ADV. THEOR. MATH. PHYS.
Volume 24, Number 8, 1931–2041, 2020

Unification of integrability in supersymmetric gauge theories

Kevin Costello and Junya Yagi

A four-dimensional analog of Chern–Simons theory produces integrable lattice models from Wilson lines and surface operators. We show that this theory describes a quasi-topological sector of maximally supersymmetric Yang–Mills theory in six dimensions, topologically twisted and subjected to an Ω-deformation. By realizing the six-dimensional theory in string theory and applying dualities, we unify various phenomena in which the eight-vertex model and the XYZ spin chain, as well as variants thereof, emerge from supersymmetric gauge theories.

1. Introduction

Over the past decade there has been considerable progress in our understanding of connections between quantum field theories and quantum integrable systems. Many phenomena have been discovered in which structures

of integrable quantum spin chains and integrable lattice models emerge from quantum field theories in diverse spacetime dimensions, in most cases supersymmetric ones.

Among these phenomena, there are several instances where the same family of integrable systems appears. The most notable example is the XXX spin chain and its generalizations the XXZ and XYZ spin chains, or equivalently, the six- and eight-vertex models [1]. These spin chains and lattice models have been found to arise in two-, three- and four-dimensional supersymmetric gauge theories with four supercharges [2, 3], four-, five- and six-dimensional supersymmetric gauge theories with eight supercharges [4–8], three-dimensional $\mathcal{N} = 4$ supersymmetric gauge theories [9, 10], four-dimensional supersymmetric gauge theories in the presence of surface operators [11–17], and a four-dimensional analog of Chern–Simons theory [18–21].

Then a question comes to mind: why does a single family of integrable systems makes appearances in multiple contexts?

In this paper we provide an answer to this question. We argue that these field theory setups are actually different descriptions of one and the same physical system, all related by dualities in string theory.

Another, closely related, aim of the paper is to better understand four-dimensional Chern–Simons theory. This bosonic theory has a fairly direct connection with integrable lattice models, which can be elegantly deduced solely from its topological–holomorphic nature. Yet, this is by far the strangest of the theories listed above. For one thing, it can only be defined on a product $\Sigma \times C$ of two surfaces, with C being either the complex plane $\mathbb{C}$, the punctured complex plane $\mathbb{C}^\times = \mathbb{C} \setminus \{0\}$ or an elliptic curve $E = \mathbb{C}/(\mathbb{Z} + \tau\mathbb{Z})$; the three choices correspond to the three levels of the rational–trigonometric–elliptic hierarchy of integrable systems. Moreover, it has a complex gauge group and a complex action functional.

One of the main results of this paper is that four-dimensional Chern–Simons theory in fact has an origin in six dimensions: it describes a six-dimensional topological–holomorphic theory, subjected to a so-called "Ω-deformation" [22–24]. This six-dimensional theory is a partial topological twist [25] of maximally supersymmetric Yang–Mills theory, and the restriction on the choice of C comes from the requirement for unbroken supercharges. The complex gauge group and the complex action functional naturally arise when the path integral is partially carried out to yield a four-dimensional description.

In turn, the six-dimensional construction allows us to realize four-dimensional Chern–Simons theory and its observables using branes in string theory. Various chains of dualities then relate the brane configuration thus

obtained to different but physically equivalent configurations which realize the other relevant theories, thereby unifying the connections between quantum field theories and the eight-vertex model mentioned above.

Since this paper is somewhat long and at times technical, let us give a brief overview here before proceeding to detailed discussions.

We begin in Section 2 by formulating the six-dimensional topological–holomorphic theory. The theory is defined on a product $M \times C$, and is topological on the four-manifold M and holomorphic on C. For $M = D \times \Sigma$, we may regard the theory as a B-twisted gauge theory [26, 27] on the surface D, with an infinite-dimensional gauge group and infinite-dimensional matter representations. It turns out that in this two-dimensional description, the theory has a superpotential which coincides with the action of four-dimensional Chern–Simons theory.

In Section 3, we turn to general B-twisted gauge theories and explain how to introduce Ω-deformations to these theories [28, 29]. By localization of the path integral, we show that when the spacetime D is $\mathbb{R}^2$, the quasi-topological sector of an Ω-deformed B-twisted gauge theory is equivalent to a zero-dimensional gauge theory with complex gauge group, whose action is given by the superpotential of the two-dimensional theory [28–30]. The integration domain of the path integral for this zero-dimensional theory consists of the gradient flow trajectories generated by the superpotential, terminating on a Lagrangian submanifold chosen in a relevant moduli space.

Then we apply this result to the six-dimensional topological–holomorphic theory, viewing it as a B-twisted gauge theory. We are immediately led to the conclusion that the topological–holomorphic theory on $\mathbb{R}^2 \times \Sigma \times C$, with an Ω-deformation on $\mathbb{R}^2$, is equivalent to four-dimensional Chern–Simons theory on $\Sigma \times C$. This is done in Section 4.

Section 5 is devoted to discussions on the relations between four-dimensional Chern–Simons theory and integrable lattice models in the case when $C = E$. We explain how a lattice model [31, 32] whose Boltzmann weights are given by Felder's dynamical R-matrix [33, 34] arises from a lattice of Wilson lines, and how certain surface operators transform this R-matrix to the Baxter–Belavin R-matrix [35–37] for the eight-vertex model and its $\mathfrak{sl}_N$ generalization. We also use these surface operators to define two kinds of L-operators, which may be thought of as R-matrices associated with a pair of finite- and infinite-dimensional representations. The section ends with some discussions on framing anomaly and junctions of Wilson lines; these lie outside the main line of argument and are not strictly necessary for understanding of the rest of the paper.

In Section 6, we present a string theory realization of the Ω-deformed six-dimensional topological–holomorphic theory. This realization involves a stack of D5-branes placed in a background with a nontrivial Ramond–Ramond (RR) two-form field. Wilson lines are created by fundamental strings ending on the D5-branes, whereas surface operators are produced by D3-branes forming bound states with the D5-branes. Applying string dualities, we map this brane configuration to those realizing brane tiling [38, 39] and class-$\mathcal{S}_k$ [14, 40, 41] theories, linear quiver theories [42], and theories related to the cotangent bundles of partial flag manifolds. In each dual picture we identify how the structures of lattice models and spin chains arise.

There are many directions for future research. One important question which we hope this paper will shed some light on is the origin of the chiral Potts model and its higher genus curve for spectral parameters. The mysterious coincidence between the chiral Potts model and magnetic monopoles, pointed out by Atiyah [43] in 1990, hints that we are on the right track. Indeed, we have necessary ingredients in our construction: monopoles create surface operators, and crossings of surface operators produce [44–46] a variant of the Bazhanov–Sergeev model [47, 48] which is known to reduce to the chiral Potts model in a special limit. It is plausible that the higher genus curve emerges in low energy physics as a geometric object, in a way similar to how the Seiberg–Witten curve does when a D4–NS5 brane configuration is lifted to M-theory [42].

Finally, we remark that another string theory construction of four-dimensional Chern–Simons theory has been proposed recently in [49]. Their construction appears to be related to a T-dual version of ours. Also, a string theory realization was discussed by Nikita Nekrasov in his talk on his joint work with Samson Shatashvili and Mina Aganagic at the conference String–Math 2017, where we also announced our results.

2. Six-dimensional topological–holomorphic theory

In this section we formulate the six-dimensional topological–holomorphic theory, from which four-dimensional Chern–Simons theory arises via an Ω-deformation. After explaining the construction, we reformulate this theory as a two-dimensional gauge theory, in a form more suited for the application of the Ω-deformation.

2.1. $\mathcal{N} = (1,1)$ super Yang–Mills theory

The topological–holomorphic theory is defined as a topological twist of $\mathcal{N} = (1,1)$ super Yang–Mills theory in six dimensions, which in turn can be constructed from super Yang–Mills theory in ten dimensions by dimensional reduction. So let us quickly review these super Yang–Mills theories. We mainly follow the convention of [50].

To describe spinors in ten dimensions, we use the gamma matrices Γ_I, $I = 0, \ldots, 9$, obeying the anticommutation relation

$$\{\Gamma_I, \Gamma_J\} = 2\eta_{IJ}\,, \tag{1}$$

where $\eta = -(\mathrm{d}x^0)^2 + (\mathrm{d}x^1)^2 + \cdots + (\mathrm{d}x^9)^2$ is the ten-dimensional Minkowski metric. They can be chosen to be real 32×32 matrices. We let $\Gamma_{I_1 \ldots I_k}$ be the matrix that equals $\Gamma_{I_1} \cdots \Gamma_{I_k}$ if $I_1, \ldots, I_k$ are all different and vanishes otherwise.

The generators of the Lorentz group $\mathrm{Spin}(9,1)$ are represented on $\mathbb{R}^{32}$ by the matrices Γ_{IJ}. The chirality operator $\Gamma_{0123456789}$ squares to 1 and anticommutes with Γ_I. Its eigenspaces therefore furnish irreducible spinor representations of $\mathrm{Spin}(9,1)$ on which the chirality operator acts as multiplication by $+1$ or -1; let S^+ and S^- denote the space of positive chirality spinors and that of negative chirality spinors, respectively. There is a charge conjugation matrix C such that $C\Gamma_{IJ}C^{-1} = -\Gamma_{IJ}^T$, and the map

$$\alpha \mapsto \bar{\alpha} = \alpha^T C \tag{2}$$

sends $\alpha \in S^\pm$ to its dual $\bar{\alpha} \in (S^\pm)^*$. For α, $\beta \in S^\pm$, the product $\bar{\alpha}\Gamma_{I_1 \ldots I_k}\beta$ transforms under $\mathrm{Spin}(9,1)$ like the corresponding component of a k-form.

The fields of ten-dimensional super Yang–Mills theory with gauge group G are the gauge field A and a fermionic field Ψ, the latter being a positive chirality spinor in the adjoint representation. More precisely, A is a connection of a principal G-bundle P over Minkowski spacetime $\mathbb{R}^{9,1}$, and Ψ is a section of $S^+ \otimes \mathrm{ad}(P)$, where $\mathrm{ad}(P)$ is the adjoint bundle of P. The theory is governed by the action

$$-\frac{1}{e^2}\int \mathrm{d}^{10}x\ \mathrm{Tr}\left(\frac{1}{2}F^{IJ}F_{IJ} - \mathrm{i}\overline{\Psi}\Gamma^I D_I \Psi\right). \tag{3}$$

Here e is the gauge coupling, $F = \mathrm{d}A + A \wedge A$ is the field strength and $D = \mathrm{d} + A$ is the covariant derivative. The symbol Tr denotes an invariant symmetric bilinear form on the Lie algebra $\mathfrak{g}$ of the compact Lie group

G. We have chosen A in such a way that it is antihermitian in a unitary representation of G, and Tr to be negative definite. We pick generators T_a, $a = 1, \dots, \dim G$, of $\mathfrak{g}$ such that $\mathrm{Tr}(T_a T_b) = -\delta_{ab}$ so that we can write, for example, $A_I = \sum_{a=1}^{\dim G} A_I^a T_a$ with real coefficients A_I^a.

The action (3) is invariant under the supersymmetry variation

$$\delta_\epsilon A_I = \mathrm{i}\bar{\epsilon}\Gamma_I \Psi\,, \qquad \delta_\epsilon \Psi = \frac{1}{2} F_{IJ}\Gamma^{IJ}\epsilon\,, \tag{4}$$

whose parameter ϵ is a constant spinor of positive chirality. Hence, the theory has sixteen supercharges, which is the maximum amount of supercharges for theories without gravity. Let Q_ϵ be the supercharge generating the above transformation. Up to equations of motion, the supercharges obey the anticommutation relation

$$\{Q_\epsilon, Q_\eta\} = \bar{\epsilon}\Gamma^I \eta P_I\,, \tag{5}$$

where the momentum P_I generates translations in the x^I-direction.

Now, let us demand the fields to be independent of the coordinates x^6, x^7, x^8, x^9. Then we obtain a six-dimensional gauge theory, which is $\mathcal{N} = (1,1)$ super Yang–Mills theory.

Under the splitting of $\mathbb{R}^{9,1}$ into $\mathbb{R}^{5,1}$ and $\mathbb{R}^4$, the ten-dimensional Lorentz group $\mathrm{Spin}(9,1)$ decomposes into the product $\mathrm{Spin}(5,1) \times \mathrm{Spin}(4)_R$. The first factor is the six-dimensional Lorentz group, while the second is the R-symmetry group of the six-dimensional theory. The ten-dimensional gauge field $\sum_{I=0}^{9} A_I \mathrm{d}x^I$ descends to a gauge field $\sum_{I=0}^{5} A_I \mathrm{d}x^I$ and four adjoint scalar fields

$$\phi_\mu\,, \quad \mu = 0, \dots, 3\,, \tag{6}$$

in the six-dimensional theory, where the latter come from the components $A_{\mu+6}$ and transform in the vector representation $\mathbf{4}$ of $\mathrm{Spin}(4)_R$. Upon the dimensional reduction the bosonic part of the action becomes

$$-\frac{1}{e^2}\int \mathrm{d}^6 x\, \mathrm{Tr}\left(\frac{1}{2}\sum_{I,J=0}^{5} F^{IJ}F_{IJ} + \sum_{I=0}^{5}\sum_{\mu=0}^{3} D^I\phi^\mu D_I\phi_\mu + \frac{1}{2}\sum_{\mu=0}^{3}[\phi^\mu,\phi^\nu][\phi_\mu,\phi_\nu]\right). \tag{7}$$

To understand what the fermion Ψ becomes in six dimensions, we recall that $\mathrm{Spin}(4)$ is isomorphic to $\mathrm{SU}(2) \times \mathrm{SU}(2)$. In the case of $\mathrm{Spin}(4)_R$, we

can take the first SU(2) factor to be generated by

$$\frac{1}{2}(\Gamma_{67}+\Gamma_{89})\,,\qquad \frac{1}{2}(\Gamma_{68}+\Gamma_{97})\,,\qquad -\frac{1}{2}(\Gamma_{69}+\Gamma_{78}) \tag{8}$$

and the second to be generated by

$$\frac{1}{2}(\Gamma_{67}-\Gamma_{89})\,,\qquad \frac{1}{2}(\Gamma_{68}-\Gamma_{97})\,,\qquad \frac{1}{2}(\Gamma_{69}-\Gamma_{78})\,. \tag{9}$$

Acting on these generators with the chirality operator Γ_{6789} for $\mathrm{Spin}(4)_R$, we see that the irreducible spinor representations of $\mathrm{Spin}(4)_R$ of positive and negative chirality are the representations $(\mathbf{1},\mathbf{2})$ and $(\mathbf{2},\mathbf{1})$ of $\mathrm{SU}(2)\times\mathrm{SU}(2)$, respectively. The chirality operator for $\mathrm{Spin}(5,1)$ is Γ_{012345}, so S^+ of $\mathrm{Spin}(9,1)$ decomposes with respect to $\mathrm{Spin}(5,1)\times\mathrm{Spin}(4)_R$ as

$$\big(\mathbf{4}_+,(\mathbf{1},\mathbf{2})\big)\oplus\big(\mathbf{4}_-,(\mathbf{2},\mathbf{1})\big)\,, \tag{10}$$

where $\mathbf{4}_\pm$ are the spinor representations of $\mathrm{Spin}(5,1)$ with the chirality indicated by the subscripts. Thus, in six dimensions, Ψ becomes two sets of spinors which have opposite chirality and are doublets of different SU(2) factors of $\mathrm{Spin}(4)_R$.

The six-dimensional super Yang–Mills theory inherits the sixteen supercharges from ten dimensions. Since the parameter ϵ of supersymmetry variations transforms in the same way as Ψ does, the theory has two SU(2) doublets of supercharges with opposite chirality, generating $\mathcal{N}=(1,1)$ supersymmetry in six dimensions.

2.2. Topological–holomorphic theory

The topological–holomorphic theory we are going to construct is a topological twist of the Euclidean version of six-dimensional $\mathcal{N}=(1,1)$ super Yang–Mills theory, and can be defined on a product $M\times C$, with M being a four-manifold and C either $\mathbb{C}$, $\mathbb{C}^\times$ or $\mathbb{C}/(\mathbb{Z}+\tau\mathbb{Z})$. The relevant topological twist is essentially the GL-twist of $\mathcal{N}=4$ super Yang–Mills theory in four dimensions, which plays an important role in a gauge theoretic approach to the geometric Langlands duality [50].

The Euclidean theory is obtained from the Lorentzian one by the Wick rotation $x^0\mapsto -ix^0$. Correspondingly, we get gamma matrices in Euclidean signature by making the replacement $\Gamma_0\mapsto \mathrm{i}\Gamma_0$ in those in Lorentzian signature.

For a moment, suppose that M is a spin manifold. The structure group of the spinor bundle over $M \times C$ is $\mathrm{Spin}(4)_M \times \mathrm{Spin}(2)_C$. To implement the topological twist in question, we turn on a background gauge field for $\mathrm{Spin}(4)_R$ whose value is equal to the spin connection of M, and interpret the diagonal subgroup $\mathrm{Spin}(4)'_M$ of $\mathrm{Spin}(4)_M \times \mathrm{Spin}(4)_R$ as a new rotation group on M. Under $\mathrm{Spin}(4)'_M$, the scalars ϕ_μ transform as the components of a one-form

$$\phi = \phi_\mu \mathrm{d}x^\mu \tag{11}$$

since they originated from the components of the ten-dimensional gauge field along the directions rotated by $\mathrm{Spin}(4)_R$.

The transformation properties of the fermions can be identified as follows. In Minkowski signature, we can take the chirality operators for $\mathrm{Spin}(3,1)$, $\mathrm{Spin}(2)$ and $\mathrm{Spin}(5,1)$ to be $-\mathrm{i}\Gamma_{0123}$, $-\mathrm{i}\Gamma_{45}$ and $\Gamma_{012345} = -(-\mathrm{i}\Gamma_{0123})(-\mathrm{i}\Gamma_{45})$, respectively. Then, $\mathbf{4}_+$ of $\mathrm{Spin}(5,1)$ transforms under the subgroup $\mathrm{Spin}(3,1) \times \mathrm{Spin}(2)$ as $\mathbf{2}_+^{-1} \oplus \mathbf{2}_-^{1}$, while $\mathbf{4}_-$ transforms as $\mathbf{2}_+^{1} \oplus \mathbf{2}_-^{-1}$. Here the superscripts indicate the charges under $\mathrm{Spin}(2) \cong \mathrm{U}(1)$, measured by $-\mathrm{i}\Gamma_{45}$. In Euclidean signature, $\mathrm{Spin}(3,1)$ is replaced with $\mathrm{Spin}(4) \cong \mathrm{SU}(2) \times \mathrm{SU}(2)$, and $\mathbf{2}_+$ becomes $(\mathbf{1},\mathbf{2})$ and $\mathbf{2}_-$ becomes $(\mathbf{2},\mathbf{1})$. Using the decomposition $\mathbf{2} \otimes \mathbf{2} = \mathbf{1} \oplus \mathbf{3}$ of $\mathrm{SU}(2)$, we find that the fermions transform under $\mathrm{Spin}(4)'_M \times \mathrm{Spin}(2)_C$ as

$$2(\mathbf{1},\mathbf{1})^{-1} \oplus (\mathbf{1},\mathbf{3})^{-1} \oplus (\mathbf{3},\mathbf{1})^{-1} \oplus 2(\mathbf{2},\mathbf{2})^{1}\,. \tag{12}$$

The first three summands represent two scalars and one two-form on M, transforming as negative chirality spinors on C:

$$\xi,\, \xi' \in \Gamma\big(\Lambda^0_M \otimes \overline{K}_C^{-1/2} \otimes \mathrm{ad}(P)\big)\,, \quad \chi \in \Gamma\big(\Lambda^2_M \otimes \overline{K}_C^{-1/2} \otimes \mathrm{ad}(P)\big)\,. \tag{13}$$

The last summand gives two one-forms on M which are positive chirality spinors on C:

$$\psi\,, \psi' \in \Gamma\big(\Lambda^1_M \otimes \overline{K}_C^{1/2} \otimes \mathrm{ad}(P)\big)\,. \tag{14}$$

Here Λ^p_M is the bundle of p-forms on M and $\overline{K}_C^{\pm 1/2}$ are the bundles of spinors on C with positive and negative chirality, all pulled back to $M \times C$. We have used the same symbol P as in the ten-dimensional case to denote the gauge bundle.

Since the twisted theory does not contain any spinors on M, at this point we can relax the assumption that M is spin. The twisted theory can be defined on any four-manifold M.

Looking at the transformation properties of the fermions, we see that the twisted theory has two supercharges that are invariant under $\mathrm{Spin}(4)'_M$. They generate supersymmetry transformations whose parameters are constant scalars on M and constant spinors on C, and as such are present for any choice of M. In contrast, the other supercharges are broken unless M admits covariantly constant one-forms or two-forms.

Let us describe the supercharges that are scalars on M more explicitly. The relevant supersymmetry parameters are annihilated by the generators of $\mathrm{Spin}(4)'_M$:

$$(\Gamma_{\mu\nu} + \Gamma_{\mu+6,\nu+6})\epsilon = 0\,, \quad \mu,\, \nu = 0,\, \ldots,\, 3\,. \tag{15}$$

These equations can be rewritten as

$$\epsilon = \Gamma_{\mu\nu\mu+6,\nu+6}\epsilon\,, \tag{16}$$

and impose three independent constraints on ϵ. Each of them reduces the dimension of the parameter space by half, so there are $16 \times (1/2)^3 = 2$ independent solutions, as expected.

We can single out a supercharge by further demanding

$$\epsilon = -\mathrm{i}\Gamma_{\mu,\mu+6}\epsilon\,. \tag{17}$$

These equations are compatible with the condition (16) and the chirality condition

$$\mathrm{i}\Gamma_{0123456789}\epsilon = \epsilon \tag{18}$$

in Euclidean signature since $\Gamma_{\mu,\mu+6}$ commute with $\Gamma_{0123456789}$ and $\Gamma_{\mu\nu\mu+6,\nu+6}$. We are left with a unique solution up to rescaling, and call the corresponding supercharge Q. Similarly, imposing the condition

$$\epsilon = \mathrm{i}\Gamma_{\mu,\mu+6}\epsilon \tag{19}$$

we obtain another supercharge Q'.

An important property of the supercharges thus defined is that they square to zero:

$$Q^2 = (Q')^2 = 0\,. \tag{20}$$

It is clear that P_μ cannot appear in Q^2 or $(Q')^2$ because these supercharges are $\mathrm{Spin}(4)'_M$-invariant. To see that P_I for any $I = 0, \ldots, 9$ makes no appearance either, say in Q^2, we pick μ such that $I \neq \mu,\ \mu+6$ and note $\mathrm{i}\bar\epsilon\Gamma_{\mu,\mu+6} = -\mathrm{i}\epsilon^T\Gamma^T_{\mu,\mu+6}C = \bar\epsilon$. So we have

$$\bar\epsilon\Gamma_I\epsilon = \bar\epsilon\Gamma_I(-\mathrm{i}\Gamma_{\mu,\mu+6}\epsilon) = -\mathrm{i}\bar\epsilon\Gamma_{\mu,\mu+6}\Gamma_I\epsilon = -\bar\epsilon\Gamma_I\epsilon \tag{21}$$

and $Q^2 = \bar\epsilon\Gamma^I\epsilon P_I = 0$.

From the constraints (16) and the chirality condition (18), it follows

$$\epsilon = \Gamma_{0167}\Gamma_{2389}\epsilon = \mathrm{i}\Gamma_{45}\epsilon\,. \tag{22}$$

Since the parameters ϵ under consideration have charge -1 with respect to the $\mathrm{U}(1)$ symmetry generated by $-\mathrm{i}\Gamma_{45}$, the corresponding supercharges have charge $+1$. As can be seen from the transformation property of $\bar\alpha(\Gamma_4 - \mathrm{i}\Gamma_5)\beta$, the linear combination $P_4 - \mathrm{i}P_5$ has charge 2 and is the only translation generator with that charge. Hence, $\{Q, Q'\} \propto P_4 - \mathrm{i}P_5$. Introducing the complex coordinate

$$z = \frac{1}{2}(x^4 - \mathrm{i}x^5)\,, \tag{23}$$

we normalize the supercharges in such a way that

$$\{Q, Q'\} = P_{\bar z}\,. \tag{24}$$

Comparing the equation $(\Gamma_4 - i\Gamma_5)\epsilon = 0$ with the formula (4) for supersymmetry variations, we see that $A_{\bar z}$ is invariant under the action of Q and Q'. The extra condition (17) says $(\Gamma_\mu + \mathrm{i}\Gamma_{\mu+6})\epsilon = 0$, meaning that $A_\mu + \mathrm{i}A_{\mu+6}$ is annihilated by Q. The twisted theory therefore has the Q-invariant partial connection

$$\mathcal{A} = \mathcal{A}_\mu \mathrm{d}x^\mu + A_{\bar z}\mathrm{d}\bar z\,, \qquad \mathcal{A}_\mu = A_\mu + \mathrm{i}\phi_\mu\,. \tag{25}$$

By the same token, if we define

$$\overline{\mathcal{A}} = (A_\mu - \mathrm{i}\phi_\mu)\mathrm{d}x^\mu + A_z\mathrm{d}z\,, \tag{26}$$

then $\overline{\mathcal{A}}_\mu \mathrm{d}x^\mu + A_{\bar{z}} \mathrm{d}\bar{z}$ is Q'-invariant. We write $\mathcal{D}$ and $\mathcal{F}$ for the covariant derivative and curvature of $\mathcal{A} + A_z \mathrm{d}z$, and $\overline{\mathcal{D}}$ and $\overline{\mathcal{F}}$ for those of $\overline{\mathcal{A}} + A_{\bar{z}} \mathrm{d}\bar{z}$.

We can readily write down the action of Q on the rest of the fields. From the Q-invariance of $\mathcal{A}$, the transformation properties of the fields under $\mathrm{Spin}(4)'_M \times \mathrm{Spin}(2)_C$, and the fact that the supersymmetry variation of a fermion is a linear combination of the field strength F_{IJ} in ten dimensions, we deduce that the Q-action can be written as

$$
\begin{aligned}
\delta \mathcal{A}_\mu &= 0\,, & \delta \overline{\mathcal{A}}_\mu &= \psi_\mu\,,\\
\delta A_z &= \xi\,, & \delta A_{\bar{z}} &= 0\,,\\
\delta \xi &= 0\,, & \delta \xi' &= \mathsf{P}\,,\\
\delta \psi_\mu &= 0\,, & \delta \psi'_\mu &= \mathcal{F}_{\mu\bar{z}}\,,\\
\delta \mathsf{P} &= 0\,, & \delta \chi_{\mu\nu} &= \mathcal{F}_{\mu\nu}\,.
\end{aligned}
\tag{27}
$$

We have introduced an auxiliary bosonic scalar P in order to realize the correct supersymmetry algebra off-shell.

The variation δ' under the action of Q' can be identified with the help of relation (24). Setting $\{\delta, \delta'\}\mathcal{A}_\mu = \mathcal{F}_{\bar{z}\mu}$ leads to $\delta\delta'\mathcal{A}_\mu = -\delta\psi'_\mu$, so we have $\delta'\mathcal{A}_\mu = -\psi'_\mu$ up to Q-invariant terms which we can absorb in the definition of ψ'_μ. Likewise, the relation $\{\delta, \delta'\}\overline{\mathcal{A}}_\mu = \overline{\mathcal{F}}_{\bar{z}\mu}$ implies $\delta'\psi_\mu = -\overline{\mathcal{F}}_{\mu\bar{z}}$. Redefining ξ' and P if necessary, we can set $\delta' A_z = -\xi'$. Then, $\{\delta, \delta'\}A_z = F_{\bar{z}z}$ gives $\delta'\xi = \mathsf{P} - F_{z\bar{z}}$.

To have $\{\delta, \delta'\}\chi = D_{\bar{z}}\chi$, we need to use the equation of motion for χ. Let us postulate that the part of the action that contains χ is given by

$$
\begin{aligned}
S_1 &= \frac{1}{e^2}\int_{M\times C} \mathrm{d}^2 z\, \mathrm{Tr}\Big(-\delta\big(\chi \wedge \star_M \overline{\mathcal{F}}\big) + \chi \wedge \mathcal{D}\psi' + \frac{1}{2}\chi \wedge D_{\bar{z}}\chi\Big)\\
&= \frac{1}{e^2}\int_{M\times C} \mathrm{d}^2 z\, \mathrm{Tr}\Big(-\mathcal{F} \wedge \star_M \overline{\mathcal{F}} + \chi \wedge \star_M \overline{\mathcal{D}}\psi\\
&\qquad\qquad + \chi \wedge \mathcal{D}\psi' + \frac{1}{2}\chi \wedge D_{\bar{z}}\chi\Big)\,,
\end{aligned}
\tag{28}
$$

where $\mathrm{d}^2 z = -2\mathrm{i}\,\mathrm{d}z \wedge \mathrm{d}\bar{z} = \mathrm{d}x^4 \wedge \mathrm{d}x^5$. Note that we have chosen a metric g_M on M to define the Hodge star $\star_M$ on M. The above expression is Q-invariant thanks to the Bianchi identity $\mathcal{D}\mathcal{F} = 0$, provided that boundary terms do not arise in the Q-variation. The equation of motion for χ derived from this action is

$$
D_{\bar{z}}\chi = -\delta\big(\star_M \overline{\mathcal{F}}\big) + \delta'\delta\chi\,. \tag{29}
$$

Equating the left-hand side with $\{\delta, \delta'\}\chi$, we find $\delta'\chi = -\star_M \overline{\mathcal{F}}$ up to Q-invariant terms. Because we did not redefine χ, its supersymmetry variation is still given by a linear combination of F_{IJ}, and the Q-invariant terms, if exist, must be constructed from $\mathcal{F}$. Such terms are not compatible with $\delta'^2\chi = 0$ and should be absent (unless we are willing to use other equations of motion). The Q'-variations of the remaining fields can be fixed by the requirement $\delta'^2 = 0$.

We have thus obtained the following formula for the supersymmetry transformation generated by Q':

$$\begin{aligned} \delta'\mathcal{A}_\mu &= -\psi'_\mu\,, & \delta'\overline{\mathcal{A}}_\mu &= 0\,,\\ \delta' A_z &= -\xi'\,, & \delta' A_{\bar z} &= 0\,,\\ \delta'\xi &= \mathsf{P} - F_{z\bar z}\,, & \delta'\xi' &= 0\,,\\ \delta'\psi_\mu &= -\overline{\mathcal{F}}_{\mu\bar z}\,, & \delta'\psi'_\mu &= 0\,,\\ \delta'\mathsf{P} &= D_{\bar z}\xi'\,, & \delta'\chi_{\mu\nu} &= -(\star_M \overline{\mathcal{F}})_{\mu\nu}\,. \end{aligned} \tag{30}$$

With these transformation rules, we can write S_1 as

$$S_1 = \frac{1}{e^2}\int_{M\times C} \mathrm{d}^2 z\, \mathrm{Tr}\left(\delta'(\mathcal{F}\wedge\chi) + \chi\wedge\star_M \overline{\mathcal{D}}\psi + \frac{1}{2}\chi\wedge D_{\bar z}\chi\right). \tag{31}$$

This is Q'-invariant, again up to boundary contributions. Hence, S_1 is invariant under both Q and Q'.

