hep-th/yymmnnn
CERN-TH/2003-xxx

Lectures on Chern-Simons theory, matrix models and topological strings

Marcos Mariño

Theory Division, CERN
Geneva 23, CH-1211 Switzerland

Marcos.Marino.Beiras@cern.ch

Abstract

This is a PRELIMINARY and UNFINISHED set of lectures giving a pedagogical introduction to recent developments on topological string theory and its relation to Chern-Simons theory. Your comments are welcome.

Contents

1 Introduction

2 Chern-Simons theory and knot invariants

In this section we make a short review of Chern-Simons theory and its relations to knot invariants.

2.1 Chern-Simons theory: basic ingredients

In [89], Witten showed that the quantum field theory in three dimensions whose action is solely given by the Chern-Simons term turns out to provide a physical description of a wide class of invariants of three-manifolds and of knots and links in three-manifolds.

The Chern-Simons action with gauge group G on a generic three-manifold M is defined by

$$S = \frac{k}{4\pi}\int_M \mathrm{Tr}\Big(A\wedge dA + \frac{2}{3}A\wedge A\wedge A\Big) \tag{2.1}$$

Here, k is the coupling constant, and A is a G-gauge connection on the trivial bundle over M. We will assume for simplicity that G is a simply-laced group, unless otherwise stated. As noticed by Witten, [89], since this action does not involve the metric, the resulting quantum theory is topological, at least formally. In particular, the partition function

$$Z(M) = \int [\mathcal{D}A] \mathrm{e}^{iS} \tag{2.2}$$

should define a topological invariant of the manifold M. A detailed analysis [89] shows that this is in fact the case, with an extra subtlety: the invariant depends on the three-manifold *and* of a choice of framing (*i.e.* a choice of trivialization of the bundle $TM\oplus TM$). As explained in [9], for every three-manifold there is a canonical choice of framing, and the different choices are labeled by an integer $s \in \mathbf{Z}$ in such a way that $s = 0$ corresponds to the canonical framing. In the following all the results for the partition functions will be presented in the canonical framing.

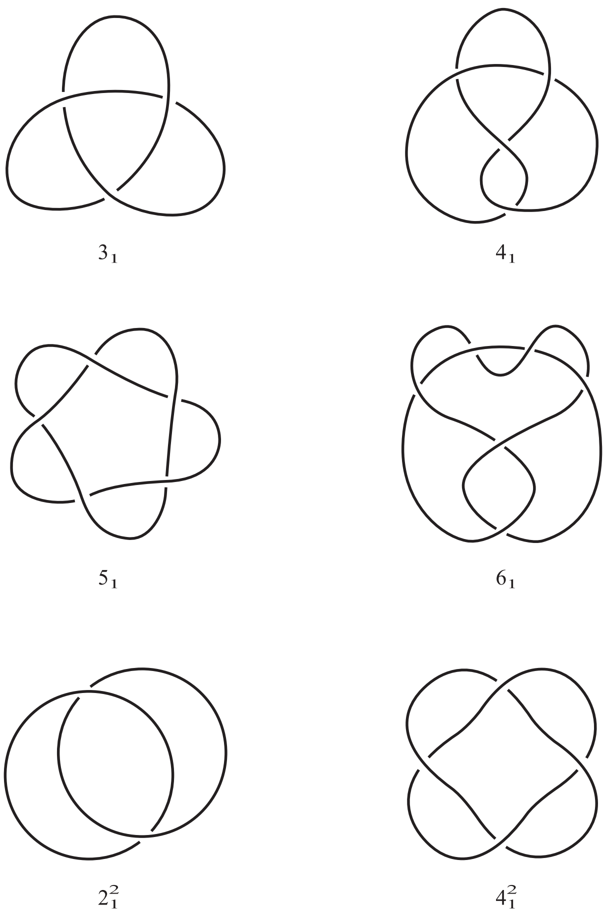


Figure 1: Some knots and links. In the notation x^L_n, x indicates the number of crossings, L the number of components (in case it is a link with $L > 1$) and n is a number used to enumerate knots and links in a given set characterized by x and L. The knot 3_1 is also known as the trefoil knot, while 4_1 is known as the figure-eight knot. The link 2^2_1 is called the Hopf link.

Besides providing invariants of three-manifolds, Chern-Simons theory also provides invariants of knots and links inside three-manifolds (for a survey of modern knot theory, see [61, 79]). Some examples of knots and links are depicted in Fig. 1. Given an oriented knot $\mathcal{K}$ in $\mathbf{S}^3$, we can consider the trace of the holonomy of the gauge connection around $\mathcal{K}$

in a given irrreducible representation R of $SU(N)$, which gives the Wilson loop operator:

$$W_R^{\mathcal{K}}(A) = \mathrm{Tr}_R U_{\mathcal{K}} \tag{2.3}$$

where

$$U_{\mathcal{K}} = \mathrm{P}\ \exp \oint_{\mathcal{K}} A \tag{2.4}$$

is the holonomy around the knot. (2.3) is a gauge invariant operator whose definition does not involve the metric on the three-manifold. The irreducible representations of $SU(N)$ and $U(N)$ will be labeled by highest weights or equivalently by the lengths of rows in a Young tableau, l_i, where $l_1 \geq l_2 \geq \cdots$. If we now consider a link $\mathcal{L}$ with components $\mathcal{K}_i$, $i = 1, \cdots, L$, we can in principle compute the correlation function,

$$W_{(R_1,\cdots,R_L)}(\mathcal{L}) = \langle W_{R_1}^{\mathcal{K}_1} \cdots W_{R_L}^{\mathcal{K}_L} \rangle = \frac{1}{Z(M)} \int [\mathcal{D}A] \Big(\prod_{i=1}^{L} W_{R_i}^{\mathcal{K}_i} \Big) \mathrm{e}^{iS}. \tag{2.5}$$

The topological character of the action, and the fact that the Wilson loop operators can be defined without usint any metric on the three-manifold, indicate that (2.5) is a topological invariant of the link $\mathcal{L}$. Notice that our knots and links are oriented, and this makes a difference. If $\mathcal{K}^{-1}$ denotes the knot obtained from $\mathcal{K}$ by inverting its orientation, we have that

$$\mathrm{Tr}_R U_{\mathcal{K}^{-1}} = \mathrm{Tr}_R U_{\mathcal{K}}^{-1} = \mathrm{Tr}_{\overline{R}} U_{\mathcal{K}}, \tag{2.6}$$

where $\overline{R}$ denotes the conjugate representation. For further use we notice that, given two linked oriented knots $\mathcal{K}_1$, $\mathcal{K}_2$, one can define a elementary topological invariant, the *linking number*, by

$$\mathrm{lk}(\mathcal{K}_1, \mathcal{K}_2) = \frac{1}{2} \sum_{\mathrm{p}} \epsilon(p), \tag{2.7}$$

where the sum is over all crossing points, and $\epsilon(p) = \pm 1$ is a sign associated to the crossings as indicated in Fig. 2.

2.2 Perturbative approach

The partition function and correlators of Wilson loops in Chern-Simons theory can be computed in a variety of ways. One can for example use standard perturbation theory. In the computation of the partition function in perturbation theory, we have first to find the classical solutions of the Chern-Simons equations of motion. If we write $A = \sum_a A^a T_a$, where T_a is a basis of the Lie algebra, we find:

$$\frac{\delta S}{\delta A_\mu^a} = \frac{k}{4\pi} \epsilon^{\mu\nu\rho} F_{\nu\rho}^a,$$

Figure 2: When computing the linking number of two knots, the crossings are assigned a sign ± 1 as indicated in the figure.

therefore the classical solutions are just flat connections on M. Flat connections are in one-to-one correspondence with group homomorphisms

$$\pi_1(M) \to G. \tag{2.8}$$

For example, if $M = \mathbf{S}^3/\mathbf{Z}_p$ is the lens space $L(p,1)$, one has $\pi_1(L(p,1)) = \mathbf{Z}_p$, and flat connections are labeled by homomorphisms $\mathbf{Z}_p \to G$. Let us assume that these are a discrete set of points (this happens, for example, if M is a rational homology sphere, since in that case $\pi_1(M)$ is a finite group). In that situation, one expresses $Z(M)$ as a sum of terms associated to stationary points:

$$Z(M) = \sum_c Z^{(c)}(M), \tag{2.9}$$

where c labels the different flat connections $A^{(c)}$ on M. Each of the $Z^{(c)}(M)$ will be an asympotic series in $1/k$ of the form

$$Z^{(c)}(M) = Z^{(c)}_{1-\text{loop}}(M). \exp\left\{\sum_{\ell=1}^{\infty} S_\ell^{(c)} x^\ell\right\}. \tag{2.10}$$

In this equation, x is the effective expansion parameter:

$$x = \frac{2\pi i}{k+y}, \tag{2.11}$$

and y is the dual Coxeter of the group. The one-loop correction $Z^{(c)}_{1-\text{loop}}(M)$ was first analyzed in [89], and has been studied in big detail since then. It has the form,

$$Z^{(c)}_{1-\text{loop}}(M) = \frac{(2\pi x)^{\frac{1}{2}(\dim H^0_c - \dim H^1_c)}}{\text{vol}(H_c)} \mathrm{e}^{-\frac{1}{x} S_{\text{CS}}(A^{(c)}) - \frac{i\pi}{4}\varphi} \sqrt{|\tau_R^{(c)}|}, \tag{2.12}$$

where $H_c^{0,1}$ are the cohomology groups with values in the Lie algebra of G associated to the flat connection $A^{(c)}$, $\tau_R^{(c)}$ is the Reidemeister-Ray-Singer torsion of $A^{(c)}$, H_c is the isotropy group of $A^{(c)}$, and φ is a certain phase. In (2.12), $S_{\rm CS}(A^{(c)})/x$ is the classical action evaluated at the flat connection, so we have also included in (2.12) the tree-level part. More details about the structure of this term can be found in [89, 33, 47, 80]. Notice that, when $A^{(c)}=0$, the trivial flat connection, one has that $\dim H_c^0 = \dim G$, $\dim H_c^1 = 0$, and $H_c = G$.

The terms $S_\ell^{(c)}$ in (2.10) correspond to connected diagrams with $\ell/2$ vertices. In Chern-Simons theory the vertex is trivalent, so $S_\ell^{(c)}$ is the contribution to the free energy at $\ell+1$ loops. The contributions of the trivial connection will be denoted simply by S_ℓ, and their diagrammatic content for the first few ℓ is as follows :

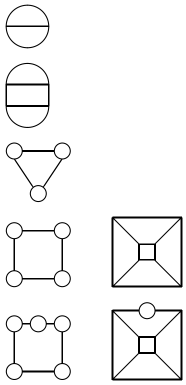
$\ell = 1:$

$\ell = 2:$

$\ell = 3:$

$\ell = 4:$

$\ell = 5:$



(2.13)

Notice that the computation of $S_\ell^{(c)}$ involves the evaluation of group factors of Feynman diagrams, so they depend explicitly on the gauge group G (in the $SU(N)$ or $U(N)$ case, they are polynomials in N). For example, when $G=U(N)$ and one expands around the trivial connection,

$$S_1 \sim 2N(N^2-1) \tag{2.14}$$

which is the group factor of the corresponding diagram.

The perturbative evaluation of Wilson loop correlators can also be done using the standard procedures. First of all one has to expand the holonomy operator as

$$W_R^{\mathcal{K}}(A) = {\rm Tr}_R\Big[\mathbf{1} + \oint_{\mathcal{K}} dx^\mu A_\mu(x) + \oint_{\mathcal{K}} dx^\mu \int^x dy^\nu A_\nu(y) A_\mu(x) + \cdots\Big], \tag{2.15}$$

where $A_\mu = \sum_a A_\mu^a T_a$. Then, after gauge fixing, one can proceed and evaluate the correlation functions in standard perturbation theory. The perturbative study of Wilson loops was started in [40]. A nice review of its development can be found in [54]. In these lectures we will rather focus on the nonperturbative approach to Chern-Simons theory, which we now explain.

2.3 Canonical quantization and surgery

As Witten showed in [89], one can solve Chern-Simons theory exactly by using nonperturbative methods and the relation to the Wess-Zumino-Witten (WZW) model.

We start with some basic facts about the canonical quantization of Chern-Simons theory. Let M be a three-manifold with boundary given by a Riemann surface Σ. We can insert a general operator $\mathcal{O}$ in M, which will be in general a product of Wilson loops along different knots and in arbitrary representations of the gauge group. We will consider the case in which the Wilson loops do not intersect the surface Σ. The path integral over the three-manifold with boundary M gives a wavefunction $\Psi_{M,\mathcal{O}}(\mathcal{A})$ which is a functional of the values of the field at Σ. Schematically we have:

$$\Psi_{M,\mathcal{O}}(\mathcal{A}) = \langle \mathcal{A}|\Psi_{M,\mathcal{O}}\rangle = \int_{A|_\Sigma=\mathcal{A}} \mathcal{D}A\, \mathcal{O}\, \mathrm{e}^{iS}. \tag{2.16}$$

In fact, associated to the Riemann surface we have a Hilbert space $\mathcal{H}(\Sigma)$, which can be obtained by doing canonical quantization of Chern-Simons theory on $\Sigma_g \times \mathbf{R}$. Before spelling out in detail the structure of these Hilbert spaces, let us make some general considerations about the computation of physical quantities.

In the context of canonical quantization, the partition function can be computed as follows. We first perform a Heegard splitting of the three-manifold, *i.e.* we represent it as the connected sum of two three-manifolds M_1 and M_2 sharing a common boundary Σ, where Σ is a Riemann surface. If $f : \Sigma \to \Sigma$ is an homeomorphism, we will write $M = M_i \cup_f M_2$, so that M is obtained by gluing M_1 to M_2 through their common boundary and using the homeomorphism f. This is represented in Fig. 3. We can then compute the full path integral (2.2) over M by computing first the path integral over M_1 to obtain a state $|\Psi_{M_1}\rangle$ in $\mathcal{H}(\Sigma)$. The boundary of M_2 is also Σ, but with opposite orientation, so its Hilbert space is the dual space $\mathcal{H}^*(\Sigma)$. The path integral over M_2 produces then a state $\langle\Psi_{M_2}|$. The homeomorphism $f : \Sigma \to \Sigma$ will be represented by an operator acting on $\mathcal{H}(\Sigma)$,

$$U_f : \mathcal{H}(\Sigma) \to \mathcal{H}(\Sigma). \tag{2.17}$$

and the partition function can then be evaluated as

$$Z(M) = \langle \Psi_{M_2}|U_f|\Psi_{M_1}\rangle. \tag{2.18}$$

Therefore, if we know explicitly what are the wavefunctions and the operators associated to homeomorphisms, we can compute the partition function.

One of the most fundamental results of Witten in [89] is in fact a precise description of $\mathcal{H}(\Sigma)$: it is the space of conformal blocks of a WZW model on Σ with gauge group G and

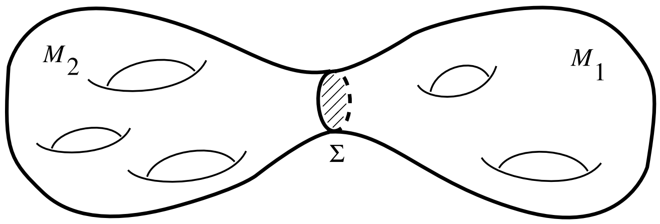


Figure 3: Heegard splitting of a three-manifold M into two three manifolds M_1 and M_2 with a common boundary Σ.

level k (for an extensive review of the WZW model, see for example [25]). In particular, $\mathcal{H}(\Sigma)$ has finite dimension. We will not review here the derivation of this fundamental result. Instead we will use the relevant information from the WZW model in order to solve CS theory.

The description of the space of conformal blocks on Riemann surfaces can be made very explicit when Σ is a sphere or a torus. For $\Sigma = \mathbf{S}^2$, the space of conformal blocks is one-dimensional, so $\mathcal{H}(\mathbf{S}^2)$ is spanned by a single element. For $\Sigma = \mathbf{T}^2$, the space of conformal blocks is in one to one correspondence with the integrable representations of the affine Lie algebra associated to G at level k [1]. We recall that a representation given by a highest weight Λ is integrable if the weight $\rho + \Lambda$ is in the fundamental chamber $\mathcal{F}_l$, where $l = k + y$ (ρ denotes as usual the Weyl vector, given by the sum of the fundamental weights). The fundamental chamber is given by $\Lambda_{\rm w}/l\Lambda_{\rm r}$ modded out by the action of the Weyl group. For example, in $SU(N)$ a weight $p = \sum_{i=1}^r p_i \lambda_i$ is in $\mathcal{F}_l$ if

$$\sum_{i=1}^{r} p_i < l, \quad \text{and} \quad p_i > 0,\, i = 1, \cdots, r. \tag{2.19}$$

In the following, the basis of integrable representations will be labeled by the weights in $\mathcal{F}_l$, and the states in the Hilbert state of the torus $\mathcal{H}(\mathbf{T}^2)$ will be denoted by $|p\rangle = |\rho + \Lambda\rangle$ where $\rho + \Lambda$, as we have stated, is an integrable representation of the WZW model at level k. We will also denote these states by $|R\rangle$, where R is the representation associated to Λ. The state $|\rho\rangle$ will be denoted by $|0\rangle$. The states $|R\rangle$ can be chosen to be orthonormal [31, 57], so we have

$$\langle R|R'\rangle = \delta_{RR'}. \tag{2.20}$$

[1]We will use the following notations in the following: the fundamental weights of G will be denoted by λ_i, and the simple roots by α_i, with $i = 1, \cdots, r-1$, and r denotes the rank of G. The weight and root lattices of G are denoted by $\Lambda_{\rm w}$ and $\Lambda_{\rm r}$, respectively, and $|\Delta_+$ denotes the number of positive roots.

There is a special class of homeomorphisms of $\mathbf{T}^2$ that have a simple expression as operators in $\mathcal{H}(\mathbf{T}^2)$, given by $\mathrm{Sl}(2,\mathbf{Z})$ transformations. Recall that the group $\mathrm{Sl}(2,\mathbf{Z})$ consists of 2×2 matrices with integer entries and unit determinant. If $(1,0)$ and $(0,1)$ denote the two one-cycles of $\mathbf{T}^2$, we can specify the action of an $\mathrm{Sl}(2,\mathbf{Z})$ transformation on the torus by giving its action on this homology basis. The $\mathrm{Sl}(2,\mathbf{Z})$ group is generated by the transformations T and S, which are given by

$$T=\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \qquad S=\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \tag{2.21}$$

Notice that the S transformations exchanges the one-cycles of the torus. These transformations can be lifted to $\mathcal{H}(\mathbf{T}^2)$, and they have the following matrix elements in the basis of integrable representations:

$$\begin{aligned} T_{pp'} &= \delta_{p,p'} \mathrm{e}^{2\pi i (h_p - c/24)}, \\ S_{pp'} &= \frac{i^{|\Delta_+|}}{(k+y)^{r/2}} \left(\frac{\mathrm{Vol}\,\Lambda^w}{\mathrm{Vol}\,\Lambda^r}\right)^{\frac{1}{2}} \sum_{w\in\mathcal{W}} \exp\Bigl(-\frac{2\pi i}{k+y} p\cdot w(p')\Bigr). \end{aligned} \tag{2.22}$$

In the first equation, h_p is the conformal weight of the primary field associated to p:

$$h_p=\frac{p^2-\rho^2}{2(k+y)}, \tag{2.23}$$

where we remind that p is of the form $\rho+\Lambda$. In the second equation, the sum over w is a sum over the elements of the Weyl group.

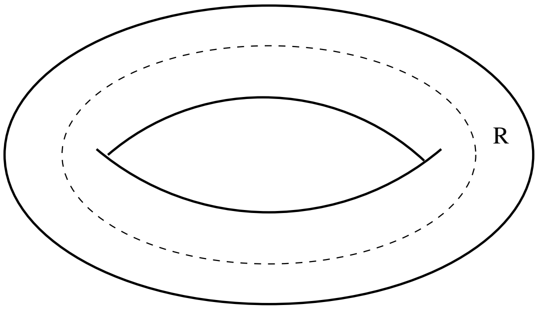


Figure 4: Performing the path integral on a solid torus with a Wilson line in representation R gives the state $|R\rangle$ in $\mathcal{H}(\mathbf{T}^2)$.

What is the description of these states from the point of view of canonical quantization? Consider the solid torus $\mathcal{T}=D\times\mathbf{S}^1$, where D is a disc in $\mathbf{R}^2$. This is a three-manifold whose boundary is a $\mathbf{T}^2$, and it has a noncontractible cycle given by the $\mathbf{S}^1$. Let us now consider the Chern-Simons path integral on the solid torus, with the insertion of the operator $\mathcal{O}_R = \mathrm{Tr}_R U$ given by a Wilson loop in the representation R around the noncontractible cycle, as shown in Fig. 4. In this way one obtains a state in $\mathcal{H}(\mathbf{T}^2)$, and one has

$$|\Psi_{\mathcal{T},\mathcal{O}_R}\rangle = |R\rangle. \tag{2.24}$$

In particular, the path integral over the solid torus with no operator insertion gives $|0\rangle$, the "vacuum" state.

These results allow us to compute the partition function of any three-manifold that admits a Heegard splitting along a torus. Imagine that we take two solid tori and we glue them along their boundary. Since a solid torus is a disc times a circle $D \times \mathbf{S}^1$, by performing this operation we get a manifold which is $\mathbf{S}^1$ times the two discs glued together along their boundaries. Therefore, with this surgery we obtain $\mathbf{S}^2 \times \mathbf{S}^1$, and (2.18) gives then

$$Z(\mathbf{S}^2 \times \mathbf{S}^1) = \langle 0|0\rangle = 1. \tag{2.25}$$

If we do the gluing, however, after performing an S-transformation on the $\mathbf{T}^2$ the resulting manifold is instead $\mathbf{S}^3$. To see this, notice that the complement to a solid torus inside $\mathbf{S}^3$ is indeed another solid torus whose noncontractible cycle is homologous to the contractible cycle in the first torus. We then find

$$Z(\mathbf{S}^3) = \langle 0|S|0\rangle = S_{00}. \tag{2.26}$$

By using Weyl's denominator formula,

$$\sum_{w\in\mathcal{W}} \epsilon(w) \mathrm{e}^{w(\rho)} = \prod_{\alpha>0} 2\sinh\frac{\alpha}{2}, \tag{2.27}$$

one finds

$$Z(\mathbf{S}^3) = \frac{1}{(k+y)^{r/2}} \left(\frac{\mathrm{Vol}\,\Lambda^w}{\mathrm{Vol}\,\Lambda^r}\right)^{\frac{1}{2}} \prod_{\alpha>0} 2\sin\Big(\frac{\pi(\alpha\cdot\rho)}{k+y}\Big). \tag{2.28}$$

The above result can be generalized in order to compute the vacuum expectation value of $\mathbf{S}^3$ with some knots and links inserted. Consider a solid torus where a Wilson line in representation R has been inserted. The corresponding state is $|R\rangle$, as we explained before. If we now glue this to an empty solid torus after an S-transformation, we obtain an unknot in $\mathbf{S}^3$. The path integral with the insertion is then,

$$Z(\mathbf{S}_3, \mathcal{O}_R) = \langle 0|S|R\rangle. \tag{2.29}$$

It follows that the normalized vev for the unknot in $\mathbf{S}^3$, in representation R, is given by

$$W_R(\text{unknot}) = \frac{S_{0R}}{S_{00}} = \frac{\sum_{w\in\mathcal{W}} \epsilon(w) \mathrm{e}^{-\frac{2\pi i}{k+y}\rho\cdot w(\Lambda+\rho)}}{\sum_{w\in\mathcal{W}} \epsilon(w) \mathrm{e}^{-\frac{2\pi i}{k+y}\rho\cdot w(\rho)}}. \tag{2.30}$$

This is just the character of the representation R, evaluated on a particular element, so the vacuum expectation value can be written as

$$W_R(\text{unknot}) = \mathrm{ch}_R\Big[-\frac{2\pi i}{k+y}\rho\Big]. \tag{2.31}$$

Moreover, using (2.27), we finally write

$$W_R(\text{unknot}) = \prod_{\alpha>0} \frac{\sin\Big(\frac{\pi}{k+y}\alpha\cdot(\Lambda+\rho)\Big)}{\sin\Big(\frac{\pi}{k+y}\alpha\cdot\rho\Big)} \tag{2.32}$$

This quantity is often called the *quantum dimension* of R, and it is denoted by $\mathrm{dim}_q R$.

Exercise. *Quantum dimensions for $SU(N)$.* Consider Chern-Simons theory with gauge group $SU(N)$. Define the q-numbers:

$$[x] = q^{\frac{x}{2}} - q^{-\frac{x}{2}}, \quad [x]_\lambda = \lambda^{\frac{1}{2}} q^{\frac{x}{2}} - \lambda^{-\frac{1}{2}} q^{-\frac{x}{2}} \tag{2.33}$$

where

$$q = \exp\Big(\frac{2\pi i}{k+N}\Big), \quad \lambda = q^N. \tag{2.34}$$

If R corresponds to a Young tableau with c_R rows of lengths l_i, $i = 1, \cdots, c_R$, show that the quantum dimension is given by:

$$\mathrm{dim}_q R = \prod_{1\le i<j\le c_R} \frac{[l_i - l_j + j - i]}{[j-i]} \prod_{i=1}^{c_R} \frac{\prod_{v=-i+1}^{l_i - i}[v]_\lambda}{\prod_{v=1}^{l_i}[v-i+c_R]}. \tag{2.35}$$

Verify that in the limit $k+N \to \infty$ (*i.e.* in the semiclassical limit) the quantum dimension becomes the dimension of the representation R. Notice that the normalized vev is a rational function of $q^{\pm 1/2}$, $\lambda^{\pm 1/2}$. This is a general property of *all* normalized vevs of knots and links in $\mathbf{S}^3$.

We can also consider a solid torus with Wilson loop in representation R, glued to another solid torus with the representation R' through an S-transformation. What we obtain is clearly a link in $\mathbf{S}^3$ with two components, which is the Hopf link shown in Fig. 1. Taking into account the orientation carefully, we find that this is the Hopf link with linking number -1. The path integral with this insertion is:

$$Z(\mathbf{S}^3, \mathcal{O}_R \mathcal{O}_{R'}) = \langle R'|S|R\rangle, \tag{2.36}$$

so the normalized vev is

$$W_{RR'}(\mathrm{Hopf}^{-1}) = \frac{S_{RR'}}{S_{00}}, \tag{2.37}$$

where the superscript -1 refers to the linking number. Therefore, the Chern-Simons invariant of the Hopf link is essentially an S-matrix element. In order to obtain the invariant of the Hopf link with linking number $+1$, we notice that the two Hopf links can be related by changing the orientation of one of the components. We then have

$$\mathcal{W}_{RR'} \equiv W_{RR'}(\mathrm{Hopf}^{+1}) = \frac{S_{\overline{R}R'}}{S_{00}} = \frac{S^{-1}_{RR'}}{S_{00}}, \tag{2.38}$$

where we have used the property (2.6) and the fact that $S_{\overline{R}R'} = S^{-1}_{RR'}$ (see for example [25]).

Although the S-matrix elements can be evaluated through the explicit expression (2.22), or for example by using the relation between $U(N)$ characters and Schur functions, an explicit expression of (2.37) as a rational function of q, λ is more difficult to find. We present now an algorithm due to Morton and Lukac [70] which is useful in this respect. Their formula is as follows. Let μ be a Young tableau, and let $\mu^\vee$ denote its transposed tableau (remember that this tableau is obtained from μ by exchanging rows and columns). The Schur polynomial in the variables $(x_1, \cdots, x_N)$ corresponding to μ (which is the character of the diagonal $SU(N)$ matrix $(x_1, \cdots, x_N)$ in the representation corresponding to μ), will be denoted by s_μ. They can be written in terms of elementary symmetric polynomials $e_i(x_1, \cdots, x_N)$, $i \geq 1$, as follows [62]:

$$s_\mu = \det M_\mu \tag{2.39}$$

where

$$M^{ij}_\mu = (e_{\mu^\vee_i + j - i})$$

M_μ is an $r \times r$ matrix, with $r = d(\mu^\vee)$. To evaluate s_μ we put $e_0 = 1$, $e_k = 0$ for $k < 0$. The expression (2.39), known sometimes as the Jacobi-Trudy identity, can be formally extended to give the Schur polynomial $s_\mu(E(t))$ associated to any formal power series $E(t) = 1 + \sum_{n=1}^\infty a_i t^i$. To obtain this, we simply use the Jacobi-Trudy formula (2.39), but where e_i denote now the coefficients of the series $E(t)$, i.e. $e_i = a_i$. Morton and Lukac define the series $E_\emptyset(t)$ as follows:

$$E_\emptyset(t) = 1 + \sum_{n=1}^\infty c_n t^n, \tag{2.40}$$

where the coefficients c_n are defined by

$$c_n = \prod_{i=1}^n \frac{1 - \lambda^{-1} q^{i-1}}{q^i - 1}. \tag{2.41}$$

They also define a formal power series associated to a tableau μ, $E_\mu(t)$, as follows:

$$E_\mu(t) = E_\emptyset(t) \prod_{j=1}^{d(\mu)} \frac{1+q^{\mu_j - j}t}{1+q^{-j}t}. \tag{2.42}$$

One can then consider the Schur function of the power series (2.42), $s_\mu(E_{\mu'}(t))$, for any pair of tableaux μ, μ', by expanding $E_{\mu'}(t)$ and substituting its coefficients in the Jacobi-Trudy formula (2.39). It turns out that this Schur function is essentially the invariant we were looking for. More precisely, one has

$$\mathcal{W}_{R_1,R_2}(q,\lambda) = (\dim_q R_1)(\lambda q)^{\frac{\ell_2}{2}} s_{\mu_2}(E_{\mu_1}(t)), \tag{2.43}$$

where $\mu_{1,2}$ are the tableaux corresponding to $R_{1,2}$, and ℓ_2 is the number of boxes of μ_2.

