Chern-Simons theory and topological strings

Marcos Mariño*

Department of Physics, CERN, Geneva 23, Switzerland

(Published 4 August 2005)

A review of the relation between Chern-Simons gauge theory and topological string theory on noncompact Calabi-Yau spaces is given. This relation has made it possible to give an exact solution of topological string theory on these spaces to all orders in the string coupling constant. Here the focus is on the construction of this solution, which is encoded in the topological vertex, and the implications of the physics of string/gauge theory duality for knot theory and for the geometry of Calabi-Yau manifolds.

CONTENTS

I. INTRODUCTION

Even though string theory has not found yet a clear place in our understanding of Nature, it has already established itself as a source of fascinating results and research directions in mathematics. In recent years, string theory and some of its close cousins (such as conformal field theory and topological field theory) have had an enormous impact in representation theory, differential geometry, low-dimensional topology, and algebraic geometry.

One mathematical area which has been deeply influenced by conformal field theory and topological field theory is knot theory. Witten (1989) found that many topological invariants of knots and links discovered in the 1980s (like the Jones and the HOMFLY polynomials) could be reinterpreted as correlation functions of Wilson loop operators in Chern-Simons theory, a gauge theory in three dimensions with topological invariance. Witten also showed that the partition function of this theory provided a new topological invariant of three-manifolds, and by working out the exact solution of Chern-Simons gauge theory he made a connection between these knot and three-manifold invariants and con-

*Also at Departamento de Matemática, IST, Lisboa, Portugal. Electronic address: marcos@mail.cern.ch

formal field theory in two dimensions (in particular, the Wess-Zumino-Witten model).

In a seemingly unrelated development, it was found that the study of string theory on Calabi-Yau manifolds (which was triggered by the phenomenological interest of the resulting four-dimensional models) provided new insights in the geometry of these spaces. Some correlation functions of string theory on Calabi-Yau manifolds turn out to compute numbers of holomorphic maps from the string worldsheet to the target, therefore they contain information about the enumerative geometry of the Calabi-Yau spaces. This led to the introduction of Gromov-Witten invariants in mathematics as a way to capture this information. Moreover, the existence of a powerful duality symmetry of string theory in Calabi-Yau spaces—mirror symmetry—allowed the computation of generating functions for these invariants, and made possible to solve with physical techniques difficult enumerative problems [see Hori *et al.* (2003) for a review of these developments]. The existence of a topological sector in string theory which captured the enumerative geometry of the target space led also to the construction of simplified models of string theory which kept only the topological information of the more complicated, physical theory. These models are called topological string theories and turn out to provide in many cases exactly solvable models of string dynamics.

The key idea that allowed one to build a bridge between topological string theory and Chern-Simons theory was the gauge theory–string theory correspondence. It is an old idea, going back to 't Hooft (1974), that gauge theories can be described in the $1/N$ expansion by string theories. This idea has been difficult to implement, but in recent years some spectacular progress was made thanks to the work of Maldacena (1998), who found a duality between type-IIB string theory on $\mathrm{AdS}_5\times\mathbf{S}^5$ and $\mathcal{N}=4$ super-Yang-Mills with gauge group $U(N)$. It is then natural to ask if gauge theories which are simpler than $\mathcal{N}=4$ Yang-Mills—like, for example, Chern-Simons theory—also admit a string theory description. It was shown by Gopakumar and Vafa (1999) that Chern-Simons gauge theory on the three-sphere has in fact a closed string description in terms of topological string theory propagating on a particular Calabi-Yau target, the so-called resolved conifold.

The result of Gopakumar and Vafa has three important consequences. First of all, it provides a toy model of the gauge theory–string theory correspondence which makes it possible to test in detail general ideas about this duality. Second, it gives a stringy interpretation of invariants of knots in the three-sphere. More precisely, it establishes a relation between invariants of knots based on quantum groups and Gromov-Witten invariants of open strings propagating on the resolved conifold. These are *a priori* two very different mathematical objects, and in this way the physical idea of a correspondence between gauge theories and strings gives new and fascinating results in mathematics that we are only starting to unveil. Finally, one can use the results of Gopakumar and Vafa to completely solve topological string theory on certain Calabi-Yau threefolds in a closed form. As we will see, this gives the all-genus answer for certain string amplitudes, and it is in fact one of the few examples in string theory where such an answer is available. The all-genus solution to the amplitudes also encodes the information about all the Gromov-Witten invariants for those threefolds. Since the solution involves building blocks from Chern-Simons theory, it suggests yet another bridge between knot invariants and Gromov-Witten theory.

In this review we will focus on this last aspect. The organization of the review is the following: in Sec. II we give an introduction to the relevant aspects of Chern-Simons theory that will be needed for the applications to Calabi-Yau geometry. In particular, we give detailed results for the computation of the relevant knots and link invariants. In Sec. III we give a short review on the $1/N$ expansion of Chern-Simons theory, which is the approach that makes possible the connection to string theory. Section IV contains a review of closed and open topological string theory on Calabi-Yau threefolds, and we construct in full detail the geometry of noncompact, toric Calabi-Yau spaces, since these are the manifolds that we will be able to study by using the gauge theory–string theory correspondence. In Sec. V we establish the correspondence between Chern-Simons theory on the three-sphere and closed topological string theory on a resolved conifold. In Sec. VI we show how the arguments of Sec. V can be extended to construct gauge theory duals of topological string theory on more complicated noncompact, toric Calabi-Yau manifolds. In Sec. VII we complete this program by defining the topological vertex, an object that allows one to solve topological string theory on all noncompact, toric Calabi-Yau threefolds by purely combinatorial methods. We also give a detailed derivation of the topological vertex from Chern-Simons theory, and we give various applications of the formalism. The last section contains some conclusions and open directions for further research. A short appendix contains some elementary facts about the theory of symmetric polynomials that are used in the review.

There are many issues that we have not analyzed in detail in this review. For example, we have not discussed the mirror-symmetric side of the story, and we do not address in detail the relation between topological string amplitudes and type-II superstring amplitudes. We refer the reader to the excellent book by Hori *et al.* (2003) for an introduction to these topics. Other reviews of the topics discussed here can be found in Grassi and Rossi (2002) and Mariño (2002b).

II. CHERN-SIMONS THEORY AND KNOT INVARIANTS

A. Chern-Simons theory: Basic ingredients

In a groundbreaking paper, Witten (1989) showed that Chern-Simons gauge theory, which is a quantum field theory in three dimensions, provides a physical descrip-

tion of a wide class of invariants of three-manifolds and of knots and links in three-manifolds.[1] The Chern-Simons action with gauge group G on a generic three-manifold M is defined by[2]

$$S=\frac{k}{4\pi}\int_M \mathrm{Tr}\left(A\wedge dA+\frac{2}{3}A\wedge A\wedge A\right). \tag{1}$$

Here, k is the coupling constant, and A is a G-gauge connection on the trivial bundle over M. In this review we will mostly consider Chern-Simons theory with gauge group $G=U(N)$. As noticed by Witten (1989), since this action does not involve the metric, the resulting quantum theory is topological, at least formally. In particular, the partition function

$$Z(M)=\int [\mathcal{D}A]e^{iS} \tag{2}$$

should define a topological invariant of the manifold M. A detailed analysis shows that this is in fact the case, with an extra subtlety: the invariant depends not only on the three-manifold but also on a choice of framing (i.e., a choice of trivialization of the bundle $TM\oplus TM$). As explained by Atiyah (1990), 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 results for the partition functions will be presented in the canonical framing.

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 Lickorish (1998) and Prasolov and Sossinsky (1997)]. 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 irreducible representation R of $U(N)$, which gives the Wilson loop operator:

$$W_R^{\mathcal{K}}(A)=\mathrm{Tr}_R U_{\mathcal{K}}, \tag{3}$$

where

$$U_{\mathcal{K}}=P\exp\left(\oint_{\mathcal{K}} A\right) \tag{4}$$

is the holonomy around the knot. Equation (3) is a gauge invariant operator whose definition does not involve the metric on the three-manifold. The irreducible representations of $U(N)$ 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 $\mathcal{K}_\alpha$, $\alpha=1,\ldots,L$, we can in principle compute the correlation function,

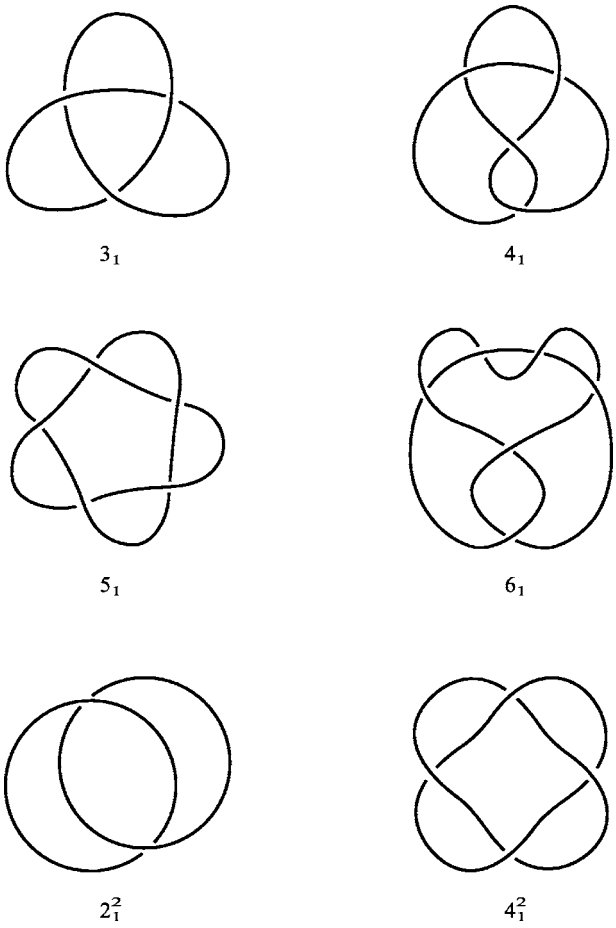

FIG. 1. Some knots and links. In the notation x_n^L, x indicates the number of crossings, L the number of components (in case it is a link with $L>1$), and n is a number used to enumerate knots and links in a given set characterized by x and L. The knot 3_1 is also known as the trefoil knot, while 4_1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

$$W_{R_1\cdots R_L}(\mathcal{L})=\langle W_{R_1}^{\mathcal{K}_1}\cdots W_{R_L}^{\mathcal{K}_L}\rangle = \frac{1}{Z(M)}\int [\mathcal{D}A]\left(\prod_{\alpha=1}^{L} W_{R_\alpha}^{\mathcal{K}_\alpha}\right)e^{iS}. \tag{5}$$

The topological character of the action and the fact that the Wilson loop operators can be defined without using any metric on the three-manifold indicate that Eq. (5) is a topological invariant of the link $\mathcal{L}$. Note that we are taking the knots and links to be oriented, and this makes a difference. If $\mathcal{K}^{-1}$ denotes the knot obtained from $\mathcal{K}$ by inverting its orientation, we have that

$$\mathrm{Tr}_R U_{\mathcal{K}^{-1}}=\mathrm{Tr}_R U_{\mathcal{K}}^{-1}=\mathrm{Tr}_{\bar{R}} U_{\mathcal{K}}, \tag{6}$$

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

[1]This was also conjectured by Schwarz (1987).

[2]The notation for integrals over differential forms may be found in Polchinski (1998, App. B).

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

$$\mathrm{lk}(\mathcal{K}_1,\mathcal{K}_2)=\frac{1}{2}\sum_{\mathrm{p}}\epsilon(p), \tag{7}$$

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. The linking number of a link $\mathcal{L}$ with components $\mathcal{K}_\alpha$, $\alpha=1,\ldots,L$, is defined by

$$\mathrm{lk}(\mathcal{L})=\sum_{\alpha<\beta}\mathrm{lk}(\mathcal{K}_\alpha,\mathcal{K}_\beta). \tag{8}$$

Some of the correlation functions of Wilson loops in Chern-Simons theory turn out to be closely related to important polynomial invariants of knots and links. For example, one of the most important polynomial invariants of a link $\mathcal{L}$ is the HOMFLY polynomial $P_{\mathcal{L}}(q,\lambda)$, which depends on two variables q and λ and was introduced by Freyd *et al.* (1985). This polynomial turns out to be related to the correlation function (5) when the gauge group is $U(N)$ and all the components are in the fundamental representation $R_\alpha=\square$. More precisely, we have

$$W_{\square\cdots\square}(\mathcal{L})=\lambda^{\mathrm{lk}(\mathcal{L})}\left(\frac{\lambda^{1/2}-\lambda^{-1/2}}{q^{1/2}-q^{-1/2}}\right)P_{\mathcal{L}}(q,\lambda), \tag{9}$$

where $\mathrm{lk}(\mathcal{L})$ is the linking number of $\mathcal{L}$, and the variables q and λ are related to the Chern-Simons variables as

$$q=\exp\left(\frac{2\pi i}{k+N}\right),\quad \lambda=q^N. \tag{10}$$

When $N=2$ the HOMFLY polynomial reduces to a one-variable polynomial, the Jones polynomial. When the gauge group of Chern-Simons theory is $SO(N)$, $W_{\square\cdots\square}(\mathcal{L})$ is closely related to the Kaufmann polynomial. For the mathematical definition and properties of these polynomials, see, for example, Lickorish (1998).

B. Perturbative approach

The partition function and correlation functions 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 to find first the classical solutions of the Chern-Simons equations of motion. If we write $A=\Sigma_a A^a T_a$, where T_a is a basis of the Lie algebra, we find

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

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

$$\pi_1(M)\to G. \tag{11}$$

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:

$$Z(M)=\sum_c Z^{(c)}(M), \tag{12}$$

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

$$Z^{(c)}(M)=Z^{(c)}_{\text{1-loop}}(M)\exp\left\{\sum_{\ell=1}^{\infty}S_\ell^{(c)}x^\ell\right\}. \tag{13}$$

In this equation, x is the effective expansion parameter:

$$x=\frac{2\pi i}{k+y}, \tag{14}$$

and y is the dual Coxeter number of the group [for $G=U(N)$, $y=N$]. The one-loop correction $Z^{(c)}_{\text{1-loop}}(M)$ was first analyzed by Witten (1989), and has been studied in great detail since then (Freed and Gompf, 1991; Jeffrey, 1992; Rozansky, 1995). It has the form

$$Z^{(c)}_{\text{1-loop}}(M)\propto\frac{\sqrt{|\tau_R^{(c)}|}}{\mathrm{vol}(H_c)}, \tag{15}$$

where $\tau_R^{(c)}$ is the Reidemeister-Ray-Singer torsion of $A^{(c)}$ and H_c is the isotropy group of $A^{(c)}$. Note that, for the trivial flat connection $A^{(c)}=0$, $H_c=G$.

The terms $S_\ell^{(c)}$ in Eq. (13) correspond to connected diagrams with 2ℓ 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. Note that the computation of $S_\ell^{(c)}$ involves the evaluation of group factors of Feynman diagrams, and therefore they depend explicitly on the gauge group G. When $G=U(N)$, they are polynomials in N. For example, S_1 contains the group factor $2N(N^2-1)$.

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

$$W_R^{\mathcal{K}}(A) = \mathrm{Tr}_R\left[\mathbf{1} + \oint_{\mathcal{K}} dx^\mu A_\mu(x) + \oint_{\mathcal{K}} dx^\mu \int^x dy^\nu A_\nu(y) A_\mu(x) + \cdots\right], \tag{16}$$

where $A_\mu = \Sigma_a A_\mu^a T_a$. Then, after gauge fixing, one can proceed and evaluate the correlation functions in standard perturbation theory. The perturbative study of Wilson loops was started by Guadagnini, Martellini, and Mintchev (1990). A nice review of its development can be found in Labastida (1999). Here we will rather focus on the nonperturbative approach to Chern-Simons theory, which we now explain.

C. Canonical quantization and surgery

As shown by Witten (1989), Chern-Simons theory is exactly solvable using nonperturbative methods and the relation to the Wess-Zumino-Witten (WZW) model. In order to present this solution, it is convenient to recall some basic facts about the canonical quantization of the model.

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 wave function $\Psi_{M,\mathcal{O}}(\mathcal{A})$ which is a functional of the values of the field at Σ. Schematically we have

$$\Psi_{M,\mathcal{O}}(\mathcal{A}) = \langle \mathcal{A}|\Psi_{M,\mathcal{O}}\rangle = \int_{A|_\Sigma = \mathcal{A}} \mathcal{D}A e^{iS}\mathcal{O}. \tag{17}$$

In fact, associated with the Riemann surface Σ we have a Hilbert space $\mathcal{H}(\Sigma)$, which can be obtained by doing canonical quantization of Chern-Simons theory on $\Sigma \times \mathbf{R}$. Before providing 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 Heegaard 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 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 then compute the full path integral (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 then produces a state $\langle \Psi_{M_2}| \in \mathcal{H}^*(\Sigma)$. The homeomorphism $f:\Sigma \to \Sigma$ is represented by an operator acting on $\mathcal{H}(\Sigma)$,

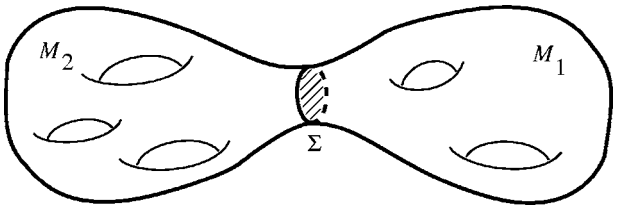

FIG. 3. Heegaard splitting of a three-manifold M into two three-manifolds M_1 and M_2 with a common boundary Σ.

$$U_f : \mathcal{H}(\Sigma) \to \mathcal{H}(\Sigma), \tag{18}$$

and the partition function can be finally evaluated as

$$Z(M) = \langle \Psi_{M_2}|U_f|\Psi_{M_1}\rangle. \tag{19}$$

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

One of the most fundamental results of Witten (1989) is in fact a precise description of $\mathcal{H}(\Sigma)$: it is the space of conformal blocks of a Wess-Zumino-Witten model on Σ with gauge group G and level k [for an extensive review of the Wess-Zumino-Witten model, see, for example, Di Francesco *et al.* (1997)]. 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 Wess-Zumino-Witten model in order to solve Chern-Simons 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 with G at level k. We will use the following notations: the fundamental weights of G will be denoted by λ_i, and the simple roots by α_i, with $i = 1, \ldots, r$, and r denotes the rank of G. The weight and root lattices of G are denoted by Λ^w and Λ^r, respectively, and $|\Delta_+|$ denotes the number of positive roots. The fundamental chamber $\mathcal{F}_l$ is given by $\Lambda^w / l\Lambda^r$, modded out by the action of the Weyl group. For example, in $SU(N)$ a weight $p = \Sigma_{i=1}^r p_i \lambda_i$ is in $\mathcal{F}_l$ if

$$\sum_{i=1}^r p_i < l, \quad \text{and} \quad p_i > 0,\ i = 1, \ldots, r. \tag{20}$$

We recall that a representation given by a highest weight Λ is integrable if $\rho + \Lambda$ 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). In the following, 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 Wess-Zumino-Witten 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 (Elitzur *et al.*, 1989; Labastida and Ramallo, 1989; Witten, 1989), so we have

$$\langle R|R'\rangle = \delta_{RR'}. \tag{21}$$

There is a special class of homeomorphisms of $\mathbf{T}^2$ that has a simple expression as operators in $\mathcal{H}(\mathbf{T}^2)$; these are the $\mathrm{Sl}(2,\mathbf{Z})$ 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

$$T=\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}, \quad S=\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}. \tag{22}$$

Note that the S transformation 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:

$$T_{pp'} = \delta_{p,p'} e^{2\pi i(h_p - c/24)},$$

$$S_{pp'} = \frac{i^{|\Delta_+|}}{(k+y)^{r/2}}\left(\frac{\mathrm{Vol}\ \Lambda^w}{\mathrm{Vol}\ \Lambda^r}\right)^{1/2} \times \sum_{w\in\mathcal{W}} \epsilon(w)\exp\left(-\frac{2\pi i}{k+y} p\cdot w(p')\right). \tag{23}$$

In the first equation, c is the central charge of the Wess-Zumino-Witten model, and h_p is the conformal weight of the primary field associated to p:

$$h_p = \frac{p^2-\rho^2}{2(k+y)}, \tag{24}$$

where 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 $\mathcal{W}$, $\epsilon(w)$ is the signature of the element w, and Vol $\Lambda^{w(r)}$ denote, respectively, the volume of the weight (root) lattice. We often write $S_{RR'}$ for $S_{pp'}$, where $p=\rho+\Lambda$, $p'=\rho+\Lambda'$, and Λ,Λ' are the highest weights corresponding to the representations R, R'.

What is the description of the states $|R\rangle$ in $\mathcal{H}(\mathbf{T}^2)$ from the point of view of canonical quantization? Consider the solid torus $\mathcal{T}=D\times\mathbf{S}^1$, where D is a disk in $\mathbf{R}^2$. This is a three-manifold whose boundary is a $\mathbf{T}^2$, and it has a noncontractible cycle given by $\mathbf{S}^1$. 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

$$|\Psi_{\mathcal{T},\mathcal{O}_R}\rangle = |R\rangle. \tag{25}$$

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 Heegaard splitting along a torus. Imagine, for example, that we take two solid tori and we glue them along their boundary with the identity map. Since a solid torus is a disk times a circle $D\times\mathbf{S}^1$, by performing this operation we get a manifold which is $\mathbf{S}^1$ times the two disks glued together along their boundaries. Therefore, with this operation we obtain $\mathbf{S}^2\times\mathbf{S}^1$, and Eq. (19) gives

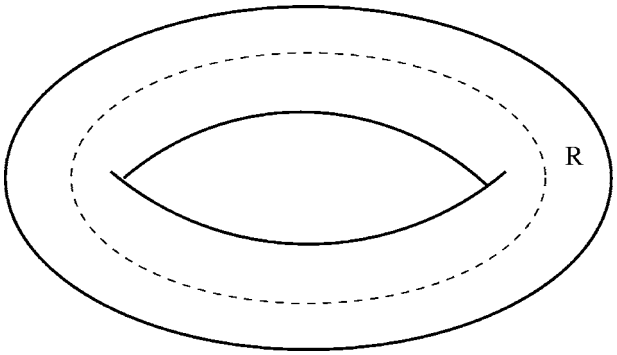

FIG. 4. Performing the path integral on a solid torus with a Wilson line in representation R gives the state $|R\rangle$ in $\mathcal{H}(\mathbf{T}^2)$.

$$Z(\mathbf{S}^2\times\mathbf{S}^1) = \langle 0|0\rangle = 1. \tag{26}$$

If we preform the gluing, however, after an S transformation on the $\mathbf{T}^2$ the resulting manifold is instead $\mathbf{S}^3$. To see this, note 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

$$Z(\mathbf{S}^3) = \langle 0|S|0\rangle = S_{00}. \tag{27}$$

By using Weyl's denominator formula,

$$\sum_{w\in\mathcal{W}} \epsilon(w) e^{w(\rho)} = \prod_{\alpha>0} 2\sinh\frac{\alpha}{2}, \tag{28}$$

where $\alpha>0$ are positive roots, one finds

$$Z(\mathbf{S}^3) = \frac{1}{(k+y)^{r/2}}\left(\frac{\mathrm{Vol}\ \Lambda^w}{\mathrm{Vol}\ \Lambda^r}\right)^{1/2} \prod_{\alpha>0} 2\sin\left(\frac{\pi(\alpha\cdot\rho)}{k+y}\right). \tag{29}$$

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

$$Z(\mathbf{S}_3,\mathcal{O}_R) = \langle 0|S|R\rangle. \tag{30}$$

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

$$W_R(\text{unknot}) = \frac{S_{0R}}{S_{00}} = \frac{\sum_{w\in\mathcal{W}} \epsilon(w) e^{-[2\pi i/(k+y)]\rho\cdot w(\Lambda+\rho)}}{\sum_{w\in\mathcal{W}} \epsilon(w) e^{-[2\pi i/(k+y)]\rho\cdot w(\rho)}}. \tag{31}$$

Recall that the character of the representation R, evaluated on an element $a \in \Lambda_w \otimes \mathbf{R}$, is defined by

$$\mathrm{ch}_R(a) = \sum_{\mu\in M_R} e^{a\cdot\mu}, \tag{32}$$

where M_R is the set of weights associated to the irreducible representation R. By using Weyl's character formula we can write

$$W_R(\text{unknot}) = \mathrm{ch}_R\left[-\frac{2\pi i}{k+y}\rho\right]. \tag{33}$$

Moreover, using Eq. (28), we finally obtain

$$W_R(\text{unknot}) = \prod_{\alpha>0} \frac{\sin\left(\frac{\pi}{k+y}\alpha\cdot(\Lambda+\rho)\right)}{\sin\left(\frac{\pi}{k+y}\alpha\cdot\rho\right)}. \tag{34}$$

This quantity is often called the *quantum dimension* of R, and it is denoted by $\dim_q R$.

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 the orientation carefully into account, we find that this is the Hopf link with linking number +1. The path integral with this insertion is

$$Z(\mathbf{S}^3, \mathcal{O}_R\mathcal{O}_{R'}) = \langle R'|S|R\rangle, \tag{35}$$

so the normalized vacuum expectation value is

$$\mathcal{W}_{RR'} \equiv W_{RR'}(\text{Hopf}^{+1}) = \frac{S_{\bar{R}'R}}{S_{00}} = \frac{S^{-1}_{R'R}}{S_{00}}, \tag{36}$$

where the superscript +1 refers to the linking number. Here we have used that the bras $\langle R|$ are canonically associated to conjugate representations $\bar{R}$, and that $S_{\bar{R}'R} = S^{-1}_{R'R}$ [see, for example, Di Francesco *et al.* (1997)]. 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 note that the two Hopf links can be related by changing the orientation of one of the components. We then have

$$W_{RR'}(\text{Hopf}^{-1}) = \frac{S_{R'R}}{S_{00}}, \tag{37}$$

where we have used the property (6).

