\Phi \sim \Phi + 2\pi$. Such a periodicity is achieved by adding multiples of $2\pi$ to the eigenvalues $\alpha_I$ of $\Phi$. This addition of multiple images is the familiar way to compactify transverse directions for D-branes or for matrix models in M-theory [20]. For example, in this way the Vandermonde determinants in the measure become after regularization

$$\prod_{n \in \mathbb{Z}} (\alpha_I - \alpha_J + 2\pi n) = \sin\left(\frac{\alpha_I - \alpha_J}{2}\right).$$

The unitary matrix model describes a collection of particles in a potential $W(\alpha)$ on the unit circle interacting through a Coulomb potential. The equation of motion of the unitary model is

$$W'(\alpha_I) - 2g_s \sum_j \cot\left(\frac{\alpha_I - \alpha_J}{2}\right) = 0$$

The large $N$ solution proceeds exactly as in the uncompactified case. One introduces again a resolvent, this time defined as

$$\omega(x) = -\frac{1}{N} \sum_I \cot\left(\frac{x - \alpha_I}{2}\right),$$
that satisfies a quadratic loop equation that can be derived in exactly the same way as in
the hermitian case. In the limit $N \to \infty$ with $g_s N = \mu$ fixed this loop equation takes
the following familiar form, when written in terms of the variable $y = W'(x) + 2\mu \omega(x),$

$$y^2 - W'(x)^2 + 4\mu f(x) = 0.$$ \hfill (2.4)

The quantum correction $f(x)$ in this unitary case is given by the expression

$$f(x) = \frac{1}{N} \sum I (W'(x) - W'(\alpha_I)) \cot \left( \frac{x - \alpha_I}{2} \right)$$ \hfill (2.5)

2.3. The Gross-Witten model

Our candidate unitary matrix model will be a much studied one, namely the so-called Gross-Witten model \cite{13} with potential $W(\alpha) \sim \cos \alpha$. This model was originally introduced as a lattice discretization of two-dimensional (non-supersymmetric) Yang-Mills theory. In such a lattice model to each plaquette with holonomy $U$ around the edge one associates the Wilson action (in other contexts known as the Toda potential)

$$S(U) = \frac{\epsilon}{2g_s} \Tr(U + U^{-1}) = \frac{\epsilon}{g_s} \Tr \cos(\Phi)$$

Here $\Phi$ can be thought of as the lattice approximation to the gauge field strength $F_{\mu\nu}$, and in the limit $\Phi \to 0$ this gives the quadratic Yang-Mills action $\Tr \Phi^2$. (The parameter $\epsilon$ we introduce for convenience. It can of course be absorbed by rescaling $g_s = g_{YM}^2$). In a general lattice model one integrates over a collection of plaquettes, but in two dimensions a single plaquette suffices to compute for instance a Wilson loop action.

The GW model has two critical points

$$W'(\alpha) = -\epsilon \sin(\alpha) = 0$$

at $a_1 = 0$ and $a_2 = \pi$. Note that for real and positive $\epsilon$ (the case relevant for Yang-Mills theory) the second point is an unstable critical point. But that issue is irrelevant for the holomorphic matrix models that we are considering here. The parameter $\epsilon$ can be complex, and our eigenvalues $e^{i\alpha}$ are allowed to move off the unit circle into the punctured complex plane $\mathbb{C}^*$. In fact, following our general philosophy, we will consider the perturbative expansion of the matrix integral (2.3) around the saddle point where $N_1$ eigenvalues are at the first critical point $a_1$ and $N_2 = N - N_1$ are at the second point $a_2$. So we are dealing with a two-cut, meta-stable solution to the matrix integral. These two cuts introduce a
second parameter, besides the overall ’t Hooft coupling $\mu = g_s N$, namely the relative filling fraction

$$\nu = (N_1 - N_2)/N.$$ 

If we introduce the separate ’t Hooft couplings for the two critical points

$$\mu_1 = g_s N_1, \quad \mu_2 = g_s N_2,$$

then the difference of these couplings is related to the filling fraction

$$\mu' = \mu_1 - \mu_2 = g_s (N_1 - N_2) = \mu \cdot \nu.$$ 

We will be interested in computing the planar limit of the free energy $F$ as a function of the two coupling $\mu_1$ and $\mu_2$, or equivalently as a function of the ’t Hooft coupling $\mu$ and the filling fraction $\nu$. Although $\nu$ takes values in the interval $[-1, 1]$ the final result will turn out to be a holomorphic function of $\nu$.