2.4 Framing dependence

In the above discussion on the correlation functions of Wilson loops we have missed an important ingredient. We mentioned that, in order to define the partition function of Chern-Simons theory at the quantum level, one has to specify a framing of the three-manifold. It turns out that the evaluation of correlation functions like (2.5) also involves a choice of framing of the knots, as Witten discovered in [89]. Since this is important in the duality with topological strings, we will explain it in some detail.

A good starting point to understand the framing is to take Chern-Simons theory with gauge group $U(1)$. This is also useful to understand $U(N)$ versus $SU(N)$ Chern-Simons theory, and to get a concrete feeling of how to deal with correlation functions like (2.5). The abelian Chern-Simons theory turns out to be extremely simple, since the cubic term in (2.1) drops out, and we are left with a Gaussian theory [78]. The different representations are labeled by integers, and in particular the vevs of Wilson loop operators can be computed exactly. In order to compute them, however, one has to choose a framing for each of the knots $\mathcal{K}_i$. This arises as follows: in evaluating the vev, contractions of the holonomies corresponding to different $\mathcal{K}_i$ produce the following integral:

$$\mathrm{lk}(\mathcal{K}_i, \mathcal{K}_j) = \frac{1}{4\pi} \oint_{\mathcal{K}_i} dx^\mu \oint_{\mathcal{K}_j} dy^\nu \epsilon_{\mu\nu\rho} \frac{(x-y)^\rho}{|x-y|^3}. \tag{2.44}$$

This is in fact a topological invariant, *i.e.* it is invariant under deformations of the knots $\mathcal{K}_i$, $\mathcal{K}_j$, and it is in fact their linking number $\mathrm{lk}(\mathcal{K}_i, \mathcal{K}_j)$. On the other hand, contractions of the holonomies corresponding to the same knot $\mathcal{K}$ involve the integral

$$\phi(\mathcal{K}) = \frac{1}{4\pi} \oint_{\mathcal{K}} dx^\mu \oint_{\mathcal{K}} dy^\nu \epsilon_{\mu\nu\rho} \frac{(x-y)^\rho}{|x-y|^3}. \tag{2.45}$$

This integral is well-defined and finite (see, for example, [40]), and it is called the cotorsion or *writhe* of $\mathcal{K}$. It is in fact given by the self-linking number of $\mathcal{K}$: if we project $\mathcal{K}$ on a plane, and we denote by $n_{\pm}(\mathcal{K})$ the number of positive (negative) crossings as indicated in Fig. 2.44, then we have that

$$\phi(\mathcal{K}) = n_+(\mathcal{K}) - n_-(\mathcal{K}). \tag{2.46}$$

The problem is that the cotorsion is not invariant under deformations of the knot. In order to preserve topological invariance one has to choose another definition of the composite operator $(\oint_{\mathcal{K}} A)^2$ by means of a framing. A framing of the knot consists of choosing another knot $\mathcal{K}^f$ around $\mathcal{K}$, specified by a normal vector field n. The cotorsion $\phi(\mathcal{K})$ becomes then

$$\phi_f(\mathcal{K}) = \frac{1}{4\pi}\oint_{\mathcal{K}} dx^\mu \oint_{\mathcal{K}^f} dy^\nu \epsilon_{\mu\nu\rho}\frac{(x-y)^\rho}{|x-y|^3} = \mathrm{lk}(\mathcal{K},\mathcal{K}^f). \tag{2.47}$$

The correlation function that we obtain in this way is a topological invariant (a linking number) but the price that we have to pay is that our regularization depends on a set of integers $p_i = \mathrm{lk}(\mathcal{K}_i, \mathcal{K}_i^f)$ (one for each knot). The vev (2.5) can now be computed, after choosing the framings, as follows:

$$\langle \prod_i \exp(n_i \oint_{\gamma_i} A)\rangle = \exp\Big(\frac{\pi i}{k}\sum_i n_i^2 p_i + \frac{\pi i}{k}\sum_{i\neq j} n_i n_j \mathrm{lk}(\mathcal{K}_i,\mathcal{K}_j)\Big). \tag{2.48}$$

This regularization is nothing but the 'point-splitting' method familiar in the context of QFT's.

Let us now consider Chern-Simons theory with gauge group $SU(N)$, and suppose that you want to compute a correlation function like (2.5). If you try to do it in perturbation theory, for example, you will find very soon that self-contractions of the holonomies lead to the same kind of ambiguities that we found in the abelian case, *i.e.* you will have to make a choice of framing for each knot $\mathcal{K}_i$. The only difference is that the self contraction comes with a group factor $\mathrm{Tr}_{R_i}(T_a T_a)$ for each knot $\mathcal{K}_i$, where T_a is a basis of the Lie algebra [40]. The precise result can be better stated as the effect on the correlation function (2.5) under a change of framing, and it says that, under a change of framing of $\mathcal{K}_i$ by p_i units, the vev of the product of Wilson loops changes as follows [89]:

$$W_{(R_1,\cdots,R_L)} \rightarrow \exp\Big[2\pi i \sum_i p_i h_{R_i}\Big] W_{(R_1,\cdots,R_L)}, \tag{2.49}$$

In this equation, h_R is the conformal weight of the WZW primary field corresponding to the representation R. In (2.23) we label R through $\alpha = \rho + \Lambda$, where Λ is the highest weight of R. In fact, one can write (2.23) as

$$h_R = \frac{C_R}{2(k+N)}, \tag{2.50}$$

where $C_R = \mathrm{Tr}_R(T_a T_a)$ is the quadratic Casimir in the representation R. For $SU(N)$, one has

$$C_R^{SU(N)} = N\ell + \kappa_R - \frac{\ell^2}{N}, \tag{2.51}$$

where ℓ is the total number of boxes in the tableau, and

$$\kappa_R = \ell + \sum_i (l_i^2 - 2il_i). \tag{2.52}$$

We then see that the evaluation of vacuum expectation values of Wilson loop operators in Chern-Simons theory depends on a choice of framing for knots. It turns out that for knots and links in $\mathbf{S}^3$, there is a *standard* or canonical framing, defined by asking that the self-linking number is zero. The expression we have given for the Chern-Simons invariants of the unknot is in the standard framing. Once the value of the invariant is known in the standard framing, the value in any other framing specified by nonzero integers p_i can be easily obtained from (2.49). Notice that the operator that implements the change of framing is the T operator: the state $T^p|R\rangle$ corresponds to a Wilson line insertion in a solid torus with framing p, as one can easily check by computing the normalized vev $\langle 0|ST^p|R\rangle / \langle 0|ST^p|0\rangle$.

Let us now consider a $U(N)$ Chern-Simons theory. The $U(1)$ factor decouples from the $SU(N)$ theory, and all the vevs factorize into an $U(1)$ and an $SU(N)$ piece. Representations of $U(N)$ are also labeled by Young tableaux, and they decompose into a representation of $SU(N)$ corresponding to that tableau, and a representation of $U(1)$ with charge:

$$n = \frac{\ell}{\sqrt{N}}, \tag{2.53}$$

where ℓ is the number of boxes in the Young tableau. In order to compute the vevs associated to the $U(1)$ of $U(N)$, one has to take also into account that the coupling constant k is shifted as $k \to k + N$. We then find that the vev of a product of $U(N)$ Wilson loops in representations R_i is given by:

$$W^{U(N)}_{(R_1,\cdots,R_L)} = \exp\Big(\frac{\pi i}{N(k+N)} \sum_i \ell_i^2 p_i + \frac{\pi i}{N(k+N)} \sum_{i\neq j} \ell_i \ell_j \mathrm{lk}(\mathcal{K}_i, \mathcal{K}_j)\Big) W^{SU(N)}_{(R_1,\cdots,R_L)}, \tag{2.54}$$

where the $SU(N)$ vev is computed in the framing specified by p_i. Notice that, in the case of knots, the $SU(N)$ and $U(N)$ computations differ in a factor which only depends on the choice of framing, while for links the answers also differ in a topological piece involving the linking numbers. The change of framing for vacuum expectation values in the $U(N)$ theory is again governed by (2.49) and (2.50), but now the quadratic Casimir is given by

$$C_R^{U(N)} = N\ell + \kappa_R, \tag{2.55}$$

Notice that the difference between the change of $SU(N)$ and $U(N)$ vevs under the change of framing is consistent with (2.59). In terms of the variables (2.34) we see that $U(N)$ vevs change, under the change of framing, as

$$W_{(R_1,\cdots,R_L)} \to q^{\frac{1}{2}\sum_i \kappa_{R_i} p_i} \lambda^{\frac{1}{2}\sum_i \ell_i p_i} W_{(R_1,\cdots,R_L)}. \tag{2.56}$$

2.5 More results on Wilson loops

In this subsection we will mostly list some useful results for the computation of vacuum expectation values of Wilson loops. These results are behaviour under parity, factorization of disjoint links, and fusion rules. We will also give some results for the invariant of a Hopf link that will be useful in the applications to topological strings.

Chern-Simons theory is a theory of oriented links, and under a parity transformation a link $\mathcal{L}$ will transform into its mirror $\mathcal{L}^*$. The mirror of $\mathcal{L}$ is obtained from its planar projection simply by changing undercrossings by overcrossings, and viceversa. On the other hand, parity changes the sign of the Chern-Simons action, in other words $k+N \to -(k+N)$. We then find that vacuum expectation values transform as

$$W_{(R_1,\cdots,R_L)}(\mathcal{L}^*)(q,\lambda) = W_{(R_1,\cdots,R_L)}(\mathcal{L})(q^{-1},\lambda^{-1}). \tag{2.57}$$

This is interesting from a knot-theoretic point of view, since it implies that Chern-Simons invariants of links can distinguish in principle a link from its mirror image. As an example of this property, notice for example that the framed unknot is identical to its mirror image, therefore quantum dimensions satisfy:

$$(\dim_q R)(q^{-1},\lambda^{-1}) = (\dim_q R)(q,\lambda). \tag{2.58}$$

The factorization property for disjoint links states the following. Let $\mathcal{L}$ be a link with L components $\mathcal{K}_1, \cdots, \mathcal{K}_L$ which are *disjoint* knots, and let us attach the representation R_α to the α-th component. Then one has

$$W_{(R_1,\cdots,R_L)}(\mathcal{L}) = \prod_{\alpha=1}^{L} W_{R_\alpha}(\mathcal{K}_\alpha). \tag{2.59}$$

This property is easy to prove in Chern-Simons theory. It only involves some elementary surgery and the fact that $\mathcal{H}(\mathbf{S}^2)$ is one-dimensional. A proof can be found in [89].

Let us now explain the most elementary fusion rule in Chern-Simons theory. Consider a vacuum expectation value of the form

$$\langle \mathrm{Tr}_{R_1} U \mathrm{Tr}_{R_2} U \rangle, \tag{2.60}$$

where U is the holonomy of the gauge field around a knot $\mathcal{K}$. The product of traces satisfies the fusion rule

$$\mathrm{Tr}_{R_1}U\mathrm{Tr}_{R_2}U = \sum_R N_{R_1 R_2}^R \mathrm{Tr}_R U, \tag{2.61}$$

where $N_{R_1 R_2}^R$ are tensor product coefficients which can be obtained, for example, by using the Littlewood-Richardson rule. In Chern-Simons theory, these coefficients become the fusion coefficients of the WZW model, and this can be easily understood from the fact that the admissible representations that appear in the theory are the integrable ones, so one has to truncate the list of "classical" representations, and this implies in particular that the fusion rules of classical traces have to be modified.

However, in the computation of knot invariants in $SU(N)$ or $U(N)$ Chern-Simons theory it is natural to work in a setting in which k and N are much larger than any of the representations involved, in such a way that no truncations appear in the fusion rules. In that case, the vacuum expectation values of the theory satisfy

$$\langle \mathrm{Tr}_{R_1}U\mathrm{Tr}_{R_2}U \rangle = \sum_R N_{R_1 R_2}^R \langle \mathrm{Tr}_R U \rangle. \tag{2.62}$$

Imagine now that we want to compute $\langle \mathrm{Tr}_{R_1}U_1\mathrm{Tr}_{R_2}U_2 \rangle$, where $U_{1,2}$ are holonomies around disjoint unknots with zero framing. We can take the unknots to be very close, in such a way that the paths along which we compute the holonomy are the same. In that case, this vacuum expectation value becomes exactly the l.h.s of (5). Using also the factorization property (2.59), we deduce the following fusion rule:

$$\dim_q R_1 \dim_q R_2 = \sum_R N_{R_1 R_2}^R \dim_q R. \tag{2.63}$$

In fact, this fusion rule is just a particular case of the multiplication rule of characters. Another application of the fusion rules is to vacuum expectation values involving two disjoint unknots in a link, like in Fig. 5. Fusing the two parallel unknots we find

$$\langle \mathrm{Tr}_{R_1}U_1\mathrm{Tr}_{R_2}U_2\mathrm{Tr}_R U_3 \rangle = \sum_{R'} N_{R_1 R_2}^{R'} \langle \mathrm{Tr}_{R'}U\mathrm{Tr}_R U_3 \rangle \tag{2.64}$$

where U_1, U_2 correspond to the holonomies of the parallel unknots. The computation is therefore reduced to a Hopf link with two components in higher representations. A more general fusing rule comes from considering two unknots linked between them through m twists. In Fig. 6 we show two unknots linked through one twist in order to obtain the Hopf link. In this case, the fusion rule is (see for example [41])

$$\langle \mathrm{Tr}_{R_1}U_1^{(m)}\mathrm{Tr}_{R_2}U_2^{(m)} \rangle = \sum_R N_{R_1 R_2}^R \langle \mathrm{Tr}_R U^{(m)} \rangle, \tag{2.65}$$

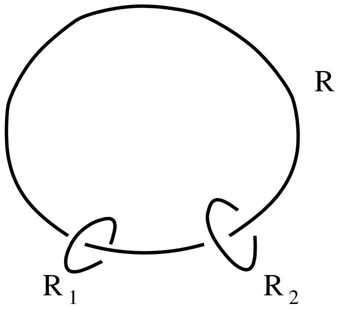


Figure 5: A link with two disjoint unknots.

where the m superscript denotes the framing of the various components, and equals the twisting. Notice that here we fuse two framed and linked unknots into a framed unknot. For the $U(N)$ Chern-Simons invariant of the Hopf link with linking number $+1$ (whose invariant is given by $S^{-1}_{R_1 R_2}/S_{00}$) we obtain in particular the following expression in terms of quantum dimensions:

$$\mathcal{W}_{R_1 R_2} = \sum_R N^R_{R_1 R_2} q^{(\kappa_R - \kappa_{R_1} - \kappa_{R_2})/2} \mathrm{dim}_q R. \tag{2.66}$$

Notice that the Hopf link with linking number -1 is the mirror image of the Hopf link with opposite linking number, therefore we find

$$\frac{S_{R_1 R_2}}{S_{00}} = \sum_R N^R_{R_1 R_2} q^{(\kappa_{R_1} + \kappa_{R_2} - \kappa_R)/2} \mathrm{dim}_q R. \tag{2.67}$$

Exercise. Compute $\mathcal{W}_{\square\,\square}$, $\mathcal{W}_{\square\!\square\,\square}$, $\mathcal{W}_{\begin{smallmatrix}\square\\ \square\end{smallmatrix},\square}$, using both (2.66) and (2.43).

As we will see, the relation between Chern-Simons theory and string theory involves the vacuum expectation values for arbitrary irreducible representations of $U(N)$, so it is convenient to have a generating functional that encodes all the information about them. We will for simplicity consider the case in which one has just a single knot. We then have to find a suitable basis for the Wilson loop operators. There are two natural basis for the problem: the basis labelled by representations R, and the basis labelled by conjugacy classes $C(\vec{k})$ of the symmetric group. Let U be the holonomy of the gauge connection around the knot $\mathcal{K}$, and consider the operator $\Upsilon_{\vec{k}}(U)$ defined as in (4.43). The vevs of these operators give the "$\vec{k}$-basis" for the vacuum expectation values of the Wilson loops:

$$W_{\vec{k}} = \langle \Upsilon_{\vec{k}}(U) \rangle = \sum_R \chi_R(C(\vec{k})) W_R \tag{2.68}$$

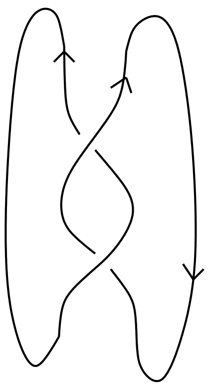


Figure 6: The Hopf link can be regarded as two unknots linked through one relative twist. The two unknots can be fused together into an unknot with framing one.

where χ_R are characters of the permutation group S_ℓ in the representation R, and we have used Frobenius formula

$$\mathrm{Tr}_R(U) = \sum_{\vec{k}} \frac{\chi_R(C(\vec{k}))}{z_{\vec{k}}} \Upsilon_{\vec{k}}(U), \tag{2.69}$$

with

$$z_{\vec{k}} = \prod_j k_j! j^{k_j}. \tag{2.70}$$

If V is a $U(M)$ matrix (a "source" term), one can define the following operator, which was introduced in [73] and is known in this context as the Ooguri-Vafa operator:

$$Z(U,V) = \exp\Big[\sum_{n=1}^{\infty} \frac{1}{n} \mathrm{Tr}\, U^n \,\mathrm{Tr}\, V^n\Big]. \tag{2.71}$$

When expanded, this operator can be written in the k-basis as follows,

$$Z(U,V) = 1 + \sum_{\vec{k}} \frac{1}{z_{\vec{k}}} \Upsilon_{\vec{k}}(U) \Upsilon_{\vec{k}}(V). \tag{2.72}$$

We see that $Z(U,V)$ includes all possible Wilson loop operators $\Upsilon_{\vec{k}}(U)$ associated to a knot $\mathcal{K}$. One can also use Frobenius formula to show that

$$Z(U,V) = \sum_R \mathrm{Tr}_R(U) \mathrm{Tr}_R(V), \tag{2.73}$$

where the sum over representations starts with the trivial one. In $Z(U,V)$ we assume that U is the holonomy of a dynamical gauge field and that V is a source. The vacuum expectation value $Z(V) = \langle Z(U,V) \rangle$ has then information about the vevs of the Wilson loop operators, and by taking its logarithm one can define the connected vacuum expectation values $W^{(c)}_{\vec{k}}$:

$$F_{\rm CS}(V) = \log Z(V) = \sum_{\vec{k}} \frac{1}{z_{\vec{k}}!} W^{(c)}_{\vec{k}} \Upsilon_{\vec{k}}(V) \tag{2.74}$$

One has, for example:

$$W^{(c)}_{(2,0,\cdots)} = \langle ({\rm Tr} U)^2 \rangle - \langle {\rm Tr} U \rangle^2 = W_{\square\!\square} + W_{\substack{\square \\ \square}} - W^2_{\square}.$$

The free energy $F_{\rm CS}(V)$, which is a generating functional for connected vevs $W^{(c)}_{\vec{k}}$, will be the relevant object for the duality with topological strings.

3 The $1/N$ expansion and Chern-Simons theory

3.1 The $1/N$ expansion: general aspects and examples

In quantum field theory, the usual perturbative expansion gives a series in powers of the coupling constants of the model. However, in theories with a $U(N)$ or $SU(N)$ gauge symmetry there is an extra parameter that enters into the game, namely N, and it turns out that there is a way to express the free energy and the correlation functions of the theory as power series in $1/N$. This is the $1/N$ expansion introduced by 't Hooft in [84], in the context of QCD.

A good starting point to construct the $1/N$ expansion is the usual perturbative expansion. The N dependence of the perturbative expansion comes from the group factors of Feynman diagrams, but it is clear that a single Feynman diagram gives rise to a polynomial in N involving different powers of N. Therefore, the standard Feynman diagrams, which are good in order to keep track of powers of the coupling constants, are not good in order to keep track of powers of N. What we have to do is to "split" each diagram into different pieces which correspond to a definite power of N. To do that, one writes the Feynman diagrams of the theory as "fatgraphs" or ribbon graphs [84]. Let us explain this

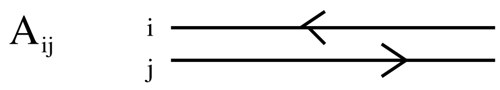


Figure 7: The index structure of the fields in the adjoint representation of $U(N)$ is represented through a double line.

in some detail. Let us assume that we have a quantum field theory with a $U(N)$ gauge symmetry, in such a way that the fundamental field A_{ij} is in the adjoint representation. Since the adjoint representation is the tensor product of the fundamental N and the antifundamental $\overline{N}$, we can look at i (j) as an index of the fundamental (antifundamental) representation, respectively. We will represent this double-index structure by a double line notation as shown in Fig. 7. The only thing we have to do now is to rewrite the Feynman rules of the theory concerning the group structure while taking into account this double-line notation. For example, the kinetic term of the theory is of the form

$$\frac{1}{x}\mathrm{Tr}\, A\mathcal{D}A = \frac{1}{x}A_{ij}\mathcal{D}A_{ji} \tag{3.1}$$

where x is the coupling constant, and $\mathcal{D}$ is a differential operator (in Chern-Simons theory, $\mathcal{D} = d$, the exterior differential). This means that, at the level of group structure, the propagator of the theory is of the form

$$\langle A_{ij}A_{kl}\rangle = x\delta_{il}\delta_{jk} \tag{3.2}$$

and can be represented in the double line notation as in Fig. 8. Next, we consider the

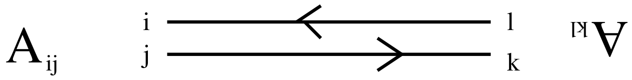
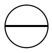


Figure 8: The propagator in the double line notation.

vertices of the theory. The trivalent vertex of Chern-Simons theory is given by,

$$\frac{1}{x}\mathrm{Tr}(A\wedge A\wedge A) = \frac{1}{x}A_{ij}\wedge A_{jk}\wedge A_{ki}, \tag{3.3}$$

which in the double line notation can be represented as in Fig. 9, where again we have just focused on the group structure of the interaction.

Once we have rewritten the Feynman rules in the double-line notation, we can construct the corresponding graphs, which look like ribbons and are called ribbon graphs or fatgraphs. It is clear that in general a usual Feynman diagram can give rise to many different fatgraphs. Consider for example the one-loop vacuum diagram  in Chern-Simons theory, which comes from contracting two cubic vertices. In the double line notation the contraction can be done in two different ways. The first one is illustrated in Fig. 10 and gives a group factor

$$\sum_{ijkmnp}\langle A_{ij}A_{mn}\rangle\langle A_{jk}A_{pm}\rangle\langle A_{ki}A_{np}\rangle = N^3. \tag{3.4}$$

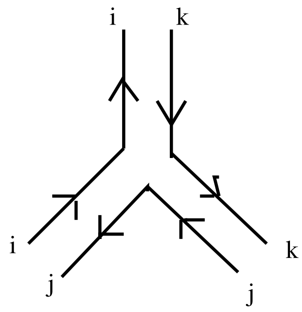


Figure 9: The cubic vertex in the double line notation.

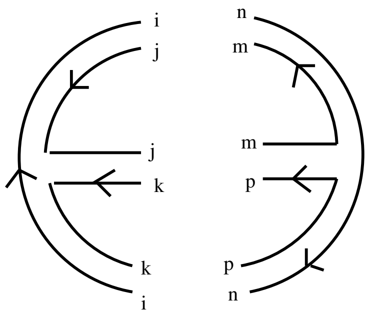
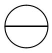


Figure 10: Contracting two cubic vertices in the double line notation: the N^3 contribution.

The second one is shown in Fig. 11 and gives a group factor

$$\sum_{ijkmnp} \langle A_{ij} A_{mn} \rangle \langle A_{jk} A_{np} \rangle \langle A_{ki} A_{pm} \rangle = N. \tag{3.5}$$

The group factor of the original Feynman diagram  is given by $2N(N^2-1)$. The N^3 piece comes from the fatgraph in Fig. 10, while the N piece comes from Fig. 11. In this way we have split the original diagram into two different fatgraphs with a well-defined power of N associated to them. The number of factors of N is simply equal to the number of closed loops in the graph: there are three closed lines in the fatgraph resulting from the contractions in Fig. 10 (see the first graph in Fig. 13), while there is only one in the diagram resulting from Fig. 11 (see Fig. 14). In general, fatgraphs turn out to be characterized topologically by the number of propagators E, the number of vertices V, and finally the number of closed loops, h. Each propagator gives a power of x, while each interaction vertex gives a power of x^{-1}. As we saw in the examples above, each closed

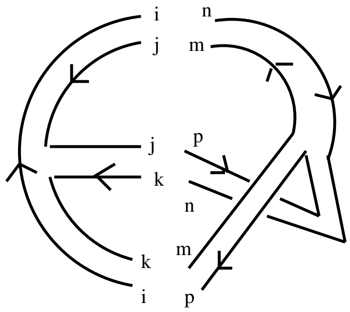


Figure 11: Contracting two cubic vertices in the double line notation: the N contribution.

loop gives a power of N and therefore the diagram will have in total a factor

$$x^{E-V}N^h. \tag{3.6}$$

The key point now is to regard the fatgraph as a Riemann surface with holes, in which each closed loop represents the boundary of a hole. The genus of such a surface is determined by the elementary topological relation

$$2g-2=E-V+h \tag{3.7}$$

therefore we can write (3.6) as

$$x^{2g-2+h}N^h=x^{2g-2}t^h \tag{3.8}$$

where we have introduced the 't Hooft parameter $t=Nx$. A fatgraph corresponding to a Riemann surface with genus one and $h=9$ holes is shown in Fig. 12. The fatgraphs with $g=0$ are called *planar*, while the ones with $g>0$ are called *nonplanar*. The graph giving the N^3 contribution in Fig. 10 is planar: it has $E=3$, $V=2$ and $h=3$, therefore $g=0$, and it is a sphere with three holes. The graph in Fig. 11 is nonplanar: it has $E=3$, $V=2$ and $h=1$, therefore $g=1$, and represents a torus with one hole (it is easy to see this by drawing the diagram on the surface of a torus). We can now organize the computation of the different quantities in the field theory in terms of fatgraphs. For example, the computation of the free energy is given in the usual perturbative expansion by connected vacuum bubbles. When the vacuum bubbles are written in the double line notation, we find that the perturbative expansion of the free energy is given by

$$F=\sum_{g=0}^{\infty}\sum_{h=1}^{\infty}F_{g,h}x^{2g-2}t^h, \tag{3.9}$$

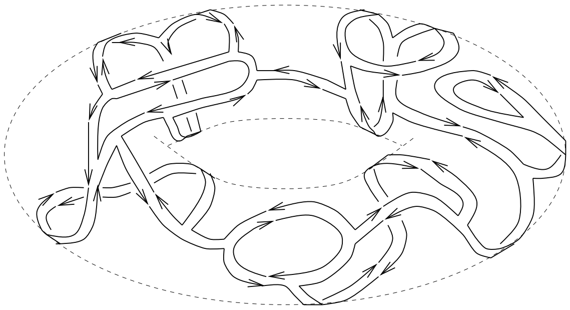


Figure 12: This figure shows a fatgraph with $h = 9$ and $g = 1$.

where $F_{g,h}$ is simply a number that can be computed by the usual rules of perturbation theory. In (3.9) we have written the diagrammatic series as an expansion in x around $x = 0$, keeping the 't Hooft parameter $t = xN$ fixed. Equivalently, we can regard it as an expansion in $1/N$, keeping t fixed, and then the N dependence appears as N^{2-2g}. Therefore, for t fixed and N large, the leading contribution comes from planar diagrams, which go like $\mathcal{O}(N^2)$. The nonplanar diagrams give subleading corrections. Notice that the contribution of a given order in N (or in x) is given by an infinite series where we sum over all possible numbers of holes h, weighted by t^h.

In Chern-Simons theory we are particularly interested in computing the vacuum expectation values of Wilson loop operators. The $1/N$ expansion of Wilson loops can be easily analyzed (see for example [21]), and it turns out that the vevs that have a well-defined behaviour in the $1/N$ expansion are the connected vevs $W_{\vec{k}}^{(c)}$ introduced in (2.74). One finds that these vevs admit an expansion of the form,

$$W_{\vec{k}}^{(c)} = \sum_{g=0}^{\infty} W_{g,\vec{k}}(t) x^{2g-2+|\vec{k}|}, \tag{3.10}$$

where $W_{g,\vec{k}}(t)$ is a function of the 't Hooft parameter.

3.2 Matrix models

A simple context to study the $1/N$ expansion is a gauge theory in zero dimensions or *matrix model* (for a detailed exposition on matrix models, see for example [24, 23]). In matrix models the complications of spacetime dependence in the computation of the perturbative expansion are absent, and the $1/N$ expansion is reduced to its bare bones: the combinatorics of graphs. Moreover, as we will see later, in matrix models there is a beautiful way of resumming the diagrams of fixed genus over all numbers of holes. A third reason to consider matrix models in our context is that Chern-Simons theory on certain three-manifolds can be represented as a matrix model [63, 2], as we will see later on.