The rest of the action of the twisted theory is

$$\begin{aligned} S_2 &= \frac{1}{e^2}\int_{M\times C}\sqrt{g}\,\mathrm{d}^6x\,\delta\delta'\,\mathrm{Tr}(-\xi\xi' + 2\mathrm{i}F_{\mu z}\phi^\mu)\\ &= \frac{1}{e^2}\int_{M\times C}\sqrt{g}\,\mathrm{d}^6x\,\delta\,\mathrm{Tr}\big((-\mathsf{P} + F_{z\bar z} + 2\mathrm{i}D^\mu\phi_\mu)\xi' - \overline{\mathcal{F}}_{\mu z}\psi'^\mu\big)\,, \end{aligned} \tag{32}$$

where we have ignored boundary terms in going to the last expression. To define the volume form $\sqrt{g}\,\mathrm{d}^6x$ we have endowed C with the metric $g_C = (\mathrm{d}x^4)^2 + (\mathrm{d}x^5)^2$; the total metric on $M\times C$ is $g = g_M \oplus g_C$ and we have $g_{z\bar z} = g_{\bar z z} = 2$. This action is manifestly Q- and Q'-invariant. Explicitly, we have

$$\begin{aligned} S_2 = \frac{1}{e^2}\int_{M\times C}\sqrt{g}\,\mathrm{d}^6x\,\mathrm{Tr}\big(&-\mathsf{P}(\mathsf{P} - F_{z\bar z} - 2\mathrm{i}D^\mu\phi_\mu) - 2\overline{\mathcal{F}}^{\mu\bar z}\mathcal{F}_{\mu\bar z}\\ &+ \xi' D_{\bar z}\xi + \xi'\mathcal{D}_\mu\psi^\mu + \psi'_\mu\overline{\mathcal{D}}^\mu\xi + D_z\psi^\mu\psi'_\mu\big)\,. \end{aligned} \tag{33}$$

The bosonic part of the full action $S_1 + S_2$ is

$$
\begin{aligned}
(34)\quad & \frac{1}{e^2}\int_{M\times C}\sqrt{g}\,\mathrm{d}^6x\,\mathrm{Tr}\biggl(-\frac{1}{2}\overline{\mathcal{F}}^{\mu\nu}\mathcal{F}_{\mu\nu}-\mathsf{P}(\mathsf{P}-F_{z\bar z}-2\mathrm{i}D^\mu\phi_\mu)-2\overline{\mathcal{F}}^{\mu\bar z}\mathcal{F}_{\mu\bar z}\biggr)\\
&\quad=\frac{1}{e^2}\int_{M\times C}\sqrt{g}\,\mathrm{d}^6x\,\mathrm{Tr}\biggl(-\Bigl(\mathsf{P}-\frac{1}{2}F_{z\bar z}-\mathrm{i}D^\mu\phi_\mu\Bigr)^2-\frac{1}{2}F^{\mu\nu}F_{\mu\nu}\\
&\qquad-2F^{\mu z}F_{\mu z}-F^{z\bar z}F_{z\bar z}-D^\mu\phi^\nu D_\mu\phi_\nu-2D^z\phi^\mu D_z\phi_\mu\\
&\qquad-\frac{1}{2}[\phi^\mu,\phi^\nu][\phi_\mu,\phi_\nu]-R_{\mu\nu}\phi^\mu\phi^\nu\biggr),
\end{aligned}
$$

where $R_{\mu\nu}$ is the Ricci curvature of g. For $M=\mathbb{R}^4$, it reproduces the bosonic part (7) of the action for $\mathcal{N}=(1,1)$ super Yang–Mills theory.

Now that we have constructed a theory with two supercharges that square to zero, let us pick one of them, say Q, and consider the Q-invariant sector of the theory. The correlation function of a Q-exact operator vanishes because in the path integral representation it is the integral of a "total derivative" over an infinite-dimensional supermanifold. Therefore, the correlation function of a Q-invariant operator depends only on the Q-cohomology class of that operator. We can also define the Q-cohomology of states. This is a module over the Q-cohomology of operators, and the partition function with Q-closed states specified on the boundary components of spacetime depends only on the Q-cohomology classes of those states. (More generally, correlation functions of Q-closed operators with Q-closed states have a similar property.)

Since the dependence of the action on the metric of M is completely buried in the Q-exact part, the theory becomes topological on M once we restrict it to the Q-invariant sector. Similarly, the Q-invariant sector of the theory depends on the complex structure of C but not on the metric. The anticommutation relation (24) shows that $P_{\bar z}=0$ in the Q-cohomology, so correlation functions of Q-closed operators supported at points on C vary holomorphically on C. In this sense, the twisted theory is a topological–holomorphic theory on $M\times C$.

An example of a Q-closed operator is a Wilson line constructed from $\mathcal{A}$, supported along a closed curve $K\subset M$ and a point $z\in C$:

$$
(35)\qquad \mathrm{Tr}_V\,P\exp\biggl(\oint_{K\times\{z\}}\mathcal{A}\biggr).
$$

The trace is taken in some representation $G_{\mathbb{C}} \to \mathrm{GL}(V)$ of the complexification $G_{\mathbb{C}}$ of G. Such Wilson lines form one of the two classes of observables from which we construct integrable lattice models.

2.3. Two-dimensional formulation

Suppose that C is an elliptic curve. If we make C very small and discard the Kaluza–Klein modes, the topological–holomorphic theory on $M \times C$ reduces to a topological theory on M. This is the GL-twisted $\mathcal{N} = 4$ super Yang–Mills theory [50]. We can further take M to be the product of a surface D and a torus, and make the torus very small. Then, after discarding the Kaluza–Klein modes, we obtain a topologically twisted $\mathcal{N} = (8, 8)$ super Yang–Mills theory on D.

In this series of reduction from six to two dimensions, we could as well keep all Kaluza–Klein modes. If we chose to do so, we would end up with a formulation of the six-dimensional topological–holomorphic theory as a two-dimensional gauge theory. Let us describe this two-dimensional formulation concretely, as we will use it later when we introduce an Ω-deformation of the theory.

2.3.1. Two-dimensional supersymmetry. Recall that the topological twist of the six-dimensional theory replaces the generators $\Gamma_{\mu\nu}$ of the rotation group Spin(6) with $\Gamma_{\mu\nu} + \Gamma_{\mu+6,\nu+6}$. The supercharges Q, Q' are characterized by three conditions on the parameter ϵ of supersymmetry transformation:

$$(\Gamma_{01} + \Gamma_{67})\epsilon = (\Gamma_{12} + \Gamma_{78})\epsilon = (\Gamma_{23} + \Gamma_{89})\epsilon = 0\,. \tag{36}$$

Requiring the additional condition

$$\epsilon = -\mathrm{i}\Gamma_{39}\epsilon \tag{37}$$

then picks out the supercharge Q used to define the topological–holomorphic theory. To describe this procedure in two-dimensional terms, let us impose the above conditions in a different order.

First, we demand $(\Gamma_{23} + \Gamma_{89})\epsilon = (1 + \mathrm{i}\Gamma_{39})\epsilon = 0$. These equations have four independent solutions that are eigenvectors of the two-dimensional chirality operator $-\mathrm{i}\Gamma_{01}$. The action of Γ_{06} leaves the space of solutions invariant but changes chirality, so there are equal number of positive and negative chirality solutions. The corresponding supercharges generate $\mathcal{N} = (2, 2)$ supersymmetry in two dimensions.

Next, we impose $(\Gamma_{01}+\Gamma_{67})\epsilon=0$, which reduces the number of independent solutions to two. There are two U(1) R-symmetries that rotate the supercharges of $\mathcal{N}=(2,2)$ supersymmetry, $\mathrm{U}(1)_V$ generated by $-\mathrm{i}\Gamma_{45}$ and $\mathrm{U}(1)_A$ generated by $-\mathrm{i}\Gamma_{67}$. This condition means that we twist the two-dimensional rotation group $\mathrm{U}(1)_D$ by replacing it with the diagonal subgroup $\mathrm{U}(1)'_D$ of $\mathrm{U}(1)_D\times\mathrm{U}(1)_A$, and keep only those supercharges that are scalars under $\mathrm{U}(1)'_D$. The R-symmetry $\mathrm{U}(1)_A$ used in this twisting acts on the scalars ϕ_6, ϕ_7, which are part of the Q-invariant connection $\sum_{i=0}^{1}\mathcal{A}_i\mathrm{d}x^i$ in two dimensions and belong to the vector multiplet of $\mathcal{N}=(2,2)$ supersymmetry. It is known as the axial U(1) R-symmetry, and the topological twist with respect to it is called the *B-twist* [26, 27]. Since $\epsilon=\mathrm{i}\Gamma_{45}\epsilon$, the scalar supercharges have charge 1 under the other R-symmetry $\mathrm{U}(1)_V$, referred to as the vector U(1) R-symmetry. The topological twist using $\mathrm{U}(1)_V$ is called the A-twist.

Finally, the condition $(\Gamma_{12}+\Gamma_{78})\epsilon=0$ picks out a particular linear combination of the scalar supercharges, which we have been calling Q. We can choose another scalar supercharge $\widetilde{Q}$ (which is different from Q') such that Q and $\widetilde{Q}$ obey the relations

$$Q^2=\widetilde{Q}^2=\{Q,\widetilde{Q}\}=0 \tag{38}$$

up to central charges. These supercharges do not have definite chirality since the last condition is not compatible with the chirality condition $\epsilon=\pm\mathrm{i}\Gamma_{01}\epsilon$.

2.3.2. B-twisted gauge theory. Our task is therefore to describe the six-dimensional topological–holomorphic theory as a B-twisted gauge theory in two dimensions. To this purpose we briefly review the construction of the latter theory.

We write $\mathcal{G}$ for the gauge group of a B-twisted gauge theory to distinguish it from the gauge group of the six-dimensional theory. We pick generators $\mathcal{T}_a$, $a=1,\ldots,\dim\mathcal{G}$, of the Lie algebra $\mathrm{Lie}(\mathcal{G})$ of $\mathcal{G}$ that are orthonormal with respect to the minus of an invariant symmetric bilinear form Tr. The spacetime of the theory is a surface D, and the gauge bundle is a principal $\mathcal{G}$-bundle $\mathcal{P}\to D$.

The basic ingredients of a B-twisted gauge theory are vector multiplets and chiral multiplets. A vector multiplet consists of a gauge field A of $\mathcal{P}$, bosonic fields

$$\sigma\in\Gamma\big(\Lambda^1_D\otimes\mathrm{ad}(\mathcal{P})\big)\,,\quad \mathsf{D}\in\Gamma\big(\Lambda^0_D\otimes\mathrm{ad}(\mathcal{P})\big)\,, \tag{39}$$

and fermionic fields

$$\alpha \in \Gamma\left(\Lambda^0_D \otimes \mathrm{ad}(\mathcal{P})\right), \quad \lambda \in \Gamma\left(\Lambda^1_D \otimes \mathrm{ad}(\mathcal{P})\right), \quad \zeta \in \Gamma\left(\Lambda^2_D \otimes \mathrm{ad}(\mathcal{P})\right). \tag{40}$$

A chiral multiplet is valued in a unitary representation R of $\mathcal{G}$. It consists of bosonic fields

$$\varphi \in \Gamma\left(\Lambda^0_D \otimes R(\mathcal{P})\right), \quad \mathsf{F} \in \Gamma\left(\Lambda^2_D \otimes R(\mathcal{P})\right), \tag{41}$$

and fermionic fields

$$\bar{\eta} \in \Gamma\left(\Lambda^0_D \otimes \overline{R}(\mathcal{P})\right), \quad \rho \in \Gamma\left(\Lambda^1_D \otimes R(\mathcal{P})\right), \quad \bar{\mu} \in \Gamma\left(\Lambda^2_D \otimes \overline{R}(\mathcal{P})\right), \tag{42}$$

where $R(\mathcal{P})$ denotes the vector bundle associated to $\mathcal{P}$ constructed from R, and $\overline{R}$ is the complex conjugate of R which we also regard as the dual of R by a hermitian form on the representation space.

As mentioned already, a B-twisted $\mathcal{N} = (2,2)$ supersymmetric theory has two scalar supercharges, Q and $\widetilde{Q}$. Under the action of Q, the vector multiplet transforms as

$$\begin{aligned} \delta\mathcal{A} &= 0\,, & \delta\overline{\mathcal{A}} &= \lambda\,, \\ \delta\lambda &= 0\,, & \delta\alpha &= \mathsf{D}\,, \\ \delta\mathsf{D} &= 0\,, & \delta\zeta &= \mathcal{F}\,, \end{aligned} \tag{43}$$

while the chiral multiplet transforms as

$$\begin{aligned} \delta\varphi &= 0\,, & \delta\bar{\varphi} &= \bar{\eta}\,, \\ \delta\rho &= \mathcal{D}\varphi\,, & \delta\bar{\eta} &= 0\,, \\ \delta\mathsf{F} &= \mathcal{D}\rho - \zeta\varphi\,, & \delta\overline{\mathsf{F}} &= 0\,, \\ & & \delta\bar{\mu} &= \overline{\mathsf{F}}\,. \end{aligned} \tag{44}$$

Here we have introduced the notation

$$\mathcal{A} = A + \mathrm{i}\sigma\,, \qquad \overline{\mathcal{A}} = A - \mathrm{i}\sigma\,. \tag{45}$$

The fields $\bar{\varphi}$ and $\overline{\mathsf{F}}$ are the hermitian conjugates of φ and F.

The other supercharge $\widetilde{Q}$ depends on a choice of a metric on D. It acts on the vector multiplet by

$$
\begin{aligned}
&\tilde{\delta}\mathcal{A} = \star\lambda\,, && \tilde{\delta}\overline{\mathcal{A}} = 0\,,\\
&\tilde{\delta}\lambda = 0\,, && \tilde{\delta}\alpha = \star\overline{\mathcal{F}}\,,\\
&\tilde{\delta}\mathsf{D} = -\star\overline{\mathcal{D}}\lambda\,, && \tilde{\delta}\zeta = -\star\mathsf{D} + 2\mathrm{i}D\star\sigma
\end{aligned}
\tag{46}
$$

and on the chiral multiplet by

$$
\begin{aligned}
&\tilde{\delta}\varphi = 0\,, && \tilde{\delta}\bar{\varphi} = -\star\bar{\mu}\,,\\
&\tilde{\delta}\rho = -\star\overline{\mathcal{D}}\varphi\,, && \tilde{\delta}\bar{\eta} = \star\overline{\mathsf{F}}\,,\\
&\tilde{\delta}\mathsf{F} = \overline{\mathcal{D}}\star\rho - \star\alpha\varphi\,, && \tilde{\delta}\overline{\mathsf{F}} = 0\,,\\
& && \tilde{\delta}\bar{\mu} = 0\,.
\end{aligned}
\tag{47}
$$

The main part of the action is exact with respect to both Q and $\widetilde{Q}$. The action governing the dynamics of the vector multiplet is

$$
\begin{aligned}
S_{\mathrm{V}} &= \int_D \delta\tilde{\delta}\,\mathrm{Tr}(\zeta\alpha)\\
&= \int_D \delta\,\mathrm{Tr}\big((-\star\mathsf{D} + 2\mathrm{i}D\star\sigma)\alpha - \zeta\star\overline{\mathcal{F}}\big)\\
&= \int_D \mathrm{Tr}\big(-\overline{\mathcal{F}}\star\mathcal{F} - \mathsf{D}\star(\mathsf{D} - 2\mathrm{i}\star D\star\sigma) + \alpha\mathcal{D}\star\lambda + \zeta\star\overline{\mathcal{D}}\lambda\big)\,.
\end{aligned}
\tag{48}
$$

The action for the chiral multiplet is

$$
\begin{aligned}
S_{\mathrm{C}} &= \int_D \delta\tilde{\delta}(-\bar{\varphi}\mathsf{F})\\
&= \int_D \delta\big(-\bar{\varphi}(\overline{\mathcal{D}}\star\rho - \star\alpha\varphi) + \bar{\mu}\star\mathsf{F}\big)\\
&= \int_D \big(-\bar{\varphi}\overline{\mathcal{D}}\star\mathcal{D}\varphi + \star\bar{\varphi}\mathsf{D}\varphi + \overline{\mathsf{F}}\star\mathsf{F}\\
&\qquad - \bar{\eta}\overline{\mathcal{D}}\star\rho - \bar{\mu}\star\mathcal{D}\rho + \star\bar{\eta}\alpha\varphi - \bar{\varphi}\lambda\wedge\star\rho + \bar{\mu}\star\zeta\varphi\big)\,.
\end{aligned}
\tag{49}
$$

In addition, we can turn on a superpotential W, which is a gauge invariant holomorphic function of the chiral multiplet scalars. It generates the

interaction terms given by

$$S_W = \int_D \left(\mathsf{F} \frac{\partial W}{\partial \varphi} + \frac{1}{2} \rho \wedge \rho \frac{\partial^2 W}{\partial \varphi \partial \varphi} + \star \delta \tilde{\delta} \overline{W} \right) = \int_D \left(\mathsf{F} \frac{\partial W}{\partial \varphi} + \frac{1}{2} \rho \wedge \rho \frac{\partial^2 W}{\partial \varphi \partial \varphi} - \overline{\mathsf{F}} \frac{\partial \overline{W}}{\partial \bar{\varphi}} - \bar{\eta} \bar{\mu} \frac{\partial^2 \overline{W}}{\partial \bar{\varphi} \partial \bar{\varphi}} \right). \tag{50}$$

Unlike S_{V} and S_{C}, this is neither Q-exact nor $\widetilde{Q}$-exact. Furthermore, it is not automatically invariant under Q or $\widetilde{Q}$ if D has a boundary. The Q-invariance requires

$$\int_{\partial D} \rho \frac{\partial W}{\partial \varphi} = 0\,, \tag{51}$$

while for the $\widetilde{Q}$-invariance we need

$$\int_{\partial D} \star \rho \frac{\partial W}{\partial \varphi} = 0\,. \tag{52}$$

Appropriate boundary conditions must be imposed for the supercharges to be unbroken.

2.3.3. Topological–holomorphic theory as a B-twisted gauge theory. Now we take $M = D \times \Sigma$ and describe the six-dimensional topological–holomorphic theory on $D \times \Sigma \times C$ as a B-twisted gauge theory on D. We use letters i, j, ... for indices for D and m, n, ... for those for Σ. For simplicity we assume $D \times \Sigma \times C$ has no boundary (or impose appropriate boundary conditions so that all boundary terms arising from integration by parts vanish).

We need to organize the fields of the six-dimensional theory into supermultiplets of B-twisted gauge theory. Clearly, the theory has a single vector multiplet whose gauge field $A_i \mathrm{d}x^i$ is part of the six-dimensional gauge field. Comparing the transformation rules (43) and (27) in two and six dimensions, we identify the other fields in the vector multiplet as

$$\sigma_i = \phi_i\,, \qquad \mathsf{D} = \mathsf{P}\,, \qquad \alpha = \xi'\,, \qquad \lambda_i = \psi_i\,, \qquad \zeta_{ij} = \chi_{ij}\,. \tag{53}$$

In order to lift the vector multiplet action (48) to six dimensions, we interpret the bilinear form Tr on $\mathrm{Lie}(\mathcal{G})$ as

$$\frac{1}{e^2} \int_{\Sigma \times C} \star_{\Sigma \times C} \mathrm{Tr}\,, \tag{54}$$

where Tr in the integrand stands for the bilinear form on $\mathfrak{g}$. This gives

$$S_{\mathrm{V}} = \frac{1}{e^2}\int_{D\times\Sigma\times C}\sqrt{g}\,\mathrm{d}^6x\,\delta\,\mathrm{Tr}\Big((-\mathsf{P}+2\mathrm{i}D^i\phi_i)\xi' - \frac{1}{2}\chi^{ij}\overline{\mathcal{F}}_{ij}\Big)\,. \tag{55}$$

As can be seen from the bosonic fields annihilated by Q, we have three chiral multiplets in the adjoint representation, whose scalar components are $\mathcal{A}_m$ and $A_{\bar z}$. We name the fields of these multiplets as $(\mathcal{A}_m, \mathsf{F}_m, \bar\eta_m, \rho_m, \bar\mu_m)$ and $(A_{\bar z}, \mathsf{F}_{\bar z}, \bar\eta_z, \rho_{\bar z}, \bar\mu_z)$.

To lift formula (44) to six dimensions, what we have to do is essentially to replace the scalar fields with the corresponding covariant derivatives so that when we perform dimensional reduction on $\Sigma\times C$, we get back to the same formula. In this way we obtain

$$\begin{aligned}
&\delta\mathcal{A}_m = 0\,, && \delta\overline{\mathcal{A}}_m = \bar\eta_m\,,\\
&\delta\rho_{mi} = \mathcal{F}_{im}\,, && \delta\bar\eta_m = 0\,,\\
&\delta\mathsf{F}_{mij} = \mathcal{D}_i\rho_{mj} - \mathcal{D}_j\rho_{mi} + \mathcal{D}_m\zeta_{ij}\,, && \delta\overline{\mathsf{F}}_{mij} = 0\,,\\
& && \delta\bar\mu_{mij} = \overline{\mathsf{F}}_{mij}
\end{aligned} \tag{56}$$

and

$$\begin{aligned}
&\delta A_{\bar z} = 0\,, && \delta A_z = \bar\eta_z\,,\\
&\delta\rho_{\bar z i} = \mathcal{F}_{i\bar z}\,, && \delta\bar\eta_z = 0\,,\\
&\delta\mathsf{F}_{\bar z ij} = \mathcal{D}_i\rho_{\bar z j} - \mathcal{D}_j\rho_{\bar z i} + D_{\bar z}\zeta_{ij}\,, && \delta\overline{\mathsf{F}}_{zij} = 0\,,\\
& && \delta\bar\mu_{zij} = \overline{\mathsf{F}}_{zij}\,.
\end{aligned} \tag{57}$$

From the Q-variations involving the gauge field, we see

$$\bar\eta_m = \psi_m\,,\qquad \rho_{mi} = \chi_{im}\,,\qquad \bar\eta_z = \xi\,,\qquad \rho_{\bar z i} = \psi'_i\,. \tag{58}$$

With this identification, we can write $\delta\mathsf{F}_{mij} = (\mathcal{D}\chi)_{mij}$ and $\delta\mathsf{F}_{\bar z ij} = (\mathcal{D}\psi')_{ij} + D_{\bar z}\chi_{ij}$. On the other hand, from the six-dimensional action we derive the equations of motion

$$\mathcal{D}_M\chi = -\star_M D_z\psi + \star_M\mathcal{D}_M\xi = -\star_M\iota_{\partial_z}\delta\overline{\mathcal{F}} \tag{59}$$

and

$$\mathcal{D}_M\psi' + D_{\bar z}\chi = -\star_M\overline{\mathcal{D}}_M\psi = -\star_M\delta\overline{\mathcal{F}}\,, \tag{60}$$

with $\mathcal{D}_M = \mathcal{D}_\mu \mathrm{d}x^\mu$. Combining these equations we deduce the on-shell relations

$$\mathsf{F}_{mij} = -\big(\star_M \iota_{\partial_z} \overline{\mathcal{F}}\big)_{mij}\,, \qquad \mathsf{F}_{\bar{z}ij} = -\big(\star_M \overline{\mathcal{F}}\big)_{ij}\,. \tag{61}$$

Then, we have $\overline{\mathsf{F}}_{mij} = -(\star_M \iota_{\partial_{\bar{z}}} \mathcal{F})_{mij}$ and $\overline{\mathsf{F}}_{zab} = -(\star_M \mathcal{F})_{ij}$ on shell (note the sign; the on-shell value of $\overline{\mathsf{F}}$ is minus the hermitian conjugate of F) and

$$\bar{\mu}_{mij} = (\star_M \psi')_{mij}\,, \qquad \bar{\mu}_{zij} = -(\star_M \chi)_{ij}\,. \tag{62}$$

Lifting the chiral multiplet action (49) to six dimensions is also straightforward. For example, the term $\delta(-\bar{\varphi}\overline{\mathcal{D}} \star \rho)$ in the integrand on the second line can be converted to $\delta(\overline{\mathcal{D}}\bar{\varphi} \wedge \star\rho)$ by integration by parts and lifted to $\delta \operatorname{Tr}(-\overline{\mathcal{F}}^{im}\chi_{im} - \overline{\mathcal{F}}_{iz}\psi'^i)$. Also, since φ is in the adjoint representation, $\delta(\bar{\varphi} \star \alpha\varphi)$ can be written as $\delta \operatorname{Tr}(\star[\varphi, \bar{\varphi}]\alpha)$ and is lifted to $\delta \operatorname{Tr}((2\mathrm{i}D^m\phi_m + F_{z\bar{z}})\xi')$. For comparison with the six-dimensional description, it is useful to express the chiral multiplet action as

$$\begin{aligned} S_{\mathrm{C}} = \frac{1}{e^2}\int_{D\times\Sigma\times C} \sqrt{g}\,\mathrm{d}^6x \operatorname{Tr}\bigg(& \delta\big(-\overline{\mathcal{F}}^{im}\chi_{im} - \overline{\mathcal{F}}_{iz}\psi'^i + (2\mathrm{i}D^m\phi_m + F_{z\bar{z}})\xi'\big) \\ & + \frac{1}{2}\overline{\mathsf{F}}^{mij}\mathsf{F}_{mij} + \frac{1}{2}\overline{\mathsf{F}}_z^{ij}\mathsf{F}_{\bar{z}ij}\bigg) \\ + \frac{1}{e^2}\int_{D\times\Sigma\times C} \mathrm{d}^2z \operatorname{Tr}\big(-\mathcal{D}\chi \wedge \psi'_\Sigma + \chi_\Sigma \wedge (\mathcal{D}\psi' + D_{\bar{z}}\chi)\big)\,. \end{aligned} \tag{63}$$

Here we have defined

$$\chi_\Sigma = \frac{1}{2}\chi_{mn}\mathrm{d}x^m \wedge \mathrm{d}x^n\,, \qquad \psi'_\Sigma = \psi'_m \mathrm{d}x^m\,. \tag{64}$$

We also need to determine the superpotential. In order to reproduce the equations of motion (61), the superpotential must be

$$W = -\frac{\mathrm{i}}{e^2}\int_{\Sigma\times C} \mathrm{d}z \wedge \mathrm{CS}(\mathcal{A})\,, \tag{65}$$

where

$$\mathrm{CS}(\mathcal{A}) = \operatorname{Tr}\bigg(\mathcal{A} \wedge \mathrm{d}\mathcal{A} + \frac{2}{3}\mathcal{A} \wedge \mathcal{A} \wedge \mathcal{A}\bigg) \tag{66}$$

is the Chern–Simons three-form constructed from $\mathcal{A}$. The corresponding superpotential terms are

$$
\begin{aligned}
(67)\qquad S_W = \frac{1}{e^2}\int_{D\times\Sigma\times C}\sqrt{g}\,\mathrm{d}^6x\,\mathrm{Tr}\Big(&\frac{1}{2}\mathsf{F}^{mij}(\star_M\,\iota_{\partial_{\bar z}}\mathcal{F})_{mij}+\frac{1}{2}\mathsf{F}^{ij}_{\bar z}(\star_M\,\mathcal{F})_{ij}\\
&+\frac{1}{2}\overline{\mathsf{F}}^{mij}(\star_M\,\iota_{\partial_z}\overline{\mathcal{F}})_{mij}+\frac{1}{2}\overline{\mathsf{F}}^{ij}_{z}(\star_M\,\overline{\mathcal{F}})_{ij}\\
&-\delta\overline{\mathcal{F}}_{mz}\psi'^m+\frac{1}{2}\chi^{mn}\delta\overline{\mathcal{F}}_{mn}\Big)\\
+\frac{1}{e^2}\int_{D\times\Sigma\times C}\mathrm{d}^2z\,\mathrm{Tr}\Big(&\frac{1}{2}\chi_{D|\Sigma}\wedge D_{\bar z}\chi+\chi_{D|\Sigma}\wedge\mathcal{D}\psi'_D\Big),
\end{aligned}
$$

where

$$
(68)\qquad \chi_{D|\Sigma}=\chi_{im}\,\mathrm{d}x^i\wedge\mathrm{d}x^m\,,\qquad \psi'_D=\psi'_i\,\mathrm{d}x^i\,.
$$

The superpotential (65) is not quite gauge invariant, but this is not a problem because the resulting action is gauge invariant.

While the sum $S_{\mathrm{V}}+S_{\mathrm{C}}+S_W$ reproduces the fermionic part of the six-dimensional action, they lack the terms

$$
(69)\qquad \frac{1}{e^2}\int_{D\times\Sigma\times C}\sqrt{g}\,\mathrm{d}^6x\,\mathrm{Tr}\Big(-\frac{1}{2}\overline{\mathcal{F}}^{mn}\mathcal{F}_{mn}-2\overline{\mathcal{F}}^{m\bar z}\mathcal{F}_{m\bar z}\Big)
$$

from the bosonic part. These missing terms are supplied when the auxiliary fields are integrated out. Thus, we have obtained a two-dimensional formulation of the six-dimensional topological–holomorphic theory, which is applicable for $M=D\times\Sigma$.

3. Ω-deformation of B-twisted gauge theories

Once we reformulate the six-dimensional topological–holomorphic theory as a two-dimensional B-twisted gauge theory, we can subject it to an Ω-deformation [22–24] following the construction of [29]. Via localization of the path integral, the Ω-deformation reduces the topological sector of the B-twisted gauge theory to a zero-dimensional gauge theory with complex gauge group [28–30]. In this section we discuss this localization mechanism for a general B-twisted gauge theory.

3.1. Ω-deformation

As we have seen above, a B-twisted gauge theory has two scalar supercharges Q and $\widetilde{Q}$. If the spacetime D is flat, the theory additionally has a one-form supercharge $G = G_i \mathrm{d}x^i$, satisfying $\{Q, G_i\} = P_i$ and $\{G_i, G_j\} = 0$ in some coordinates. More generally, if V is a parallel vector field on D (which may or may not be curved), there is an associated fermionic symmetry and hence the corresponding supercharge $\iota_V G$. The linear combination $Q + \iota_V G$ is then a supercharge which squares to V.

If V is not covariantly constant, $\iota_V G$ does not exist in general. Nevertheless, if V is a Killing vector field generating an isometry of D, we can construct a deformation of the theory such that it has a supercharge Q_V that squares to the generator of the isometry and reduces to Q for $V = 0$. This deformation is what we call an Ω-deformation of the B-twisted gauge theory.

Specifically, the deformed supercharge Q_V acts on the vector multiplet by

$$\begin{aligned} &\delta_V \mathcal{A} = \iota_V \zeta\,, && \delta_V \overline{\mathcal{A}} = \lambda - \iota_V \zeta\,, \\ &\delta_V \lambda = 2\iota_V F - 2\mathrm{i} D \iota_V \sigma\,, && \delta_V \zeta = \mathcal{F}\,, \\ &\delta_V \alpha = \mathsf{D}\,, && \delta_V \mathsf{D} = \iota_V \mathcal{D}\alpha\,, \end{aligned} \tag{70}$$

and on the chiral multiplet by

$$\begin{aligned} &\delta_V \varphi = \iota_V \rho\,, && \delta_V \bar{\varphi} = \bar{\eta}\,, \\ &\delta_V \rho = \mathcal{D}\varphi + \iota_V \mathsf{F}\,, && \delta_V \bar{\eta} = \iota_V \mathcal{D}\bar{\varphi}\,, \\ &\delta_V \mathsf{F} = \mathcal{D}\rho - \zeta\varphi\,, && \delta_V \overline{\mathsf{F}} = \mathcal{D}\iota_V \bar{\mu}\,, \\ & && \delta_V \bar{\mu} = \overline{\mathsf{F}}\,. \end{aligned} \tag{71}$$

Its square is essentially the Lie derivative $\mathcal{L}_V = \mathrm{d}\iota_V + \iota_V \mathrm{d}$, but made covariant with respect to the complexified gauge symmetry:

$$Q_V^2 = \mathcal{D}\iota_V + \iota_V \mathcal{D}\,. \tag{72}$$

The right-hand side equals $\mathcal{L}_V$ plus the infinitesimal gauge transformation generated by $\iota_V \mathcal{A}$.

Being a generator of an isometry, V is a real vector field. More generally, we allow V to be a complex Killing vector field that commutes with its complex conjugate $\overline{V}$. Also, $V|_{\partial D}$ must be tangent to ∂D so that the isometry preserves the boundary.

The action of the Ω-deformed theory is again of the form $S_{\mathrm{V}} + S_{\mathrm{C}} + S_W$, each term being a Q_V-invariant deformation of the corresponding term in the undeformed action. As in the undeformed case, we can take S_{V} and S_{C} to be Q_V-exact. A minimal choice is

$$
\begin{aligned}
(73) \qquad S_{\mathrm{V}} &= \delta_V \int_D \mathrm{Tr}\big((-\star\mathsf{D} + 2\mathrm{i}D\star\sigma)\alpha - \zeta\star\overline{\mathcal{F}}\big) \\
&= \int_D \mathrm{Tr}\big(-\overline{\mathcal{F}}\star\mathcal{F} - \mathsf{D}\star(\mathsf{D} - 2\mathrm{i}\star D\star\sigma) \\
&\qquad\qquad + \alpha\mathcal{D}\star\tilde{\lambda} + \zeta\star\overline{\mathcal{D}}\tilde{\lambda} + \alpha\mathrm{d}V^\flat\star\zeta\big) \\
&\quad + \int_{\partial D} \mathrm{Tr}(\star\zeta\star\iota_V\star\alpha)
\end{aligned}
$$

and

$$
\begin{aligned}
(74) \qquad S_{\mathrm{C}} &= \delta_V \int_D \big((\overline{\mathcal{D}}\bar{\varphi} + \iota_{\overline{V}}\overline{\mathsf{F}}) \wedge\star\rho + \star\bar{\varphi}\alpha\varphi + \bar{\mu}\star\mathsf{F}\big) \\
&= \int_D \big((\overline{\mathcal{D}}\bar{\varphi} + \iota_{\overline{V}}\overline{\mathsf{F}}) \wedge\star\big(\mathcal{D}\varphi + \iota_V\mathsf{F}\big) + \star\bar{\varphi}\mathsf{D}\varphi + \overline{\mathsf{F}}\star\mathsf{F} + \overline{\mathcal{D}}\bar{\eta}\star\rho \\
&\qquad - \bar{\mu}\star\mathcal{D}\rho + \star\bar{\eta}\alpha\varphi - \bar{\varphi}\tilde{\lambda}\wedge\star\rho + \bar{\mu}\star\zeta\varphi + \iota_{\overline{V}}\mathcal{D}\iota_V\bar{\mu}\star\rho\big),
\end{aligned}
$$

where $V^\flat$ is the one-form dual to V with respect to the metric on D and

$$
(75) \qquad \tilde{\lambda} = \lambda - \iota_V\zeta - \star\iota_V\star\alpha\,.
$$

It is important here that V is a Killing vector field and $[V,\overline{V}] = 0$. The former property means that $\mathcal{L}_V$ annihilates the metric and commutes with $\star$, while the latter implies $[\mathcal{L}_V, \iota_{\overline{V}}] = 0$. Together with the identity $[\mathcal{L}_V, \iota_V] = 0$, these properties ensure the Q_V-invariance of S_{V} and S_{C}.