The planar limit can be computed solving the loop equation (2.4). To this end we have to compute the quantum correction $f(x)$ defined in (2.5). For our choice of potential $W(x) = \epsilon \cos x$ this becomes the following average over the eigenvalues

$$f(x) = \frac{1}{N} \sum_I \epsilon (\sin x - \sin \alpha_I) \cot\left(\frac{x - \alpha_I}{2}\right)$$

$$= -\frac{1}{N} \sum_I \epsilon (\cos x + \cos \alpha_I)$$

$$= -\epsilon (\cos x + u) \quad (2.6)$$

Here the constant $u$ is defined as the average

$$u = \frac{1}{N} \sum_I \cos \alpha_I.$$ 

In the semi-classical approximation $g_s \rightarrow 0$ we have

$$u \approx \frac{N_1 - N_2}{N} = \nu.$$ 

since there are $N_1$ eigenvalues at the critical point $\alpha_I = 0$ and $N_2$ eigenvalues at $\alpha_I = \pi$, that contribute respectively $+1$ and $-1$ to the average of $\cos \alpha_I$.

Inserting our expression for $f(x)$ into (2.4) gives the spectral curve

$$y^2 - \epsilon^2 \sin^2 x + 4 \mu \epsilon (\cos x + u) = 0. \quad (2.7)$$
We see that the original double points $a_1 = 0$ and $a_2 = \pi$ now split up in four branch points $a_1^\pm, a_2^\pm$. The two branch cuts

$$A_1 = [a_1^-, a_1^+], \quad A_2 = [a_2^-, a_2^+]$$

describe the condensation of eigenvalues around the two critical points. The discontinuities in $y$ give the eigenvalue density $\rho(\alpha)$. Integrating the one-form $ydx$ around the branch cuts $A_i$ gives the filling fractions $N_i/N$. These relations allow one to express the filling fraction, or more precisely the relative coupling $\mu' = \mu \nu$, in terms of the period

$$\mu' = \frac{1}{2\pi i} \oint_A y(x)dx$$

where $A$ is a one-cycle homologous to $A_1 - A_2$. This equation describes the exact relation between the variables $\mu'$ and $u$. Integrating the same differential over the conjugated $B$-cycle encircling the cut $[\alpha_+^{(1)}, \alpha_-^{(2)}]$ computes the variation of the free energy $\mathcal{F}$ under a change of the relative number of eigenvalues in the two cuts,

$$\frac{\partial \mathcal{F}}{\partial \mu'} = \int_B y(x)dx.$$

Note that the GW solution assumes that all eigenvalues center around the stable vacuum $a_1 = 0$, in which case $N_2 = 0$ and $\nu = 1$. In fact we claim that for this vacuum we have exactly $u = 1$, so that the spectral curve is given by

$$y^2 - \epsilon^2 \sin^2 x + 4\mu\epsilon(\cos x + 1) = 0.$$

Indeed, since one only fills the stable critical point $a_1 = 0$ with eigenvalues, only this double point will get resolved into two branch points. The second critical point $a_2 = \pi$ remains unresolved. Inserting $u = 1$ gives the GW solution for the large $N$ eigenvalue density that can be written as in [13]

$$\rho(\alpha) \sim \cos\left(\frac{\alpha}{2}\right) \sqrt{\frac{\mu}{2\epsilon} - \sin^2\left(\frac{\alpha}{2}\right)}.$$
2.4. Double scaling limit

Now we will take a double scaling limit of the GW model to obtain the SW solution relevant for $\mathcal{N} = 2$ supersymmetric gauge theory. In this limit we will send $N \to \infty$ and at the same time $\epsilon \to 0$ and $\nu \to 0$, keeping $\epsilon N, \nu/\epsilon$ and $g_s$ fixed—or, equivalently, we will send the ’t Hooft coupling $\mu \to \infty$ keeping the difference of the two couplings $\mu_1$ and $\mu_2$

$$\mu' = g_s(N_1 - N_2) = \mu \nu$$

fixed. In this limit the absolute difference in eigenvalues $N_1 - N_2$ remains finite, but $N_1$ and $N_2$ become both infinite, and therefore the relative filling fraction $\nu = (N_1 - N_2)/N$ goes to zero.