When we take $G=U(N)$, the above vacuum expectation values for unknots and Hopf links can be evaluated explicitly in terms of Schur polynomials. It is well known that the character of the unitary group in the representation R is given by the Schur polynomial s_R [see, for example, Fulton and Harris (1991)]. There is a precise relation between the element a where one evaluates the character in Eq. (32) and the variables entering the Schur polynomial. Let μ_i, $i=1,\ldots,N$, be the weights associated to the fundamental representation of $U(N)$. Note that if R is given by a Young tableau whose rows have lengths $l_1 \geqslant \cdots \geqslant l_N$, then $\Lambda_R = \Sigma_i l_i \mu_i$. We also have

$$\rho = \sum_{i=1}^{N} \frac{1}{2}(N-2i+1)\mu_i. \tag{38}$$

Let $a \in \Lambda^w \otimes \mathbf{R}$ be given by

$$a = \sum_{i=1}^{N} a_i \mu_i. \tag{39}$$

Then,

$$\mathrm{ch}_R[a] = s_R(x_i = e^{a_i}). \tag{40}$$

For example, in the case of the quantum dimension, one has $\dim_q R = \dim_q \bar{R}$, and we find

$$\dim_q R = s_R(x_i = q^{(1/2)(N-2i+1)}), \tag{41}$$

where q is given in Eq. (10). By using that s_R is homogeneous of degree $\ell(R)$ in the coordinates x_i we finally obtain

$$\dim_q R = \lambda^{\ell(R)/2} s_R(x_i = q^{-i+1/2}),$$

where $\lambda = q^N$ as in Eq. (10), and there are N variables x_i. The quantum dimension can be written explicitly in terms of the q numbers:

$$[x] = q^{x/2} - q^{-x/2}, \quad [x]_\lambda = \lambda^{1/2} q^{x/2} - \lambda^{-1/2} q^{-x/2}. \tag{42}$$

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

$$\dim_q R = \prod_{1\leqslant i<j\leqslant c_R} \frac{[l_i - l_j + j - i]}{[j-i]} \prod_{i=1}^{c_R} \frac{\prod_{v=-i+1}^{l_i-i} [v]_\lambda}{\prod_{v=1}^{l_i} [v-i+c_R]}. \tag{43}$$

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

The S-matrix elements that appear in Eqs. (36) and (37) can be evaluated through the explicit expression (23), by using the relation between $U(N)$ characters and Schur functions that we explained above. Note first that

$$\frac{S^{-1}_{R_1R_2}}{S_{00}} = \mathrm{ch}_{R_1}\left[\frac{2\pi i}{k+y}(\Lambda_{R_2}+\rho)\right]\mathrm{ch}_{R_2}\left[\frac{2\pi i}{k+y}\rho\right]. \tag{44}$$

If we denote by $l_i^{R_2}$, $i=1,\ldots,c_{R_2}$ the lengths of rows for the Young tableau corresponding to R_2, it is easy to see that

$$\mathcal{W}_{R_1R_2}(q,\lambda) = (\lambda q)^{\ell(R_1)/2} s_{R_1}(x_i = q^{l_i^{R_2}-i})\dim_q R_2, \tag{45}$$

where we set $l_i^{R_2}=0$ for $i>c_{R_2}$. A convenient way to evaluate $s_{R_1}(x_i=q^{l_i^R-i})$ for a partition $\{l_i^R\}_{\{i=1,\ldots,c_R\}}$ associated with R is to use the Jacobi-Trudy formula (A6). It is easy to show that the generating functional of elementary symmetric functions (A2) for this specialization is given by

$$E_R(t) = E_\emptyset(t)\prod_{j=1}^{c_R}\frac{1+q^{l_j^R-j}t}{1+q^{-j}t}, \tag{46}$$

where

$$E_\emptyset(t) = 1+\sum_{n=1}^{\infty} a_n t^n, \tag{47}$$

and the coefficients a_n are defined by

$$a_n\prod_{r=1}^{n}\frac{1-\lambda^{-1}q^{r-1}}{q^r-1}. \tag{48}$$

The formula (45), together with the expressions above for $E_R(t)$, provides an explicit expression for Eq. (36) as a rational function of $q^{\pm 1/2}$, $\lambda^{\pm 1/2}$, and it was first written down by Morton and Lukac (2003).

D. Framing dependence

In the above discussion on the correlation functions of Wilson loops we have missed an important ingredient. We already 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 Eq. (5) also involves a choice of framing of the knots, as discovered by Witten (1989). Since this is important in the context of 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 $U(1)$. The Abelian Chern-Simons theory turns out to be extremely simple, since the cubic term in Eq. (1) drops out, and we are left with a Gaussian theory (Polyakov, 1988). $U(1)$ representations are labeled by integers, and the correlation function (5) can be computed exactly. In order to do that, however, one has to choose a framing for each of the knots $\mathcal{K}_\alpha$. This arises as follows: in evaluating the correlation function, contractions of the holonomies corresponding to different $\mathcal{K}_\alpha$ produce the following integral:

$$\mathrm{lk}(\mathcal{K}_\alpha,\mathcal{K}_\beta) = \frac{1}{4\pi}\oint_{\mathcal{K}_\alpha} dx^\mu \oint_{\mathcal{K}_\beta} dy^\nu \epsilon_{\mu\nu\rho}\frac{(x-y)^\rho}{|x-y|^3}. \tag{49}$$

This is a topological invariant, i.e., it is invariant under deformations of the knots $\mathcal{K}_\alpha,\mathcal{K}_\beta$, and it is in fact the Gauss integral representation of their linking number $\mathrm{lk}(\mathcal{K}_\alpha,\mathcal{K}_\beta)$ defined in Eq. (7). On the other hand, contractions of the holonomies corresponding to the same knot $\mathcal{K}$ involve the integral

$$\phi(\mathcal{K}) = \frac{1}{4\pi}\oint_{\mathcal{K}} dx^\mu \oint_{\mathcal{K}} dy^\nu \epsilon_{\mu\nu\rho}\frac{(x-y)^\rho}{|x-y|^3}. \tag{50}$$

This integral is well defined and finite [see, for example, Guadagnini, Martellini, and Mintchev (1990)], and it is called the *cotorsion* or *writhe* of $\mathcal{K}$. It gives the self-linking number of $\mathcal{K}$: if we project $\mathcal{K}$ on a plane, and we denote $n_\pm(\mathcal{K})$ the number of positive (negative) crossings as indicated in Fig. 2, then we have that

$$\phi(\mathcal{K}) = n_+(\mathcal{K}) - n_-(\mathcal{K}). \tag{51}$$

The problem is that the cotorsion is not invariant under deformations of the knot. In order to preserve topological invariance of the correlation function, one has to choose another definition of the composite operator $(\oint_{\mathcal{K}} A)^2$ by means of 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

$$\phi_f(\mathcal{K}) = \frac{1}{4\pi}\oint_{\mathcal{K}} dx^\mu \oint_{\mathcal{K}^f} dy^\nu \epsilon_{\mu\nu\rho}\frac{(x-y)^\rho}{|x-y|^3} = \mathrm{lk}(\mathcal{K},\mathcal{K}^f). \tag{52}$$

The correlation function that we obtain in this way is topological invariant (since it only involves linking numbers) but the price that we have to pay is that our regularization depends on a set of integers $p_\alpha=\mathrm{lk}(\mathcal{K}_\alpha,\mathcal{K}_\alpha^f)$ (one for each knot). The correlation function (5) can now be computed, after choosing the framings, as follows:

$$\left\langle \prod_\alpha \exp\left(n_\alpha \oint_{\mathcal{K}_\alpha} A\right)\right\rangle = \exp\left\{\frac{\pi i}{k}\left(\sum_\alpha n_\alpha^2 p_\alpha + \sum_{\alpha\neq\beta} n_\alpha n_\beta \mathrm{lk}(\mathcal{K}_\alpha,\mathcal{K}_\beta)\right)\right\}. \tag{53}$$

This regularization is nothing but the "point-splitting" method familiar in the context of quantum field theory.

Let us now consider Chern-Simons theory with gauge group $U(N)$, and suppose that we are interested in the computation of Eq. (5), in the context of perturbation theory. It is easy to see that self-contractions of the holonomies lead to the same kind of ambiguities that we found in the Abelian case, i.e., a choice of framing has to be made for each knot $\mathcal{K}_\alpha$. The only difference with the Abelian case is that the self-contraction of $\mathcal{K}_\alpha$ gives a group factor $\mathrm{Tr}_{R_\alpha}(T_a T_a)$, where T_a is a basis of the Lie algebra [see, for example, Guadagnini, Martellini, and

Mintchev (1990)]. The precise result can be better stated as the effect on the correlation function (5) under a change of framing, and it states that, under a change of framing of $\mathcal{K}_\alpha$ by p_α units, the vacuum expectation value of the product of Wilson loops changes as follows [Witten (1989)]:

$$W_{R_1\cdots R_L} \to \exp\left[2\pi i \sum_{\alpha=1}^{L} p_\alpha h_{R_\alpha}\right] W_{R_1\cdots R_L}. \tag{54}$$

In this equation, h_R is the conformal weight of the Wess-Zumino-Witten primary field corresponding to the representation R. One can write Eq. (24) as

$$h_R = \frac{C_R}{2(k+N)}, \tag{55}$$

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

$$C_R = N\ell(R) + \kappa_R, \tag{56}$$

where $\ell(R)$ is the total number of boxes in the tableau, and

$$\kappa_R = \ell(R) + \sum_i (l_i^2 - 2il_i). \tag{57}$$

In terms of the variables (10) the change under framing (54) can be written as

$$W_{R_1\cdots R_L} \to q^{(1/2)\Sigma_{\alpha=1}^{L}\kappa_{R_\alpha}p_\alpha}\lambda^{(1/2)\Sigma_{\alpha=1}^{L}\ell(R_\alpha)p_\alpha} W_{R_1\cdots R_L}. \tag{58}$$

Therefore, 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 requiring that the self-linking number is zero. The expressions we have given before for the Chern-Simons invariant of the unknot and the Hopf link are all 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_α can be easily obtained from Eq. (54).

E. More results on Wilson loops

In this subsection we discuss some useful results for the computation of vacuum expectation values of Wilson loops. Most of these results can be found, for example, in Guadagnini (1992).

The first property we want to state is the *factorization property* for the vacuum expectation values of disjoint links, which says the following. Let $\mathcal{L}$ be a link with L components $\mathcal{K}_1,\ldots,\mathcal{K}_L$ which are *disjoint* knots, and let us attach the representation R_α to the αth component. Then one has

$$W_{R_1\cdots R_L}(\mathcal{L}) = \prod_{\alpha=1}^{L} W_{R_\alpha}(\mathcal{K}_\alpha). \tag{59}$$

This property is easy to prove in Chern-Simons theory. It only involves some elementary procedure and the fact that $\mathcal{H}(\mathbf{S}^2)$ is one dimensional. A proof can be found in Witten (1989).

The second property we will consider is *parity symmetry*. 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 vice versa. 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

$$W_{R_1\cdots R_L}(\mathcal{L}^*)(q,\lambda) = W_{R_1\cdots R_L}(\mathcal{L})(q^{-1},\lambda^{-1}). \tag{60}$$

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, note, for example, that the unknot is identical to its mirror image; therefore, quantum dimensions satisfy

$$(\dim_q R)(q^{-1},\lambda^{-1}) = (\dim_q R)(q,\lambda). \tag{61}$$

Let us now discuss the simplest example of a *fusion rule* in Chern-Simons theory. Consider a vacuum expectation value of the form

$$\langle \mathrm{Tr}_{R_1} U \, \mathrm{Tr}_{R_2} U\rangle, \tag{62}$$

where U is the holonomy of the gauge field around a knot $\mathcal{K}$. The classical operator $\mathrm{Tr}_{R_1}U\,\mathrm{Tr}_{R_2}U$ can always be written as

$$\mathrm{Tr}_{R_1}U\,\mathrm{Tr}_{R_2}U = \mathrm{Tr}_{R_1\otimes R_2}U = \sum_R N_{R_1R_2}^R \mathrm{Tr}_R U, \tag{63}$$

where $R_1\otimes R_2$ denotes the tensor product, and $N_{R_1R_2}^R$ are tensor product coefficients. In Chern-Simons theory, the quantum Wilson loop operators satisfy a very similar relation, with the only difference that the coefficients become the fusion coefficients for integrable representations of the Wess-Zumino-Witten model. This can be easily understood if we take into account 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 product rules of classical traces have to be modified. However, in the computation of knot invariants in $U(N)$ Chern-Simons theory it is natural to work in a setting in which both k and N are much larger than any of the representations involved. In that case, the vacuum expectation values of the theory satisfy

$$\langle \mathrm{Tr}_{R_1}U\,\mathrm{Tr}_{R_2}U\rangle = \sum_R N_{R_1R_2}^R \langle \mathrm{Tr}_R U\rangle, \tag{64}$$

where $N_{R_1R_2}^R$ are the Littlewood-Richardson coefficients of $U(N)$. As a simple application of the fusion rule, imagine 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 left-hand side of Eq.

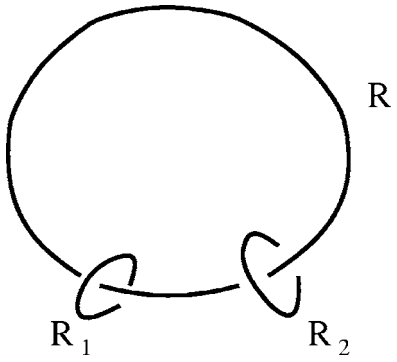

FIG. 5. A link with two disjoint unknots.

(64). Using also the factorization property (59), we deduce the following fusion rule:

$$\dim_q R_1 \dim_q R_2 = \sum_R N^R_{R_1 R_2} \dim_q R. \tag{65}$$

The last property we will state is the behavior of correlation functions under *direct sum*. This operation is defined as follows. Let us consider two links $\mathcal{L}_1, \mathcal{L}_2$ with components $\mathcal{K}_1, \mathcal{K}$ and $\mathcal{K}_2, \mathcal{K}$, respectively, i.e., the component knot $\mathcal{K}$ is the same in $\mathcal{L}_1$ and $\mathcal{L}_2$. The direct sum $\mathcal{L} = \mathcal{L}_1 \sharp \mathcal{L}_2$ is a link of three components which is obtained by joining $\mathcal{L}_1$ and $\mathcal{L}_2$ through $\mathcal{K}$. It is not difficult to prove that the Chern-Simons invariant of $\mathcal{L}$ is given by [Witten (1989)]

$$W_{R_1 R_2 R}(\mathcal{L}) = \frac{W_{R_1 R}(\mathcal{L}_1) W_{R_2 R}(\mathcal{L}_2)}{W_R(\mathcal{K})}. \tag{66}$$

As an application of this rule, let us consider the three-component link in Fig. 5. This link is a direct sum of two Hopf links whose common component is an unknot in representation R, and the knots $\mathcal{K}_1, \mathcal{K}_2$ are unknots in representations R_1, R_2. Equation (66) expresses the Chern-Simons invariant of $\mathcal{L}$ in terms of invariants of Hopf links and quantum dimensions. Note that the invariant of the link in Fig. 5 can be also computed using fusion rules. If we fuse the two parallel unknots with representations R_1, R_2, we find

$$W_{R_1 R_2 R}(\mathcal{L}) = \sum_{R'} N^{R'}_{R_1 R_2} \langle \mathrm{Tr}_{R'} U' \mathrm{Tr}_R U \rangle, \tag{67}$$

where U is the holonomy around the unknot in representation R, and U' is the holonomy around the unknot which is obtained by fusing the two parallel unknots in Fig. 5. Equation (67) expresses the invariant (66) in terms of the invariants of a Hopf link with representations R', R.

F. Generating functionals for knot and link invariants

In the applications of Chern-Simons theory, we will need the invariants of knots and links in arbitrary representations of the gauge group. It is then natural to consider generating functionals for Wilson loop operators in arbitrary representations.

There are two natural bases for the set of Wilson loop operators: the basis labeled by representations R, which is the one that we have considered so far, and the basis labeled by conjugacy classes of the symmetric group. Wilson loop operators in the basis of conjugacy classes are constructed as follows. Let U be the holonomy of the gauge connection around the knot $\mathcal{K}$. Let $\vec{k} = (k_1, k_2, \ldots)$ be a vector of infinite entries, almost all of which are zero. This vector defines naturally a conjugacy class $C(\vec{k})$ of the symmetric group S_ℓ with

$$\ell = \sum_j j k_j. \tag{68}$$

We will also denote

$$|\vec{k}| = \sum_j k_j. \tag{69}$$

The conjugacy class $C(\vec{k})$ is simply the class that has k_j cycles of length j. We now define the operator

$$\Upsilon_{\vec{k}}(U) = \prod_{j=1}^{\infty} (\mathrm{Tr}\, U^j)^{k_j}, \tag{70}$$

which gives the Wilson loop operator in the conjugacy class basis. It is a linear combination of the operators $\mathrm{Tr}_R U$ labeled by representations

$$\Upsilon_{\vec{k}}(U) = \sum_R \chi_R(C(\vec{k})) \mathrm{Tr}_R U, \tag{71}$$

where $\chi_R(C(\vec{k}))$ are the characters of the symmetric group S_ℓ in the representation R evaluated at the conjugacy class $C(\vec{k})$. The above formula can be inverted as

$$\mathrm{Tr}_R(U) = \sum_{\vec{k}} \frac{\chi_R(C(\vec{k}))}{z_{\vec{k}}} \Upsilon_{\vec{k}}(U), \tag{72}$$

with

$$z_{\vec{k}} = \prod_j k_j!\, j^{k_j}. \tag{73}$$

If U is a diagonal matrix $U = \mathrm{diag}(x_1, \ldots, x_N)$, it is an elementary result in the representation theory of the unitary group that $\mathrm{Tr}_R U$ is the Schur polynomial in x_i:

$$\mathrm{Tr}_R U = s_R(x). \tag{74}$$

It is immediate to see that

$$\Upsilon_{\vec{k}}(U) = P_{\vec{k}}(x), \tag{75}$$

where the Newton polynomials are defined in Eq. (A7). The relation (71) is nothing but the Frobenius formula, which relates the two bases of symmetric polynomials given by the Schur and Newton polynomials. The vacuum expectation values of the operators (70) will be denoted by

$$W_{\vec{k}} = \langle \Upsilon_{\vec{k}}(U) \rangle. \tag{76}$$

If V is a $U(M)$ matrix (a "source" term), one can define the following operator, which was introduced by Ooguri and Vafa (2000) and is known in this context as the Ooguri-Vafa operator:

$$Z(U,V)=\exp\left[\sum_{n=1}^{\infty}\frac{1}{n}\mathrm{Tr}\,U^n\mathrm{Tr}\,V^n\right]. \tag{77}$$

When expanded, this operator can be written in the $\vec{k}$ basis as follows:

$$Z(U,V)=1+\sum_{\vec{k}}\frac{1}{z_{\vec{k}}}\Upsilon_{\vec{k}}(U)\Upsilon_{\vec{k}}(V). \tag{78}$$

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

$$Z(U,V)=\sum_{R}\mathrm{Tr}_R(U)\mathrm{Tr}_R(V), \tag{79}$$

where the sum over representations starts with the trivial one. Note that in the above equation R is regarded as a Young tableau, and since we are taking both N and M to be large, it can be regarded as a representation of both $U(N)$ and $U(M)$. 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 $Z_{\mathrm{CS}}(V)=\langle Z(U,V)\rangle$ then has information about the vacuum expectation values of the Wilson loop operators, and by taking its logarithm one can define the connected vacuum expectation values $W_{\vec{k}}^{(c)}$:

$$F_{\mathrm{CS}}(V)=\ln Z_{\mathrm{CS}}(V)=\sum_{\vec{k}}\frac{1}{z_{\vec{k}}!}W_{\vec{k}}^{(c)}\Upsilon_{\vec{k}}(V). \tag{80}$$

One has, for example,

$$W_{(2,0,\ldots)}^{(c)}=\langle(\mathrm{Tr}U)^2\rangle-\langle\mathrm{Tr}U\rangle^2=W_{\square\square}+W_{\square}-W_{\square}^2.$$

The free energy $F_{\mathrm{CS}}(V)$, which is a generating functional for connected vacuum expectation values $W_{\vec{k}}^{(c)}$, is an important quantity when one considers the string/gauge theory correspondence, as we will see.

III. THE 1/N EXPANSION AND CHERN-SIMONS THEORY

A. The 1/N expansion

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 (1974) 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 "fat graphs" or ribbon graphs ('t Hooft, 1974).

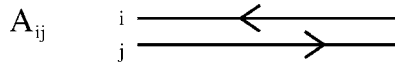

FIG. 6. The index structure of the fields in the adjoint representation of $U(N)$ is represented by a double line.

In the fat-graph approach to perturbation theory, the propagator of the gluon field A_{ij} is represented by a double line, as shown in Fig. 6. The indices $i,j=1,\ldots,N$ are gauge indices for the adjoint representation. Similarly, the trivalent vertex of Chern-Simons theory is represented in this notation as in Fig. 7. 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 fat graphs. A usual Feynman diagram can give rise to many different fat graphs. For example, the two-loop vacuum diagram $\ominus$ in Chern-Simons theory, which comes from contracting two cubic vertices, gives rise to two fat graphs. The first one, which is shown in Fig. 8, gives a group factor $2N^3$, while the second one, which is shown in Fig. 9, gives $-2N$. The advantage of introducing fat graphs is precisely that each of them gives a definite power of N: fat graphs are characterized topologically by the number of propagators E, the number of vertices V, and finally the number of closed loops h. If we denote by x the coupling constant, each propagator gives a power of x, each interaction vertex gives a power of x^{-1}, and each closed loop gives a power of N, so that every fat graph will give a contribution in x and N given by

$$x^{E-V}N^h. \tag{81}$$

The key point now is to regard the fat graph as a Riemann surface with holes, in which each closed loop rep-

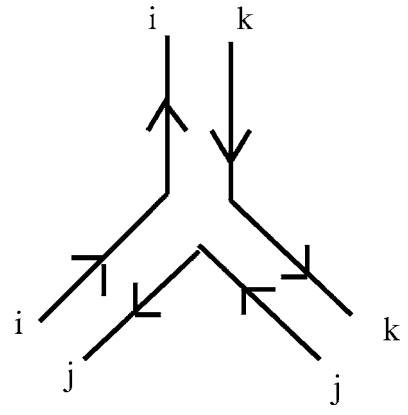

FIG. 7. The cubic vertex in the double-line notation.

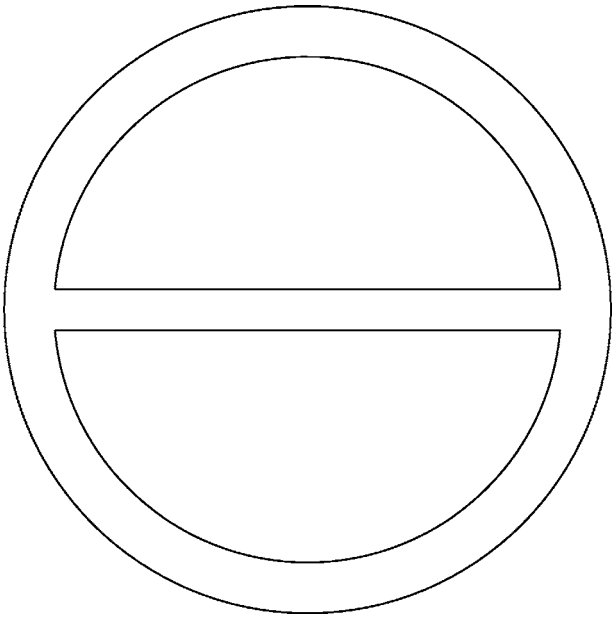

FIG. 8. A planar diagram with $h=3$.

resents the boundary of a hole. The genus of such a surface is determined by the elementary topological relation

$$2g-2=E-V-h, \tag{82}$$

therefore we can write Eq. (81) as

$$x^{2g-2+h}N^h=x^{2g-2}t^h, \tag{83}$$

where we have introduced the 't Hooft parameter $t=Nx$. Fat graphs with $g=0$ are called *planar*, while the ones with $g>0$ are called *nonplanar*. The diagram in Fig. 8, for example, is planar: it has $E=3$, $V=2$, and $h=3$, therefore $g=0$. The diagram in Fig. 9 is nonplanar: it has $E=3$, $V=2$, and $h=1$, therefore $g=1$.

We can now organize the computation of the different quantities in the field theory in terms of fat graphs. 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 can be written as

$$F^p=\sum_{g=0}^{\infty}\sum_{h=1}^{\infty}F^p_{g,h}x^{2g-2}t^h, \tag{84}$$

where $F^p_{g,h}$ is simply a number that can be computed by the usual rules of perturbation theory. The superscript p refers to the perturbative free energy. As we will see, the total free energy may have a nonperturbative contribution which is not captured by Feynman diagrams. In Eq. (84) 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 as $\mathcal{O}(N^2)$. The nonplanar diagrams give subleading corrections. Note 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 also 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, Coleman (1988)], and it turns out that correlation functions that have well-defined behavior are the connected vacuum expectation values $W^{(c)}_{\vec{k}}$ introduced in Eq. (80). They admit an expansion of the form

$$W^{(c)}_{\vec{k}}=\sum_{g=0}^{\infty}W_{\vec{k},g}(t)x^{2g-2+|\vec{k}|}, \tag{85}$$

where $W_{\vec{k},g}(t)$ is a function of the 't Hooft parameter and $|\vec{k}|$ is defined in Eq. (69).

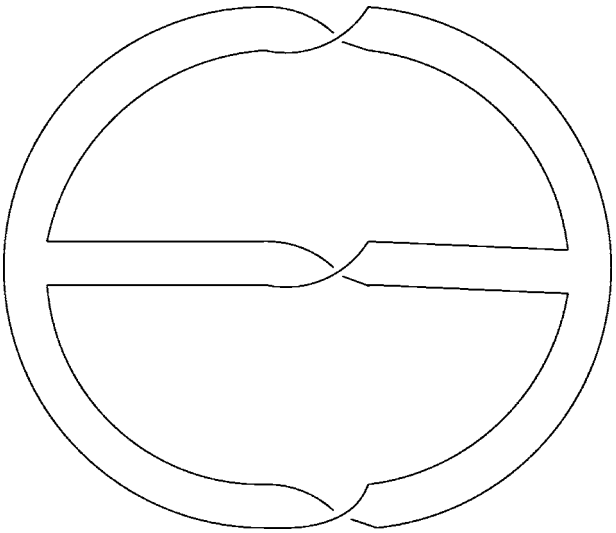

FIG. 9. A nonplanar diagram with $g=h=1$.

B. The 1/*N* expansion in Chern-Simons theory

The above considerations on the $1/N$ expansion are valid for any $U(N)$ gauge theory, and in particular one should be able to expand the free energy of Chern-Simons theory on the three-sphere as in Eq. (84).[3] Of course, the computation of $F_{g,h}$ in Chern-Simons theory directly from perturbation theory is difficult, since it involves the evaluation of integrals of products of propagators over the three-sphere. However, since we know the exact answer for the partition function, we just have to expand it to obtain Eq. (84) and the explicit expression for $F_{g,h}$.

[3]For earlier work on the $1/N$ expansion of Chern-Simons theory, see Camperi *et al.* (1990), Periwal (1993), and Correale and Guadagnini (1994).

The partition function of CS with gauge group $U(N)$ on the three-sphere can be obtained from the formula (29) for $SU(N)$ after multiplying it by an overall factor $N^{1/2}/(k+N)^{1/2}$, which is the partition function of the $U(1)$ factor. The final result is

$$Z = \frac{1}{(k+N)^{N/2}} \prod_{\alpha>0} 2\sin\left(\frac{\pi(\alpha\cdot\rho)}{k+N}\right). \tag{86}$$

Using the explicit description of the positive roots of $SU(N)$, one gets

$$F = \ln Z = -\frac{N}{2}\ln(k+N) + \sum_{j=1}^{N-1}(N-j)\ln\left[2\sin\frac{\pi j}{k+N}\right]. \tag{87}$$

We can now write the sin as

$$\sin \pi z = \pi z \prod_{n=1}^{\infty}\left(1-\frac{z^2}{n^2}\right), \tag{88}$$

and we find that the free energy is the sum of two pieces. We will call the first one the *nonperturbative* piece:

$$F^{\rm np} = -\frac{N^2}{2}\ln(k+N) + \frac{1}{2}N(N-1)\ln 2\pi + \sum_{j=1}^{N-1}(N-j)\ln j, \tag{89}$$

and the other piece will be called the *perturbative* piece:

$$F^{p} = \sum_{j=1}^{N}(N-j)\sum_{n=1}^{\infty}\ln\left[1-\frac{j^2 g_s^2}{4\pi^2 n^2}\right], \tag{90}$$

where we have denoted

$$g_s = \frac{2\pi}{k+N}, \tag{91}$$

which, as we will see later, coincides with the open string coupling constant under the gauge/string theory duality.

To see that Eq. (89) has a nonperturbative origin, we notice that [see, for example, Ooguri and Vafa (2002)]

$$\mathrm{vol}[U(N)] = \frac{(2\pi)^{(1/2)N(N+1)}}{G_2(N+1)}, \tag{92}$$

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

$$G_2(z+1) = \Gamma(z)G_2(z), \quad G_2(1) = 1. \tag{93}$$

One then finds

$$F^{\rm np} = \ln\frac{(2\pi g_s)^{(1/2)N^2}}{\mathrm{vol}[U(N)]}. \tag{94}$$

This indeed corresponds to the volume of the gauge group in the one-loop contribution (15), where $A^{(c)}$ is in this case the trivial flat connection. Therefore, $F^{\rm np}$ is the logarithim of the prefactor of the path integral, which is not captured by Feynman diagrams.