In a matrix model, the basic field is simply a $U(N)$ matrix M, which we will take to be Hermitian, and we will consider an action of the form:

$$S(M) = \frac{1}{2x}\mathrm{Tr}M^2 + \frac{1}{3x}\mathrm{Tr}M^3, \tag{3.11}$$

where the first term, which is quadratic in M, plays the role of the kinetic term, and the second term is a cubic interaction term. The theory has the obvious gauge symmetry

$$M \to UMU^{-1}, \tag{3.12}$$

where U is a $U(N)$ matrix. The partition function of the theory is given by

$$Z = \frac{1}{\mathrm{vol}(U(N))}\int dM\, e^{-S(M)} \tag{3.13}$$

where the factor $\mathrm{vol}(U(N))$ is the usual volume factor of the gauge group that arises after fixing the gauge. In other words, we are considering here a *gauged* matrix model. The measure in the "path integral" is the Haar measure

$$dM = \prod_i dM_{ii} \prod_{i,j} d\mathrm{Re}M_{ij} d\mathrm{Im}M_{ij}. \tag{3.14}$$

The partition function of the cubic matrix model can be evaluated in perturbation theory by expanding around the Gaussian point, bringing down powers of $\mathrm{Tr}M^3$, and using Wick's theorem. Let us spell this out in some detail. The Gaussian matrix model is defined by the partition function

$$Z_{\mathrm{Gaussian}} = \frac{1}{\mathrm{vol}(U(N))}\int dM\, e^{-\frac{1}{2x}\mathrm{Tr}M^2}, \tag{3.15}$$

We will denote by

$$\langle f(M)\rangle = \frac{\int dM\, f(M)\, e^{-\mathrm{Tr}M^2/2x}}{\int dM\, e^{-\mathrm{Tr}M^2/2x}} \tag{3.16}$$

the normalized vevs of a functional $f(M)$ in the Gaussian matrix model. By expanding $\exp(\mathrm{Tr}M^3/3x)$ in (3.13), we can compute the partition function perturbatively in terms of vevs like (3.16). Of course, we can interpret this computation in terms of Feynman diagrams, and as usual the perturbative expansion of the free energy

$$F = \log Z$$

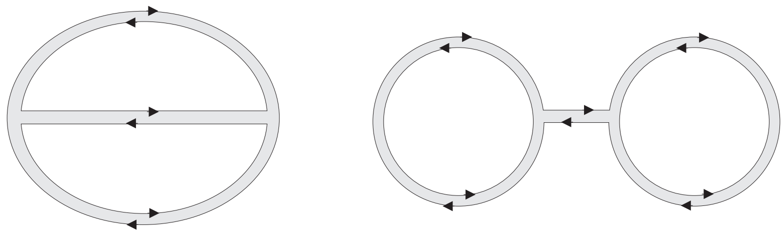


Figure 13: Two planar diagrams in the cubic matrix model.

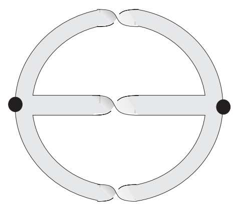


Figure 14: A nonplanar diagram in the cubic matrix model.

will only involve connected vacuum bubbles, but written in the double-line notation.

Exercise. Show that in the cubic matrix model the expansion of the free energy reads:

$$F - F_{\rm Gaussian} = \frac{2}{3}xN^3 + \frac{1}{6}xN + \cdots \tag{3.17}$$

where the first term corresponds to the two planar diagrams shown in Fig. 13, and the second term corresponds to the nonplanar diagram shown in Fig. 14.

One can also consider more general actions of the form

$$S(M) = \frac{1}{2x}{\rm Tr}M^2 + \frac{1}{x}{\rm Tr}V(M), \qquad V(M) = \sum_{p\geq 3} \frac{g_p}{p}M^p. \tag{3.18}$$

In this case we have vertices of various types (cubic, quartic, and so on). In the evaluation of the free energy we consider all possible connected vacuum diagrams in the double line notation. Denote by V_p the number of vertices with p-legs in such a diagram, $V = \sum_p V_p$ the total number of vertices, P the number of propagators, and h the number of loops. The weight of such a diagram will then be given by

$$x^{2g-2}t^h \prod_p g_p^{V_p} \tag{3.19}$$

where we used the topological relation $2 - 2g = V - P + h$ again.

There is an alternative way of writing the matrix model partition function which is very useful. The original matrix model variable has N^2 real parameters, but using the gauge symmetry we can see that, after modding by gauge transformations, there are only N parameters left. We can for example use our gauge freedom to diagonalize the matrix:

$$M \to UMU^{-1} = D, \tag{3.20}$$

with $D = \mathrm{diag}(\lambda_1, \cdots, \lambda_N)$, and impose this as a gauge choice. We can now use standard Faddeev-Popov techniques in order to compute the gauge-fixed integral (see for example [15]). The gauge fixing (3.20) leads to consider the delta-function constraint

$$\delta({}^U M) = \prod_{i<j} \delta^{(2)}({}^U M_{ij}) \tag{3.21}$$

where ${}^U M = UMU^\dagger$. We then introduce

$$\Delta^{-2}(M) = \int dM \, \delta({}^U M). \tag{3.22}$$

It then follows that the integral of any gauge-invariant function $f(M)$ can be written as

$$\int dM \, f(M) \Delta^2(M) \int dU \, \delta({}^U M) = \Omega_N \int \prod_{i=1}^N d\lambda_i \Delta^2(\lambda) f(\lambda), \tag{3.23}$$

where we have used the gauge invariance of $\Delta(M)$, and

$$\Omega_N = \int dU \tag{3.24}$$

is proportional to the volume of the gauge group $U(N)$, and we will compute it shortly. We just have to evaluate the the factor $\Delta(\lambda)$. Notice that, if we write $U = \mathbf{1} + A + \cdots$, where A is a antiHermitian matrix, we find

$$f_{ij}(A) = (UDU^\dagger)_{ij} = A_{ij}(\lambda_i - \lambda_j) + \cdots. \tag{3.25}$$

The standard Faddeev-Popov formula

$$\Delta^2(M) = \det \Big| \frac{\delta f}{\delta A} \Big|_{f=0} \tag{3.26}$$

leads immediately to

$$\Delta^2(\lambda) = \prod_{i<j} (\lambda_i - \lambda_j)^2, \tag{3.27}$$

the square of the Vandermonde determinant. Finally, we fix the factor Ω_N as follows. The Gaussian partition function can be computed explicitly by using the Haar measure

$$\int dM \, e^{-\frac{1}{2g_s} \mathrm{Tr}\, M^2} = 2^{N/2} (\pi g_s)^{N^2/2}. \tag{3.28}$$

On the other hand, this should equal

$$\Omega_N \int \prod_{i=1}^{N} \Delta^2(\lambda) e^{-\frac{1}{2g_s}\sum_{i=1}^{N}\lambda_i^2} \tag{3.29}$$

The integral over eigenvalues can be evaluated in various ways, using for example the Selberg function or Hermite polynomials [67], and its value is

$$g^{N^2/2}(2\pi)^{N/2} G_2(N+2) \tag{3.30}$$

where $G_2(z)$ is the Barnes function, defined by

$$G_2(z+1) = \Gamma(z) G_2(z), \quad G_2(1) = 1. \tag{3.31}$$

Comparing these results, we find that

$$\Omega_N = \frac{2^N \pi^{\frac{N(N+1)}{2}}}{G_2(N+2)} = 2^{-\frac{N(N-1)}{2}} \frac{\mathrm{vol}(U(N))}{N!} \tag{3.32}$$

where (see for example [74]):

$$\mathrm{vol}(U(N)) = \frac{(2\pi)^{\frac{1}{2}N(N+1)}}{G_2(N+1)}. \tag{3.33}$$

We then see that, up to a numerical factor,

$$\frac{1}{\mathrm{vol}(U(N))} \int dM\, f(M) = \frac{1}{N!} \int \prod_{i=1}^{N} d\lambda_i\, \Delta^2(\lambda) f(\lambda). \tag{3.34}$$

The factor $N!$ in the r.h.s. of (3.34) has an obvious interpretation: after fixing the gauge symmetry of the matrix integral by fixing the diagonal gauge, there is still a residual symmetry given by the Weyl symmetry of $U(N)$, which is the symmetric group S_N acting as permutation of the eigenvalues. The "volume" of this discrete gauge group is just its order, $|S_N| = N!$, and since we are considering gauged matrix models we have to divide by it as shown in (3.34).

3.3 The $1/N$ expansion in Chern-Simons theory

In the previous subsection we illustrated the $1/N$ expansion in a very simple model - gauge theory in zero dimensions. However, as we explained in detail, the structure of the expansion (3.9) is valid for any $U(N)$ gauge theory, in particular one should be able to expand the free energy of Chern-Simons theory on the three-sphere in this way. Of course, the computation of $F_{g,h}$ in Chern-Simons theory directly from perturbation theory is more difficult, since it involves not only combinatorial factors, but also integrals of products of propagators over the three-sphere. However, since we now the exact answer for the partition function, we just have to expand it to obtain (3.9) and the explicit expression for $F_{g,h}$.

The partition function of CS with gauge group $U(N)$ on the sphere is a slight modification of (2.28):

$$Z = \frac{1}{(k+N)^{N/2}} \prod_{\alpha>0} 2\sin\Big(\frac{\pi(\alpha\cdot\rho)}{k+N}\Big). \tag{3.35}$$

and differs from it in an overall factor $N^{1/2}/(k+N)^{1/2}$ which is the partition function for the $U(1)$ factor (recall that $U(N) = U(1) \otimes SU(N)/\mathbf{Z}_N$). Using the explicit description of the positive roots of $SU(N)$, one gets

$$F = \log Z = -\frac{N}{2}\log(k+N) + \sum_{j=1}^{N-j} \log\Big[2\sin\frac{\pi j}{k+N}\Big]. \tag{3.36}$$

We can now write the sin as

$$\sin \pi z = \pi z \prod_{n=1}^{\infty}\Big(1 - \frac{z^2}{n^2}\Big), \tag{3.37}$$

and we find that the free energy is the sum of two pieces. One of them is the nonperturbative piece:

$$F^{\rm np} = -\frac{N^2}{2}\log(k+N) + \frac{1}{2}N(N-1)\log 2\pi + \sum_{j=1}^{N-1}(N-j)\log j, \tag{3.38}$$

and the other piece is the perturbative one:

$$F^{\rm p} = \sum_{j=1}^{N-1}(N-j)\sum_{n=1}^{\infty}\log\Big[1 - \frac{j^2 g_s^2}{4\pi^2 n^2}\Big], \tag{3.39}$$

where g_s corresponds to the open string coupling constant and it is given by (5.16). To see that (3.38) corresponds to the nonperturbative piece of the free energy, we recall (3.33) and notice that

$$F^{\rm np} = \log \frac{(2\pi g_s)^{\frac{1}{2}N^2}}{{\rm vol}(U(N))} \tag{3.40}$$

so it is given by the log of (2.12), where $A^{(c)}$ is in this case the trivial flat connection. Therefore, $F^{\rm np}$ is the log of the prefactor of the path integral, which is not captured by Feynman diagrams.

Let us work out the perturbative piece (3.39). By expanding the log, using that $\sum_{n=1}^{\infty} n^{-2k} = \zeta(2k)$, and the formula

$$\sum_{j=1}^{N-1} j^{2k} = -\frac{N^{2k}}{2} + \sum_{l=0}^{k} \binom{2k+1}{2l} \frac{B_{2l}}{2k+1} N^{2k+1-2l} \tag{3.41}$$

we find that (3.39) can be written as

$$F^{\rm p} = \sum_{g=0}^{\infty} \sum_{h=2}^{\infty} F^{\rm p}_{g,h} g_s^{2g-2+h} N^h, \tag{3.42}$$

where $F^{\rm p}_{g,h}$ is given by:

$$\begin{aligned} F^{\rm p}_{0,h} &= \frac{B_{h-2}}{(h-2)h!} \\ F^{\rm p}_{1,h} &= \frac{1}{12}\frac{B_h}{h\, h!} \\ F^{\rm p}_{g,h} &= 2\frac{\zeta(2g-2+h)}{(2\pi)^{2g-2+h}} \binom{2g-3+h}{h} \frac{B_{2g}}{2g(2g-2)}, \quad g \geq 2. \end{aligned} \tag{3.43}$$

This gives the contribution of connected diagrams with two loops and beyond to the free energy of Chern-Simons on the sphere. The nonperturbative piece also admits an asymptotic expansion that can be can be easily worked out by expanding the Barnes function [76, 74]. One finds:

$$F^{\rm np} = \frac{N^2}{2}\Big(\log(N g_s) - \frac{3}{2}\Big) - \frac{1}{12}\log N + \zeta'(-1) + \sum_{g=2}^{\infty} \frac{B_{2g}}{2g(2g-2)} N^{2-2g}. \tag{3.44}$$

3.4 The string interpretation of the $1/N$ expansion

The expansion (3.9) of the free energy in a $U(N)$ gauge theory looks very much like the perturbative expansion of an *open* string theory with $U(N)$ Chan-Paton factors, where x is the open string coupling constant, and $F_{g,h}$ corresponds to some amplitude on a Riemann surface of genus g with h holes. There is in fact a way to produce a *closed* string theory interpretation of the free energy of gauge theories: let us introduce the function

$$F_g(t) = \sum_{h=1}^{\infty} F_{g,h} t^h. \tag{3.45}$$

The free energy can be written now as

$$F = \sum_{g=0}^{\infty} x^{2g-2} F_g(t), \tag{3.46}$$

which looks now like a *closed* string expansion where t is some modulus of the theory. Notice that in writing (3.45) we have grouped together all open Riemann surfaces with the same bulk topology but with different number of holes, so by "summing over all holes" we are "filling up the holes" to produce a closed Riemann surface. This leads to 't Hooft idea [84] that, given a gauge theory, one should be able to find a string theory interpretation in the way we have described, namely, the fatgraph expansion of the free energy is resummed to give a function of the 't Hooft parameter $F_g(t)$ at every genus which is then interpreted as a closed string amplitude. For example, the planar sector of the gauge theory corresponds to the closed string at tree level.

Once the closed string intepretation is available, the $1/N$ expansion (3.10) can be regarded as an open string expansion, where $F_{g,\vec{k}}(t)$ are interpreted as amplitudes in an open string theory at genus g and with $h = |\vec{k}|$ holes. The vector k specifies the winding numbers of the holes around a one-cycle in the target space of the theory, according to the rule we gave in subsection 2.3. So the Wilson loops "creates" a one-cycle in the target space where the boundaries of Riemann surfaces can end, and the generating functional for connected vevs (2.74) is interpreted as the total free energy of an open string, as in (4.40). These open strings shouldn't be confused with the ones that we associated to the expansion (3.9). The open strings underlying (3.10) should be regarded as an open string sector in the closed string theory associated to the resummed expansion (3.46).

So far the construction of the functions $F_g(t)$ seems purely formal, and the definition in (3.46) express them as a power series in t, whose coefficients have to be computed order by order in perturbation theory. In Chern-Simons theory on $\mathbf{S}^3$ we know the exact answer for the partition function, so the expansion can be done easily and we have obtained general results for $F_{g,h}$. In the case of matrix models, one can in fact compute $F_g(t)$ in closed form for many different potentials. In order to do that, we will have to introduce some more matrix model technology.

3.5 Matrix model technology I

In this section we will show that the planar free energy $F_0(t)$, which is a resummation of all planar fatgraphs, can be computed in a closed form, following the seminal paper [16].

Let us consider a general matrix model with action $V(M)$, and let us write it in the eigenvalue basis as follows:

$$Z = \frac{1}{N!} \int \prod_{i=1}^{N} d\lambda_i e^{N^2 S_{\rm eff}(\lambda)} \tag{3.47}$$

where the effective action is given by

$$S_{\text{eff}}(\lambda) = \frac{1}{tN}\sum_{i=1}^{N} V(\lambda_i) + \frac{1}{N^2}\sum_{i\neq j}\log(\lambda_i - \lambda_j). \tag{3.48}$$

Notice that, since a sum over N eigenvalues is roughly of order N, the effective action is of order $\mathcal{O}(1)$. We can now regard N^2 as a sort of $\hbar^{-1}$ in such a way that, as $N \to \infty$, the integral (3.47) will be dominated by the saddle-point configuration that extremizes the effective action. Varying $S_{\text{eff}}(\lambda)$ w.r.t. the eigenvalue λ_i, we obtain the equation

$$\frac{1}{t}V'(\lambda_i) = \frac{1}{N}\sum_{j\neq i}\frac{1}{\lambda_i - \lambda_j}. \tag{3.49}$$

In the large N limit, it is also reasonable to expect that the configuration of eigenvalues λ_i that solves the above equation becomes a continuous distribution $\rho(\lambda)$. Formally, the eigenvalue distribution is defined for finite N as

$$\rho(\lambda) = \frac{1}{N}\sum_{i=1}^{N}\delta(\lambda - \lambda_i). \tag{3.50}$$

In the large N limit, the distribution will become a function $\rho(\lambda)$ defined on some interval on the real line $\mathcal{C}$. We will assume that $\mathcal{C}$ is just made of one connected interval (the so-called one-cut solutions).

Qualitatively, what is going on is the following. Assume for simplicity that $V(x)$, the potential, has only one minimum x_*. We can regard the eigenvalues as coordinates of a system of N classical particles moving in the real line. The equation (3.49) says that these particles are subject to an effective potential

$$V_{\text{eff}}(\lambda_i) = \frac{1}{t}V(\lambda_i) - \frac{1}{N}\sum_{j\neq i}\log|\lambda_i - \lambda_j| \tag{3.51}$$

which involves a logarightmic Coulomb repulsion between eigenvalues. For small ‘t Hooft parameter, the potential term dominates over the Coulomb repulsion, and the particles tend to be in the minimum x_* of the potential $V'(x_*) = 0$. This means that, for $t = 0$, the interval $\mathcal{C}$ will be collapsed to the point x_*. As t grows, the Coulomb repulsion will force the eigenvalues to be apart from each other and to spread out over an interval $\mathcal{C}$.

We can now write the saddle-point equation in terms of continuum quantities, by using the standard rule

$$\frac{1}{N}\sum_{i=1}^{N} f(\lambda_i) \to \int_{\mathcal{C}} f(\lambda)d\lambda. \tag{3.52}$$

Notice first that the distribution of eigenvalues $\rho(\lambda)$ satisfies the normalization condition

$$\int_{\mathcal{C}} \rho(\lambda) d\lambda = 1. \tag{3.53}$$

The equation (3.49) becomes then

$$\frac{1}{2t} V'(\lambda) = \mathrm{P} \int \frac{\rho(\lambda')d\lambda'}{\lambda - \lambda'} \tag{3.54}$$

where P denotes the principal value of the integral. The above equation is an integral equation that allows in principal to compute $\rho(\lambda)$ given the potential $V(\lambda)$, as a function of the 't Hooft parameter t and the coupling constants. Once $\rho(\lambda)$ is known, one can easily compute $F_0(t)$: in the saddle-point approximation, the free energy is given by

$$F = N^2 S_{\mathrm{eff}}(\rho) + \mathcal{O}(1), \tag{3.55}$$

where the effective action in the continuum limit is a functional of ρ that can be easily obtained to be

$$S_{\mathrm{eff}}(\rho) = \frac{1}{t} \int_{\mathcal{C}} d\lambda \rho(\lambda) V(\lambda) + \int_{\mathcal{C}\times\mathcal{C}} d\lambda d\lambda' \rho(\lambda)\rho(\lambda') \log(\lambda - \lambda'). \tag{3.56}$$

Therefore, the planar free energy is given by

$$F_0(t) = t^2 S_{\mathrm{eff}}(\rho), \tag{3.57}$$

where the effective action is evaluated on the distribution of eigenvalues which solves (3.54).

The question now is how to find the solution to the saddle-point equation (3.54). This equation is a singular integral equation which has been studied in detail in other contexts of physics (see, for example, [71]). The way to solve it is to introduce an auxiliary function called the *resolvent*. The resolvent is defined as a correlator in the matrix model:

$$W(p) = \frac{1}{N} \langle \mathrm{Tr} \frac{1}{p - M} \rangle, \tag{3.58}$$

which is in fact a generating functional of correlation functions of "Wilson loops":

$$W(p) = \frac{1}{N} \sum_{k=0}^{\infty} \langle \mathrm{Tr} M^k \rangle p^{-k-1} \tag{3.59}$$

Being a generating functional of connected correlators, it admits a $1/N$ expansion of the form:

$$W(p) = \sum_{g=0}^{\infty} N^{-2g} W_g(p), \tag{3.60}$$

and the genus zero piece can be written in terms of the eigenvalue density as

$$W_0(p) = \int d\lambda \frac{\rho(\lambda)}{p-\lambda} \tag{3.61}$$

The genus zero resolvent has three important properties. First of all, as a function of p it is an analytic function on the whole complex plane except on the interval $\mathcal{C}$, since if $\lambda \in \mathcal{C}$ one has a singularity at $p = \lambda$. Second, due to the normalization property of the eigenvalue distribution, it has the asymptotic behaviour

$$W_0(p) \sim \frac{1}{p}, \qquad p \to \infty. \tag{3.62}$$

Finally, one can compute the discontinuity of $W_0(p)$ as one crosses the interval $\mathcal{C}$. This is just the residue at $p = \lambda$, and one then finds the key equation

$$\rho(\lambda) = \frac{1}{2\pi i}(W_0(p+i\epsilon) - W_0(p-i\epsilon)). \tag{3.63}$$

Therefore, if the resolvent at genus zero is known, the eigenvalue distribution follows from (3.63), and therefore one can compute the planar free energy. On the other hand, by looking again at the resolvent as we approach the discontinuity, we see that the r.h.s. of (3.54) is given by $-(W_0(p+i\epsilon)+W_0(p-i\epsilon))/2$, and we then find the equation

$$W_0(p+i\epsilon) + W_0(p-i\epsilon) = -\frac{1}{t}V'(p), \tag{3.64}$$

which determines the resolvent in terms of the potential. There is in fact a closed expression for the resolvent in terms of a contour integral [68] which is very useful. Let $\mathcal{C}$ be given by the interval $b \leq \lambda \leq a$. Then, one has:

$$W_0(p) = \frac{1}{2t}\oint_{\mathcal{C}} \frac{d\omega}{2\pi i}\frac{V'(\omega)}{p-\omega}\Big(\frac{(p-a)(p-b)}{(\omega-a)(\omega-b)}\Big)^{\frac{1}{2}}. \tag{3.65}$$

This equation is easily proven by converting (3.64) into a discontinuity equation:

$$\widehat{W}_0(p+i\epsilon) - \widehat{W}_0(p-i\epsilon) = -\frac{1}{t}\frac{V'(p)}{\sqrt{(p-a)(p-b)}}, \tag{3.66}$$

where $\widehat{W}_0(p) = W_0(p)/\sqrt{(p-a)(p-b)}$. This equation determines $W_0(p)$ to be given by (3.65) up to regular terms. Because of the asymptotics (3.62), these regular terms are absent.

The asymptotics of $W_0(p)$ gives also two more conditions. By taking $p \to \infty$, one finds that the r.h.s. of (3.65) behaves like $cp + d + \mathcal{O}(1/p)$. Requiring the vanishing of c and d imposes

$$\begin{aligned} \oint_{\mathcal{C}} \frac{V'(\omega)}{\sqrt{(\omega-a)(\omega-b)}} &= 0, \\ \oint_{\mathcal{C}} \frac{\omega V'(\omega)}{\sqrt{(\omega-a)(\omega-b)}} &= 2t. \end{aligned} \tag{3.67}$$

These equations are enough to determine the endpoints of the cuts, a and b.

One can now apply this technology to the simplest case, the Gaussian model with $V(M) = M^2/2$. The way to perform the integrals (3.65) and (3.67) is to deform the integration contour to infinity. Let us first look for the the position of the endpoints from (3.67). Deforming the contour to infinity and introducing the variable $z = 1/\omega$, we find that the first equation in (3.67) becomes:

$$\oint_0 \frac{dz}{2\pi i} \frac{1}{z^2} \frac{1}{\sqrt{(1-az)(1-bz)}} = 0, \tag{3.68}$$

where the contour is now around $z = 0$. Therefore $a + b = 0$, in accord with the symmetry of the potential. Taking this into account, the second equation becomes:

$$\oint_0 \frac{dz}{2\pi i} \frac{1}{z^3} \frac{1}{\sqrt{1-a^2z^2}} = 2t, \tag{3.69}$$

which gives

$$a = 2\sqrt{t}. \tag{3.70}$$

We see that the interval $\mathcal{C} = [-a, a] = [-2\sqrt{t}, 2\sqrt{t}]$ opens up as the 't Hooft parameter grows up, as expected. Let us now compute the resolvent. Deforming the contour to infinity we pick up a pole at $\omega = p$, and we find:

$$W_0(p) = \frac{1}{2}V'(p) - \frac{1}{2t}\sqrt{(p-a)(p-b)}M(p), \tag{3.71}$$

where

$$M(p) = \oint_0 \frac{dz}{2\pi i} \frac{V'(1/z)}{1-pz} \frac{1}{\sqrt{(1-az)(1-bz)}}. \tag{3.72}$$

The expression (3.71) for the one-cut resolvent is in fact valid for any polynomial potential $V(M)$. For the Gaussian model we immediately find

$$W_0(p) = \frac{1}{2t}\left(p - \sqrt{p^2-4t}\right), \tag{3.73}$$

and from the discontinuity equation we derive the density of eigenvalues

$$\rho(\lambda) = \frac{1}{2\pi t}\sqrt{4t - \lambda^2}. \tag{3.74}$$

The graph of this function is a semicircle of radius $2t^{\frac{1}{2}}$, and the above eigenvalue distribution is the famous Wigner-Dyson semicircle law.

Exercise. *Resolvent for the cubic matrix model.* Consider the cubic matrix model with potential $V(M) = M^2/2 + gM^3/3$. Derive an expression for the endpoints of the one-cut solution as a function of t, g, and find the resolvent and the planar free energy.

3.6 A matrix model for Chern-Simons theory on $\mathbf{S}^3$

The partition function of Chern-Simons theory on $\mathbf{S}^3$ is known exactly, and moreover one can expand it in the form required by the $1/N$ expansion. Moreover, as we will see in 5.6, one can resum the expansion at fixed genus g in order to find the functions $F_g(t)$. It also turns out that the partition function of Chern-Simons on $\mathbf{S}^3$, as well as on many other three-manifolds, can be represented as a matrix integral [63]. In the case of $\mathbf{S}^3$ this is not a big deal from a technical point of view since most of the physical information in $Z(\mathbf{S}^3)$ can be obtained by other means. For other three-manifolds like lens spaces and Seifert spaces, the matrix model representation is actually crucial in order to extract perturbative results around the trivial connection [63] and specially around nontrivial flat connections [2]. It is still instructive to consider the Chern-Simons matrix model on $\mathbf{S}^3$, and we will do that here to illustrate the considerations on matrix models that we presented before.

Although there are many ways to derive the Chern-Simons matrix models, in the case of $\mathbf{S}^3$ the easiest way is through direct computation. Consider the following integral:

$$Z = \frac{e^{-\frac{x}{12}N(N^2-1)}}{N!}\int \prod_{i=1}^{N}\frac{d\beta_i}{2\pi}\, e^{-\sum_i \beta_i^2/2x}\prod_{i<j}\Big(2\sinh\frac{\beta_i-\beta_j}{2}\Big)^2 \tag{3.75}$$

where x is given by (2.11). It can be easily seen that this reproduces the partition function of $U(N)$ Chern-Simons theory on $\mathbf{S}^3$, given in (3.35), and the derivation is left as an exercise.

Exercise. Use Weyl's formula (2.27) to write (3.75) as a Gaussian integral, and show that it reproduces (3.35).

The measure factor in (3.75)

$$\prod_{i<j}\Big(2\sinh\frac{\beta_i-\beta_j}{2}\Big)^2 \tag{3.76}$$

is not the standard Vandermonde discriminant, although it reduces to it for small separations among the eigenvalues. In fact, for very small x, the Gaussian potential in (3.75) will be narrow and one can expand the sinh in power series. At leading order we find the usual Gaussian matrix model, while the corrections to it can be evaluated systematically by computing correlators in the Gaussian theory. It is easy to see that this expansion is nothing but the original perturbative expansion of Chern-Simons theory, see [63] for details.

Here we will take a slightly different route in order to analyze the model. First of all, we want to write the above integral as a standard matrix integral with the usual Vandermonde discriminant. This can be achieved with the change of variables [85]

$$\exp(\beta_i + Nx) = \lambda_i. \tag{3.77}$$

It is easy to see that the above integral becomes, up to a constant,

$$Z = \frac{1}{N!} \int \prod_{i=1}^{N} d\lambda_i \, \Delta^2(\lambda) \, \exp\Bigl(-\sum_i (\log \lambda_i)^2/2x\Bigr), \tag{3.78}$$

therefore we are considering the matrix model

$$\frac{1}{\mathrm{vol}(U(N))} \int dM \, e^{-\frac{1}{2x}(\log M)^2}. \tag{3.79}$$

Matrix integrals with logarithmic potentials are somewhat exotic, but have appeared before in connection with the Penner model [75] and with the $\mathbb{P}^1$ model [30]. As explained in detail in [85], the above matrix integral can be solved with orthogonal polynomials, the so-called Stieltjes-Wiegert polynomials. Here we are interested in solving the planar limit of the matrix integral (3.75). This was first done in [2], and here we will give a detailed computation following the techniques we introduced in 3.5.