Remarkably, the Ω-deformation allows S_W to be Q_V-invariant without resorting to any boundary conditions. Suppose, for simplicity, that there is only one boundary component in D, and parametrize this boundary circle by an angular coordinate θ. Then

$$
\begin{aligned}
(76) \qquad S_W &= \int_D \left(\mathsf{F}\frac{\partial W}{\partial\varphi} + \frac{1}{2}\rho\wedge\rho\frac{\partial^2 W}{\partial\varphi\partial\varphi} - \delta_V\left(\bar{\mu}\frac{\partial\overline{W}}{\partial\bar{\varphi}}\right)\right) - \int_{\partial D} W\frac{\mathrm{d}\theta}{V^\theta} \\
&= \int_D \left(\mathsf{F}\frac{\partial W}{\partial\varphi} + \frac{1}{2}\rho\wedge\rho\frac{\partial^2 W}{\partial\varphi\partial\varphi} - \overline{\mathsf{F}}\frac{\partial\overline{W}}{\partial\bar{\varphi}} - \bar{\eta}\bar{\mu}\frac{\partial^2\overline{W}}{\partial\bar{\varphi}\partial\bar{\varphi}}\right) - \int_{\partial D} W\frac{\mathrm{d}\theta}{V^\theta}
\end{aligned}
$$

is a Q_V-invariant superpotential action.

Since Q_V squares to zero on operators and states that are invariant under the gauge symmetry and the isometry, we can define its cohomology

in the spaces of such states and operators. Unlike the undeformed case, the Q_V-invariant sector of the theory is not quite topological: it is invariant under deformations of the metric only if V remains as a Killing vector field. For this reason, we refer to the Ω-deformed B-twisted gauge theory as a quasi-topological theory.

3.2. Localization on a disk

As we have just seen, an Ω-deformation can be applied to a B-twisted gauge theory whenever the spacetime D has an isometry. A basic example is when D is a disk of finite radius, equipped with a rotation invariant metric, and V is a generator of rotations. We now show that for a suitable boundary condition, the quasi-topological sector of the Ω-deformed theory is in this case equivalent to a zero-dimensional theory, whose path integral is performed over a domain specified by the boundary condition.

To be concrete, we endow D with the metric of a hemisphere of unit area. In terms of polar coordinates (r, θ), the metric takes the form

$$g(r,\theta) = g_{rr}(r)\mathrm{d}r^2 + g_{\theta\theta}(r)\mathrm{d}\theta^2 \tag{77}$$

and we have

$$V = \epsilon\partial_\theta \tag{78}$$

for some $\epsilon \in \mathbb{C}$. We use hatted indices $(\hat{r}, \hat{\theta})$ to denote components of tensors with respect to the orthonormal vectors $\partial_{\hat{\theta}} = \sqrt{g^{\theta\theta}}\,\partial_\theta$, $\partial_{\hat{r}} = \sqrt{g^{rr}}\,\partial_r$ and their duals $\mathrm{d}\hat{\theta} = \sqrt{g_{\theta\theta}}\,\mathrm{d}\theta$, $\mathrm{d}\hat{r} = \sqrt{g_{rr}}\,\mathrm{d}\hat{r}$. For example, $V^{\hat{\theta}} = \sqrt{g_{\theta\theta}}\,V^\theta$ and $|V^{\hat{\theta}}|$ equals the norm $\|V\|$ of V.

3.2.1. Boundary conditions. To begin with, let us figure out what sort of boundary condition should be imposed. In general, a good boundary condition ensures that the boundary terms vanish in the variation of the action so that the classical equations of motion are obtained from the variational principle. In our case, we moreover want the boundary condition to be Q_V-invariant.

We start with the vector multiplet. Integrating D out and varying the gauge field, we see that the boundary terms in the variation of the action vanish if either the Dirichlet condition $\delta A_\theta = 0$ or the Neumann condition $F_{r\theta} = 0$ is satisfied on the boundary. The former breaks the gauge symmetry on the boundary. For our applications we look for a boundary condition that

preserves the gauge symmetry, so we pick the Neumann condition. We can also choose a gauge in which

$$A_r = 0 \tag{79}$$

on the boundary. Then, the Neumann condition reads

$$\partial_r A_\theta = 0\,. \tag{80}$$

Next, varying σ, we find that each component σ_i of σ should obey either the Dirichlet condition $\delta\sigma_i = 0$ or the Neumann condition $\partial_r\sigma_i = 0$. In view of the fact that the gauge field appears in Q_V^2 through the combination $\mathcal{A} = A + \mathrm{i}\sigma$, it is natural to choose

$$\sigma_r = \partial_r\sigma_\theta = 0\,. \tag{81}$$

Letting Q_V act on the boundary conditions we have so far, we get

$$\zeta_{r\theta} = \lambda_r = \partial_r\lambda_\theta = 0\,. \tag{82}$$

These conditions already ensure that the boundary terms vanish under variations of the fermions.

Since λ_θ does not vanish on the boundary, the action should have a term that contains the boundary value of λ_θ. The only term that may not vanish on the boundary is $\alpha\mathcal{D}^\theta\lambda_\theta$. So we require α to obey the Neumann condition

$$\partial_r\alpha = 0\,, \tag{83}$$

just as λ_θ does. Then we must have

$$\partial_r\mathsf{D} = 0 \tag{84}$$

for Q_V-invariance.

The set of boundary conditions for the vector multiplet thus obtained is Q_V-invariant. Repeated action of Q_V does not lead to any further conditions since Q_V^2 just generates translations on the boundary.

On the chiral multiplet, we impose a boundary condition of brane type. The target space X for the chiral multiplet is the representation space of the representation R. We choose a submanifold γ in X, and demand the

boundary value of the scalar field to lie in γ:

$$\varphi \in \gamma\,. \tag{85}$$

The Q_V-action on this condition yields

$$(\iota_V \rho, \bar{\eta}) \in T_\varphi \gamma \otimes \mathbb{C}\,. \tag{86}$$

We require γ to be G-invariant so that the gauge symmetry is preserved. Furthermore, we assume that $\mathrm{Re}(W/\epsilon)$ is bounded above on γ so that the boundary term in the superpotential action (76) does not render the path integral divergent.

Varying the fermions we get the boundary terms

$$\int_{\partial D} \mathrm{d}\hat{\theta} \left(-\frac{1}{\|V\|^2} \delta(\iota_V \rho)(\iota_{\overline{V}} \bar{\mu})_{\hat{r}} + \delta\bar{\eta}\rho_{\hat{r}} \right). \tag{87}$$

For these terms to vanish, we should have

$$\left(\rho_{\hat{r}}, -\|V\|^{-2}(\iota_{\overline{V}}\bar{\mu})_{\hat{r}}\right) \in N_\varphi\gamma \otimes \mathbb{C}\,, \tag{88}$$

where $N\gamma$ is the normal bundle of γ with respect to the Kähler metric

$$g_X = \mathrm{Re}(\mathrm{d}\varphi \otimes \mathrm{d}\bar{\varphi}) \tag{89}$$

of X. The Q_V-variation of this condition, together with the gauge condition (79), gives

$$\left(\partial_{\hat{r}}\varphi + (\iota_V \mathsf{F})_{\hat{r}}, -\|V\|^{-2}(\iota_{\overline{V}}\overline{\mathsf{F}})_{\hat{r}},\right) \in N_\varphi\gamma \otimes \mathbb{C}\,, \tag{90}$$

which completes a Q_V-invariant set of boundary conditions on the chiral multiplet.

The equations of motion for F and $\overline{\mathsf{F}}$ are

$$\begin{aligned} \mathsf{F}_{\hat{r}\hat{\theta}} &= \frac{1}{1+\|V\|^2}\left(\overline{V}^{\hat{\theta}}\mathcal{D}_{\hat{r}}\varphi + \frac{\partial\overline{W}}{\partial\bar{\varphi}}\right), \\ \overline{\mathsf{F}}_{\hat{r}\hat{\theta}} &= \frac{1}{1+\|V\|^2}\left(V^{\hat{\theta}}\overline{\mathcal{D}}_{\hat{r}}\bar{\varphi} - \frac{\partial W}{\partial\varphi}\right). \end{aligned} \tag{91}$$

Plugging these equations into the boundary condition (90), we get

$$\left(\partial_{\hat{r}}\varphi - V^{\hat{\theta}}\frac{\partial\overline{W}}{\partial\bar{\varphi}}, \partial_{\hat{r}}\bar{\varphi} - \frac{1}{V^{\hat{\theta}}}\frac{\partial W}{\partial\varphi}\right) \in N_\varphi\gamma \otimes \mathbb{C}\,. \tag{92}$$

As a check, let us verify that boundary terms are absent in the variation of the action under this boundary condition. After F and $\overline{\mathsf{F}}$ are integrated out, the bosonic terms in $S_{\mathrm{C}} + S_W$ are given by

$$\int_D \mathrm{d}\hat{r}\,\mathrm{d}\hat{\theta}\Big(\frac{1}{1+\|V\|^2}\Big(\overline{\mathcal{D}}_{\hat{r}}\bar{\varphi}\mathcal{D}_{\hat{r}}\varphi + \overline{V}^{\hat{\theta}}\partial_{\hat{r}}W - V^{\hat{\theta}}\partial_{\hat{r}}\overline{W} + \frac{\partial W}{\partial\varphi}\frac{\partial\overline{W}}{\partial\bar{\varphi}}\Big) + \overline{\mathcal{D}}_{\hat{\theta}}\bar{\varphi}\mathcal{D}_{\hat{\theta}}\varphi + \bar{\varphi}\mathsf{D}\varphi\Big) - \int_{\partial D} W\frac{\mathrm{d}\hat{\theta}}{V^{\hat{\theta}}}\,. \tag{93}$$

Varying the scalars, we see that the boundary terms indeed vanish.

In the undeformed case $\epsilon = 0$, the boundary condition (92) implies that W is locally constant on γ. The same condition then requires $(\partial_r\varphi, \partial_r\bar{\varphi}) \in N_\varphi\gamma$ on the boundary. If γ is a complex submanifold, this is (part of) a "B-brane" boundary condition for a B-twisted Landau–Ginzburg model [27], which preserves half of $\mathcal{N} = (2,2)$ supersymmetry. For our application, however, we will actually take γ to be, roughly speaking, a *Lagrangian* submanifold.

We remark that the boundary condition described here depends on $\|V\|^2$. As a consequence, the presence of boundary mildly breaks the quasi-topological invariance of the theory. We are still allowed to deform the metric as long as we continue to impose the same boundary condition defined with respect to the original metric.

3.2.2. Gauge fixing. Performing the path integral requires gauge fixing. We do this by adapting the BRST gauge fixing procedure to the present setting.

We enlarge the set of fields with additional fermionic fields b, c and auxiliary bosonic field B, all transforming in the adjoint representation:

$$b,\, c,\, \mathsf{B} \in \Gamma\big(\mathrm{ad}(\mathcal{P})\big)\,. \tag{94}$$

Then we introduce the BRST symmetry that acts on these fields by

$$\delta_{\mathrm{B}} b = \mathsf{B}\,, \qquad \delta_{\mathrm{B}}\mathsf{B} = 0\,, \qquad \delta_{\mathrm{B}} c = \frac{1}{2}\{c,c\}\,. \tag{95}$$

On the other fields the BRST symmetry acts by the gauge transformation generated by c; for instance, $\delta_{\mathrm{B}}\varphi = c\varphi$. Since the action of the theory is gauge invariant, it is invariant under the BRST symmetry.

The conserved charge Q_{B} for the BRST symmetry squares to zero. In the standard BRST gauge fixing, one adds Q_{B}-exact gauge fixing terms to

the action and considers the Q_{B}-cohomology. However, such terms will not be Q_V-invariant and breaks the quasi-topological invariance of the theory.

To remedy this problem we combine Q_{B} with Q_V. Let us postulate that Q_V acts on b, c and B by

$$\delta_V b = 0\,, \qquad \delta_V \mathsf{B} = \iota_V \mathrm{d} b\,, \qquad \delta_V c = -\iota_V \mathcal{A}\,. \tag{96}$$

With this definition of the action of Q_V, the combined charge $Q_{V+\mathrm{B}} = Q_V + Q_{\mathrm{B}}$ satisfies

$$Q_{V+\mathrm{B}}^2 = \iota_V \mathrm{d} + \mathrm{d}\iota_V\,. \tag{97}$$

The right-hand side is the ordinary Lie derivative instead of a gauge covariant one, so we can define the cohomology with respect to the action of $Q_{V+\mathrm{B}}$ on rotation invariant states and operators which are not necessarily gauge invariant.

After gauge fixing, therefore, what we should consider is not the Q_V-cohomology, but the $Q_{V+\mathrm{B}}$-cohomology in the spaces of rotation invariant states and operators. Since S_{V} and S_{C} are Q_V-commutators of gauge invariant expressions, they are automatically exact with respect to $Q_{V+\mathrm{B}}$. The quasi-topological invariance of the theory is thus maintained.

Now we can perform gauge fixing as in the usual BRST procedure, treating $Q_{V+\mathrm{B}}$ as the BRST operator. We pick a suitable $\mathrm{Lie}(\mathcal{G})$-valued function f_{GF} constructed from the original set of fields, and add to the action the $Q_{V+\mathrm{B}}$-exact term

$$\delta_{V+\mathrm{B}} \int_D \star \operatorname{Tr}(2\mathrm{i} b f_{\mathrm{GF}}) = \int_D \star \operatorname{Tr}(2\mathrm{i}\mathsf{B} f_{\mathrm{GF}} - 2\mathrm{i} b\, \delta_{V+\mathrm{B}} f_{\mathrm{GF}})\,. \tag{98}$$

Integrating over B produces a delta function which imposes the gauge fixing condition

$$f_{\mathrm{GF}} = 0\,. \tag{99}$$

The fermionic part can be written as

$$2\mathrm{i} \int_D \star (\mathcal{T}_b \cdot f_{\mathrm{GF}})^a b^a \tilde{c}^b \tag{100}$$

for some fermion

$$\tilde{c} = c + \cdots\,, \tag{101}$$

where $\mathcal{T}_b \cdot f_{\mathrm{GF}}$ denotes the infinitesimal gauge transformation of f_{GF} by $\mathcal{T}_b \in \mathrm{Lie}(\mathcal{G})$. This is possible because for a sensible choice of f_{GF}, the matrix-valued function $((\mathcal{T}_b \cdot f_{\mathrm{GF}})^a)_{a,b=1}^{\dim \mathcal{G}}$ represents an invertible operator on the field space. The integration over the fermions produces the Faddeev–Popov determinant for the gauge fixing.

For the convenience of computation, we actually make another choice of gauge fixing terms:

$$S_{\mathrm{GF}} = \delta_{V+\mathrm{B}} \int_D \star \mathrm{Tr}(-b\mathsf{B} + 2\mathrm{i}b f_{\mathrm{GF}}) \\ = \int_D \star \mathrm{Tr}\big(-\mathsf{B}^2 + b\iota_V \mathrm{d}b + 2\mathrm{i}\mathsf{B} f_{\mathrm{GF}} - 2\mathrm{i}b\, \delta_{V+\mathrm{B}} f_{\mathrm{GF}}\big)\,. \tag{102}$$

This is a $Q_{V+\mathrm{B}}$-exact deformation of the previous gauge fixing action, so it leads to the same result. With this choice, integrating B out yields the potential term $\mathrm{Tr}(-f_{\mathrm{GF}}^2)$ rather than setting $f_{\mathrm{GF}} = 0$.

In the present situation where D is endowed with a hemisphere metric, the gauge field on D can be obtained from one on the sphere S^2 by restriction. Then, Uhlenbeck's theorem [51] guarantees that there exists a representative satisfying the Lorenz gauge condition

$$\nabla^i A_i = 0 \tag{103}$$

in each gauge equivalence class (at least when the field strength is sufficiently small), and it is unique up to constant gauge transformations. Further, the Lorenz gauge is compatible with the boundary conditions (79) and (80) on the equator ∂D.[1]

We can remove the residual gauge freedom by imposing an additional gauge fixing condition:

$$f_{\mathrm{GF},0}(\varphi_0, \bar{\varphi}_0) = 0\,. \tag{104}$$

---

[1]Let (θ, ϕ) be the spherical coordinates on S^2, and A a gauge field on D obeying the boundary condition $A_\phi = \partial_\phi A_\theta = 0$ at $\phi = \pi/2$. We extend A to S^2 by setting $A_\theta(\theta, \phi) = A_\theta(\theta, \pi - \phi)$ and $A_\phi(\theta, \phi) = -A_\phi(\theta, \pi - \phi)$ for $\phi \geq \pi/2$. Suppose that a gauge transformed connection $A^g = gAg^{-1} - g^{-1}\mathrm{d}g$ satisfies the Lorenz gauge condition, and let $\tilde{g}(\theta, \phi) = g(\theta, \pi - \phi)$. Then, $A^{\tilde{g}}_\theta(\theta, \phi) = A^g_\theta(\theta, \pi - \phi)$ and $A^{\tilde{g}}_\phi(\theta, \phi) = -A^g_\phi(\theta, \pi - \phi)$, and $A^{\tilde{g}}$ also satisfies the Lorenz gauge condition. By the uniqueness property we have $\tilde{g} = g_0 g$ for some constant element g_0. This implies that at $\phi = \pi/2$, we have $-g^{-1}\partial_\phi g = -\tilde{g}^{-1}\partial_\phi \tilde{g} = +g^{-1}\partial_\phi g$ and hence $A^g_\phi = -g^{-1}\partial_\phi g = 0$ and $\partial_\phi A^g_\theta = g F_{\phi\theta} g^{-1} = 0$.

Here $f_{\mathrm{GF},0}$ is an appropriate $\mathrm{Lie}(\mathcal{G})$-valued function on X, and φ_0 is the constant part of φ, defined as the average of φ with respect to the volume form of D:

$$\varphi_0 = \int_D \star\,\varphi\,. \tag{105}$$

Hence, we take our gauge fixing function to be

$$f_{\mathrm{GF}} = \nabla^i A_i + f_{\mathrm{GF},0}(\varphi_0, \bar{\varphi}_0)\,. \tag{106}$$

The corresponding gauge fixing action is

$$\begin{aligned} S_{\mathrm{GF}} = \int_D \star\,\mathrm{Tr}\big(-(\nabla^i A_i)^2 + b\iota_V \mathrm{d}b + 2\mathrm{i}b\nabla^i D_i c - \mathrm{i}b\nabla^i \lambda_i\big) \\ - \mathrm{Tr}\big(f_{\mathrm{GF},0}(\varphi_0,\bar{\varphi}_0)^2 + 2\mathrm{i}b_0\,\delta_{V+\mathrm{B}} f_{\mathrm{GF},0}(\varphi_0,\bar{\varphi}_0)\big)\,, \end{aligned} \tag{107}$$

where b_0 is the constant part of b.

Before proceeding, we need to specify the boundary conditions for b, c and B. For the $Q_{V+\mathrm{B}}$-action to preserve the boundary gauge condition (79), c must obey the Neumann condition. Then, for the nondegeneracy of S_{GF} on the boundary b should also obey the Neumann condition, and so should B for $Q_{V+\mathrm{B}}$-invariance. Thus, we impose the boundary condition

$$\partial_r b = \partial_r c = \partial_r \mathsf{B} = 0\,. \tag{108}$$

The set of all boundary conditions is then invariant under Q_{B} as well as Q_V.

3.2.3. Localization. We are ready to demonstrate the equivalence between the $Q_{V+\mathrm{B}}$-invariant sector of the Ω-deformed B-twisted gauge theory and a zero-dimensional theory. To this end we take advantage of the invariance of the theory under $Q_{V+\mathrm{B}}$-exact deformations and rescale the kinetic terms by large factors. Such a rescaling makes the oscillating modes of fields very massive and yields an effective description in terms of constant modes.

There are various ways of doing this by a $Q_{V+\mathrm{B}}$-exact deformation, but perhaps the most transparent is to rescale the metric as

$$g \to t^{-2} g \tag{109}$$

and send t to a large value. (A disadvantage of this choice of deformation is that the metric appearing in the action no longer matches the one that enters the boundary condition.) In other words, we shrink the spacetime D

by a large factor so that the excited modes get large masses. To cancel the accompanied rescaling of the volume form, at the same time we also rescale the $Q_{V+\mathrm{B}}$-exact part of the action by a factor of t^2:

$$S_{\mathrm{V}} + S_{\mathrm{C}} + S_{\mathrm{GF}} + S_W \to t^2(S_{\mathrm{V}} + S_{\mathrm{C}} + S_{\mathrm{GF}}) + S_W \,. \tag{110}$$

We want to show that in the limit $t \to \infty$, the path integral with respect to this deformed action reduces to the path integral for a zero-dimensional theory.

After this rescaling, the bosonic part of the action, with the auxiliary fields integrated out, becomes

$$\begin{aligned}
\int_D \mathrm{d}\hat{r}\,\mathrm{d}\hat{\theta}\Bigg(& t^4 \operatorname{Tr}\Bigg(-F_{\hat{r}\hat{\theta}}^2 - D^i\sigma^j D_i\sigma_j - R_{ij}\sigma^i\sigma^j - [\sigma_{\hat{r}},\sigma_{\hat{\theta}}]^2 - (\nabla^i A_i)^2 \Bigg) \\
& + t^2\Bigg(\frac{t^2 - \|V\|^2}{t^2+\|V\|^2}\big(|D_{\hat{r}}\varphi|^2 + |\sigma_{\hat{r}}\varphi|^2\big) \\
& \qquad + \frac{\|V\|^2}{t^2+\|V\|^2}|\overline{\mathcal{D}}_{\hat{r}}\varphi|^2 + |D_{\hat{\theta}}\varphi|^2 + |\sigma_{\hat{\theta}}\varphi|^2 \Bigg) \\
& - \frac{1}{4}(\bar{\varphi}\mathcal{T}_a\varphi)^2 + \frac{2\mathrm{i}}{t^2+\|V\|^2}\operatorname{Im}\big(\overline{V}^{\hat{\theta}}\partial_{\hat{r}}W\big) \\
& + \frac{t^{-2}}{t^2+\|V\|^2}\left|\frac{\partial W}{\partial\varphi}\right|^2 \Bigg) \\
& - \frac{1}{\epsilon}\int_{\partial D} W\mathrm{d}\theta - \operatorname{Tr}\big(f_{\mathrm{GF},0}(\varphi_0,\bar{\varphi}_0)^2\big)\,,
\end{aligned} \tag{111}$$

where R_{ij} is the Ricci curvature of g. The metric used in this expression is the original one before the rescaling.

The real part of the integrand of the bulk integral is a sum of squares, while the boundary integral is bounded below by the assumption on the boundary condition for φ. Looking at the terms multiplied by positive powers of t, we find that as $t \to \infty$, the action diverges away from the field configurations such that

$$F_{ij} = \nabla^i A_i = D_i\sigma_j = [\sigma_i,\sigma_j] = \operatorname{Tr}(R_{ij}\sigma^i\sigma^j) = D_i\varphi = \sigma_i\varphi = 0\,. \tag{112}$$

The path integral therefore localizes in this limit to the locus of the field space defined by these equations.[2]

[2]It is crucial here that D is compact. If D were noncompact, $D_r\varphi$, for example, could vanish as t^{-1} in the limit $t \to \infty$ but φ could still vary by a finite amount over D.

A general solution $(A_0, \sigma_0, \varphi_0)$ of the localization equations can be easily identified. The obvious solution of $F_{ij} = 0$ is the vanishing gauge field, and any flat connection on D is gauge equivalent to it under the Neumann boundary condition which preserves the gauge symmetry. This solution also satisfies the Lorenz gauge condition $\nabla^i A_i = 0$, so we have

$$A_0 = 0\,. \tag{113}$$

Also, as we have equipped D with a hemisphere metric which has a positive Ricci curvature, $\mathrm{Tr}(R_{ij}\sigma^i\sigma^j) = 0$ implies

$$\sigma_0 = 0\,. \tag{114}$$

Given the vanishing of the gauge field, $D_i\varphi = 0$ simply means that φ_0 is a constant, which by the boundary condition must belong to the submanifold γ of X:

$$\varphi_0 \in \gamma\,. \tag{115}$$

We can evaluate the path integral by perturbation theory around these localization configurations. To facilitate the calculation we write

$$A = A_0 + t^{-2}A'\,, \qquad \sigma = \sigma_0 + t^{-2}\sigma'\,, \qquad \varphi = \varphi_0 + t^{-1}\varphi'\,, \tag{116}$$

and rescale the fermions by the usual scale transformation:

$$(\alpha, \lambda, \zeta) \to (\alpha, t^{-1}\lambda, t^{-2}\zeta)\,, \qquad (\bar\eta, \rho, \bar\mu) \to (\bar\eta, t^{-1}\rho, t^{-2}\bar\mu)\,. \tag{117}$$

The rescaling suppresses the fermionic terms that contain V.

Further, for each fermion Ψ, let Ψ_0 be the part of Ψ that is a zero mode of the Laplace–de Rham operator $\Delta_{\mathrm{d}} = (\mathrm{d} - \star\,\mathrm{d}\,\star)^2$ and satisfies the relevant boundary condition, and write

$$\Psi = \begin{cases} \Psi_0 + t^{-1/2}\Psi' & (\Psi \notin \{b, c\})\,; \\ \Psi_0 + t^{-1}\Psi' & (\Psi \in \{b, c\})\,. \end{cases} \tag{118}$$

There are no zero modes for λ and ρ since there are no harmonic one-forms on a hemisphere. A zero mode for ζ would be proportional to the volume

form but this is killed by the boundary condition $\zeta_0 = 0$. Thus we have

$$\lambda_0 = \zeta_0 = \rho_0 = 0\,. \tag{119}$$

Then, η_0 and $\star\mu_0$ are constants satisfying the boundary conditions

$$(0, \bar{\eta}_0) \in T_{\varphi_0}\gamma \otimes \mathbb{C}\,, \qquad (0, \star\,\bar{\mu}_0) \in N_{\varphi_0}\gamma \otimes \mathbb{C}\,. \tag{120}$$

The zero modes b_0, c_0 of b, c are constant and not affected by the boundary conditions $\partial_r b_0 = \partial_r c_0 = 0$.

In terms of the field variables introduced above, the action reads, to the zeroth order in t,

$$\begin{aligned}\int_D \Big(& \mathrm{Tr}\big(-A' \star \Delta_{\mathrm{d}} A' - \sigma' \star \Delta_{\mathrm{d}} \sigma' + \lambda' \wedge \star(\mathrm{d} - \star\,\mathrm{d}\,\star)(\alpha' - \zeta')\big) \\ & + \frac{1}{2}(\bar{\varphi}' \star \Delta_{\mathrm{d}} \varphi' + \Delta_{\mathrm{d}} \varphi' \star \bar{\varphi}') \\ & - \rho' \wedge \star(\mathrm{d} - \star\,\mathrm{d}\,\star)(\bar{\eta}' - \bar{\mu}') - 2\mathrm{i} b' \star \Delta_{\mathrm{d}} c' \Big) \\ & + \int_{\partial D} \mathrm{d}\hat{\theta} \left(\frac{1}{2}(\bar{\varphi}' \partial_{\hat{r}} \varphi' + \partial_{\hat{r}} \bar{\varphi}' \varphi') - \bar{\mu}'_{\hat{r}\hat{\theta}} \rho'_{\hat{\theta}} \right) + S_0\,. \end{aligned} \tag{121}$$

The last term contains only φ_0, $\bar{\varphi}_0$ and the fermion zero modes:

$$\begin{aligned} S_0 = & -\frac{2\pi}{\epsilon} W(\varphi_0) + (f^a_{\mathrm{GF},0})^2 \\ & + 2\mathrm{i}(\mathcal{T}_b \cdot f^a_{\mathrm{GF},0}) b^a_0 \tilde{c}^b_0 - \frac{1}{4}(\bar{\varphi}_0 \mathcal{T}_a \varphi_0)^2 + \bar{\eta}_0 \alpha_0 \varphi_0\,. \end{aligned} \tag{122}$$

Here $\tilde{c}_0 = c_0 + \cdots$ is the constant part of the fermion $\tilde{c}$ defined earlier. The contributions from the higher order terms vanish in the limit $t \to \infty$.

Thus, to the order relevant in the limit we are interested in, the bosonic and fermionic nonzero modes (the primed variables) enter the action quadratically and can be integrated out exactly. In general, the one-loop determinant $\Delta_{\text{1-loop}}(\varphi_0, \bar{\varphi}_0)$ produced by this integration is a function on γ: even though the quadratic terms are independent of the point $\varphi_0 \in \gamma$ around which we are expanding φ, the boundary conditions do depend on it.

Now that the nonzero modes have been integrated out, we are only left with the integration over the zero modes. This step can be expressed

schematically as

$$\int_\gamma \mathrm{vol}_\gamma \int \mathrm{d}\alpha_0\, \mathrm{d}\bar{\mu}_0\, \mathrm{d}\bar{\eta}_0\, \mathrm{d}b_0\, \mathrm{d}\tilde{c}_0\, \Delta_{\text{1-loop}}(\varphi_0, \bar{\varphi}_0) e^{-S_0}\,, \tag{123}$$

where vol_γ is a volume form of γ. The final expression may be thought of as the path integral for a zero-dimensional theory, which is what we wanted to obtain.

3.2.4. Lagrangian branes and complex gauge symmetry. The first thing to notice about the integral (123) is that $\bar{\mu}_0$ is absent from the action (122); as ζ has no zero modes, the interaction term $\bar{\mu} \star \zeta\varphi$ dropped out in the localization process. For the integral to be nonvanishing, then, $\bar{\mu}$ should have no zero modes either. In the same way, the number of zero modes for $\bar{\eta}$ should equal that of α or the integral vanishes.

The numbers of zero modes for $\bar{\eta}$ and $\bar{\mu}$ depend on the boundary conditions (120). If we wish to have a nontrivial result, we should choose the submanifold γ for the support of the brane appropriately so that both of the above requirements are satisfied.

First, suppose that the gauge symmetry is trivial. Then, the theory has no vector multiplet, and we want the boundary conditions to kill $\bar{\eta}_0$ and $\bar{\mu}_0$ completely. This is achieved if we take γ to be a Lagrangian submanifold of the target space X, with respect to the Kähler form

$$\omega_X = \frac{\mathrm{i}}{2} \mathrm{d}\varphi \wedge \mathrm{d}\bar{\varphi}\,. \tag{124}$$

An interesting property of a Lagrangian submanifold of a Kähler manifold is that the action of the complex structure J interchanges the tangent and normal bundles. It follows that $(0, -\mathrm{i}\bar{\eta}_0) \in N_{\varphi_0}\gamma \otimes \mathbb{C}$ and $(0, -\mathrm{i} \star \bar{\mu}_0) \in T_{\varphi_0}\gamma \otimes \mathbb{C}$, hence $\bar{\eta}_0 = \bar{\mu}_0 = 0$, as desired.

Now suppose that the theory has a nontrivial gauge symmetry. In this case, the action contains the potential

$$(f^a_{\mathrm{GF},0})^2 - \frac{1}{4}(\bar{\varphi}_0 \mathcal{T}_a \varphi_0)^2\,. \tag{125}$$

Actually, we can rescale this potential by an arbitrarily large factor without affecting the localization argument; we just have to rescale the $Q_{V+\mathrm{B}}$-exact part of the action by that factor. Hence, for a nontrivial result, γ must intersect with the zero locus of the potential.