Let us work out now the perturbative piece (90), following Gopakumar and Vafa (1998a, 1999). By expanding the logarithim, using that $\Sigma_{n=1}^{\infty} n^{-2k} = \zeta(2k)$, and the formula

$$\sum_{j=1}^{N} j^k = \frac{1}{k+1}\sum_{l=1}^{k+1}(-1)^{\delta_{lk}}\binom{k+1}{l}B_{k+1-l}N^l, \tag{95}$$

where B_n are Bernoulli numbers, we find that Eq. (90) can be written as

$$F^{p} = \sum_{g=0}^{\infty}\sum_{h=2}^{\infty} F^{p}_{g,h} g_s^{2g-2+h} N^h, \tag{96}$$

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

$$F^{p}_{0,h} = -\frac{|B_{h-2}|}{(h-2)h!}, \quad h \geqslant 4,$$

$$F^{p}_{1,h} = \frac{1}{12}\frac{|B_h|}{hh!} \tag{97}$$

($F^{p}_{0,h}$ vanishes for $h \leqslant 3$), and for $g \geqslant 2$ one finds

$$F^{\rm p}_{g,h} = \frac{\zeta(2g-2+h)}{(2\pi)^{2g-3+h}}\binom{2g-3+h}{h}\frac{B_{2g}}{2g(2g-2)}. \tag{98}$$

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 easily worked out by expanding the Barnes function (Periwal, 1993; Ooguri and Vafa, 2002). One finds

$$F^{\rm np} = \frac{N^2}{2}\left(\ln(Ng_s) - \frac{3}{2}\right) - \frac{1}{12}\ln N + \zeta'(-1) + \sum_{g=2}^{\infty}\frac{B_{2g}}{2g(2g-2)}N^{2-2g}. \tag{99}$$

C. The string interpretation of the 1/*N* expansion

The expansion (84) of the free energy in a $U(N)$ gauge theory looks very similar to 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

$$F^{p}_{g}(t) = \sum_{h=1}^{\infty} F^{p}_{g,h} t^h. \tag{100}$$

The perturbative free energy can be written as

$$F^{p} = \sum_{g=0}^{\infty} x^{2g-2} F^{p}_{g}(t), \tag{101}$$

which looks like a closed string expansion where t is some modulus of the theory. Note that Eq. (100) con-

tains the contribution of all open Riemann surfaces that appear in the perturbative series with the same bulk topology (specified by the genus g), but with a different number of holes. Therefore, by "summing over all holes" we are "filling up the holes" to produce a closed Riemann surface of genus g. This leads to the 't Hooft idea (1974) that, given a gauge theory, one should be able to find a string theory interpretation in the way we have described, namely, the fat-graph expansion of the free energy is resummed to give a function of the 't Hooft parameter $F_g^p(t)$ at every genus, which is then interpreted as a closed string amplitude. For example, the planar sector of the gauge theory corresponds to a closed string theory at tree level (i.e., at genus $g=0$). Although we are only considering here the perturbative piece of the free energy, we will see that in the Chern-Simons case the nonperturbative piece is crucial to obtain the closed string picture.

Once a closed string intepretation is available, the $1/N$ expansion (85) can be regarded as an open string expansion, where $W_{\vec{k},g}(t)$ are interpreted as amplitudes in an open string theory at genus g and with $h=|\vec{k}|$ holes. According to this interpretation, the Wilson loop creates a one-cycle in the target space where the boundaries of Riemann surfaces can end. The vector $\vec{k}$ specifies the winding numbers for the boundaries as follows: there are k_j boundaries wrapping j times the one-cycle associated to the Wilson loop. The generating functional for connected vacuum expectation values (80) is interpreted as the total free energy of an open string. The open strings that are relevant to the string interpretation of Wilson loop amplitudes should not be confused with the open strings that we associated with the expansion (84). The open strings underlying Eq. (85) should be regarded as an open string sector in the closed string theory associated to the resummed expansion (101).

From the point of view of perturbation theory, the functions $F_g^p(t)$ are rather formal, and the definition (100) expresses them as a power series in t whose coefficients have to be computed order by order in perturbation theory. In some cases the series can be exactly summed up in h and the functions $F_g^p(t)$ can then be obtained in closed form (this is the case, for example, in some matrix models). We will see later that in the case of Chern-Simons theory the $F_g^p(t)$ can be also resummed to give a function of the 't Hooft coupling t.

Of course, the main problem of the 't Hooft program is to identify the closed string theory underlying a gauge theory. This program has been sucessful in some cases, and string theory descriptions have been found for two-dimensional Yang-Mills theory (Gross, 1993; Gross and Taylor, 1993) and for $\mathcal{N}=4$ Yang-Mills theory in four dimensions (Maldacena, 1998). As we will see in this review, Chern-Simons theory also admits a string theory description in terms of topological strings, which we now introduce.

IV. TOPOLOGICAL STRINGS

String theory can be regarded, at the algebraic level, as a two-dimensional conformal field theory coupled to two-dimensional gravity. When conformal field theory is in addition a topological field theory (i.e., a theory whose correlation functions do not depend on the metric on the Riemann surface), the resulting string theory turns out to be very simple and in many cases can be completely solved. A string theory which is constructed in this way is called a *topological string theory*.

The starting point to obtain a topological string theory is therefore a conformal field theory with topological invariance. Such theories are called topological conformal field theories and can be constructed out of $\mathcal{N}=2$ superconformal field theories in two dimensions by a procedure called twisting [see Dijkgraaf *et al.* (1991) for a review of these topics]. For example, one can take the $\mathcal{N}=2$ minimal models to obtain the so-called topological strings in $d<1$. These models are very beautiful and interesting and are deeply related to noncritical string theories. In this review we will consider a more complicated class of topological string theories, where the topological field theory is taken to be a topological sigma model with target space a Calabi-Yau manifold. We will first review the topological sigma model and then explain its coupling to gravity in order to obtain a topological string. We will also introduce some ingredients of toric geometry which are needed to fully understand the class of models that we will consider in this review.

A. Topological sigma models

The topological sigma model was introduced and studied by Witten in a series of papers (Witten, 1988, 1990, 1991a, 1991b) and can be constructed by twisting the $\mathcal{N}=2$ superconformal sigma model in two dimensions [see also Labastida and Llatas (1992)]. A detailed review of topological sigma models and topological strings can be found in Hori *et al.* (2003).

The field content of the topological sigma 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 $i=1,\ldots,2d$. Since we have a complex structure, we will also have holomorphic and antiholomorphic indices, that we will denote, respectively, by $I,\bar{I}$, where $I,\bar{I}=1,\ldots,d$. We also have Grassmann fields $\chi\in x^*(TX)$, which are scalars on Σ_g, and a Grassmannian one-form ψ_α with values in $x^*(TX)$. This last field satisfies a self-duality condition which implies that only its nonzero components are $\psi_{\bar{z}}^I\in x^*(T^{(1,0)}X)$ and $\psi_z^{\bar{I}}\in x^*(T^{(0,1)}X)$, where we have picked up local coordinates z, $\bar{z}$ on Σ_g. The action for the theory is

$$\mathcal{L}=2t\int_{\Sigma_g} d^2z\left(\frac{1}{2}G_{ij}\partial_z x^i\partial_{\bar z}x^j+iG_{I\bar J}\psi^I_{\bar z}D_z\chi^{\bar J}\right.$$
$$\left.+iG_{\bar I J}\psi^{\bar J}_z D_{\bar z}\chi^J-R_{I\bar I J\bar J}\psi^I_{\bar z}\psi^{\bar J}_z\chi^J\chi^{\bar J}\right), \tag{102}$$

where d^2z is the measure $-idz\wedge d\bar z$, t is a parameter that plays the role of $1/\hbar$, and the covariant derivative D_α is given by

$$D_\alpha\chi^i=\partial_\alpha\chi^i+\partial_\alpha x^j\Gamma^i_{jk}\chi^k. \tag{103}$$

The theory also has a BRST, or topological, charge $\mathcal{Q}$ which acts on the fields according to

$$\{\mathcal{Q},x\}=\chi,$$
$$\{\mathcal{Q},\chi\}=0,$$
$$\{\mathcal{Q},\psi^I_{\bar z}\}=i\partial_{\bar z}x^I-\chi^J\Gamma^I_{JK}\psi^K_{\bar z},$$
$$\{\mathcal{Q},\psi^{\bar I}_z\}=i\partial_z x^{\bar I}-\chi^{\bar J}\Gamma^{\bar I}_{\bar J\bar K}\psi^{\bar K}_z. \tag{104}$$

One can show that $\mathcal{Q}^2=0$ on-shell (i.e., modulo the equations of motion). Finally, we also have a $U(1)$ ghost number symmetry, in which x, χ, and ψ have ghost numbers 0, 1, and −1, respectively. Note that the Grassmannian charge $\mathcal{Q}$ then has ghost number 1.

The action (102) turns out to be $\mathcal{Q}$ exact, up to a topological term and terms that vanish on-shell (Witten, 1988, 1991b):

$$\mathcal{L}=-i\{\mathcal{Q},V\}-t\int_{\Sigma_g}x^*(J), \tag{105}$$

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

$$V=t\int_{\Sigma_g}d^2zG_{I\bar J}(\psi^I_{\bar z}\partial_z x^{\bar J}+\partial_{\bar z}x^I\psi^{\bar J}_z), \tag{106}$$

and

$$\int_{\Sigma_g}x^*(J)=i\int_{\Sigma_g}dz\wedge d\bar zG_{I\bar J}(\partial_z x^I\partial_{\bar z}x^{\bar J}-\partial_{\bar z}x^I\partial_z x^{\bar J}). \tag{107}$$

Note that this term in Eq. (105) is 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

$$-it\int_{\Sigma_g}x^*(B), \tag{108}$$

which will replace the Kähler form by the complexified Kähler form $\omega=J+iB$. It is easy to covariantize Eqs. (102) and (106) to introduce an arbitrary metric $g_{\alpha\beta}$ on Σ_g. Since the last term in Eq. (105) is topological, the energy-momentum tensor of this theory is given by

$$T_{\alpha\beta}=\{\mathcal{Q},b_{\alpha\beta}\}, \tag{109}$$

where $b_{\alpha\beta}=\delta V/\delta g^{\alpha\beta}$ and with ghost number −1. The fact that the energy-momentum tensor is $\mathcal{Q}$ 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

$$Z=\int\mathcal{D}\phi e^{-\mathcal{L}}, \tag{110}$$

where ϕ 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

$$\frac{\delta Z}{\delta g^{\alpha\beta}}=-\langle\{\mathcal{Q},b_{\alpha\beta}\}\rangle, \tag{111}$$

where the brackets denote an unnormalized vacuum expectation value. Since $\mathcal{Q}$ is a symmetry of the theory, the above vacuum expectation value vanishes, and we find that Z is metric independent, at least formally.

The $\mathcal{Q}$ 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*. Note 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. The different instanton sectors are classified topologically by the homology class

$$\beta=x_*[(\Sigma_g)]\in H_2(X,\mathbf{Z}). \tag{112}$$

Sometimes it is also useful to introduce a basis $[S_i]$ of $H_2(X,\mathbf{Z})$, where $i=1,\ldots,b_2(X)$, in such a way that we can expand $\beta=\Sigma_i n_i[S_i]$ and the instanton sectors are labeled by $b_2(X)$ integers n_i.

What are the operators to consider in this theory? Since the most interesting aspect of this model is the independence with respect to 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 $\mathcal{Q}$ do the job: topological invariance requires them to be $\mathcal{Q}$ closed, and, on the other hand, they cannot be $\mathcal{Q}$ exact, since otherwise their correlation functions will vanish. One can also check that the $\mathcal{Q}$ cohomology is given by operators of the form

$$\mathcal{O}_\phi=\phi_{i1\cdots ip}\chi^{i1}\cdots\chi^{ip}, \tag{113}$$

where $\phi=\phi_{i1\cdots ip}dx^{i1}\cdots dx^{ip}$ is a closed p-form representing a nontrivial class in $H^p(X)$. Therefore, in this case the $\mathcal{Q}$ cohomology is in one-to-one correspondence with the de Rham cohomology of the target manifold X. Also note 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 $\langle\mathcal{O}_{\phi1}\cdots\mathcal{O}_{\phi l}\rangle$ vanishes unless

$$\sum_{k=1}^{l} \deg(\mathcal{O}_{\phi k}) = 2d(1-g) + 2\int_{\Sigma_g} x^*(c_1(X)), \tag{114}$$

where $\deg(\mathcal{O}_{\phi k}) = \deg(\phi_k)$ and $c_1(X)$ is the first Chern class of the Kähler manifold X. This selection rule corresponds to the fact that the $U(1)$ current is anomalous, and the anomaly is given by the right-hand side of Eq. (114), which calculates the number of zero modes of the Dirac operator (in other words, the right-hand side 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.

In what follows we will focus on Calabi-Yau threefolds, i.e., Kähler manifolds of complex dimension 3, and such that $c_1(X)=0$. For these manifolds the selection rule says that, at genus $g=0$ (i.e., when the Riemann surface is a sphere), the correlation function of three operators associated to two-forms is generically nonvanishing. Since, as we have seen, the semiclassical approximation is exact, the correlation function can be evaluated by counting semiclassical configurations, or, in other words, by summing over worldsheet instantons. In the trivial sector (i.e., when $\beta=0$ and 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 two-forms, while the nontrivial instanton sectors give an infinite series. The final answer looks, schematically, like

$$\langle \mathcal{O}_{\phi 1} \mathcal{O}_{\phi 2} \mathcal{O}_{\phi 3} \rangle = (D_1 \cap D_2 \cap D_3) + \sum_{\beta} I_{0,3,\beta}(\phi_1, \phi_2, \phi_3) Q^{\beta}. \tag{115}$$

The notation is as follows: we have introduced the complexified Kähler parameters

$$t_i = \int_{S_i} \omega, \quad i = 1, \ldots, b_2(X), \tag{116}$$

where ω is the complexified Kähler form of X, and S_i is a basis of $H_2(X)$. We also define $Q_i = e^{-t_i}$, and if $\beta = \Sigma_i n_i [S_i]$, then Q^{β} denotes $\Pi_i Q_i^{n_i}$. The coefficient $I_{0,3,\beta}(\phi_1, \phi_2, \phi_3)$ "counts" in an appropriate sense the number of holomorphic maps from the sphere to the Calabi-Yau manifold that send the point of insertion of $\mathcal{O}_{\phi_i}$ to the divisor D_i. It can be shown that the coefficients $I_{0,3,\beta}(\phi_1, \phi_2, \phi_3)$ can be written as

$$I_{0,3,\beta}(\phi_1, \phi_2, \phi_3) = N_{0,\beta} \int_{\beta} \phi_1 \int_{\beta} \phi_2 \int_{\beta} \phi_3 \tag{117}$$

in terms of invariants $N_{0,\beta}$ that encode all the information about the three-point functions (115) of the topological sigma model. The invariants $N_{0,\beta}$ are our first example of *Gromov-Witten invariants*. It is convenient to put all these invariants together in a generating functional called the *prepotential*:

$$F_0(t) = \sum_{\beta} N_{0,\beta} Q^{\beta}. \tag{118}$$

What happens if we go beyond $g=0$? For $g=1$ and $c_1(X)=0$, the selection rule (114) states that the only quantity that may lead to a nontrivial answer is the partition function itself, while for $g>1$ all correlation functions vanish. 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 consider the topological string theory made out of the topological sigma model, i.e., we have to couple the theory to two-dimensional gravity and to consider all possible metrics on the Riemann surface.

B. Closed topological strings

1. Coupling to gravity

To couple the topological sigma model to gravity, we use the fact pointed out by Dijkgraaf *et al.* (1991), Bershadsky *et al.* (1994), and Witten (1995) that the structure of the twisted theory is tantalizingly close to that of the bosonic string. In the bosonic string, there is a nilpotent BRST operator, $\mathcal{Q}_{\mathrm{BRST}}$, and the energy-momentum tensor turns out to be a $\mathcal{Q}_{\mathrm{BRST}}$ commutator: $T(z) = \{\mathcal{Q}_{\mathrm{BRST}}, b(z)\}$. In addition, there is a ghost number with anomaly $3\chi(\Sigma_g) = 6-6g$, which results that $\mathcal{Q}_{\mathrm{BRST}}$ and $b(z)$ have ghost number 1 and -1, respectively. This is precisely the same structure that we found in Eq. (109), and the composite 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 $g \geqslant 1$ free energy as follows:

$$F_g = \int_{\bar{M}_g} \left\langle \prod_{k=1}^{6g-6} (b, \mu_k) \right\rangle, \tag{119}$$

where

$$(b, \mu_k) = \int_{\Sigma_g} d^2 z [b_{zz} (\mu_k)^z_{\bar{z}} + b_{\bar{z}\bar{z}} (\bar{\mu}_k)^{\bar{z}}_z], \tag{120}$$

and μ_k are the usual Beltrami differentials. The vacuum expectation value in Eq. (119) refers to the path integral over the fields of the topological sigma model, and gives a differential form on the moduli space of Riemann surfaces of genus g, $\bar{M}_g$, which is then integrated over. Note that it is precisely when the target space is a Calabi-Yau threefold that the anomaly (114) 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.

It turns out that the free energies F_g, $g \geqslant 1$, can be also evaluated as a sum over instanton sectors, as in the topological sigma model. Therefore they have the structure

$$F_g(t) = \sum_{\beta} N_{g,\beta} Q^{\beta}, \tag{121}$$

where $N_{g,\beta}$ "count" in an appropriate sense the number of curves of genus g and in the two-homology class β.

We will refer to $N_{g,\beta}$ as the Gromov-Witten invariant of the Calabi-Yau X at genus g and in the class β. They generalize the Gromov-Witten invariants at genus 0 that were introduced in Eq. (117).

2. Mathematical description

The Gromov-Witten invariants that we introduced in Eq. (121) can be defined in a rigorous mathematical way, and have played an important role in algebraic geometry and symplectic geometry. We will now give a short summary of the main mathematical ideas involved in Gromov-Witten theory.

The coupling of the model to gravity involves the moduli space of Riemann surfaces $\bar{M}_g$, as we have just seen. In order to construct the Gromov-Witten invariants in full generality we also need the moduli space of possible metrics (or, equivalently, complex structures) on a Riemann surface with punctures, which is the famous Deligne-Mumford space $\bar{M}_{g,n}$ of n-pointed stable curves [the definition of what stable means can be found, for example, in Harris and Morrison (1998)]. The relevant moduli space in the theory of topological strings $\bar{M}_{g,n}(X,\beta)$ is a generalization of $\bar{M}_{g,n}$, and depends on a choice of a two-homology class β in X. Very roughly, a point in $\bar{M}_{g,n}(X,\beta)$ can be written as $(f,\Sigma_g,p_1,\ldots,p_n)$ and is given by (i) a point in $\bar{M}_{g,n}$, i.e., a Riemann surface with n punctures $(\Sigma_g,p_1,\ldots,p_n)$, together with a choice of complex structure on Σ_g, and (ii) a map $f:\Sigma_g\to X$ which is holomorphic with respect to this choice of complex structure and such that $f_*[\Sigma_g]=\beta$. The set of all such points forms a good moduli space provided a certain number of conditions are satisfied [see, for example, Cox and Katz (1999) and Hori *et al.* (2003) for a detailed discussion of these issues]. $\bar{M}_{g,n}(X,\beta)$ is the basic moduli space we will need in the theory of topological strings. Its complex virtual dimension is given by

$$(1-g)(d-3)+n+\int_{\Sigma_g} f^*(c_1(X)), \tag{122}$$

which is given by the right-hand side of Eq. (114) plus $3g-3+n$, which is the dimension of $\bar{M}_{g,n}$ and takes into account the extra moduli that come from the coupling to two-dimensional gravity. We also have two natural maps,

$$\pi_1:\bar{M}_{g,n}(X,\beta)\to X^n,$$

$$\pi_2:\bar{M}_{g,n}(X,\beta)\to \bar{M}_{g,n}. \tag{123}$$

The first map is easy to define: given a point $(f,\Sigma_g,p_1,\ldots,p_n)$ in $\bar{M}_{g,n}(X,\beta)$, we just compute $(f(p_1),\ldots,f(p_n))$. The second map essentially sends $(f,\Sigma_g,p_1,\ldots,p_n)$ to $(\Sigma_g,p_1,\ldots,p_n)$, i.e., forgets the information about the map and keeps the information about the punctured curve. We can now formally define the Gromov-Witten invariant $I_{g,n,\beta}$ as follows. Let us consider cohomology classes $\phi_1,\ldots,\phi_n$ in $H^*(X)$. If we pull-back their tensor product to $H^*(\bar{M}_{g,n}(X,\beta))$ via π_1, we get a differential form on the moduli space of maps that we can integrate (as long as there is a well-defined fundamental class for this space):

$$I_{g,n,\beta}(\phi_1,\ldots,\phi_n)=\int_{\bar{M}_{g,n}(X,\beta)}\pi_1^*(\phi_1\otimes\cdots\otimes\phi_n). \tag{124}$$

The Gromov-Witten invariant $I_{g,n,\beta}(\phi_1,\ldots,\phi_n)$ vanishes unless the degree of the form equals the dimension of the moduli space. Therefore, we have the following constraint:

$$\frac{1}{2}\sum_{i=1}^{n}\deg(\phi_i)=(1-g)(d-3)+n+\int_{\Sigma_g} f^*(c_1(X)). \tag{125}$$

Note 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 Eq. (115). When $n=0$, one gets an invariant $I_{g,0,\beta}$ which does not require any insertions. This is precisely the Gromov-Witten invariant $N_{g,\beta}$ that appears in Eq. (121). Note that these invariants are in general *rational*, due to the orbifold character of the moduli spaces involved.

By using the Gysin map $\pi_{2!}$, one can reduce any integral of the form (124) to an integral over the moduli space of curves $\bar{M}_{g,n}$. The resulting integrals involve two types of differential forms. The first type of forms are the Mumford classes ψ_i, $i=1,\ldots,n$, which are constructed as follows. We first define the line bundle $\mathcal{L}_i$ over $\bar{M}_{g,n}$ to be the line bundle whose fiber over each curve Σ_g is the cotangent space of Σ_g at p_i (where p_i is the ith marked point). We then have

$$\psi_i=c_1(\mathcal{L}_i),\quad i=1,\ldots,n. \tag{126}$$

The second type of differential forms are the Hodge classes λ_j, $j=1,\ldots,g$, which are defined as follows. On $\bar{M}_{g,n}$ there is a complex vector bundle $\mathbb{E}$ of rank g, called the Hodge bundle, whose fiber at a point Σ_g is $H^0(\Sigma_g,K_{\Sigma_g})$ (i.e., the space of holomorphic sections of the canonical line bundle K_{Σ_g} of Σ_g). The Hodge classes are simply the Chern classes of this bundle,

$$\lambda_j=c_j(\mathbb{E}). \tag{127}$$

The integrals of the ψ classes can be obtained by the results of Witten (1991a) and Kontsevich (1992), while the integrals involving ψ and λ classes (the so-called Hodge integrals) can be in principle computed by reducing them to pure ψ integrals (Faber, 1999). Explicit formulas for some Hodge integrals can be found, for example, in Getzler and Pandharipande (1998) and Faber

and Pandharipande (2000). As we will see, one of the outcomes of the string/gauge correspondence for Chern-Simons theory is an explicit formula for a wide class of Hodge integrals.

3. Integrality properties

The free energies F_g of topological string theory, which contain information about the Gromov-Witten invariants of the Calabi-Yau manifold X, turn out to play an important role in type-IIA string theory: they capture certain couplings in the four-dimensional $\mathcal{N}=2$ supergravity which is obtained when type-IIA theory is compactified on X. For example, the prepotential F_0 encodes information about the effective action for vector multiplets up to two derivatives. As shown by Antoniadis *et al.* (1994) and Bershadsky *et al.* (1994), the higher genus amplitudes F_g with $g\geqslant 1$ can be also interpreted as couplings in the four-dimensional supergravity theory, involving the curvature and the graviphoton field strength.

This connection between topological strings and usual type-II superstrings has been a source of insights for both models, and in particular has indicated a hidden integrality structure in the Gromov-Witten invariants $N_{g,\beta}$. In order to make manifest this structure it is useful to introduce a generating functional for the all-genus free energy:

$$F(g_s,t)=\sum_{g=0}^{\infty} F_g(t) g_s^{2g-2}. \tag{128}$$

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. Gopakumar and Vafa (1998b) showed that the generating functional (128) can be written as a generalized index that counts BPS states in the type-IIA superstring theory compactified on X, and this leads to the following structure result for $F(g_s,t)$:

$$F(g_s,t)=\sum_{g=0}^{\infty}\sum_{\beta}\sum_{d=1}^{\infty} n_\beta^g \frac{1}{d}\left(2\sin\frac{dg_s}{2}\right)^{2g-2} Q^{d\beta}, \tag{129}$$

where n_β^g, known as Gopakumar-Vafa invariants, are *integer* numbers. Therefore, Gromov-Witten invariants of closed strings, which are in general rational, can be written in terms of integer invariants. In fact, by knowing the Gromov-Witten invariants we can explicitly compute the Gopakumar-Vafa invariants from Eq. (129). The Gopakumar-Vafa invariants can also be computed in some circumstances directly in terms of the geometry of embedded curves, and in many cases their computation only involves elementary algebraic geometry (Katz, Klemm, and Vafa, 1999). However, a rigorous mathematical definition of the invariants is not known yet.

There is also a contribution of constant maps to F_g for $g\geqslant 2$ which has not been included in Eq. (129) and is given by $N_{g,0}$. It was shown by Bershadsky *et al.* (1994) [see also Getzler and Pandharipande (1998)] that this contribution can be written as a Hodge integral,

$$N_{g,0}=(-1)^g\frac{\chi(X)}{2}\int_{\bar{M}_g} c_{g-1}^3(\mathbb{E}), \quad g\geqslant 2, \tag{130}$$

where $\chi(X)$ is the Euler characteristic of the Calabi-Yau manifold X. The above integral can be evaluated explicitly to give [Faber and Pandharipande (2000)]

$$N_{g,0}=\frac{(-1)^g\chi(X)|B_{2g}B_{2g-2}|}{4g(2g-2)(2g-2)!}. \tag{131}$$

This can be also deduced from the physical picture of Gopakumar and Vafa (1998b) and from type-IIA/heterotic string duality (Mariño and Moore, 1999).

It is easy to show that the Gopakumar-Vafa formula (129) predicts the following expression for $F_g(t)$:

$$F_g(t)=\sum_{\beta}\left(\frac{|B_{2g}|n_\beta^0}{2g(2g-2)!}+\frac{2(-1)^g n_\beta^2}{(2g-2)!}\pm\cdots-\frac{g-2}{12}n_\beta^{g-1}+n_\beta^g\right)\mathrm{Li}_{3-2g}(Q^\beta), \tag{132}$$

and Li_j is the polylogarithm of index j defined by

$$\mathrm{Li}_j(x)=\sum_{n=1}^{\infty}\frac{x^n}{n^j}. \tag{133}$$

The appearance of the polylogarithm of order $3-2g$ in F_g was first predicted from type-IIA/heterotic string duality by Mariño and Moore (1999).

C. Open topological strings

One can extend many of the previous results to 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 analyzed in detail by Witten (1995). The main issue is of course to specify boundary conditions for the maps $f:\Sigma_{g,h}\to X$. It turns out that the relevant boundary conditions are Dirichlet and are specified by Lagrangian submanifolds of the Calabi-Yau manifold X. A Lagrangian submanifold $\mathcal{L}$ is a cycle where the Kähler form vanishes:

$$J|_{\mathcal{L}}=0. \tag{134}$$

If we denote by C_i, $i=1,\ldots,h$, the boundaries of $\Sigma_{g,h}$ we have to pick a Lagrangian submanifold $\mathcal{L}$, and consider holomorphic maps such that

$$f(C_i)\subset\mathcal{L}. \tag{135}$$

These boundary conditions are a consequence of requiring $\mathcal{Q}$ invariance at the boundary. One also has boundary conditions on the Grassmann fields of the topological sigma model, which require that χ and ψ at the boundary C_i take values on $f^*(T\mathcal{L})$.