We want to obtain the resolvent of the model in the planar limit, $W_0(p)$, and this amounts to computing the integral (3.65) with

$$V'(\omega) = \frac{\log \omega}{\omega}. \tag{3.80}$$

Notice that this potential has a minimum at $\omega = 1$. We then expect a one-cut solution where the endpoints of the interval $a(t)$, $b(t)$ will satisfy $a(0) = b(0) = 1$. In order to compute the integral (3.65) we deform the integration contour. In the cases of polynomial potentials, we picked a residue at $\omega = p$ and at infinity. Here, since the logarithm has a branch cut, we can not push the contour to infinity. Instead, we deform the contour as indicated in Fig. 15: we pick the pole at $p = \omega$, and then we surround the cut of the

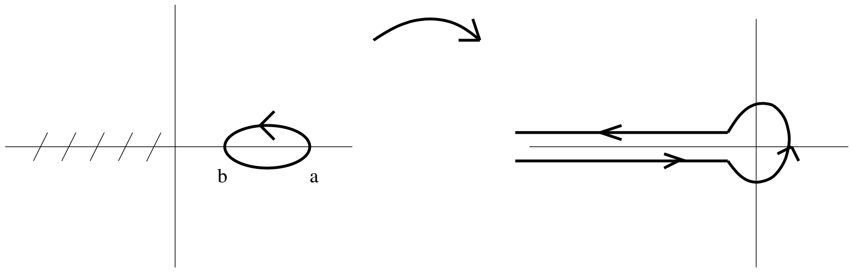


Figure 15: This shows the deformation of the contour needed to compute the planar resolvent of the Chern-Simons matrix integral. We pick a residue at $\omega = p$, and we have to encircle the singularity at the origin as well as the branch cut of the logarithm, which on the left hand side is represented by the dashed lines.

logarithm along the negative real axis and the singularity at $\omega = 0$ with a small circle C_ϵ of radius ϵ. This kind of integrals have also appeared in the solution of matrix models with the character expansion [50]. The resulting integrals are:

$$\frac{1}{2t}\left\{-\int_{-\infty}^{-\epsilon}\frac{ds}{s(s-h)\sqrt{(s-a)(s-b)}}+\oint_{C_\epsilon}\frac{ds\,\log s}{s(s-h)\sqrt{(s-a)(s-b)}}\right\}. \tag{3.81}$$

Both integrals are singular as $\epsilon \to 0$, but singularities cancel, and after some easy computations one finds:

$$W_0(p) = -\frac{1}{2tp}\log\left[\frac{(\sqrt{a}\sqrt{p-b}-\sqrt{b}\sqrt{p-a})^2}{(\sqrt{p-a}-\sqrt{p-b})^2p^2}\right]+\frac{\sqrt{(p-a)(p-b)}}{2tp\sqrt{ab}}\log\left[\frac{4ab}{2\sqrt{ab}+a+b}\right]. \tag{3.82}$$

In order to satisfy the asymptotics (3.62) the second term must vanish, and the first one must go like $1/p$. This implies:

$$\begin{aligned} 4ab &= 2\sqrt{ab}+a+b, \\ \sqrt{a}+\sqrt{b} &= 2\mathrm{e}^t, \end{aligned} \tag{3.83}$$

and from here we obtain the positions of the endpoints of the cut a, b as a function of the 't Hooft parameter:

$$\begin{aligned} a(t) &= 2e^{2t}-e^t+2e^{\frac{3t}{2}}\sqrt{e^t-1}, \\ b(t) &= 2e^{2t}-e^t-2e^{\frac{3t}{2}}\sqrt{e^t-1}. \end{aligned} \tag{3.84}$$

Notice that, for $t = 0$, $a(0) = b(0) = 1$, as expected. For large 't Hooft parameter, the endpoints of the cut behave like

$$a(t) \sim 4e^{2t}, \qquad , b(t) \sim \frac{1}{4}+\frac{e^{-t}}{8}. \tag{3.85}$$

The final expression for the resolvent is then:

$$W_0(p) = -\frac{1}{tp} \log\left[\frac{1 + e^{-t}p + \sqrt{(1 + e^{-t}p)^2 - 4p}}{2p}\right], \tag{3.86}$$

and from here we can easily find the density of eigenvalues

$$\rho(\lambda) = \frac{1}{\pi t \lambda} \tan^{-1}\left[\frac{1 + e^{-t}\lambda + \sqrt{(1 + e^{-t}\lambda)^2 - 4\lambda}}{2\lambda}\right]. \tag{3.87}$$

This expression for $\rho(\lambda)$ agrees with the density of zeroes of Stieltjes-Wiegert polynomials [53], and this is consistent with the fact that these polynomials solve the original matrix integral. One can now use the density to find $F_0(t)$, to find

$$F_0(t) = \sum_{n=1}^{\infty} \frac{e^{-nt}}{n^3} = \mathrm{Li}_3(e^{-t}) \tag{3.88}$$

where the polylogarithm of index j is given by:

$$\mathrm{Li}_j(x) = \sum_{n=1}^{\infty} \frac{x^n}{n^j}. \tag{3.89}$$

If we now define

$$v(\omega) = \mathrm{P} \int \frac{\rho(\omega)\omega d\omega}{p - \omega} \tag{3.90}$$

we find that

$$W_0(p) = \frac{1 + v(p)}{p},$$

and e^v solves the quadratic equation

$$(e^v - 1)(e^{v+u} - 1) + e^t - 1 = 0 \tag{3.91}$$

where we put $p = e^{u+t}$. The fact that v as defined in (3.90) satisfies (3.91) was found in [2] by a similar analysis. As we will see, (3.91) will have a beautiful string theory interpretation, when we understand the topological string dual of Chern-Simons theory on $\mathbf{S}^3$ in the $1/N$ expansion. We also mention that one can easily write a matrix integral for Chern-Simons theory for other gauge groups [63], and the corresponding models have been analyzed in [42]. As a final remark, notice that the consistency of the solution requires that $a(t), b(t) \geq 0$, so that the logarithmic potential of the matrix integral is well-defined. One could have singular behaviour if for some value of the 't Hooft coupling $b(t) = 0$, as it happens for example in $O(n)$ matrix models [32]. But since $b(t) > 0$ for all t this matrix model has no critical behaviour, therefore no double-scaling limit [17] can be taken, as it was argued in [29].

4 Topological strings

In this section we give a brief review of topological strings (the A-model) and what they compute.

4.1 Topological sigma models

The starting point to construct topological strings is an $\mathcal{N} = (2,2)$ superconformal field theory, the $\mathcal{N} = (2,2)$ nonlinear sigma model. This model can be twisted in two ways in order to produce a topological field theory [88, 55, 90], which are usually called the A and the B model. We will focus here on the A-model.

The field content of this model is the following. First, we have a map $x : \Sigma_g \to X$ from a Riemann surface of genus g to a target space X, that will be a Kähler manifold of complex dimension d. Indices in the tangent space of X will be denoted by i, with $i = 1, \cdots, 2d$. Since we have a complex structure, we will also have holomorphic and antiholomorphic indices, that we will denote respectively by $I, \overline{I}$, $i = 1, \cdots, d$. We also have fermions $\chi \in x^*(TX)$, which are scalars on Σ_g, and a fermionic one form ψ_α with values in $x^*(TX)$. This last field satisfies a selfduality condition which implies that its only nonzero components are $\psi^I_{\bar{z}} \in x^*(T^{(1,0)}X)$ and $\psi^{\overline{I}}_z \in x^*(T^{(0,1)}X)$. The action for the theory is:

$$\mathcal{L} = 2t \int_{\Sigma_g} d^2z \Big(\frac{1}{2} G_{ij} \partial_z x^i \partial_{\bar{z}} x^j + iG_{I\overline{J}} \psi^I_{\bar{z}} D_z \chi^{\overline{J}} + iG_{\overline{I}J} \psi^{\overline{I}}_z D_{\bar{z}} \chi^J - R_{I\overline{I}J\overline{J}} \psi^I_{\bar{z}} \psi^{\overline{I}}_z \chi^J \chi^{\overline{J}} \Big), \quad (4.1)$$

where d^2z is the measure $-idz \wedge d\bar{z}$, and t is a parameter that plays the role of $1/\hbar$. The theory also has a BRST, or topological, charge Q which acts on the fields according to

$$\begin{aligned} \{Q, x\} &= \chi, \\ \{Q, \chi\} &= 0, \\ \{Q, \psi^I_{\bar{z}}\} &= i\partial_{\bar{z}} x^I - \chi^J \Gamma^I_{JK} \psi^K_{\bar{z}}, \\ \{Q, \psi^{\overline{I}}_z\} &= i\partial_z x^{\overline{I}} - \chi^{\overline{J}} \Gamma^{\overline{I}}_{\overline{JK}} \psi^{\overline{K}}_{\bar{z}}. \end{aligned} \quad (4.2)$$

Finally, we also have a $U(1)$ ghost number symmetry, in which x, χ and ψ have ghost numbers 0, 1 and -1, respectively. Notice that the fermionic charge Q has then ghost number 1.

The action (4.1) turns out to be Q-exact, up to a topological term and terms that vanish on-shell [88, 90]:

$$\mathcal{L} = -i\{Q, V\} - t \int_{\Sigma_g} x^*(J), \quad (4.3)$$

where $J = iG_{I\overline{J}}dx^I \wedge dx^{\overline{J}}$ is the Kähler class of X, V (sometimes called the gauge fermion) is given by

$$V = t\int_{\Sigma_g} d^2z G_{I\overline{J}}(\psi^I_{\bar{z}}\partial_z c x^{\overline{J}} + \partial_{\bar{z}}x^I\psi^{\overline{J}}_z). \tag{4.4}$$

and

$$\int_{\Sigma_g} x^*(J) = i\int_{\Sigma_g} dz\wedge d\bar{z}\, G_{I\overline{J}}(\partial_z x^I\partial_{\bar{z}}x^{\overline{J}} - \partial_{\bar{z}}x^I\partial_z x^{\overline{J}}). \tag{4.5}$$

Notice that this term in (4.3) is a topological invariant characterizing the homotopy type of the map $x : \Sigma_g \to X$. We can also add a coupling to a B-field into the action,

$$-it\int_{\Sigma_g} x^*(B), \tag{4.6}$$

which will replace the Kähler form by the complexified Kähler form $\omega = J + iB$. Since the last term in (4.3) is topological, the energy-momentum tensor of this theory is given by:

$$T_{\alpha\beta} = \{Q, b_{\alpha\beta}\}, \tag{4.7}$$

where $b_{\alpha\beta} = \delta V/\delta g^{\alpha\beta}$ and has ghost number -1. The fact that the energy-momentum tensor is Q-exact means that the theory is topological, in the sense that the partition function does not depend on the background two-dimensional metric. This is easily proven: the partition function is given by

$$Z = \int \mathcal{D}X e^{-\mathcal{L}}, \tag{4.8}$$

where X denotes the set of fields of the theory, and we compute it in the background of a two-dimensional metric on the Riemann surface $g_{\alpha\beta}$. Since $T_{\alpha\beta} = \delta\mathcal{L}/\delta g^{\alpha\beta}$, we find that

$$\frac{\delta Z}{\delta g^{\alpha\beta}} = -\langle\{Q, b_{\alpha\beta}\}\rangle, \tag{4.9}$$

where the bracket denotes an unnormalized vacuum expectation value (vev). Since Q is a symmetry of the theory, the above vev vanishes, and we find that Z is metric-independent, at least formally.

Exercise. Show that $Q^2 = 0$ on-shell (*i.e.* modulo the equations of motion) and verify (4.3).

The Q-exactness of the action itself has also an important consequence: the same argument that we used above implies that the partition function of the theory is independent of t. Now, since t plays the role of $1/\hbar$, the limit of t large corresponds to the semiclassical approximation. Since the theory does not depend on t, the semiclassical approximation is *exact*. Notice that the classical configurations for the above action are holomorphic maps $x : \Sigma_g \to X$. These are the instantons of the nonlinear sigma model with a Kähler target, and minimize the bosonic action.

What are the operators to consider in this theory? Since the most interesting aspect of this model is the independence w.r.t. the metric, we want to look for operators whose correlation functions satisfy this condition. It is easy to see that the operators in the cohomology of Q do the job: topological invariance requires them to be Q-closed, and on the other hand they cannot be Q-exact, since otherwise their correlation functions will vanish. One can also check that the Q-cohomology is given by operators of the form

$$\mathcal{O}_\phi = \phi_{i_1 \cdots i_p} \chi^{i_1} \cdots \chi^{i_p}, \tag{4.10}$$

where $\phi = \phi_{i_1 \cdots i_p} dx^{i_1} \cdots dx^{i_p}$ is a closed p-form representing a nontrivial class in $H^p(X)$. Therefore, in this case the Q-cohomology is in one-to-one correspondence with the de Rham cohomology of the target manifold X. Also notice that the degree of the differential form corresponds to the ghost number of the operator. Moreover, one can derive a selection rule for correlation functions of such operators: the vacuum expectation value $\langle \mathcal{O}_{\phi_1} \cdots \mathcal{O}_{\phi_\ell} \rangle$ vanishes unless

$$\sum_{k=1}^{\ell} \deg(\mathcal{O}_{\phi_k}) = 2d(1-g) + \int_{\Sigma_g} x^*(c_1(X)), \tag{4.11}$$

where $\deg(\mathcal{O}_{\phi_k}) = \deg(\phi)$. This selection rule corresponds to the fact that the $U(1)$ current is anomalous, and the anomaly is given by the r.h.s. of (4.11), which calculates the number of zero modes of the Dirac operator (in other words, the r.h.s. is minus the ghost number of the vacuum). As usual in quantum field theory, the operators with nontrivial vacuum expectation values have to soak up the zero modes associated to the anomaly. From the mathematical point of view, the number of zero modes is nothing but the virtual dimension of the moduli space of holomorphic maps, $\mathcal{M}^{\rm hol}_{\Sigma_g \to X}$. Since the operators (4.10) can be interpreted as differential forms on this moduli space, the above selection rule just says that we have to integrate top forms.

In the case of a Calabi-Yau manifold of complex dimension 3, we have $c_1(X) = 0$, and the selection rule says that at genus $g = 0$ (*i.e.* when the Riemann surface is a sphere) we have to insert three operators associated to 2-forms. The correlation functions can be evaluated by summing over worldsheet instantons. In the trivial sector (*i.e.* when the image of the sphere is a point in the target), the correlation function is just the classical intersection number $D_1 \cap D_2 \cap D_3$ of the three divisors D_i, $i = 1, 2, 3$, associated to the 2-forms, while the nontrivial instanton sectors give an infinite series. The final answer looks, schematically,

$$\langle \mathcal{O}_{\phi_1} \mathcal{O}_{\phi_2} \mathcal{O}_{\phi_3} \rangle = (D_1 \cap D_2 \cap D_3) + \sum_\beta I_{0,3,\beta}(\phi_1, \phi_2, \phi_3) q^\beta \tag{4.12}$$

The notation is as follows: let $\omega = \sum_{i=1}^{b_2} t_i \omega_i$, be the complexified Kähler form of X, where ω_i is a basis for $H^2(X)$, and t_i are the complexified Kähler parameters. Set $q_i = e^{-t_i}$. If $\beta = \sum_i n_i \omega_i$, then q^β denotes $\prod_i q_i^{n_i}$. The coefficients $I_{0,3,\beta}(\phi_1, \phi_2, \phi_3)$ "count" in some appropriate way the number of holomorphic maps from the sphere to the Calabi-Yau, in such a way that the point of insertion of $\mathcal{O}_{\phi_i}$ gets mapped to the divisor D_i. These is an example of a Gromov-Witten invariant, although to get the general picture we have to couple the model to gravity, as we will see very soon.

When $c_1(X) > 0$, the evaluation of the correlation functions has the structure of (4.12): the trivial sector gives just the classical intersection number of the cohomology ring, and then there are quantum corrections associated to the worldsheet instantons. One important aspect of the case $c_1(X) > 0$ is that, due to the selection rule for the insertions, only a finite number of topological sectors contribute. The correlation function is then the sum of a classical intersection number plus a finite number of "quantum" corrections, and this leads to the definition of the quantum cohomology of X, see [22] for details.

4.2 Closed topological strings

In the above considerations on topological sigma models we have focused on $g = 0$. For $g = 1$ and a Calabi-Yau manifold, the only vev that may lead to a nontrivial answer is that of the unit operator, *i.e.* the partition function itself, while for $g > 1$ the virtual dimension of the moduli space is negative and the above theory is no longer useful to study the target space X. This corresponds mathematically to the fact that, for a generic metric on the Riemann surface Σ_g, there are no holomorphic maps at genus $g > 1$. In order to circumvent this problem, we have to couple the theory to two-dimensional gravity, which means considering all possible metrics on the Riemann surface. The resulting object is called a *topological string*. We will start by giving a general idea from a more mathematical point of view (see [22] for a rigorous discussion), and then we will present the physical construction.

The moduli space of possible metrics (or equivalently, complex structures) on a Riemann surface with punctures is the famous Deligne-Mumford space $\overline{M}_{g,n}$ of n-pointed stable curves (the definition of what stable means can be found for example in [43]). The moduli space we have to consider in the theory of topological strings also involves maps. It consists on one hand of a point in $\overline{M}_{g,n}$, *i.e.* a Riemann surface with n punctures, $(\Sigma_g, p_1, \cdots, p_n)$, and a choice of complex structure on Σ_g. On the other hand, we have a map $f : \Sigma_g \to X$ which is holomorphic with respect to the choice of complex structure on Σ_g.

Let us now fix the topological sector of the holomorphic map, *i.e.* the homology class

$\beta = f_*[\Sigma_g]$. In general, there will be many maps in this sector. The set given by the possible data $(f, \Sigma_g, p_1, \cdots, p_n)$ associated to the class β can be promoted to a moduli space $\overline{M}_{g,n}(X, \beta)$, provided a certain number of conditions are satisfied. This is the basic moduli space we will need in the theory of topological strings. Its virtual dimension is given by:

$$(1-g)(d-3) + n + \int_{\Sigma_g} x^*(c_1(X)). \tag{4.13}$$

The moduli space $\overline{M}_{g,n}(X, \beta)$ comes equipped with the natural maps

$$\begin{aligned} \pi_1 : \overline{M}_{g,n}(X, \beta) &\longrightarrow X^n, \\ \pi_2 : \overline{M}_{g,n}(X, \beta) &\longrightarrow \overline{M}_{g,n}. \end{aligned} \tag{4.14}$$

The first map is easy to define: given a point $(f, \Sigma_g, p_1, \cdots, p_n)$ in $\overline{M}_{g,n}(X, \beta)$, we just compute $(f(p_1), \cdots, f(p_n))$. The second map is essentially sends $(f, \Sigma_g, p_1, \cdots, p_n)$ to $(\Sigma_g, p_1, \cdots, p_n)$, *i.e.* forgets the information about the map and leaves the punctured curve (there are some subtleties with this map, associated to the stability conditions; see [22]). We can now formally define the Gromov-Witten invariant $I_{g,n,\beta}$ as follows. Let us consider cohomology classes $\phi_1, \cdots, \phi_m$ in $H^*(X)$. If we pullback them to $H^*\overline{M}_{g,n}(X, \beta)$ via π_1, we get differential forms on the moduli space of maps that we can integrate (as long as there is a well-defined fundamental class for this space):

$$I_{g,n,\beta}(\phi_1, \cdots, \phi_m) = \int_{\overline{M}_{g,n}(X,\beta)} \pi^*(\phi_1 \wedge \cdots \wedge \phi_m). \tag{4.15}$$

By using the Gysin map $\pi_{2!}$, one can reduce this to an integral over the moduli space of curves $\overline{M}_{g,n}$. The Gromov-Witten invariant $I_{g,n,\beta}(\phi_1, \cdots, \phi_m)$ vanishes unless the degree of the form equals the dimension of the moduli space. Therefore, we have the following constraint:

$$\frac{1}{2}\sum_{i=1}^{n} \deg(\phi_i) = (1-g)(d-3) + n + \int_{\Sigma_g} x^*(c_1(X)) \tag{4.16}$$

Notice that Calabi-Yau threefolds play a special role in the theory, since for those targets the virtual dimension only depends on the number of punctures, and therefore the above condition is always satisfied if the forms ϕ_i have degree 2. These invariants generalize the invariants obtained from topological sigma models. In particular, $I_{0,3,\beta}$ are the invariants involved in the evaluation of correlation functions of the topological sigma model with a Calabi-Yau threefold as its target in (4.12). Notice that, when $n = 0$, one gets an invariant $N_{g,\beta} = I_{g,0,\beta}$ which does not require any insertions. We will refer to this as the Gromov-Witten invariant of the Calabi-Yau X at genus g and in the class β. These are the only (closed) Gromov-Witten invariants that we will deal with here. It can be also shown that, for genus 0,

$$I_{0,3,\beta}(\phi_1, \phi_2, \phi_3) = N_{0,\beta} \int_\beta \phi_1 \int_\beta \phi_2 \int_\beta \phi_3 \tag{4.17}$$

so from these Gromov-Witten invariants one can recover as well the information about the three-point functions of the topological sigma model.

The physical point of view on the Gromov-Witten invariants $N_{g,\beta}$ comes about as follows. It is clear that we have to couple the topological sigma model to two dimensional gravity in order to get nontrivial invariants. To do that, one realizes [28, 91, 14] that the structure of the twisted theory is tantalizingly close to the bosonic string. In the bosonic string, there is a nilpotent BRST operator, $Q_{\rm BRST}$, and the energy-momentum tensor turns out to be a $Q_{\rm BRST}$-commutator: $T(z) = \{Q_{\rm BRST}, b(z)\}$. In addition, there is a ghost number with anomaly $3\chi(\Sigma_g) = 6 - 6g$, in such a way that $Q_{\rm BRST}$ and $b(z)$ have ghost number 1 and -1, respectively. This is precisely the same structure that we found in (4.7), so the field $b_{\alpha\beta}$ plays the role of an antighost. Therefore, one can just follow the prescription of coupling to gravity for the bosonic string and define a genus g free energy as follows:

$$F_g = \int_{\overline{M}_g} \langle \prod_{k=1}^{6g-6} (b, \mu_k) \rangle, \tag{4.18}$$

where

$$(b, \mu_k) = \int d^2 z (b_{zz} (\mu_k)_{\bar z}{}^{z} + b_{\bar z \bar z} (\overline{\mu}_k)_{z}{}^{\bar z}), \tag{4.19}$$

and μ_k are the usual Beltrami differentials. The vev in (4.18) refers to the path integral over the fields of the twisted sigma model. The result, which depends on the choice of complex structure of the Riemann surface, is then integrated over the moduli space $\overline{M}_g$. F_g can be evaluated again, like in the topological sigma model, as a sum over instanton sectors. It turns out [14] that F_g is a generating functional for the Gromov-Witten invariants $N_{g,\beta}$, or more precisely,

$$F_g(t) = \sum_\beta N_{g,\beta} q^\beta. \tag{4.20}$$

It is also useful to introduce a generating functional for the all-genus free energy:

$$F(g_s, t) = \sum_{g=0}^\infty F_g(t) g_s^{2g-2}. \tag{4.21}$$

The parameter g_s can be regarded as a formal variable, but in the context of type II strings it is nothing but the string coupling constant. As a final remark, notice that precisely when the target space is a Calabi-Yau threefold the anomaly is exactly the one of the usual bosonic string. In that sense, one can say that topological strings whose target is a Calabi-Yau threefold are critical.

Gromov-Witten invariants turn out to have an underlying integral structure which appears when the generating functional (4.21) is written in terms of contributions associated to BPS states in compactifications of type IIA superstring theory. As shown in [36], $F(g_s, t)$ can be written as

$$F(g_s,t) = \sum_{g=0}^{\infty}\sum_{\beta}\sum_{d=1}^{\infty} n_\beta^g \frac{1}{d}\Big(2\sin\frac{dg_s}{2}\Big)^{2g-2} q^{d\beta}, \tag{4.22}$$

where n_β^g, known as Gopakumar-Vafa invariants, are *integer* numbers. Therefore, Gromov-Witten invariants of closed strings can be written in terms of these new, integer invariants. In fact, by knowing the Gromov-Witten invariants we can explicitly compute the Gopakumar-Vafa invariants from (4.22). There is also a contribution of constant maps to F_g which is proportional to the Euler characteristic of the manifold and has the form

$$N_{g,0} = \frac{(-1)^g \chi(X) |B_{2g}B_{2g-2}|}{4g(2g-2)(2g-2)!}. \tag{4.23}$$

Exercise. Show that the Gopakumar-Vafa formula (4.22) predicts the following expression for $F_g(q)$:

$$F_g(t) = \sum_\beta \Big(\frac{|B_{2g}|n_\beta^0}{2g(2g-2)!} + \frac{2(-1)^g n_\beta^2}{(2g-2)!} \pm \cdots - \frac{g-2}{12} n_\beta^{g-1} + n_\beta^g\Big)\mathrm{Li}_{3-2g}(q^\beta), \tag{4.24}$$

where B_n denote the Bernoulli numbers and Li_j is the polylogarithm of index j.

4.3 Examples of closed string amplitudes

Gromov-Witten invariants can be computed by using the localization techniques pioneered by Kontsevich [52]. For compact Calabi-Yau manifolds, only $N_{0,\beta}$ can be defined rigorously, since a good definition of the so-called virtual fundamental class of the moduli space hasn't been found yet. For non-compact Calabi-Yau manifolds (the so-called local case) one can compute $N_{g,\beta}$ for arbitrary genus using these localization techniques. We will now give some concrete examples of Calabi-Yau manifolds and we will present results for their topological string amplitudes F_g.

The simplest Calabi-Yau manifold is probably the so-called resolved conifold, which is the total space of the bundle $\mathcal{O}(-1)\oplus\mathcal{O}(-1)\to\mathbb{P}^1$. This manifold is toric, and it has a concrete algebraic description as the zero locus of

$$|z_1|^2 + |z_4|^2 - |z_2|^2 - |z_3|^2 = s \tag{4.25}$$

quotiented by a $U(1)$ that acts as

$$z_1, z_2, z_3, z_4 \to e^{i\alpha} z_1, e^{-i\alpha} z_2, e^{-i\alpha} z_3, e^{i\alpha} z_4 \tag{4.26}$$

This description in terms of a symplectic quotient appears naturally in the linear sigma model of [92]. In that context, the resolved conifold is just the classical Higgs branch of a supersymmetric $U(1)$ gauge theory in two dimensions with four charged matter multiplets. The $U(1)$ symmetry in (4.26) is nothing but the $U(1)$ gauge symmetry, and from there we can read the charges of the different multiplets. Notice that, for $z_2 = z_3 = 0$, (4.25) describes a $\mathbb{P}^1$ whose area is proportional to s. Therefore, (z_1, z_4) can be taken as homogeneous coordinates of the $\mathbb{P}^1$ which is the basis of the fibration, while z_2, z_3 can be regarded as coordinates for the fibers.

The resolved conifold has $b_2(X) = 1$, corresponding to the $\mathbb{P}^1$ in the base, and its total free energy can be computed by various means to be

$$F(g_s, t) = \sum_{d=1}^{\infty} \frac{1}{d\left(2 \sin \frac{d g_s}{2}\right)^2} q^d, \tag{4.27}$$

where $q = e^{-t}$ and t is the complexified area of the $\mathbb{P}^1$. We see that the only nonzero Gopakumar-Vafa invariant is $n_1^0 = 1$. On the other hand, this model already has an infinite number of nontrivial $N_{g,\beta}$ invariants, which can be obtained by expanding the above expression in powers of g_s.

Let us now consider a more complicated example, namely the non-compact Calabi-Yau manifold $\mathcal{O}(-3) \to \mathbb{P}^2$. This is the total space of $\mathbb{P}^2$ together with its anticanonical bundle. We can give again a symplectic description in terms of four complex variables, $z_0, \cdots, z_3$, satisfying the equation

$$|z_1|^2 + |z_2|^2 + |z_3|^2 - 3|z_0|^2 = t \tag{4.28}$$

and quotiented under the $U(1)$ action

$$z_0, z_1, z_2, z_3 \to e^{-3i\alpha} z_0, e^{i\alpha} z_1, e^{i\alpha} z_2, e^{i\alpha} z_3. \tag{4.29}$$

Notice that $z_{1,2,3}$ describe the basis $\mathbb{P}^2$, while z_0 parametrizes the complex direction of the fiber. The space $\mathcal{O}(-3) \to \mathbb{P}^2$ has $b_2 = 1$, corresponding to the hyperplane class of $\mathbb{P}^2$. Therefore, the class β is labeled by a single integer. By using the localization techniques of Kontsevich, adapted to the noncompact case, one finds [20, 51]:

$$\begin{aligned} F_0(q) &= -\frac{t^3}{18} + 3\,q - \frac{45\,q^2}{8} + \frac{244\,q^3}{9} - \frac{12333\,q^4}{64} \cdots \\ F_1(q) &= -\frac{t}{12} + \frac{q}{4} - \frac{3\,q^2}{8} - \frac{23\,q^3}{3} + \frac{3437\,q^4}{16} \cdots \\ F_2(q) &= \frac{\chi(X)}{5720} + \frac{q}{80} + \frac{3\,q^3}{20} + \frac{514\,q^4}{5} \cdots \end{aligned} \tag{4.30}$$

and so on. In (4.30), t is the Kähler class of the manifold, and $\chi(X) = 2$ is the Euler characteristic of the local $\mathbb{P}^2$. The first term in F_0 is the intersection number H^3 of the hyperplane class, divided by 3!, while the first term in F_1 is the intersection number between H and $c_2(X)$. The first term in F_2 is the contribution of constant maps.