After rescaling $y$ appropriately, the spectral curve reduces in this limit exactly to the SW curve (with $\Lambda = 1$)

$$y^2 + \cos x + u = 0.$$ 

Note that the double scaled curve depends only on a single parameter $u$ that at weak coupling could be identified with the filling fraction $\nu = (N_1 - N_2)/N$ of the matrix model. Of course the limit we are taking is at a strong coupling point and the relation between $u$ and matrix model modulus $\mu'$ is more complicated, as discussed above.

As we have already mentioned, the prepotential is now computed by the periods of the differential $ydx = ydz/z$ along the $A$ and $B$ cycles. This allows us to identify the SW periods as

$$a = \oint_A y(x)dx = \mu' \tag{2.8}$$

$$a_D = \oint_B y(x)dx = \frac{\partial F}{\partial \mu'}$$

From the original four branch points $a_1^\pm, a_2^\pm$ obtained in resolving the two double points $a_1, a_2$, our double scaling limit takes the branch points $a_1^-, a_2^+$ to infinity while keeping $a_1^+, a_2^-$ at finite distance. This leaves two homology one-cycles: the $B$-cycle that runs around the cut $[a_1^+, a_2^-]$ and the dual $A$-cycle that is homologous to $A_1 - A_2$. This behaviour of the branch points is exactly the behaviour in the double scaling limit one takes in the old matrix models [21,22], as we also noted in [1]. For example, the $(2,3)$ critical point of the one-matrix model was obtained starting from a curve of the form $y^2 = x^6 + \ldots$ and the double scaling limit got rid of the all monomials with power more than $x^3$ giving an equation of the form $y^2 = x^3 + \ldots$ (Note that these branch points in the
The double scaling is very analogous to the limit that was used in the A-model topological string in [23]. In fact, when the scaling limit is embedded in type II string theory, the resulting CY geometry based on (2.7) will have RR flux through the compact cycles \(A_1\) and \(A_2\). It is crucial that the remaining compact cycle \(B\) does not carry any Ramond flux. We are thus engineering a large \(N\) dual of a geometry without fluxes.

The GW model has a famous third order phase transition at (in our convention) \(\mu/\epsilon = 2\). This signals a transition of the eigenvalue distribution in which the single cut changes topology and starts to cover the whole unit circle. After the GW phase transition the eigenvalue distribution is given by \(\rho(\alpha) \sim \cos \alpha + \mu/2\). Geometrically speaking, in that phase all four branch points are on top of each other.

This phase transition is however not relevant in our model. First of all, we are studying a more general question by considering a stationary phase approximation around a metastable state with two clusters of eigenvalues and consequently have to work with a two-dimensional phase diagram \((\mu, \nu)\). As we argued the GW solution puts \(\nu = 1\) and that is very far away from our double scaling limit in which \(\nu\) tends to zero. Indeed in our limit the number of eigenvalues in the two cuts is roughly equal. Secondly, we are dealing with an holomorphic object, and holomorphy excludes any phase transitions, one can just go around the singularity. The GW phase transition is just a (very special) real slice of our complex phase diagram.

Finally it would be interesting to connect this approach to the beautiful semi-classical computation of the SW solution and its gravitational counterparts in [24]. That computation was inspired by matrix integrals appearing in D-branes formulas.

2.5. Generalizations for general groups

It is not difficult to guess how the SW solution for gauge group \(SU(n)\) can be engineered. In this case the curve associated to the local CY is of the form

\[
\cos x + P_n(y) = 0, 
\]

with \(P_n(y)\) a polynomial of degree \(n\). More generally we can consider a chain of \(U(n_i)\) gauge theories with bifundamental matter, for which the corresponding curve has been obtained.
from the M5 brane viewpoint in [25] and from the viewpoint of geometric engineering in [17]. This will give rise to a curve of the form

\[ F(e^{ix}, y) = 0 \]

where \( F \) is a polynomial in \( e^{\pm ix} \) and \( y \). In particular if we consider the rank of all the gauge groups to be equal to \( n \), then \( F \) is a polynomial in \( y \) of degree \( n \). Moreover the difference in power of \( e^{ix} \) between the highest and lowest powers is the number of \( U(n) \) gauge groups plus 1. As we will discuss in greater detail in the next section in the context of hermitian multi-matrix models such curves are typically produced by a multi-matrix model consisting of \( n - 1 \) matrices. The choice of the coefficients in \( F \) will be related to the choice of the action and some suitable double scaling limit, as we studied in the context of \( SU(2) \) gauge theory here.