The zero locus is characterized by the equations $f_{\mathrm{GF},0} = 0$ and $\mathrm{i}\bar{\varphi}_0 \mathcal{T}_a \varphi_0/2 = 0$. The former is the gauge fixing condition, so we can drop

it and instead undo the gauge fixing. This puts us in a situation where we have the Kähler manifold X, endowed with a $\mathcal{G}$-action and the $\mathcal{G}$-invariant Kähler form ω_X. The quantities

$$\mu_a(\varphi_0, \bar{\varphi}_0) = \frac{\mathrm{i}}{2}\bar{\varphi}_0 \mathcal{T}_a \varphi_0 \tag{126}$$

which we are setting to zero are the moment map $\mu\colon X \to \mathrm{Lie}(\mathcal{G})^*$ for the $\mathcal{G}$-action evaluated on $\mathcal{T}_a$. By μ being the moment map, we mean that $\mathrm{d}\mu_a = \iota_{v_a}\omega_X$, where $v_a = \mathcal{T}_a\varphi_0\partial_{\varphi_0} - \bar{\varphi}_0\overline{\mathcal{T}}_a\partial_{\bar{\varphi}_0}$ is the vector field on X generated by $\mathcal{T}_a$.

As μ is $\mathcal{G}$-equivariant (that is, $\langle \mu(g\cdot x), \mathcal{T}_a\rangle = \langle \mu(x), g^{-1}\mathcal{T}_a g\rangle$), the level set $\mu^{-1}(0)$ is $\mathcal{G}$-invariant. The zero locus of the potential is homeomorphic to the quotient

$$\mathcal{M} = \mu^{-1}(0)/\mathcal{G}\,. \tag{127}$$

This is the symplectic reduction of X by the $\mathcal{G}$-action and itself a symplectic manifold. The symplectic form of $\mathcal{M}$ is naturally induced from ω_X since $\omega_X(v_a, v) = v(\mu_a) = 0$ for any vector field v tangent to $\mu^{-1}(0)$.

The equation $\mu = 0$, like the other equation $f_{\mathrm{GF},0} = 0$, can be regarded as a gauge fixing condition, albeit for a complex gauge symmetry. The $\mathcal{G}$-action on X naturally extends to a holomorphic action of the complexified gauge group $\mathcal{G}_{\mathbb{C}}$, whose Lie algebra $\mathrm{Lie}(\mathcal{G}_{\mathbb{C}})$ is spanned by $\{\mathcal{T}_a, \mathrm{i}\mathcal{T}_a\}$. The vector fields Jv_a generated by $\mathrm{i}\mathcal{T}_a$ are normal to $\mu^{-1}(0)$: for $v \in \Gamma(T\mu^{-1}(0))$, we have $g_X(Jv_a, v) = \omega_X(v_a, v) = 0$. Hence, the $\mathcal{G}_{\mathbb{C}}$-orbit $\mathcal{G}_{\mathbb{C}}\cdot x$ of a point $x \in \mu^{-1}(0)$ intersects the $\mathcal{G}$-orbit $\mathcal{G}\cdot x \subset \mu^{-1}(0)$ orthogonally. Moreover, it can be shown that every $\mathcal{G}_{\mathbb{C}}$-orbit contains in its closure at most a single $\mathcal{G}$-orbit inside $\mu^{-1}(0)$.

A point of X such that the closure of its $\mathcal{G}_{\mathbb{C}}$-orbit has a nonempty intersection with $\mu^{-1}(0)$ is said to be *semistable*. The fact just mentioned implies that $\mathcal{M}$ is homeomorphic to the quotient of the set X^{ss} of semistable points by the $\mathcal{G}_{\mathbb{C}}$-action:

$$\mathcal{M} \simeq X^{\mathrm{ss}}/\mathcal{G}_{\mathbb{C}}\,. \tag{128}$$

Put differently, imposing the condition $\mu = 0$, roughly speaking, gauge fixes the noncompact part of the complex gauge symmetry generated by $\{\mathrm{i}\mathcal{T}_a\}$. Being a quotient by a holomorphic $\mathcal{G}_{\mathbb{C}}$-action, $\mathcal{M}$ is complex, hence Kähler.

At low energies, the theory effectively becomes one without gauge symmetry whose target space is the curved Kähler manifold $\mathcal{M}$. Then, an argument similar to what we have given for flat target spaces would show that $\bar{\eta}_0$

and $\bar{\mu}_0$ should vanish when pushed forward by the projection $\pi\colon X^{\mathrm{ss}} \to \mathcal{M}$. Thus, we take γ to be the preimage of a Lagrangian submanifold $\mathcal{L}$ of $\mathcal{M}$:

$$\gamma = \pi^{-1}(\mathcal{L})\,. \tag{129}$$

This is indeed a good choice. The kernel of π_* is spanned by the vectors $(\mathcal{T}_a\varphi_0, 0)$ and $(0, \bar{\varphi}_0\mathcal{T}_a)$. These lie in $T_{\varphi_0}\gamma \otimes \mathbb{C}$, so we still have $\bar{\mu}_0 = 0$. However, $\bar{\eta}_0$ no longer needs to vanish and can be anything of the form

$$\bar{\eta}_0 = \bar{\varphi}_0\beta_0\,, \tag{130}$$

with $\beta_0 \in \mathrm{Lie}(\mathcal{G})$. The number of zero modes for $\bar{\eta}_0$ is therefore $\dim \mathcal{G}$, just as for α_0.

We call a boundary condition of the type described above a *Lagrangian brane*. Its support γ gives a Lagrangian submanifold in the symplectic reduction $\mathcal{M}$ of X.

Putting together what we have found, we conclude that the localized path integral is given by

$$\int_\gamma \mathrm{vol}_\gamma \int \prod_a \left(\mathrm{d}\alpha_0^a\, \mathrm{d}\beta_0^a\, \mathrm{d}b_0^a\, \mathrm{d}\tilde{c}_0^a\right) \Delta_{\text{1-loop}} \exp\left(\frac{2\pi}{\epsilon} W - S_0'\right), \tag{131}$$

with

$$S_0' = (f^a_{\mathrm{GF},0})^2 + 2\mathrm{i}(\mathcal{T}_b \cdot f^a_{\mathrm{GF},0}) b_0^a \tilde{c}_0^b + \mu_a^2 + (\bar{\varphi}_0 \mathcal{T}_a \mathcal{T}_b \varphi_0)\beta_0^a \alpha_0^b\,. \tag{132}$$

The measure for the fermion zero modes is the natural one induced by the metric on $\mathrm{Lie}(\mathcal{G})$.

The superpotential W is a holomorphic function of φ_0 and gauge invariant, and as such invariant under $\mathcal{G}_{\mathbb{C}}$. The domain γ of the bosonic integration is also $\mathcal{G}_{\mathbb{C}}$-invariant. The emergence of complex gauge symmetry is suggestive. Sure enough, the above integral may be interpreted as the path integral for a zero-dimensional gauged sigma model with gauge group $\mathcal{G}_{\mathbb{C}}$. This is a gauge theory described by a map φ_0 from a point to γ, and its action is given by $-2\pi W/\epsilon$. The Ω-deformation parameter ϵ plays the role of the Planck constant, so the undeformed limit $\epsilon \to 0$ is the classical limit.

Since the integral (131) is supposed to be a gauge fixed form of the path integral for this bosonic theory, the fermionic piece S_0' in the exponent must be a gauge fixing action. As we explained already, the complex gauge symmetry can be gauge fixed by the condition $f_{\mathrm{GF},0} = \mu = 0$. Denoting the ghosts

for the real and imaginary parts of $\mathcal{G}_{\mathbb{C}}$ by (b_0, c_0) and (β_0, α_0), respectively, we can write the corresponding gauge fixing action as

$$S_{\mathrm{GF},0} = (f^a_{\mathrm{GF},0})^2 + 2\mathrm{i}(\mathcal{T}_b \cdot f^a_{\mathrm{GF},0}) b^a_0 \tilde{c}^b_0 + \mu_a^2 + 2\mathrm{i}\big((\mathrm{i}\mathcal{T}_b) \cdot \mu_a\big)\beta^a_0 \alpha^b_0 \,. \tag{133}$$

Similarity between S'_0 and $S_{\mathrm{GF},0}$ is obvious, but they do not precisely match. The last term in $S_{\mathrm{GF},0}$ is $-\mathrm{i}\bar{\varphi}_0\{\mathcal{T}_a, \mathcal{T}_b\}\varphi_0\beta^a_0\alpha^b_0$, so up to a trivial rescaling of the ghosts, it differs from the last term in S'_0 by a quantity which vanishes on $\mu^{-1}(0)$. However, the effect of this discrepancy, if any, should be offset by the one-loop determinant, as we can argue as follows.

The idea is to rescale the potential μ_a^2 by a large factor via a $Q_{V+\mathrm{B}}$-exact deformation (say, by rescaling the bilinear form Tr). Then, the integral localizes to $\mu^{-1}(0)$ where the discrepancy disappears. Now we show that $\Delta_{\text{1-loop}}$ is constant on $\mu^{-1}(0)$.[3]

First, we note that the intersection $\gamma \cap \mu^{-1}(0)$ is a Lagrangian submanifold of X. This is because $\gamma \cap \mu^{-1}(0)$ is the union of $\mathcal{G}$-orbits in $\mu^{-1}(0)$ that make up the Lagrangian submanifold $\mathcal{L} \subset \mathcal{M}$. Having an isotropic image under the symplectic reduction, $\gamma \cap \mu^{-1}(0)$ is itself isotropic. Furthermore, it has dimension $\dim \mathcal{G} + \dim \mathcal{L} = \dim X/2$.

Next, pick $\varphi_0 \in \gamma \cap \mu^{-1}(0)$ and choose an orthonormal basis (e_j), $j = 1, \ldots, \dim \mathcal{G}$, of $T_{\varphi_0}(\mathcal{G} \cdot \varphi_0) \subset T_{\varphi_0}(\gamma \cap \mu^{-1}(0))$. We can extend it to an orthonormal basis (e_k), $k = 1, \ldots, \dim \mathcal{G} + \dim \mathcal{L}$, of $T_{\varphi_0}(\gamma \cap \mu^{-1}(0))$. As we saw earlier, the vectors Je_j are normal to $\mu^{-1}(0)$. They are also tangent to γ, so (e_k, Je_j) is an orthonormal basis of $T_{\varphi_0}\gamma$. On the other hand, (Je_k) is an orthonormal basis of $N_{\varphi_0}(\gamma \cap \mu^{-1}(0))$. Then, (e_k, Je_k) is an orthonormal basis of $T_{\varphi_0}X$, and $((e_k - \mathrm{i}Je_k)/\sqrt{2})$ is a unitary basis of $T^{1,0}_{\varphi_0}X$. We require the bases constructed here to vary smoothly over $\gamma \cap \mu^{-1}(0)$.

In terms of this unitary basis, the boundary conditions are described in a uniform manner, irrespective of the choice of the point $\varphi_0 \in \gamma \cap \mu^{-1}(0)$. For example, the condition $(\iota_V\rho, \bar{\eta}) \in T_{\varphi_0}\gamma \otimes \mathbb{C}$ says that $\iota_V\rho^l = \bar{\eta}^l$ for $l = \dim \mathcal{G} + 1, \ldots, \dim \mathcal{G} + \dim \mathcal{L}$. Also, the quadratic terms in the nonzero modes, from which the one-loop determinant is calculated, has a uniform expression in a unitary basis. Therefore, the one-loop determinant is independent of φ_0.

[3]In general, the argument given below cannot be applied globally, and the one-loop determinant is not a constant but a flat section of a line bundle. We will not address this issue here.

3.3. Localization on a plane

We have just seen that when the spacetime is a disk of finite radius and the boundary condition is given by a Lagrangian brane, the quasi-topological sector of the Ω-deformed B-twisted gauge theory is equivalent to a zero-dimensional gauged sigma model with complex gauge group whose target space is the support of the brane.

The case when the spacetime is a plane is similar but qualitatively different. It is similar in that the Ω-deformed B-twisted gauge theory in this case is still equivalent to a zero-dimensional gauged sigma model with the same complex gauge group. The target space, however, is different due to the noncompactness of the spacetime; it is no longer given by the brane itself. Rather, it consists of gradient flows generated by the superpotential [30], as we now show.

3.3.1. Path integral on a semi-infinite cylinder. Let us deform the spacetime $D = \mathbb{R}^2$ into the shape of a cigar, consisting of a semi-infinite cylinder capped with a hemisphere. We split the path integral on the cigar into two parts. One is performed on the hemisphere, and we already understand it well. The other is on the cylinder. Our strategy is to impose some boundary condition at infinity and see what state the latter path integral yields at the other end of the cylinder. Subsequently we feed this state into the former path integral to deduce the result of the path integral on the whole cigar.

Let D_0 and D_∞ be the hemisphere and cylinder parts of D, respectively. As usual, we can deform the action by Q_V-exact terms since this does not change the Q_V-cohomology class of the state at the end. Using this freedom we choose the metric to be such that

$$g_{\theta\theta} = \frac{1}{|\epsilon|^2} \tag{134}$$

on D_∞ so that we have $\|V\| = 1$ on the cylinder. Moreover, we make $g_{rr}(r)$ decay sufficiently fast so that D has a finite area.

The action on D is the sum of two Q_V-invariant integrals, S_{D_0} on D_0 and S_{D_∞} on D_∞. It turns out that S_{D_∞} is Q_V-exact: the part of the action that depends on W can be written as

$$\delta_V \int_{D_\infty} \frac{\mathrm{d}\theta}{V^\theta} \rho \frac{\partial W}{\partial \varphi}\,. \tag{135}$$

For the above choice of metric, the bosonic part of S_{D_∞} is given by

$$\int_{D_\infty} \mathrm{d}\hat{r}\,\mathrm{d}\hat{\theta}\Big(\mathrm{Tr}\Big(-\frac{1}{2}F^{ij}F_{ij} - D^i\sigma^j D_i\sigma_j - \frac{1}{2}[\sigma^i,\sigma^j][\sigma_i,\sigma_j]\Big) + \frac{1}{2}\Big|\overline{\mathcal{D}}_{\hat{r}}\varphi - \frac{\epsilon}{|\epsilon|}\frac{\partial\overline{W}}{\partial\bar{\varphi}}\Big|^2 + |D_{\hat{\theta}}\varphi|^2 + |\sigma_{\hat{\theta}}\varphi|^2 - \frac{1}{4}(\bar{\varphi}\mathcal{T}_a\varphi)^2\Big)\,. \tag{136}$$

Note that there are no boundary terms in this expression.

Let us "squash" the cigar in the longitudinal direction in such a way that g_{rr} is rescaled on D_∞ as

$$g_{rr} \to t^{-2} g_{rr}\,. \tag{137}$$

At the same time, we also rescale W as

$$W \to tW\,. \tag{138}$$

Both of these deformations are Q_V-exact. In the limit $t \to \infty$, the path integral on D_∞ localizes to the locus where

$$F = D_r\sigma = D_\theta\sigma_r = [\sigma_r,\sigma_\theta] = \overline{\mathcal{D}}_{\hat{r}}\varphi - \frac{\epsilon}{|\epsilon|}\frac{\partial\overline{W}}{\partial\bar{\varphi}} = 0\,. \tag{139}$$

For the path integral to be nonvanishing, the boundary condition at $r = \infty$ must be compatible with the localization equations. Then, for the vector multiplet, we should take the Neumann condition as we did in the hemisphere case. We can choose the gauge $A_r = 0$ on D_∞, and in this gauge and with this boundary condition on the vector multiplet, the above equations reduce to

$$\mathcal{A}_r = \partial_r\mathcal{A}_\theta = \partial_{\hat{r}}\varphi - \frac{\epsilon}{|\epsilon|}\frac{\partial\overline{W}}{\partial\bar{\varphi}} = 0\,. \tag{140}$$

If D_∞ were compact, with r varying over a finite interval, we would be able to shrink it to a very short cylinder so that the localization equation for φ would imply that φ does not vary in the longitudinal direction. In the case at hand, however, r is not bounded above and takes values in $[r_0,\infty)$ on D_∞ for some $r_0 > 0$. If we introduce a new coordinate

$$s = |\epsilon| \int_{r_0}^{r} \sqrt{g_{rr}}\,\mathrm{d}r\,, \tag{141}$$

which ranges from 0 to ∞, then in terms of this coordinate the equation becomes

$$\partial_s \varphi - \frac{1}{\bar{\epsilon}} \frac{\partial \overline{W}}{\partial \bar{\varphi}} = 0\,. \tag{142}$$

Therefore, φ localizes to a solution of the gradient flow equation generated by the function $\mathrm{Re}(W/\epsilon)$ on X with respect to the Kähler metric (89).

We have found that the chiral multiplet scalar should approach a gradient flow as $t \to \infty$. As we will see, for the convergence of path integral the flow must terminate at a fixed point. So we pick a submanifold γ_∞ of the critical locus $\mathrm{Crit}(W)$ of W and demand

$$\varphi \in \gamma_\infty \tag{143}$$

at $r = \infty$. The boundary condition at $r = \infty$ for the chiral multiplet is the brane-type condition characterized by γ_∞.

Now we turn our attention to the state produced by the path integral at the other boundary of D_∞. The wavefunction of this state is sharply peaked on the localization locus. In the limit $t \to \infty$, the effect of including this wavefunction in the path integral on D_0 is to impose a boundary condition that forces the bosonic fields to lie on the localization locus. For the vector multiplet this is the Neumann condition.

For the chiral multiplet, the boundary condition is a brane-type condition whose support γ_0 consists of all points $p \in X$ such that there exists a gradient flow $\varphi_s \colon \mathbb{R}_{\geq 0} \to X$ with $\varphi_0 = p$ and $\varphi_\infty \in \gamma_\infty$, namely the union of all gradient flow trajectories terminating on γ_∞. When γ_∞ is a nondegenerate critical point, γ_0 is known as a Lefschetz thimble.

3.3.2. Gradient flow trajectories as Lagrangian branes. We have reduced the path integral on a plane to the path integral on the hemisphere D_0 with a particular brane boundary condition. For the path integral on D_0 to be sensible, $\mathrm{Re}(W/\epsilon)$ had better be bounded above on the brane support γ_0 so that the boundary term does not diverge. Furthermore, for the path integral to be nonvanishing, γ_0 should be a Lagrangian brane, that is, there should be a Lagrangian submanifold $\mathcal{L}_0$ of $\mathcal{M}$ such that $\gamma_0 = \pi^{-1}(\mathcal{L}_0)$.

These requirements are satisfied if we choose the brane support γ_∞ at $r = \infty$ appropriately [52, 53]. Since W is invariant under $\mathcal{G}_\mathbb{C}$ and so is the gradient flow equation (142), gradient flows in X define gradient flows in $\mathcal{M}$ which are generated by $\mathrm{Re}(W/\epsilon)$ as a function on $\mathcal{M}$. We pick a compact

Lagrangian submanifold $\mathcal{L}_\infty$ of $\mathrm{Crit}(W) \subset \mathcal{M}$ and set

$$\gamma_\infty = \pi^{-1}(\mathcal{L}_\infty)\,. \tag{144}$$

Then we have $\gamma_0 = \pi^{-1}(\mathcal{L}_0)$, with $\mathcal{L}_0$ being the union of all gradient flow trajectories in $\mathcal{M}$ that terminate on $\mathcal{L}_\infty$.

First of all, $\mathrm{Re}(W/\epsilon)$ is bounded above on γ_0 because it is nondecreasing along a gradient flow:

$$\partial_s \,\mathrm{Re}\left(\frac{W}{\epsilon}\right) = \frac{1}{|\epsilon|^2}\frac{\partial W}{\partial \varphi}\frac{\partial \overline{W}}{\partial \bar{\varphi}} \geq 0\,. \tag{145}$$

As such, it attains the maximum value along each gradient flow when it reaches γ_∞, but this value is locally constant on $\mathrm{Crit}(W)$.

We can show that $\mathcal{L}_0$ is a middle-dimensional submanifold of $\mathcal{M}$ as follows. By the holomorphic Morse–Bott lemma, in a neighborhood of any point $p \in \mathrm{Crit}(W) \subset \mathcal{M}$ we can find local holomorphic coordinates $(z^i)_{i=1}^{\dim_{\mathbb{C}} \mathcal{M}}$ such that

$$W = W(p) + \sum_{i=1}^{n} (z^i)^2 \tag{146}$$

for some n. The Hessian of $\mathrm{Re}(W/\epsilon)$ at p has n positive, n negative and $(\dim_{\mathbb{R}} \mathcal{M} - 2n)$ zero eigenvalues. Hence, the union of gradient flows trajectories terminating at p is a submanifold of dimension n. Since $\mathcal{L}_0$ is the union of such submanifolds as p varies over the $(\dim_{\mathbb{R}} \mathcal{M}/2 - n)$-dimensional submanifold $\mathcal{L}_\infty$, it has dimension $\dim_{\mathbb{R}} \mathcal{M}/2$.

To show that $\mathcal{L}_0$ is isotropic, we use the fact that gradient flows are Hamiltonian flows generated by $\mathrm{Im}(W/\epsilon)$. Indeed, if $v = \partial_s\varphi\partial_\varphi + \partial_s\bar{\varphi}\partial_{\bar{\varphi}}$ is a vector field generating a gradient flow, we have

$$\iota_v \omega_X = \frac{\mathrm{i}}{2}\left(\frac{1}{\bar{\epsilon}}\frac{\partial \overline{W}}{\partial \bar{\varphi}}\mathrm{d}\bar{\varphi} - \mathrm{d}\varphi\frac{1}{\epsilon}\frac{\partial W}{\partial \varphi}\right) = \mathrm{d}\,\mathrm{Im}\left(\frac{W}{\epsilon}\right). \tag{147}$$

It follows that ω_X is preserved along the flows: $\mathcal{L}_v\omega_X = (\mathrm{d}\iota_v + \iota_v\mathrm{d})\omega_X = 0$. On the other hand, any differential form on $\mathcal{L}_0$ is mapped to a differential form on $\mathcal{L}_\infty$ upon pullback to $\mathcal{L}_\infty$ by gradient flows. Since ω_X vanishes on $\mathcal{L}_\infty$ by construction, the invariance of ω_X under gradient flows implies that ω_X vanishes when restricted to $\mathcal{L}_0$.

Combining what we have just found and the localization of the path integral on the hemisphere, we arrive at the main result of this section: The

quasi-topological sector of a B-twisted gauge theory with gauge group $\mathcal{G}$, subjected to an Ω-deformation on $\mathbb{R}^2$, is equivalent to a zero-dimensional gauged sigma model with gauge symmetry $\mathcal{G}_{\mathbb{C}}$ whose action is $-2\pi W/\epsilon$ and target space is a Lagrangian brane $\gamma_0 = \pi^{-1}(\mathcal{L}_0)$. The Lagrangian submanifold $\mathcal{L}_0$ of the Kähler quotient $\mathcal{M}$ consists of the gradient flow trajectories generated by $\mathrm{Re}(W/\epsilon)$, terminating on a chosen compact Lagrangian submanifold $\mathcal{L}_\infty$ of $\mathrm{Crit}(W) \subset \mathcal{M}$.

4. Four-dimensional Chern–Simons theory from six dimensions

Let us apply the result obtained in the previous section to the six-dimensional topological–holomorphic theory on $D \times \Sigma \times C$, viewing it as a B-twisted gauge theory on D.

The chiral multiplet scalars of the theory form a partial $G_{\mathbb{C}}$-connection

$$\mathcal{A} = \mathcal{A}_m \mathrm{d}x^m + A_{\bar{z}} \mathrm{d}\bar{z} \tag{148}$$

on $\Sigma \times C$. The target space X is therefore the space of such connections, with $(\mathcal{A}_m, A_{\bar{z}})$ providing holomorphic coordinates. The gauge group $\mathcal{G}$ is the group of maps from $\Sigma \times C$ to G, which is the group of gauge transformations that are constant on D. Looking at the chiral multiplet action, we see that X is endowed with the $\mathcal{G}$-invariant Kähler metric

$$g_X = -\frac{1}{2e^2} \int_{\Sigma \times C} \sqrt{g_\Sigma}\, \mathrm{d}^2x\, \mathrm{d}^2z\, \mathrm{Tr}\big(\delta\mathcal{A}^m \otimes \delta\overline{\mathcal{A}}_m + \delta\overline{\mathcal{A}}^m \otimes \delta\mathcal{A}_m + \delta A_{\bar{z}} \otimes \delta A_z + \delta A_z \otimes \delta A_{\bar{z}}\big)\,, \tag{149}$$

where $\sqrt{g_\Sigma}\, \mathrm{d}^2x$ is the volume form of Σ. The superpotential is given by the integral (65). This is not a fully gauge invariant expression; we will address this point later.

Now we take $D = \mathbb{R}^2$ and turn on an Ω-deformation using the rotation symmetry. Then, by localization the path integral reduces to the integral

$$\int \mathcal{D}\mathcal{A} \exp\left(\frac{\mathrm{i}}{\pi\hbar} \int_{\Sigma \times C} \mathrm{d}z \wedge \mathrm{CS}(\mathcal{A})\right), \tag{150}$$

with

$$\hbar = -\frac{\epsilon e^2}{2\pi^2}\,. \tag{151}$$

This is precisely the path integral for four-dimensional Chern–Simons theory with gauge group $G_{\mathbb{C}}$ [18–21].

Thus we conclude: the Ω-deformed topological–holomorphic theory on $\mathbb{R}^2 \times \Sigma \times C$ is equivalent to four-dimensional Chern–Simons theory on $\Sigma \times C$.

We still have to identify the integration domain for the localized path integral (150). With application to integrable lattice models in mind, let us do so in the case when Σ is a flat torus T^2 and C is an elliptic curve $E = \mathbb{C}/(\mathbb{Z} + \tau\mathbb{Z})$. Moreover, we take G to be either $\mathrm{U}(N)$ or a connected and simply connected compact Lie group. We parametrize Σ with periodic Cartesian coordinates (x, y), and let $\mathcal{C}_x$ and $\mathcal{C}_y$ denote the homology cycles represented by loops in the x- and y-directions. The one-cycles along E are denoted by $\mathcal{C}_a$ and $\mathcal{C}_b$, with the former corresponding to a path from $z = 0$ to 1 and the latter a path from $z = 0$ to τ.

The main task is to understand the critical locus of W in the Kähler quotient $\mathcal{M}$ of X. The Kähler form of X is

$$\omega_X = -\frac{\mathrm{i}}{2e^2} \int_{\Sigma\times C} \sqrt{g_\Sigma}\, \mathrm{d}^2x\, \mathrm{d}^2z\, \mathrm{Tr}\big(\delta\mathcal{A}^m \wedge \delta\overline{\mathcal{A}}_m + \delta A_{\bar z} \wedge \delta A_z\big)\,. \tag{152}$$

A simple computation shows that the moment map μ for the $\mathcal{G}$-action is given by the formula

$$\langle \mu, \varepsilon\rangle = -\frac{1}{e^2} \int_{\Sigma\times C} \sqrt{g_\Sigma}\, \mathrm{d}^2x\, \mathrm{d}^2z\, \mathrm{Tr}\left(\varepsilon\left(D^m\phi_m - \frac{\mathrm{i}}{2}F_{z\bar z}\right)\right), \tag{153}$$

where $\varepsilon \in \mathrm{Lie}(\mathcal{G})$ is the parameter of gauge transformation. Hence, the zero locus of μ is described by the condition

$$D^m\phi_m - \frac{\mathrm{i}}{2}F_{z\bar z} = 0\,, \tag{154}$$

and $\mathcal{M}$ is the quotient by the $\mathcal{G}$-action of the space of partial $G_{\mathbb{C}}$-connections satisfying this condition. The critical locus of W is where the equations of motion hold:

$$\mathcal{F}_{mn} = \mathcal{F}_{m\bar z} = 0\,. \tag{155}$$

Imposing the conditions (154) and (155) is the same as requiring the vanishing of the integral

$$-\int_{\Sigma\times C}\sqrt{g_\Sigma}\,\mathrm{d}^2x\,\mathrm{d}^2z\,\mathrm{Tr}\bigg(\frac{1}{2}\overline{\mathcal{F}}^{mn}\mathcal{F}_{mn}+\overline{\mathcal{F}}^{m}{}_{z}\mathcal{F}_{m\bar{z}} + \bigg(D^m\phi_m-\frac{\mathrm{i}}{2}F_{z\bar{z}}\bigg)^2\bigg). \tag{156}$$

By integration by parts we can rewrite this integral as

$$-\int_{\Sigma\times C}\sqrt{g_\Sigma}\,\mathrm{d}^2x\,\mathrm{d}^2z\,\mathrm{Tr}\bigg(\frac{1}{2}F^{mn}F_{mn}+F^{m}{}_{z}F_{m\bar{z}}-\frac{1}{4}F_{z\bar{z}}^2 + D^m\phi^nD_m\phi_n+D_z\phi^mD_{\bar{z}}\phi_m+\frac{1}{2}[\phi^m,\phi^n][\phi_m,\phi_n]\bigg). \tag{157}$$

The integrand is again a sum of nonnegative terms and must vanish separately. On $\mu^{-1}(0)\cap\mathrm{Crit}(W)$, therefore, A is a flat connection and ϕ satisfies

$$D_m\phi_n=D_z\phi_m=[\phi_m,\phi_n]=0\,. \tag{158}$$

A flat connection A on a principal G-bundle $P\to T^2\times E$ is characterized, up to gauge transformation, by the holonomies around the one-cycles in the base,

$$P\exp\bigg(\int_{\mathcal{C}_\bullet}A\bigg),\quad \bullet=x,\,y,\,a,\,b\,. \tag{159}$$

Since the fundamental group of $T^2\times E$ is abelian, the holonomies form a commuting quadruple of elements of G.

For $G=\mathrm{U}(N)$, the elements of the quadruple can be diagonalized simultaneously. Things are a little more complicated if G is not unitary. In this case, these elements can be pairwise conjugated to lie in a given maximal torus $\mathbb{T}$ of G, but in general it is not possible to put all of them into $\mathbb{T}$. Still, the moduli space of commuting quadruples has a component in which all four elements belong to the same maximal torus, and if the holonomies are generic they fall in this component. We will restrict our attention to this generic situation.

The equations $D_m\phi_n=D_z\phi_m=0$ imply that ϕ_m are left invariant by the holonomies. Under the genericity assumption, this condition requires ϕ_m to be valued in the Lie algebra $\mathfrak{t}$ of $\mathbb{T}$, and the remaining equation $[\phi_m,\phi_n]=0$

is satisfied. If we choose a gauge such that A is represented by a constant $\mathfrak{t}$-valued one-form, the same equations imply that ϕ is also a constant $\mathfrak{t}$-valued one-form.

The constant $\mathfrak{t}_{\mathbb{C}}$-valued one-form $\mathcal{A}$ define local holomorphic coordinates on $\mathrm{Crit}(W) \subset \mathcal{M}$. A better set of local holomorphic coordinates is given by

$$\tau_x = \frac{1}{2\pi\mathrm{i}} \int_{\mathcal{C}_x} \mathcal{A}\,, \qquad \tau_y = \frac{1}{2\pi\mathrm{i}} \int_{\mathcal{C}_y} \mathcal{A} \tag{160}$$

and

$$\lambda = -\frac{1}{2\pi\mathrm{i}} \left(\tau \int_{\mathcal{C}_a} A - \int_{\mathcal{C}_b} A \right) = \frac{1}{4\pi} \int_E A_{\bar{z}}\, \mathrm{d}^2 z\,. \tag{161}$$

These quantities are invariant under topologically trivial $\mathbb{T}_{\mathbb{C}}$-valued gauge transformations.

Globally, topologically nontrivial gauge transformations induce the identifications

$$\tau_x \sim \tau_x + u\,, \qquad \tau_y \sim \tau_y + u\,, \qquad \lambda \sim \lambda + u \sim \lambda + \tau u\,, \tag{162}$$

with $2\pi\mathrm{i}u \in \mathfrak{t}$ being an element of the kernel of the exponential map $\exp\colon \mathfrak{g} \to G$. Lastly, we must identify values related by the action of the Weyl group $W(G)$ of G, which is part of the gauge symmetry. Altogether, the relevant part of $\mathrm{Crit}(W) \subset \mathcal{M}$ is isomorphic to

$$\big((\mathbb{C}^\times)^r \times (\mathbb{C}^\times)^r \times E^r\big)/W(G)\,, \tag{163}$$

where r is the rank of G.