We can also couple the theory to Chan-Paton degrees of freedom on the boundaries, giving rise to a $U(N)$ gauge symmetry. The model can then be interpreted as a topological open string theory in the presence of N *topological D-branes* wrapping the Lagrangian submani-

fold $\mathcal{L}$. The Chan-Paton factors give rise to a boundary term in the presence of a gauge connection. If A is a $U(N)$ connection on $\mathcal{L}$, then the path integral has to be modified by inserting

$$\prod_i \mathrm{Tr\ P}\exp\left(\oint_{C_i} f^*(A)\right), \tag{136}$$

where we pullback the connection to C_i through the map f, restricted to the boundary. In contrast to physical D branes in Calabi-Yau manifolds, which wrap special Lagrangian submanifolds (Becker *et al.*, 1995; Ooguri *et al.*, 1996), 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 by summing over topological sectors, similarly to what we did in the closed case. 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 require the boundaries of $f_*[\Sigma_{g,h}]$ to end on $\mathcal{L}$. Therefore, we will set

$$f_*[\Sigma_{g,h}] = \beta \in H_2(X,\mathcal{L}). \tag{137}$$

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

$$f_*[C_i] = w_i\gamma, \quad w_i \in \mathbf{Z}, \quad i=1,\ldots,h, \tag{138}$$

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

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 and the winding numbers, and sum over all bulk classes. This produces the following generating functional of open Gromov-Witten invariants:

$$F_{w,g}(t) = \sum_\beta F_{w,g,\beta} Q^\beta. \tag{139}$$

In this equation, the sum is over relative homology classes $\beta \in H_2(X,\mathcal{L})$, and the notation is as in Eq. (115). The quantities $F_{w,g,\beta}$ are open Gromov-Witten invariants. They "count" in an appropriate sense the number of holomorphically embedded Riemann surfaces of genus g in X with Lagrangian boundary conditions specified by $\mathcal{L}$, and in the class represented by β, w. They are in general rational numbers.

In order to consider all topological sectors, we have to introduce a matrix V which makes it possible to take into account different sets of winding numbers w. The total free energy is defined by

$$F(V) = \sum_{g=0}^\infty \sum_{h=1}^\infty \sum_{w_1,\ldots,w_h} \frac{i^h}{h!} g_s^{2g-2+h} F_{w,g}(t) \mathrm{Tr}\ V^{w_1}\cdots \mathrm{Tr}\ V^{w_h}, \tag{140}$$

where g_s is the string coupling constant. The factor i^h is introduced for convenience, while the factor $h!$ is a symmetry factor which takes into account that the holes are indistinguishable.

In the case of open topological strings one can also write the open Gromov-Witten invariants in terms of a new set of integer invariants that we denote as $n_{w,g,\beta}$. The integrality structure of open Gromov-Witten invariants was derived by Ooguri and Vafa (2000) and by Labastida, Mariño, and Vafa (2000) following arguments similar to those of Gopakumar and Vafa (1998b). According to this structure, the free energy of open topological string theory in the sector labeled by w can be written in terms of the integer invariants $n_{w,g,\beta}$ as follows:

$$\sum_{g=0}^\infty g_s^{2g-2+h} F_{w,g}(t) = \frac{1}{\prod_i w_i} \sum_{g=0}^\infty \sum_\beta \sum_{d|w} (-1)^{h+g} n_{w/d,g,\beta} d^{h-1} \times \left(2\sin\frac{dg_s}{2}\right)^{2g-2} \prod_{i=1}^h \left(2\sin\frac{w_i g_s}{2}\right) Q^{d\beta}. \tag{141}$$

Note there is one such identity for each w. In this expression, the sum is over all integers d which satisfy that $d|w_i$ for all $i=1,\ldots,h$. When this is the case, we define the h-uple w/d whose ith component is w_i/d. The expression (141) can be expanded to give formulas for the different genera. For example, at $g=0$ one simply finds

$$F_{w,g=0,\beta} = (-1)^h \sum_{d|w} d^{h-3} n_{w/d,0,\beta/d}, \tag{142}$$

where the integer d has to divide the vector w (in the sense explained above) and it is understood that $n_{w_d,0,\beta/d}$ is zero if β/d is not a relative homology class. Formulas for higher genera can be easily worked out from Eq. (141), see Mariño and Vafa (2002) for examples.

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

$$h = |\vec{k}|, \quad l = \sum_i w_i = \sum_j j k_j. \tag{143}$$

Note that a given $\vec{k}$ will correspond to many w's which differ by a permutation of their 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

$$F(V)=\sum_{g=0}^{\infty}\sum_{\vec{k}}\frac{i^{|\vec{k}|}}{\prod_j k_j!}g_s^{2g-2+h}F_{\vec{k},g}(t)\Upsilon_{\vec{k}}(V), \tag{144}$$

where $\Upsilon_{\vec{k}}(V)$ was introduced in Eq. (70).

We have considered for simplicity the case in which the boundary conditions are specified by a single Lagrangian submanifold with a single nontrivial one-cycle. In case there are more one-cycles in the geometry, say L, providing possible boundary conditions for the open strings, the above formalism has to be generalized in an obvious way: one needs to specify L sets of winding numbers $w^{(\alpha)}$, and the generating functional (144) depends on L different matrices V_α, $\alpha=1,\ldots,L$. It is useful to write the free energy (144) as

$$F(V)=\sum_R F_R(g_s,t)\mathrm{Tr}_R V \tag{145}$$

by using Frobenius formula (72). The total partition function $Z=e^F$ can then be written as

$$Z(V)=\sum_R Z_R(g_s,t)\mathrm{Tr}_R V \tag{146}$$

by simply expanding Eq. (145) as a formal power series in V. One has, for example, $F_{\square}=Z_{\square}$, $F_{\square\square}=Z_{\square\square}-Z_{\square}^2/2$, and so on. When there are L one-cycles in the target geometry providing boundary conditions, the total partition function has the structure

$$Z(V_i)=\sum_{R_1,\ldots,R_L}Z_{R_1\cdots R_L}(g_s,t)\prod_{\alpha=1}^{L}\mathrm{Tr}_{R_\alpha}V_\alpha. \tag{147}$$

It turns out that the integer invariants $n_{w,g,\beta}$ appearing in Eq. (141) are not the most fundamental ones. As we have seen, if all the winding numbers are positive we can represent w by a vector $\vec{k}=(k_1,k_2,\ldots)$. As we explained in Sec. II.F, such a vector can be interpreted as a label for a conjugacy class $C(\vec{k})$ of the symmetric group S_l, where l is the total winding number. The invariant $n_{w,g,\beta}$ will be denoted in this case as $n_{\vec{k},g,\beta}$. It turns out that this invariant can be written as (Labastida, Mariño, and Vafa, 2000)

$$n_{\vec{k},g,\beta}=\sum_R \chi_R(C(\vec{k}))N_{R,g,\beta,} \tag{148}$$

where $N_{R,g,\beta}$ are integer numbers labeled by representations of the symmetric group, i.e., by Young tableaux, and χ_R is the character of S_l in the representation R. The above relation is invertible, since by orthonormality of the characters one has

$$N_{R,g,\beta}=\sum_{\vec{k}}\frac{\chi_R(C(\vec{k}))}{z_{\vec{k}}}n_{\vec{k},g,\beta}, \tag{149}$$

where $z_{\vec{k}}$ is given in Eq. (73). Note that integrality of $N_{R,g,\beta}$ implies integrality of $n_{\vec{k},g,\beta}$, but not the other way around. In that sense, the integer invariants $N_{R,g,\beta}$ are the most fundamental ones. When there are both positive and negative winding numbers, we can introduce two sets of vectors $\vec{k}^{(1)},\vec{k}^{(2)}$ associated to the positive and negative winding numbers, respectively, and following the same steps we can define BPS invariants $n_{\vec{k}^{(1)},\vec{k}^{(2)},g,\beta}$ and $N_{R_1,R_2,g,\beta}$.

In contrast to conventional Gromov-Witten invariants, a rigorous theory of open Gromov-Witten invariants is not yet available. However, localization techniques make it possible to compute them in some situations (Graber and Zaslow, 2002; Katz and Liu, 2002; Li and Song, 2002; Mayr, 2002).

D. Some toric geometry

So far we have considered topological string theory on general Calabi-Yau threefolds. We will now restrict ourselves to a particular class of geometries, namely, noncompact, toric Calabi-Yau threefolds. These are threefolds that have the structure of a fibration with torus fibers. In particular, the manifolds we will be interested in have the structure of a fibration of $\mathbb{R}^3$ by $\mathbf{T}^2\times\mathbb{R}$. It turns out that the geometry of these threefolds can be packaged in a two-dimensional graph which encodes the information about the degeneration locus of the fibration. We will often call these graphs the *toric diagrams* of the corresponding Calabi-Yau manifolds. Instead of relying on general ideas of toric geometry [which can be found, for example, in Cox and Katz (1999) and in Hori *et al.* (2003)] we will use the approach developed by Leung and Vafa (1998), Aganagic and Vafa (2001), and specially Aganagic, Klemm, Mariño, and Vafa (2003).

1. $\mathbf{C}^3$

In the approach to toric geometry developed by Aganagic, Klemm, Mariño, and Vafa (2003), noncompact toric Calabi-Yau threefolds are constructed out of an elementary building block, namely, $\mathbf{C}^3$. We will now exhibit its structure as a $\mathbf{T}^2\times\mathbb{R}$ fibration over $\mathbb{R}^3$, and 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 or Hamiltonians,

$$\begin{aligned} r_\alpha(z)&=|z_1|^2-|z_3|^2,\\ r_\beta(z)&=|z_2|^2-|z_3|^2,\\ r_\gamma(z)&=\mathrm{Im}(z_1z_2z_3). \end{aligned} \tag{150}$$

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

$$\partial_v z_i=\{r_v,z_i\}_\omega,\quad v=\alpha,\beta,\gamma. \tag{151}$$

This gives the fibration structure that we were looking for: the base of the fibration $\mathbb{R}^3$ is parametrized by the Hamiltonians (150), 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

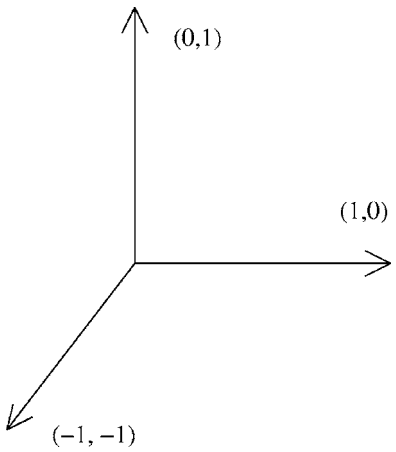

FIG. 10. The degeneration locus graph of the $\mathbf{T}^2\times\mathbb{R}$ fibration of $\mathbf{C}^3$ in the base $\mathbb{R}^3$ parametrized by $(r_\alpha, r_\beta, r_\gamma)$.

$$e^{i\alpha r_\alpha + i\beta r_\beta}:(z_1,z_2,z_3)\to(e^{i\alpha}z_1, e^{i\beta}z_2, e^{-i(\alpha+\beta)}z_3), \qquad (152)$$

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

Note 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 $r_\alpha=r_\gamma=0, r_\beta\geqslant 0$. Similarly, over $z_2=0=z_3$ the (1,0) cycle degenerates over the subspace $r_\beta=r_\gamma=0$ and $r_\alpha\geqslant 0$. Finally, the one-cycle parametrized by $\alpha+\beta$ degenerates over $z_1=0=z_2$, where $r_\alpha-r_\beta=0=r_\gamma$ and $r_\alpha\leqslant 0$.

We will represent the $\mathbf{C}^3$ geometry by a graph which encodes the degeneration 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_α-r_β plane. The degeneration loci will then be straight lines described by the equation $pr_\alpha+qr_\beta=$const. Over this line the $(-q,p)$ cycle of the $\mathbf{T}^2$ 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. 10, drawn in the $r_\gamma=0$ plane.

There is a symmetry in the $\mathbf{C}^3$ geometry that makes it possible to represent it by different toric graphs. These graphs are characterized by three-vectors v_i that are obtained from the ones in Fig. 10 by an $\mathrm{Sl}(2,\mathbf{Z})$ transformation. The vectors have to satisfy

$$\sum_i v_i = 0. \qquad (153)$$

The $\mathrm{Sl}(2,\mathbf{Z})$ symmetry is inherited from the $\mathrm{Sl}(2,\mathbf{Z})$ symmetry of $\mathbf{T}^2$ that appeared in Sec. II.C in a very different context. In the above discussion the generators $H_1(\mathbf{T}^2)$ have been chosen to be the one-cycles associated with r_α and r_β, but there are other choices that differ from this one by an $\mathrm{Sl}(2,\mathbf{Z})$ transformation on $\mathbf{T}^2$. For example, we can choose r_α to generate a (p,q) one-cycle and r_β a (t,s) one-cycle, provided that $ps=qt=1$. These different choices give different trivalent graphs. As we will see in the examples below, the construction of general toric geometries requires in fact these more general graphs representing $\mathbf{C}^3$.

2. More general geometries

The noncompact, toric Calabi-Yau threefolds that we study can be described as symplectic quotients. Let us consider the complex linear space $\mathbf{C}^{N+3}$, described by $N+3$ coordinates $z_1, \ldots, z_{N+3}$, and let us introduce N real equations of the form

$$\mu_A = \sum_{j=1}^{N+3} Q_A^j |z_j|^2 = t_A, \quad A=1,\ldots,N. \qquad (154)$$

In this equation, Q_A^j are integer numbers satisfying

$$\sum_{j=1}^{N+3} Q_A^j = 0. \qquad (155)$$

Furthermore, we consider the action of the group $G_N=U(1)^N$ on the z's where the Ath $U(1)$ acts on z_j by

$$z_j\to\exp(iQ_A^j\alpha_A)z_j.$$

The space defined by Eqs. (154), quotiented by the group action G_N,

$$X=\bigcap_{A=1}^{N}\mu_A^{-1}(t_A)/G_N, \qquad (156)$$

turns out to be a Calabi-Yau manifold [it can be seen that the condition (155) is equivalent to the Calabi-Yau condition]. The N parameters t_A are Kähler moduli of the Calabi-Yau manifold. This mathematical description of X appears in the study of the two-dimensional linear sigma model with $\mathcal{N}=(2,2)$ supersymmetry (Witten, 1993). The theory has $N+3$ chiral fields, whose lowest components are the z's and are charged under N vector multiplets with charges Q_A^j. Equations (154) are the D-term equations, and after dividing by the $\mathrm{U}(1)^N$ gauge group we obtain the Higgs branch of the theory.

The Calabi-Yau manifold X defined in Eq. (156) can be described by $\mathbf{C}^3$ geometries glued together in an appropriate way. Since each of these $\mathbf{C}^3$'s is represented by the trivalent vertex depicted in Fig. 10, we will be able to encode the geometry of Eq. (156) into a trivalent graph. In order to provide this description, we must first find a decomposition of the set of all coordinates $\{z_j\}_{j=1}^{N+3}$ 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 $r_\gamma=\mathrm{Im}(\Pi_{j=1}^{N+3}z_j)$, which is manifestly gauge invariant and, moreover, patch by patch, can be identified with the coordinate used in the $\mathbf{C}^3$ example above. Equation (154) can then be used to find the action of $r_{\alpha,\beta}$ on the other patches.

We will now exemplify this procedure with two important examples: the resolved conifold and the local $\mathbb{P}^2$ geometry.

3. The resolved conifold

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 has a description of the form (156), with $N=1$. There is only one constraint given by

$$|z_1|^2+|z_4|^2-|z_2|^2-|z_3|^2=t \tag{157}$$

and the $U(1)$ group acts as

$$z_1,\ z_2,\ z_3,\ z_4\to e^{i\alpha}z_1,\ e^{-i\alpha}z_2,\ e^{-i\alpha}z_3,\ e^{i\alpha}z_4. \tag{158}$$

Note that, for $z_2=z_3=0$, Eq. (157) describes a $\mathbb{P}^1$ whose area is proportional to t. 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.

Let us now give a description in terms of $\mathbf{C}^3$ patches glued together. The first patch will be defined by $z_4\neq 0$. Using Eq. (157) we can solve for the modulus of z_4 in terms of the other coordinates, and using the $U(1)$ action we can fix its phase. Therefore, the patch will be parametrized by $U_4=(z_1,z_2,z_3)$. The Hamiltonians will be in this case

$$r_\alpha(z)=|z_2|^2-|z_1|^2,$$

$$r_\beta(z)=|z_3|^2-|z_1|^2, \tag{159}$$

which generate the actions

$$e^{i\alpha r_\alpha+i\beta r_\beta}:(z_1,z_2,z_3)\to(e^{-i(\alpha+\beta)}z_1,e^{i\alpha}z_2,e^{i\beta}z_3). \tag{160}$$

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 Eq. (157) to rewrite the Hamiltonians as

$$r_\alpha(z)=|z_4|^2-|z_3|^2-t,$$

$$r_\beta(z)=|z_4|^2-|z_2|^2-t, \tag{161}$$

generating the action

$$e^{i\alpha r_\alpha+i\beta r_\beta}:(z_4,z_2,z_3)\to(e^{i(\alpha+\beta)}z_4,e^{-i\beta}z_2,e^{-i\alpha}z_3). \tag{162}$$

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. 11. Note that the common edge of the graphs represents the $\mathbb{P}^1$ of the resolved conifold: along this edge, one of the $\mathbf{S}^1$'s of $\mathbf{T}^2$ has degenerated, while the other only degenerates at the end points. An $\mathbf{S}^1$ fibration of an interval which degenerates at its end points is nothing but a two-sphere. The length of the edge is t, the Kähler parameter associated to the $\mathbb{P}^1$.

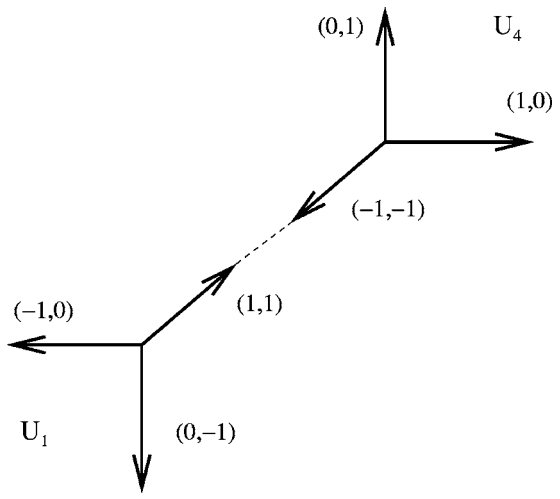

FIG. 11. The resolved conifold $\mathcal{O}(-1)\oplus\mathcal{O}(-1)\to\mathbb{P}^1$ and associated graph. This manifold is made out of two $\mathbf{C}^3$ patches glued through a common edge.

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

Let us now consider a more complicated example, namely, the noncompact Calabi-Yau manifold $\mathcal{O}(-3)\to\mathbb{P}^2$. This is the total space of $\mathbb{P}^2$ together with its anticanonical bundle, and it is often called local $\mathbb{P}^2$. We can describe it again as in Eq. (156) with $N=1$. There are four complex variables, $z_0,\ldots,z_3$, and the constraint (154) now reads

$$|z_1|^2+|z_2|^2+|z_3|^2-3|z_0|^2=t. \tag{163}$$

The $U(1)$ action on the z's is

$$z_0,z_1,z_2,z_3\to e^{-3i\alpha}z_0,e^{i\alpha}z_1,e^{i\alpha}z_2,e^{i\alpha}z_3. \tag{164}$$

Note that $z_{1,2,3}$ describe the basis $\mathbb{P}^2$, while z_0 parametrizes the complex direction of the fiber.

Let us now give a description in terms of glued $\mathbf{C}^3$ patches. 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 nonzero 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 parametrize $\mathbf{C}^3$: $U_3=(z_0,z_1,z_2)$. A similar statement holds for the other two patches. Let us now construct the corresponding degeneration graph. In the $U_3=(z_0,z_1,z_2)$ patch we take as our Hamiltonians

$$r_\alpha=|z_1|^2-|z_0|^2,$$

$$r_\beta=|z_2|^2-|z_0|^2. \tag{165}$$

The graph of the degenerate fibers in the r_α-r_β plane is the same as in the $\mathbf{C}^3$ example, see Fig. 10. The third direction in the base r_γ is now given by the gauge invariant product $r_\gamma=\mathrm{Im}(z_0z_1z_2z_3)$. The same two Hamilto-

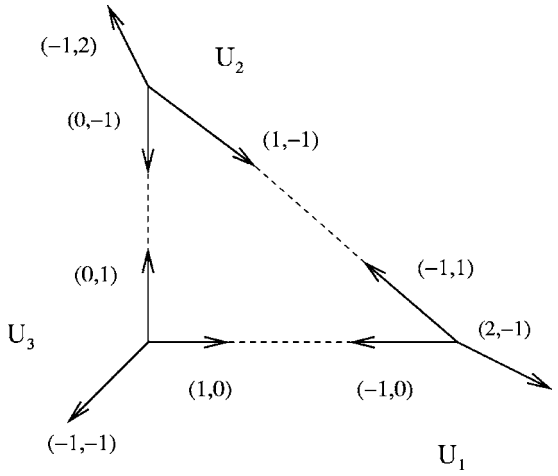

FIG. 12. The $\mathcal{O}(-3)\to\mathbb{P}^2$ graph. This manifold is built out of three $\mathbf{C}^3$ patches.

nians $r_{\alpha,\beta}$ generate the action in the $U_2=(z_0,z_1,z_3)$ patch, and we use the constraint (163) 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

$$r_\alpha=|z_1|^2-|z_0|^2,$$

$$r_\beta=t+2|z_0|^2-|z_1|^2-|z_3|^2, \tag{166}$$

hence

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

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. 12.

5. Lagrangian submanifolds

In order to consider open string amplitudes in the above Calabi-Yau geometries, we have to construct Lagrangian submanifolds providing boundary conditions, as explained in Sec. IV.C. Let us start by considering the $\mathbf{C}^3$ geometry discussed above. In this case, one can easily construct Lagrangian submanifolds following the work of Harvey and Lawson (1982). In terms of the Hamiltonians in Eq. (150), we have three types:

$$\begin{aligned}
&L_1:\quad r_\alpha=0,\quad r_\beta=r_1,\quad r_\gamma\geqslant 0,\\
&L_2:\quad r_\alpha=r_2,\quad r_\beta=0,\quad r_\gamma\geqslant 0,\\
&L_3:\quad r_\alpha=r_\beta=r_3,\quad r_\gamma\geqslant 0,
\end{aligned} \tag{167}$$

where r_i, $i=1,2,3$, are constants. It is not difficult to check that the above submanifolds are indeed Lagrangian (they turn out to be special Lagrangian as well). In terms of the graph description we developed above, they correspond to points in the edges of the planar graph spanned by (r_α,r_β), and they project to semi-infinite straight lines on the basis of the fibration $\mathbf{R}^3$ parametrized by $r_\gamma\geqslant 0$. Since they are located at the edges, where one of the circles of the fibration degenerates, they have the topology of $\mathbf{C}\times\mathbf{S}^1$.

It is easy to generalize the construction to other toric geometries, such as the resolved conifold or local $\mathbb{P}^2$: Lagrangian submanifolds with the topology of $\mathbf{C}\times\mathbf{S}^1$ are given by points on the edges of the planar graphs. Such Lagrangian submanifolds were first considered in the context of open topological string theory by Aganagic and Vafa (2000), and further studied by Aganagic, Klemm, and Vafa (2002).

E. Examples of closed string amplitudes

Gromov-Witten invariants of Calabi-Yau threefolds can be computed in a variety of ways. A powerful technique which can be made mathematically rigorous is the localization technique pioneered by Kontsevich (1995). For compact Calabi-Yau manifolds, only $N_{g=0,\beta}$ have been computed rigorously. For noncompact, toric Calabi-Yau manifolds one can compute $N_{g,\beta}$ for arbitrary genus using these localization techniques. We will now present some results for the topological string amplitudes F_g of the geometries we described above.

The resolved conifold $\mathcal{O}(-1)\oplus\mathcal{O}(-1)\to\mathbb{P}^1$ has one single Kähler parameter t corresponding to the $\mathbb{P}^1$ in the base, and its total free energy is given by

$$F(g_s,t)=\sum_{d=1}^{\infty}\frac{1}{d[2\sin(dg_s/2)]^2}Q^d, \tag{168}$$

where $Q=e^{-t}$. 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. The above closed expression was obtained in Gromov-Witten theory by Faber and Pandharipande (2000).

The space $\mathcal{O}(-3)\to\mathbb{P}^2$ has also one single Kähler parameter, corresponding to the hyperplane class of $\mathbb{P}^2$. By using the localization techniques of Kontsevich, adapted to the noncompact case, one finds (Chiang *et al.*, 1999; Klemm and Zaslow, 2001)

$$F_0(t)=-\frac{t^3}{18}+3\,Q-\frac{45\,Q^2}{8}+\frac{244\,Q^3}{9}-\frac{12\,333\,Q^4}{64}+\cdots,$$

$$F_1(t)=-\frac{t}{12}+\frac{Q}{4}-\frac{3\,Q^2}{8}-\frac{23\,Q^3}{3}+\frac{3437\,Q^4}{16}+\cdots,$$

$$F_2(t) = \frac{\chi(X)}{5720} + \frac{Q}{80} + \frac{3\,Q^3}{20} + \frac{514\,Q^4}{5} + \cdots, \tag{169}$$

and so on. In Eq. (169), t is the Kähler class of the manifold, $Q=e^{-t}$, and $\chi(X)=2$ is the Euler characteristic of the local $\mathbb{P}^2$. The first term in F_0 is proportional to the intersection number H^3 of the hyperplane class, while the first term in F_1 is proportional to the intersection number between H and $c_2(X)$. The first term in F_2 is the contribution of constant maps.

As 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:

$$f(z,Q) = \sum_{g,\beta} n_\beta^g z^g Q^\beta, \tag{170}$$

where z is a formal parameter. For local $\mathbb{P}^2$ we find

$$f(z,Q) = 3\,Q - 6\,Q^2 + (27 - 10\,z)Q^3 - (192 - 231\,z + 102\,z^2 - 15\,z^3)Q^4 + \mathcal{O}(Q^5). \tag{171}$$

It should be mentioned that there is of course a very powerful method to compute the amplitude F_g, namely, mirror symmetry. In the mirror symmetric computation, the F_g amplitudes are closely 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 Bershadsky *et al.* (1993, 1994). Gromov-Witten invariants of noncompact, toric Calabi-Yau threefolds have been computed with mirror symmetry by Chiang *et al.* (1999), Katz, Klemm, and Vafa (1999), and Klemm and Zaslow (2001).

V. CHERN-SIMONS THEORY AS A STRING THEORY

In this section we show that the 't Hooft program to interpret the $1/N$ expansion of a gauge theory in terms of a string theory can be realized in detail in the case of Chern-Simons theory on the three-sphere.

A. 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 of the $1/N$ expansion of the free energy (100). This was done by Witten (1995) in a remarkable paper, 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

$$J = \sum_{a=1}^{3} dp_a \wedge dq_a. \tag{172}$$

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, therefore it is a Calabi-Yau manifold. In fact, $T^*\mathbf{S}^3$ is a well-known Calabi-Yau manifold, namely, the *deformed conifold*. The deformed conifold is usually described by the algebraic equation

$$\sum_{\mu=1}^{4} \eta_\mu^2 = a. \tag{173}$$

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, and let us take a to be real. We find the two equations

$$\sum_{\mu=1}^{4} (x_\mu^2 - v_\mu^2) = a,$$

$$\sum_{\mu=1}^{4} x_\mu v_\mu = 0. \tag{174}$$

The first equation indicates that the locus $v_\mu=0$, $\mu=1,\ldots,4$, 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, Eq. (173) is nothing but $T^*\mathbf{S}^3$.