As we explained above, we can express the closed string amplitudes in terms of Gopakumar-Vafa invariants. Let us introduce a generating functional for integer invariants as follows:

$$f(z,q) = \sum_{g,\beta} n^g_\beta z^g q^\beta. \tag{4.31}$$

Then, for local $\mathbb{P}^2$ we easily extract from (4.22) the following expression:

$$f(z,q) = 3\,q - 6\,q^2 + (27 - 10\,z)\,q^3 - (192 - 231\,z + 102\,z^2 - 15\,z^3)\,q^4 + \mathcal{O}(q^5). \tag{4.32}$$

It should be mentioned that there is of course a very powerful method to compute F_g, namely mirror symmetry (the B-model). In the B-model, the F_g amplitudes are deeply related to the variation of complex structures on the Calabi-Yau manifold (Kodaira-Spencer theory) and can be computed through the holomorphic anomaly equations of [14]. B-model computations of Gromov-Witten invariants can be found in [14, 20, 51, 48].

4.4 Open topological strings

Let us now consider open topological strings. The natural starting point is a topological sigma model in which the worldsheet is now a Riemann surface $\Sigma_{g,h}$ of genus g with h holes. Such models were considered for the first time in [44] and analyzed in detail in [91]. The main issue is of course to specify boundary conditions for the maps $f : \Sigma_{g,h} \to X$. It turns out that, for the A-model, the relevant boundary conditions are Dirichlet, supported on Lagrangian submanifolds of the Calabi-Yau X. Recall that a Lagrangian submanifold L is a cycle where the Kähler form vanishes:

$$J|_L = 0. \tag{4.33}$$

If we denote by C_i, $i = 1, \cdots, h$ the holes of $\Sigma_{g,h}$ (*i.e.* the disconnected components of the boundary $\partial\Sigma_{g,h}$), we have to pick Lagrangian submanifolds $\mathcal{L}_i$, and consider maps such that

$$f(C_i) \subset \mathcal{L}_i. \tag{4.34}$$

These boundary conditions are a consequence of requiring Q-invariance at the boundary. One also has boundary conditions on the fermionic fields of the theory, which require that χ and ψ at the boundary C_i take values on $x^*(T\mathcal{L}_i)$.

We can also couple the theory to Chan-Paton degrees of freedom on the boundaries, giving rise to a $\otimes_i U(N_i)$ gauge symmetry, where the i-th factor $U(N_i)$ corresponds to the i-th boundary C_i. The model can then be interpreted as a topological open string theory in the presence of N_i topological D-branes wrapping the Lagrangian submanifolds $\mathcal{L}_i$. The Chan-Paton factors give rise to a boundary term in the presence of a gauge connections. If $A = A_i dx^i$ is a $U(N_i)$ connection on L, then the path integral has to be modified by inserting

$$\prod_i \mathrm{Tr}\, \mathrm{P}\, \exp \oint_{C_i} x^*(A), \tag{4.35}$$

where we pullback the connection to C_i through the map x, restricted to the boundary. Notice that, in contrast to physical D-branes in Calabi-Yau manifolds, which wrap special Lagrangian submanifolds [12, 72], in the topological framework the conditions are relaxed to just Lagrangian.

Once boundary conditions have been specified, one can define the free energy of the topological string theory similarly to what we did in the closed case. Let us consider for simplicity the case in which one has a single Lagrangian submanifold $\mathcal{L}$, so that all the boundaries of $\Sigma_{g,h}$ are mapped to $\mathcal{L}$. Now, in order to specify the topological sector of the map, we have to give two different kinds of data: the boundary part and the bulk part. For the bulk part, the topological sector is labeled by relative homology classes, since we are requiring the boundaries of $f_*[\Sigma_{g,h}]$ to end on $\mathcal{L}$. Therefore, we will set

$$f_*[\Sigma_{g,h}] = \mathcal{Q}, \quad \mathcal{Q} \in H_2(X, \mathcal{L}) \tag{4.36}$$

To specify the topological sector of the boundary, we will assume that $b_1(\mathcal{L}) = 1$, so that $H_1(\mathcal{L})$ is generated by a nontrivial one cycle γ. We then have

$$f_*[C_i] = w_i \gamma, \quad w_i \in \mathbf{Z}, \quad i = 1, \cdots, h, \tag{4.37}$$

in other words, w_i is the winding number associated to the map f restricted to C_i. We will collect these integers into a single vector h-uple denoted by $w = (w_1, \cdots, w_h)$.

There are various generating functionals that we can consider, depending on the topological data that we want to keep fixed. It is very useful to fix g, h and the winding numbers, and sum over all bulk classes. This produces the following generating functional of open Gromov-Witten invariants:

$$F_{w,g}(t) = \sum_Q F_{w,g}^Q e^{-Q \cdot t}. \tag{4.38}$$

In this equation, we have written

$$\int_{\mathcal{Q}} \omega = Q \cdot t, \tag{4.39}$$

where $t = (t_1, \cdots, t_{b_2})$ are the complexified Kähler parameters of the Calabi-Yau manifold, and Q is a vector that labels the relative homology class of the embedded Riemann surface. In many examples relevant to knot theory, the entries Q are naturally chosen to be half-integers. Finally, the quantities $F^Q_{w,g}$ are the open string Gromov-Witten invariants, and they "count" in an appropriate sense the number of holomorphically embedded Riemann surfaces of genus g in X with Lagrangian boundary conditions specified by $\mathcal{L}$ with the class represented by $Q, \vec{w}$. These are in general rational numbers.

We can now consider the total free energy, which is the generating functional for all topological sectors:

$$F(V) = \sum_{g=0}^{\infty} \sum_{h=1}^{\infty} \sum_{w_1, \cdots, w_h} \frac{i^h}{h!} g_s^{2g-2+h} F_{g,w}(t) \mathrm{Tr}\, V^{w_1} \cdots \mathrm{Tr}\, V^{w_h}, \tag{4.40}$$

where g_s is the string coupling constant. The factor i^h is very convenient in order to compare to the Chern-Simons free energy, as we will see later. The factor $h!$ is a symmetry factor which takes into account that the holes are indistinguishable (or one could have absorbed them into the definition of $F_{g,w}$).

When all w_i are positive, one can label w in terms of a vector $\vec{k}$. Given an h-uple $w = (w_1, \cdots, w_h)$, we define a vector $\vec{k}$ as follows: the i-th entry of $\vec{k}$ is the number of w_j's which take the value i. For example, if $w_1 = w_2 = 1$ and $w_3 = 2$, then this corresponds to $\vec{k} = (2, 1, 0, \cdots)$. In terms of $\vec{k}$, the number of holes and the total winding number are given by

$$h = |\vec{k}| \equiv \sum_j k_j, \quad \ell = \sum_i w_i = \sum_j j k_j. \tag{4.41}$$

Note that a given $\vec{k}$ will correspond to many w's which differ by permutation of entries. In fact there are $h!/\prod_j k_j!$ h-uples w which give the same vector $\vec{k}$ (and the same amplitude). We can then write the total free energy for positive winding numbers as:

$$F(V) = \sum_{g=0}^{\infty} \sum_{\vec{k}} \frac{i^{|\vec{k}|}}{\prod_j k_j!} g_s^{2g-2+h} F_{g,\vec{k}}(t) \Upsilon_{\vec{k}}(V) \tag{4.42}$$

where

$$\Upsilon_{\vec{k}}(V) = \prod_{j=1}^{\infty} (\mathrm{Tr} V^j)^{k_j} \tag{4.43}$$

Although a rigorous theory of open Gromov-Witten invariants is not available, localization techniques make possible to compute them in various situations [49, 60, 39, 66]. It is also possible to use mirror symmetry to compute disc invariants (*i.e.* when $g = 0$, $h = 1$), as it was first shown in [6] and subsequently explored in [4, 65, 58, 46, 38].

5 Chern-Simons theory as a string theory

5.1 Topological open strings on T^*M

In order to give a string theory interpretation of Chern-Simons theory on $\mathbf{S}^3$, a good starting point is to give an open string interpretation to the $1/N$ expansion of the free energy (3.45). This was done by Witten in a remarkable paper [91], and we will follow his analysis very closely.

Let M be an arbitrary (real) three-dimensional manifold, and consider the six-dimensional space given by M together by its cotangent bundle, T^*M. This space is a symplectic manifold. If we pick local coordinates q_a on M, $a = 1, 2, 3$, and local coordinates for the fiber p_a, the symplectic form can be written as

$$J = \sum_{a=1}^{3} dp_a \wedge dq_a. \tag{5.1}$$

One can find a complex structure on T^*M such that J is a Kähler form, so T^*M can be regarded as a Kähler manifold. Since the curvature of the cotangent bundle cancels exactly the curvature of M, it is Ricci-flat, so it is a Calabi-Yau manifold. In fact $T^*\mathbf{S}^3$ is a well-known Calabi-Yau, namely the *deformed conifold*. The deformed conifold is usually described by the algebraic equation

$$\sum_{\mu=1}^{4} \eta_\mu^2 = a. \tag{5.2}$$

To see that this describes $T^*\mathbf{S}^3$, let us write $\eta_\mu = x_\mu + iv_\mu$, where x_μ, v_μ are real coordinates. we find the two equations

$$\begin{aligned} \sum_{\mu=1}^{4}(x_\mu^2 - v_\mu^2) &= a, \\ \sum_{\mu=1}^{4} x_\mu v_\mu &= 0. \end{aligned} \tag{5.3}$$

The first equation indicates that the locus $v_\mu = 0$, $\mu = 1, \cdots, 4$, describes a sphere $\mathbf{S}^3$ of radius $R^2 = a$, and the second equation shows that the v_μ are coordinates for the cotangent space. Therefore, (5.2) is nothing but $T^*\mathbf{S}^3$.

It is obvious that M is a Lagrangian submanifold in T^*M, since J vanishes along $p_\mu = \text{const}$. Since we have a Calabi-Yau manifold together with a Lagrangian submanifold in it, we can consider a system of N topological D-branes wrapping M, thus providing Dirichlet boundary conditions for topological open strings on T^*M. Our goal now is to obtain a spacetime action describing the dynamics of these topological D-branes.

5.2 Open string field theory

In order to obtain the spacetime description of open strings on T^*M, we have to recall the story for the usual bosonic strings. It was shown in [87] that the spacetime dynamics of open bosonic strings is captured by a cubic string field theory. In string field theory, we consider the worldsheet to be an infinite strip parameterized by a spatial coordinate $0 \leq \sigma \leq \pi$ and a time coordinate $-\infty < \tau < \infty$, with the flat metric $ds^2 = d\sigma^2 + d\tau^2$. We then consider maps $x : I \to T^*M$, with $I = [0, \pi]$. The string field is a functional of open string configurations $\Psi[x(\sigma)]$, of ghost number one. We then define two operations on the space of string functionals. The first one is the integration of string functionals, and it is defined formally by folding the string around its midpoint and gluing the two halves:

$$\int \Psi = \int \mathcal{D}x(\sigma) \prod_{0 \leq \sigma \leq \pi/2} \delta[x(\sigma) - x(\pi - \sigma)]\Psi[x(\sigma)]. \tag{5.4}$$

The integration has ghost number -3, which is the ghost number of the vacuum. This corresponds to the usual fact that in open string theory on the disc one has to soak up three zero modes. One also defines an associative, noncommutative star product $\star$ of string functionals through the following equation:

$$\int \Psi_1 \star \cdots \star \Psi_N = \int \prod_{i=1}^{N} \mathcal{D}x_i(\sigma) \prod_{i=1}^{N} \prod_{0 \leq \sigma \leq \pi/2} \delta[x_i(\sigma) - x_{i+1}(\pi - \sigma)]\Psi_i[x_i(\sigma)], \tag{5.5}$$

where $x_{N+1} \equiv x_1$. The star product simply glues the string together by folding them around their midpoints, and gluing the first half of one with the second half of the following (see for example the recent review [83] for more details), and it doesn't change the ghost number. In terms of these geometric operations, the string field action is given by

$$S = \frac{1}{g_s} \int \Big(\frac{1}{2} \Psi \star Q_{\mathrm{BRST}} \Psi + \frac{1}{3} \Psi \star \Psi \star \Psi \Big). \tag{5.6}$$

Notice that the integrand has ghost number 3, while the integration has ghost number -3, so that the action has zero ghost number. It is interesting to point out that there is another possible normalization which arises when the ghost vacuum of the bc system is assigned the ghost number $-1/2$ (as in [77] and [87]). In that case, the integration operation, the star product and the string field have ghost numbers $-3/2$, $1/2$ and $-1/2$, respectively.

If we add Chan-Paton factors, the string field is promoted to a $U(N)$ matrix of string fields, and the integration includes Tr. This action has all the information about the spacetime dynamics of open bosonic strings, with or without D-branes. In particular, one can derive the Born-Infeld action describing the dynamics of D-branes from the cubic string field theory (see for example [82]).

We will not need all the technology of string field theory in order to understand open topological strings. The only piece of relevant information is the following: the string functional is a function of the zero mode of the string (which corresponds to the position of the string midpoint), and of the higher oscillators. If we decouple all the oscillators, the string functional just becomes an ordinary function of spacetime, the $\star$ product becomes the usual product of functions, and the integral is the usual integration of functions (see [69] for a closely related discussion on this issue). Notice that decoupling the oscillators is in fact the pointlike limit of string theory. As we will see, this is the relevant limit for topological open strings on T^*M.

5.3 Chern-Simons theory as an open string theory

We can now exploit again the analogy between open topological strings and the open bosonic string that we actually used to define the coupling of topological sigma models to gravity (*i.e.*, that both have a nilpotent BRST operator and an energy-momentum tensor that is $Q_{\rm BRST}$-exact). Since both theories have a similar structure, the spacetime dynamics of topological D-branes in T^*M should be governed as well by (5.6), where $Q_{\rm BRST}$ is given in this case by the topological charge defined in (4.2), and where the star product and the integration operation are as in the bosonic string. Sometimes it is useful to consider the ghost number to be shifted by $-d/2$ (where d is the real dimension of the target). Remember that the topological sigma model comes from a twisted $\mathcal{N}=2$ superconformal theory, and the ghost number shifted in this way corresponds to the $U(1)_R$ charge of the superconformal algebra in the Ramond sector. The $U(1)_R$ charge corresponds then to the ghost number in the normalization of [87], in which the string field has ghost number $-1/2$.

The above spacetime description is subtle. Open topological string theory is a theory that describes open string instantons with Lagrangian boundary conditions. On the other hand, it is easy to see that instantons $x : \Sigma \rightarrow T^*M$ such that $x(\partial\Sigma) \subset M$ are necessarily constant. To see this, notice that $J = d\rho$, where

$$\rho = \sum_{a=1}^{3} p_a dq^a, \tag{5.7}$$

and p_a vanishes on M. Since x is a holomorphic map, the instanton action equals the topological piece $-\int_\Sigma x^*(J)$. This can be evaluated to be

$$\int_\Sigma x^*(J) = \int_{\partial\Sigma} x^*(\rho) = 0 \tag{5.8}$$

since $x(\partial\Sigma) \subset M$. Therefore, holomorphic maps with the above boundary conditions are necessarily constant, and there are no "honest" worldsheet instantons! However, as explained by Witten, there are degenerate instantons of zero area in which the Riemann surface degenerates to a graph. It is in fact well-known that the moduli space of open Riemann surfaces contains this type of configurations. This moduli space can be constructed out of strips of variable lengths. It is precisely when these lengths go to infinity that the Riemman surfaces degenerate to become graphs. The string field theory action (5.6) describes then the dynamics of these degenerate instantons when the underlying matter plus ghost system is the one provided by the topological sigma model coupled to gravity.

In order to be more precise about the string field theory description, we have to figure out what is exactly the string field, the $\star$ algebra and so on in the context of topological open strings. To do that, as in the conventional string field theory of the bosonic string, we have to consider the Hamiltonian description of topological open strings. We then take Σ to be an infinite strip parameterized by a spatial coordinate $0 \leq \sigma \leq \pi$ and a time coordinate $-\infty < \tau < \infty$, and consider maps $x : I \to T^*M$, with $I = [0, \pi]$, such that ∂I is mapped to M. The fermion field ψ, being a one-form on Σ, can be split as $\psi = \psi_\sigma d\sigma + \psi_\tau d\tau$, but due to the self-duality condition only one of them, say ψ_τ, is independent. The canonical commutation relations can be read out of the Lagrangian:

$$\begin{aligned} \Big[\frac{dx^i}{d\tau}(\sigma), x^j(\sigma')\Big] &= -\frac{i}{t} G^{ij}\delta(\sigma - \sigma'), \\ \{\psi_\tau(\sigma), \chi(\sigma')\} &= \frac{1}{t}\delta(\sigma - \sigma'). \end{aligned} \tag{5.9}$$

The Hilbert space is made out of functionals $\Phi(x, \cdots)$, where x is a map from the interval as we have just described, and the $\cdots$ refer to the fermion fields. The Hamiltonian is obtained, as usual in string theory, as

$$L_0 = \int_0^\pi d\sigma T_{00}. \tag{5.10}$$

The bosonic piece of T_{00} is just

$$tG_{ij}\Big(\frac{dx^i}{d\sigma}\frac{dx^j}{d\sigma} + \frac{dx^i}{d\tau}\frac{dx^j}{d\tau}\Big), \tag{5.11}$$

and using the canonical commutation relations we find:

$$L_0 = \int_0^\pi d\sigma\Big(-\frac{1}{t}G^{ij}\frac{\delta^2}{\delta^i x(\sigma)\delta^j x(\sigma)} + tG_{ij}\frac{dx^i}{d\sigma}\frac{dx^j}{d\sigma}\Big). \tag{5.12}$$

We then see that string functionals with $dx^i/d\sigma \neq 0$ can not contribute: since the physics is t-independent, we can take $t \to \infty$, where they get infinitely massive and decouple from the spectrum. Therefore, the map $x : I \to T^*M$ has to be constant and in particular it must be a point in M. The conclusion of this analysis is that the relevant string functionals are just functions of the bosonic zero modes. A similar analysis holds for the fermions as well, and we finally obtain that the string functionals are functions of the bosonic and the fermionic zero modes. Denoting them by q^a, χ^a, we see that the string functional reduces to

$$\mathcal{A} = A^{(0)}(q) + \sum_{p=1}^{3} \chi^{\mu_1} \cdots \chi^{\mu_p} A^{(p)}_{\mu_1 \cdots \mu_p}. \tag{5.13}$$

These functionals can be interpreted as differential forms on M, and a differential form of degree p will have ghost number p, or equivalently, $U(1)_R$ charge $p-3/2$. If we have N D-branes wrapping M, the above differential forms take values in the adjoint representation of the gauge group (*i.e.* they are valued in the $U(N)$ Lie algebra). On these functionals, the Q symmetry acts as the exterior differential, and $\{Q, \mathcal{A}\} = 0$ if the differential forms are closed. Of course in string field theory we do not restrict ourselves to functionals in the Q cohomology. Rather we compute the string field action for arbitrary functionals, and then the condition of being in the Q-cohomology arises as a linearized equation of motion. This means that when we consider the string field theory as a quantum theory we will not impose the equations of motion, therefore we do not impose any condition on the functionals (5.13).

We are now ready to write the string field action for topological open strings on T^*M with Lagrangian boundary conditions specified by M. We have seen that the relevant string functionals are of the form (5.13). Since in string field theory the string field has ghost number one (equivalently, $U(1)_R$ charge $-1/2$), we see that

$$\Phi = \chi^\mu A_\mu(q), \tag{5.14}$$

where $A_\mu(q)$ is a Hermitian matrix. In other words, the string field is just a $U(N)$ gauge connection on M. Since the string field only depends on bosonic and fermionic zero modes, the integration of string functionals becomes ordinary integration of forms on M, and the star product becomes the usual wedge products of forms. We then have the following dictionary:

$$\begin{aligned} \Psi \to A, &\qquad Q_{\rm BRST} \to d \\ \star \to \wedge, &\qquad \int \to \int_M . \end{aligned} \tag{5.15}$$

The string field action (5.6) is then the usual Chern-Simons action for A, and by comparing with (2.1) we have the following relation between the string coupling constant and the Chern-Simons coupling:

$$g_s = \frac{2\pi}{k+N}, \tag{5.16}$$

after accounting for the usual shift $k \to k + N$. Notice that, in the open bosonic string, the string field involves an infinite tower of string excitations. For the open topological string, the topological character of the model implies that all excitations decouple, except for the lowest lying one. In other words, the usual reduction to a finite number of degrees of freedom that occurs in topological theories downsizes the string field to a single excitation. In physical terms, what is happening is that string theory reduces in this context to its pointlike limit, since the only relevent degree of freedom of the string is its zero mode, which describes the motion of a pointlike particle. As expected, the dynamics reduces then to a usual quantum field theory. The reasons for this is that, as explained above, the topological nature of the model allows to scale the metric of the target in such a way that α' corrections are supressed. The only possible stringy feature that could remaind would be worldsheet instantons, but as we explained these are also absent for targets of the form T^*M.

Before, we showed that topological open string theory on T^*M has only contributions coming from degenerate instantons. Now we see that these instantons correspond in fact to the fatgraphs of Chern-Simons theory! In particular, to characterize topologically these degenerate instantons we just need their genus g and number of holes h, which are of course the same ones of the associated fatgraph. The perturbative invariant $F_{g,h}$ that we calculated in (3.43) is in fact the free energy of topological open strings in T^*M of genus g and h holes. It would be interesting to derive this directly by using the computational tools for open Gromov-Witten invariants which have been developed in the last two years [49, 39].

5.4 Topological open strings on more general Calabi-Yau manifolds

Above we have presented an explicit description of open topological strings on T^*M, following [91]. What about more general Calabi-Yau manifolds? Consider a Calabi-Yau manifold X together with some Lagrangian submanifolds $M_i \subset X$, with N_i D-branes wrapped over M_i. In this case the topological open strings will have contributions from degenerate holomorphic curves, which are captured by Chern-Simons theories in the way we explained for T^*M, but as pointed out in [91] there will be as well some honest holomorphic curves. An embedded holomorphic curve S will intersect the Lagrangian submanifolds M_i along one-dimensional curves $\mathcal{K}_i(S)$, which can be in fact knots inside M_i. We know from (4.35) that the boundary of such a curve will give a Wilson loop insertion in the spacetime action $\prod_i \mathrm{Tr} U_{\mathcal{K}_i(S)}$, where U is the holonomy of the A connection. In addition, this instanton will be weighted by its area (which corresponds to the closed string background). We can then take into account the contributions of all curves by including the corresponding Chern-Simons theories $S_{\rm CS}(A_i)$, which account for the degenerate curves, coupled in an appropriate way with the honest holomorphic curves. The spacetime action will then have the form

$$S(A_i) = \sum_i S_{\rm CS}(A_i) + \sum_S e^{-\int_S \omega} \prod_i {\rm Tr} U_{\mathcal{K}_i(S)} \tag{5.17}$$

where the second sum is over "honest" holomorphic curves S. Notice that all the Chern-Simons theories $S_{\rm CS}(A_i)$ have the same coupling constant, equal to the string coupling constant. More precisely,

$$\frac{2\pi}{k_i + N_i} = g_s. \tag{5.18}$$

In the action (5.17), the honest holomorphic maps are put "by hand" and in principle one has to solve a nontrivial enumerative problem to find them. Once they are included in the action, the path integral over the Chern-Simons connections will join degenerate instantons to these honest worldsheet instantons: if we have a honest worlsheet instantons ending on a knot $\mathcal{K}$, it will give rise to a Wilson loop operator in (5.17), and the $1/N$ evaluation of the vacuum expectation value will generate all possible fatgraphs Γ joined to the knot $\mathcal{K}$, as it is well-known in Chern-Simons perturbation theory in the presence of Wilson loops. The fatgraphs are interpreted as degenerate instantons. An example of this situation is depicted in Fig. 16. In other words, the action (5.17) will be a sum of contributions coming from partial degenerations of Riemann surfaces, in which a surface $\Sigma_{g,h}$ degenerates to another surface $\Sigma_{g',h'}$ whose boundary ends on a knot $\mathcal{K}$, together with a fatgraph whose external legs end in $\mathcal{K}$ as well. This more complicated scenario was explored in [7, 26, 27, 5], and we will provide examples of (5.17) in section 6.

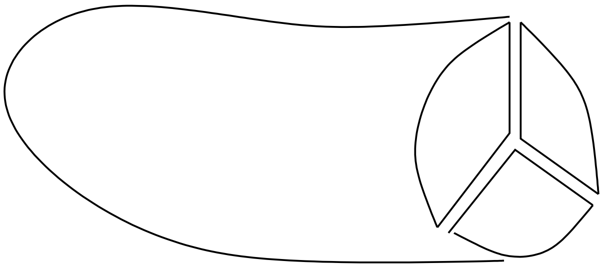


Figure 16: This figure shows a partially degenerated worldsheet instanton of genus $g = 0$ and with $h = 3$ ending on an unknot. The instanton is made out of a honest holomorphic disk and the degenerate piece, which is a fatgraph.

5.5 The conifold transition and the large N duality

We know now, in particular, that Chern-Simons theory on $\mathbf{S}^3$ is a topological open string theory on $T^*\mathbf{S}^3$. The next step is to see if there is a *closed* string theory leading to the resummation (3.46). As shown by Gopakumar and Vafa in another important paper [37], the answer is yes.

One way to motivate their result is as follows: since the holes of the Riemann surfaces are due to the presence of D-branes, "filling the holes" to get the closed strings means getting rid of the D-branes. But this is precisely what happens in another large N duality, namely the AdS/CFT correspondence [8], where type IIB theory in flat space in the presence of D-branes is conjectured to be equivalent to type IIB theory in $\mathrm{AdS}_5 \times \mathbf{S}^5$ with no D-branes, and where the radius of the $\mathbf{S}^5$ is related to the number of D-branes. The reason for that is that, at large N, the presence of the D-branes can be traded by a deformation of the background geometry. In other words, we can make the branes dissapear if we change the background geometry at the same time: as emphasized by Gopakumar and Vafa, large N dualities relating open and closed strings should be associated to transitions in the geometry. This reasoning suggests to look for a transition involving the background $T^*\mathbf{S}^3$. It turns out that such a transition is well-known in the physical and the mathematical literature, and it is called the conifold transition (see for example [18]). Let us explain this in detail.

The algebraic equation describing the deformed conifold is (5.2). It is useful to rewrite this equation as follows. Introduce the following complex coordinates:

$$\begin{aligned} x &= \eta_1 + i\eta_2, \qquad & v &= i(\eta_3 - i\eta_4), \\ u &= i(\eta_3 + i\eta_4), \qquad & y &= \eta_1 - i\eta_2. \end{aligned} \tag{5.19}$$

The deformed conifold can be now written as

$$xy = uv + a \tag{5.20}$$

Notice that in this parametrization the geometry has a $\mathbf{T}^2$ fibration

$$x, y, u, v \to xe^{i\theta_a}, ye^{-i\theta_a}, ue^{i\theta_b}, ve^{-i\theta_b} \tag{5.21}$$

where the θ_a and θ_b actions above can be taken to generate the $(1,0)$ and $(0,1)$ cycles of the $\mathbf{T}^2$. The $\mathbf{T}^2$ fiber can degenerate to $\mathbf{S}^1$ by collapsing one of its one-cycles. In the equation above, for example, the $U(1)_a$ action fixes $x = 0 = y$ and therefore fails to generate a circle there. In the total space, the locus where this happens, i.e. the $x = 0 = y$ subspace of X, is a cylinder $uv = -a$. Similarly, the locus where the other circle collapses, $u = 0 = v$, gives another cylinder $xy = a$. Therefore, we can regard the whole geometry

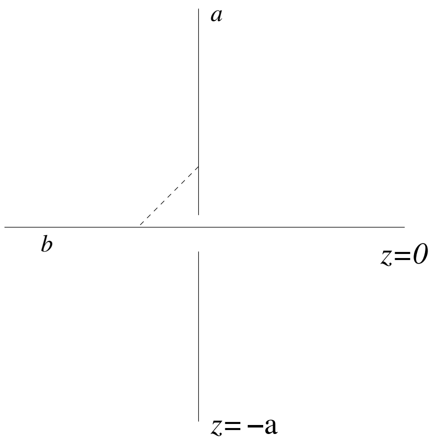


Figure 17: This figure represents $T^*\mathbf{S}^3$, regarded as a $\mathbf{T}^2 \times \mathbb{R}$ fibration of $\mathbb{R}^3$. Two of the directions represent the axes of the two cylinders, and the thir direction represents the real axis of the z-plane.

as a $\mathbf{T}^2 \times \mathbb{R}$ fibration over $\mathbb{R}^3$: if we define $z = uv$, the $\mathbb{R}^3$ of the base is given by $\mathrm{Re}(z)$ and the axes of the two cylinders. The fiber is given by the circles of the two cylinders, and by $\mathrm{Im}(z)$.

It is very useful to represent the above geometry by depicting the singular loci of the torus action in the base $\mathbb{R}^3$. The loci where the cycles of the torus collapse, which are cylinders, project to lines in the base space. Notice that $\mathbf{S}^3$ can be regarded as a torus fibration over an interval, with singular loci at the endpoints. In Fig. 5.5, the three-sphere of the deformed conifold geometry is represented by a dashed line in the z-plane between $z = 0$ and $z = -a$, together with the θ_a and the θ_b circles that degenerate over the endpoints.