2.6. Connections with A-model topological strings

It is natural to ask if there is a connection with A-model topological strings, and in particular for A-models on local toric CY. As was demonstrated in [26,27] the B-model mirrors are given by

\[ vv' + F(e^{ix}, e^{iy}) = 0 \]

for some \( F \). This is analogous to an infinite matrix model version of the unitary matrix models, as follows from our discussion above. In this case the A-model has a gauge theory dual involving certain correlations functions of Chern-Simons theory [23,28]. It would be interesting to connect these matrix models directly with the Chern-Simons gauge theory computation, thus completing the circle of ideas. There are some hints that this idea indeed works [29].

3. Quiver matrix models

We will now turn to a related generalization of [1] where we will connect superpotential computations of quiver gauge theories to multi-matrix models.
3.1. $\mathcal{N} = 1$ quiver gauge theories and topological strings

We will restrict our discussion here to the $ADE$ quivers, in particular the $A_r$ case, although one can also include the affine quivers based on the extended Dynkin diagrams $\hat{A}\hat{D}\hat{E}$. Let $r$ denote the rank of the quiver $G$ and consider a partition

$$N = N_1 + \ldots + N_r.$$  

In the associated $\mathcal{N} = 2$ quiver gauge theory we assign to each of the $r$ vertices $v_i$ of the Dynkin diagram of $G$ a $U(N_i)$ gauge field and to links connecting vertices $v_i$ and $v_j$ we associate bifundamentals $Q_{ij}$ transforming in the representation $(N_i, \overline{N}_j)$ with a hermiticity condition $Q_{ij}^\dagger = Q_{ji}$.

For such a gauge theory we can write a general tree-level superpotential

$$W(\Phi, Q) = \sum_{i,j} s_{ij} \text{Tr} Q_{ij} \Phi_j Q_{ji} + \sum_i \text{Tr} W_i(\Phi_i),$$  

with $s_{ij} = -s_{ji} = 1$ (for some ordering $i < j$), if the vertices $v_i$ and $v_j$ are linked in the Dynkin diagram (we will write this relation also as $\langle i, j \rangle$), and $s_{ij} = 0$ otherwise. Here the first term is the standard superpotential of the $\mathcal{N} = 2$ theory with bifundamental matter. The additional potentials $W_i(\Phi_i)$ are introduced to break the supersymmetry down to $\mathcal{N} = 1$.

Within type II string theory these quiver gauge theories are obtained by wrapping D5-branes over a particular CY geometry that is a fibration of the corresponding $ADE$ singularity over the complex plane [7]. This geometry contains $r$ intersecting $\mathbf{P}^1$’s. According to [7,8] in the large $N$ limit the geometry undergoes a transition to a deformed geometry where these $\mathbf{P}^1$’s are blown down and a number of $S^3$’s with RR flux are “blown up.” The corresponding smooth CY geometry gives a dual description of the gauge theory system.

In the context of B-model topological strings, the deformed CY geometry is dual to a two-dimensional large $N$ gauge system, obtained from a collection of B-branes wrapped on the intersecting $\mathbf{P}^1$’s. The world-volume theory of these branes consists of open topological strings. So each $\mathbf{P}^1$ gives rise to a two-dimensional field theory with Lagrangian [30]

$$S(\Phi) = \frac{1}{g_s} \int_{\mathbf{P}^1} \text{Tr} \left( \Phi_i^1 \overline{D}_A \Phi_i^0 + W_i(\Phi_i^0) \omega \right).$$  

(3.2)
where \( \omega \) is some volume form on \( \mathbb{P}^1 \), and \( \Phi_i^0 \) and \( \Phi_i^1 \) are adjoint fields of respectively spin 0 and spin 1 coupled to an \( U(N_i) \) holomorphic gauge field. Here we included the effect of the superpotential \( W_i(\Phi_0) \). The open topological strings connecting different \( \mathbb{P}^1 \)’s give as physical fields the bifundamentals \( Q_{ij} \). Since the different \( \mathbb{P}^1 \)’s intersect in points, the action of these bifundamental scalar fields localizes to the intersection point \( x \) and is given by

\[
S(Q) = \sum_{(i,j) \in \mathbb{P}_i^1 \cap \mathbb{P}_j^1} \sum_{x} \text{Tr}(Q_{ij}(x)\Phi_j^0(x)Q_{ji}(x) - Q_{ji}(x)\Phi_i^0(x)Q_{ij}(x)).
\]

(Compare the similar computation for the coupling of open topological strings connecting Lagrangians A-branes intersection along one-dimensional curves in [31].)