As discussed in the previous section, the integration domain for the localized path integral is the union of the gradient flow trajectories terminating on a chosen compact Lagrangian submanifold $\mathcal{L}_\infty$ of $\mathrm{Crit}(W) \subset \mathcal{M}$. An obvious Lagrangian submanifold of the moduli space (163) is the product of closed curves in each factor. For example, for the torus part $(\mathbb{C}^\times)^r \times (\mathbb{C}^\times)^r$, we can set $\mathrm{Im}\,\tau_x$ and $\mathrm{Im}\,\tau_y$ to constant elements of $\mathfrak{t}$.

It will prove useful to interpret the last factor of the moduli space (163) in the language of holomorphic vector bundles. At each point in Σ, the gauge bundle P restricts to a principal G-bundle over E. Pick a unitary representation of G and consider the vector bundle associated to this representation. The G-action extends to a $G_{\mathbb{C}}$-action, making it a $G_{\mathbb{C}}$-bundle. Since the integrability condition $\bar{\partial}_A^2 = 0$ is trivially satisfied for a dimensional reason, given a connection A there always exists a holomorphic structure on this bundle

in which the Dolbeault operator $\bar\partial$ coincides with $\bar\partial_A$. On $\mu^{-1}(0)\cap \mathrm{Crit}(W)$ where A is flat, the bundle has degree 0 (that is, topologically trivial) and is semistable. Conversely, a semistable holomorphic vector bundle of degree 0 arises in this way from a flat unitary connection, according to the Narasimhan–Seshadri theorem [54].

The relation between the coordinates on $\mathcal{M}$ and the holomorphic structure is as follows. The associated vector bundle in question is a quotient of a flat bundle over the universal cover $\mathbb{C}$ of E. Let us take a gauge in which $A_{\bar z}$ is constant and valued in $\mathfrak{t}_{\mathbb{C}}$. Then, choosing a basis $(s_i(0))_{i=1}^n$ consisting of eigenvectors of $A_{\bar z}$ in the fiber at $z=0$, we can define holomorphic sections

$$s_i(z) = \exp\bigl((z-\bar z)A_{\bar z}\bigr)s_i(0) = \exp\left(-2\pi\mathrm{i}\lambda_i\frac{\operatorname{Im} z}{\operatorname{Im}\tau}\right)s_i(0)\,, \tag{164}$$

where λ_i is the eigenvalue of λ associated with $s_i(0)$. These sections provide a basis for a local holomorphic frame. They obey the monodromy relations

$$s_i(z+1) = s_i(z)\,, \qquad s_i(z+\tau) = \exp(-2\pi\mathrm{i}\lambda_i)s_i(z)\,, \tag{165}$$

which determine the corresponding holomorphic transition functions. Thus, the parameter λ of the flat connection specifies the holomorphic structure via monodromy of holomorphic sections.

Now that we have understood the integration domain, let us come back to the more fundamental question: how do we make sense of the superpotential in the first place when it lacks gauge invariance? Fortunately, no problem arises if G is connected, which we assume.

The point is that given a homotopy $\widetilde{\mathcal{A}}\colon [0,1]\to X$ between two connections $\mathcal{A}_0$ and $\mathcal{A}_1$, we can define the difference of W evaluated for $\mathcal{A}_0$ and $\mathcal{A}_1$ in a gauge invariant manner:

$$W(\mathcal{A}_1) - W(\mathcal{A}_0) = \frac{\mathrm{i}}{e^2}\int_{[0,1]\times\Sigma\times C} \mathrm{d}z\wedge \mathrm{Tr}\bigl(\widetilde{\mathcal{F}}\wedge\widetilde{\mathcal{F}}\bigr)\,. \tag{166}$$

Here $\widetilde{\mathcal{F}}$ is the curvature of $\widetilde{\mathcal{A}}$, regarded as a connection on $[0,1]\times\Sigma\times C$. By assumption, for any two gauge equivalent connections $\mathcal{A}_0$ and $\mathcal{A}_1$ satisfying the equations of motion (155), there is a path $\tilde g\colon [0,1]\to G$ such that $\tilde g(0)$ is the identity element and $\mathcal{A}_1$ is the gauge transform of $\mathcal{A}_0$ by $\tilde g(1)$. For the homotopy $\widetilde{\mathcal{A}}$ generated by the action of $\tilde g$ on $\mathcal{A}_0$, the right-hand side of the above formula vanishes since the components of $\mathrm{d}z\wedge\widetilde{\mathcal{F}}$ along $\Sigma\times C$ are zero throughout the interval $[0,1]$. Hence, W can be made gauge invariant for connections in $\mathrm{Crit}(W)$. Also by the same formula, the value of $W(\mathcal{A}_1)$ for

a connection $\mathcal{A}_1$ equipped with a homotopy to a connection $\mathcal{A}_0$ in $\mathrm{Crit}(W)$ is determined from $W(\mathcal{A}_0)$. We only have to deal with such connections because $\mathcal{A}$ must approach a point on γ_∞ as $r \to \infty$, and γ_∞ is a submanifold of $\mathrm{Crit}(W)$.

If we choose γ_∞ inside a connected component of $\mathrm{Crit}(W)$, the definition of W on the relevant part of X boils down to a choice of a single constant as the value of W in that component. This constant may be thought of as an overall normalization factor for the path integral.

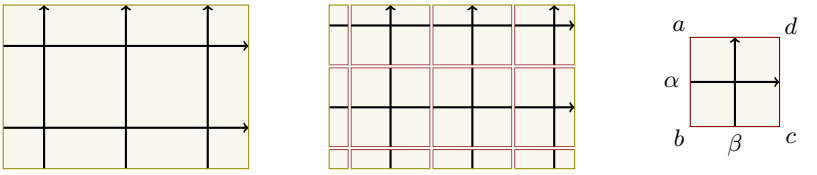
5. Integrable lattice models from four-dimensional Chern–Simons theory

Now that we have understood the six-dimensional origin of four-dimensional Chern–Simons theory, let us focus on this theory itself and explore its physical properties. In this section we explain how integrable lattice models and related mathematical structures arise from nonlocal observables of the theory. Throughout this section we take $C = E$, except for the argument in Section 5.1 which works for all choices $C = \mathbb{C}$, $\mathbb{C}^\times$ and E. Also, we take $\Sigma = T^2$ whenever the topology of Σ matters.

5.1. Line operators and integrable lattice models

As in the ordinary Chern–Simons theory, the basic observables in four-dimensional Chern–Simons theory are Wilson lines. Recall that in the six-dimensional topological–holomorphic theory there are Q-invariant Wilson lines constructed from the partial $G_{\mathbb{C}}$-connection $\mathcal{A}$, which lie in the four-manifold $M = D \times \Sigma$ and are supported at points on C. For $D = \mathbb{R}^2$ or a disk, these Wilson lines remain as good observables even after the Ω-deformation is turned on (that is, they are Q_V-invariant) if they are supported on closed curves in Σ and placed at the origin of D. They descend to Wilson lines in four-dimensional Chern–Simons theory.

In the present setup, these Wilson lines wind around various one-cycles of $\Sigma = T^2$. More generally, suppose that there are $m + n$ line operators L_α, $\alpha = 1, \ldots, m + n$, the first m of which are supported on the horizontal lines located at $(y, z) = (y_\alpha, z_\alpha)$, while the last n are supported on the vertical lines at $(x, z) = (x_\alpha, z_\alpha)$. These line operators form an $m \times n$ square lattice on T^2. The case with $(m, n) = (2, 3)$ is illustrated in Figure 1(a).



(a) (b) (c)

Figure 1: (a) A lattice formed by line operators on T^2. (b) Decomposition of the lattice into square pieces. (c) A single square piece with boundary conditions specified on the corners.

We are interested in the correlation function

$$\left\langle \prod_{\alpha=1}^{m+n} L_\alpha \right\rangle . \tag{167}$$

In order to compute this quantity, we break T^2 up into square pieces, each containing precisely two intersecting segments of line operators [46, 55]. See Figure 1(b) for an example of this decomposition.

Take a single such piece, containing line operators L_α and L_β. On the corners we pick boundary conditions,[4] which we label a, b, c and d, as in Figure 1(c). This determines Hilbert spaces assigned to the sides of the square. Let $\mathbb{V}_{ab,\alpha}$ be the Hilbert space of states on an interval with boundary conditions a on the left end and b on the right end, intersected by L_α in the middle. The path integral on the square piece produces a linear map

$$\check{R}_{\alpha\beta}\begin{pmatrix} a & d \\ b & c \end{pmatrix} : \mathbb{V}_{ab,\alpha} \otimes \mathbb{V}_{bc,\beta} \to \mathbb{V}_{ad,\beta} \otimes \mathbb{V}_{dc,\alpha} . \tag{168}$$

We call this operator an *R-matrix*.

After computing the path integral on each square piece, we can glue the pieces back together by composing the resulting R-matrices in an appropriate way. Finally, we sum over the boundary conditions specified on the

[4] To handle surfaces with corners in the framework of open-closed topological field theory, one may imagine cutting out the corners and replacing them with branes on which open strings have ends. For each corner we are choosing a boundary condition that specifies the type of the brane sitting there [55].

corners so that the fields are allowed to have all possible behaviors at those points.[5]

This procedure for computing the correlation function of line operators may be thought of as defining the partition function of a lattice model in statistical mechanics. In this lattice model, state variables (or "spins") are placed on the faces and edges of the lattice of line operators. The boundary conditions on the corners are identified with the face variables, whereas basis vectors of the Hilbert spaces on the sides of the squares are the edge variables. The matrix elements of an R-matrix encode the local Boltzmann weights for various configurations of states around a vertex of the lattice. The partition function of the lattice model is the product of the local Boltzmann weights, summed over all allowed state configurations. This is precisely what we have to calculate to reconstruct the path integral on the whole torus from those on the square pieces.

The crucial property that makes this interpretation useful is that the theory is topological on T^2.[6] This property ensures that the state space and local Boltzmann weights of the lattice model are independent of the locations of the lines or how we cut T^2 into pieces; only topology matters.

So far we have only used the structure of a two-dimensional topological field theory to establish that a collection of line operators gives rise to a lattice model. Actually, our theory has more than just this structure. It is really four-dimensional, and this fact has a profound implication [18, 19].

The two-dimensional topological invariance guarantees that the partition function of the lattice model remains unchanged when one of the lines, say a horizontal one, is moved up and down. This is true as long as it does not pass another horizontal line, at which point the topology of the lattice changes. In general, one excepts the partition function of a quantum field theory to behave badly at a singular configuration where two line operators sit on top of each other. In the present case, however, the line operators are generically located at different points on C, and the partition function should be perfectly smooth even when two lines coincide on T^2 since they are separated on C. The topological invariance on T^2 then implies that the

[5]Here we are assuming that the vacuum state of the Hilbert space for a closed string (which is mapped to the identity operator under the state–operator correspondence) can be expanded in boundary states describing branes [55].

[6]In reality, as we will see later, the topological invariance is broken due to a framing anomaly [20, 21]. For the purpose of this discussion it suffices to consider the situation where the lines making up the lattice are straight and therefore the framing anomaly plays no role.

partition function is left intact when the positions of two lines are interchanged.

Another important point is that each line in the lattice carries a continuous complex parameter, namely its coordinate on C. In the context of lattice models, this parameter is called the *spectral parameter* of the line. Hence, $\check{R}_{\alpha\beta}$ depends on the spectral parameters z_α and z_β, and by translation invariance it is a function of the difference $z_\alpha - z_\beta$. Since the theory is holomorphic on C, it should satisfy

$$[D_{\bar{z}}, \check{R}_{\alpha\beta}] = 0 \tag{169}$$

so that gauge invariant quantities constructed from the R-matrices are holomorphic in the spectral parameters.

These two properties – the commutativity of any two parallel lines and the existence of a spectral parameter assigned to each line – are what make a lattice model *integrable*. Let us quickly explain why.

Formally, we can reformulate the above lattice model in such a way that it no longer carries state variables on the faces: we simply introduce big Hilbert spaces

$$\mathbb{V}_\alpha = \bigoplus_{a,b} \mathbb{V}_{ab,\alpha} \tag{170}$$

and extend the R-matrix (168) to a linear map

$$\check{R}_{\alpha\beta}(z_\alpha - z_\beta) \colon \mathbb{V}_\alpha \otimes \mathbb{V}_\beta \to \mathbb{V}_\beta \otimes \mathbb{V}_\alpha \,, \tag{171}$$

setting the excess matrix elements to zero. With this reformulation, we can introduce the row-to-row *monodromy matrices*

$$\begin{aligned} T_\alpha(z_\alpha; z_{m+1}, \dots, z_{m+n}) \\ = \check{R}_{\alpha,m+n}(z_\alpha - z_{m+n}) \circ_{\mathbb{V}_\alpha} \cdots \circ_{\mathbb{V}_\alpha} \check{R}_{\alpha,m+1}(z_\alpha - z_{m+1}) \end{aligned} \tag{172}$$

and *transfer matrices*

$$t_\alpha(z_\alpha; z_{m+1}, \dots, z_{m+n}) = \mathrm{Tr}_{\mathbb{V}_\alpha} T_\alpha(z_\alpha; z_{m+1}, \dots, z_{m+n}) \,, \tag{173}$$

where the compositions and trace are taken in the space $\mathbb{V}_\alpha$ assigned to the horizontal edges in the αth row. These are endomorphisms of $\mathbb{V}_\alpha \otimes \mathbb{V}_{m+1} \otimes \cdots \otimes \mathbb{V}_{m+n}$ and $\mathbb{V}_{m+1} \otimes \cdots \otimes \mathbb{V}_{m+n}$, respectively. Graphically, a monodromy matrix is a horizontal line traversing segments of vertical lines,

(a) (b)

Figure 2: (a) A monodromy matrix and (b) a transfer matrix.

and a transfer matrix is obtained when the horizontal line makes a loop and comes back to the starting point; see Figure 2.

Using the transfer matrices we can express the partition function as a trace:

$$\left\langle \prod_{\alpha=1}^{m+n} L_\alpha \right\rangle = \mathrm{Tr}_{\mathbb{V}_{m+1}\otimes\cdots\otimes\mathbb{V}_{m+n}}(t_m \circ \cdots \circ t_1)\,. \tag{174}$$

If we think of the vertical direction as a time direction, we may regard the transfer matrices $t_1, \ldots, t_m$ as a sequence of discrete "time evolution operators" acting on the "total Hilbert space" $\mathbb{V}_{m+1} \otimes \cdots \otimes \mathbb{V}_{m+n}$ of the lattice model.

The commutativity of horizontal lines means that transfer matrices commute:

$$[t_\alpha(z_\alpha), t_\beta(z_\beta)] = 0\,. \tag{175}$$

(Here we have suppressed the dependence of the transfer matrices on the spectral parameters assigned to the vertical lines.) If we expand $t_\alpha(z_\alpha)$ in the powers of z_α, the expansion coefficients are themselves operators on the total Hilbert space. In this way we obtain an infinite number of "conserved charges" which commute with the time evolution operator $t_\beta(z_\beta)$. Further expanding $t_\beta(z_\beta)$ in z_β, we learn that these conserved charges mutually commute. In this sense the lattice model is said to be integrable.

To recapitulate, the correlation function of a lattice of line operators in four-dimensional Chern–Simons theory is the partition function of an integrable lattice model defined on the same lattice. The integrability is a consequence of the topological invariance on T^2 and the existence of the extra dimensions C.

In fact, we can make a stronger statement. A similar argument as above leads to the conclusion that the R-matrices satisfy the *unitarity relation*

$$\sum_e \check{R}_{\beta\alpha}\left(\begin{array}{cc} a & d \\ e & c \end{array}\,\middle|\, z_\beta - z_\alpha\right)\check{R}_{\alpha\beta}\left(\begin{array}{cc} a & e \\ b & c \end{array}\,\middle|\, z_\alpha - z_\beta\right) = \delta_{bd}\,\mathrm{id}_{\mathbb{V}_{ab,\alpha}\otimes\mathbb{V}_{bc,\beta}} \tag{176}$$

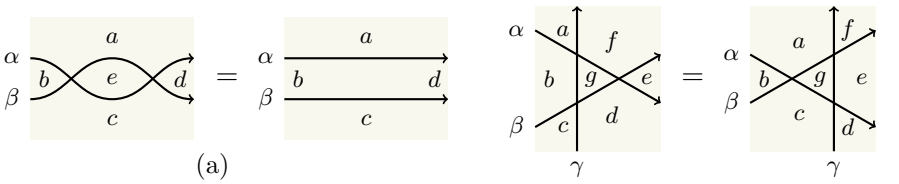


Figure 3: (a) The unitarity relation and (b) the Yang–Baxter equation.

and the *Yang–Baxter equation*

$$\sum_g \check{R}_{\alpha\beta}\left(\begin{array}{cc} f & e \\ g & d \end{array}\middle|\, z_\alpha - z_\beta\right)\check{R}_{\alpha\gamma}\left(\begin{array}{cc} a & f \\ b & g \end{array}\middle|\, z_\alpha - z_\gamma\right)\check{R}_{\beta\gamma}\left(\begin{array}{cc} b & g \\ c & d \end{array}\middle|\, z_\beta - z_\gamma\right) \tag{177}$$
$$= \sum_g \check{R}_{\beta\gamma}\left(\begin{array}{cc} a & f \\ g & e \end{array}\middle|\, z_\beta - z_\gamma\right)\check{R}_{\alpha\gamma}\left(\begin{array}{cc} g & e \\ c & d \end{array}\middle|\, z_\alpha - z_\gamma\right)\check{R}_{\alpha\beta}\left(\begin{array}{cc} a & g \\ b & c \end{array}\middle|\, z_\alpha - z_\beta\right).$$

The latter is an equality between two linear maps from $\mathbb{V}_{ab,\alpha} \otimes \mathbb{V}_{bc,\beta} \otimes \mathbb{V}_{cd,\gamma}$ to $\mathbb{V}_{af,\gamma} \otimes \mathbb{V}_{fe,\beta} \otimes \mathbb{V}_{ed,\alpha}$, and each R-matrix is implicitly tensored with an identity operator. These relations imply the commutativity of transfer matrices, hence integrability. Their graphical representations are shown in Figure 3.

5.2. Wilson lines and dynamical R-matrices

What kinds of R-matrices do we get if L_α are Wilson lines

$$W_\alpha = \mathrm{Tr}_{V_\alpha} P \exp\left(\oint \mathcal{A}\right) \tag{178}$$

in representations $G \to \mathrm{GL}(V_\alpha)$? To answer this question, we recall how we defined the path integral for our theory. The computation is done in two steps. First, we fix a $\mathfrak{t}_\mathbb{C}$-valued gauge field $\mathcal{A}^\infty$ that represents a point in the Lagrangian submanifold $\mathcal{L}_\infty$ of the moduli space (163), and integrate over the gradient flow trajectories generated by the real part of the action. The result is a function on $\mathcal{L}_\infty$. Subsequently, we integrate this function over $\mathcal{L}_\infty$.

The first step can be well approximated by perturbation theory around the background $\mathcal{A}^\infty$. In perturbation theory, the contributions to the correlation function come from the exchange of gluons between Wilson lines. (There are also vacuum and self-energy diagrams which should be taken

care of by renormalization.) The fluctuations from $\mathcal{A}^\infty$ that we integrate over are massive by construction. So if we take advantage of the topological invariance of the theory and rescale the metric on T^2 by a very large factor, the contributions from gluons traveling a finite distance in T^2 are suppressed. This argument might fail if the coupling constant increases as we take the large volume limit, but this does not happen as our theory is actually infrared free. Thus, quantum effects get localized in the vicinity of the crossings of Wilson lines. Accordingly, the correlation function factorizes into the product of local contributions associated to the vertices of the lattice. These local contributions are the R-matrices of the lattice model.

While quantum effects are important only for interactions between nearby Wilson lines, classical effects are not confined to short distances. A Wilson line may be thought of as a heavy, electrically charged particle moving along a curve. The state of this particle is labeled by a weight $\omega \in \mathfrak{t}^*_{\mathbb{C}}$ of its representation. When two such particles encounter, they exchange gluons and their states may change. Hence, a state of the system under consideration is specified by a set of weights assigned to the edges of the lattice. Each of these edges sources an electromagnetic field, which does affect charged objects at distant places.

As an example, consider a Wilson line in the state ω along a horizontal line K at $y = z = 0$. The part of the Wilson line felt by faraway objects is

$$\exp\left(\int_K \omega(\mathcal{A}^\infty)\right) = \exp\left(\int_{K\times I\times E} \mathrm{Tr}(\omega\mathcal{A}^\infty)\mathrm{d}\theta(y)\delta^2(z,\bar{z})\,\mathrm{d}^2z\right), \tag{179}$$

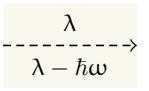
where I is an interval in the y-direction such that $K \times I \times E$ contains the Wilson line and the objects under consideration, $\theta(y)$ is a step function such that $\partial_y\theta(y) = \delta(y)$, and we have identified $\mathfrak{t}_{\mathbb{C}}$ and $\mathfrak{t}^*_{\mathbb{C}}$ via the bilinear form Tr. The presence of this factor in the path integral has the same effect on those objects as shifting $A^\infty_{\bar{z}}$ by $-\pi\hbar\omega\theta(y)\delta^2(z,\bar{z})$ over $K \times I \times E$. We must take this shift into account when computing the R-matrices.

The above analysis shows that the R-matrices depend on the effective background gauge field which differs from $\mathcal{A}^\infty$ by a shift due to the combined effect of all Wilson lines present in the system. By gauge symmetry, the R-matrices are functions of the parameter

$$\lambda \in \mathfrak{t}^*_{\mathbb{C}} \tag{180}$$

for the effective background, defined by formula (161). Its value jumps by $\hbar\omega$ across a segment of Wilson line carrying the state ω. Drawing a Wilson

line with a dashed line, we can express this jump rule graphically as follows:

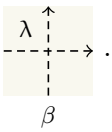
(181) 

In lattice models, this parameter λ is called the *dynamical parameter*. An R-matrix that has a dynamical parameter is known as a *dynamical R-matrix*. The appearance of dynamical R-matrices from Wilson lines was argued in [20] based on considerations in an effective two-dimensional abelian gauge theory.

When one refers to an R-matrix depending on a dynamical parameter, there is a potential confusion as to which point one is evaluating the dynamical parameter at because its value varies from place to place. We define the R-matrix

$$\check{R}_{\alpha\beta}(z_\alpha - z_\beta, \lambda) \colon V_\alpha \otimes V_\beta \to V_\beta \otimes V_\alpha \,, \tag{182}$$

arising from the crossing of two Wilson lines W_α and W_β, with respect to the dynamical parameter on the top-left face:

(183) $\check{R}_{\alpha\beta}(z_\alpha - z_\beta, \lambda) =$ 

The dynamical parameters on the other three faces are determined once states are chosen on the edges. Consistency at the bottom-right face requires that the R-matrix has zero weight, that is, $\check{R}_{\alpha\beta}$ commutes with the action of $\mathfrak{t}_{\mathbb{C}}$.

A priori, the R-matrix (183) also depends on the components of $\mathcal{A}^\infty$ along T^2. However, by a $\mathbb{T}_{\mathbb{C}}$-gauge transformation we can make $\mathcal{A}^\infty_x$ and $\mathcal{A}^\infty_y$ vanish everywhere except in the neighborhood of a single x-coordinate and a single y-coordinate, respectively. In this gauge the sole effect of these components is to twist the periodic boundary conditions with the gauge transformations by the corresponding holonomies. Hence, we conclude that the transfer matrices and the partition function are given by

$$t_\alpha = \mathrm{Tr}_{V_\alpha}\big(\exp(2\pi \mathrm{i}\tau_x)\check{R}_{\alpha,m+n} \circ_{V_\alpha} \cdots \circ_{V_\alpha} \check{R}_{\alpha,m+1}\big) \tag{184}$$

and

$$\left\langle \prod_{\alpha=1}^{m+n} W_\alpha \right\rangle = \mathrm{Tr}_{V_{m+1}\otimes\cdots\otimes V_{m+n}}\big(\exp(2\pi\mathrm{i}\tau_y)t_m \circ \cdots \circ t_1\big)\,, \tag{185}$$

where $\check{R}_{\alpha\beta}$ now refers to the R-matrix in the background with $\mathcal{A}_x^\infty = \mathcal{A}_y^\infty = 0$. The zero-weight property of $\check{R}_{\alpha\beta}$ implies that the transfer matrix has zero weight. In turn, this ensures that the partition function is independent of the choice of the row in which the twist $\exp(2\pi\mathrm{i}\tau_y)$ is inserted, as it must be by gauge invariance. By symmetry the same can be said about the choice of the column for $\exp(2\pi\mathrm{i}\tau_x)$.

Keeping track of how the dynamical parameter changes in the graphical representation of the Yang–Baxter equation, we find that the R-matrices arising from the crossings of Wilson lines obey

$$\begin{aligned}\check{R}_{\alpha\beta}(z_\alpha - z_\beta, \lambda - \hbar h_\gamma)\check{R}_{\alpha\gamma}(z_\alpha - z_\gamma, \lambda)\check{R}_{\beta\gamma}(z_\beta - z_\gamma, \lambda - \hbar h_\alpha) \\ = \check{R}_{\beta\gamma}(z_\beta - z_\gamma, \lambda)\check{R}_{\alpha\gamma}(z_\alpha - z_\gamma, \lambda - \hbar h_\beta)\check{R}_{\alpha\beta}(z_\alpha - z_\beta, \lambda)\,.\end{aligned} \tag{186}$$

Here the notation h_α means that it is to be replaced with ω when the R-matrices act on a state with weight ω in V_α. The Yang–Baxter equation of this form is known as the *dynamical Yang–Baxter equation*.

Four-dimensional Chern–Simons theory thus produces a dynamical R-matrix $\check{R}_{\alpha\beta}$, specified by a choice of the gauge group G and a pair of representations (V_α, V_β) of G. This R-matrix has zero weight and satisfies the unitarity relation

$$\check{R}_{\beta\alpha}(z_\beta - z_\alpha, \lambda)\check{R}_{\alpha\beta}(z_\alpha - z_\beta, \lambda) = \mathrm{id}_{V_\alpha\otimes V_\beta}\,. \tag{187}$$

Furthermore, by perturbation theory we can compute the R-matrix order by order in $\hbar$. At each order (except the zeroth), $\check{R}_{\alpha\beta}(z_\alpha - z_\beta, \lambda)$ diverges at $z_\alpha - z_\beta = 0$, which is the point corresponding to the situation where W_α and W_β intersect in the four-dimensional spacetime. At the first order, the divergence comes from a diagram in which a single gluon travels between the two Wilson lines in a neighborhood of the intersection, without going around one-cycles of E. Hence, if we gauge away $A_{\bar{z}}^\infty$ in this neighborhood, the singular behavior of the R-matrix is independent of the dynamical parameter to first order in $\hbar$.

5.3. Dynamical R-matrices for $G = \mathrm{U}(N)$ and $\mathrm{SU}(N)$

For $G = \mathrm{U}(N)$ and $(V_\alpha, V_\beta) = (\mathbb{C}^N, \mathbb{C}^N)$, with Tr taken to be the trace in the vector representation $\mathbb{C}^N$, Etingof and Varchenko [56] showed that a dynamical R-matrix with the properties described above is unique to all orders in perturbation theory, up to certain simple transformations and perturbative corrections to τ. It is Felder's R-matrix for the elliptic quantum group for $\mathfrak{sl}_N$ [33, 34], which first appeared as the Boltzmann weight for an integrable lattice model discovered by Jimbo, Miwa and Okado [31, 32]. For $N = 2$, the Jimbo–Miwa–Okado model reduces to the eight-vertex solid-on-solid model [57].

To state the result of [56] more precisely, we need a little preparation.

First of all, let us introduce some notations. For $G = \mathrm{U}(N)$, the complexified Cartan subalgebra $\mathfrak{t}_{\mathbb{C}}$ is the space of complex diagonal matrices. The standard basis for $\mathfrak{t}_{\mathbb{C}}$ consists of the matrices E_{ii}, $i = 1, \ldots, N$, which have 1 in the (i, i) entry and 0 elsewhere. The trace Tr identifies E_{ii} with its dual E_{ii}^*, so we can write the dynamical parameter as

$$\lambda = \sum_{i=1}^{N} \lambda_i E_{ii}^* \,, \tag{188}$$

using an N-tuple of complex numbers $(\lambda_1, \ldots, \lambda_N)$. The standard basis vector e_i of $\mathbb{C}^N$ has weight $\omega_i = E_{ii}^*$. The matrix elements of an endomorphism R of $\mathbb{C}^N \otimes \mathbb{C}^N$ are defined by $R(e_i \otimes e_j) = \sum_{k,l=1}^{N} e_k \otimes e_l R_{ij}^{kl}$.

We also need Jacobi's first theta function $\theta_1(z) = \theta_1(z|\tau)$. In terms of the theta function with characteristics

$$\theta\begin{bmatrix} a \\ b \end{bmatrix}(z|\tau) = \sum_{n=-\infty}^{\infty} e^{\pi \mathrm{i}(n+a)^2\tau + 2\pi\mathrm{i}(n+a)(z+b)} \,, \tag{189}$$

this is given by

$$\theta_1(z|\tau) = -\theta\begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix}(z|\tau) \,. \tag{190}$$

It is an odd function:

$$\theta_1(-z) = -\theta_1(z) \,. \tag{191}$$

From the identities

$$\begin{aligned}\theta\begin{bmatrix}a\\b\end{bmatrix}(z+1|\tau)&=e^{2\pi\mathrm{i}a}\theta\begin{bmatrix}a\\b\end{bmatrix}(z|\tau)\,,\\ \theta\begin{bmatrix}a\\b\end{bmatrix}(z+c\tau|\tau)&=e^{-\pi\mathrm{i}c^2\tau-2\pi\mathrm{i}c(z+b)}\theta\begin{bmatrix}a+c\\b\end{bmatrix}(z|\tau)\,,\end{aligned}\tag{192}$$

it follows that θ_1 has the following quasi-periodicity property:

$$\theta_1(z+1)=-\theta_1(z)\,,\qquad \theta_1(z+\tau)=-e^{-\pi\mathrm{i}\tau-2\pi\mathrm{i}z}\theta_1(z)\,.\tag{193}$$

We can now define Felder's R-matrix R^{F}. This is an $\mathrm{End}(\mathbb{C}^N\otimes\mathbb{C}^N)$-valued meromorphic function on $\mathbb{C}\times\mathfrak{t}^*_{\mathbb{C}}$ such that

$$\check{R}^{\mathrm{F}}(z,\lambda)=PR^{\mathrm{F}}(z,\lambda)\tag{194}$$

satisfies the dynamical Yang–Baxter equation (186) and the unitarity relation (187). Here $P\in\mathrm{End}(\mathbb{C}^N\otimes\mathbb{C}^N)$ is the swap isomorphism: $P(v\otimes w)=w\otimes v$. The matrix elements $R^{\mathrm{F}}(z,\lambda)^{kl}_{ij}$ vanishes unless $\{i,j\}=\{k,l\}$. The nonzero matrix elements are

$$\begin{aligned}R^{\mathrm{F}}(z,\lambda)^{ii}_{ii}&=1\,,\\ R^{\mathrm{F}}(z,\lambda)^{ij}_{ij}&=\frac{\theta_1(z)\theta_1(\lambda_{ij}+\hbar)}{\theta_1(z-\hbar)\theta_1(\lambda_{ij})}\,,\\ R^{\mathrm{F}}(z,\lambda)^{ji}_{ij}&=\frac{\theta_1(\hbar)\theta_1(z-\lambda_{ij})}{\theta_1(z-\hbar)\theta_1(\lambda_{ij})}\,,\end{aligned}\tag{195}$$

where $i\neq j$ and $\lambda_{ij}=\lambda_i-\lambda_j$.

Finally, let $\check{R}^{\mathrm{U}(N)}$ be the R-matrix for the crossing of two Wilson lines in the vector representation of $\mathrm{U}(N)$, and $R^{\mathrm{U}(N)}=P\check{R}^{\mathrm{U}(N)}$. As our aim is to relate $R^{\mathrm{U}(N)}$ and R^{F}, we must identify $R^{\mathrm{U}(N)}$ with an $\mathrm{End}(\mathbb{C}^N\otimes\mathbb{C}^N)$-valued meromorphic function on $\mathbb{C}\times\mathfrak{t}^*_{\mathbb{C}}$. We do this by choosing a trivialization for the rank-N holomorphic vector bundle $V_\lambda\to E$ corresponding to a flat gauge field characterized by the dynamical parameter λ.