The conifold singularity appears when $a = 0$ and the three-sphere collapses. This is described by the equation

$$xy = uv. \tag{5.22}$$

In algebraic geometry, singularities can be avoided in two ways, in general. The first way is to deform the complex geometry, and in our case this leads to the deformed conifold (5.2). The other way is to resolve the singularity, for example by performing a blow up, and this leads to the resolved conifold geometry (see for example [18]). The resolution of the geometry can be explained as follows. When $a = 0$, (5.20) says that $xy = uv$. This can be solved by

$$x = \lambda v, \quad u = \lambda y \tag{5.23}$$

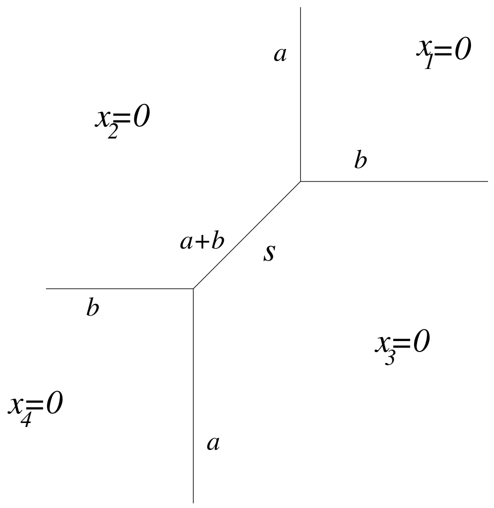


Figure 18: This figure represents the resolved conifold $\mathcal{O}(-1) \oplus \mathcal{O}(-1) \to \mathbb{P}^1$ and the fixed point loci of the $\mathbf{T}^2$ action.

where λ is an inhomogeneous coordinate in $\mathbb{P}^1$. (5.23) can be interpreted as defining the bundle $\mathcal{O}(-1) \oplus \mathcal{O}(-1) \to \mathbb{P}^1$ [18]. To make contact with the toric description given in (4.25), we put $x = x_1 x_3$, $y = x_2 x_4$, $u = x_1 x_2$ and $v = x_3 x_4$. We then see that $\lambda = x_1/x_4$ is the inhomogenous coordinate for the $\mathbb{P}^1$ described in (4.25) by $|x_1|^2 + |x_4|^2 = 0$. We therefore have a *conifold transition* in which the three-sphere of the deformed conifold shrinks to zero size as a goes to zero, and then a two-sphere of size s blows up giving the resolved conifold.

In terms of the coordinates $x_1, \cdots, x_4$, the $\mathbf{T}^2$ action (5.21) becomes

$$x_1, x_2, x_3, x_4 \to e^{i(\theta_a + \theta_b)} x_1, e^{-i\theta_a} x_2, e^{-i\theta_b} x_3, x_4. \tag{5.24}$$

The above structure of $\mathbf{T}^2$ fibration will be crucial in what follows. Notice that the torus action has fixed point loci inherited from the resolved conifold geometry. It is useful to represent the resolved conifold geometry and the fixed point loci by means of the planar graph shown in Fig. 18. Although the full machinery to construct these graphs will be presented in the next section, let us explain in some detail how Fig. 18 is obtained. First of all, notice that a two-sphere can be regarded as an interval fibered by a circle that pinches off at the endpoints, and the diagonal edge in Fig. 18 represents precisely the interval associated to the $\mathbb{P}^1$ of the resolved conifold. The endpoints are $x_1 = 0$ and $x_4 = 0$. The lines coming out of this edge represent the noncompact directions of the resolved conifold. For example, when $x_1 = x_2 = 0$, we have a line $|x_4|^2 = |x_3|^2 + s$ where

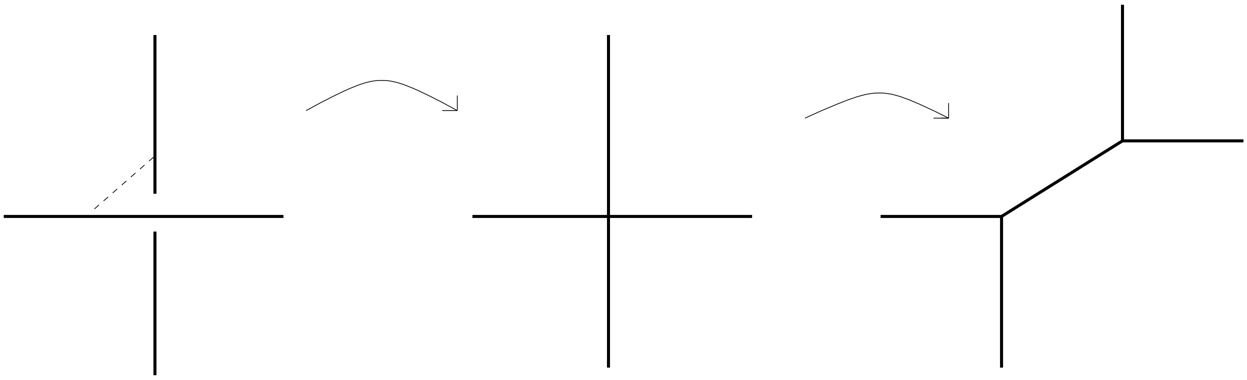


Figure 19: A pictorial representation of the conifold transition. The figure on the left represents the deformed conifold with an $\mathbf{S}^3$, the figure on the center is the conifold singularity, and the figure on the right is the resolved conifold.

the θ_a cycle degenerates. The conifold transition can be represented schematically as in Fig. 19, where the figure on the left represents the deformed conifold with an $\mathbf{S}^3$. The three-sphere collapses to zero size to produce the conifold singularity represented by the figure in the center. Blowing-up the singularity produces the resolved conifold with a $\mathbb{P}^1$, which is the figure on the right.

Now we are ready to state the conjecture of Gopakumar and Vafa in [36]. We know that Chern-Simons theory is an open topological string on the deformed conifold geometry with N topological D-branes wrapping the three-sphere. The conjecture is that at large N the D-branes induce a conifold transition in the background geometry, so that we end up with the resolved conifold and no D-branes. But in the absence of D-branes that enforce boundary conditions we just have a theory of closed topological strings. Therefore, *Chern-Simons theory on* $\mathbf{S}^3$ *is equivalent to closed topological string theory on the resolved conifold.* As we will see, the relation between the parameters is the following: the closed string coupling constant is the open string coupling constant, or equivalently the Chern-Simons effective coupling constant as in (5.16). The size of the $\mathbb{P}^1$ in the resolved geometry is given by the 't Hooft coupling of Chern-Simons theory,

$$t = ig_sN = xN. \tag{5.25}$$

This conjecture has been proved by embedding the duality in type II superstring theory [86] and lifting it to M-theory [1, 10]. A worldsheet derivation of the duality has been presented in [74]. In the next subsection, we will give evidence for the conjecture at the level of the free energy.

5.6 A test of the duality: the partition function

A nontrivial test of the duality advocated by Gopakumar and Vafa is to verify that the free energy of $U(N)$ Chern-Simons theory on the sphere agrees with the free energy of closed topological strings on the resolved conifold. So far, what we have uncovered is the open string expansion of Chern-Simons theory, which is (order by order in x) determined by the perturbative expansion. In order to find a closed string interpretation, we have to sum over the holes, as in (3.45). The 't Hooft parameter t is given by (5.25) and

$$F_g^{\rm p}(t) = \sum_{h=1}^{\infty} F_{g,h}^{\rm p}(-it)^h. \tag{5.26}$$

Let us first focus on $g \geq 2$. To perform the sum explicitly, we write again the ζ function as $\zeta(2g-2+2p) = \sum_{n=1}^{\infty} n^{2-2g-2p}$, and use the binomial series,

$$\frac{1}{(1-z)^q} = \sum_{n=0}^{\infty} \binom{q+n-1}{n} z^n \tag{5.27}$$

to obtain:

$$F_g^{\rm p}(t) = \frac{(-1)^g |B_{2g}B_{2g-2}|}{2g(2g-2)(2g-2)!} + \frac{B_{2g}}{2g(2g-2)} \sum_{n\in\mathbf{Z}}{}' \frac{1}{(-it+2\pi n)^{2g-2}}, \tag{5.28}$$

where $'$ means that we omit $n=0$. Now we notice that, if we write

$$F^{\rm np} = \sum_{g=0}^{\infty} F_g^{\rm np}(t) g_s^{2g-2} \tag{5.29}$$

then for, $g \geq 2$, $F_g^{\rm np}(t) = B_{2g}/(2g(2g-2)(-it)^{2g-2}$, which is precisely the $n=0$ term missing in (5.28). We then define:

$$F_g(t) = F_g^{\rm p}(t) + F_g^{\rm np}(t). \tag{5.30}$$

Finally, since

$$\sum_{n\in\mathbf{Z}} \frac{1}{n+z} = \frac{2\pi i}{1-e^{-2\pi i z}}, \tag{5.31}$$

by taking derivatives w.r.t. z we can write

$$F_g(t) = \frac{(-1)^g |B_{2g}B_{2g-2}|}{2g(2g-2)(2g-2)!} + \frac{|B_{2g}|}{2g(2g-2)!} {\rm Li}_{3-2g}(e^{-t}), \tag{5.32}$$

again for $g \geq 2$. If we now compare to (4.24), we see that This has *precisely* the structure of the free energy of a closed topological string, with $n_1^0 = 1$, and the rest of the Gopakumar-Vafa invariants being zero. Also, from the first term, which gives the contribution of the constant maps, we find that $\chi(X) = 2$. In fact, (5.32) is precisely the F_g amplitude of the resolved conifold. This is a remarkable check of the conjecture.

5.7 Incorporating Wilson loops

How do we incorporate Wilson loops in the picture? As we have discussed above, once one has a closed string description of the $1/N$ expansion, Wilson loops are related to open strings in the closed string geometry. Since the string description involves topological strings, it is natural to assume that Wilson loops are going to be described by open topological strings in the resolved conifold, and this means that we need a Lagrangian submanifold specifying the boundary conditions for the strings.

These issues were addressed in an important paper by Ooguri and Vafa [73]. In order to give boundary conditions for the open strings in the resolved conifold, Ooguri and Vafa constructed a natural Lagrangian submanifold $\widetilde{\mathcal{C}}_{\mathcal{K}}$ in $T^*\mathbf{S}^3$ for any knot $\mathcal{K}$ in $\mathbf{S}^3$. This construction is rather canonical, and it is called the conormal bundle of $\mathcal{K}$. The details are as follows: suppose that knot is parametrized by a curve $q(s)$, where $s \in [0, 2\pi)$, for example. The conormal bundle of $\mathcal{K}$ is the space

$$\widetilde{\mathcal{C}}_{\mathcal{K}} = \Big\{ (q(s), p) \in T^*\mathbf{S}^3 | \sum_i p_i \dot{q}_i = 0, \;\; 0 \le s \le 2\pi \Big\} \tag{5.33}$$

where p_μ are coordinates for the cotangent bundle, and $\dot{q}_i$ denote the derivatives w.r.t. s. This space is an $\mathbb{R}^2$-fibration of the knot itself, where the fiber on the point $q9s)$ is given by the two-dimensional subspace of $T^*_q\mathbf{S}^3$ of planes orthogonal to $\dot{q}(s)$. $\widetilde{\mathcal{C}}_{\mathcal{K}}$ has in fact the topology of $\mathbf{S}^1 \times \mathbb{R}^2$, and intersects $\mathbf{S}^3$ along the knot $\mathcal{K}$.

One can now consider, together with the N branes wrapping $\mathbf{S}^3$, a set of M probe branes wrapping $\widetilde{\mathcal{C}}_{\mathcal{K}}$. As usual when we have two sets of D-branes, we have three different types of strings: the strings with both ends on the N branes are described by $U(N)$ Chern-Simons theory on $\mathbf{S}^3$, as we argued before. In the same way, the strings with both ends on the M branes are described by $U(M)$ Chern-Simons theory on $\widetilde{\mathcal{C}}_{\mathcal{K}}$. But there is a new sector due to strings strechted between the N branes and the M branes. To study these strings, we can follow an analysis similar to the one we did before: we have to impose again that $dx^i/d\sigma = 0$. Therefore, x^i has to be a constant, and since the endpoints of the string lie on different submanifolds, the only possibility is that $x^i \in \mathbf{S}^3 \cup \widetilde{\mathcal{C}}_{\mathcal{K}} = \mathcal{K}$. A similar analysis holds for the fermions, and we then find that the string functionals describing the new sector of strings are a function of a single bosonic zero mode q parametrizing $\mathcal{K}$, and a single fermionic zero mode χ. In other words,

$$\mathcal{A} = \phi(q) + \chi \xi(q). \tag{5.34}$$

where ϕ is a complex scalar field in the bifundamental representation $(N, \overline{M})$, and living in the intersection of the two branes, $\mathcal{K}$. The fact that the scalar is complex is due to the fact that our strings are oriented, and we have to consider both a real scalar in representation

$(N, \overline{M})$ together with another real scalar in representation $(\overline{N}, M)$, which we can put together as a complex scalar in one of the representations. Again, the Q operator is just the exterior differential d on $\mathbf{S}^1$.

As we explained above, the string field is the piece of the above functional with $U(1)_R$ charge $-1/2$. However, now the $U(1)_R$ charge assignment is different from the one in $\mathbf{S}^3$, and it is given for a differential form of degree p by $p - 1/2$. This is because the target is now $\mathbf{S}^1$ with real dimension 1. Therefore, the surviving field is in this case the scalar $\phi(q)$. This is consistent with the fact that, since the spacetime dynamics takes place now on a circle, and since $Q = d$, the kinetic term for the string field action (5.6) is only non-trivial if the string field is a scalar. The full action for $\phi(q)$ is simply $\oint_{\mathcal{K}} \bar{\phi} d\phi$. However, there are also two background gauge fields that interact with the Chan-Paton factors at the endpoints of the strings. These are the $U(N)$ gauge connection A on $\mathbf{S}^3$, and the $U(M)$ gauge connection $\widetilde{A}$ on $\widetilde{\mathcal{C}}_{\mathcal{K}}$. The interaction is as prescribed by (4.35), and the action for the scalar is finally given by

$$\oint_{\mathcal{K}} \mathrm{Tr}\, \bar{\phi}\, (d + A - \widetilde{A})\, \phi. \tag{5.35}$$

Here we regard $\widetilde{A}$ as a source. We can now proceed to integrate out ϕ [73]. In this way we obtain the effective action for the A field

$$S_{\mathrm{CS}}(A) + \sum_{n=1}^{\infty} \frac{1}{n} \mathrm{Tr} U^n \mathrm{Tr} V^{-n} \tag{5.36}$$

where $S_{\mathrm{CS}}(A)$ is of course the Chern-Simons action for A, and U, V are the holonomies of A, $\widetilde{A}$ around the knot $\mathcal{K}$. Therefore, in the presence of the probe branes, the action involves an insertion of the Ooguri-Vafa operator that was introduced in (2.71). Since we are regarding the M branes as a probe, the holonomy V is an arbitrary source. Notice that the extra piece in (5.36) can be interpreted as coming from an annulus of zero length interpolating between the two sets of D-branes, and in fact one can see using the results of [73, 56] that is has the right structure for an open string amplitude. Later on we will find a simple generalization for an annulus of finite lenght.

Let us now follow this system through the geometric transition. The N branes dissapear, and the background geometry becomes the resolved conifold. However, the M probe branes are still there. It is natural to conjecture that they are now wrapping a Lagrangian submanifold $\mathcal{C}_{\mathcal{K}}$ of $\mathcal{O}(-1) \oplus \mathcal{O}(-1) \to \mathbb{P}^1$ that can be obtained from $\widetilde{\mathcal{C}}_{\mathcal{K}}$ through the geometric transition. The final outcome is the existence of a map

$$\{\text{knots in } \mathbf{S}^3\} \to \{\text{Lagrangian submanifolds in}\, \mathcal{O}(-1) \oplus \mathcal{O}(-1) \to \mathbb{P}^1\} \tag{5.37}$$

which sends

$$\mathcal{K} \to \mathcal{C}_{\mathcal{K}}. \tag{5.38}$$

Moreover, one has $b_1(\mathcal{C}_\mathcal{K}) = 1$. This conjecture is cleary well-motivated in the physics. In [73] Ooguri and Vafa constructed $\mathcal{C}_\mathcal{K}$ explicitly when $\mathcal{K}$ is the unknot, and [56] proposed Lagrangian submanifolds for certain algebraic knots and links (including torus knots). Finally, [81] constructs in detail a map from knots to Lagrangian submanifolds in the resolved conifold for a wide class of knots. Later on we will discuss the case of the unknot in some detail.

Now, the resulting Lagrangian submanifold in the resolved geometry gives precisely the open string sector that is needed in order to extend the large N duality to Wilson loops. According to [73], the free energy of open topological strings (4.40) with boundary conditions specified by $\mathcal{C}_\mathcal{K}$ is identical to the free energy of the deformed Chern-Simons theory with action (5.36), which is nothing but (2.74):

$$F_{\rm string}(V) = F_{\rm CS}(V). \tag{5.39}$$

Notice that, since $b_1(\mathcal{C}_\mathcal{K}) = 1$, the topological sectors of maps with positive winding numbers correspond to vectors $\vec{k}$ labeling the connected vevs, and one finds

$$i^{|\vec{k}|}\sum_{g=0}^{\infty} F_{g,\vec{k}}(t) g_s^{2g-2+|\vec{k}|} = -\frac{1}{\prod_j j^{k_j}} W_{\vec{k}}^{(c)}. \tag{5.40}$$

It is further assumed that there is an analytic continuation of $F_{\rm string}(V)$ from negative to positive winding numbers in such a way that the equality (5.39) holds in general.

When $\mathcal{K}$ is the unknot in the three-sphere, the conjecture of Ooguri and Vafa can be tested in full detail. The first thing we need is a construction of the Lagrangian submanifold $\mathcal{C}_\mathcal{K}$ that corresponds to the unknot in $\mathbf{S}^3$. This was done by Ooguri and Vafa in [73]. The construction goes as follows. Let us start with $T^*\mathbf{S}^3$ expressed as (5.2) , and consider the following anti-holomorphic involution on it.

$$\eta_{1,2} = \bar\eta_{1,2}, \quad \eta_{3,4} = -\bar\eta_{3,4}. \tag{5.41}$$

The symplectic form ω changes its sign under the involution, therefore its fixed point set is a Lagrangian submanifold of $T^*\mathbf{S}^3$. If we write $\eta_\mu = x_\mu + ip_\mu$, the invariant locus of the action (5.41) is

$$p_{1,2} = 0, \quad x_{3,4} = 0 \tag{5.42}$$

and intersects the deformed conifold at

$$x_1^2 + x_2^2 = a + p_3^2 + p_4^2. \tag{5.43}$$

Therefore, the fixed point locus intersects $\mathbf{S}^3$ along the equator, which is an unknot described by the equations.

$$x_1^2 + x_2^2 = a, \quad x_3 = x_4 = 0.$$

We conclude that, if we denote by $\mathcal{U}$ the unknot in $\mathbf{S}^3$, the above fixed point locus defined by (5.41) is the Lagrangian submanifold $\widehat{C}_{\mathcal{U}}$. Now we want to construct the Lagrangian submanifold $\mathcal{C}_{\mathcal{U}}$, obtained from $\widehat{C}_{\mathcal{U}}$ after the conifold transition. To do that, we continue to identify it with the invariant locus of the anti-holomorphic involution. We can describe this explicitly by using the coordinates $(x, u, , z)$ or (y, v, z^{-1}) defined in (5.19) and and (5.23). In these coordinates, $\widehat{C}_{\mathcal{U}}$ is characterized by

$$x = \bar{y}, \quad u = \bar{v}, \tag{5.44}$$

and the conifold equation (5.22) restricted on $\mathcal{C}$ becomes

$$x\bar{x} = u\bar{u}. \tag{5.45}$$

The complex coordinate on the base $\mathbb{P}^1$ defined by (5.23) is

$$z = \frac{x}{\bar{u}}, \tag{5.46}$$

but since $|x| = |u|$, z is a phase. We then find that $\mathcal{C}_{\mathcal{U}}$ is a line bundle over the equator $|z| = 1$ of $\mathbb{P}^1$, and the fiber is the subspace of $\mathcal{O}(-1)+\mathcal{O}(-1)$ given by $u = z\bar{u}$ (remember that x, u are complex coordinates for the fibers). In particular, $\mathcal{C}_{\mathcal{U}}$ intersects with the $\mathbb{P}^1$ at the base along $|z| = 1$. The resulting geometry is shown in Fig. 20. The open

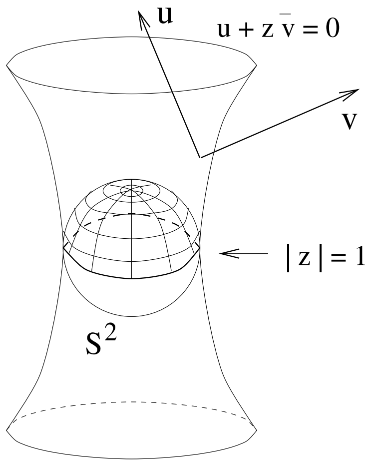


Figure 20: This figure represents the Lagrangian submanifold in $\mathcal{O}(-1) \oplus \mathcal{O}(-1) \rightarrow \mathbb{P}^1$ that corresponds to the unknot in $\mathbf{S}^3$.

Gromov-Witten invariants associated to open string in $\mathcal{O}(-1) \oplus \mathcal{O}(-1)$ whose boundaries end in the above Lagrangian submanifold have been computed in [49] (see also [60, 39]).

6 Topological string amplitudes from Chern-Simons theory

The duality between Chern-Simons on $\mathbf{S}^3$ and closed topological strings on the resolved conifold give a surprising point of view of Chern-Simons invariants of knots and links. However, from the "gravity" point of view we do not learn much about the closed string geometry, since the resolved conifold is quite simple (remember that it only has one nontrivial Gopakumar-Vafa invariant). It would be very interesting to find a topological gauge theory dual to more complicated geometries, in such a way that we could use our knowledge of gauge theory to learn about enumerative invariants of closed strings, and about closed strings in general.

Such a program was started by Aganagic and Vafa in [6]. Their basic idea is to construct geometries that locally contain $T^*\mathbf{S}^3$'s, and then follow geometric transitions to dual geometries where the deformed conifolds are replaced by resolved conifolds. Remarkably, a large class of non-compact toric manifolds can be realized in this way, as it was made clear in [5]. In this section we will present various examples where closed string amplitudes can be computed in that way. As we will see, the method in [5], although very powerful, can be improved, and this will lead us to introduce the trivalent vertex as a building block for string amplitudes [3]. Before all this, however, we have to discuss some issues related to toric geometry.

6.1 Some toric geometry

The geometries that we will consider are noncompact, toric Calabi-Yau manifolds. In order to compute the topological string amplitudes we need some very explicit descriptions of these manifolds as toric fibrations. In the above examples, we have seen that manifolds like the deformed conifold or the resolved conifold have the structure of toric fibrations. Here we will develop this idea systematically.

6.1.1 $\mathbf{C}^3$

Our first example is, in a very concrete sense, the building block of the geometries we are considering: $\mathbf{C}^3$. We want to exhibit its structure as a $\mathbf{T}^2 \times \mathbb{R}$ fibration over $\mathbb{R}^3$. Moreover, we will encode this information in a simple trivalent, planar graph.

Let z_i be complex coordinates on $\mathbf{C}^3$, $i = 1, 2, 3$. We introduce three functions which will be interpreted as Hamiltonians:

$$r_\alpha(z) = |z_1|^2 - |z_3|^2,$$

$$
\begin{aligned}
r_\beta(z) &= |z_2|^2 - |z_3|^2, \\
r_\gamma(z) &= \mathrm{Im}(z_1 z_2 z_3).
\end{aligned} \tag{6.1}
$$

These Hamiltonians generate three flows on $\mathbf{C}^3$ via the standard symplectic form $\omega = \sum_i dz_i \wedge d\overline{z}_i$ on $\mathbf{C}^3$ and the Poisson brackets

$$\partial_\epsilon z_i = \{\epsilon \cdot r, z_i\}_\omega. \tag{6.2}$$

This gives the fibration structure that we were looking for: the base of the fibration, $\mathbb{R}^3$, is parametrized by the Hmailtonians (6.1), while the fiber $\mathbf{T}^2 \times \mathbb{R}$ is parametrized by the flows associated to the Hamiltonians. In particular, the $\mathbf{T}^2$ fiber is generated by the circle actions

$$\exp(i\alpha r_\alpha + i\beta r_\beta): \quad (z_1, z_2, z_3) \;\rightarrow\; (e^{i\alpha} z_1, e^{i\beta} z_2, e^{-i(\alpha+\beta)} z_3), \tag{6.3}$$

while r_γ generates the real line $\mathbb{R}$. We will call the cycle generated by r_α the $(0,1)$ cycle, and the cycle generated by r_β the $(1,0)$ cycle.

Notice that the $(0,1)$ cycle degenerates over the subspace of $\mathbf{C}^3$ described by $z_1 = 0 = z_3$, which is the subspace of the base $\mathbb{R}^3$ given by $r_\alpha = r_\gamma = 0$, $r_\beta \geq 0$. Similarly, over $z_2 = 0 = z_3$ the $(1,0)$-cycle degenerates over the subspace $r_\beta = r_\gamma = 0$ and $r_\alpha \geq 0$. Finally, the one-cycle parameterized by $\alpha+\beta$ degenerates over $z_1 = 0 = z_2$, where $r_\alpha - r_\beta = 0 = r_\gamma$ and $r_\alpha \leq 0$. We will represent the $\mathbf{C}^3$ geometry by a graph which encodes the degeneration

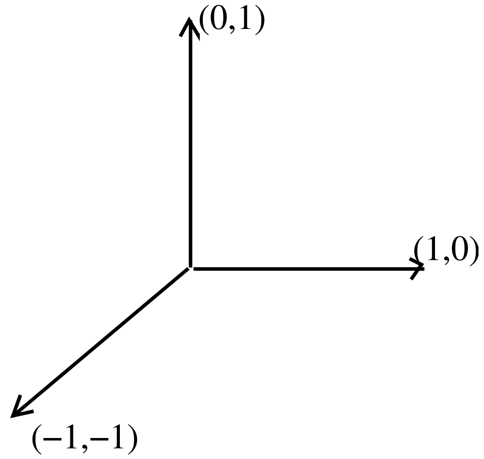


Figure 21: The figure depicts the degeneration locus of the $\mathbf{T}^2 \times \mathbb{R}$ fibration of $\mathbf{C}^3$ in the base $\mathbb{R}^3 = (r_\alpha, r_\beta, r_\gamma)$. This locus is a graph Γ. The labels (p_i, q_i) correspond to the cycles of $\mathbf{T}^2$ which vanish over the corresponding edge.

loci in the $\mathbb{R}^3$ base. In fact, it is useful to have a planar graph by taking $r_\gamma = 0$ and drawing the lines in the $r_\alpha - r_\beta$ plane. The degeneration loci will then be straigth lines described by $pr_\alpha + qr_\beta = \text{const.}$, and over this line we have a $(-q, p)$ cycle of the $\mathbf{T}^2$ that degenerates. Therefore we correlate the degenerating cycles unambiguously with the lines in the graph (up to $(q,p) \to (-q,-p)$). This yields the graph in Fig. 21 drawn in the $r_\gamma = 0$ plane. As we will see, the graphs we will find are planar and trivalent. The graph representing $\mathbf{C}^3$ will be in fact the building block of the geometry, and it corresponds to the trivalent vertex. If we denote the charges associated to the i-th edge by $v_i = (p_i, q_i)$, we will choose the signs in such a way that

$$\sum_i v_i = 0 \tag{6.4}$$

where the sum is over all edges coming out of a given vertex.

In the above discussion we have made a choice for generators of $H_1(T^2)$ to be the one-cycles generated by r_α and r_β. Other choices will differ from this one by an $\mathrm{SL}(2,\mathbf{Z})$ transformation that acts on the $\mathbf{T}^2$, *i.e.* we can have that r_α generates a (p,q) one-cycle and r_β a (t,s) one-cycle such that $ps - qt = 1$. This of course is a symmetry of $\mathbf{C}^3$. However, when gluing different $\mathbf{C}^3$'s together, as we will discuss below, the relative choices will matter and will give rise to different geometries.

6.1.2 The resolved conifold

In order to represent more general geometries, we have to glue together various $\mathbf{C}^3$ patches. The geometries we will be considering are described by $N+3$ coordinates $z_1, \cdots, z_{N+3}$ and N constraints of the form

$$\sum_i Q_i^A |z_i|^2 = t^A \tag{6.5}$$

as in (4.25) and (4.28). In this equation, Q_A, $A = 1, \ldots N$ are integral charges satisfying

$$\sum_i Q_i^A = 0 \tag{6.6}$$

which is the Calabi-Yau constraint. Finally, one divides the space of solutions to (6.6) by $U(1)^N$ action on $z's$ where the A-th $U(1)$ acts on z_i by

$$z_i \to \exp(i\, Q_i^A\, \theta_A) z_i.$$

The N parameters t^A are Kähler moduli of the Calabi-Yau. This mathematical construction arises in the physical context of two-dimensional linear sigma model with $\mathcal{N} = (2,2)$ supersymmetry on the Higgs branch [92]. The theory has $N+3$ chiral fields, whose lowest components are z's, which are charged under N vector multiplets with charges Q_i^A. The equations (6.5) are the D-term equations, and after dividing by the $U(1)^N$ gauge group we obtain the Higgs branch of the theory. When (6.6) holds, the theory flows to a two dimensional conformal sigma model in the infrared.