As in [1] one can see that in the end this two-dimensional topological field theory can be completely reduced to the zero modes of the fields \( \Phi_i^0(x) = \Phi_i \) and \( Q_{ij} \). Thereby the path-integral reduces to the “quiver matrix integral”

\[
Z = \int \prod_i d\Phi_i \prod_{(i,j)} dQ_{ij} \exp \left( -\frac{1}{g_s} \text{Tr} W(\Phi, Q) \right),
\]

(3.3)

where \( \Phi_i \) (with \( i = 1, \ldots, r \)) is an \( N_i \times N_i \) hermitian matrix and for every linked indices \( (i,j) \) the variable \( Q_{ij} \) is a \( N_i \times N_j \) rectangular matrix, satisfying \( Q_{ij}^\dagger = Q_{ji} \).

The generalization of the conjecture in [1] will now identify the free energy of the large \( N \) quiver matrix model, for given filling fractions of the saddle points, with the closed topological string partition function in the corresponding deformed CY geometry, and this in turn with the effective superpotential of the quiver gauge theory.

### 3.2. Saddle points and dual CY geometry

The saddle points of the quiver superpotential have been discussed extensively in [7,8,9] following the mathematical literature. The eigenvalues \( x \) of the adjoints \( \Phi_i \) have to satisfy a series of equations: one for every positive roots \( \alpha_k \) of \( G \). If that root is expressed in the simple roots \( e_i \) as

\[
\alpha_k = \sum_i n_k^i e_i,
\]

then the associated condition reads

\[
\sum_i n_k^i W_i'(x) = 0.
\]

(3.4)
The saddle points can be labeled as $x_{a,k}$ with $a = 1, \ldots, d$ and $d$ the maximal degree occurring in (3.4). If such a critical point appears with multiplicity $N_{a,k}$ then the total number of eigenvalues of the matrix $\Phi_i$ in this saddle point is $N_{a,k} \cdot n_k^i$. A general saddle point is therefore parametrized by the filling fractions $\nu_{a,k} = N_{a,k}/N$. We will consider the matrix integral in the limit where both the $N_{a,k}$ and $N$ tend to infinity keeping the filling fractions and the 't Hooft coupling finite.

In the case of an $A_r$ quiver there is a more straightforward description of the saddle points [7,8,9]. Introduce the $r + 1$ potentials

$$t_0(x) = 0, \quad t_i(x) = \sum_{j=1}^{i} W'_j(x), \quad i = 1, \ldots, r.$$ 

The associated rational planar curves

$$y - t_i(x) = 0$$

intersect in various double points given by

$$t_j(x) - t_i(x) = \sum_{k=i+1}^{j} W'_k(x) = 0.$$ 

The saddle points of the $A_r$ quiver matrix potential correspond exactly to these double points.

In this $A_r$ case the original singular CY geometry is given by

$$uv + \prod_{i=0}^{r} (y - t_i(x)) = 0$$

which after reduction over $u,v$ gives precisely this collection of nodal curves. After the deformation the corresponding smooth Riemann surface is given by

$$\prod_{i=0}^{r} (y - t_i(x)) + f(x, y) = 0. \quad (3.5)$$

for a suitable normalizable quantum deformation $f(x, y)$. Again every double point gets resolved into two branch points. The resulting quantum curve is now an $r + 1$ fold cover of the $x$-plane. By moving around in the $x$-plane these sheets will be exchanged through Weyl reflections acting on the parameters $t_i$. 

15
The analogues of the meromorphic one-form are constructed by reducing the holomorphic three-form of the local CY over various cycles and for $A_r$ have the following description [7,8]. Write the curve (3.5) in the factorized form
\[ \prod_{i=0}^{r} (y - a_i(x)) = 0. \]
Then we have a basis of $r-1$ meromorphic one-forms $\eta_1, \ldots, \eta_{r-1}$ that is in one-to-one correspondence with a basis of positive roots for $A_{r-1}$ given by
\[ \eta_i = (a_{i+1}(x) - a_i(x)) dx. \] (3.6)
In the undeformed case, with $a_i(x) = t_i(x)$, this gives $\eta_i = W'_i(x) dx = dW_i$. Different choices of positive roots correspond to Weyl reflections, which are in fact the generalized Seiberg dualities of gauge theories (which should also have some direct interpretation as dualities of the matrix models). Note that in the case of $A_1$ we have
\[ (y - a)(y + a) = 0, \quad a^2 = W'(x)^2 + f(x) \]
and this gives the usual definition $\eta = \pm y dx$.