Let us treat the dynamical parameter on the left side of a Wilson line as the background gauge field experienced by the charged particle; for instance, the Wilson line in diagram (181) is a charged particle moving in the background λ, which itself sources a gauge field and shifts the background to $\lambda-\hbar\omega$ on the right side. If the spectral parameter of this line is z, a state of the charged particle is a point in the fiber $V_\lambda|_z$. We identify the local holomorphic frame $(s_i)_{i=1}^N$ of V_λ defined by formula (164) and the standard

frame $(e_i)_{i=1}^N$ of the trivial bundle $\mathbb{C} \times \mathbb{C}^N$. With respect to this trivialization, $R^{\mathrm{U}(N)}$ is an $\mathrm{End}(\mathbb{C}^N \otimes \mathbb{C}^N)$-valued function, and its matrix elements are meromorphic functions of the spectral parameter since $D_{\bar{z}} s_i = 0$ and $[D_{\bar{z}}, R^{\mathrm{U}(N)}] = 0$. The matrix element $R^{\mathrm{U}(N)}(z_1 - z_2, \lambda)_{ij}^{kl}$ describes the process in which the state $s_i(z_1) \otimes s_j(z_2)$ in $V_\lambda|_{z_1} \otimes V_{\lambda - \hbar\omega_i}|_{z_2}$ evolves into the state $s_k(z_1) \otimes s_l(z_2)$ in $V_{\lambda - \hbar\omega_l}|_{z_1} \otimes V_\lambda|_{z_2}$.

By choosing this gauge, we have set $A_{\bar{z}}^\infty = 0$ and let the monodromies of s_i encode the dynamical parameter. For a generic value of λ, all we can do now is to rescale s_i by separate factors, so the residual gauge symmetry (apart from the Weyl group action) is given by $\mathbb{T}_C$-valued gauge transformations that are constant on E. Since we have also gauged away $\mathcal{A}_x^\infty$ and $\mathcal{A}_y^\infty$, these gauge transformations must be constant on Σ as well.

According to a theorem of Etingof and Varchenko [56], $R^{\mathrm{U}(N)}$, regarded as an $\mathrm{End}(\mathbb{C}^N \otimes \mathbb{C}^N)$-valued function as above, is related to R^{F} by a sequence of transformations. Some of these transformations can be understood as $\mathbb{T}_{\mathbb{C}}$-valued gauge transformations which are meromorphic and possibly multivalued on E. On a dynamical R-matrix $R\colon \mathbb{C} \times \mathfrak{t}_{\mathbb{C}}^* \to \mathrm{End}(\mathbb{C}^N \otimes \mathbb{C}^N)$, the gauge transformation $R \mapsto g \cdot R$ by a $\mathbb{T}_{\mathbb{C}}$-valued meromorphic function g on $\mathbb{C} \times \mathfrak{t}_{\mathbb{C}}^*$ acts by

$$g \cdot R(z_1 - z_2, \lambda) = \big(g(z_1, \lambda - \hbar h_2) \otimes g(z_2, \lambda)\big) R(z_1 - z_2, \lambda) \times \big(g(z_1, \lambda)^{-1} \otimes g(z_2, \lambda - \hbar h_1)^{-1}\big)\,. \tag{196}$$

Under gauge transformations a unitary R-matrix is mapped to a unitary R-matrix.

One of the transformations relevant for the theorem is the gauge transformation by a multivalued function of the form

$$g(z, \lambda)_j^i = \delta_j^i e^{-z(\psi(\lambda) - \psi(\lambda - \hbar\omega_i))}\,, \tag{197}$$

with $\psi(\lambda)$ being a meromorphic function on $\mathfrak{t}_{\mathbb{C}}^*$.

Another transformation involves a closed meromorphic multiplicative two-form φ on $\mathfrak{t}_{\mathbb{C}}^*$, which is a set $\{\varphi_{ij}\}$ of meromorphic functions on $\mathfrak{t}_{\mathbb{C}}^*$ such that $\varphi_{ij} = \varphi_{ji}^{-1}$ and

$$\frac{\varphi_{ij}(\lambda)\varphi_{jk}(\lambda)\varphi_{ki}(\lambda)}{\varphi_{ij}(\lambda - \hbar\omega_k)\varphi_{jk}(\lambda - \hbar\omega_i)\varphi_{ki}(\lambda - \hbar\omega_j)} = 1\,. \tag{198}$$

Its action $R \mapsto \varphi \cdot R$ is given by

$$\varphi \cdot R(z, \lambda)_{ij}^{ij} = \varphi_{ij}(\lambda) R(z, \lambda)_{ij}^{ij}\,, \tag{199}$$

with the other matrix elements unchanged. This transformation is also a gauge transformation, at least locally on $\mathfrak{t}^*_{\mathbb{C}}$. Indeed, locally we can write φ as an exact form [58]; in other words, there exist meromorphic functions $\{\xi_i\}$ such that

$$\varphi_{ij}(\lambda) = \frac{\xi_i(\lambda)\xi_j(\lambda - \hbar\omega_i)}{\xi_i(\lambda - \hbar\omega_j)\xi_j(\lambda)}\,. \tag{200}$$

Thus, the action of φ is locally the gauge transformation with $g(z,\lambda)^i_j = \delta^i_j \xi_i(\lambda)^{-1}$.

The other relevant transformations are the maps

$$R(z,\lambda) \mapsto \sigma \otimes \sigma R(z, \sigma^{-1}\cdot\lambda)(\sigma\otimes\sigma)^{-1}\,, \tag{201}$$

with σ being an element of the symmetric group S_N, acting on $\mathfrak{t}^*_{\mathbb{C}}$ and $\mathbb{C}^N$ in the obvious ways;

$$R(z,\lambda) \mapsto f(z)R(z,\lambda)\,, \tag{202}$$

with f a meromorphic function on $\mathbb{C}$ such that $f(z)f(-z) = 1$; and

$$R(z,\lambda) \mapsto R(bz, c\lambda + \mu)\,, \tag{203}$$

with b, $c \in \mathbb{C}^\times$ and $\mu \in \mathfrak{t}^*_{\mathbb{C}}$.

The map (201) is simply the action of the Weyl group, under which our R-matrix should be invariant. For $c \neq 1$, the map (203) changes the amount by which the dynamical parameter jumps across Wilson lines. So we have $c = 1$.

To constrain the remaining freedom, we look at the quasi-periodicity of $R^{\mathrm{U}(N)}$. In the gauge we are using, the holomorphic sections s_i obey the monodromy relations (165). In view of these relations, the matrix elements of $R^{\mathrm{U}(N)}$ have the quasi-periodicity property

$$\begin{aligned} R^{\mathrm{U}(N)}(z+1,\lambda)^{kl}_{ij} &= R^{\mathrm{U}(N)}(z,\lambda)^{kl}_{ij}\,,\\ R^{\mathrm{U}(N)}(z+\tau,\lambda)^{kl}_{ij} &= e^{2\pi\mathrm{i}(\lambda_{ki} - \hbar(\omega_l)_k)} R^{\mathrm{U}(N)}(z,\lambda)^{kl}_{ij}\\ &= e^{-2\pi\mathrm{i}(\lambda_{lj} + \hbar(\omega_i)_j)} R^{\mathrm{U}(N)}(z,\lambda)^{kl}_{ij}\,. \end{aligned} \tag{204}$$

Here $\omega_i = \sum_{j=1}^N (\omega_i)_j E^*_{jj}$, or $(\omega_i)_j = \delta_{ij}$. For this to be the case, $R^{\mathrm{U}(N)}$ should take the form[7]

$$R^{\mathrm{U}(N)}(z,\lambda) = f(z)\varphi \cdot R^{\mathrm{F}}(-z,\lambda)\,, \tag{205}$$

with f satisfying the quasi-periodicity relations

$$f(z+1) = f(z)\,, \qquad f(z+\tau) = e^{-2\pi \mathrm{i}\hbar} f(z)\,. \tag{206}$$

At this point there is nothing that constrains φ.

Let us turn to the case when the gauge group is $\mathrm{SU}(N)$. In this case $\mathfrak{t}_{\mathbb{C}}$ is the space of complex traceless diagonal matrices, so the dynamical parameter for $\mathrm{SU}(N)$, which we call $\bar\lambda$, obeys the constraint

$$\sum_{i=1}^N \bar\lambda_i = 0\,. \tag{207}$$

The weight $\bar\omega_i = \sum_{j=1}^N (\bar\omega_i)_j E^*_{jj}$ of e_i is given by $(\bar\omega_i)_j = \delta_{ij} - 1/N$. We refer to the background field configuration specified by a dynamical parameter $\bar\lambda$ as an $(N,0)$ *background*, for a reason that will become clear later.

To identify the R-matrix $R^{(N,0)}$ for the vector representation of $\mathrm{SU}(N)$, we consider four-dimensional Chern–Simons theory for $G = \mathrm{U}(N)$ and split the gauge field into the overall $\mathrm{U}(1)$ part and the $\mathrm{SU}(N)$ part:

$$\mathcal{A} = \mathcal{A}^{\mathrm{U}(1)} + \mathcal{A}^{\mathrm{SU}(N)}\,. \tag{208}$$

Correspondingly, the dynamical parameter splits as

$$\lambda = \lambda_0 I^* + \bar\lambda\,, \qquad \lambda_0 = \frac{1}{N}\sum_{i=1}^N \lambda_i\,, \tag{209}$$

where I^* is the dual of the identity matrix. Since $\mathcal{A}^{\mathrm{U}(1)}$ and $\mathcal{A}^{\mathrm{SU}(N)}$ are decoupled in the theory, the total R-matrix is the product of the R-matrices

[7]For $R^{\mathrm{U}(N)}(z)^{ii}_{ii}$ to have the correct quasi-periodicity, $\psi(\lambda) - 2\psi(\lambda - \hbar\omega_i) + \psi(\lambda - 2\hbar\omega_i)$ must be independent of λ for all i. A function $\psi(\lambda)$ that has this property and is invariant under the Weyl group action is a multiple of the trace $\sum_{i=1}^N \lambda_i$, but the corresponding gauge transformation acts trivially on the R-matrix. Then, the quasi-periodicity of $R^{\mathrm{U}(N)}(z)^{ii}_{ii}$ fixes that of f, and the quasi-periodicity of the other components determines the value of b and tells that μ_{ij} are integers. Shifting λ_{ij} by integers does not affect R^{F}.

$R^{\mathrm{U}(1)}$ for the U(1) part and $R^{(N,0)}$ for the SU(N) part:

$$R^{\mathrm{U}(N)}(z,\lambda) = R^{\mathrm{U}(1)}(z,\lambda_0)R^{(N,0)}(z,\bar{\lambda})\,. \tag{210}$$

The U(1) part is a scalar function of z and λ_0, while the SU(N) part depends on z and $\bar{\lambda}$.

We know $R^{\mathrm{U}(N)}(z,\lambda)^{ii}_{ii} = f(z)$ and therefore $R^{\mathrm{U}(1)}$ is independent of λ_0. The formula (205) for $R^{\mathrm{U}(N)}$ then implies that $\varphi \cdot R^{\mathrm{F}}$ is a function of $\bar{\lambda}$ and not of λ_0. As R^{F} is independent of λ_0, so is φ. Thus, we can write

$$R^{(N,0)}(z,\bar{\lambda}) = f^{(N,0)}(z)\xi^{-1} \cdot R^{\mathrm{F}}(-z,\bar{\lambda})\,, \tag{211}$$

where we have expressed the action of φ as the gauge transformation by a diagonal matrix $\xi^{-1} = \mathrm{diag}(\xi_1^{-1},\dots,\xi_N^{-1})$ of meromorphic functions of $\bar{\lambda}$. By considering the monodromies of $R^{(N,0)}$ as in the U(N) case, we deduce

$$\begin{aligned} &f^{(N,0)}(z+1) = f^{(N,0)}(z)\,,\\ &f^{(N,0)}(z+\tau) = e^{-2\pi\mathrm{i}\hbar(N-1)/N} f^{(N,0)}(z)\,. \end{aligned} \tag{212}$$

The unitarity relation requires

$$f^{(N,0)}(z)f^{(N,0)}(-z) = 1\,. \tag{213}$$

In Sections 5.7 and 5.8, we will obtain more conditions on $f^{(N,0)}$ from considerations on framing anomaly and junctions of Wilson lines.

5.4. Surface operators and nondynamical R-matrices

Just as an electrically charged particle moving in spacetime creates a Wilson line, the worldline of a magnetically charged particle is also a line operator. This operator is called an *'t Hooft line operator* if the particle is a magnetic monopole. More generally, a dyon, which carries both electric and magnetic charges, creates a Wilson–'t Hooft line operator [59].

Suppose that in addition to Wilson lines, we have 't Hooft lines lying in Σ and supported at points on E in four-dimensional Chern–Simons theory. The inclusion of 't Hooft lines in the path integral means that the gauge field has a prescribed behavior such that as the distance from any of these lines tends to zero, the gauge field approaches the corresponding monopole configuration.

For monopoles to originate from Q_V-invariant configurations in six dimensions and have a classical interpretation as a particle, their field configurations, away from the points at which they are located, should be solutions of the semistability condition (154) and the equations of motion (155). Although we analyzed these equations in Section 4, there we assumed that all fields were nonsingular, which may not be the case in the presence of monopoles. We have to reexamine the analysis to incorporate possible singularities.

Let U be the union of small disks in E, each centered at the location of a monopole where some fields may become singular. Performing integration by parts on the integral (156) as before, but this time taking $C = E \setminus U$, we find that this integral equals the bulk integral (157) plus the boundary term

$$-\int_{\Sigma\times\partial C}\sqrt{g_\Sigma}\,\mathrm{d}^2x\,\mathrm{Tr}\big(2\phi^m(F_{mz}\mathrm{d}z+F_{m\bar z}\mathrm{d}\bar z)\big)\,. \tag{214}$$

For solutions of the equations of motion, this term equals

$$\begin{aligned}\mathrm{i}\int_{\Sigma\times\partial C}\sqrt{g_\Sigma}\,\mathrm{d}^2x(\mathrm{d}z\partial_z-\mathrm{d}\bar z\partial_{\bar z})\,\mathrm{Tr}(\phi^m\phi_m)\\ =-\int_{\Sigma\times\partial U}\sqrt{g_\Sigma}\,\mathrm{d}^2x\,\mathrm{d}\theta\,r\partial_r\,\mathrm{Tr}(\phi^m\phi_m)\,,\end{aligned} \tag{215}$$

where (r,θ) are the polar coordinates around the monopoles (defined by $2\bar z = re^{\mathrm{i}\theta}$; recall the definition (23) of z).

We know that ϕ is constant in the absence of monopoles, so $-\mathrm{Tr}(\phi^m\phi_m)$ should decay to a constant as r increases. Then the boundary term is nonpositive. There are two possibilities: either the boundary term remains nonzero as we send the radii of the disks to zero, or it vanishes in this limit.

In the former case, the previous argument based on the positivity of the terms in the integrand fails. As a result, the characterization of semistable solutions of the equations of motion is altered, a complication we want to avoid. We will not pursue this possibility in this paper.

Therefore we consider the latter possibility. The previous argument then goes through, and the semistable solutions are still parametrized by the same data as in the case with no monopoles, as long as we stay away from singularities. In particular, the curvature of the gauge field vanishes everywhere except at the points on E where the 't Hooft lines are placed. Such tightly confined magnetic fluxes are familiar: they are Dirac strings attached to the monopoles.

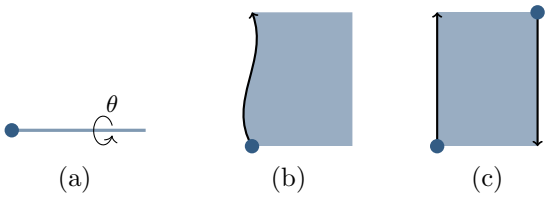


Figure 4: (a) A Dirac string emanating from a monopole. (b) The motion of the monopole creates an 't Hooft operator bounding a surface operator. (c) A surface operator formed by a Dirac string stretched between a monopole–antimonopole pair. Here the antimonopole is represented as a monopole moving in the reverse direction.

As the monopoles move, their Dirac strings sweep out surfaces. Hence, these 't Hooft lines are really the boundaries of *surface operators*. Since the spacetime is compact in the present setup, every Dirac string emanating from a monopole must be eventually absorbed by other monopoles. For example, a Dirac string may be suspended between a pair of monopoles with opposite charges. The introduction of 't Hooft lines thus divides Σ into distinct regions supporting various surface operators. See Figure 4 for illustrations.

The signature of a confined magnetic flux is the Aharonov–Bohm effect, the phase shift in the wavefunction as an electrically charged particle travels around the flux. Near the location of a Dirac string in E, the gauge field behaves as

$$A = \mathrm{i}\alpha\,\mathrm{d}\theta + \cdots\,, \tag{216}$$

where $\mathrm{i}\alpha \in \mathfrak{t}$ and the ellipsis refers to terms less singular than $1/r$ as $r \to 0$. The gauge transformation by $g = \exp(\mathrm{i}u\theta)$, with $2\pi\mathrm{i}u \in \ker\exp|_{\mathfrak{t}}$, shifts α by u, so the singular behavior of the gauge field is characterized by the holonomy $\exp(2\pi\mathrm{i}\alpha)$ around the singularity. Surface operators that induce nontrivial monodromies in fields are often called Gukov–Witten surface operators [60].

In the familiar story of monopoles, one requires this monodromy to be the identity so that the Dirac sting is unobservable, and this leads to the quantization of monopole charges. Here, the quantization condition needs not be satisfied. If the monodromy is nontrivial, the Dirac string is physical, hence so is the surface operator it creates. In that case the 't Hooft line is not a genuine line operator as it cannot exist by itself without having to bound a physical surface operator.

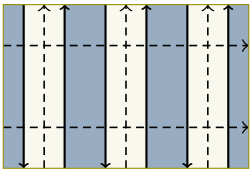


Figure 5: A lattice of Wilson lines in the presence of surface operators.

To better understand these surface operators, consider first a situation in which none of them are present in the system. Part of the data specifying a semistable solution of the equations of motion is a flat G-bundle over E. Such a bundle is characterized by the holonomies a, b of the gauge field around the one-cycles $\mathcal{C}_a$, $\mathcal{C}_b$ of E. They satisfy the relation

$$aba^{-1}b^{-1} = e\,, \tag{217}$$

where e is the identity element of G.

Now suppose that we put surface operators at a point $p \in E$, covering some region of Σ whose boundaries extend in the y-direction, as in Figure 5. As a result of the introduction of the surface operators, the holonomies are modified in this region, where instead of the above relation they obey

$$aba^{-1}b^{-1} = \exp(2\pi \mathrm{i}\alpha)\,. \tag{218}$$

While a pair (a, b) satisfying this modified relation corresponds to a flat G-bundle over $E \setminus \{p\}$, this bundle cannot be extended to a flat G-bundle over all of E. The right-hand side becomes the identity only if we project the equation to the quotient of G by a normal subgroup N containing $\exp(2\pi \mathrm{i}\alpha)$. This means that (a, b) still describes a flat bundle over E only if the structure group can be reduced to G/N. The surface operators thus modify the gauge bundle in a rather drastic way.

This modification of the gauge bundle is of a special kind [50]. The surface operators map a solution of the equations of motion to another solution. In particular, we have

$$\mathcal{F}_{x\bar{z}} = F_{x\bar{z}} - \mathrm{i}D_{\bar{z}}\phi_x = 0 \tag{219}$$

throughout Σ, provided that we are away from $p \in E$. In a gauge in which $A_x = 0$, this equation reads

$$\partial_x A_{\bar{z}} = \mathrm{i}D_{\bar{z}}\phi_x\,. \tag{220}$$

This shows that along the x-direction $A_{\bar{z}}$ varies by gauge transformations, and the holomorphic structure defined by $A_{\bar{z}}$ remains unchanged. Therefore, the holomorphic vector bundles associated to a unitary representation of G, before and after the modification, are isomorphic on $E \setminus \{p\}$.

If the normal subgroup N acts trivially in the chosen representation, the modified bundle can be extended to E as a $(G/N)_{\mathbb{C}}$-bundle. In this situation the surface operators modify a holomorphic $(G/N)_{\mathbb{C}}$-bundle over E to another holomorphic $(G/N)_{\mathbb{C}}$-bundle over E, which is isomorphic to the original one on $E \setminus \{p\}$. Such a modification of a holomorphic vector bundle over a Riemann surface is known as a *Hecke modification*.

In relation to integrable lattice models, the case of particular interest is when $G = \mathrm{SU}(N)$ and

$$\alpha = \mathrm{diag}\left(1 - \frac{1}{N}, -\frac{1}{N}, \ldots, -\frac{1}{N}\right), \tag{221}$$

or any choice of α related to this one by a permutation of the diagonal entries.[8] In this case the monodromies are represented by $N \times N$ matrices A, B with determinant 1, satisfying the relation

$$ABA^{-1}B^{-1} = e^{-2\pi \mathrm{i}/N} I\,. \tag{222}$$

The right-hand side of this relation is a generator of the center $\mathbb{Z}_N$ of $\mathrm{SU}(N)$. Thus, the pair (A, B) defines a flat vector bundle over E with structure group $\mathrm{PSU}(N) = \mathrm{SU}(N)/\mathbb{Z}_N$.

The reason this surface operator is interesting is that relation (222) determines (A, B) uniquely up to gauge transformation: we can take them to be the matrices defined by

$$Ae_k = e^{\pi \mathrm{i}(N-1)/N} e^{-2\pi \mathrm{i}k/N} e_k\,, \qquad Be_k = e_{k+1}\,. \tag{223}$$

In other words, the flat $\mathrm{PSU}(N)$-bundle over E has no moduli. Consequently, the R-matrix arising from the crossing of a pair of Wilson lines in this

[8] In the rest of this section we only consider 't Hooft lines and surface operators whose charges have $\pm 1 - 1/N$ in the first entry. The distinction between surface operators with charges related by permutations is meaningful in a generic background.

background has no dynamical parameter, and satisfies the ordinary Yang–Baxter equation

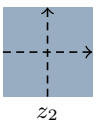
$$\begin{aligned}(224)\qquad &\check{R}_{\alpha\beta}(z_\alpha - z_\beta)\check{R}_{\alpha\gamma}(z_\alpha - z_\gamma)\check{R}_{\beta\gamma}(z_\beta - z_\gamma)\\ &\quad = \check{R}_{\beta\gamma}(z_\beta - z_\gamma)\check{R}_{\alpha\gamma}(z_\alpha - z_\gamma)\check{R}_{\alpha\beta}(z_\alpha - z_\beta)\,.\end{aligned}$$

With respect to holomorphic frames on the holomorphic vector bundles associated to the representations of the Wilson lines, the R-matrix is represented by a matrix of meromorphic function on E.

Although the form of the R-matrix generally depends on the location $p \in E$ of the surface operator in which the Wilson lines are placed, for a suitable choice of holomorphic frames this dependence disappears. (We implicitly assumed that such a choice was made when we wrote down the Yang–Baxter equation above.) This is because if we change the location of the surface operator from p to p', the associated bundles over $E \setminus \{p\}$ change to new bundles over $E \setminus \{p'\}$, but the two sets of bundles are isomorphic on $E \setminus \{p, p'\}$ since they are both isomorphic there to the set of bundles we originally had before the introduction of the surface operator. It follows that there exists a choice of holomorphic frames on the relevant bundles with respect to which the form of the R-matrix remains unchanged under the shift in the location, at all point in $E \setminus \{p, p'\}$, hence on the whole E.

Let us call the field configuration for this surface operator the $(N, 1)$ *background*, and let $R^{(N,1)}$ denote the R-matrix for the crossing of two Wilson lines in the vector representation in this background:

$$(225)\qquad R^{(N,1)}(z_1 - z_2) = \quad$$



.

The associated bundle over $E \setminus \{p\}$ has holomorphic sections

$$(226)\qquad \tilde{s}_i(z) = P\exp\left(-\int_0^z A\right)\tilde{s}_i(0)\,.$$

(We have to be a little careful about the choice of the counter for the integral in the exponent because of the singularity of A.) With respect to the holomorphic frame $(\tilde{s}_i)_{i=1}^N$, we have $A_{\bar{z}} = 0$ identically and the dependence on the location of the surface operator disappears. In this frame $R^{(N,1)}$ is an $\mathrm{End}(\mathbb{C}^N \otimes \mathbb{C}^N)$-valued meromorphic function with the quasi-periodicity

property

$$
\begin{aligned}
R^{(N,1)}(z+1) &= A_1 R^{(N,1)}(z) A_1^{-1} = A_2^{-1} R^{(N,1)}(z) A_2 \,, \\
R^{(N,1)}(z+\tau) &= B_1 R^{(N,1)}(z) B_1^{-1} = B_2^{-1} R^{(N,1)}(z) B_2 \,.
\end{aligned} \tag{227}
$$

We have introduced the notation $X_1 = X \otimes I$ and $X_2 = I \otimes X$ for $X \in \mathrm{End}(\mathbb{C}^N)$.

There is a well-known R-matrix which almost has the same quasi-periodicity. The Baxter–Belavin R-matrix R^{B} [35–37] is a unitary solution of the Yang–Baxter equation (224) satisfying the relations

$$
\begin{aligned}
R^{\mathrm{B}}(z+1) &= A_1 R^{\mathrm{B}}(z) A_1^{-1} = A_2^{-1} R^{\mathrm{B}}(z) A_2 \,, \\
R^{\mathrm{B}}(z+\tau) &= e^{2\pi \mathrm{i}\hbar(N-1)/N} B_1 R^{\mathrm{B}}(z) B_1^{-1} \\
&= e^{2\pi \mathrm{i}\hbar(N-1)/N} B_2^{-1} R^{\mathrm{B}}(z) B_2 \,.
\end{aligned} \tag{228}
$$

It is an $\mathrm{End}(\mathbb{C}^N \otimes \mathbb{C}^N)$-valued meromorphic function whose matrix elements are given by [61]

$$
R^{\mathrm{B}}(z)^{kl}_{ij} = \delta_{i+j,k+l} \frac{\theta_1(\hbar)}{\theta_1(z+\hbar)} \frac{\theta^{(k-l)}(z+\hbar)}{\theta^{(k-i)}(\hbar)\theta^{(i-l)}(z)} \frac{\prod_{m=0}^{N-1} \theta^{(m)}(z)}{\prod_{n=1}^{N-1} \theta^{(n)}(0)} \,, \tag{229}
$$

where the indices are understood modulo N and

$$
\theta^{(j)}(z|\tau, N) = \theta \begin{bmatrix} 1/2 - j/N \\ 1/2 \end{bmatrix} (z|N\tau) \,. \tag{230}
$$

For $N = 2$, these matrix elements reduce to the local Boltzmann weights for the eight-vertex model [35, 36].

Comparing the quasi-periodicity of $R^{(N,1)}$ and R^{B}, it is fairly natural to identify these two R-matrices:

$$
R^{(N,1)}(z) = f^{(N,1)}(z) R^{\mathrm{B}}(z) \,. \tag{231}
$$

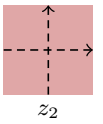
Here $f^{(N,1)}$ is a function that accounts for the slight discrepancy in the quasi-periodicity, and satisfies $f^{(N,1)}(z) f^{(N,1)}(-z) = 1$ so that the unitarity is preserved. In fact, as explained in [21], a theorem proved by Belavin and Drifeld [62] on the classification of the solutions of the classical Yang–Baxter equation ensures that $R^{(N,1)}$ must be of this form to all orders in $\hbar$, up to reparametrizations of $\hbar$.

We can also consider the $(N,-1)$ *background* created by the surface operator with the opposite charge,

$$\alpha = \operatorname{diag}\left(-1+\frac{1}{N},\frac{1}{N},\dots,\frac{1}{N}\right), \tag{232}$$

and identify the R-matrix $R^{(N,-1)}$ that arises from the crossing of Wilson lines in this background. Graphically we distinguish the $(N,-1)$ background from the $(N,1)$ background by using a different color:

$$R^{(N,-1)}(z_1-z_2) = \quad z_1 \quad z_2 \,. \tag{233}$$



Since $A^{-1}BAB^{-1} = e^{2\pi \mathrm{i}/N} I$, in an appropriate gauge this R-matrix should obey the quasi-periodicity relations

$$\begin{aligned} R^{(N,-1)}(z+1) &= A_1^{-1} R^{(N,-1)}(z) A_1 = A_2 R^{(N,-1)}(z) A_2^{-1}\,, \\ R^{(N,-1)}(z+\tau) &= B_1 R^{(N,-1)}(z) B_1^{-1} = B_2^{-1} R^{(N,-1)}(z) B_2\,. \end{aligned} \tag{234}$$

Noting that $A^T = A$ and $B^T = B^{-1}$, we see that $R^{(N,-1)}$ can be written as

$$R^{(N,-1)}(z) = f^{(N,-1)}(z) R^{\mathrm{B}}(z)^T\,, \tag{235}$$

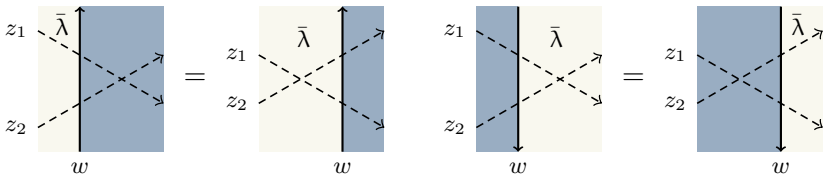
where $R^{\mathrm{B}}(z)^T$ is the transpose of $R^{\mathrm{B}}(z)$.

We have already encountered a function that has the right properties to be $f^{(N,1)}$ or $f^{(N,-1)}$: the function $f^{(N,0)}$ which enters the definition (211) of the dynamical R-matrix $R^{(N,0)}$. We will argue in Section 5.5 that the three functions are actually equal.

The relation between modification of bundles and that of R-matrices discussed here had been previously considered in [63, 64]. In particular, the R-matrices in the presence of more general surface operators were studied in [64]. In general, the R-matrices depend on l moduli, with $0 \le l \le N-1$.

5.5. Intertwining operators and vertex–face correspondences

Once we have new line operators, we can construct new R-matrices. Especially interesting are the R-matrices that correspond to a Wilson line crossing an 't Hooft line and moving into a surface operator. The Yang–Baxter equations involving two Wilson lines and one 't Hooft line, such

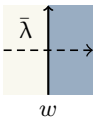


(a) (b)

Figure 6: Vertex–face correspondences between $R^{(N,0)}$ and $R^{(N,1)}$.

as the ones illustrated in Figure 6, show that these R-matrices intertwine the dynamical R-matrix and nondynamical ones. This kind of relation between dynamical and nondynamical R-matrices is known as a *vertex–face correspondence* [31, 32, 57], for the two R-matrices may be regarded as the Boltzmann weights for lattice models of "face type" and "vertex type," respectively.

Let S be the intertwining operator between $R^{(N,0)}$ and $R^{(N,1)}$ that arises from the the crossing of a Wilson line in the vector representation and an 't Hooft line of charge $\mathrm{diag}(1-1/N, -1/N, \dots, -1/N)$:

(236) $$S(z-w,\bar{\lambda}) = $$ 

By translation invariance S is a function of the difference of the spectral parameters of the two lines, which we have written as z and w here; unlike the location of the bulk of the surface operator, that of the 't Hooft line is a physical parameter. It also depends on the value of the dynamical parameter $\bar{\lambda}$ in the region adjacent to the 't Hooft line.

With respect to the local holomorphic frames we have been using for the relevant bundles, S is an $\mathrm{End}(\mathbb{C}^N)$-valued meromorphic function and satisfies the quasi-periodicity relations

(237) $$S(z+1,\bar{\lambda}) = AS(z,\bar{\lambda})\,, \qquad S(z+\tau,\bar{\lambda}) = BS(z,\bar{\lambda})\exp(-2\pi\mathrm{i}\bar{\lambda})\,.$$

In perturbation theory, we expect $S(z,\bar{\lambda})$ to have poles at $z=0$ where the Wilson and 't Hooft lines intersect in the four-dimensional spacetime.

A matrix $\Phi(z,\bar\lambda)$ that has the right quasi-periodicity and pole structure is given by [63]

$$\Phi(z,\bar\lambda)_i^j = \frac{\theta^{(j)}(z+N\bar\lambda_i)}{\theta_1(z)^{1/N}}\,. \tag{238}$$

In fact, Φ is an intertwining operator relating R^{F} and R^{B} [31]:

$$\begin{aligned} R^{\mathrm{B}}(z_1-z_2)\Phi_1(z_1,\bar\lambda)\Phi_2(z_2,\bar\lambda-\hbar h_1) \\ = \Phi_2(z_2,\bar\lambda)\Phi_1(z_1,\bar\lambda-\hbar h_2)\Theta^{-1}\cdot R^{\mathrm{F}}(z_2-z_1,\bar\lambda)\,. \end{aligned} \tag{239}$$

Here $\Theta=\mathrm{diag}(\Theta_1,\dots,\Theta_N)$ is the diagonal matrix of meromorphic functions with

$$\Theta_i(\lambda) = \prod_{j(\neq i)} \theta_1(\lambda_{ij})\,, \tag{240}$$

acting on R^{F} by the gauge transformation (196).