We will now present a method to construct a graph associated to the Calabi-Yau manifold X described by (6.5), which we will exemplify in the case of the resolved conifold and the local $\mathbb{P}^2$ geometry. First, we must find a decomposition of the set of all coordinates $\{z_i\}_{i=1}^{N+3}$ into triplets $U_a = (z_{i_a}, z_{j_a}, z_{k_a})$ that correspond to the decomposition of X into $\mathbf{C}^3$ patches. We pick one of the patches and we associate to it two Hamiltonians r_α, r_β as we did for $\mathbf{C}^3$ before. These two coordinates will be *global* coordinates in the base $\mathbb{R}^3$, therefore they will generate a globally defined $\mathbf{T}^2$ fiber. The third coordinate in the base is $r_\gamma = \mathrm{Im}(\prod_{i=1}^{N+3} z_i)$, which is manifestly gauge invariant and moreover, patch by patch, can be identified with the coordinate used in the $\mathbf{C}^3$ example above.We can call the cycle generated by r_α the $(1,0)$ cycle, and that generated by r_β the $(0,1)$ cycle. The equation (6.5) can then be used to find the action of $r_{\alpha,\beta}$ on the other patches.

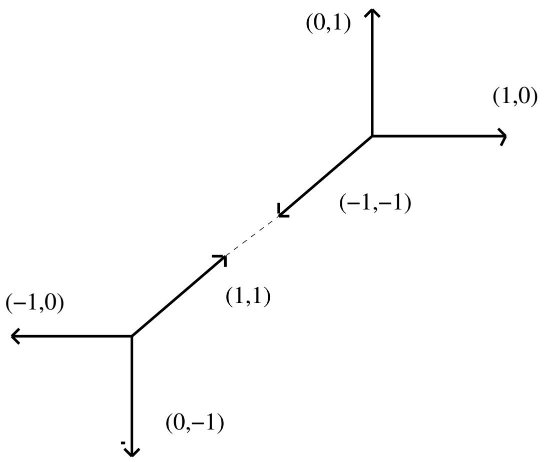


Figure 22: The graph associated to the resolved conifold $\mathcal{O}(-1) \oplus \mathcal{O}(-1) \to \mathbb{P}^1$. The manifold is made out of two patches $\mathbf{C}^3$ glued through a common edge.

Let us illustrate the procedure with the resolved conifold, whose defining equation is (4.25). The first patch will be defined by $z_4 \neq 0$. Using (4.25) we can solve for the modulus of z_4 in terms of the other coordinates, and using the $U(1)$ action we can gauge away its phase. Therefore, the patch will be parameterized by $U_4 = (z_1, z_2, z_3)$. The Hamiltonians will be in this case

$$\begin{aligned} r_\alpha(z) &= |z_2|^2 - |z_1|^2, \\ r_\beta(z) &= |z_3|^2 - |z_1|^2, \end{aligned} \tag{6.7}$$

which generate the actions

$$\exp(i\alpha r_\alpha + i\beta r_\beta): \quad (z_1, z_2, z_3) \;\rightarrow\; (e^{-i(\alpha+\beta)}z_1, e^{i\alpha}z_2, e^{i\beta}z_3). \tag{6.8}$$

This patch will be represented by the same graph that we found for $\mathbf{C}^3$. The other patch will be defined by $z_1 \neq 0$, therefore we can write it as $U_1 = (z_4, z_2, z_3)$. However, in this patch z_1 is no longer a natural coordinate, but we can use (4.26) to rewrite the Hamiltonians as

$$\begin{aligned} r_\alpha(z) &= |z_4|^2 - |z_2|^2 - t, \\ r_\beta(z) &= |z_4|^2 - |z_3|^2 - t, \end{aligned} \tag{6.9}$$

generating the action

$$\exp(i\alpha r_\alpha + i\beta r_\beta): \quad (z_4, z_2, z_3) \;\rightarrow\; (e^{i(\alpha+\beta)}z_4, e^{-i\alpha}z_2, e^{-i\beta}z_3). \tag{6.10}$$

The degeneration loci in this patch are the following: i) $z_4 = 0 = z_2$, corresponding to the line $r_\beta = -t$ where a $(-1, 0)$ cycle degenerates; ii) $z_4 = 0 = z_3$, corresponding to the line $r_\alpha = -t$, and with a $(0, 1)$ cycle degenerates; iii) finally, $z_2 = 0 = z_3$, where $r_\alpha - r_\beta = 0$, and a cycle $(1, 1)$ degenerates. This patch is identical to the first one, and they are joined together through the common edge where $z_2 = 0 = z_3$. The full construction is represented in Fig. 22.

6.1.3 $\mathcal{O}(-3) \to \mathbb{P}^2$

Our last example will be the $\mathcal{O}(-3)$ bundle over $\mathbb{P}^2$. In this case there are four coordinates z_0, z_1, z_2, z_3, and the D-term constraint is (4.28). There are three patches U_i defined by $z_i \neq 0$, for $i = 1, 2, 3$, since at least one of these three coordinates must be non-zero in X. All of these three patches look like $\mathbf{C}^3$. For example, for $z_3 \neq 0$, we can "solve" again for z_3 in terms of the other three unconstrained coordinates which then parameterize $\mathbf{C}^3$: $U_3 = (z_0, z_1, z_2)$. Similar statement holds for the other two patches. Let us now construct the corresponding degeneration graph Γ. In the $U_3 = (z_0, z_1, z_2)$ patch we take as our Hamiltonians

$$\begin{aligned} r_\alpha &= |z_2|^2 - |z_0|^2, \\ r_\beta &= |z_2|^2 - |z_0|^2. \end{aligned} \tag{6.11}$$

The graph of the degenerate fibers in the $r_\alpha - r_\beta$ plane is the same as in the $\mathbf{C}^3$ example, Fig. 21 (the third direction in the base, r_γ is now given by the gauge invariant product $r_\gamma = \mathrm{Im}(z_0 z_1 z_2 z_3)$). The same two Hamiltonians $r_{\alpha,\beta}$ generate the action in the $U_2 =$

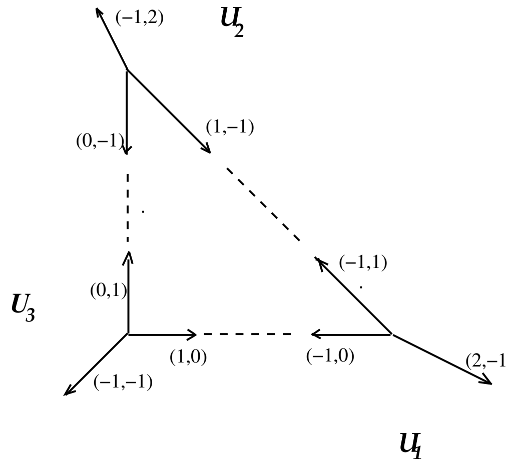


Figure 23: The figure depicts the graph of $\mathcal{O}(-3) \to \mathbb{P}^2$. The manifold is built out of three $\mathbf{C}^3$ patches. The transition functions correspond to $\mathrm{SL}(2,\mathbf{Z})$ transformations of the $\mathbf{T}^2$ fibers as one goes from one patch to the next.

(z_0, z_1, z_3) patch, where we use the (4.28) constraint to rewrite them as follows: since both z_0 and z_1 are coordinates of this patch r_α does not change. On the other hand, r_β must be rewritten since z_2 is not a natural coordinate here. We then find:

$$\begin{aligned} r_\alpha &= |z_2|^2 - |z_0|^2, \\ r_\beta &= t + 2|z_0|^2 - |z_1|^2 - |z_3|^2, \end{aligned} \tag{6.12}$$

and hence

$$\exp\left(i\alpha\, r_\alpha + i\beta r_\beta\right) : (z_0, z_1, z_3) \to (e^{i(-\alpha+2\beta)} z_0, e^{i(\alpha-\beta)} z_1, e^{-i\beta} z_3),$$

We see from the above that the fibers degenerate over three lines: i) $r_\alpha + r_\beta = t$, corresponding to $z_0 = 0 = z_3$, and where a $(-1,1)$ cycle degenerates; ii) there is a line over which a $(-1,2)$ cycle degenerates where $z_1 = 0 = z_3$, $2r_\alpha + r_\beta = t$, and finally, iii) there is a line over which $r_\alpha = 0$, and a $(0,1)$-cycle degenerates. The U_1 patch is similar, and we end up with the graph for $\mathcal{O}(-3) \to \mathbb{P}^2$ shown in Fig. 23. Since at least two of the z's have to be zero for the fibers to degenerate, the graph lies in the $r_\gamma = 0$ plane.

6.2 Geometric transitions for toric manifolds

The geometries we have just discussed are $\mathbf{T}^2 \times \mathbb{R}$ fibrations of $\mathbb{R}^3$ which contain two-spheres (represented by the compact edges of the geometry). We can also construct geometries that contain three-spheres, and as we will see these geometries can be related by a geometric transition to some of the geometries that we just analyzed.

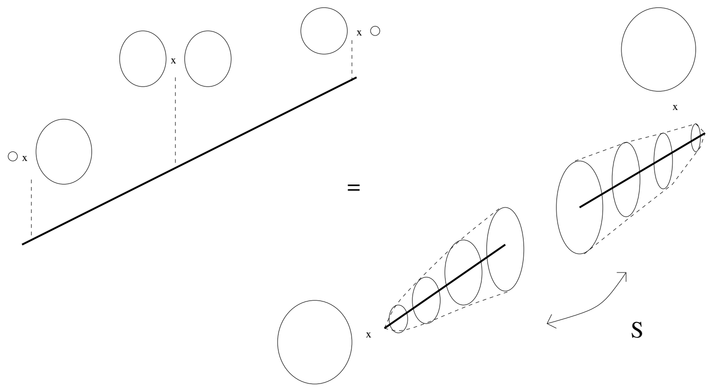


Figure 24: On the left hand side, we represent $\mathbf{S}^3$ as a $\mathbf{T}^2$ fibration over the interval. One of the circles of the torus degenerates over one endpoint, while the other circle degenerates over the opposite endpoint. Each of the degenerating circles fibers over half the interval to produce a disc D, and on the right hand side the three-sphere is equivalently realized as two $\mathbf{S}^1 \times D$ glued through an S transformation.

Recall from our discussion in section 5 that the deformed conifold can be represented by a nonplanar graph where one indicates the degeneration loci of the cycles of the torus fiber, as shown in Fig. 5.5. These loci project onto different values of the variable $\mathrm{Re}(z)$, and the $\mathbf{S}^3$ of the resolved conifold is represented by a dashed line in between this loci. In this way, $\mathbf{S}^3$ is realized as a $\mathbf{T}^2$ fibration over an interval, where at one of the endpoints the $(1,0)$ cycle collapses, while at the other endpoint the collapsing cycle is the $(0,1)$. This description is in fact equivalent to the description of $\mathbf{S}^3$ in terms of a Heegard splitting along solid tori. To see this, let us cut the three-sphere in two pieces by cutting the interval I in two smaller intervals $I_{1,2}$ through its midpoint. Each of the halves is a fibration of $\mathbf{T}^2 = \mathbf{S}^1 \times \mathbf{S}^1_c$ over an interval I_i, where $\mathbf{S}^1_c$ denotes the collapsing cycle. Of course the nontrivial part of the fibration refers to the collapsing cycle, so we can see each of the halves as $\mathbf{S}^1$ times the fibration of the collapsing cycle over I_i, which is nothing but a disk. In other words, we are constructing the three-sphere by gluing two manifolds of the form $\mathbf{S}^1 \times D$. These are of course two solid tori, which are glued after exchanging the two cycles, *i.e.* after performing an S transformation. This is shown in Fig. 24.

We can now construct more general $\mathbf{T}^2 \times \mathbb{R}$ fibrations of $\mathbb{R}^3$ by specifying degeneration loci in the basis through a diagram. A simple example of this is shown in Fig. 25. Notice that this geometry contains two $\mathbf{S}^3$'s, located at the crossing points of the lines. These

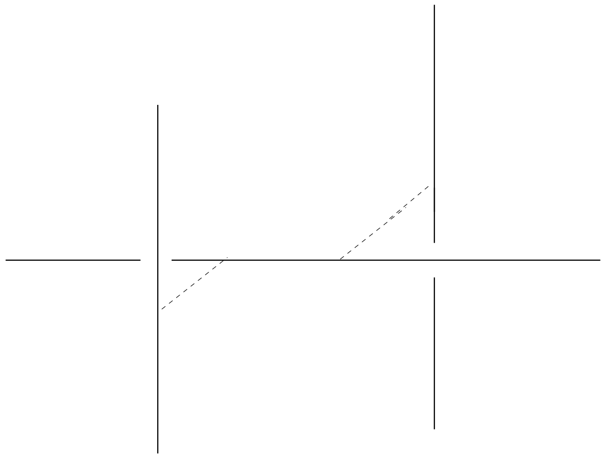


Figure 25: This shows a Calabi-Yau which is a $\mathbf{T}^2 \times \mathbb{R}$ fibration of $\mathbb{R}^3$. The dashed lines represent $\mathbf{S}^3$'s.

three-spheres are also constructed as torus fibrations over the interval, and the cycles that degenerate at the endpoints can be read from the graph. In fact, both are described by a $\mathbf{T}^2$ fibration where the $(0,1)$ cycle collapses at one endpoint, and the $(1,0)$ cycle collapses at the other endpoint. As before, this gives a Heegard splitting of the three-sphere along solid tori that are glued through an S transformation that relates one of the collapsing cycles to the other.

The general situation can be formalized as follows. Given two overlapping lines in a graph like the one depicted in Fig. 25, the dashed line in between them will represent a three-manifold M obtained as a torus fibration over an interval. Let (p_L, q_L) be the cycle of the $\mathbf{T}^2$ fiber that degenerates over the left half on M, and let (p_R, q_R) be the cycle that degenerates over the right half. The gluing matrix V_M can be written as

$$V_M = V_L^{-1} V_R, \tag{6.13}$$

where

$$V_{L,R} = \begin{pmatrix} p_{L,R} & s_{L,R} \\ q_{L,R} & t_{L,R} \end{pmatrix} \in \mathrm{SL}(2, \mathbf{Z}). \tag{6.14}$$

In the case of $M = \mathbf{S}^3$ in the resolved conifold, since $(1,0)$ degenerates in the left half of M and $(0,1)$ in the right half, $V_M = S$, as we explained before. As a small modification, we could make $(p,1)$ degenerate over the left half instead, so that $V_L = T^p S$ is a lens space $\mathbf{S}^3/\mathbf{Z}_p = L(p,1)$ and V is $S^{-1}T^{-p}S$.

Given a graph like Fig. 25, one can try to use the conifold transition "locally," as it was first explained in [6]. The above geometry, for example, contains two deformed conifolds

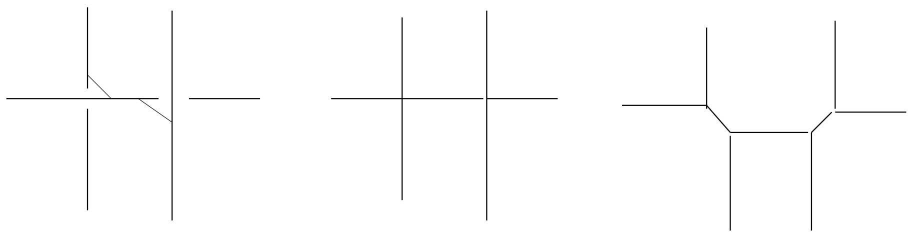


Figure 26: This shows the geometric transition of the Calabi-Yau shown in Fig. 25.In the leftmost geometry there are two minimal 3-cycles. The lengths of the dashed lines are proportional to their sizes. The intermediate geometry is singular, and the figure on the right shows the planar graph associated to the smooth toric Calabi-Yau after the transition.

with their corresponding three-spheres, therefore there is a geometric transition where the three-spheres go to zero size and then the corresponding singularities are blown-up to give a resolved geometry. This geometric transition is depicted in Fig. 26. The resolved geometry is clearly toric, and it has three Kähler classes corresponding to the two blown-up two-spheres, together with an extra sphere associated to the intermediate, horizontal leg.

Although we have focused on the example depicted in Fig. 25, it is clear what is the general philosophy: one considers a "deformed" geometry and performs geometric transitions "locally." The resulting "resolved" geometry will be a toric Calabi-Yau manifold, and if needed one can send some of the Kähler parameters to infinity in order to obtain more simple manifolds. The planar graph describing the resolved geometry can be easily reconstructed from the nonplanar graph describing the deformed geometry. The question now is to know if this procedure allows one to compute the topological string amplitudes on these backgrounds. As we will see in a moment, the answer is yes.

6.3 Closed string amplitudes and geometric transitions

We will now use the generalized geometric transition that we found in the last subsection in order to compute the topological string amplitudes. Let us first wrap N_i branes, $i = 1, 2$, around the two $\mathbf{S}^3$'s of the deformed geometry depicted in Fig. 25. What is the effective topological action describing the resulting open strings? Since this geometry is not globally of the form T^*M, we are in the situation described in 5.4: for open strings with both ends on the same $\mathbf{S}^3$, the dynamics is described by Chern-Simons theory with gauge group $U(N_i)$, therefore we will have two Chern-Simons theories with groups $U(N_1)$

and $U(N_2)$. However, there is a new sector of open strings stretched between the two three-spheres, giving nondegenerate instantons as explained in [91].

Instead of describing these open strings in geometric terms, it is better to use the spacetime physics associated to these strings. In fact, a similar situation was considered when we analyzed the incorporation of Wilson loops in the large N duality. There we had two sets of intersecting D-branes, giving a massless complex scalar field living in the intersection and in the bifundamental representation of the gauge groups. Now, if we focus, say, on the N_1, N_2 branes, we will get again a complex scalar ϕ in the representation $(N_1, \overline{N}_2)$ corresponding to the bifundamental strings stretched between the two sets of D-branes. The difference with the situation that we were considering before is that this complex scalar is now massive, since the strings have a finite length, and its mass is proportional to the "distance" between the two three-spheres. This length is measured by a complexified Kähler parameter that will be denote by r, therefore the kinetic term of the complex scalar will be given by

$$\oint_{\mathbf{S}^1} \bar{\phi}(d - A_1 + A_2 - r)\phi. \tag{6.15}$$

We can now integrate out this complex scalar field to obtain the correction to the Chern-Simons actions on the three-spheres due to the presence of the new sector of open strings, which is given by:

$$\begin{aligned} \mathcal{O}(U_1, U_2; r) &= \exp\Big[-\mathrm{Tr}\ \log(e^{r/2}U_1^{-1/2} \otimes U_2^{1/2} - e^{-r/2}U_1^{1/2} \otimes U_2^{-1/2})\Big] \\ &= \exp\Big\{\sum_{n=1}^{\infty} \frac{e^{-nr}}{n}\ \mathrm{Tr}U_1^n\ \mathrm{Tr}U_2^{-n}\Big\}, \end{aligned} \tag{6.16}$$

where $U_{1,2}$ are the holonomies of the corresponding gauge fields around a loop. The operator $\mathcal{O}$ can also be interpreted as the amplitude for a primitive annulus of size r together with its multicovers, which are labeled by n [73, 56]. This annulus "connects" the two $\mathbf{S}^3$'s, *i.e.* one of its boundaries is in one three-sphere, and the other boundary is in the other sphere. The exponent in (6.16) encodes precisely the contribution of this annulus to the second term in (5.17), and r is the complexified area of the annulus.

The problem now is to determine how many configurations like this one contribute to the full amplitude. It turns out that the only contributions come from open strings stretching along the degeneracy locus, *i.e.* along the edges of the graph that represents the geometry. This was found by Diaconescu, Florea and Grassi [27] using localization arguments, and derived in [5] by exploiting invariance under deformation of complex structures. This result simplifies the problem enormously, and gives a precise description of all the nondegenerate instantons contributing in this geometry: they are annuli

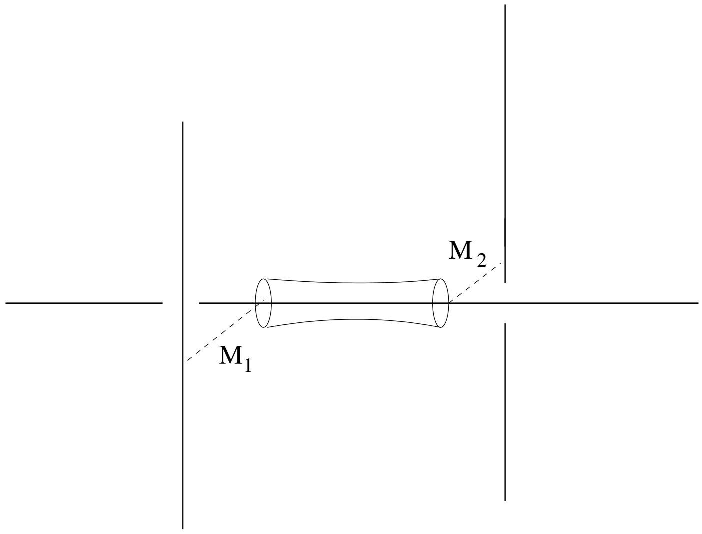


Figure 27: The only nondegenerate instantons contributing to the geometry depicted here come from an annulus stretching along the degeneracy locus.

stretching along the fixed lines of the $\mathbf{T}^2$ action, together with their multicoverings. This is illustrated in Fig. 27. The action describing the dynamics of topological D-branes in the example we are considering above is then:

$$S = S_{\rm CS}(A_1) + S_{\rm CS}(A_2) + \sum_{n=1}^{\infty} \frac{e^{-nr}}{n} \, {\rm Tr} U_1^n \, {\rm Tr} U_2^{-n}, \tag{6.17}$$

where the A_i are $U(N_i)$ gauge connections on $M_i = \mathbf{S}^3$, $i = 1, 2$, and U_i are the corresponding holonomies around loops. There is a very convenient way to write the free energy of the theory with the above action. First notice that, by following the same steps that led to (2.73), one can write the operator (6.16) as

$$\mathcal{O}(U_1, U_2; r) = \sum_R {\rm Tr}_R U_1 e^{-\ell r} {\rm Tr}_R U_2^{-1}, \tag{6.18}$$

where ℓ denotes the number of boxes of the representation R. In the situation depicted in Fig. 27, we see that the boundaries of the annulus gives a knot in M_1, and another knot in M_2. Therefore, the total free energy can be written as:

$$F = F_{\rm CS}(N_1, g_s) + F_{\rm CS}(N_2, g_s) + \log \sum_R e^{-\ell r} W_R(\mathcal{K}_1) W_R(\mathcal{K}_2), \tag{6.19}$$

where $F_{\rm CS}(N_i, g_s)$ denotes the free energy of Chern-Simons theory withgauge group $U(N_i)$. These correspond to the degenerate instantons that come from each of the three-spheres.

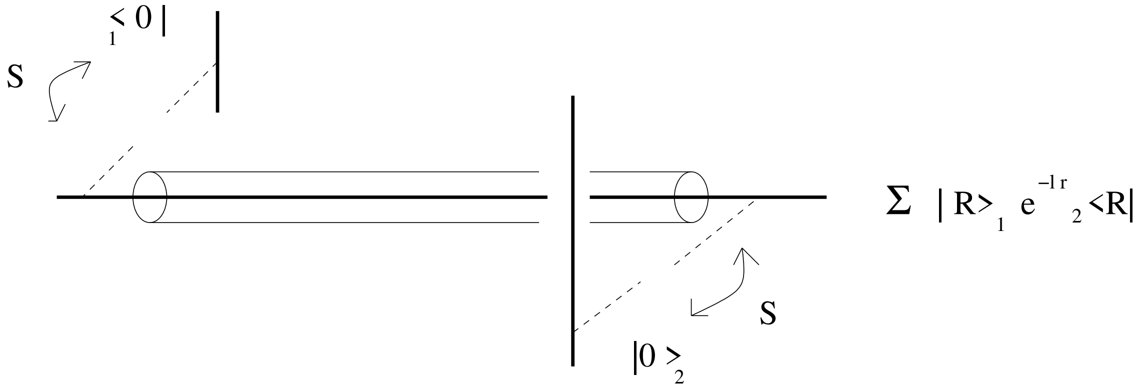


Figure 28: The geometry of Fig. 25 can be chopped into three pieces.

Of course, in order to compute (6.23) we need some extra information: we have to know what are, topologically, the knots $\mathcal{K}_i$, and also if there is some framing induced by the geometry. It turns out that these questions can be easily answered if we evaluate the path integral by chopping the geometry into pieces. The geometry of the knots is then encoded in the geometry of the degeneracy locus.

The evaluation proceeds as follows: we chop the geometry into three pieces, as indicated in Fig. 28, by Heegard splitting the two three-spheres into solid tori. The first piece comes from a solid torus embedded in the total geometry with no insertion, obtained by splitting M_1. This gives the state ${}_1\langle 0|$ in $\mathcal{H}_1^*(\mathbf{T}^2)$, where the subscript 1 refers to the Hilbert space of the $U(N_1)$ Chern-Simons theory on M_1. Similarly, the third piece is another solid torus from the splitting of M_2, and gives the state $|0\rangle_2$. The path integral with the insertion of $\mathcal{O}(U_1, U_2; r)$ produces the following operator in the canonical formalism:

$$\mathcal{O} = \sum_R |R\rangle_1 \, e^{-\ell r} {}_2\langle R| \in \mathcal{H}_1(\mathbf{T}^2) \otimes \mathcal{H}_2^*(\mathbf{T}^2) \tag{6.20}$$

where $|R\rangle$ is the Chern-Simons state that we constructed in section 2, we have introduced subscripts for the labels of the different Hilbert spaces. The gluing is made, as before, through the S transformation in both sides, and the total partition function is then given by:

$$Z(N_1, N_2, g_s, r) = {}_1\langle 0|S\mathcal{O}S|0\rangle_2 = \sum_R {}_1\langle 0|S|R\rangle_1 e^{-\ell r} {}_2\langle {}_2R|S|0\rangle_2. \tag{6.21}$$

What happens now if we go through the geometric transition of Fig. 26? As in the case originally studied by Gopakumar and Vafa, the string coupling constant gives the Chern-Simons "effective" coupling constant $g_s = 2\pi/(k_i + N_i)$ (which is the same for the three theories, see (5.18)), while the 't Hooft parameters $t_i = g_s N_i$ correspond to the

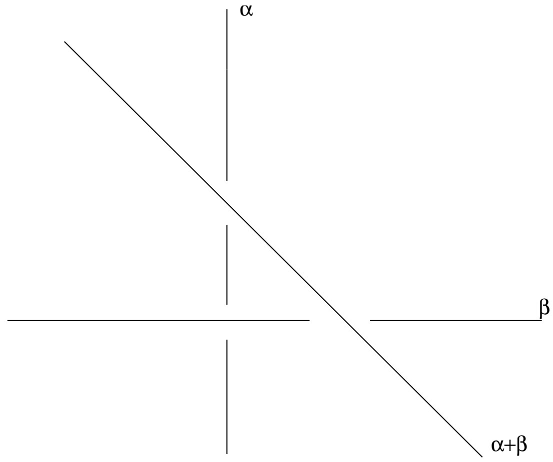


Figure 29: This shows a geometry containing three $\mathbf{S}^3$'s.

sizes of the two outward legs of the toric diagram on the right side of Fig. 26. The free energy (6.19) is, due to the large N transition, the free energy of topological *closed* strings propagating in that toric geometry.

6.4 A geometric transition for local $\mathbb{P}^2$

In the previous subsections we considered a geometric transition involving the geometry depicted in Fig. 25, and we explained how to compute the topological string amplitudes on the closed side by considering the amplitudes of the open side. The resulting geometry was still relatively simple, in the sense that it only had a finite number of nontrivial Gopakumar-Vafa invariants. We will now analyzed a more complicated example: the local $\mathbb{P}^2$ geometry. The starting point is the "deformed" geometry depicted in Fig. 29. Again, we can implement the conifold transition in a local way, as shown in Fig. 30, to obtain the resolved geometry depicted there. This resolved geometry is clearly toric. We see that it contains a $\mathcal{O}(-3) \to \mathbb{P}^2$ together with three extra compact edges that correspond to the $\mathbb{P}^1$'s that appear after the geometric transition. In fact, this geometry is essentially the one obtained by performing three blowups of $\mathcal{O}(-3) \to \mathbb{P}^2$, *i.e.* the noncompact Calabi-Yau manifold given by the del Pezzo surface $\mathbb{B}_3$ together with its canonical bundle. In any case, it is easy to see that to recover the local $\mathbb{P}^2$ geometry one just has to send the sizes of the three $\mathbb{P}^1$'s to infinity.