We will argue that this process of smoothing out the singular curve including the meromorphic differentials is exactly described by the large $N$ dynamics of the quiver matrix integral.

3.3. ADE matrix models

Quite remarkably it turns out that the quiver matrix integrals (3.3) (up to some minor details) have already been studied in the context of the “old matrix models.” They have been used to describe the coupling of ADE conformal minimal models to two-dimensional gravity by Kostov [32], see also the reviews [33,34], and they have naturally emerged in the study of matrix models and integrable systems in the work of the ITEP group [35]. We will follow closely these works in presenting the main results, leaving the details to the literature.

First of all, one can immediately integrate out the bifundamental fields $Q_{ij}$ in the quiver matrix integral to give an effective interaction between the adjoint fields $\Phi_i$ and $\Phi_j$
\[ \det(\Phi_i \otimes 1 - 1 \otimes \Phi_j)^{-1} \]
After expressing everything in terms of the eigenvalues of the remaining hermitian matrices $\Phi_i$

$$\Phi_i \sim \text{diag}(\lambda_{i,1}, \ldots, \lambda_{i,N_i})$$

the quiver matrix integral reduces to

$$Z = \int \prod_{i,I} d\lambda_{i,I} \cdot \frac{\prod_{i,I,J} (\lambda_{i,I} - \lambda_{i,J})^2}{\prod_{i<J} (\lambda_{i,I} - \lambda_{j,J})^{s_{ij}}} \exp \left( -\frac{1}{g_s} \sum_{i,I} W_i(\lambda_{i,I}) \right).$$

Note that if we introduce the Cartan matrix of $G$

$$C_{ij} = 2\delta_{ij} - |s_{ij}| = e_i \cdot e_j,$$

then the quiver eigenvalue measure, that generalizes the usual “fermionic” Vandermonde determinants of the one matrix model

$$\Delta(\lambda)^2 = \prod_{I<J} (\lambda_I - \lambda_J)^2,$$

can now be written as

$$\prod_{(j,J) \neq (i,I)} (\lambda_{i,I} - \lambda_{j,J})^{e_i \cdot e_j / 2}.$$  

The resemblance to a correlation function of vertex operators is not accidental—it was in fact the main motivation to study these kind of matrix models in [35] since it allows one to express the partition function as a particular state in a two-dimensional chiral CFT, to be more precise the level one realization of the corresponding $ADE$ current algebra.

In the large $N$ limit this many-flavor Dyson gas of eigenvalues will spread out in cuts around the saddle points and will form a continuum of eigenvalues described by a series of densities functions

$$\rho_i(x) = \frac{1}{N} \sum_I \delta(x - \lambda_{i,I}).$$

The solution of the model proceeds again through the resolvents or loop operators of the matrices $\Phi_i$

$$\omega_i(x) = \frac{1}{N} \text{Tr} \left( \frac{1}{x - \Phi_i} \right) = \frac{1}{N} \sum_I \frac{1}{x - \lambda_{i,I}}.$$

The jump of $\omega_i(x)$ across a branch cut measure the eigenvalue density $\rho_i(x)$.
In fact, it is natural to work with a closely related object—the derivative of the matrix model action $S$ with respect to an eigenvalue of type $i$ evaluated at a general position $x$ in the complex plane (away from the cuts) as it interacts with all the other eigenvalues,

$$y_i(x) = g_s \partial_i S = W'_i(x) - 2\mu C^{ij}\omega_j(x).$$  \hspace{1cm} (3.7)

As we already mentioned in [1] for a multi-matrix model we want to identify the one-forms $\eta_i$ (3.6) coming from the local CY geometry with the expressions $y_i(x)dx$, up to a possible change of basis. The most powerful and general technique that can be used to relate the CY geometry to the matrix model are the loop equations.

3.4. Loop equations and collective fields

Before we discuss the loop equations of the multi-matrix models, let us first rewrite the solution of the one-matrix model as used in [1] in a more suggestive form, that is actually a standard technique in matrix model technology. Here we found among others the reviews [33,34,36] very helpful.