Given the expression (211) for $R^{(N,0)}$, the above consideration suggests that we have

$$S(z,\bar\lambda) = \Phi(z+d,\bar\lambda)(\Theta^{-1}\xi)(\bar\lambda)\bar g(z,\bar\lambda)\,, \tag{241}$$

where $d\in\mathbb{C}$ and $\bar g$ is a diagonal matrix of meromorphic functions on $E\times\mathfrak{t}^*$ that acts trivially on $R^{(N,0)}$. Let us further assume that the two functions $f^{(N,0)}$ and $f^{(N,1)}$ in formulas (211) and (231) are equal:

$$f^{(N,0)} = f^{(N,1)}\,. \tag{242}$$

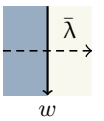
Then, with this form of S, the following vertex–face correspondence holds:

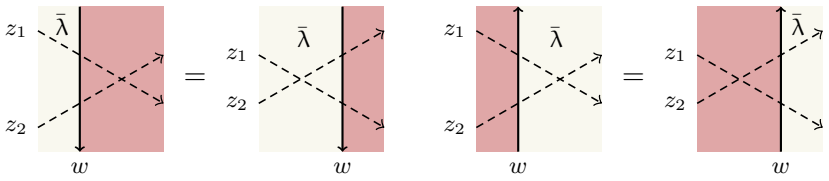
$$\begin{aligned} R^{(N,1)}(z_1-z_2)S_1(z_1-w,\bar\lambda)S_2(z_2-w,\bar\lambda-\hbar h_1) \\ = S_2(z_2-w,\bar\lambda)S_1(z_1-w,\bar\lambda-\hbar h_2)R^{(N,0)}(z_1-z_2,\bar\lambda)\,. \end{aligned} \tag{243}$$

The two sides of this relation are represented by the diagrams in Figure 6(a).

A Wilson line coming out of the surface operator produces another intertwining operator:

$$S'(z-w,\bar\lambda) = \tag{244}$$





(a) (b)

Figure 7: Vertex–face correspondences between $R^{(N,0)}$ and $R^{(N,-1)}$.

It satisfies the relation

$$R^{(N,0)}(z_1-z_2,\bar\lambda)S_2'(z_2-w,\bar\lambda-\hbar h_1)S_1'(z_1-w,\bar\lambda) = S_1'(z_1-w,\bar\lambda-\hbar h_2)S_2'(z_2-w,\bar\lambda)R^{(N,1)}(z_1-z_2)\,, \tag{245}$$

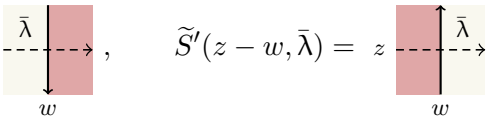
which is the vertex–face correspondence in Figure 6(b). This relation suggests that S' is essentially the inverse of S. Hence, we propose that it can be written as

$$S'(z,\bar\lambda) = \chi(z,\bar\lambda)S(z+\delta,\bar\lambda)^{-1}\,, \tag{246}$$

where $\delta\in\mathbb{C}$ and χ is a diagonal matrix of meromorphic functions that acts trivially on $R^{(N,0)}$. We will determine δ and χ in Sections 5.7 and 5.8.

The intertwining operators involving the $(N,-1)$ background can be identified in a similar manner. Let us write

$$\widetilde S(z-w,\bar\lambda) = \quad , \qquad \widetilde S'(z-w,\bar\lambda) = \quad , \tag{247}$$



and define a matrix $\widetilde\Phi(z,\bar\lambda)$ by

$$\widetilde\Phi(z,\bar\lambda)_i^j = \Phi(z,-\bar\lambda+\hbar\bar\omega_j)_j^i\,. \tag{248}$$

If we assume

$$f^{(N,0)} = f^{(N,-1)} \tag{249}$$

and

$$\begin{aligned}\widetilde S(z,\bar\lambda) &= \widetilde S'(z+\tilde\delta,\bar\lambda)^{-1}\tilde\chi(z,\bar\lambda)\,,\\ \widetilde S'(z,\bar\lambda) &= \tilde g(z,\bar\lambda)\xi^{-1}(\bar\lambda)\widetilde\Phi(z+\tilde d,\bar\lambda)\end{aligned} \tag{250}$$

for some diagonal matrices $\tilde{g}(z,\bar{\lambda})$ and $\tilde{\chi}(z,\bar{\lambda})$ acting trivially on $R^{(N,0)}$, then using the identities

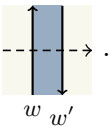
$$\begin{aligned} R^{\mathrm{F}}(z,-\bar{\lambda})^T &= \Theta^{-1}\cdot R^{\mathrm{F}}(z,\bar{\lambda})\,, \\ R^{\mathrm{F}}\big(z,\bar{\lambda}+\hbar(h_1+h_2)\big) &= R^{\mathrm{F}}(z,\bar{\lambda}) \end{aligned} \tag{251}$$

we can verify that $\widetilde{S}$ and $\widetilde{S}'$ furnish the vertex–face correspondences between $R^{(N,0)}$ and $R^{(N,-1)}$, shown in Figure 7.

5.6. L-operators

Now consider a surface operator stretched between two antiparallel 't Hooft lines, and a Wilson line traversing it. This configuration defines an *L-operator* $L^{(N,0)}$:

$$L^{(N,0)}(z-w,z-w') = \quad z \qquad . \tag{252}$$



The Wilson line shifts the dynamical parameters on the two sides of the surface operator by amounts depending on the states on the left and right edges. Hence, we may think of $L^{(N,0)}$ as a matrix whose entries are difference operators.

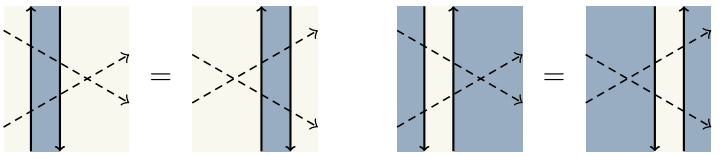
More precisely, we define the matrix element $L^{(N,0)}(z-w,z-w')^j_i$ to be a difference operator acting on a Weyl-invariant meromorphic function f on $\mathfrak{t}^*_{\mathbb{C}}\times\mathfrak{t}^*_{\mathbb{C}}$ as

$$\begin{aligned} &L^{(N,0)}(z-w,z-w')^j_i f(\bar{\lambda},\bar{\mu}) \\ &\qquad = S'(z-w,\bar{\mu})^j_k S(z-w',\bar{\lambda})^k_i f(\bar{\lambda}-\hbar\bar{\omega}_i,\bar{\mu}-\hbar\bar{\omega}_j)\,. \end{aligned} \tag{253}$$

Then, the vertex–face correspondences (243) and (245) imply that $L^{(N,0)}$ satisfies the following *RLL relation* with $R^{(N,0)}$:

$$\begin{aligned} &R^{(N,0)}(z_1-z_2,\bar{\mu})L_1^{(N,0)}(z_1-w,z_1-w')L_2^{(N,0)}(z_2-w,z_2-w') \\ &\quad = :L_2^{(N,0)}(z_2-w,z_2-w')L_1^{(N,0)}(z_1-w,z_1-w')R^{(N,0)}(z_1-z_2,\bar{\lambda}):\,. \end{aligned} \tag{254}$$

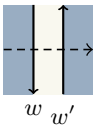
The normal ordering sign : : means that the matrix elements of $R^{(N,0)}$ should be placed in the leftmost position so as not to be acted on by the L-operators. This relation is depicted in Figure 8(a).



(a) (b)

Figure 8: RLL relations for (a) $R^{(N,0)}$ and (b) $R^{(N,1)}$.

Interchanging the intertwining operators we get another L-operator:

(255) $$L^{(N,1)}(z-w,z-w') = $$ 

This is a matrix of difference operators acting on Weyl-invariant meromorphic functions on $\mathfrak{t}^*_{\mathbb{C}}$ by

$$(256)\quad L^{(N,1)}(z-w,z-w')^j_i f(\bar\lambda) = S(z-w,\bar\lambda)^j_k S'(z-w',\bar\lambda)^k_i f(\bar\lambda-\hbar\bar\omega_k)\,.$$

It satisfies the RLL relation

$$\begin{aligned}(257)\quad & R^{(N,1)}(z_1-z_2)L^{(N,1)}_1(z_1-w,z_1-w')L^{(N,1)}_2(z_2-w,z_2-w')\\ &= L^{(N,1)}_2(z_2-w,z_2-w')L^{(N,1)}_1(z_1-w,z_1-w')R^{(N,1)}(z_1-z_2)\,,\end{aligned}$$

which is the relation shown in Figure 8(b).

A good way to think about the L-operators is that they are R-matrices associated with the crossings of Wilson lines and "thick" line operators, where the latter are composed of pairs of antiparallel 't Hooft lines and carry infinite-dimensional representations. For example, $L^{(N,1)}$ is an R-matrix whose vertical line carries an infinite-dimensional representation on the space of Weyl-invariant meromorphic functions on $\mathfrak{t}^*_{\mathbb{C}}$.

Being constructed from the same intertwining operators, the two L-operators $L^{(N,0)}$ and $L^{(N,1)}$ lead to the same transfer matrix:

$$(258)\qquad \mathrm{Tr}_{\mathbb{C}^N}\left(L^{(N,0)}_k\cdots L^{(N,0)}_1\right) = \mathrm{Tr}_{\mathbb{C}^N}\left(L^{(N,1)}_k\cdots L^{(N,1)}_1\right).$$

This is a difference operator acting on the space of Weyl-invariant meromorphic functions on $(\mathfrak{t}^*_{\mathbb{C}})^{\otimes k}$. By considering Wilson lines in various representations, we get a number of such difference operators which commute

with each other. For $k = 1$, these difference operators are [65] the conserved charges of the elliptic Ruijsenaars–Schneider model of type A_{N-1} [66].

The RLL relation (257) is, roughly speaking, the defining relation for the elliptic quantum algebra $\mathcal{A}_{q,p}(\widehat{\mathfrak{sl}}_N)$ [67–69] at level zero, with $(q,p) = (e^{2\pi i\hbar}, e^{2\pi i\tau})$. (It should be supplemented with the relation that sets the quantum determinant of the L-operator to 1.) The algebra $\mathcal{A}_{q,p}(\widehat{\mathfrak{sl}}_N)$ is generated by the matrix elements of the L-operator, and is the elliptic counterpart of the Yangian double $\mathcal{D}Y_\hbar(\mathfrak{sl}_N)$ and the quantum affine algebra $U_q(\widehat{\mathfrak{sl}}_N)$. The coalgebra structure making $\mathcal{A}_{q,p}(\widehat{\mathfrak{sl}}_N)$ a quantum group was given in [70].

If it is further required that the dependence of the L-operator on the spectral parameter takes a certain special form, the RLL relation encodes the defining relations for the $\mathbb{Z}_N$ Sklyanin algebra [71, 72]. This is a two-parameter deformation of the universal enveloping algebra $U(\mathfrak{sl}_N)$ of $\mathfrak{sl}_N$, and reduces to the quantum group $U_q(\mathfrak{sl}_N)$ in the limit $\tau \to i\infty$ [73–75]. Essentially, our L-operator $L^{(N,1)}$ gives an infinite-dimensional representation of the $\mathbb{Z}_N$ Sklyanin algebra in terms of difference operators [76–78]. For $N = 2$, this representation corresponds to a Verma module of $\mathfrak{sl}_2$ whose highest weight is determined by the difference $w - w'$ of the spectral parameters of the two 't Hooft lines [71].

In a similar way, the other L-operator $L^{(N,0)}$ provides [16] an infinite-dimensional representation of Felder's elliptic quantum group $E_{q,p}(\mathfrak{sl}_N)$ [33, 34]. Alternative formulations of (a central extension of) $E_{q,p}(\mathfrak{sl}_N)$ are discussed in [70, 79–81].

5.7. Framing anomaly

Up until now we have discussed four-dimensional Chern–Simons theory on $\Sigma \times C$ assuming it is perfectly topological on Σ, as suggested by the form of the action which makes no reference to a metric on Σ. As a matter of fact, this assumption is a little too naive. When it comes to actually performing the path integral, one needs to introduce a metric on Σ for gauge fixing and regularization. The introduction of metric can potentially spoil the topological invariance. This is indeed what happens, but in a somewhat subtle manner.

A manifestation of this quantum anomaly is the fact that the equation that seemingly represents the equivalence between two diagrams, shown in Figure 9(a), does not quite hold:

$$R^{(N,1)}(z_1 - z_2)^{ml}_{kj} R^{(N,1)}(z_2 - z_1)^{nk}_{li} \neq \delta^m_i \delta^n_j \,. \tag{259}$$

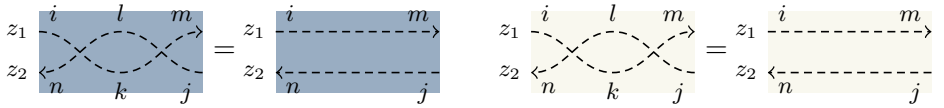


Figure 9: Crossing–unitarity relations in (a) the $(N,1)$ background and (b) an $(N,0)$ background.

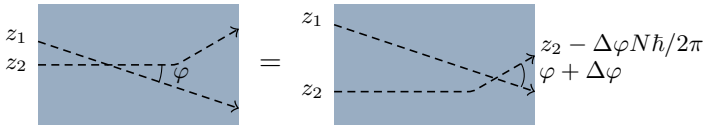


Figure 10: Translation of Wilson lines leads to the same operator.

Instead, $R^{(N,1)}$ satisfies the *crossing–unitarity relation* [61]

$$R^{(N,1)}\left(z-\frac{1}{2}N\hbar\right)^{ml}_{kj} R^{(N,1)}\left(-z-\frac{1}{2}N\hbar\right)^{nk}_{li} = \delta_i^m \delta_j^n \,, \tag{260}$$

provided that we have

$$f^{(N,0)}(z) f^{(N,0)}(-z-N\hbar) = \frac{\theta_1(z+\hbar)\theta_1(z+(N-1)\hbar)}{\theta_1(z)\theta_1(z+N\hbar)} \,. \tag{261}$$

Somehow the arguments of the R-matrices used in this relation have to be shifted by $-N\hbar/2$ compared to the ordinary unitarity relation.

This shift is due to *framing anomaly*. As an analysis carried out in [20] revealed, an anomaly breaks the gauge invariance of a Wilson line when the line curves in the $(N,1)$ background. For this anomaly to be canceled, the spectral parameter must be shifted by $-\Delta\varphi N\hbar/2\pi$, where $\Delta\varphi$ is the angle by which the Wilson line bends. Note that in order to talk about the angle of a curve, one must endow Σ with a framing, that is, a choice of a trivialization of the tangent bundle. The only closed surface that admits a framing is T^2, hence our choice $\Sigma = T^2$.

In turn, the framing anomaly implies, under the assumption that the topological invariance on Σ is otherwise unbroken, that the R-matrix should really depend on the angle at which two Wilson lines cross. This is because as these lines curve, the R-matrix should change by shifting the argument so as to compensate for the shift in the spectral parameters. In Figure 10,

two diagrams are shown in which a straight Wilson line intersects another Wilson line which initially goes straight but at one point bends by angle $\Delta\varphi$. If the straight Wilson lines in the two diagrams are parallel, these diagrams should represent the same operator. From this equality we deduce that the R-matrix $R_\varphi^{(N,1)}$ for Wilson lines crossing at angle φ satisfies the relation

$$R_\varphi^{(N,1)}(z) = R_0^{(N,1)}\Big(z - \frac{\varphi}{2\pi}N\hbar\Big). \tag{262}$$

The unitarity relation (187), as we formulated it, does not involve any shift in the spectral parameter. This is possible only if the equation refers to the situation where the two lines are almost parallel. Therefore, the R-matrix $R^{(N,1)}$ that appears in this equation corresponds to the case $\varphi = 0$:

$$R^{(N,1)} = R_0^{(N,1)}. \tag{263}$$

The crossing–unitarity relation (260), on the other hand, corresponds to the case when two lines are almost antiparallel, which explains the shift by $-N\hbar/2$.

It turns out that the framing anomaly in an $(N,0)$ background is more complicated. To see why, consider the crossing–unitarity relation shown in Figure 9(b). For $i \neq n$, the left-hand side is nonvanishing only when $i = k = m$ and $j = l = n$. If the sole effect of the framing anomaly were to shift the spectral parameter just as in the $(N,1)$ background, then the left-hand side in this case would be

$$\begin{aligned} &R^{(N,0)}\Big(z_1 - z_2 - \frac{1}{2}N\hbar, \bar{\lambda} + \hbar\bar{\omega}_j\Big)_{ij}^{ij} \\ &\quad \times R^{(N,0)}\Big(z_2 - z_1 - \frac{1}{2}N\hbar, \bar{\lambda} + \hbar\bar{\omega}_j\Big)_{ji}^{ji}. \end{aligned} \tag{264}$$

This equals

$$\frac{\theta_1(\bar{\lambda}_{ij})\theta_1(\bar{\lambda}_{ij} - 2\hbar)}{\theta_1(\bar{\lambda}_{ij} - \hbar)^2} \tag{265}$$

and not 1 as required by the relation.

Apparently, the matrix elements of the R-matrix $R_\pi^{(N,0)}$ for $\varphi = \pi$ differs from those of $R^{(N,0)} = R_0^{(N,0)}$ not only by the shift in the spectral parameter, but also by some factors which are ratios of theta functions containing $\bar{\lambda}$. Let us determine these factors.

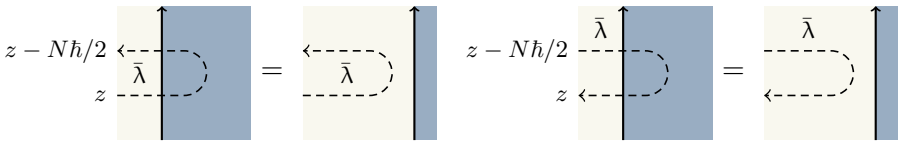


Figure 11: Unitarity relations involving a Wilson line and an 't Hooft line.

First, consider the equality between two diagrams shown in Figure 11(a). On the left-hand side, a Wilson line enters the $(N,1)$ background and makes a left turn. The spectral parameter gets shifted by $-N\hbar/2$, and the line comes out to an $(N,0)$ background. The right-hand side would be the identity operator if it were placed in the $(N,1)$ background. In the $(N,0)$ background, however, the framing anomaly replaces it with a diagonal matrix $\mathrm{diag}(\chi_1,\dots,\chi_N)$, which is a function of $\bar\lambda$ but not of z because of translation invariance. So we get the equality

$$S'\left(z-\frac{1}{2}N\hbar,\bar\lambda\right)^k_j S(z,\bar\lambda)^j_i = \chi_k(\bar\lambda)\delta^k_i\,. \tag{266}$$

Comparing this equation with the expression (246) for S', we see

$$\delta=\frac{1}{2}N\hbar\,,\qquad \chi(z,\bar\lambda)^k_i=\chi_k(\bar\lambda)\delta^k_i\,. \tag{267}$$

It should be emphasized here that we have defined the intertwining operators S, S' using Wilson and 't Hooft lines crossing at the right angle.

Next, suppose that the Wilson line instead makes a right turn, as in Figure 11(b). Then, the right-hand side is replaced with $\chi(\bar\lambda)^{-1}$ because one can straighten out a line that makes successive left and right turns, without altering the initial and the final directions. Thus we get another relation between S and S':

$$S'(z,\bar\lambda-\hbar\bar\omega_i+\hbar\bar\omega_k)^k_j S\left(z-\frac{1}{2}N\hbar,\bar\lambda\right)^j_i = \chi_k(\bar\lambda)^{-1}\delta^k_i\,. \tag{268}$$

These two relations imply

$$S^{-1}(z+N\hbar,\bar\lambda+\hbar\bar\omega_k)^k_j S(z,\bar\lambda+\hbar\bar\omega_i)^j_i = \chi_k(\bar\lambda+\hbar\bar\omega_k)^{-2}\delta^k_i\,. \tag{269}$$

The left-hand side of this equation contains

$$\Phi^{-1}(z+N\hbar+d,\bar{\lambda}+\hbar\bar{\omega}_k)^k_j\Phi(z+d,\bar{\lambda}+\hbar\bar{\omega}_i)^j_i\,. \tag{270}$$

According to the formula [65]

$$\begin{aligned}\frac{\theta_1(z)^{1/N}}{\theta_1(z+N\hbar)^{1/N}}&\Phi^{-1}(z+N\hbar,\bar{\mu})^k_j\Phi(z,\bar{\lambda})^j_i\\ &=\frac{\theta_1(z+(N-1)\hbar+\bar{\lambda}_i-\bar{\mu}_k-(N-1)/2)}{\theta_1(z+N\hbar-(N-1)/2)}\prod_{l(\neq k)}\frac{\theta_1(\bar{\lambda}_i-\bar{\mu}_l-\hbar)}{\theta_1(\bar{\mu}_{kl})}\,,\end{aligned} \tag{271}$$

this factor vanishes for $i\neq k$. Setting $i=k$, we find

$$\begin{aligned}\frac{\bar{g}_k(z+N\hbar,\bar{\lambda}+\hbar\bar{\omega}_k)}{\bar{g}_k(z,\bar{\lambda}+\hbar\bar{\omega}_k)}&=\frac{\theta_1(z+N\hbar+d)^{1/N}}{\theta_1(z+d)^{1/N}}\\ &\times\frac{\theta_1(z+(N-1)\hbar+d-(N-1)/2)}{\theta_1(z+N\hbar+d-(N-1)/2)}\\ &\times\chi_k(\bar{\lambda}+\hbar\bar{\omega}_k)^2\prod_{l(\neq k)}\frac{\theta_1(\bar{\lambda}_{kl})}{\theta_1(\bar{\lambda}_{kl}+\hbar)}\,.\end{aligned} \tag{272}$$

Since z and $\bar{\lambda}$ appear in separate factors on the right-hand side, $\bar{g}_k$ takes the form[9]

$$\bar{g}_k(z,\bar{\lambda})=h_k(z)\eta_k(\bar{\lambda})\,. \tag{273}$$

Then we have

$$\chi_k(\bar{\lambda})^2=C_k\prod_{l(\neq k)}\frac{\theta_1(\bar{\lambda}_{kl})}{\theta_1(\bar{\lambda}_{kl}-\hbar)}\,, \tag{274}$$

$$\begin{aligned}\frac{h_k(z+N\hbar)}{h_k(z)}&=C_k\frac{\theta_1(z+N\hbar+d)^{1/N}}{\theta_1(z+d)^{1/N}}\\ &\times\frac{\theta_1(z+(N-1)\hbar+d-(N-1)/2)}{\theta_1(z+N\hbar+d-(N-1)/2)}\end{aligned} \tag{275}$$

for some constants C_k.

[9]If we write $\bar{g}_k(z,\bar{\lambda})=h_k(z)\eta_k(z,\bar{\lambda})$, with h_k as given below, then η_k is a doubly periodic meromorphic function of z satisfying $\eta_k(z,\bar{\lambda})=\eta_k(z+N\hbar,\bar{\lambda})$. Assuming that any pair from 1, τ and $N\hbar$ are linearly independent in $\mathbb{C}$, this implies that η_k is independent of z as there are no triply periodic meromorphic functions other than constants.

The requirement that $\bar{g}$ acts trivially on $R^{(N,0)}$ translates to the constraints

$$\frac{h_i(z_1)h_j(z_2)}{h_j(z_1)h_i(z_2)} = \frac{\eta_i(\lambda)\eta_j(\lambda-\hbar\omega_i)}{\eta_i(\lambda-\hbar\omega_j)\eta_j(\lambda)}\,. \tag{276}$$

This equation tells that the left-hand side cannot depend on z_1 or z_2, so we have

$$h_k(z) = c_k h(z) \tag{277}$$

for some function h and constants c_k. Absorbing c_k into η_k, we can set

$$h_k = h\,, \qquad C_k = C \tag{278}$$

for some constant C.

We will see in Section 5.8 that $C = 1$. Then we can write

$$\chi_k(\bar{\lambda}) = \prod_{i<j} \frac{\theta_1(\bar{\lambda}_{ij})^{1/2}}{\theta_1((\bar{\lambda}-\hbar\bar{\omega}_k)_{ij})^{1/2}}\,. \tag{279}$$

This shows that in the definition (256) of the difference operator $L^{(N,1)}$, what the factor χ contained in S' does is just to apply conjugation with the operator that acts on a function $f(\bar{\lambda})$ by multiplication by $\prod_{i<j}\theta_1(\bar{\lambda}_{ij})^{1/2}$. Therefore, it does not affect the algebra generated by the L-operator.

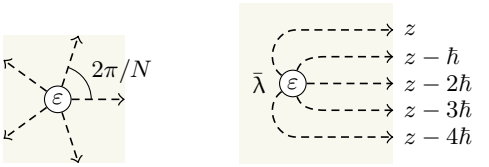
Having determined χ, we finally consider the same relation as in Figure 10 but placed in an $(N,0)$ background. Taking $\varphi = 0$ and $\Delta\varphi = \pi$, we conclude

$$R_\pi^{(N,0)}(z,\bar{\lambda})_{ij}^{kl} = \frac{\chi_l(\bar{\lambda})}{\chi_j(\bar{\lambda}-\hbar\bar{\omega}_i)} R^{(N,0)}\left(z-\frac{1}{2}N\hbar,\bar{\lambda}\right)_{ij}^{kl}. \tag{280}$$

The prefactor on the right-hand side cancels the extra factor (265) in the crossing–unitarity relation, as it should. The unitarity relation for $R_\pi^{(N,0)}$ also readily follows from this relation.

5.8. Junctions of Wilson lines

Although the Yang–Baxter equations and various other relations put strong constraints on the forms of the R-matrices and the intertwining operators, we have not been able to fix some ambiguities. While the determination of the



(a) (b)

Figure 12: (a) A junction of Wilson lines for $N = 5$. (b) The spectral parameters are shifted due to the framing anomaly when the Wilson lines are bent.

matrix ξ is not so crucial as it drops out from gauge invariant expressions, the function $f^{(N,0)} = f^{(N,\pm 1)}$ does affect physical quantities. We can determine this function by considering junctions of Wilson lines [21].

In gauge theory, one can join Wilson lines by contracting the ends of the lines with an invariant tensor of the gauge group. In the case of $G = \mathrm{SU}(N)$, we use a completely antisymmetric tensor ε to construct a junction of N Wilson lines in the vector representation:

$$\varepsilon^{i_1 \dots i_N} (W_1)^{j_1}_{i_1} \cdots (W_N)^{j_N}_{i_N} . \tag{281}$$

An example for $N = 5$ is shown in Figure 12(a).

While in the path integral the junction is described by the constant tensor ε, it can receive quantum corrections in the effective description we are using. This is natural because states on the Wilson lines participating in a junction live in holomorphic vector bundles that are inequivalent due to the jumps of the spectral parameter, and the notion of determinant has to be modified.

Now, take a junction and bend the Wilson lines so that they all extend horizontally to the right, as in Figure 12(b). At the junction the lines have the same spectral parameter, but as they curve their spectral parameters get shifted because of the framing anomaly. It was found in [20] that quantum mechanically a configuration of Wilson lines suffers from an anomaly unless the lines make equal angles at the junctions. Therefore, in the region where the lines are horizontal, the spectral parameters of adjacent lines must differ by $\hbar$. Let these parameters be $z, z - \hbar, \dots, z - (N-1)\hbar$ from top to bottom.

As we have seen already, in addition to the shifts in the spectral parameters, bending of Wilson lines in an $(N, 0)$ background also induces some

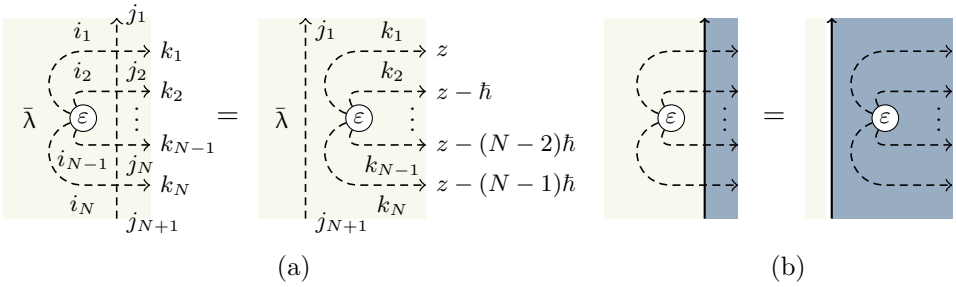


Figure 13: (a) Quantum determinant relation for $R^{(N,0)}$ (b) Quantum determinant relation with a surface operator.

factors of theta functions containing the dynamical parameter. We have determined these factors only in the case when the lines make 180-degree turns, which can be useful only for $N = 2$.

Rather than trying to determine the quantum corrections to the junction and the framing anomaly for general angles separately, let us encapsulate both of these effects into a single tensor $\varepsilon_\hbar(\bar{\lambda})$. This is the operator representing the diagram in Figure 12(b). It is still totally antisymmetric since the contributions to the path integral from terms in the junction (281) vanish if $i_m = i_n$ for some (m, n).

To this collection of Wilson lines let us introduce an additional Wilson line, almost parallel to the horizontal lines. The familiar field theory argument then suggests that the relation shown in Figure 13(a) should hold. (For the ease of visualization we have drawn the additional Wilson line vertically.) The left-hand side of this relation, evaluated for $(k_1, k_2, \ldots, k_N) = (1, 2, \ldots, N)$, is the *quantum determinant* of $R^{(N,0)}$.

To determine $\varepsilon_\hbar(\bar{\lambda})$, we look at a similar relation, in which the vertical Wilson line is replaced with an 't Hooft line; see Figure 13(b). The right-hand side of this relation contains a junction in the $(N, 1)$ background. Since the $(N, 1)$ background has no moduli, the antisymmetric tensor can only receive quantum corrections that rescale it by an overall factor, which can be absorbed by rescaling of the antisymmetric tensor used to define the junction in the path integral. Thus, we have

$$\varepsilon_\hbar(\bar{\lambda})^{i_1 \ldots i_N} \prod_{n=1}^{N} S\Big(z - (n-1)\hbar, \bar{\lambda} - \hbar \sum_{m=1}^{n-1} \bar{\omega}_{i_m}\Big)_{i_n}^{k_n} = \varepsilon^{k_1 \ldots k_N} , \tag{282}$$

or

$$\varepsilon_\hbar(\bar\lambda)^{i_1\dots i_N}\prod_{n=1}^{N} h\big(z-(n-1)\hbar\big)\frac{\theta^{(k_n)}(z+N\bar\lambda_{i_n}+d)}{\theta_1(z-(n-1)\hbar+d)^{1/N}} \times(\Theta^{-1}\xi\eta)\Big(\bar\lambda-\hbar\sum_{m=1}^{n-1}\bar\omega_{i_m}\Big)_{i_n}=\varepsilon^{k_1\dots k_N}\,. \tag{283}$$

A key to crack this equation is the following determinant formula [82]:

$$\det\big(\theta^{(j)}(z+N\bar\lambda_i)\big)_{i,j=1,\dots,N}=C_{N,\tau}\theta_1\big(z-(N-1)/2\big)\prod_{i<j}\theta_1(\bar\lambda_{ij})\,. \tag{284}$$

Here $C_{N,\tau}$ is a constant that depends only on N and τ. From this formula it follows

$$\begin{aligned}&\frac{1}{\theta_1(z+d-(N-1)/2)}\prod_{n=1}^{N}\frac{\theta_1(z+d-(n-1)\hbar)^{1/N}}{h(z-(n-1)\hbar)}\\&=\frac{1}{N!}\varepsilon_{i_1\dots i_N}C_{N,\tau}\prod_{i<j}\theta_1(\bar\lambda_{ij})\varepsilon_\hbar(\bar\lambda)^{i_1\dots i_N}\prod_{n=1}^{N}(\Theta^{-1}\xi\eta)\Big(\bar\lambda-\hbar\sum_{m=1}^{n-1}\bar\omega_{i_m}\Big)_{i_n}\,,\end{aligned} \tag{285}$$

and the two sides are equal to some constant D which can depend on $\hbar$. Thus we get

$$\begin{aligned}&\varepsilon_\hbar^{i_1\dots i_N}(\bar\lambda)\prod_{n=1}^{N}(\Theta^{-1}\xi\eta)_{i_n}\Big(\bar\lambda-\hbar\sum_{m=1}^{n-1}\bar\omega_{i_m}\Big)\\&\qquad=C_{N,\tau}^{-1}D\varepsilon^{i_1\dots i_N}\prod_{i<j}\theta_1(\bar\lambda_{ij})^{-1}\end{aligned} \tag{286}$$

and

$$\frac{1}{\theta_1(z+d-(N-1)/2)}\prod_{n=1}^{N}\frac{\theta_1(z-(n-1)\hbar+d)^{1/N}}{h(z-(n-1)\hbar)}=D\,. \tag{287}$$

The last equation is consistent with relation (275) only if

$$C_k=1 \tag{288}$$

for all k.