It is obvious that M is a Lagrangian submanifold in T^*M, since J vanishes along $p_a=$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, and as we shall see this action is nothing but Chern-Simons theory on M. This will prove the sought-for realization of Chern-Simons theory in terms of open strings.

B. Open string field theory

In order to obtain the spacetime description of open strings on T^*M we use string field theory. We briefly summarize here some basic ingredients of the cubic string field theory introduced by Witten (1986) to describe the spacetime dynamics of open bosonic strings, since we use the same model to describe topological strings.

In bosonic open string field theory, we consider the worldsheet of the string to be an infinite strip parametrized by a spatial coordinate $0\leqslant\sigma\leqslant\pi$ and a time coordinate $-\infty<\tau<\infty$, and we pick the flat metric $ds^2=d\sigma^2+d\tau^2$. We then consider maps $x: I\to X$, with $I=[0,\pi]$ and X the target of the string. The string field is a functional of open string configurations $\Psi(x(\sigma))$, with ghost number 1 (although we will not indicate it explic-

itly, this string functional depends as well on the ghost fields). Witten (1986) defines two operations on the space of string functionals. The first one is the *integration*, which is defined formally by folding the string around its midpoint and gluing the two halves:

$$\int \Psi = \int \mathcal{D}x(\sigma) \prod_{0\leq\sigma\leq\pi/2} \delta[(\sigma) - x(\pi-\sigma)]\Psi[x(\sigma)]. \tag{175}$$

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 disk one has to soak up three zero modes. One also defines an associative, noncommutative *star product* $\star$ of string functionals through the following equation:

$$\int \Psi_1 \star \cdots \star \Psi_N = \int \prod_{i=1}^{N} \mathcal{D}x_i(\sigma) \prod_{i=1}^{N} \prod_{0\leq\sigma\leq\pi/2} \delta[x_i(\sigma) - x_{i+1}(\pi-\sigma)]\Psi_i[x_i(\sigma)], \tag{176}$$

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 review of Taylor and Zwiebach (2003) for more details], and it does not change the ghost number. In terms of these geometric operations, the string field action is given by

$$S = \frac{1}{g_s} \int \left(\frac{1}{2} \Psi \star Q_{\mathrm{BRST}} \Psi + \frac{1}{3} \Psi \star \Psi \star \Psi \right). \tag{177}$$

Note that the integrand has ghost number 3, while the integration has ghost number −3, so that the action (177) has ghost number 0. If we add Chan-Paton factors, the string field is promoted to a $U(N)$ matrix of string fields, and the integration in Eq. (177) includes a trace Tr. The action (177) contains all the information on 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 above action (Taylor, 2000).

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 oscillators, the string functional becomes an ordinary function of spacetime, the $\star$ product becomes the usual product of functions, and the integral is the usual integration of functions. The decoupling of 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.

C. 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 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 $\mathcal{Q}_{\mathrm{BRST}}$ exact). Since both theories have a similar structure, the spacetime dynamics of topological D branes in T^*M is governed as well by Eq. (177), where $\mathcal{Q}_{\mathrm{BRST}}$ is given in this case by the topological charge defined in Eq. (104), and where the star product and the integration operation are given by the bosonic string. The construction of the cubic string field theory also requires the existence of a ghost number symmetry, which is also present in the topological sigma model, as discussed in Sec. IV.A. It is convenient to consider the ghost number charge shifted by $-d/2$ with respect to the assignment presented in Sec. IV.A (here, d is the dimension of the target). The shifted ghost number is actually the axial charge of the original $\mathcal{N}=2$ superconformal theory in the Ramond sector. When $d=3$ this corresponds to the normalization used by Witten (1986) in which the ghost vacuum of the bc system is assigned the ghost number $-1/2$.

In order to provide the string field theory description of open topological strings on T^*M, we have to determine the precise content of the string field, the $\star$ algebra, and the integration of string functionals for this particular model. 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 and consider maps $x: I \to T^*M$, with $I=[0,\pi]$, such that ∂I is mapped to M. The Grassmann 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 extracted from the Lagrangian (102):

$$\left[\frac{dx^i}{d\tau}(\sigma), x^j(\sigma') \right] = -\frac{i}{t} G^{ij} \delta(\sigma - \sigma'),$$

$$\{\psi_\tau(\sigma), \chi(\sigma')\} = \frac{1}{t} \delta(\sigma - \sigma'). \tag{178}$$

The Hilbert space is made out of functionals $\Psi[(\sigma), \ldots]$, where x is a map from the interval as we have just described, and the … refer to the Grassmann fields (which represent here the role of ghost fields). The Hamiltonian is obtained, as usual in string theory, by

$$L_0 = \int_0^\pi d\sigma T_{00}. \tag{179}$$

The bosonic piece of T_{00} is just

$$t G_{ij} \left(\frac{dx^i}{d\sigma} \frac{dx^j}{d\sigma} + \frac{dx^i}{d\tau} \frac{dx^j}{d\tau} \right), \tag{180}$$

and using the canonical commutation relations we find

$$L_0 = \int_0^\pi d\sigma \left(-\frac{1}{t} G^{ij} \frac{\delta^2}{\delta x^i(\sigma)\delta x^j(\sigma)} + t G_{ij} \frac{dx^i}{d\sigma}\frac{dx^j}{d\sigma} \right). \quad (181)$$

We then see that string functionals with $dx^i/d\sigma \neq 0$ cannot 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. A similar analysis holds for the Grassmann fields as well, and the conclusion is that the string functionals are functions of the commuting and anticommuting zero modes. Denoting them by q^a, χ^a, the string functional reduces to

$$\Psi = A^{(0)}(q) + \sum_{p=1}^{3} \chi^{a_1} \cdots \chi^{a_p} A^{(p)}_{a_1 \cdots a_p}. \quad (182)$$

These functionals can be interpreted as differential forms on M. A differential form of degree p will have ghost number p, or, equivalently, shifted ghost number 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 $\mathcal{Q}$ symmetry acts as the exterior differential, and $\{\mathcal{Q}, \Psi\} = 0$ if the differential forms are closed. Of course in string field theory we do not restrict ourselves to functionals in the $\mathcal{Q}$ cohomology. We compute rather the string field action for arbitrary functionals, and then the condition of being in the $\mathcal{Q}$ cohomology arises as a linearized equation of motion.

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 (182). Since in string field theory the string field has ghost number 1 [equivalently, $U(1)_R$ charge $-1/2$], we see that

$$\Psi = \chi^a A_a(q), \quad (183)$$

where $A_a(q)$ 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 commuting and anticommuting 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:

$$\Psi \to A, \quad \mathcal{Q}_{\mathrm{BRST}} \to d,$$

$$\star \to \wedge, \quad \int \to \int_M. \quad (184)$$

The string field action (177) is then the usual Chern-Simons action for A, and by comparing with Eq. (1) we have the following relation between the string coupling constant and the Chern-Simons coupling:

$$g_s = \frac{2\pi}{k+N}, \quad (185)$$

after taking into account the shift $k \to k+N$.

This result is certainly remarkable. In the usual 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 relevant 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.

However, as explained by Witten (1995), since open topological string theory is a theory that describes open string instantons with Lagrangian boundary conditions, we should expect to have corrections to the above result due to nontrivial worldsheet instantons. It is easy to see that instantons $x: \Sigma \to T^*M$ such that $x(\partial\Sigma) \subset M$ are necessarily constant. Note first that $J = d\rho$, where

$$\rho = \sum_{a=1}^{3} p_a dq_a, \quad (186)$$

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

$$\int_\Sigma x^*(J) = \int_{\partial\Sigma} x^*(\rho) = 0, \quad (187)$$

since $x(\partial\Sigma) \subset M$. Holomorphic maps with the above boundary conditions are necessarily constant, and there are no worldsheet instantons in the geometry. Therefore, there are no instanton corrections to the Chern-Simons action that we derived above.

One of the immediate consequences of the Chern-Simons spacetime description of open topological strings on T^*M is that the coefficient $F_{g,h}$ in the perturbative expansion (84) of Chern-Simons theory on M is given by the free energy of the topological string theory at genus g and h holes. What is then the interpretation of the fat graph associated to $F_{g,h}$ from the point of view of the topological string theory on T^*M? Even though there are no "honest" worldsheet instantons in this geometry, there are degenerate instantons of zero area in which the Riemann surface degenerates to a graph in M. It is well known that the moduli space of open Riemann surfaces contains this type of configurations. In the case at hand, the fat graphs appearing in the $1/N$ expansion of Chern-Simons theory on M are precisely the graphs that describe the degenerate instantons of the geometry. This model gives then a very concrete realization of the string picture of the $1/N$ expansion discussed in Sec. III.

D. More general Calabi-Yau manifolds

In the previous section we have presented an explicit description of open topological strings on T^*M, following Witten (1995). What happens if the target is a more general Calabi-Yau manifold?

Let us 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 spacetime description of topological open strings will have two contributions. First of all, we have the contributions of degenerate holomorphic curves. These are captured by Chern-Simons theories on the manifolds M_i, following the same mechanism that we described for T^*M. However, as pointed out by Witten (1995), for a general Calabi-Yau manifold X we may also have *honest* open string instantons contributing to the spacetime description, which will be embedded holomorphic Riemann surfaces with boundaries ending on the Lagrangian submanifolds M_i. An open string instanton β will intersect the M_i along one-dimensional curves $\mathcal{K}_i(\beta)$, which are in general knots inside M_i. We know from Eq. (136) that the boundary of such an instanton will give a Wilson loop insertion in the spacetime action of the form $\Pi_i \mathrm{Tr} U_{\mathcal{K}_i(\beta)}$, where $U_{\mathcal{K}_i(\beta)}$ is the holonomy of the Chern-Simons connection on M_i along the knot $\mathcal{K}_i(\beta)$. 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 instantons by including the corresponding Chern-Simons theories $S_{\mathrm{CS}}(A_i)$, which account for the degenerate instantons, coupled in an appropriate way with the "honest" holomorphic instantons. The spacetime action will then have the form

$$S(A_i) = \sum_i S_{\mathrm{CS}}(A_i) + \sum_\beta e^{-\int_\beta \omega} \prod_i \mathrm{Tr} U_{\mathcal{K}_i(\beta)}, \tag{188}$$

where ω is the complexified Kähler form and the second sum is over honest holomorphic instantons β. Notice that all the Chern-Simons theories $S_{\mathrm{CS}}(A_i)$ have the same coupling constant, equal to the string coupling constant. More precisely,

$$\frac{2\pi}{k_i + N_i} = g_s. \tag{189}$$

In the action (188), the honest holomorphic instantons are put in "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 worldsheet instanton ending on a knot $\mathcal{K}$, it will give rise to a Wilson loop operator in Eq. (188), and the $1/N$ evaluation of the vacuum expectation value will generate all possible fat graphs Γ joined to the knot $\mathcal{K}$, producing in this way partially degenerate worldsheet instantons (the fat graphs are interpreted, as before, as degenerate instantons). An example of this situation is depicted in Fig. 13. This more complicated scenario was explored by Aganagic and Vafa (2001), Diaconescu, Florea, and Grassi (2003a, 2003b), and Aganagic, Mariño, and Vafa (2004). We will give examples of Eq. (188) in Sec. VI.

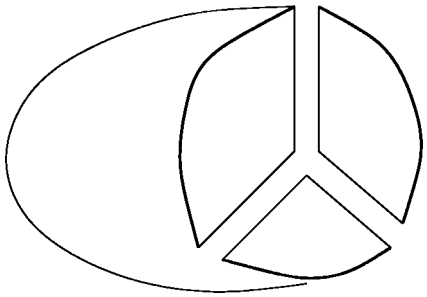

FIG. 13. A partially degenerate worldsheet instanton of genus $g=0$ and with $h=3$ ending on an unknot. The instanton is made out of an honest holomorphic disk and the degenerate piece, which is a fat graph.

E. The conifold transition and the large *N* duality

We know now 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 (101). As shown by Gopakumar and Vafa (1999) in an important paper 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 (Maldacena, 1998), 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 disappear 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, Candelas and de la Ossa (1990)]. Let us explain this in detail.

The algebraic equation describing the deformed conifold is Eq. (173). It is useful to rewrite this equation as follows. Introduce the following complex coordinates:

$$x = \eta_1 + i\eta_2, \quad v = i(\eta_3 - i\eta_4),$$

$$u = i(\eta_3 + i\eta_4), \quad y = \eta_1 - i\eta_2. \tag{190}$$

The deformed conifold can now be written as

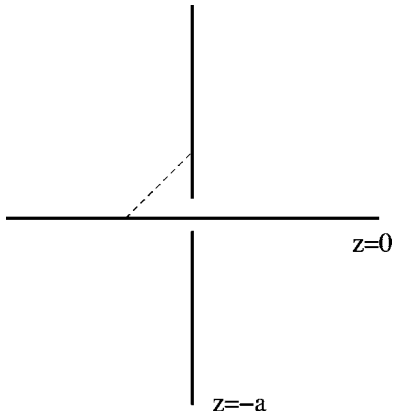

FIG. 14. The deformed conifold $T^*\mathbf{S}^3$, regarded as a $\mathbf{T}^2\times\mathbb{R}$ fibration of $\mathbb{R}^3$. Two of the directions represent the axes of the two cylinders, and the third direction represents the real axis of the z plane. The dashed line represents the $\mathbf{S}^3$ cycle.

$$xy = uv + a. \tag{191}$$

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

$$x,y,u,v \to xe^{-i\alpha}, ye^{i\alpha}, ue^{-i\beta}, ve^{i\beta}, \tag{192}$$

where the α and β actions above can be taken to generate the (0,1) and (1,0) cycles of the $\mathbf{T}^2$, respectively. The $\mathbf{T}^2$ fiber can degenerate to $\mathbf{S}^1$ by collapsing one of its one-cycles. In Eq. (192), for example, the $U(1)_\alpha$ 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 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 $\mathrm{Im}(z)$. The $U(1)_\alpha$ fibration degenerates at $z=-a$, while the $U(1)_\beta$ fibration degenerates at $z=0$. This is the same kind of fibration structure that we found when discussing the geometries of the form (156).

As presented in Sec. IV, 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. This is shown in Fig. 14. Note that the $\mathbf{S}^3$ of the deformed conifold geometry is realized in this picture as a $\mathbf{T}^2$ fibration over an interval I. This interval is represented in Fig. 14 by a dashed line in the z plane between $z=-a$ [where the (0,1) cycle collapses] and $z=0$ [where the collapsing cycle is the (1,0)]. The geometric description of $\mathbf{S}^3$ obtained this way is in fact equivalent to the description given in Sec. II in terms of a Heegaard 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. 15.

The conifold singularity appears when $a=0$ in Eq. (191) and the three-sphere collapses. This is described by

$$xy = uv. \tag{193}$$

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 de-

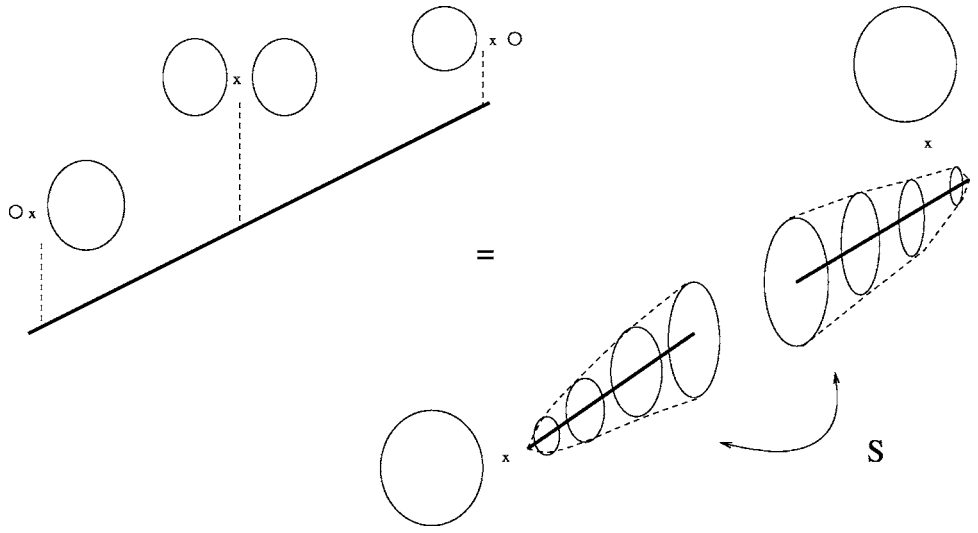

FIG. 15. On the left-hand side, we represent $\mathbf{S}^3$ as a $\mathbf{T}^2$ fibration over the interval. One of the circles of the torus degenerates over one end point, while the other circle degenerates over the opposite end point. Each of the degenerating circles fibers over half the interval to produce a disk 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.

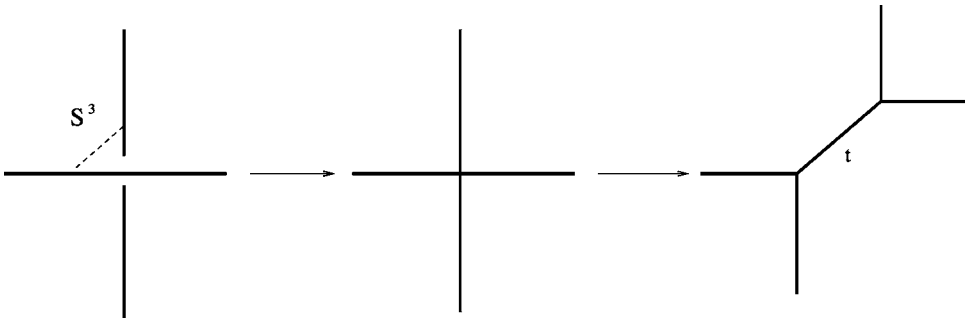

FIG. 16. A pictorial representation of the conifold transition. The deformed conifold (left) with an $\mathbf{S}^3$, the conifold singularity (center), and the resolved conifold (right).

formed conifold (173). 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, Candelas and de la Ossa (1990)]. The resolution of the geometry can be explained as follows. Equation (193) can be solved by

$$x = \lambda v, \quad u = \lambda y, \tag{194}$$

where λ is an inhomogeneous coordinate in $\mathbb{P}^1$. Equation (194) can be interpreted as defining the bundle $\mathcal{O}(-1)\oplus\mathcal{O}(-1)\to\mathbb{P}^1$. To make contact with the toric description given in Eq. (157), we put $x=z_1z_3$, $y=z_2z_4$, $u=z_1z_2$, and $v=z_3z_4$. We then see that $\lambda=z_1/z_4$ is the inhomogenous coordinate for the $\mathbb{P}^1$ described in Eq. (157) by $|z_1|^2+|z_4|^2=t$. 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 t grows giving the resolved conifold. In terms of the coordinates $z_1,\ldots,z_4$, the $\mathbf{T}^2$ action (192) becomes

$$z_1,z_2,z_3,z_4 \to e^{-i(\alpha+\beta)}z_1, e^{i\alpha}z_2, e^{i\beta}z_3, z_4. \tag{195}$$

This $\mathbf{T}^2$ fibration is precisely Eq. (160). Note that the singular loci of fibration of the resolved conifold which is encoded in the trivalent graph of Fig. 11 is inherited from the singular loci depicted in Fig. 14. The transition from the deformed to the resolved conifold can then be represented pictorially as in Fig. 16.

Now we are ready to state the conjecture of Gopakumar and Vafa (1999). 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. However, in the absence of D-branes that enforce boundary conditions, we have a theory of closed topological strings only. Therefore, *Chern-Simons theory on* $\mathbf{S}^3$ *is equivalent to closed topological string theory on the resolved conifold*. As we shall 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 Eq. (91). The size of the $\mathbb{P}^1$ in the resolved geometry is given by the 't Hooft coupling of Chern-Simons theory,

$$t = ig_sN = xN. \tag{196}$$

This conjecture has been proved by embedding the duality in type-II superstring theory (Vafa, 2001a) and lifting it to M theory (Acharya, 2000; Atiyah, Maldacena, and Vafa, 2001). There is also a worldsheet derivation of the duality due to Ooguri and Vafa (2002). In the next subsection, we will give evidence for the conjecture at the level of the free energy.

F. 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 three-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 Eq. (100). The 't Hooft parameter t is given by Eq. (196) and

$$F_g^p(t) = \sum_{h=1}^{\infty} F_{g,h}^p(-it)^h. \tag{197}$$

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

$$\frac{1}{(1-z)^q} = \sum_{n=0}^{\infty}\binom{q+n-1}{n}z^n \tag{198}$$

to obtain

$$F_g^p(t) = \frac{(-1)^g|B_{2g}B_{2g-2}|}{2g(2g-2)(2g-2)!} + \frac{B_{2g}}{2g(2g-2)}\sum_{n\in\mathbf{Z}}{}'\frac{1}{(-it+2\pi n)^{2g-2}}, \tag{199}$$

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

$$F^{\rm np} = \sum_{g=0}^{\infty} F_g^{\rm np}(t) g_s^{2g-2}, \tag{200}$$

then for $g \geqslant 2$, $F_g^{\rm np}(t) = B_{2g}/2g(2g-2)(-it)^{2g-2}$, which is precisely the $n=0$ term missing in Eq. (199). We then define

$$F_g(t) = F_g^p(t) + F_g^{\rm np}(t). \tag{201}$$

Finally, since

$$\sum_{n\in\mathbf{Z}} \frac{1}{n+z} = \frac{2\pi i}{1-e^{-2\pi i z}}, \tag{202}$$

by taking derivatives with respect to z we can write

$$F_g(t) = \frac{(-1)^g |B_{2g} B_{2g-2}|}{2g(2g-2)(2g-2)!} + \frac{|B_{2g}|}{2g(2g-2)!} \mathrm{Li}_{3-2g}(e^{-t}), \tag{203}$$

again for $g \geqslant 2$. If we now compare to Eq. (132), we see that Eq. (203) 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 constant maps, we find that $\chi(X) = 2$. In fact, Eq. (203) is precisely the F_g amplitude of the resolved conifold. This is a remarkable check of the conjecture.

G. Incorporating Wilson loops

As we have extensively discussed, most of the wealth of Chern-Simons theory on $\mathbf{S}^3$ is due to the Wilson loop operators along knots. How do we incorporate Wilson loops in the string picture that we have just developed? In Sec. III.C we saw that, 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 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 (2000). In order to give boundary conditions for the open strings in the resolved conifold, Ooguri and Vafa constructed a natural Lagrangian submanifold $\tilde{\mathcal{C}}_{\mathcal{K}}$ 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 a knot $\mathcal{K}$ is parametrized by a curve $q(s)$, where $s \in [0, 2\pi)$. The conormal bundle of $\mathcal{K}$ is the space

$$\tilde{\mathcal{C}}_{\mathcal{K}} = \left\{ (q(s), p) \in T^*\mathbf{S}^3 \Big| \sum_a p_a \dot{q}_a = 0,\ 0 \leqslant s \leqslant 2\pi \right\}, \tag{204}$$

where p_a are coordinates for the cotangent bundle, and $\dot{q}_a$ denote the derivatives with respect to s. This space is an $\mathbb{R}^2$ fibration of the knot itself, where the fiber on the point $q(s)$ is given by the two-dimensional subspace of $T^*_{q}\mathbf{S}^3$ of planes orthogonal to $\dot{q}(s)$. $\tilde{\mathcal{C}}_{\mathcal{K}}$ 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 $\tilde{\mathcal{C}}_{\mathcal{K}}$. 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 $\tilde{\mathcal{C}}_{\mathcal{K}}$. But there is a new sector due to strings stretched between the N-branes and the M-branes. To study these strings, we can make an analysis similar to the one we did in Sec. V.C above. First, we have to impose again that $dx^i/d\sigma = 0$. Therefore, x^i has to be a constant, and since the end points of the string lie on different submanifolds, the only possibility is that $x^i \in \mathbf{S}^3 \cap \tilde{\mathcal{C}}_{\mathcal{K}} = \mathcal{K}$. A similar analysis holds for the Grassmann fields, and we then find that the string functionals describing the new sector of strings are a function of a single commuting zero mode q parametrizing $\mathcal{K}$, and a single anticommuting zero mode χ. In other words,

$$\mathcal{A} = \phi(q) + \chi \xi(q), \tag{205}$$

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

As explained above, the string field is the piece of the above functional with $U(1)_R$ charge $-1/2$. However, the $U(1)_R$ charge assignment is now 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 $d = 1$. 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 $\mathcal{Q} = d$, the kinetic term for the string field action (177) is only nontrivial 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 end points of the strings. These are the $U(N)$ gauge connection A on $\mathbf{S}^3$, and the $U(M)$ gauge connection $\tilde{A}$ on $\tilde{\mathcal{C}}_{\mathcal{K}}$. The complex scalar couples to the gauge fields in the standard way,

$$\oint_{\mathcal{K}} \mathrm{Tr}(\bar{\phi} A \phi - \phi \tilde{A} \bar{\phi}). \tag{206}$$

Here we regard $\tilde{A}$ as a source. If we now integrate out ϕ, we obtain

$$\exp\left[-\ln\det\left(\frac{d}{ds}+\sum_a (A_a-\tilde{A}_a)\frac{dq_a}{ds}\right)\right], \tag{207}$$

which can be easily evaluated as

$$\exp[-\mathrm{Tr}\ln(1-U\otimes V)]=\exp\left\{\sum_{n=1}^{\infty}\frac{1}{n}\mathrm{Tr}U^n\mathrm{Tr}V^n\right\}, \tag{208}$$

where U, V^{-1} are the holonomies of $A, \tilde{A}$ around the knot $\mathcal{K}$, and we have dropped an overall constant. In this way we obtain the effective action for the A field,

$$S_{\mathrm{CS}}(A)+\sum_{n=1}^{\infty}\frac{1}{n}\mathrm{Tr}U^n\mathrm{Tr}V^n, \tag{209}$$

where $S_{\mathrm{CS}}(A)$ is of course the Chern-Simons action for A. Therefore, in the presence of the probe branes, the action involves an insertion of the Ooguri-Vafa operator that was introduced in Eq. (77). Since we are regarding the M-branes as a probe, the holonomy V is an arbitrary source. The extra piece in Eq. (209) can be interpreted as coming from an annulus of zero length interpolating between the two sets of D-branes. Later on we will consider a simple generalization of the above for an annulus of finite length.

Let us now follow this system through the geometric transition. The N-branes disappear, 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 $\tilde{\mathcal{C}}_{\mathcal{K}}$ through the geometric transition. The final outcome is the existence of a map between knots in $\mathbf{S}^3$ and Lagrangian submanifolds in $\mathcal{O}(-1)\oplus\mathcal{O}(-1)\to\mathbb{P}^1$ which sends

$$\mathcal{K}\to\mathcal{C}_{\mathcal{K}}. \tag{210}$$

Moreover, one has $b_1(\mathcal{C}_{\mathcal{K}})=1$. This conjecture is clearly well motivated in the physics. Ooguri and Vafa (2000) constructed $\mathcal{C}_{\mathcal{K}}$ explicitly when $\mathcal{K}$ is the unknot, and Labastida, Mariño, and Vafa (2000) proposed Lagrangian submanifolds for certain algebraic knots and links (including torus knots). Finally, Taubes (2001) has constructed a map from knots to Lagrangian submanifolds in the resolved conifold for a wide class of knots.

The Lagrangian submanifold $\mathcal{C}_{\mathcal{K}}$ 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 Ooguri and Vafa (2000), the free energy of open topological strings (140) with boundary conditions specified by $\mathcal{C}_{\mathcal{K}}$ is identical to the free energy of the deformed Chern-Simons theory with action (209), which is nothing but Eq. (80):

$$F(V)=F_{\mathrm{CS}}(V). \tag{211}$$

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 vacuum expectation values, and one finds

$$i^{|\vec{k}|}\sum_{g=0}^{\infty}F_{g,\vec{k}}(t)g_s^{2g-2+|\vec{k}|}=-\frac{1}{\prod_j j^{k_j}}W_{\vec{k}}^{(c)}. \tag{212}$$

It is further assumed that there is an analytic continuation of $F(V)$ from negative to positive winding numbers in such a way that the equality (211) holds in general. Another useful way to state the correspondence (211) is to use the total partition function of topological open strings (147) instead of the free energy. The duality between open string amplitudes and Wilson loop expectation values reads simply

$$Z_R=W_R, \tag{213}$$

where Z_R was introduced in Eq. (147) and W_R is the knot invariant in representation R.