The resolvent $\omega(x)$ has a natural interpretation as a loop operator. More precisely, the inverse Laplace transform

$$\int \frac{dx}{2\pi} e^{ix\ell} \omega(x) = \text{Tr}(e^{\ell \Phi})$$

is the zero-dimensional analogue of the Wilson loop. The non-linear all-genus loop equation is usually written in terms of $\omega(x)$ as

$$\oint_C \frac{dz}{2\pi i} \frac{W'(z)}{x - z} \langle \omega(z) \rangle = \mu \langle \omega(x)^2 \rangle$$  \hspace{1cm} (3.8)

where $\langle \cdots \rangle$ indicates an expectation value within the matrix integral. The contour $C$ encircles all the cuts but not the point $x$. This equation is supplemented with the boundary condition $\langle \omega(x) \rangle \sim 1/x$ at infinity. The loop equation acts as a Schwinger-Dyson equation of the matrix model. It gives a recursive relation to solve for the loop operator and the free energy. In the planar limit we have large $N$ factorization $\langle \omega(x)^2 \rangle = \langle \omega(x) \rangle^2$ and the loop equation becomes algebraic.

Loop operators are closely connected to collective fields. By integrating out the angular variables the individual eigenvalues start to behave as fermions, and the collective field is essentially constructed by bosonization of these fermion fields. In [1] we have already

18
speculated that this collective field should be identified with the Kodaira-Spencer field [18] describing the closed strings moving on the local CY geometry.

For a single matrix model the collective field is defined as the chiral two-dimensional scalar field
\[ \varphi(x) = W(x) - 2g_s \sum I \log(x - \lambda_I) \]
It clearly satisfies \( \partial \varphi(x) = y(x) \) with
\[ y = W'(x) - 2g_s \sum I \frac{1}{x - \lambda_I} = W'(x) - 2\mu \omega(x) \]
So in view of (3.7) we can identify the function \( \varphi(x) \) with the action \( S(x) \) of a single eigenvalue as a function of its position \( x \) in the complex plane in the presence of the gas of other eigenvalues \( \lambda_1, \ldots, \lambda_N \). The function \( \varphi(x) \) is multi-valued in the \( x \)-plane. It has branch cuts around which it changes sign. It is therefore only properly defined on the double cover
\[ y^2 - W'(x)^2 + f(x) = 0. \tag{3.9} \]
On this Riemann surface \( \varphi(x) \) has quantized periods around the \( A \)-cycles, given by the filling numbers \( \mu_i = g_s N_i \). Since it is a chiral field the periods around the dual \( B \)-cycles are not independent and expressed by the special geometry relations as \( \partial F / \partial \mu_i \).

Note that if we work with a general, not necessarily polynomial, superpotential
\[ W(x) = \sum_{n \geq 0} t_n x^n, \]
then the expectation value of the field \( \partial \varphi(x) \) inserted in the matrix integral can be represented by a linear differential operator in the couplings \( t_n \) acting on the partition function. For example,
\[ \langle \partial \varphi(x) \rangle = \left( \sum_{n > 0} nt_n x^{n-1} - 2g_s^2 \sum_{n \geq 0} x^{-n-1} \frac{\partial}{\partial t_n} \right) Z, \]
and similarly for multi-point functions. (Here we used that the derivative \( \partial / \partial t_n \) brings down a factor \( \frac{1}{g_s} \Tr \Phi^n \).)

With this notation there is an elegant way to write the loop equations. Introduce the holomorphic stress-tensor
\[ T(x) = (\partial \varphi)^2 = \sum_n L_n x^{-n-2} \]
Then the all-genus loop equation (3.8) of the one-matrix model can be rewritten in the suggestive form

\[
\oint_{C} \frac{dz}{2\pi i} \frac{1}{x-z} \langle T(z) \rangle = 0,
\]

(3.10)

That is, the expectation value \( \langle T(x) \rangle \) has no singular terms if \( x \to 0 \). Therefore we can also express (3.10) equivalently as the Virasoro constraints [37,38]

\[ L_n Z = 0, \quad n \geq -1. \]

The derivation of the constraints in the matrix model is completely standard—it simply expresses the Ward identities following from the invariance under infinitesimal reparametrization of \( \Phi \to \Phi + \epsilon \Phi^{n+1} \) of the matrix variable \( \Phi \).