The procedure to compute the open string amplitudes is very similar to what we found in the last subsection: we wrap N_i branes around the three different $\mathbf{S}^3$'s. There will be now three different sets of open strings stretching along the degeneration loci, and

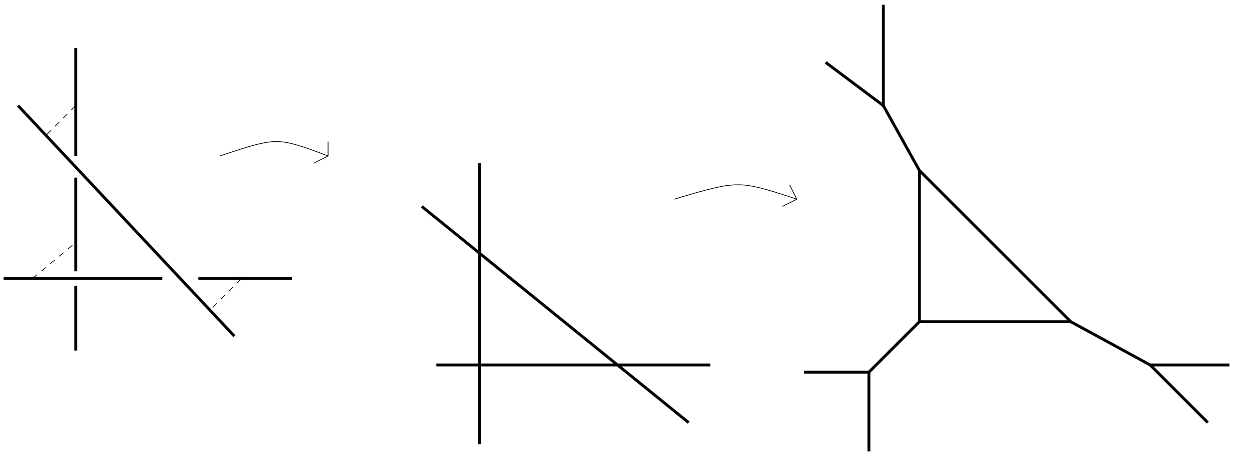


Figure 30: This shows the geometric transition of the Calabi-Yau shown in Fig. 29. In the leftmost geometry there are three minimal 3-cycles. The lengths of the dashed lines are proportional to their sizes. The intermediate geometry is singular, and the figure on the right shows the planar graph associated to the smooth toric Calabi-Yau after the transition.

connecting the spheres. The resulting annuli are depicted in Fig. 31. to annuli The action describing the dynamics of topological D-branes in the example we are considering above is then:

$$S = \sum_{i=1}^{3} S_{\rm CS}(A_i) + \sum_{n=1}^{\infty} \frac{1}{n}\Big(e^{-nr_1}\, {\rm Tr} U_1^n\, {\rm Tr} U_2^{-n} + e^{-nr_2}\, {\rm Tr} U_2^n\, {\rm Tr} U_3^{-n} + e^{-nr_3}\, {\rm Tr} U_3^n\, {\rm Tr} U_1^{-n} \Big), \tag{6.22}$$

where the A_i are $U(N_i)$ gauge connections on each of the $\mathbf{S}^3$'s, $i = 1, 2, 3$, and U_i are the corresponding holonomies around loops. Similarly to (6.19), the free energy can then be written as given by:

$$F = \sum_{i=1}^{3} F_{\rm CS}(N_i, g_s) + \log\Big\{ \sum_{R_1,R_2,R_3} e^{-\sum_{i=1}^{3} \ell_i r_i} W_{R_1,R_2}(\mathcal{L}_1) W_{R_2,R_3}(\mathcal{L}_2) W_{R_3,R_1}(\mathcal{L}_3) \Big\}, \tag{6.23}$$

where ℓ_i is the number of boxes in the representation R_i, and $F_{\rm CS}(N_i, g_s)$ denotes the free energy of Chern-Simons theory with gauge group $U(N_i)$. These correspond to the degenerate instantons that come from each of the three-spheres. We need again some extra information in order to compute the free energy: we have to know what are, topologically, the links $\mathcal{L}_i$, and also if there is some framing induced by the geometry, and to answer these questions we follow the procedure that we used before. We will denote the $\mathbf{S}^3$'s

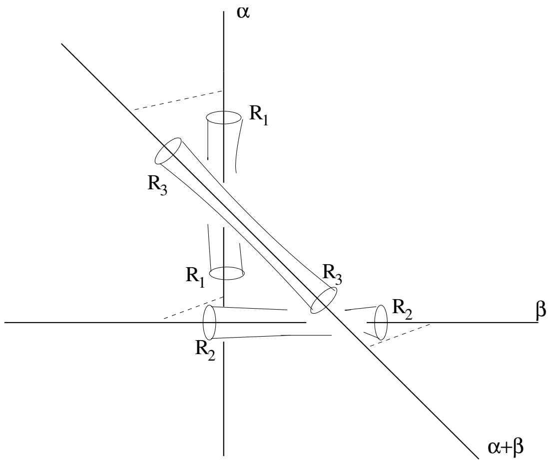


Figure 31: This shows the annuli stretched along the degeneration locus for the Calabi-Yau shown in Fig. 29.

in Fig. 29 by M_i, where M_i is the $\mathbf{S}^3$ where the annuli with representations R_i, R_{i+1} end (and we take $R_4 \equiv R_1$). These thee-manifolds are all $\mathbf{S}^3$'s constructed by gluing two solid tori through an $\mathrm{SL}(2, \mathbf{Z})$ transformation, and the gluing matrix associated to M_i will be denoted by V_{M_i}. Again, the effect of the annulus operator is creating Wilson lines in representation R in the corresponding solid tori, and the partition function can be written as

$$Z = \sum_{R_1, R_2, R_3} \langle R_1 | V_{M_3} | R_3 \rangle e^{-\ell_3 r_3} \langle R_3 | V_{M_2} | R_2 \rangle e^{-\ell_2 r_2} \langle R_2 | V_{M_1} | R_1 \rangle e^{-\ell_1 r_1}. \tag{6.24}$$

One important constraint on the operators, which follows from the geometry, is that when going around the loop from one degeneration locus to the other and then back to the first one we must end up with the identity operator, in other words

$$V_{M_3} V_{M_2} V_{M_1} = 1.$$

Looking at Fig. 31 we can read off,

$$V_{M_1} = S^{-1}, \quad V_{M_2} = S T^{-1} S, \quad V_{M_3} = T S^{-1}, \tag{6.25}$$

since for example, the three-sphere between the α and the β degenerating loci in Fig. 31 comes from gluing two tori with an S^{-1} transformation, which maps the α cycle into the β cycle. We finally find

$$W_{R_1 R_2}(\mathcal{L}_1) \quad = \quad \frac{\langle R_2 | S^{-1} | R_1 \rangle}{Z_{M_1}} = \mathcal{W}_{R_1 R_2},$$

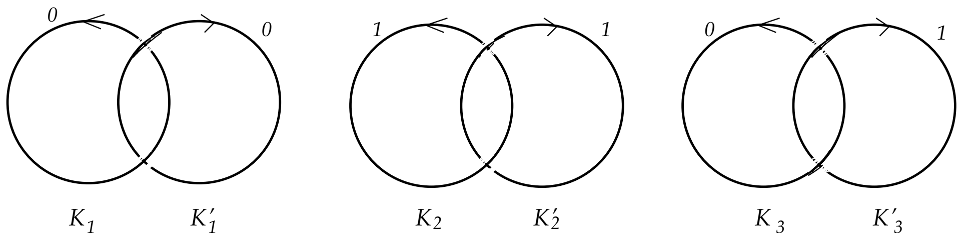


Figure 32: The figure shows the Hopf links $\mathcal{L}_i$, $i = 1, 2, 3$. The numbers indicate the framing of each knot.

$$\begin{aligned} W_{R_2R_3}(\mathcal{L}_2) &= \frac{\langle R_3|TS^{-1}T|R_2\rangle}{Z_{M_2}} = q^{\kappa_{R_2}+\kappa_{R_3}}\mathcal{W}_{R_2R_3}, \\ W_{R_3R_1}(\mathcal{L}_3) &= \frac{\langle R_1|TS^{-1}|R_3\rangle}{Z_{M_3}} = q^{\kappa_{R_1}}\mathcal{W}_{R_3R_1}, \end{aligned} \tag{6.26}$$

where $\mathcal{W}_{R_1R_2}$ is the normalized invariant for the Hopf link with linking number $+1$ given in (2.38), and the T operators give a framing factor. The piece of the exponent of T which is proportional to $N\ell$ can be absorbed in a redefinition of the r_i. We see that the $\mathcal{L}_i$ are all Hopf links, and some of the components do actually have nontrivial framing. If we denote the components of $\mathcal{L}_i$ by $\mathcal{K}_i$ and $\mathcal{K}_i'$, $i = 1, 2, 3$, the framings turn out to be the following: $\mathcal{K}_1$, $\mathcal{K}_1'$ and $\mathcal{K}_3$ have framing zero, while the remaining knots have framing $p = 1$. This means that $\mathcal{L}_1$ is in the canonical framing, in $\mathcal{L}_2$ both components are framed, while in $\mathcal{L}_3$ only one of the components, $\mathcal{K}_3'$, is framed. This is depicted in Fig. 32.

The resulting free energy is the total free energy of the toric geometry depicted in the right hand side of Fig. 30, and the ‘ t Hooft parameters of the three Chern-Simons theories become the Kähler parameters corresponding to the three $\mathbb{P}^1$'s associated to the outward legs of Fig. 30. In order to recover just local $\mathbb{P}^2$, we have to take these 't Hooft parameters to infinity, and "tune" the sizes of the annuli at the same time. To see this, consider the expansion of the second term in (6.23) in terms of the number of boxes of the representations. For one box in total there are three possible terms ($(R_1, R_2, R_3) = \square\cdot\cdot$ and two permutations), and we find:

$$e^{-r_1}W_{\square}(\mathcal{K}_1)W_{\square}(\mathcal{K}_3') + e^{-r_2}W_{\square}(\mathcal{K}_2)W_{\square}(\mathcal{K}_1') + e^{-r_3}W_{\square}(\mathcal{K}_3)W_{\square}(\mathcal{K}_2'). \tag{6.27}$$

The knots involved in this computation are just framed unknots. Since the invariant of an unknot with framing p in the fundamental representation is

$$(-1)^p\frac{\lambda^{-\frac{1}{2}} - \lambda^{\frac{1}{2}}}{q^{-\frac{1}{2}} - q^{\frac{1}{2}}} \tag{6.28}$$

we see that in order to have a finite limit as $t_i \to \infty$, we must have:

$$r = r_1 - \frac{t_1+t_3}{2} = r_2 - \frac{t_1+t_2}{2} = r_3 - \frac{t_2+t_3}{2} \tag{6.29}$$

so we take the limit $t_i \to \infty$ and at the same time $r_i \to \infty$ in such a way that the above combinations remain finite and equal to the closed string Kähler parameter r of local $\mathbb{P}^2$. If we define

$$W_{R_1R_2}(q) = \lim_{N\to\infty} \lambda^{-\frac{\ell_1+\ell_2}{2}} \mathcal{W}_{R_1R_2}(q,\lambda) \tag{6.30}$$

where $\ell_{1,2}$ denotes the number of boxes in R_i, and the limit $N \to \infty$ means that we keep g_s fixed, and take $t \to \infty$, hence $\lambda \to \infty$. We see that in this limit the second term in (6.23) becomes

$$Z_{\mathbf{P}^2} = \sum_{R_i} q^{\sum_i \kappa_{R_i}} (-1)^{\sum_i \ell_i} W_{R_1R_2} W_{R_2R_3} W_{R_3R_1} e^{-\sum_i \ell_i r}. \tag{6.31}$$

Of course, in order for this procedure to work we need (6.30) to be finite and nontrivial. Using the explicit expressions (2.43) and (2.66) one can check that this is indeed the case. Since $W_{R_1R_2}(q)$ depends only on q, we see that in the limit $t_i \to \infty$ the free energy has the structure:

$$F_{\mathbf{P}^2} = \log\Big\{1 + \sum_{\ell=1}^{\infty} a_\ell(q) e^{-\ell r}\Big\} = \sum_{\ell=1}^{\infty} a_\ell^{(c)}(q) e^{-\ell r}. \tag{6.32}$$

The coefficients $a_\ell(q)$, $a_\ell^{(c)}(q)$ can be easily obtained in terms of the invariants of the Hopf link in arbitrary representations. One finds, for example [5],

$$\begin{aligned} a_1(q) &= -\frac{3}{(q^{\frac{1}{2}} - q^{-\frac{1}{2}})^2}, \\ a_2^{(c)}(q) &= \frac{6}{(q^{\frac{1}{2}} - q^{-\frac{1}{2}})^2} + \frac{1}{2} a_1(q^2). \end{aligned} \tag{6.33}$$

If we compare to (4.22) and take into account the effects of multicovering, we find the following values for the Gopakumar-Vafa invariants of $\mathcal{O}(-3) \to \mathbb{P}^2$:

$$\begin{aligned} n_1^0 &= 3, \qquad & n_1^g &= 0 \ \text{ for } g > 0, \\ n_2^0 &= -6, \qquad & n_2^g &= 0 \ \text{ for } g > 0, \end{aligned} \tag{6.34}$$

in agreement with the results listed in (4.32). In fact, one can go much further with this method and compute the Gopakumar-Vafa invariants to high degree. The advantage of this procedure is that, in contrast to both the A and the B model computations, one gets the answer for *all genera*, see [5] for a complete listing of the invariants up to degree 12.

Exercise. Compute (6.32) up to $\ell = 2$ and verify (6.33), (6.34).

6.5 Application: the Seiberg-Witten prepotential

Another application of the above formalism is to the noncompact Calabi-Yau manifold given by $\mathbb{P}^1 \times \mathbb{P}^1$ together with its anticanonical bundle. The symplectic quotient description of this manifold is

6.6 The trivalent vertex

7 Wilson loops, matrix models, and strings

7.1 Matrix models and knot invariants

7.2 Matrix model technology III

7.3 Some computations

In the case of the unknot at the canonical framing, Ooguri and Vafa showed [73] that the generating function (2.74) can be explicitly computed to all orders. The reason is that the quantum dimension in the representation R can be regarded as the trace in the representation R of an $N \times N$ diagonal matrix U_0 whose i-th diagonal entry is

$$\exp\Big[-\frac{\pi i}{k+N}(N-2i-1)\Big]. \tag{7.1}$$

This is easily seen by remembering that ρ lives in the dual of the Cartan algebra H, and by using the natural isomorphism between H and H^* induced by the Killing form we obtain the above result from (2.31). Notice that U_0 is like a "master field" that gives the right answer by evaluating a "classical" trace. Therefore, one can compute $F(V)$ by substituting $\mathrm{Tr}U_0^n$ in (2.71), and one obtains

$$F(V) = -i\sum_{d=1}^{\infty}\frac{e^{dt/2}-e^{-dt/2}}{2d\sin\Big(\frac{dg_s}{2}\Big)}\mathrm{Tr}V^d. \tag{7.2}$$

What happens for $p \neq 0$? In that case, it is no longer possible to sum up all the correlation functions, but we can still compute the connected vevs $W^{(c)}_{\vec{k}}$ at arbitrary framing. To do that, remember that the W_R for the unknot in the canonical framing are just the quantum dimensions of the representation R given in (2.35). We have to correct them with the framing factor as prescribed in (2.49), compute the $W_{\vec{k}}$ with Frobenius formula, and then extract the connected piece

Acknowledgments

I would like to thank Mina Aganagic, Albrecht Klemm, Jose Labastida and Cumrun Vafa for enjoyable collaboration on the topics discussed in these lectures.

References

[1] B. S. Acharya, "On realising $\mathcal{N} = 1$ super Yang-Mills in M theory," hep-th/0011089.

[2] M. Aganagic, A. Klemm, M. Mariño and C. Vafa, "Matrix model as a mirror of Chern-Simons theory," arXiv:hep-th/0211098.

[3] M. Aganagic, A. Klemm, M. Mariño and C. Vafa, "The topological vertex," arXiv:hep-th/0305132.

[4] M. Aganagic, A. Klemm and C. Vafa, "Disk instantons, mirror symmetry and the duality web," hep-th/0105045, Z. Naturforsch. **A 57**, 1 (2002)

[5] M. Aganagic, M. Mariño and C. Vafa, "All loop topological string amplitudes from Chern-Simons theory," hep-th/0206164.

[6] M. Aganagic and C. Vafa, "Mirror symmetry, D-branes and counting holomorphic discs," hep-th/0012041.

[7] M. Aganagic and C. Vafa, "G_2 manifolds, mirror symmetry and geometric engineering," hep-th/0110171.

[8] O. Aharony, S. S. Gubser, J. M. Maldacena, H. Ooguri and Y. Oz, "Large N field theories, string theory and gravity," hep-th/9905111, Phys. Rept. **323** (2000) 183.

[9] M. Atiyah, "On framings of three-manifolds," Topology **29** (1990) 1.

[10] M. Atiyah, J. Maldacena and C. Vafa, "An M-theory flop as a large N duality," hep-th/0011256, J. Math. Phys. **42** (2001) 3209.

[11] S. Axelrod and I. Singer, "Chern-Simons perturbation theory," hep-th/9110056, in *Differential geometric methods in theoretical physics*, p. 3, World Scientific, 1991.

[12] K. Becker, M. Becker and A. Strominger, "Fivebranes, membranes, and nonperturbative string theory," hep-th/9507158, Nucl. Phys. **B 456** (1995) 130.

[13] M. Bershadsky, S. Cecotti, H. Ooguri and C. Vafa, "Holomorphic anomalies in topological field theories," Nucl. Phys. **B 405** (1993) 279, hep-th/9302103.

[14] 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** (1994) 311, hep-th/9309140.

[15] D. Bessis, C. Itzykson and J. B. Zuber, "Quantum Field Theory Techniques In Graphical Enumeration," Adv. Appl. Math. **1**, 109 (1980).

[16] E. Brézin, C. Itzykson, G. Parisi and J. B. Zuber, "Planar Diagrams," Commun. Math. Phys. **59**, 35 (1978).

[17] E. Brézin and V. A. Kazakov, "Exactly Solvable Field Theories Of Closed Strings," Phys. Lett. B **236** (1990) 144. M. R. Douglas and S. H. Shenker, Nucl. Phys. B **335**, 635 (1990). D. J. Gross and A. A. Migdal, "Nonperturbative Two-Dimensional Quantum Gravity," Phys. Rev. Lett. **64**, 127 (1990).

[18] P. Candelas and X. C. De La Ossa, "Comments on conifolds," Nucl. Phys. **B 342** (1990) 246.

[19] P. Candelas, X. C. De La Ossa, P. S. Green and L. Parkes, "A pair of Calabi-Yau manifolds as an exactly soluble superconformal theory," Nucl. Phys. **B 359** (1991) 21.

[20] T. M. Chiang, A. Klemm, S. T. Yau and E. Zaslow, "Local mirror symmetry: Calculations and interpretations," hep-th/9903053, Adv. Theor. Math. Phys. **3** (1999) 495.

[21] S. Coleman, *Aspects of symmetry*, Cambridge University Press.

[22] D. Cox and S. Katz, *Mirror symmetry and algebraic geometry,* AMS 1999.

[23] P. Di Francesco, "Matrix model combinatorics: Applications to folding and coloring," arXiv:math-ph/9911002.

[24] P. Di Francesco, P. Ginsparg and J. Zinn-Justin, "2-D Gravity and random matrices," Phys. Rept. **254**, 1 (1995) [arXiv:hep-th/9306153].

[25] P. Di Francesco, P. Mathieu and D. Sénéchal, *Conformal field theory*, Springer-Verlag, 1997.

[26] D. E. Diaconescu, B. Florea and A. Grassi, "Geometric transitions and open string instantons," hep-th/0205234.

[27] D. E. Diaconescu, B. Florea and A. Grassi, "Geometric transitions, del Pezzo surfaces and open string instantons," hep-th/0206163.

[28] R. Dijkgraaf, E. Verlinde and H. Verlinde, "Notes on topological string theory and two-dimensional topological gravity," in Trieste Spring School 1990.

[29] M. R. Douglas, "Chern-Simons-Witten theory as a topological Fermi liquid," arXiv:hep-th/9403119.

[30] T. Eguchi and S. K. Yang, "The Topological $\mathbb{P}^1$ model and the large N matrix integral," Mod. Phys. Lett. A **9**, 2893 (1994) [arXiv:hep-th/9407134]. T. Eguchi, K. Hori and S. K. Yang, "Topological sigma models and large N matrix integral," Int. J. Mod. Phys. A **10**, 4203 (1995) [arXiv:hep-th/9503017].

[31] S. Elitzur, G. Moore, A. Schwimmer and N. Seiberg, "Remarks on the canonical quantization of the Chern-Simons-Witten theory," Nucl. Phys. **B 326** (1989) 108.

[32] B. Eynard and J. Zinn-Justin, "The $O(n)$ model on a random surface: Critical points and large order behavior," Nucl. Phys. B **386**, 558 (1992) [arXiv:hep-th/9204082]. B. Eynard and C. Kristjansen, "Exact solution of the $O(n)$ model on a random lattice," Nucl. Phys. B **455**, 577 (1995) [arXiv:hep-th/9506193]. B. Eynard and C. Kristjansen, "More on the exact solution of the $O(n)$ model on a random lattice and an investigation of the case $|n| > 2$," Nucl. Phys. B **466**, 463 (1996) [arXiv:hep-th/9512052].

[33] D. S. Freed and R. E. Gompf, "Computer calculation of Witten's three manifold invariant," Commun. Math. Phys. **141** (1991) 79.

[34] P. Freyd, D. Yetter, J. Hoste, W.B.R. Lickorish, K. Millett and A.A. Ocneanu, "A new polynomial invariant of knots and links," Bull. Amer. Math. Soc. (N.S.) **12** (1985) 239.

[35] W. Fulton and J. Harris, *Representation theory: a first course*, Springer-Verlag.

[36] R. Gopakumar and C. Vafa, "M-theory and topological strings. I, II" hep-th/9809187, hep-th/9812127.

[37] R. Gopakumar and C. Vafa, "On the gauge theory/geometry correspondence," hep-th/9811131, Adv. Theor. Math. Phys. **3** (1999) 1415.

[38] S. Govindarajan, T. Jayaraman and T. Sarkar, "Disc instantons in linear sigma models," hep-th/0108234.

[39] T. Graber and E. Zaslow, "Open string Gromov-Witten invariants: Calculations and a mirror 'theorem'," hep-th/0109075.

[40] E. Guadagnini, M. Martellini and M. Mintchev, "Wilson lines in Chern-Simons theory and link invariants," Nucl. Phys. **B 330** (1990) 575.

[41] E. Guadagnini, "The universal link polynomial," Int. J. Mod. Phys. **A 7** (1992) 877; *The link invariants of the Chern-Simons field theory,* Walter de Gruyter, 1993.

[42] N. Halmagyi and V. Yasnov, "Chern-Simons matrix models and unoriented strings," arXiv:hep-th/0305134.

[43] J. Harris and I. Morrison, *Moduli of curves*, Springer-Verlag, 1998.

[44] P. Horava, "Equivariant topological sigma models," Nucl. Phys. B **418**, 571 (1994) [arXiv:hep-th/9309124].

[45] A. Iqbal, "All genus topological string amplitudes and 5-brane webs as Feynman diagrams," hep-th/0207114.

[46] A. Iqbal and A. K. Kashani-Poor, "Discrete symmetries of the superpotential and calculation of disk invariants," hep-th/0109214.

[47] L. C. Jeffrey, "Chern-Simons-Witten invariants of lens spaces and torus bundles, and the semiclassical approximation," Commun. Math. Phys. **147** (1992) 563.

[48] S. Katz, A. Klemm and C. Vafa, "M-theory, topological strings and spinning black holes," Adv. Theor. Math. Phys. **3** (1999) 1445, hep-th/9910181.

[49] S. Katz and C. C. Liu, "Enumerative geometry of stable maps with Lagrangian boundary conditions and multiple covers of the disc," math.ag/0103074, Adv. Theor. Math. Phys. **5** (2002) 1.

[50] V. A. Kazakov, M. Staudacher and T. Wynter, "Character expansion methods for matrix models of dually weighted graphs," Commun. Math. Phys. **177**, 451 (1996) [arXiv:hep-th/9502132].

[51] A. Klemm and E. Zaslow, "Local mirror symmetry at higher genus," hep-th/9906046, in *Winter School on Mirror Symmetry, Vector bundles and Lagrangian Submanifolds*, p. 183, American Mathematical Society 2001.

[52] M. Kontsevich, "Enumeration of rational curves via torus actions," hep-th/9405035, in *The moduli space of curves*, p. 335, Birkhäuser, 1995.

[53] A.B.J. Kuijlaars and W. Van Assche, "The asymptotic zero distribution of orthogonal polynomials with varying recurrence coefficients," J. Approximation Theory **99**, No.1, 167-197 (1999). Y. Chen and N. Lawrence, "Density of zeros of some orthogonal polynomials," Methods Appl. Anal. **5**, No.4, 367-386 (1998)

[54] J.M.F. Labastida, "Chern-Simons theory: ten years after," hep-th/9905057.

[55] J.M.F. Labastida and P.M. Llatas, "Topological matter in two dimensions," hep-th/9112051, Nucl. Phys. **B 379** (1992) 220.

[56] J.M.F. Labastida, M. Mariño and C. Vafa, "Knots, links and branes at large N," JHEP **0011**, 007 (2000), hep-th/0010102.

[57] J.M.F. Labastida and A. V. Ramallo, "Operator formalism for Chern-Simons theories," Phys. Lett. **B 227** (1989) 92. J.M.F. Labastida, P. M. Llatas and A. V. Ramallo, "Knot operators in Chern-Simons gauge theory," Nucl. Phys. **B 348** (1991) 651.

[58] W. Lerche and P. Mayr, "On $\mathcal{N}=1$ mirror symmetry for open type II strings," hep-th/0111113. W. Lerche, P. Mayr and N. Warner, "$\mathcal{N}=1$ special geometry, mixed Hodge variations and toric geometry," hep-th/0208039.

[59] N.C. Leung and C. Vafa, "Branes and toric geometry," hep-th/9711013, Adv. Theor. Math. Phys. **2** (1998) 91.

[60] J. Li and Y. S. Song, "Open string instantons and relative stable morphisms," hep-th/0103100, Adv. Theor. Math. Phys. **5** (2002) 67.

[61] W.B.R. Lickorish, *An introduction to knot theory*, Springer-Verlag, 1998.

[62] I.G. MacDonald, *Symmetric functions and Hall polynomials*, 2nd edition, Oxford University Press, 1995.

[63] M. Mariño, "Chern-Simons theory, matrix integrals, and perturbative three-manifold invariants," arXiv:hep-th/0207096.

[64] M. Mariño and C. Vafa, "Framed knots at large N," hep-th/0108064.

[65] P. Mayr, "$\mathcal{N}=1$ mirror symmetry and open/closed string duality," hep-th/0108229.

[66] P. Mayr, "Summing up open string instantons and $\mathcal{N} = 1$ string amplitudes," hep-th/0203237.

[67] M.L. Mehta, *Random matrices*, 2nd edition, Academic Press, 1991.

[68] A. Migdal, "Loop Equations And 1/N Expansion," Phys. Rept. **102**, 199 (1983).

[69] G. Moore and W. Taylor, JHEP **0201**, 004 (2002) [arXiv:hep-th/0111069].

[70] H.R. Morton and S.G. Lukac, "The HOMFLY polynomial of the decorated Hopf link," math.GT/0108011. S.G. Lukac, "HOMFLY skeins and the Hopf link," Ph.D. Thesis, June 2001, in http://www.liv.ac.uk/ su14/knotgroup.html.

[71] N.I. Muskhelishvili, *Singular integral equations*, Noordhoff, 1953.

[72] H. Ooguri, Y. Oz and Z. Yin, "D-branes on Calabi-Yau spaces and their mirrors," hep-th/9606112, Nucl. Phys. **B 477** (1996) 407.

[73] H. Ooguri and C. Vafa, "Knot invariants and topological strings," Nucl. Phys. B **577** (2000) 419, hep-th/9912123.

[74] H. Ooguri and C. Vafa, "Worldsheet derivation of a large N duality," hep-th/0205297, Nucl. Phys. **B 641** (2002) 3.

[75] R.C. Penner, "Perturbative series and the moduli space of Riemann surfaces," J. Differ. Geom. **27**, No.1, 35-53 (1988).

[76] V. Periwal, "Topological closed string interpretation of Chern-Simons theory," hep-th/9305115, Phys. Rev. Lett. **71** (1993) 1295.

[77] J. Polchinski, *String theory*, vol. 1, Cambridge University Press, 1998.

[78] A.M. Polyakov, "Fermi-Bose transmutations induced by gauge fields," Mod. Phys. Lett. **A 3** (1988) 325.

[79] V.V. Prasolov and A.B. Sossinsky, *Knots, links, braids and 3-manifolds*, American Mathematical Society, 1997.

[80] L. Rozansky, "A large k asymptotics of Witten's invariant of Seifert manifolds," hep-th/9303099, Commun. Math. Phys. **171** (1995) 279; "A Contribution To The Trivial Connection To Jones Polynomial And Witten's Invariant Of 3-D Manifolds. 1," hep-th/9401061, Commun. Math. Phys. **175**, 275 (1996).

[81] C. H. Taubes, "Lagrangians for the Gopakumar-Vafa conjecture," math.dg/0201219, Adv. Theor. Math. Phys. **5** (2001) 139.

[82] W. Taylor, "D-brane effective field theory from string field theory," hep-th/0001201, Nucl. Phys. **B 585** (2000) 171.

[83] W. Taylor, "Lectures on D-branes, tachyon condensation, and string field theory," arXiv:hep-th/0301094.

[84] G. 't Hooft, "A planar diagram theory for strong interactions," Nucl. Phys. **B 72** (1974) 461.

[85] M. Tierz, "Soft matrix models and Chern-Simons partition functions," arXiv:hep-th/0212128.

[86] C. Vafa, "Superstrings and topological strings at large N," hep-th/0008142, J. Math. Phys. **42** (2001) 2798.

[87] E. Witten, "Noncommutative geometry and string field theory," Nucl. Phys. **B 268** (1986) 253.

[88] E. Witten, "Topological sigma models," Commun. Math. Phys. **118** (1988) 411.

[89] E. Witten, "Quantum field theory and the Jones polynomial," Commun. Math. Phys. **121** (1989) 351.

[90] E. Witten, "Mirror manifolds and topological field theory," hep-th/9112056.

[91] E. Witten, "Chern-Simons gauge theory as a string theory," hep-th/9207094, Prog. Math. **133** (1995) 637.

[92] E. Witten, "Phases of $\mathcal{N} = 2$ theories in two dimensions," hep-th/9301042, Nucl. Phys. **B 403** (1993) 159.