When $\mathcal{K}$ is the unknot in the three-sphere, the conjecture of Ooguri and Vafa can be tested in full detail [Ooguri and Vafa (2000); Mariño and Vafa (2002)]. For more general knots and links, the open string free energy is not known, but one can test the duality indirectly by verifying that the Chern-Simons side satisfies the structural properties of open string amplitudes explained at the end of Sec. IV.C (Labastida, Mariño, and Vafa, 2000; Labastida and Mariño, 2001; Ramadevi and Sharkar, 2001; Labastida and Mariño, 2002; Mariño, 2002b).

VI. STRING AMPLITUDES AND CHERN-SIMONS THEORY

The duality between Chern-Simons theory on $\mathbf{S}^3$ and closed topological strings on the resolved conifold gives a very nice realization of the gauge/string theory duality. 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, similar to the ones discussed in Sec. IV, 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.

The program of extending the geometric transition of Gopakumar and Vafa was started by Aganagic and Vafa (2001). Their basic idea was to construct geometries that locally contain $T^*\mathbf{S}^3$'s, and then follow the geometric transitions to dual geometries where the deformed conifolds are replaced by resolved conifolds. Remarkably, a large class of noncompact toric manifolds can be realized in this way, as it was made clear by Diaconescu, Florea, and Grassi (2003b) and Aganagic, Mariño, and Vafa (2004). In this section we will present some examples where closed string amplitudes can be computed using this idea.

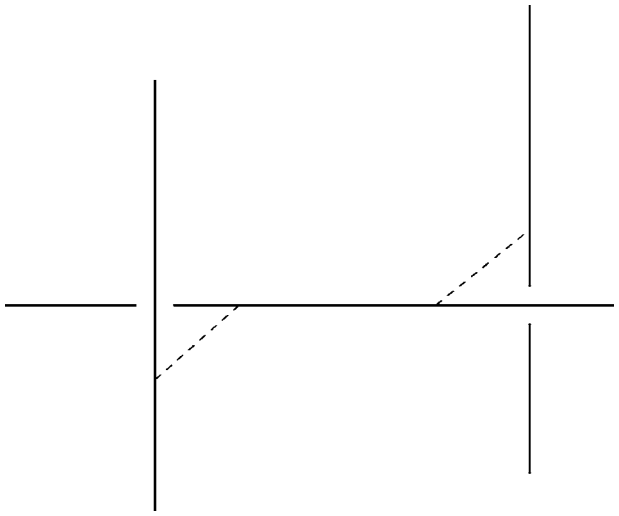

FIG. 17. A Calabi-Yau manifold which is a $\mathbf{T}^2\times\mathbb{R}$ fibration of $\mathbb{R}^3$. The dashed lines represent $\mathbf{S}^3$'s.

A. Geometric transitions for toric manifolds

The geometries discussed in Sec. IV are $\mathbf{T}^2\times\mathbb{R}$ fibrations of $\mathbb{R}^3$ which contain two-spheres (represented by the compact edges of the geometry). In this section we will construct geometries with the same fibration structure that contains three-spheres, and can be related by a geometric transition to some of the toric geometries analyzed in Sec. IV.

Recall from the discussion in Sec. V that the deformed conifold has the structure of a $\mathbf{T}^2\times\mathbb{R}$ fibration of $\mathbb{R}^3$ which can be encoded in a nonplanar graph as in Fig. 14. The degeneration loci of the cycles of the torus fiber are represented in this graph by straight lines, while the $\mathbf{S}^3$ is represented by a dashed line in between these loci. This graphical procedure can be generalized, and it is easy to construct more general $\mathbf{T}^2\times\mathbb{R}$ fibrations of $\mathbb{R}^3$ by specifying degeneration loci in a diagram that represents the $\mathbf{R}^3$ basis. A simple example is shown in Fig. 17. This geometry contains two $\mathbf{S}^3$'s, represented by dashed lines. These three-spheres are also constructed as torus fibrations over the interval, and the cycles that degenerate at the end points 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 end point, and the (1,0) cycle collapses at the other end point. As explained in Sec. V.E, this gives a Heegaard splitting of the three-sphere along solid tori. These tori are glued together through the S transformation that relates one of the collapsing cycles to the other.

We can also construct geometries which contain more general three-manifolds. If a manifold M admits a Heegaard splitting along two solid tori, it will be specified by an $\mathrm{Sl}(2,\mathbf{Z})$ matrix V_M mapping the (p_L,q_L) cycle of one $\mathbf{T}^2$ to the (p_R,q_R) cycle of the other $\mathbf{T}^2$. Equivalently, M can be obtained as a torus fibration over an interval where the (p_L,q_L) and (p_R,q_R) cycles degenerate at the end points, as explained in Sec. V.E in the simple case of the (1,0) and (0,1) cycles. The local geometry T^*M can be described by two overlapping lines with slopes $-p_L/q_L$ and $-p_R/q_R$. The dashed line in between them represents the three-manifold M.

Given a graph similar to the one in Fig. 17, one can try to use the conifold transition "locally," as first explained by Aganagic and Vafa (2001). The above geometry, for example, contains two deformed conifolds 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. 18. The resolved geometry is clearly toric, and it can be easily built up by gluing four trivalent vertices, as explained in Sec. IV.D. It has two Kähler classes corresponding to the two blown-up two-spheres, and denoted by t_1,t_2 in Fig. 18. It also contains a third two-sphere associated to the intermediate, horizontal leg, with Kähler parameter t.

Although we have focused on the example depicted in Fig. 17, it is clear what the general philosophy is: one considers a "deformed" geometry and performs geometric transitions "locally." The resulting "resolved" geometry will be a toric Calabi-Yau manifold of the type discussed in Sec. IV. The planar graph describing the resolved geometry can be easily reconstructed from the nonplanar graph describing the deformed geometry.

B. Closed string amplitudes and geometric transitions

We will now use the generalized geometric transition found in the last subsection in order to compute the

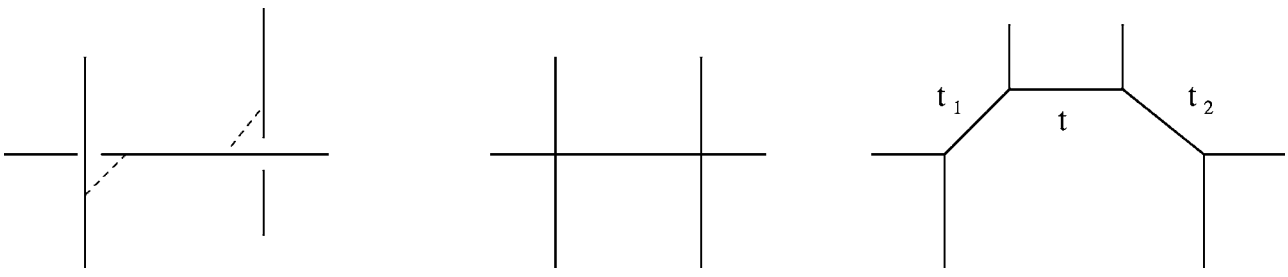

FIG. 18. The geometric transition of the Calabi-Yau manifold depicted in Fig. 17. In the leftmost geometry there are two three-spheres, represented by dashed lines. The intermediate geometry is singular, and the right shows the planar graph associated to the smooth toric Calabi-Yau manifold after the transition. It contains three $\mathbb{P}^1$'s with Kähler parameters t_1, t_2, and t.

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. 17. 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 Sec. V.D: 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: these are the nondegenerate instantons discussed in Sec. V.D following Witten (1995).

Instead of describing these open strings in geometric terms, it is better to use the spacetime physics associated with these strings. A similar situation was considered when we analyzed the incorporation of Wilson loops in the geometric transition. 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. In the situation depicted in Fig. 17, the same arguments indicate that there is a complex scalar ϕ in the representation $(N_1,\bar{N}_2)$, 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 denoted by r. The kinetic term for the complex scalar will be given by

$$\oint_{\mathbf{S}^1} \bar{\phi}(d+A_1-A_2-r)\phi. \tag{214}$$

We can now integrate out this complex scalar field as in Eq. (208) to obtain the correction to the Chern-Simons actions on the three-spheres due to the presence of the new sector of open strings:

$$\mathcal{O}(U_1,U_2;r)=\exp\left\{\sum_{n=1}^{\infty}\frac{e^{-nr}}{n}\mathrm{Tr}\,U_1^n\mathrm{Tr}\,U_2^n\right\}, \tag{215}$$

where $U_{1,2}$ are the holonomies of the corresponding gauge fields around the $\mathbf{S}^1$ in Eq. (214). The operator $\mathcal{O}$ can be also interpreted as the amplitude for a primitive annulus of area r together with its multicovers, which are labeled by n. This annulus "connects" the two $\mathbf{S}^3$'s, i.e., one of its boundaries is a circle in one three-sphere, and the other boundary is a circle in the other sphere. The sum over n in the exponent of Eq. (215) is precisely the sum over open string instantons in the second term of Eq. (188), for this particular geometry.

The problem now is to determine how many configurations similar to 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 (2003a, 2003b) using localization arguments, and derived by Aganagic, Mariño, and Vafa (2004) 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 stretching along the fixed lines of the $\mathbf{T}^2$ action, together with their multicoverings, and the $\mathbf{S}^1$ in Eq. (214) is the circle that fibers over the edge connecting M_1 and M_2. This is illustrated in Fig. 19. The action describing the dynamics of topological D-branes in this example is then

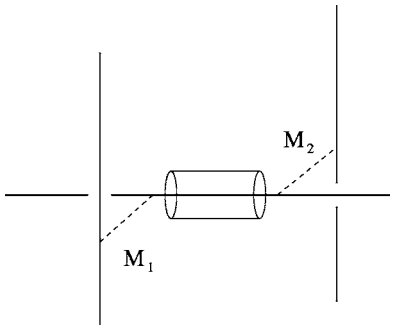

FIG. 19. The only nondegenerate instantons contributing to the geometry of Fig. 17 come from an annulus stretching along the degeneracy locus.

$$S=S_{\mathrm{CS}}(A_1)+S_{\mathrm{CS}}(A_2)+\sum_{n=1}^{\infty}\frac{e^{-nr}}{n}\mathrm{Tr}\,U_1^n\mathrm{Tr}\,U_2^n, \tag{216}$$

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 the $\mathbf{S}^1$. There is a very convenient way to write the free energy of the theory with the above action. First note that, by following the same steps that led to Eq. (79), one can write the operator (208) as

$$\mathcal{O}(U_1,U_2;r)=\sum_R \mathrm{Tr}_R U_1 e^{-\ell r}\mathrm{Tr}_R U_2, \tag{217}$$

where ℓ denotes the number of boxes of the representation R. In the situation depicted in Fig. 19, we see that the boundaries of the annulus give a knot in M_1, and another knot in M_2. Therefore, the total free energy can be written as

$$F=F_{\mathrm{CS}}(N_1,g_s)+F_{\mathrm{CS}}(N_2,g_s)+\ln\sum_R e^{-\ell r}W_R(\mathcal{K}_1)W_R(\mathcal{K}_2), \tag{218}$$

where $F_{\mathrm{CS}}(N_i,g_s)$ 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 two-spheres. Of course, in order to compute Eq. (218) we need some additional information: we have to know what knots $\mathcal{K}_i$ are topologically, 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 cutting the geometry into pieces. The geometry of the knots is then encoded in the geometry of the degeneracy locus.

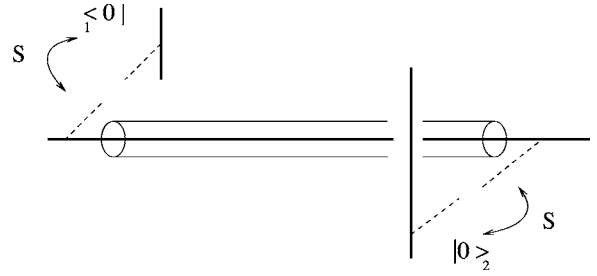

FIG. 20. The geometry of Fig. 17 can be cut into three pieces. The piece that contains the annulus gives by canonical quantization the state (219).

The evaluation proceeds as follows: we cut the geometry into three pieces, as indicated in Fig. 20, by Heegaard 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:

$$\mathcal{O}=\sum_R |R\rangle_1 e^{-\ell r}{}_2\langle R| \in \mathcal{H}_1(\mathbf{T}^2)\otimes\mathcal{H}_2^*(\mathbf{T}^2), \tag{219}$$

where $|R\rangle$ is the Chern-Simons state constructed in Sec. II, and 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 ${}_1\langle 0|S\mathcal{O}S|0\rangle_2$, so we find

$$Z(g_s,N_{1,2},r)=\sum_R {}_1\langle 0|S|R\rangle_1 e^{-\ell r}{}_2\langle R|S|0\rangle_2. \tag{220}$$

Comparing to Eq. (218), we see that

$$W_R(\mathcal{K}_i)=\frac{S_{0R}}{S_{00}}(g_s,t_i),\quad i=1,2, \tag{221}$$

where g_s is the open string coupling constant $2\pi/(k_i+N_i)$ [which is the same for the two Chern-Simons theories, see Eq. (189)] and $t_i=g_sN_i$ are the 't Hooft parameters of the $U(N_i)$ Chern-Simons theories. This means that $\mathcal{K}_{1,2}$ are unknots in the three-spheres $M_{1,2}$, respectively. Geometrically, each of the boundaries of the annulus in Fig. 19 creates a Wilson line along the noncontractible cycle of the solid torus along which we split the three-sphere.

What happens now if we go through the geometric transition of Fig. 18? As in the case originally studied by Gopakumar and Vafa, the 't Hooft parameters become the Kähler parameters t_1,t_2 in the toric diagram of Fig. 18. There is a third Kähler parameter t in the toric geometry after the transition. It turns out that this parameter is related to the parameter r appearing in Eq. (220) as follows:

$$t=r-\frac{t_1+t_2}{2}. \tag{222}$$

This relation was first suggested by Diaconescu, Florea, and Grassi (2003a). It is clearly needed in order to obtain a free energy of the expected form, with a well-defined limit as $t_{1,2}\to\infty$. The total free energy of the resulting toric manifold can be obtained from Eqs. (218) and (221), and it can be written in closed form as

$$F=\sum_{d=1}^{\infty}\frac{1}{d[2\sin(dg_2/2)]^2}\{e^{-dt_1}+e^{-dt_2}+e^{-dt}(1-e^{-dt_1})(1-e^{-dt_2})\}. \tag{223}$$

From this expression we can extract the Gopakumar-Vafa invariants of the toric manifold. Note that Eq. (223) gives the free energy of closed topological strings at all genera. In other words, the nonperturbative solution of Chern-Simons theory [which allows us to compute Eq. (221) exactly] gives us the nonperturbative answer for the topological string amplitude. This is one of the most important aspects of this approach to topological string theory.

One can consider other noncompact Calabi-Yau manifolds and obtain different closed and open string amplitudes by using these generalized geometric transitions (Diaconescu, Florea, and Grassi, 2003b; Aganagic, Mariño, and Vafa, 2004). However, this procedure becomes cumbersome, since in some cases one has to take appropriate limits of the amplitudes in order to reproduce the sought-for answers. The underlying problem of this approach is that we are taking as our basic building block for the resolved geometries the tetravalent vertex that corresponds to the resolved conifold. It is clear, however, that the true building block is the trivalent vertex shown in Fig. 10, which corresponds to $\mathbf{C}^3$. In the next section, we will see how one can define an amplitude associated to this trivalent vertex that allows one to recover any open or closed topological string amplitude for noncompact, toric geometries.

VII. THE TOPOLOGICAL VERTEX

A. Framing of topological open string amplitudes

Since the topological vertex is an open string amplitude, we have to discuss one aspect not yet addressed: the framing ambiguity of topological open string amplitudes. The framing ambiguity was discovered by Aganagic, Klemm, and Vafa (2002). They realized that when the boundary conditions are specified by noncompact Lagrangian submanifolds similar to the ones described in Eq. (167), the corresponding topological open string amplitudes are not univocally defined: they depend on a choice of an integer number (more precisely, one integer number for each boundary). For the Lagrangian submanifolds studied in Sec. IV.D, the framing ambiguity can be specified by modifying the geometry in an appropriate way. These Lagrangian submanifolds simply correspond to points in the edges of the trivalent graphs.

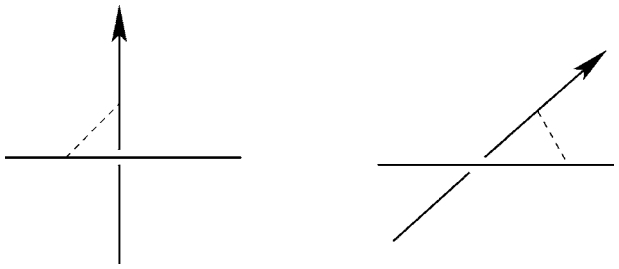

FIG. 21. Two different ways of compactifying a Lagrangian submanifold located at the horizontal edge are shown. These are specified by additional lines (drawn as arrows) where the torus fibration degenerates. The dashed lines represent the compactified submanifolds, which now have the topology of $\mathbf{S}^3$.

Their geometry can be modified by introducing additional locations in the base $\mathbf{R}^3$ where the $\mathbf{T}^2$ fiber degenerates, as we did before when we considered general deformed geometries. In this way the Lagrangian submanifolds become compact $\mathbf{S}^3$ cycles in the geometry, exactly as in Fig. 14. The additional lines are labeled by a vector $f=(p,q)$ where the $(-q,p)$ cycle degenerates. This procedure is illustrated in Fig. 21. It is useful to introduce the symplectic product of two vectors $v=(v_1,v_2)$ and $w=(w_1,w_2)$ as

$$v\wedge w = v_1 w_2 - v_2 w_1. \tag{224}$$

This product is invariant under $\mathrm{Sl}(2,\mathbf{Z})$ transformations. If the original Lagrangian submanifold is located at an edge v, the condition for the compactified cycle to be a nondegenerate $\mathbf{S}^3$ is

$$f\wedge v = 1. \tag{225}$$

Clearly, if f satisfies Eq. (225), so does $f-nv$ for any integer n. The choice of the integer n is precisely the framing ambiguity found by Aganagic, Klemm, and Vafa (2002). In the case of the Lagrangian submanifolds of $\mathbf{C}^3$ constructed in Sec. IV.D, a particular choice of compactification (therefore, of framing) that will be very important in the following is shown in Fig. 22.

What is the effect of a change of framing on open topological string amplitudes? A proposal for this was made by Aganagic, Klemm, and Vafa (2002) and further studied by Mariño and Vafa (2002), based on the duality with Chern-Simons theory. As explained in Sec. V.G, vacuum expectation values of Wilson loops in Chern-Simons theory on $\mathbf{S}^3$ compute open string amplitudes, as stated in Eq. (213). On the other hand, we explained in Sec. II.D that Wilson loop correlation functions depend on a choice of framing. This indicates that the framing ambiguity of Chern-Simons theory corresponds to the ambiguity of topological open string amplitudes that we have just described. This correspondence also suggests a very precise prescription to compute the effect of a change of framing for open string amplitudes. Let us consider for simplicity an open string amplitude involving a single Lagrangian submanifold, computed for a framing f. If we now consider the framing $f-nv$, the coefficients Z_R of the total partition function (147) change as follows:

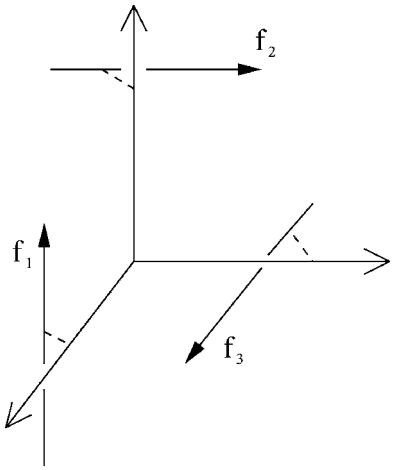

FIG. 22. The canonical choice of framing for the topological vertex.

$$Z_R \to (-1)^{n\ell(R)} q^{n\kappa_R/2} Z_R, \tag{226}$$

where κ_R was defined in Eq. (57), and $q=e^{ig_s}$. This is essentially the behavior of Chern-Simons invariants under change of framing spelled out in Eq. (58). The extra sign in Eq. (226) is crucial to guarantee integrality of the resulting amplitudes, as verified in Aganagic, Klemm, and Vafa (2002) and Mariño and Vafa (2002). If the open string amplitudes involve L boundaries, one has to specify L different framings, and Eq. (226) is generalized to

$$Z_{R_1\cdots R_L} \to (-1)^{\sum_{\alpha=1}^L n_\alpha \ell(R_\alpha)} q^{\sum_{\alpha=1}^L n_\alpha \kappa_{R_\alpha}/2} Z_{R_1\cdots R_L}. \tag{227}$$

B. Definition of the topological vertex

In Sec. IV we considered $\mathbf{C}^3$ with one Lagrangian submanifold in each of the vertices of the toric diagram. Since each of these submanifolds has the topology of $\mathbf{C}\times\mathbf{S}^1$, we can consider the topological open string amplitude associated to this geometry. The total open string partition function is given by

$$Z(V_i) = \sum_{R_1,R_2,R_3} C_{R_1R_2R_3} \prod_{i=1}^3 \mathrm{Tr}_{R_i} V_i, \tag{228}$$

where V_i is a matrix source associated to the ith Lagrangian submanifold. The amplitude $C_{R_1R_2R_3}$ is naturally a function of the string coupling constant g_s and, in the genus expansion, it contains information about maps from Riemann surfaces of arbitrary genera into $\mathbf{C}^3$ with boundaries on L_i. This open string amplitude is called the *topological vertex*, and it is the basic object from which, by gluing, one can obtain closed and open string amplitudes on arbitrary toric geometries. Since the vertex is an open string amplitude, it will depend on a choice of three different framings. As explained in the

previous section, this choice will be given by three different vectors f_1, f_2, and f_3 that specify extra degeneration loci and lead to a compactification of the L_i.

As mentioned in Sec. V.D, the $\mathbf{C}^3$ geometry can be represented by graphs involving three vectors v_i obtained from the set in Fig. 10 by an $\mathrm{Sl}(2,\mathbf{Z})$ transformation, and satisfying Eq. (153). We then introduce a topological vertex amplitude $C_{R_1R_2R_3}^{(v_i,f_i)}$ which depends on both a choice of three vectors v_i for the edges, and a choice of three vectors f_i for the framings. As a result of Eq. (225) we require

$$f_i \wedge v_i = 1.$$

We orient the edges v_i in a clockwise way. Since wedge products are preserved by $\mathrm{Sl}(2,\mathbf{Z})$, we also have

$$v_2 \wedge v_1 = v_3 \wedge v_2 = v_1 \wedge v_3 = 1. \tag{229}$$

However, not all of these choices give independent amplitudes. First of all, there is an underlying $\mathrm{Sl}(2,\mathbf{Z})$ symmetry relating the choices: if $g \in \mathrm{Sl}(2,\mathbf{Z})$, then the amplitudes are invariant under

$$(f_i, v_i) \to (g \cdot f_i, g \cdot v_i).$$

Moreover, if the topological vertex amplitude $C_{R_1R_2R_3}^{(v_i,f_i)}$ is known for a set of framings f_i, then it can be obtained for any set of the form $f_i - n_i v_i$, and it is given by the general rule (227):

$$C_{R_1R_2R_3}^{(v_i,f_i-nv_i)} = (-1)^{\sum_i n_i \ell(R_i)} q^{\sum_i n_i \kappa_{R_i}/2} C_{R_1R_2R_3}^{(v_i,f_i)}, \tag{230}$$

for all admissible choices of the vectors v_i. Since any two choices of framing can be related through Eq. (230), it is useful to pick a convenient set of f_i for any given choice of v_i which we will define as the *canonical framing* of the topological vertex. This canonical framing turns out to be

$$(f_1, f_2, f_3) = (v_2, v_3, v_1).$$

Due to the $\mathrm{Sl}(2,\mathbf{Z})$ symmetry and the transformation rule (230), any topological vertex amplitude can be obtained from the amplitude computed for a *fixed* choice of v_i in the canonical framing. A useful choice of the v_i is $v_1=(-1,-1)$, $v_2=(0,1)$, $v_3=(1,0)$, as in Fig. 10. The vertex amplitude for the canonical choice of v_i and in the canonical framing will be simply denoted by $C_{R_1R_2R_3}$. Any other choice of framing will be characterized by framing vectors of the form $f_i - n_i v_i$, and the corresponding vertex amplitude will be denoted by

$$C_{R_1R_2R_3}^{n_1,n_2,n_3}.$$

Note that $n_i = f_i \wedge v_{i+1}$ (where i runs mod 3).

One of the most important properties of $C_{R_1R_2R_3}$ is its *cyclic symmetry*. To see this, note that the $\mathrm{Sl}(2,\mathbf{Z})$ transformation $g = TS^{-1}$ takes

$$(v_i, f_i) \to (v_{i+1}, f_{i+1}),$$

where again i runs mod 3. It then follows that

$$C_{R_1R_2R_3} = C_{R_3R_1R_2} = C_{R_2R_3R_1}. \tag{231}$$

Finally, it will be sometimes useful to consider the vertex in the basis of conjugacy classes $C_{\vec{k}^{(1)}\vec{k}^{(2)}\vec{k}^{(3)}}$, which is obtained from $C_{R_1R_2R_3}$ by

$$C_{\vec{k}^{(1)}\vec{k}^{(2)}\vec{k}^{(3)}} = \sum_{R_i} \prod_{i=1}^{3} \chi_{R_i}(C(\vec{k}^{(i)})) C_{R_1R_2R_3}. \tag{232}$$

C. Gluing rules

We saw in Sec. IV that any noncompact toric geometry can be encoded in a planar graph which can be obtained by gluing trivalent vertices. It is then natural to expect that the string amplitudes associated to such a diagram can be computed by gluing the open topological string amplitudes associated to the trivalent vertices, in the same way that one computes amplitudes in perturbative quantum field theory by gluing vertices through propagators. This idea was suggested by Iqbal (2002) and Aganagic, Mariño, and Vafa (2004), and was developed into a complete set of rules by Aganagic, Klemm, Mariño, and Vafa (2003). The gluing rules for the topological vertex turn out to be quite simple. Here we will state three rules (for a change of orientation in one edge, for the propagator, and for the matching of framings in the gluing) which make it possible to compute any closed string amplitude on toric, noncompact Calabi-Yau threefolds. They also make it possible to compute open string amplitudes for Lagrangian submanifolds on edges that go to infinity. The case of Lagrangian submanifolds in inner edges is also very easy to analyze, but we refer the reader to the paper by Aganagic, Klemm, Mariño, and Vafa (2003) for the details. A mathematical point of view on the gluing rules can be found in Diaconescu and Florea (2003).