In the planar limit we can substitute the classical values for \( \partial \varphi(x) = y \) in \( T(x) \) and then equation (3.10) is a consequence of (3.9) that can now be written as

\[ T(x) = W'(x)^2 - f(x), \]

which shows that \( T(x) \) is indeed regular (even polynomial) at \( x = 0 \).

3.5. Quiver theories and W-constraints

The large \( N \) solution of the quiver matrix integral (3.3) now proceeds along similar lines [32,35]. One introduces \( r \) scalar fields \( \varphi_i(x) \) through the one-forms (3.7) as

\[ y_i(x) dx = \partial \varphi_i(x) \]

One can then show that the multi-valued fields \( \varphi_i(x) \) are actually the values of one single-valued field \( \varphi(x) \) (essentially the full matrix model action) on a \( r + 1 \) branched cover of the complex \( x \)-plane. This branched cover is the spectral curve associated to the quiver matrix integral, and turns out to be given by (3.5) in the \( A_r \) case.

The general derivation of the curve proceeds through generalized loop equations. For these multi-matrix model we do not only have the Virasoro constraints, expressing reparametrization invariance in the matrix variables \( \Phi_i \). There are also higher order relations [37,38]. The full set of loop equations are obtained by showing that the partition function \( Z \) satisfies a set of \( W \)-constraints, labeled by the Casimirs of the corresponding \( ADE \) Lie algebra, which contains the Virasoro constraints. These constraints take the form

\[
\oint_{C} \frac{dz}{2\pi i} \frac{1}{x-z} \langle W^{(s)}(z) \rangle = 0,
\]

20
where $W^{(s)}(x)$ is a spin $s$ current in the $W$-algebra. When expressed in modes these equations take the form
\[ W_n^{(s)} \cdot Z = 0, \quad n \geq 1 - s. \]

In the case of $A_r$ there is leading spin $r + 1$ current that with a suitable basis of vectors $\varphi_0, \ldots, v_r$ can be written as
\[ W^{(r+1)}(x) \sim \prod_{i=0}^{r} (v_i \cdot \partial \varphi) + \ldots \]

We claim that in the planar limit this loop equation translates directly into the curve (3.5).

To be completely explicit let us give some more detail for the simplest case of $A_2$. Here we have two matrices $\Phi_1, \Phi_2$ with potentials $W_1(\Phi_1)$ and $W_2(\Phi_2)$. The classical singular curve is after a shift in $y$ given by
\[ (y - t_1(x))(y - t_2(x))(y - t_3(x)) = 0 \]

with
\[ t_1 = -(2W'_1 + W'_2)/3, \quad t_2 = (W'_1 - W'_2)/3, \quad t_3 = (W'_1 + 2W'_2)/3, \]
all polynomials in $x$. To find the quantum curve we introduce the resolvents
\[ w_1(x) = \sum_{I} \frac{1}{x - \lambda_{1,I}}, \quad w_2(x) = \sum_{I} \frac{1}{x - \lambda_{2,I}}, \]
and the one-forms $y_i(x)dx$
\[ y_1 = W'_1 - \mu(2\omega_1 - \omega_2), \quad y_2 = W'_2 - \mu(2\omega_2 - \omega_1). \]

We now claim that the quantum curve is given by
\[ (y - a_1(x))(y - a_2(x))(y - a_3(x)) = 0, \]
where the functions $a_i(x)$ are no longer polynomials, but instead are defined as
\[ a_1 = t_1 + \mu \omega_1, \quad a_2 = t_2 - \mu(\omega_1 - \omega_2), \quad a_3 = t_3 - \mu \omega_2. \]

With this choice we have, as claimed before,
\[ a_2 - a_1 = y_1, \quad a_3 - a_2 = y_2. \]
Now after some algebra, expanding out terms like $\omega_i(x)^3$, one verifies that indeed

$$(y - a_1(x))(y - a_2(x))(y - a_3(x)) =
(y - t_1(x))(y - t_2(x))(y - t_3(x)) + f(x)y + g(x) = 0$$

with $f(x)$ and $g(x)$ polynomials.

Acknowledgements

We would like to thank J. de Boer, M. Mariño, and E. Verlinde for discussions. The research of R.D. is partly supported by FOM and the CMPA grant of the University of Amsterdam, C.V. is partly supported by NSF grants PHY-9802709 and DMS-0074329.

References

——, “Open/closed string dualities and Seiberg duality from geometric transitions in M-theory,” [arXiv:hep-th/0106040];