1. Orientation

Trivalent vertices are glued along their edges, and this corresponds to gluing curves with holes along their boundaries. In order to do that, the boundaries must have opposite orientations. This change of orientation will be represented as an inversion of the edge vector, therefore in gluing the vertices we will have an outgoing edge on one side, say v_1, and an ingoing edge on the other side, $-v_1$. What is the corresponding effect on the amplitude $C_{\vec{k}^{(1)}\vec{k}^{(2)}\vec{k}^{(3)}}$? Changing the orientation of h boundaries along the first edge gives rise to a relative factor $(-1)^h$, where $h = |\vec{k}^{(1)}|$. In the language of topological D-branes, this means that we are gluing branes to antibranes (Vafa, 2001b). If we denote Q^t as the representation whose Young tableau is transposed to the Young tableau of Q (i.e., is obtained by exchanging rows and columns), then one has the following relation between characters:

$$\chi_{Q^t}(C(\vec{k})) = (-1)^{|\vec{k}|+\ell(Q)} \chi_Q(C(\vec{k})), \tag{233}$$

and from here one can easily deduce

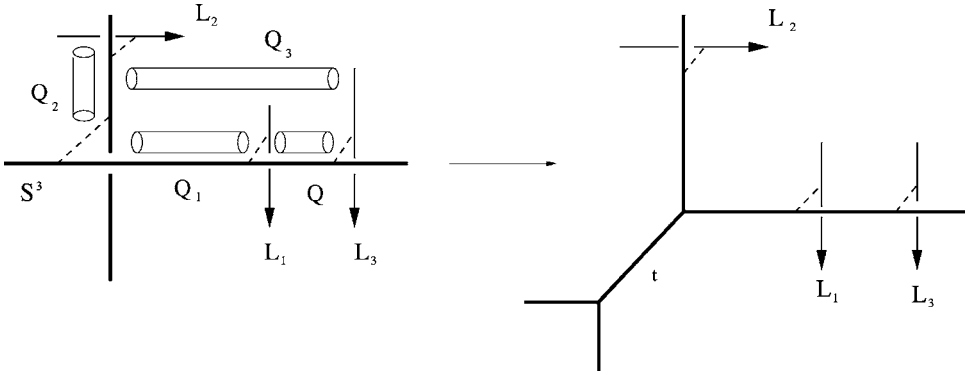

FIG. 23. The configuration used to derive the topological vertex amplitude. The "deformed" geometry shown on the left. The "resolved" geometry obtained by geometric transition shown on the right. It contains a resolved conifold and a $\mathbb{P}^1$ of size t. Also depicted on the left is the open strings stretched among the different branes which contribute to the amplitude.

$$C_{R_1R_2R_3} \to (-1)^{\ell(R_1)} C_{R_1^t R_2 R_3}$$

as we invert the orientation of v_1. Of course, a similar equation follows for the other v_i.

2. Propagator

Since gluing the edges corresponds to gluing curves with holes along their boundaries, we must have a matching number of holes and winding numbers along the edge. Therefore, the propagator must be diagonal in the $\vec{k}$ basis. After taking into account the change of orientation discussed above, and after dividing by the order of the automorphism group associated to $\vec{k}$ (which is nothing but $z_{\vec{k}}$), we find that the propagator for gluing edges with representations R_1, R_2 is given by

$$(-1)^{\ell(R_1)} e^{-\ell(R_1)t} \delta_{R_1 R_2^t}, \tag{234}$$

where t is the Kähler parameter that corresponds to the $\mathbb{P}^1$ represented by the gluing edge.

3. Framing

When gluing two vertices, the framings of the two edges involved in the gluing have to match. This means that, in general, we will have to change the framing of one of the vertices. Let us consider the case in which we glue together two vertices with outgoing vectors (v_i, v_j, v_k) and (v_i', v_j', v_k'), respectively, and let us assume that we glue them through the vectors v_i, $v_i' = -v_i$. We also assume that both vertices are canonically framed, so that $f_i = v_j$, $f_i' = v_j'$. In order to match the framings we have to change the framing of, say, v_i', so that the new framing is $-f_i$ (the opposite sign is again due to the change of orientation). There is an integer n_i such that $f_i' - n_i v_i' = -f_i$ (since $f_i \wedge v_i = f_i' \wedge v_i' = 1$, $f_i + f_i'$ is parallel to v_i), and it is immediate to check that

$$n_i = v_j' \wedge v_j.$$

The gluing of the two vertex amplitudes is then given by

$$\sum_{R_i} C_{R_j R_k R_i} e^{-\ell(R_i)t_i} (-1)^{(n_i+1)\ell(R_i)} q^{-n_i \kappa_{R_i}/2} C_{R_i^t R_j' R_k'}, \tag{235}$$

where we have taken into account the change of orientation in the (v_i', v_j', v_k') to perform the gluing, and t_i is Kähler parameter associated to the edge.

Given then a planar trivalent graph representing a noncompact Calabi-Yau manifold without D-branes, we can compute the closed string amplitude as follows: we give a presentation of the graph in terms of vertices glued together, as in Sec. IV.D. We associate the appropriate amplitude to each trivalent vertex (labeled by representations), and use the above gluing rules. The edges that go to infinity carry the trivial representation, and finally sum over all possible representations along the inner edges. The resulting quantity is the total partition function $Z_{\text{closed}} = e^F$ for closed string amplitudes. We can slightly modify this rule to compute open string amplitudes associated to D-branes, in the simple case in which the Lagrangian submanifolds are located at the outer edges of the graph (i.e., the edges that go to infinity). In this case, we compute the amplitude by associating the representations $R_1, \ldots, R_L$ to the outer edges with D-branes. The result is $Z_{\text{closed}} Z_{R_1 \cdots R_L}$, where $Z_{R_1 \cdots R_L}$ is the open string amplitude that appears in Eq. (146).

We will present some concrete examples of this procedure in a moment. Before doing that, we derive an explicit expression for the topological vertex amplitude using a geometric transition.

D. Derivation of the topological vertex

In order to derive the expression for the vertex, we consider the configuration drawn in the first picture in Fig. 23, which represents a geometry with an $\mathbf{S}^3$ together with three Lagrangian submanifolds L_1, L_2, and L_3. We also make a choice of framing for these Lagrangian submanifolds, indicated by arrows. The world-volumes of the $\mathbf{S}^3$ and of L_1, L_3 are parallel, and we consider topological D-branes wrapped on $\mathbf{S}^3$ and the L_i. The branes wrapping the L_i are probe (spectator) branes, and the large N transition of the three-sphere leads to a geom-

etry with a resolved conifold and three framed Lagrangian submanifolds. As we have seen in the examples above, the Kähler parameter of the $\mathbb{P}^1$ of the conifold t is the 't Hooft parameter of the Chern-Simons theory on $\mathbf{S}^3$. The resulting configuration is shown in the second picture of Fig. 23, and can be related to the topological vertex of Fig. 22 by (i) taking the Kähler parameter t of the $\mathbb{P}^1$ in the resolved conifold to infinity (so that the extra trivalent vertex disappears), and by (ii) moving the Lagrangian submanifold L_1 to the outgoing edge along the direction $(-1,-1)$. We first compute the total open string amplitude using the geometric transition, and then implement (i) and (ii).

The open string theory on the $\mathbf{S}^3$ is $U(N)$ Chern-Simons theory with some matter fields coming from the three noncompact Lagrangian submanifolds L_i. As discussed in Sec. VI, there are strings stretching between the $\mathbf{S}^3$ and $L_{1,2,3}$, and also strings between L_1 and L_3. These stretched strings are annuli along the degeneracy locus, and they are depicted in Fig. 23. The only spacetime excitation associated with these strings is a matter field in the bifundamental representation, and integrating it out corresponds to inserting an annulus operator similar to Eq. (215). When the two branes intersect on a circle (like the branes considered in Sec. VI, or like the $\mathbf{S}^3$ and L_2 in this situation) the matter field is a boson (a scalar field). When the branes are parallel, however, it is a fermion. This is because we can turn the two parallel branes into a brane and an antibrane intersecting along a circle. This leads to a Grassmann field, as explained by Vafa (2001b), and the resulting operator turns out to be

$$\exp\left\{-\sum_n \frac{e^{-nt}}{n}\mathrm{Tr}\, U^n \mathrm{Tr}\, V^n\right\} = \sum_R \mathrm{Tr}_R U e^{-\ell(R)t}(-1)^{\ell(R)}\mathrm{Tr}_{R^t} V. \tag{236}$$

In Fig. 23 the probe branes associated with L_2 and the dynamical branes on $\mathbf{S}^3$ intersect on a circle, while the probe branes associated with L_1, L_3 are parallel to each other and to the dynamical branes. We then have the following operators:

$$\sum_{Q_1} \mathrm{Tr}_{Q_1^t} U_1 e^{-\ell(Q_1)t}(-1)^{\ell(Q_1)}\mathrm{Tr}_{Q_1}\hat{V}_1,$$

$$\sum_R \mathrm{Tr}_{Q_2} U_2 e^{-\ell(Q_2)t}\mathrm{Tr}_{Q_2} V_2,$$

$$\sum_{Q_3} \mathrm{Tr}_{Q_3^t} U_1 e^{-\ell(Q_3)t}(-1)^{\ell(Q_3)}\mathrm{Tr}_{Q_3} V_3,$$

$$\sum_Q \mathrm{Tr}_{Q^t} V_1 e^{-\ell(Q)t}(-1)^{\ell(Q)}\mathrm{Tr}_Q V_3, \tag{237}$$

which correspond to the annuli labeled with representations Q_1, Q_2, Q_3, and Q in Fig. 23. The matrices V_2 and V_3 are sources corresponding to D-branes wrapping L_2, L_3, while $\hat{V}_1, V_1$ are sources for branes wrapping L_1 with opposite orientations, and represent Chan-Paton factors for open strings ending on opposite sides of L_1. U_1, U_2 are holonomies of the gauge connection on $\mathbf{S}^3$ around the boundaries of the annuli with representations Q_1 and Q_2 (the boundary of the annulus carrying the representation Q_3 is geometrically identical to the boundary of the annulus associated to Q_1, and it gives the holonomy U_1 as well). Putting all these ingredients together, we find that the open string amplitude on the deformed geometry is given by

$$Z(V_1,V_2,V_3) = \frac{1}{S_{00}}\sum_{Q_1,Q_2,Q_3,Q}(-1)^{\ell(Q_1)+\ell(Q_3)+\ell(Q)} \times\langle \mathrm{Tr}_{Q_2} U_2 \mathrm{Tr}_{Q_1^t} U_1 \mathrm{Tr}_{Q_3^t} U_1\rangle \times \mathrm{Tr}_{Q_1}\hat{V}_1 \mathrm{Tr}_{Q^t} V_1 \mathrm{Tr}_{Q_2} V_2 \mathrm{Tr}_{Q\otimes Q_3} V_3, \tag{238}$$

where we have factored out $1/S_{00}$, the partition function of $\mathcal{O}(-1)\oplus\mathcal{O}(-1)\to\mathbb{P}^1$. The above amplitude is an open string amplitude with three boundaries, and V_i are the corresponding sources. Note that the annuli that carry the representations Q_1, Q_3 are supported on the horizontal edge, while the annulus connecting L_2 to $\mathbf{S}^3$ lies on the vertical edge. The horizontal and the vertical edge are related by an S transformation, therefore

$$\langle \mathrm{Tr}_{Q_2} U_2 \mathrm{Tr}_{Q_1^t} U_1 \mathrm{Tr}_{Q_3^t} U_1\rangle = \sum_{Q'} N_{Q_1^t Q_3^t}^{Q'} \langle Q_2|S|Q'\rangle = \sum_{Q'} N_{Q_1^t Q_3^t}^{Q'} S^{-1}_{Q_2 Q'}, \tag{239}$$

where we have fused together the U_1 holonomies. From a geometric point of view, this means that the boundaries of the annuli give a link in $\mathbf{S}^3$ with the topology depicted in Fig. 5 (where the representations R, R_1, R_2 in Fig. 5 are now Q_2, Q_3, and Q_1, respectively), and the above expression is nothing but Eq. (67). We can also use the direct sum formula (66) for this invariant, and we finally arrive at the following expression for Eq. (238):

$$Z(V_1,V_2,V_3) = \sum_{Q_1,Q_2,Q_3,Q}(-1)^{\ell(Q_1)+\ell(Q_3)+\ell(Q)} \times\frac{\mathcal{W}_{Q_1^t Q_2}\mathcal{W}_{Q_3^t Q_2}}{\mathcal{W}_{Q_2}}\mathrm{Tr}_{Q_1}\hat{V}_1 \mathrm{Tr}_{Q^t} V_1 \mathrm{Tr}_{Q_2} V_2 \times \mathrm{Tr}_{Q\otimes Q_3} V_3, \tag{240}$$

where $\mathcal{W}_{R_1 R_2}$ is the Hopf link invariant defined in Eq. (37) and evaluated in Eq. (45). Equation (240) gives the answer for the open topological string amplitude on the geometry depicted on the left in Fig. 23. We now incorporate the two modifications which are needed in order to obtain the topological vertex. First of all, we have to take $t\to\infty$. As pointed out in the last section, in order to have a well-defined limit it is crucial to renormalize the Chern-Simons expectation values. The relation (222) suggests the definition

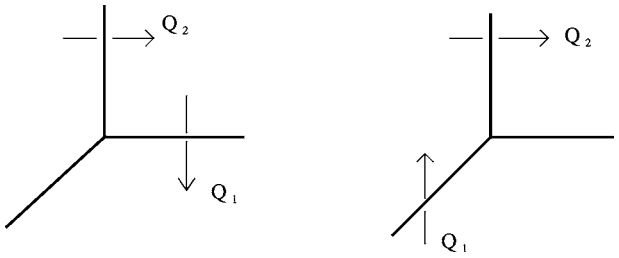

FIG. 24. Moving the Lagrangian submanifold with representation Q_1 to the outgoing edge.

$$W_{R_1R_2} = \lim_{t\to\infty} e^{-[\ell(R_1)+\ell(R_2)]t/2}\mathcal{W}_{R_1R_2}. \tag{241}$$

This limit exists, since $\mathcal{W}_{R_1R_2}$ is of the form $\lambda^{[\ell(R_1)+\ell(R_2)]/2}W_{R_1R_2}+\mathcal{O}(e^{-t})$ (remember that $\lambda=e^t$). The quantity $W_{R_1R_2}$, which is the "leading" coefficient of the Hopf link invariant (37), is the building block of the topological vertex amplitude. It is a rational function of $q^{\pm 1/2}$, therefore it only depends on the string coupling constant. We will also denote $W_R=W_{R0}$. The limit (241) was first considered by Aganagic, Mariño, and Vafa (2004).

In order to implement the second modification, we have to understand what is the effect on the amplitude of moving L_1 to the outgoing edge along $(-1,-1)$. In order to do that, consider the simplified situation depicted in Fig. 24 where we only have two stacks of D-branes wrapping $L_{1,2}$. The amplitude can be easily computed following the arguments that led to Eq. (240), and one immediately obtains

$$Z(V_1,V_2) = \sum_{Q_1,Q_2} W_{Q_2Q_1^t}(-1)^{\ell(Q_1)}\mathrm{Tr}_{Q_1}V_1\mathrm{Tr}_{Q_2}V_2.$$

On the other hand, this is a particular case of the topological vertex amplitude with R_1 trivial, $f_2=\nu_3$ and $f_3=(0,-1)$, so there is a noncanonical framing on ν_3 which corresponds to $n=-1$. We deduce

$$C_{0R_2R_1} = W_{R_2R_1^t}q^{\kappa_{R_1}/2}. \tag{242}$$

On the other hand, the amplitude on the right-hand side of Fig. 24 is the canonically framed vertex $C_{R_1R_20}$, but by cyclic symmetry this is equal to Eq. (242) with $R_2\leftrightarrow R_1$. We conclude that in going from the left- to the right-hand side of Fig. 24 we must replace

$$\begin{aligned}&(-1)^{\ell(Q_1)}W_{Q_2Q_1^t}\mathrm{Tr}_{Q_1}\hat{V}_1\mathrm{Tr}_{Q_2}V_2\\ &\quad\to W_{Q_2^tQ_1}q^{\kappa_{Q_2}/2}\mathrm{Tr}_{Q_1}V_1\mathrm{Tr}_{Q_2}V_2.\end{aligned} \tag{243}$$

After moving L_1 to the outgoing edge, all strings end on the same side of the corresponding branes, and this explains why we have replaced $\hat{V}_1$ by V_1 in the above formula. Collecting the coefficient of $\mathrm{Tr}_{R_1}V_1\mathrm{Tr}_{R_2}V_2\mathrm{Tr}_{R_3}V_3$ in the partition function we compute $C_{R_1R_2R_3}^{0,0,-1}$. We then get the following expression for the topological vertex amplitude in the canonical framing:

$$C_{R_1R_2R_3} = q^{(\kappa_{R_2}+\kappa_{R_3})/2}\sum_{Q_1,Q_3,Q} N_{QQ_1}^{R_1}N_{QQ_3}^{R_3^t}\frac{W_{R_2^tQ_1}W_{R_2Q_3}}{W_{R_2}}. \tag{244}$$

This is the final expression for the topological vertex amplitude.

Using Eq. (45) it is possible to give an explicit expression for $W_{R_1R_2}$ which is useful in computations. It is easy to see that the leading coefficient in λ in Eq. (45) is obtained by taking the λ-independent piece in Eq. (48). The generating function of elementary symmetric polynomials (46) then becomes

$$S(t)\prod_{j=1}^{c_R}\frac{1+q^{t_j^R-j}t}{1+q^{-j}t}, \tag{245}$$

where

$$S(t) = \prod_{j=1}^{\infty}(1+q^{-j}t) = 1+\sum_{r=1}^{\infty}\frac{q^{-r(r+1)/2}t^r}{\prod_{m=1}^{r}[m]}. \tag{246}$$

In terms of Schur polynomials, we find

$$W_{R_1R_2}(q) = s_{R_2}(x_i=q^{-i+1/2})s_{R_1}(x_i=q^{t_i^{R_2}-1+1/2}), \tag{247}$$

where there are now an infinite number of variables x_i with $i=1,2,\ldots$. One can also write Eq. (244) in terms of skew Schur polynomials (Okounkov, Reshetikhin, and Vafa, 2003), and using the properties of these polynomials one finds identities for the topological vertex that are very useful in computations (Hollowood, Iqbal, and Vafa, 2003; Eguchi and Kanno, 2004).

E. Some applications

We now present some examples of computation of topological string amplitudes using the topological vertex.

1. Resolved conifold

The toric diagram for the resolved conifold geometry is depicted in Fig. 11. Our rules give immediately

$$Z_{\mathbb{P}^1} = \sum_R C_{00R^t}(-1)^{\ell(R)}e^{-\ell(R)t}C_{R00}. \tag{248}$$

Since $C_{R00}=W_R=s_R(x_i=q^{-i+1/2})$, we can use the well-known formula (see, for example, Fulton and Harris, 1991; Macdonald, 1995)

$$\sum_R s_R(x)s_R(y) = \frac{1}{\prod_{i,j}(1-x_iy_j)} \tag{249}$$

to obtain

$$Z_{\mathbb{P}^1} = \exp\left\{-\sum_{d=1}^{\infty}\frac{e^{-dt}}{d(q^{d/2}-q^{-d/2})^2}\right\}, \tag{250}$$

in agreement with the known result (168).

2. Framed unknot

Let us now consider an open string amplitude, corresponding to the resolved conifold with a Lagrangian brane in one of the external legs, and in arbitrary framing p. The open string amplitude is given by

$$Z(V,p)=\sum_Q Z_Q(p)\mathrm{Tr}_Q V, \tag{251}$$

where

$$Z_Q(p)=\frac{1}{Z_{\mathbb{P}^1}}(-1)^{\ell(Q)p}q^{\kappa_Q p/2}\sum_R C_{0QR^t}(-e^{-t})^{\ell(R)}C_{R00}. \tag{252}$$

One can use various identities involving symmetric polynomials to show that

$$Z_Q(p)=(-1)^{\ell(Q)p}q^{\kappa_Q p/2}e^{-\ell(Q)t/2}(\dim_q Q). \tag{253}$$

The right-hand side is essentially the Chern-Simons invariant of the unknot. The open string free energy is given by $F(V,p)=\ln Z(V,p)$, which can be written as in Eq. (144). It turns out that the leading term of $F_{w,g}(p,t)$ as $t\to\infty$ [which we will simply denote by $F_{w,g}(p)$] can be computed in open Gromov-Witten theory by using Hodge integrals [Katz and Liu (2002); see also Li and Song (2002)]. The result is

$$F_{w,g}(p)=(-1)^{p\ell+1}[p(p+1)]^{h-1}\times\left(\prod_{i=1}^{h}\frac{\prod_{j=1}^{w_i-1}(j+w_i p)}{(w_i-1)!}\right)\mathrm{Res}_{u=0}\times\int_{\bar{M}_{g,h}}\frac{c_g(\mathbb{E}^\vee(u))c_g(\mathbb{E}^\vee[(-p-1)u])c_g(\mathbb{E}^\vee(pu))}{\prod_{i=1}^{h}(u-w_i\psi_i)}. \tag{254}$$

In this formula, $\bar{M}_{g,h}$ is the Deligne-Mumford moduli space, $\mathbb{E}$ is the Hodge bundle over $\bar{M}_{g,h}$, and its dual is denoted by $\mathbb{E}^\vee$. We have also written

$$c_g(\mathbb{E}^\vee(u))=\sum_{i=0}^{g}c_{g-i}(\mathbb{E}^\vee)u^i, \tag{255}$$

where $c_j(\mathbb{E}^\vee)$ are Chern classes, and similarly for the other two factors. On the other hand, the $t\to\infty$ limit of Z_Q is

$$(-1)^{\ell(Q)p}q^{\kappa_Q p/2}W_Q. \tag{256}$$

By equating the open Gromov-Witten result with the Chern-Simons result one finds a highly nontrivial identity that expresses the Hodge integrals appearing in Eq. (254) in terms of the W_Q's, as first noticed by Mariño and Vafa (2002). The explicit expression that one finds is

$$\sum_{g=0}^{\infty}F_{\vec{k},g}g_s^{2g-2+|\vec{k}|}=(-1)^{p\ell}i^{-|\vec{k}|-\ell}\prod_j k_j!\sum_{n\geqslant 1}\frac{(-1)^n}{n}\times\sum_{\vec{k}_1,\cdots,\vec{k}_n}\delta_{\Sigma_{\sigma=1}^n\vec{k}_\sigma,\vec{k}}\sum_{R_\sigma}\prod_{\sigma=1}^{n}\frac{\chi_{R_\sigma}(C(\vec{k}_\sigma))}{z_{\vec{k}_\sigma}}\times e^{i(p+1/2)\kappa_{R_\sigma}g_s/2}\times\prod_{1\leqslant i<j\leqslant c_{R_\sigma}}\frac{\sin[(l_i^\sigma-l_j^\sigma+j-i)g_s/2]}{\sin[(j-i)g_s/2]}\times\prod_{i=1}^{c_{R_\sigma}}\prod_{v=1}^{l_i^\sigma}\frac{1}{2\sin[(v-i+c_{R_\sigma})g_s/2]}, \tag{257}$$

where we have relabeled $w\to\vec{k}$ for positive winding numbers, as explained in Sec. IV.C. This equality is a very explicit mathematical prediction of the duality between Chern-Simons theory and topological string theory. It has been rigorously proved by Liu, Liu, and Zhou (2003a) and by Okounkov and Pandharipande (2004), and shown to have many applications in Gromov-Witten theory (Liu, Liu, and Zhou, 2003b; Zhou, 2003).

3. Local $\mathbb{P}^2$

The toric diagram is depicted in Fig. 12. Using again the rules explained above, we find the total partition function

$$Z_{\mathbb{P}^2}=\sum_{R_1,R_2,R_3}(-1)^{\Sigma_i\ell(R_i)}e^{-\Sigma_i\ell(R_i)t}q^{-\Sigma_i\kappa_{R_i}}\times C_{0R_2^tR_3}C_{0R_1^tR_2}C_{0R_3^tR_1}, \tag{258}$$

where t is the Kähler parameter corresponding to the hyperplane class in $\mathbb{P}^2$. Using that $C_{0R_2R_3^t}=W_{R_2R_3}q^{-\kappa_{R_3}/2}$ one recovers the expression for $Z_{\mathbb{P}^2}$ first obtained by Aganagic, Mariño, and Vafa (2004) using the method of geometric transition explained in Sec. VI. Note that the free energy has the structure

$$F_{\mathbb{P}^2}=\ln\left\{1+\sum_{\ell=1}^{\infty}a_\ell(q)e^{-\ell t}\right\}=\sum_{\ell=1}^{\infty}a_\ell^{(c)}(q)e^{-\ell t}. \tag{259}$$

The coefficients $a_\ell(q),a_\ell^{(c)}(q)$ can be easily obtained in terms of $W_{R_1R_2}$. One finds, for example,

$$a_1^{(c)}(q)=a_1(q)=-\frac{3}{(q^{1/2}-q^{-1/2})^2},$$

$$a_2^{(c)}(q)=\frac{6}{(q^{1/2}-q^{-1/2})^2}+\frac{1}{2}a_1(q^2). \tag{260}$$

If we compare to Eq. (129) and take into account the effects of multicovering, we find the following values for the Gopakumar-Vafa invariants of $\mathcal{O}(-3)\to\mathbb{P}^2$:

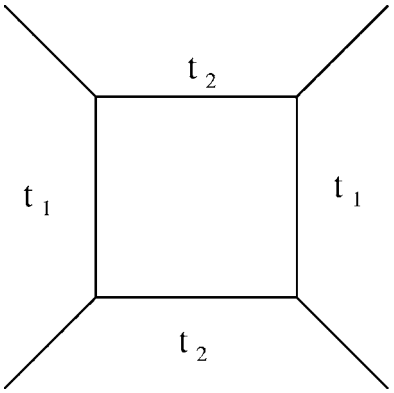

FIG. 25. The toric diagram of local $\mathbb{P}^1\times\mathbb{P}^1$.

$$n_1^0=3,\quad n_1^g=0 \text{ for } g>0,$$

$$n_2^0=-6,\quad n_2^g=0 \text{ for } g>0, \tag{261}$$

in agreement with the results listed in Eq. (171). In fact, one can go much further with this method and compute the Gopakumar-Vafa invariants to a high degree. We see again that the use of nonperturbative results in Chern-Simons theory leads to the topological string amplitudes to *all genera*. A complete listing of the Gopakumar-Vafa invariants up to degree 12 can be found in Aganagic, Mariño, and Vafa (2004). The partition function (258) can be also computed in Gromov-Witten theory by using localization techniques, and one finds indeed the same result (Zhou, 2003).

4. Local $\mathbb{P}^1\times\mathbb{P}^1$

The local $\mathbb{P}^1\times\mathbb{P}^1$ geometry is the noncompact Calabi-Yau manifold given by the four-manifold $\mathbb{P}^1\times\mathbb{P}^1$ together with its anticanonical bundle. It also admits a symplectic quotient description of the form (256), this time with $N=2$ and two Kähler parameters t_1, t_2. The charges $Q_{1,2}^j$, $j=1,\ldots,5$, can be grouped in two vectors,

$$Q_1=(-2,1,1,0,0),$$

$$Q_2=(-2,0,0,1,1). \tag{262}$$

The toric diagram for this geometry can be easily worked out from this description, and it is represented in Fig. 25. Using the gluing rules we find the closed string partition function

$$Z_{\mathbb{P}^1\times\mathbb{P}^1}=\sum_{R_i} e^{-[\ell(R_1)+\ell(R_3)]t_1-[\ell(R_2)+\ell(R_4)]t_2}q^{\Sigma_i\kappa_{R_i}/2} \times C_{0R_4R_1^t}C_{0R_1R_2^t}C_{0R_2R_3^t}C_{0R_3R_4^t}. \tag{263}$$

This amplitude can be written as

$$Z_{\mathbb{P}^1\times\mathbb{P}^1}=\sum_{R_i} e^{-[\ell(R_1)+\ell(R_3)]t_1-[\ell(R_2)+\ell(R_4)]t_2} \times W_{R_4R_1}W_{R_1R_2}W_{R_2R_3}W_{R_3R_4}. \tag{264}$$

This is the expression first obtained by Aganagic, Mariño, and Vafa (2004), and it has been shown to agree with Gromov-Witten theory by Zhou (2003). The local $\mathbb{P}^1\times\mathbb{P}^1$ geometry is interesting since it gives a string realization of the Seiberg-Witten solution (1994) of $SU(2)$, $\mathcal{N}=2$ Yang-Mills theory, as first shown by Katz, Klemm, and Vafa (1997). In particular, the prepotential of local $\mathbb{P}^1\times\mathbb{P}^1$ gives, in a certain limit, the Seiberg-Witten prepotential. This was explicitly verified by Iqbal and Kashani-Poor (2003a) by using the expression (264) and the results of Nekrasov (2002) for the Seiberg-Witten prepotential. Further applications of the topological vertex to the computation of supersymmetric gauge theory amplitudes can be found in Eguchi and Kanno (2003), Hollowood, Iqbal, and Vafa (2003), and Iqbal and Kashani-Poor (2003b).

F. Further properties of the topological vertex

Since it was first introduced by Aganagic, Klemm, Mariño, and Vafa (2003), the topological vertex has been shown to satisfy three remarkable properties: it has an underlying integrable structure (Aganagic, Dijkgraaf, Klemm, Mariño, and Vafa, 2003), it has a natural combinatorial interpretation in terms of counting tridimensional Young tableaux (Okounkov, Reshetikhin, and Vafa, 2003), and it can be also reinterpreted in terms of an appropriate counting of sheaves on $\mathbf{C}^3$ (Iqbal *et al.*, 2003; Maulik *et al.*, 2003). We briefly review each of these properties.

1. Integrable structure

If we put $\mathrm{Tr}\, V_i^n=t_n^i/n$ in Eq. (228), the resulting function of three infinite sets of "times" $Z(t_n^i)$ turns out to be a tau function of the 3-KP hierarchy as constructed, for example, by Kac and van de Leur (2003). This integrability property is better understood in the context of mirror symmetry, where the computation of the vertex can be seen to reduce to a theory of free fermions in a Riemann surface with three punctures (Aganagic, Dijkgraaf, Klemm, Mariño, and Vafa, 2003).

2. Combinatorial interpretation

Consider the problem of enumerating three-dimensional Young tableaux π (also called plane partitions) with the following boundary condition: along the edges x,y,z they end up in two-dimensional Young tableaux with the shapes R_1, R_2, and R_3, respectively. Let us introduce the partition function

$$\mathcal{C}_{R_1R_2R_3}=\sum_{\pi} q^{|\pi|}, \tag{265}$$

where $|\pi|$ is the number of boxes in π, and the sum is over plane partitions satisfying the above boundary conditions. It can be shown that, up to an overall factor independent of the R_i, the above partition function equals the topological vertex $C_{R_1R_2R_3}$ (Okounkov,

Reshetikhin, and Vafa, 2003). This combinatorial interpretation of the topological vertex makes it possible to establish a precise correspondence between quantum topological strings on local, toric Calabi-Yau manifolds and the classical statistical mechanics of melting crystals (Iqbal *et al.*, 2003; Okounkov, Reshetikhin, and Vafa, 2003; Saulina and Vafa, 2004).

3. Relation to the counting of ideal sheaves

Let X be a Calabi-Yau threefold. An ideal sheaf $\mathcal{I}$ defines a closed subscheme Y through $\mathcal{O}_Y=\mathcal{O}_X/\mathcal{I}$. This means, roughly speaking, that there is a subvariety Y of X defined by the zero locus of the equations that generate the ideal $\mathcal{I}$. Given a two-homology class β, one can consider the moduli space of ideal sheaves $I_n(X,\beta)$ such that the holomorphic Euler characteristic of Y is n and with $\mathcal{O}_Y$ supported on curves in the homology class β. This is a space of virtual dimension zero, and by counting the number of points with appropriate signs one can define the so-called Donaldson-Thomas invariant $\hat{N}_{n,\beta}$. The Donaldson-Thomas partition function is given by

$$Z_{DT}(X)=\sum_{\beta}\sum_{n\in\mathbf{Z}}\hat{N}_{n,\beta}Q^{\beta}(-q)^n, \tag{266}$$

where q is interpreted here as a formal expansion parameter, and the notation for Q^{β} is identical to the one in Eq. (115). Maulik *et al.* (2003) have shown that, when $X=\mathbf{C}^3$ (so that Z_{DT} only depends on q), the Donaldson-Thomas partition function naturally depends on three sets of representations, and agrees indeed with the topological vertex $C_{R_1R_2R_3}(q)$, where q in Eq. (266) is identified with e^{ig_s}. They have also shown that the Donaldson-Thomas partition function satisfies the same gluing rules as the topological vertex, leading to the identification of the Donaldson-Thomas partition function $Z_{\mathrm{DT}}(X)$ with the topological string all-genus partition function $Z(X)=e^F$ for all noncompact, toric Calabi-Yau manifolds X. They also conjecture that the equality holds for all Calabi-Yau threefolds. Iqbal *et al.* rephrase the Donaldson-Thomas partition function for $\mathbf{C}^3$ (which computes the topological vertex) in terms of the counting of noncommutative $U(1)$ instantons in six dimensions. These developments seem to indicate that the topological vertex, apart from providing a powerful computational tool, plays a central logical role in the theory of Gromov-Witten invariants and opens the way to connections to other moduli problems in algebraic geometry.

VIII. CONCLUSIONS AND FUTURE DIRECTIONS

The correspondence between Chern-Simons theory and topological strings provides one of the most fascinating examples of the string theory–gauge theory correspondence. It has deep mathematical implications that hold a lot of promise for the theory of knot and link invariants, as well as for the theory of Gromov-Witten invariants. The physical point of view, which culminated in the idea of the topological vertex, has allowed us to obtain a complete solution to topological string theory on a wide class of Calabi-Yau threefolds. There are, however, many questions and problems that remain open and will no doubt give us further insights on these connections. We conclude the review with some of these problems.

(i) From a mathematical point of view, some of the ingredients and results in the topological string side need further study. For example, a rigorous construction of open Gromov-Witten invariants has not been given yet, although the formal use of localization techniques leads to sensible results in agreement with the predictions of physics. One important problem is to derive the explicit expression for the topological vertex (244) in the context of Gromov-Witten theory. This will put the physical predictions on a firmer ground. Some steps in this direction have been already taken by Diaconescu and Florea (2003), and Li *et al.* (2004) have given a mathematical treatment of the topological vertex by using relative Gromov-Witten invariants.

(ii) The correspondence between knot invariants and open Gromov-Witten invariants on the resolved conifold explained in Sec. V.G is still very much uncharted. Although there are proposals for Lagrangian submanifolds in the resolved conifold associated to nontrivial knots (Labastida, Mariño, and Vafa, 2000; Taubes, 2001), no results have been obtained for open Gromov-Witten invariants with those Lagrangian boundary conditions. This correspondence is potentially very interesting from a mathematical point of view, since it gives a dictionary between two important and very different sets of invariants. It is likely that the unveiling of this correspondence will lead to deep results in the theory of knot invariants.

(iii) It would be very interesting to see if Chern-Simons theory on other three-manifolds has a string theory description as well, since this would lead in particular to a fascinating reformulation of the theory of finite-type invariants which has been so vigorously developed in the last years. So far only small steps have been taken in this direction. The geometric transition of Gopakumar and Vafa was extended to lens spaces by Aganagic, Klemm, Mariño, and Vafa (2004), and further studied by Okuda and Ooguri (2004). One of the problems faced by the extension of the correspondence to lens spaces is a typical one also encountered in the AdS-CFT correspondence: the Chern-Simons side is easily computed for small 't Hooft coupling, while the topological string theory side is better computed for large 't Hooft coupling. This makes the comparison of observables a difficult task, and in that respect more techniques are needed in order to evaluate the field theory and string theory results in other 't Hooft coupling re-

gimes. The formulation of Chern-Simons theory in terms of a matrix model given by Mariño (2002a), which was very useful for the tests performed by Aganagic, Klemm, Mariño, and Vafa (2004), may also be useful in doing Chern-Simons field theory computations at large 't Hooft coupling.

(iv) The connection between Gromov-Witten and Donaldson-Thomas invariants found by Iqbal *et al.* (2003) and Maulik *et al.* (2003) may shed a new light on many aspects of Gromov-Witten theory, and seems to be a very promising avenue in the mathematical understanding of the Gopakumar-Vafa invariants.

ACKNOWLEDGMENTS

I would like to thank Mina Aganagic, Vincent Bouchard, Robbert Dijkgraaf, Bogdan Florea, Albrecht Klemm, Jose Labastida, and Cumrun Vafa for enjoyable collaborations on the topics discussed in this review. I would also like to thank Vincent Bouchard, Brenno Carlini Vallilo, and Arthur Greenspoon for a careful reading of the manuscript.

APPENDIX: SYMMETRIC POLYNOMIALS

In this appendix we summarize some useful ingredients of the elementary theory of symmetric functions. A standard reference is Macdonald (1995).

Let $x_1,\dots,x_N$ denote a set of N variables. The *elementary symmetric polynomials* in these variables, $e_m(x)$, are defined as

$$e_m(x)=\sum_{i_1<\cdots<i_m} x_{i_1}\cdots x_{i_m}. \tag{A1}$$

The generating function of these polynomials is given by

$$E(t)=\sum_{m\geqslant 0} e_m(x)t^m=\prod_{i=1}^{N}(1+x_i t). \tag{A2}$$

The *complete symmetric function* h_m can be defined in terms of its generating function

$$H(t)=\sum_{m\geqslant 0} h_m t^m=\prod_{i=1}^{N}(1-x_i t)^{-1}, \tag{A3}$$

and one has

$$E(t)H(-t)=1. \tag{A4}$$

The products of elementary symmetric polynomials and of complete symmetric functions provide two different basis for the symmetric functions of N variables.

Another basis is given by the *Schur polynomials* $s_R(x)$ which are labeled by representations R. We express these representations in terms of Young tableaux, so R is given by a partition $(l_1,l_2,\dots,l_{c_R})$, where l_1 is the number of boxes of the ith row of the tableau, and $l_1\geqslant l_2\geqslant\cdots\geqslant l_{c_R}$. The total number of boxes of a tableau will be $\ell(R)=\Sigma_i l_i$. The Schur polynomials are defined as quotients of determinants,

$$s_R(x)=\frac{\det x_j^{l_i+N-i}}{\det x_j^{N-i}}. \tag{A5}$$

They can be written in terms of the symmetric polynomials $e_i(x_1,\dots,x_N)$, $i\geqslant 1$, as follows:

$$s_R=\det M_R, \tag{A6}$$

where

$$M_R^{ij}=(e_{l_i^t+j-i}).$$

M_R is an $r\times r$ matrix, with $r=c_{R^t}$, and R^t denotes the transposed Young tableau with row lengths l_i^t. To evaluate s_R we put $e_0=1, e_k=0$ for $k<0$. The expression (A6) is known as the Jacobi-Trudy identity.

A third set of symmetric functions is given by the *Newton polynomials* $P_{\vec{k}}(x)$. These are labeled by vectors $\vec{k}=(k_1,k_2,\dots)$, where k_j are non-negative integers, and they are defined as

$$P_{\vec{k}}(x)=\prod_{j}^{p} P_j^{k_j}(x), \tag{A7}$$

where

$$P_j(x)=\sum_{i=1}^{N} x_i^j \tag{A8}$$

are power sums. The Newton polynomials are homogeneous of degree $\ell=\Sigma_j jk_j$ and give a basis for the symmetric functions in $x_1,\dots,x_N$ with rational coefficients. They are related to the Schur polynomials through the Frobenius formula

$$P_{\vec{k}}(x)=\sum_R \chi_R(C(\vec{k}))s_R(x), \tag{A9}$$

where the sum is over all tableaux such that $\ell(R)=\ell$.

REFERENCES

Acharya, B. S., 2000, "On realising $\mathcal{N}=1$ super Yang-Mills in M theory," e-print hep-th/0011089.

Aganagic, M., R. Dijkgraaf, A. Klemm, M. Mariño, and C. Vafa, 2003, "Topological strings and integrable hierarchies," e-print hep-th/0312085.

Aganagic, M., A. Klemm, M. Mariño, and C. Vafa, 2003, "The topological vertex," e-print hep-th/0305132.

Aganagic, M., A. Klemm, M. Mariño, and C. Vafa, 2004 "Matrix model as a mirror of Chern-Simons theory," J. High Energy Phys. 0402, 010.

Aganagic, M., A. Klemm, and C. Vafa, 2002, "Disk instantons, mirror symmetry and the duality web," Z. Naturforsch., A: Phys. Sci. **57**, 1–28.

Aganagic, M., M. Mariño, and C. Vafa, 2004, "All loop topological string amplitudes from Chern-Simons theory," Commun. Math. Phys. **247**, 467–512.

Aganagic, M., and C. Vafa, 2000, "Mirror symmetry, D-branes

and counting holomorphic discs," e-print hep-th/0012041.
Aganagic, M., and C. Vafa, 2001, "G_2 manifolds, mirror symmetry and geometric engineering," e-print hep-th/0110171.
Antoniadis, I., E. Gava, K. S. Narain, and T. R. Taylor, 1994, "Topological amplitudes in string theory," Nucl. Phys. B **413**, 162–184.
Atiyah, M., 1990, "On framings of three-manifolds," Topology **29**, 1–7.
Atiyah, M., J. Maldacena, and C. Vafa, 2001, "An *M*-theory flop as a large *N* duality," J. Math. Phys. **42**, 3209–3220.
Becker, K., M. Becker, and A. Strominger, 1995, "Fivebranes, membranes, and nonperturbative string theory," Nucl. Phys. B **456**, 130–152.
Bershadsky, M., S. Cecotti, H. Ooguri, and C. Vafa, 1993, "Holomorphic anomalies in topological field theories," Nucl. Phys. B **405**, 279–304.
Bershadsky, M., S. Cecotti. H. Ooguri, and C. Vafa, 1994, "Kodaira-Spencer theory of gravity and exact results for quantum string amplitudes," Commun. Math. Phys. **165**, 311–428.
Camperi, M., F. Levstein, and G. Zemba, 1990, "The large *N* limit of Chern-Simons theory," Phys. Lett. B **247**, 549–554.
Candelas, P., and X. C. de la Ossa, 1990, "Comments on conifolds," Nucl. Phys. B **342**, 246–268.
Chiang, T. M., A. Klemm, S. T. Yau, and E. Zaslow, 1999, "Local mirror symmetry: Calculations and interpretations," Adv. Theor. Math. Phys. **3**, 495–565.
Coleman, S., 1988, "1/*N*," in *Aspects of Symmetry* (Cambridge University Press, Cambridge, England).
Correale, R., and E. Guadagnini, 1994, "Large *N* Chern-Simons field theory," Phys. Lett. B **337**, 80–85.
Cox, D., and S. Katz, 1999, *Mirror Symmetry and Algebraic Geometry* (American Mathematical Society, Providence, RI).
Diaconescu, D. E., and B. Florea, 2003, "Localization and gluing of topological amplitudes," e-print hep-th/0309143.
Diaconescu, D. E., B. Florea, and A. Grassi, 2003a, "Geometric transitions and open string instantons," Adv. Theor. Math. Phys. **6**, 619–642.
Diaconescu, D. E., B. Florea, and A. Grassi, 2003b "Geometric transitions, del Pezzo surfaces and open string instantons," Adv. Theor. Math. Phys. **6**, 643–702.
Di Francesco, P., P. Mathieu, and D. Sénéchal, 1997, *Conformal Field Theory* (Springer-Verlag, Berlin).
Dijkgraaf, R., E. Verlinde, and H. Verlinde, 1991, "Notes on topological string theory and two-dimensional topological gravity," in *String Theory and Quantum Gravity* (World Scientific, Singapore), p. 91.
Eguchi, T., and H. Kanno, 2003, "Topological strings and Nekrasov's formulas," J. High Energy Phys. 0312, 006.
Eguchi, T., and H. Kanno, 2004, "Geometric transitions, Chern-Simons gauge theory and Veneziano type amplitudes," Phys. Lett. B **585**, 163–172.
Elitzur, S., G. Moore, A. Schwimmer, and N. Seiberg, 1989, "Remarks on the canonical quantization of the Chern-Simons-Witten theory," Nucl. Phys. B **326**, 108–134.
Faber, C., 1999, "Algorithms for computing intersection numbers on moduli spaces of curves, with an application to the class of the locus of Jacobians," in *New Trends in Algebraic Geometry*, edited by K. Hulek *et al.* (Cambridge University Press, Cambridge, England).
Faber, C., and R. Pandharipande, 2000, "Hodge integrals and Gromov-Witten theory," Invent. Math. **139**, 173–199.
Freed, D. S., and R. E. Gompf, 1991, "Computer calculation of Witten's three manifold invariant," Commun. Math. Phys. **141**, 79–117.
Freyd, P., D. Yetter, J. Hoste, W. B. R. Lickorish, K. Millett, and A. A. Ocneanu, 1985, "A new polynomial invariant of knots and links," Bull., New Ser., Am. Math. Soc. **12**, 239–246.
Fulton, W., and J. Harris, 1991, *Representation Theory. A First Course* (Springer-Verlag, Berlin).
Getzler, E., and R. Pandharipande, 1998, "Virasoro constraints and the Chern classes of the Hodge bundle," Nucl. Phys. B **530**, 701–714.
Gopakumar, R., and C. Vafa, 1998a, "M-theory and topological strings, I," e-print hep-th/9809187.
Gopakumar, R., and C. Vafa, 1998b, "M-theory and topological strings, II," e-print hep-th/9812127.
Gopakumar, R., and C. Vafa, 1999, "On the gauge theory/geometry correspondence," Adv. Theor. Math. Phys. **3**, 1415–1443.
Graber, T., and E. Zaslow, 2002, "Open string Gromov-Witten invariants: Calculations and a mirror 'theorem ,' " Contemp. Math. **310**, 107–121.
Grassi, A., and M. Rossi, 2002, "Large N dualities and transitions in geometry," e-print math.AG/0209044.
Gross, D., 1993, "Two-dimensional QCD as a string theory," Nucl. Phys. B **400**, 161–180.
Gross, D., and W. Taylor, 1993, "Two-dimensional QCD is a string theory," Nucl. Phys. B **400**, 181–210.
Guadagnini, E., 1992, "The universal link polynomial," Int. J. Mod. Phys. A **7**, 877–946.
Guadagnini, E., M. Martellini, and M. Mintchev, 1990, "Wilson lines in Chern-Simons theory and link invariants," Nucl. Phys. B **330**, 575–607.
Harris, J., and I. Morrison, 1998, *Moduli of Curves* (Springer-Verlag, Berlin).
Harvey, R., and H. B. Lawson, 1982, "Calibrated geometries," Acta Math. **148**, 47–157.
Hollowood, T. J., A. Iqbal, and C. Vafa, 2003, "Matrix models, geometric engineering and elliptic genera," e-print hep-th/0310272.
Hori, K., S. Katz, A. Klemm, R. Pandharipande, R. Thomas, C. Vafa, R. Vakil, and E. Zaslow, 2003, *Mirror Symmetry* (American Mathematical Society, Providence, RI).
Iqbal, A., 2002 "All genus topological string amplitudes and 5-brane webs as Feynman diagrams," e-print hep-th/0207114.
Iqbal, A., and A. K. Kashani-Poor, 2003a, "Instanton counting and Chern-Simons theory," Adv. Theor. Math. Phys. **7**, 457–497.
Iqbal, A., and A. K. Kashani-Poor, 2003b, "SU(N) geometries and topological string amplitudes," e-print hep-th/0306032.
Iqbal, A., N. Nekrasov, A. Okounkov, and C. Vafa, 2003, "Quantum foam and topological strings," e-print hep-th/0312022.
Jeffrey, L. C., 1992, "Chern-Simons-Witten invariants of lens spaces and torus bundles, and the semiclassical approximation," Commun. Math. Phys. **147**, 563–604.
Kac, V. G., and J. W. van de Leur, 2003, "The *N*-component KP hierarchy and representation theory," J. Math. Phys. **44**, 3245–3293.
Katz, S., A. Klemm, and C. Vafa, 1997, "Geometric engineering of quantum field theories," Nucl. Phys. B **497**, 173–195.
Katz, S., A. Klemm, and C. Vafa, 1999, "M-theory, topological strings and spinning black holes," Adv. Theor. Math. Phys. **3**, 1445–1537.

Katz, S., and C. C. Liu, 2002, "Enumerative geometry of stable maps with Lagrangian boundary conditions and multiple covers of the disc," Adv. Theor. Math. Phys. **5**, 1–49.

Klemm, A., and E. Zaslow, 2001, "Local mirror symmetry at higher genus," in *Winter School on Mirror Symmetry, Vector Bundles and Lagrangian Submanifolds* (American Mathematical Society, Providence, RI), pp. 183–207.

Kontsevich, M., 1992, "Intersection theory on the moduli space of curves and the matrix Airy function," Commun. Math. Phys. **147**, 1–23.

Kontsevich, M., 1995, "Enumeration of rational curves via torus actions," Prog. Math. **129**, 335–368.

Labastida, J. M. F., 1999, "Chern-Simons theory: Ten years after," in *Trends in Theoretical Physics II*, edited by H. Falomir, R. E. Gamboa Saravi, and F. A. Schaposnik, AIP Conf. Proc. No. 484 (AIP, Woodbury, NY).

Labastida, J. M. F., and P. M. Llatas, 1992, "Topological matter in two-dimensions," Nucl. Phys. B **379**, 220–258.

Labastida, J. M. F., and M. Mariño, 2001, "Polynomial invariants for torus knots and topological strings," Commun. Math. Phys. **217**, 423–449.

Labastida, J. M. F., and M. Mariño, 2002, "A new point of view in the theory of knot and link invariants," J. Knot Theory Ramif. **11**, 173–197.

Labastida, J. M. F., M. Mariño, and C. Vafa, 2000, "Knots, links and branes at large N," J. High Energy Phys. 0011, 007.

Labastida, J. M. F., and A. V. Ramallo, 1989, "Operator formalism for Chern-Simons theories," Phys. Lett. B **227**, 92–102.

Leung, N. C., and C. Vafa, 1998, "Branes and toric geometry," Adv. Theor. Math. Phys. **2**, 91–118.

Li, J., C.-C. M. Liu, K. Liu, and J. Zhou, 2004, "A mathematical theory of the topological vertex," e-print math.AG/0408426.

Li, J., and Y. S. Song, 2002, "Open string instantons and relative stable morphisms," Adv. Theor. Math. Phys. **5**, 67–91.

Lickorish, W. B. R., 1998, *An Introduction to Knot Theory* (Springer-Verlag, Berlin).

Liu, C.-C. M., K. Liu, and J. Zhou, 2003a, "A proof of a conjecture of Mariño-Vafa on Hodge integrals," J. Diff. Geom. **65**, 289–340.

Liu, C.-C. M., K. Liu, and J. Zhou, 2003b, "Mariño-Vafa formula and Hodge integral identities," e-print math.AG/0308015.

Macdonald, I. G., 1995, *Symmetric Functions and Hall Polynomials*, 2nd ed. (Oxford University Press, New York).

Maldacena, J. M., 1998. "The large N limit of superconformal field theories and supergravity," Adv. Theor. Math. Phys. **2**, 231–252.

Mariño, M., 2002a, "Chern-Simons theory, matrix integrals, and perturbative three-manifold invariants," e-print hep-th/0207096.

Mariño, M., 2002b, "Enumerative geometry and knot invariants," e-print hep-th/0210145.

Mariño, M., and G. W. Moore, 1999, "Counting higher genus curves in a Calabi-Yau manifold," Nucl. Phys. B **543**, 592–614.

Mariño, M., and C. Vafa, 2002, "Framed knots at large N," Contemp. Math. **310**, 185–204.

Maulik, D., N. Nekrasov, A. Okounkov, and R. Pandharipande, 2003, "Gromov-Witten theory and Donaldson-Thomas theory," e-print math.AG/0312059.

Mayr, P., 2002, "Summing up open string instantons and $\mathcal{N}=1$ string amplitudes," e-print hep-th/0203237.

Morton, H. R., and S. G. Lukac, 2003, "The HOMFLY polynomial of the decorated Hopf link," J. Knot Theory Ramif. **12**, 395–416.

Nekrasov, N., 2002, "Seiberg-Witten prepotential from instanton counting," e-print hep-th/0206161.

Okounkov, A., and R. Pandharipande, 2004, "Hodge integrals and invariants of the unknot," Geom. Topol. **8**, 675–699.

Okounkov, A., N. Reshetikhin, and C. Vafa, 2003, "Quantum Calabi-Yau and classical crystals," e-print hep-th/0309208.

Okuda, T., and H. Ooguri, 2004, "D-branes and phases on string worldsheet," e-print hep-th/0404101.

Ooguri, H., Y. Oz, and Z. Yin, 1996, "D-branes on Calabi-Yau spaces and their mirrors," Nucl. Phys. B **477**, 407–430.

Ooguri, H., and C. Vafa, 2000, "Knot invariants and topological strings," Nucl. Phys. B **577**, 419–438.

Ooguri, H., and C. Vafa, 2002, "Worldsheet derivation of a large N duality," Nucl. Phys. B **641**, 3–34.

Periwal, V., 1993, "Topological closed string interpretation of Chern-Simons theory," Phys. Rev. Lett. **71**, 1295–1298.

Polchinski, J., 1998, *String Theory*, Vol. II (Cambridge University Press, Cambridge, UK).

Polyakov, A. M., 1988, "Fermi-Bose transmutations induced by gauge fields," Mod. Phys. Lett. A **3**, 325–328.

Prasolov, V. V., and A. B. Sossinsky, 1997, *Knots, Links, Braids and 3-manifolds* (American Mathematical Society, Providence, RI).

Ramadevi, P., and T. Sarkar, 2001, "On link invariants and topological string amplitudes," Nucl. Phys. B **600**, 487–511.

Rozansky, L., 1995, "A large k asymptotics of Witten's invariant of Seifert manifolds," Commun. Math. Phys. **171**, 279–322.

Saulina, N., and C. Vafa, 2004, "D-branes as defects in the Calabi-Yau crystal," e-print hep-th/0404246.

Schwarz, A., 1987, "New topological invariants arising in the theory of quantized fields," Baku International Topological Conference, Abstracts (Part 2) Baku, 1987.

Seiberg, N., and E. Witten, 1994, "Electric-magnetic duality, monopole condensation, and confinement in $\mathcal{N}=2$ supersymmetric Yang-Mills theory," Nucl. Phys. B **426**, 19–52; **430**, 485(E) (1994).

Taubes, C. H., 2001, "Lagrangians for the Gopakumar-Vafa conjecture," Adv. Theor. Math. Phys. **5**, 139–163.

Taylor, W., 2000, "D-brane effective field theory from string field theory," Nucl. Phys. B **585**, 171–192.

Taylor, W., and B. Zwiebach, 2003, "D-branes, tachyons, and string field theory," e-print hep-th/0311017.

't Hooft, G., 1974, "A planar diagram theory for strong interactions," Nucl. Phys. B **72**, 461–473.

Vafa, C., 2001a, "Superstrings and topological strings at large N," J. Math. Phys. **42**, 2798–2817.

Vafa, C., 2001b, "Brane/anti-brane systems and $U(N|M)$ supergroup," e-print hep-th/0101218.

Witten, E., 1986, "Noncommutative geometry and string field theory," Nucl. Phys. B **268**, 253–294.

Witten, E., 1988, "Topological sigma models," Commun. Math. Phys. **118**, 411–449.

Witten, E., 1989, "Quantum field theory and the Jones polynomial," Commun. Math. Phys. **121**, 351–399.

Witten, E., 1990, "On the structure of the topological phase of two-dimensional gravity," Nucl. Phys. B **340**, 281–332.

Witten, E., 1991a, "Two-dimensional gravity and intersection theory on moduli space," Surv. Diff. Geom. **1**, 243–310.

Witten, E., 1991b, "Mirror manifolds and topological field theory," e-print hep-th/9112056.
Witten, E., 1993, "Phases of $\mathcal{N}=2$ theories in two dimensions," Nucl. Phys. B **403**, 159–222.
Witten, E., 1995, "Chern-Simons gauge theory as a string theory," Prog. Math. **133**, 637–678.
Zhou, J., 2003, "Localizations on moduli spaces and free field realizations of Feynman rules," e-print math.AG/0310283.