Let us go back to the quantum determinant relation in Figure 13(a). For the calculation of the quantum determinant of $R^{(N,0)}$, we can perform a gauge transformation to put the R-matrix in a convenient form. If we apply the gauge transformation by $\Theta^{-1}\xi\eta$, the tensor used at the junction becomes precisely the left-hand side of relation (286). Moreover, since η acts on $R^{(N,0)}$ trivially, we have

$$\begin{aligned}\Theta^{-1}\xi\eta\cdot R^{(N,0)}(z,\bar\lambda)^{ij}_{ij} &= f^{(N,0)}(z)\Theta^{-1}\cdot R^{\mathrm{F}}(-z,\bar\lambda)^{ij}_{ij}\\ &= f^{(N,0)}(z)\frac{\theta_1(z)\theta_1(\bar\lambda_{ij}-\hbar)}{\theta_1(z+\hbar)\theta_1(\bar\lambda_{ij})}\end{aligned}\tag{289}$$

and therefore

$$\begin{aligned}\Theta^{-1}\xi\eta\cdot R^{(N,0)}(\hbar,\bar\lambda)^{ij}_{ij} &= f^{(N,0)}(\hbar)R^{\mathrm{F}}(-\hbar,\bar\lambda)^{ij}_{ji}\\ &= \Theta^{-1}\xi\eta\cdot R^{(N,0)}(\hbar,\bar\lambda)^{ij}_{ji}\,.\end{aligned}\tag{290}$$

From this we deduce that for generic values of $\bar\lambda$, the kernel of $\Theta^{-1}\xi\eta\cdot R^{(N,0)}(\hbar,\bar\lambda)$ is $\bigwedge^2\mathbb{C}^N$. In this gauge, the left-hand side of the quantum determinant relation is antisymmetric under an exchange of final states on adjacent horizontal lines, as we can see by making those lines cross and using the Yang–Baxter equation. Hence, it is completely antisymmetric in the final states on all horizontal lines.

Making use of this antisymmetry we can arrange the final states so that $k_1=j_1$. Then, the only contribution to the quantum determinant comes from the case when $j_n=j_1=k_1$ and $i_n=k_n$ for all n, and the quantum determinant relation reduces to the equation

$$\begin{aligned}&C_{N,\tau}^{-1}D\varepsilon^{k_1\dots k_N}\prod_{i<j}\theta_1(\bar\lambda_{ij})^{-1}\\ &\quad\times\prod_{n=1}^{N}\Theta^{-1}\xi\eta\cdot R^{(N,0)}\Bigg(z-(n-1)\hbar,\bar\lambda-\hbar\sum_{m=1}^{n-1}\bar\omega_{k_m}\Bigg)^{k_nj_1}_{k_nj_1}\\ &\qquad= C_{N,\tau}^{-1}D\varepsilon^{k_1\dots k_N}\prod_{i<j}\theta_1\big((\bar\lambda-\hbar\bar\omega_{j_1})_{ij}\big)^{-1}.\end{aligned}\tag{291}$$

All constants and functions of $\bar\lambda$ in the equation cancel out, leaving

$$\frac{\theta_1(z-(N-1)\hbar)}{\theta_1(z)}\prod_{n=1}^{N}f^{(N,0)}\big(z-(n-1)\hbar\big)=1\,.\tag{292}$$

This is consistent with the quasi-periodicity property (212), as well as with the unitarity condition (213) and the crossing–unitarity condition (261). The same equation is obtained if one sets the quantum determinant of $R^{(N,1)}$ to 1.

6. String theory realization and dualities

In the final section we discuss a realization of four-dimensional Chern–Simons theory and the associated integrable lattice models in string theory. The embedding into string theory allows us to invoke its powerful dualities. Using these dualities, we relate the field theory setup considered in the previous sections to other setups which have been extensively studied in relation to quantum integrable systems. The string theory realization thus provides a unified perspective on a number of phenomena in which the same integrable systems arise from apparently different theories.

6.1. Brane construction of the Ω-deformed topological–holomorphic theory

Consider a stack of N D5-branes in Type IIB superstring theory. If the spacetime is flat Minkowski space $\mathbb{R}^{9,1}$, the low energy dynamics of the branes is described by six-dimensional $\mathcal{N} = (1,1)$ super Yang–Mills theory with gauge group $\mathrm{U}(N)$. Discarding the decoupled degrees of freedom associated with the center-of-mass motion of the D5-branes, we obtain the theory with gauge group $\mathrm{SU}(N)$.

If, instead, the spacetime is $T^*M \times C$ and the D5-branes wrap the zero section of T^*M and C, then the effective worldvolume theory is topologically twisted along M [83]. (Here, as before, M is a four-manifold and C is either $\mathbb{C}$, $\mathbb{C}^\times$ or an elliptic curve E.) In fact, it is the twisted $\mathcal{N} = (1,1)$ super Yang–Mills theory whose Q-invariant sector is the topological–holomorphic theory on $M \times C$, constructed in Section 2. The reason is that the four bosonic fields parametrizing the positions of the branes in the fiber directions of T^*M are not scalars as in the untwisted theory. Rather, at each point on C, they are components of a one-form on M. Turning the four scalar fields into a one-form on M is precisely what the topological twisting for the topological–holomorphic theory does.

Our goal is to understand how to introduce an Ω-deformation to this brane construction of the topological–holomorphic theory. More specifically, we take $M = \mathbb{R}^2 \times \Sigma$ and $C = E$, and wish to turn on an Ω-deformation in the worldvolume theory using the rotation symmetry of $\mathbb{R}^2$.

To this end, suppose that we could realize the desired Ω-deformation, and subsequently dimensionally reduced the Ω-deformed theory on E. Then, we would obtain an Ω-deformation of the GL-twisted $\mathcal{N} = 4$ super Yang–Mills theory on $\mathbb{R}^2 \times \Sigma$. This Ω-deformation is, however, different from the one commonly considered in the study of four-dimensional $\mathcal{N} = 2$ supersymmetric gauge theories.

The standard Ω-deformation [22, 23] is compatible with the Donaldson–Witten twist [25]. Upon dimensional reduction on Σ, the Donaldson–Witten twist descends to the A-twist of $\mathcal{N} = (2, 2)$ supersymmetric theories in two dimensions [84]. On the other hand, as we have seen already, the topological–holomorphic theory reduces to a B-twisted theory in two dimensions, not an A-twisted one.

From the GL-twisted $\mathcal{N} = 4$ super Yang–Mills theory we can obtain either of these twists in two dimensions, depending on the choice of the supercharge we use to define a topological theory. In the four-dimensional theory, the two types of twists are related by *S-duality* [50, 84, 85]. This means that the Ω-deformation of the topological–holomorphic theory descends to the S-dual of the standard Ω-deformation of the GL-twisted $\mathcal{N} = 4$ super Yang–Mills theory.

A nice thing about the standard Ω-deformation of an $\mathcal{N} = 2$ supersymmetric gauge theory is that it has a transparent geometric construction. First, we lift the theory to an $\mathcal{N} = (1, 0)$ supersymmetric gauge theory in six dimensions. The lifted theory is defined on the product $M \times E$. Then, we twist this product so that when we go around the one-cycles of E, we do not come back to the point we started from, but arrive at a point that is shifted by the action of an isometry of M. Finally, we perform the dimensional reduction of the lifted theory down to four dimensions. The resulting four-dimensional theory is deformed compared to the original one because of the twisting of the product.

This procedure can be incorporated in our brane construction straightforwardly [86–89]. In our setup, the D5-branes are supported on the product $\mathbb{R}^2 \times \Sigma \times E$ sitting in the ten-dimensional spacetime $T^*\mathbb{R}^2 \times T^*\Sigma \times E$, where $\mathbb{R}^2$ is the zero section of $T^*\mathbb{R}^2$ and, for the purpose of this discussion, we can take Σ to be the zero section of $T^*\Sigma$.

If we apply T-duality on E, the D5-branes turn into D3-branes. In the limit where E shrinks to a point, the low energy dynamics of these D3-branes is described by the GL-twisted $\mathcal{N} = 4$ super Yang–Mills theory on $\mathbb{R}^2 \times \Sigma$. To introduce the Ω-deformation, we modify the geometry before applying the T-duality. Viewing $\mathbb{R}^2 \times \Sigma \times E$ as a flat $\mathbb{R}^2$-bundle over $\Sigma \times E$, we twist it so that the fiber is rotated by some angles as it is transported

along the one-cycles of E. For supersymmetry to be preserved, we must simultaneously rotate the fiber of $T^*\mathbb{R}^2$ in the opposite direction. Now, T-duality on E produces a D3-brane configuration realizing the GL-twisted $\mathcal{N} = 4$ super Yang–Mills theory, subjected to the standard Ω-deformation.

To obtain the brane setup for the Ω-deformed topological–holomorphic theory, all we have to do is to apply S-duality to this D3-brane configuration, which leaves the D3-branes intact but acts nontrivially on the background, and then T-duality on the dual elliptic curve $E^\vee$ to turn the D3-branes back into D5-branes.

Let us describe this construction more precisely, following the chain of dualities step by step. We use radial coordinates (r, ϑ) and (ρ, φ) for the base and fiber of $T^*\mathbb{R}^2$, respectively, and parametrize E with real coordinates (x^4, x^5) defined up to the identification

$$(x^4, x^5) \sim (x^4 + 2\pi R, x^5) \sim (x^4 + 2\pi R\tau_1, x^5 - 2\pi R\tau_2)\,, \tag{293}$$

with $\tau_2 > 0$. With respect to the complex coordinate $z = (x^4 - \mathrm{i}x^5)/2$, the modular parameter of E is $\tau = \tau_1 + \mathrm{i}\tau_2$.

Our starting point is the D5-branes supported on a twisted product of $T^*\mathbb{R}^2$ and E. In terms of the periodic coordinates y^1, y^2 defined by

$$x^4 = R(y^1 + \tau_1 y^2)\,, \qquad x^5 = -R\tau_2 y^2\,, \tag{294}$$

we can construct this space via the identification

$$\begin{aligned}(\vartheta, \varphi, y^1, y^2) &\sim (\vartheta + 2\pi\varepsilon_1, \varphi - 2\pi\varepsilon_1, y^1 + 2\pi, y^2) \\ &\sim (\vartheta + 2\pi\varepsilon_2, \varphi - 2\pi\varepsilon_2, y^1, y^2 + 2\pi)\,,\end{aligned} \tag{295}$$

with some parameters ε_1, $\varepsilon_2 \in \mathbb{R}$. The spacetime metric is given by

$$g = \mathrm{d}r^2 + r^2\mathrm{d}\vartheta^2 + \mathrm{d}\rho^2 + \rho^2\mathrm{d}\varphi^2 + (\mathrm{d}x^4)^2 + (\mathrm{d}x^5)^2 + g_{T^*\Sigma}\,, \tag{296}$$

where $g_{T^*\Sigma}$ is a Ricci flat metric on $T^*\Sigma$. We take the dilaton to be a constant:

$$\Phi = \Phi_0\,. \tag{297}$$

The other background fields, the Kalb–Ramond two-form field B_2 and the RR p-form fields C_p, are all set to zero.

The first step in the chain of dualities is T-duality on E. For this step it is convenient to introduce angle variables

$$\theta = \vartheta - \varepsilon_1 y^1 - \varepsilon_2 y^2\,, \qquad \phi = \varphi + \varepsilon_1 y^1 + \varepsilon_2 y^2\,, \tag{298}$$

which disentangle the identification (295):

$$(\theta, \phi, y^1, y^2) \sim (\theta, \phi, y^1 + 2\pi, y^2) \sim (\theta, \phi, y^1, y^2 + 2\pi)\,. \tag{299}$$

With these coordinates we can use the standard formulas for T-duality [90, 91].

The action of T-duality on g and B_2 can be expressed concisely in terms of the tensor $g + B_2$. We write it in the block matrix form as

$$g + B_2 = \begin{pmatrix} K & N \\ M & L \end{pmatrix}\,, \tag{300}$$

where K represents the block whose indices involve only y^1 and y^2. Under T-duality in the y^1- and y^2-directions, $g + B_2$ is transformed to $\tilde{g} + \widetilde{B}_2$, with the corresponding blocks given by

$$\widetilde{K} = K^{-1}\,, \quad \widetilde{L} = L - MK^{-1}N\,, \quad \widetilde{M} = MK^{-1}\,, \quad \widetilde{N} = -K^{-1}N\,. \tag{301}$$

The dilaton is shifted as

$$\widetilde{\Phi} = \Phi_0 - \frac{1}{2}\ln\det K\,. \tag{302}$$

Since $B_2 = 0$ initially, K and L are symmetric while $M^T = N$. Then, $\widetilde{K}$ and $\widetilde{L}$ are symmetric and $\widetilde{N} = -\widetilde{M}^T$. The T-duality thus turns the metric into a block diagonal form and induces a nonzero B-field:

$$\tilde{g} = \begin{pmatrix} \widetilde{K} & 0 \\ 0 & \widetilde{L} \end{pmatrix}\,, \qquad \widetilde{B}_2 = \begin{pmatrix} 0 & -\widetilde{M}^T \\ \widetilde{M} & 0 \end{pmatrix}\,. \tag{303}$$

An explicit calculation shows

$$\begin{aligned} \tilde{g} &= \mathrm{d}r^2 + r^2\mathrm{d}\theta^2 + \mathrm{d}\rho^2 + \rho^2\mathrm{d}\phi^2 - \frac{|\varepsilon|^2}{\Delta^2}(r^2\mathrm{d}\theta - \rho^2\mathrm{d}\phi)^2 + g_{T^*\Sigma} \\ &\quad + \frac{4}{R^4\tau_2^2\Delta^2}\Big((r^2+\rho^2)\big(\mathrm{Im}(\bar{\varepsilon}\,\mathrm{d}\zeta)\big)^2 + \mathrm{d}\zeta\,\mathrm{d}\bar{\zeta}\Big)\,, \\ \widetilde{B}_2 &= \frac{2}{R^2\tau_2\Delta^2}(r^2\mathrm{d}\theta - \rho^2\mathrm{d}\phi) \wedge \mathrm{Re}(\bar{\varepsilon}\,\mathrm{d}\zeta)\,, \\ \widetilde{\Phi} &= \Phi_0 - \ln(R^2\tau_2\Delta)\,, \end{aligned} \tag{304}$$

where we have defined

$$\varepsilon = \frac{\tau\varepsilon_1 - \varepsilon_2}{R\tau_2}\,, \qquad \zeta = \frac{R}{2}(\tau y^1 - y^2)\,, \qquad \Delta^2 = 1 + |\varepsilon|^2(r^2+\rho^2)\,. \tag{305}$$

This is the NS fluxtrap background studied in [86–89].

Next, we apply S-duality. This step changes the metric and the dilaton to

$$\hat{g} = e^{-\widetilde{\Phi}}\tilde{g}\,, \qquad \widehat{\Phi} = -\widetilde{\Phi}\,, \tag{306}$$

and exchanges the B-field and the RR two-form:

$$\widehat{B}_2 = \widetilde{C}_2\,, \qquad \widehat{C}_2 = -\widetilde{B}_2\,. \tag{307}$$

This background is called the RR fluxtrap [89].

Finally, we apply T-duality in the y^1- and y^2-directions again. The resulting metric and dilaton are

$$\begin{aligned}\check{g} = {} & R^2\tau_2\Delta e^{-\Phi_0}\Big(\mathrm{d}r^2 + r^2\mathrm{d}\theta^2 + \mathrm{d}\rho^2 + \rho^2\mathrm{d}\phi^2 \\ & \qquad - \frac{|\varepsilon|^2}{\Delta^2}(r^2\mathrm{d}\theta - \rho^2\mathrm{d}\phi)^2 + g_{T^*\Sigma}\Big) \\ & + \frac{4e^{\Phi_0}}{R^2\tau_2\Delta}\Big((r^2+\rho^2)\big(\mathrm{Im}(\bar{\varepsilon}\,\mathrm{d}z)\big)^2 + \mathrm{d}z\,\mathrm{d}\bar{z}\Big)\,, \\ \check{\Phi} = {} & \ln(R^2\tau_2\Delta)\,.\end{aligned} \tag{308}$$

On the RR two-form this step acts as a 90-degree rotation on the y^1-y^2 plane, sending $\mathrm{d}\zeta$ to $\mathrm{d}z$:

$$\check{C}_2 = \frac{2}{R^2\tau_2\Delta^2}(r^2\mathrm{d}\theta - \rho^2\mathrm{d}\phi) \wedge \mathrm{Re}(\bar{\varepsilon}\,\mathrm{d}z)\,. \tag{309}$$

Based on the argument we have given above, we claim that this is the background in which a stack of D5-branes realizes the Ω-deformed topological–holomorphic theory.

In principle, we should be able to verify this claim by comparing the Dirac–Born–Infeld (DBI) action for the worldvolume theory of the D5-branes and the action for the Ω-deformed topological–holomorphic theory. In practice, this is not as easy as it may sound because the two actions only need to coincide up to Q-exact terms and a nontrivial field redefinition. Here we content ourselves with confirming that the DBI action reproduces some important terms.

The metric on the D5-brane worldvolume is

$$\check{g}_{\mathrm{D5}} = R^2\tau_2\Delta_0 e^{-\Phi_0}\left(\mathrm{d}r^2 + \frac{r^2}{\Delta_0^2}\mathrm{d}\theta^2 + g_\Sigma\right) + \frac{4e^{\Phi_0}}{R^2\tau_2\Delta_0}\left(r^2\big(\mathrm{Im}(\bar\varepsilon\,\mathrm{d}z)\big)^2 + \mathrm{d}z\,\mathrm{d}\bar z\right), \tag{310}$$

where $\Delta_0 = 1 + |\varepsilon|^2 r^2$. For this metric to reduce at $\varepsilon = 0$ to the one we used for the topological–holomorphic theory, we must take

$$e^{\Phi_0} = R^2\tau_2\,. \tag{311}$$

Then, we have

$$\sqrt{\check{g}_{\mathrm{D5}}}\,\mathrm{d}^6x = \sqrt{g_\Sigma}\,\mathrm{d}^6x\,, \tag{312}$$

where $\mathrm{d}^6x = \mathrm{d}x^0 \wedge \cdots \wedge \mathrm{d}x^5$, with $x^0 + \mathrm{i}x^1 = re^{\mathrm{i}\theta}$ and (x^2, x^3) being coordinates on Σ.

The DBI action, expanded to quadratic order in derivatives, contains the terms

$$-\frac{(2\pi\alpha')^2}{2R^2\tau_2}T_5\int_{\mathbb{R}^2\times\Sigma\times E}\sqrt{g_\Sigma}\,\mathrm{d}^6x \times \mathrm{Tr}\left(\frac{1}{\Delta_0^2}F^{rm}F_{rm} + F^{\theta m}F_{\theta m} + \frac{1}{2\Delta_0^2}F^{mn}F_{mn}\right). \tag{313}$$

Here $(2\pi\alpha')^{-1}$ is the string tension, T_5 is the D5-brane tension, and indices are raised with respect to the metric $\mathrm{d}r^2 + r^2\mathrm{d}\theta^2 + g_\Sigma + (\mathrm{d}x^4)^2 + (\mathrm{d}x^5)^2$. We identify these terms with the kinetic terms $|D_{\hat r}\varphi|^2/(1 + \|V\|^2) + |D_{\hat\theta}\varphi|^2$ for $\varphi = \mathcal{A}_m$ and the potential term $|\partial W/\partial\varphi|^2/(1 + \|V\|^2)$ for $\varphi = A_{\bar z}$ in the bosonic part (111) (with $t = 1$) of the action for the Ω-deformed topological–holomorphic theory. Thus we find

$$\frac{1}{e^2} = \frac{(2\pi\alpha')^2}{2R^2\tau_2}T_5\,, \qquad |\epsilon| = |\varepsilon|\,. \tag{314}$$

The RR two-form induces the Wess–Zumino term

$$-\,\mathrm{i}\frac{(2\pi\alpha')^2}{2}\mu_5\int_{\mathbb{R}^2\times\Sigma\times E}\check{C}_2 \wedge \mathrm{Tr}(F\wedge F)\,, \tag{315}$$

where μ_5 is the D5-brane charge. This term contains

$$\left(\frac{(2\pi\alpha')^2}{2R^2\tau_2}e^2\mu_5\right) \times 2\mathrm{i}\,\mathrm{Im}\int_{\mathbb{R}^2} r\,\mathrm{d}r\wedge\mathrm{d}\theta\wedge\frac{\bar{\varepsilon}r}{\Delta_0^2}\partial_r\left(-\frac{\mathrm{i}}{e^2}\int_{\Sigma\times E}\mathrm{d}z\wedge\mathrm{CS}(A)\right). \tag{316}$$

We see it within the terms $2\mathrm{i}\,\mathrm{Im}(\overline{V}^{\hat{\theta}}\partial_{\hat{r}}W)/(1+\|V\|^2)$. Comparing the coefficients of $\partial_r W$, we identify

$$\epsilon = \varepsilon\,. \tag{317}$$

For the overall factor to be equal to 1, we must have $T_5 = \mu_5$. This is the BPS condition for D5-branes.

6.2. Wilson lines and surface operators

Let us construct integrable lattice models in the above string theory setup. For $\Sigma = T^2$, the ten-dimensional spacetime is

$$T^*\mathbb{R}^2\times T^*\Sigma\times E\cong\mathbb{R}^2\times T^2\times E\times\mathbb{R}^2_{67}\times\mathbb{R}^2_{89}\,, \tag{318}$$

where $\mathbb{R}^2_{67}$ and $\mathbb{R}^2_{89}$ are the fibers of $T^*\mathbb{R}^2$ and $T^*\Sigma$, respectively. The subscripts refer to the coordinates for these spaces which are consistent with the ones used in Section 2. We use coordinates (x, y) for T^2 and a complex coordinate z on E.

Four-dimensional Chern–Simons theory for $G = \mathrm{SU}(N)$ is realized by N D5-branes $\mathrm{D5}_i$, $i = 1, \ldots, N$, supported on

$$\mathbb{R}^2\times T^2\times E\times\{0\}\times\{(\phi_x^i,\phi_y^i)\}\subset\mathbb{R}^2\times T^2\times E\times\mathbb{R}^2_{67}\times\mathbb{R}^2_{89}\,. \tag{319}$$

Without loss of generality we may assume

$$\phi_x^1\le\phi_x^2\le\cdots\le\phi_x^N\,. \tag{320}$$

The coordinates (ϕ_x^i, ϕ_y^i) of $\mathrm{D5}_i$ in $\mathbb{R}^2_{89}$ determine the imaginary part of the background value of the complex gauge field $\mathcal{A}_x\mathrm{d}x + \mathcal{A}_y\mathrm{d}y$. Together with the real part, given by the values of the gauge fields on $\mathrm{D5}_i$ along T^2, they specify the twisted periodic boundary conditions of the lattice models. In the absence of the Ω-deformation, the D5-branes would preserve half of the thirty-two supercharges of Type IIB superstring theory.

The construction of integrable lattice models requires Wilson lines and surface operators bounded by 't Hooft lines. To be concrete, let us consider a lattice similar to the one illustrated in Figure 5. It consists of m horizontal and n vertical Wilson lines in the vector representation of $\mathrm{SU}(N)$, as well as k vertical strips of surface operators.

In general, Wilson lines in the worldvolume theory of a stack of N D-branes are created by fundamental strings ending on the D-branes. The end of a semi-infinite open string behaves as a charged particle with infinite mass. There are N choices for the D-brane on which the string ends, and these correspond to the possible states of the charged particle. Thus, a single open string creates a Wilson line in the vector representation. For Wilson lines in other representations, there are more elaborate constructions which involve multiple strings and additional branes [15, 92–94].

Adopting this construction, we see that the horizontal Wilson lines are realized by fundamental strings $\mathrm{F1}^{\dashrightarrow}_{\alpha}$, $\alpha = 1, \ldots, m$, ending on one of the D5-branes at $(y, z) = (y_\alpha, z_\alpha)$ and extending in the negative x^8-direction. If the αth Wilson line is in the i_αth state, $\mathrm{F1}^{\dashrightarrow}_{\alpha}$ ends on $\mathrm{D5}_{i_\alpha}$. The vertical Wilson lines are created by fundamental strings $\mathrm{F1}^{\uparrow}_{\beta}$, $\beta = 1, \ldots, n$, ending on $\mathrm{D5}_{i_\beta}$ at $(x, z) = (x_\beta, z_\beta)$ and extending in the negative x^9-direction. To be compatible with the Ω-deformation, these strings must sit at the origins of $\mathbb{R}^2$ and $\mathbb{R}^2_{67}$. In the undeformed situation, $\mathrm{F1}^{\dashrightarrow}_{\alpha}$ would break half of the sixteen supercharges preserved by the D5-branes, and $\mathrm{F1}^{\uparrow}_{\beta}$ would further break half of the surviving eight supercharges.

The brane realization for the 't Hooft lines can be identified from the fact that 't Hooft lines in $\mathcal{N} = 4$ super Yang–Mills theory in four dimensions are the S-duals of Wilson lines. As such, in the worldvolume theory of D3-branes these lines are created by D1-branes, which are the S-duals of fundamental strings. Since D3-branes are what the D5-branes become if we compactify $\mathbb{R}^2$ to a torus and apply T-duality along its one-cycles, 't Hooft lines in the worldvolume theory of the D5-branes are created by the T-duals of those D1-branes, namely D3-branes.

Therefore, the 't Hooft lines going upward in Figure 5 are created by semi-infinite D3-branes $\mathrm{D3}_\gamma$, $\gamma = 1, \ldots, k$, coming from $x^8 = -\infty$ and hitting $\mathrm{D5}_{i_\gamma}$ at $(x, z) = (x^{\uparrow}_\gamma, z^{\uparrow}_\gamma)$. The choices i_γ of the D5-branes that these D3-branes hit determine the charges of the 't Hooft lines: for $G = \mathrm{U}(N)$, the γth 't Hooft line has charge $\mathrm{diag}(0, \ldots, 0, 1, 0, \ldots, 0)$, with 1 in the i_γth entry. Throwing away the center-of-mass degrees of freedom of the D5-branes makes the charge traceless, replacing it with the fractional charge (221) or its permutation. This brane realization of monopoles is the S-dual of the one studied in [95].

Spacetime:	$\mathbb{R}^2$	$\times$	T^2	$\times$	E	$\times$	$\mathbb{R}^2_{67}$	$\times$	$\mathbb{R}^2_{89}$
$\mathrm{D5}_i$:	$\mathbb{R}^2$	$\times$	T^2	$\times$	E	$\times$	$\{0\}$	$\times$	$\{(\phi^i_x, \phi^i_y)\}$
$\mathrm{F1}^{\dashrightarrow}_\alpha$:	$\{0\}$	$\times$	$\{y = y_\alpha\}$	$\times$	$\{z_\alpha\}$	$\times$	$\{0\}$	$\times$	$\{(x^8, \phi^{i_\alpha}_y) \mid x^8 \le \phi^{i_\alpha}_x\}$
$\mathrm{F1}^{\dashuparrow}_\beta$:	$\{0\}$	$\times$	$\{x = x_\beta\}$	$\times$	$\{z_\beta\}$	$\times$	$\{0\}$	$\times$	$\{(\phi^{i_\beta}_x, x^9) \mid x^9 \le \phi^{i_\beta}_y\}$
$\mathrm{D3}^{\uparrow}_\gamma$:	$\mathbb{R}^2$	$\times$	$\{x = x^{\uparrow}_\gamma\}$	$\times$	$\{z^{\uparrow}_\gamma\}$	$\times$	$\{0\}$	$\times$	$\{(x^8, \phi^{i_\gamma}_y) \mid x^8 \le \phi^{i_\gamma}_x\}$
$\mathrm{D3}^{\downarrow}_\gamma$:	$\mathbb{R}^2$	$\times$	$\{x = x^{\downarrow}_\gamma\}$	$\times$	$\{z^{\downarrow}_\gamma\}$	$\times$	$\{0\}$	$\times$	$\{(x^8, \phi^{i_\gamma}_y) \mid x^8 \ge \phi^{i_\gamma}_x\}$

Table 1: A brane configuration for an integrable lattice model. The branes are placed in a background with nonzero RR two-form. $\mathrm{D3}_\gamma$ forms a bound state with the D5-branes in the region $\{x^{\uparrow}_\gamma \le x \le x^{\downarrow}_\gamma\}$ on T^2.

If a D3-brane creating an 't Hooft line curves in T^2, it also has to curve in $\mathbb{R}^2_{89}$ by the same angle to preserve supersymmetry. In particular, an 't Hooft line going downward is created by a D3-brane hitting one of the D5-branes from the positive x^8-direction. This observation suggests the following construction for the strips of surface operators.

When $\mathrm{D3}_\gamma$ comes from $x^8 = -\infty$ and hits $\mathrm{D5}_{i_\gamma}$, it makes a right turn to move along T^2, and the two branes form a bound state. This D3–D5 bound state creates the surface operator whose left boundary is the γth upward 't Hooft line. While maintaining the bound state, $\mathrm{D3}_\gamma$ can gradually shift its position in E. When $\mathrm{D3}_\gamma$ reaches $(x, z) = (x^{\downarrow}_\gamma, z^{\downarrow}_\gamma)$, it makes a left turn and leaves $\mathrm{D5}_{i_\gamma}$. Then $\mathrm{D3}_\gamma$ goes off to $x^8 = +\infty$, yielding the downward 't Hooft line on the right boundary of the surface operator,

The D3-branes break half of the four supercharges preserved by the other branes. We refer to the semi-infinite parts of $\mathrm{D3}_\gamma$ responsible for the upward and downward 't Hooft lines as $\mathrm{D3}^{\uparrow}_\gamma$ and $\mathrm{D3}^{\downarrow}_\gamma$, respectively.

The brane configuration realizing the integrable lattice model is summarized in Table 1.

6.3. Brane tilings and class-$\mathcal{S}_k$ theories

Tracing back the chain of dualities, we obtain another realization of the same integrable lattice model. By the application of T-duality on E, S-duality, and T-duality on E again, $\mathrm{F1}^{\dashrightarrow}_\alpha$ and $\mathrm{F1}^{\dashuparrow}_\beta$ are converted to D3-branes $\mathrm{D3}^{\dashrightarrow}_\alpha$ and $\mathrm{D3}^{\dashuparrow}_\beta$, while $\mathrm{D3}^{\uparrow}_\gamma$ and $\mathrm{D3}^{\downarrow}_\gamma$ are converted to NS5-branes $\mathrm{NS5}^{\uparrow}_\gamma$ and $\mathrm{NS5}^{\downarrow}_\gamma$. The dual brane configuration is summarized in Table 2.

Each NS5-brane forms a bound state with the D5-branes over a colored region in Figure 5. This bound state of N D5-branes and one NS5-brane is

Spacetime:	$\mathbb{R}^2$	$\times$	T^2	$\times$	E	$\times$	$\mathbb{R}^2_{67}$	$\times$	$\mathbb{R}^2_{89}$
D5$_i$:	$\mathbb{R}^2$	$\times$	T^2	$\times$	E	$\times$	$\{0\}$	$\times$	$\{(\phi_x^i, \phi_y^i)\}$
D3$^{\dashrightarrow}_{\alpha}$:	$\{0\}$	$\times$	$\{y = y_\alpha\}$	$\times$	E	$\times$	$\{0\}$	$\times$	$\{(x^8, \phi_y^{i_\alpha}) \mid x^8 \leq \phi_x^{i_\alpha}\}$
D3$^{\uparrow}_{\beta}$:	$\{0\}$	$\times$	$\{x = x_\beta\}$	$\times$	E	$\times$	$\{0\}$	$\times$	$\{(\phi_x^{i_\beta}, x^9) \mid x^9 \leq \phi_y^{i_\beta}\}$
NS5$^{\uparrow}_{\gamma}$:	$\mathbb{R}^2$	$\times$	$\{x = x_\gamma^{\uparrow}\}$	$\times$	E	$\times$	$\{0\}$	$\times$	$\{(x^8, \phi_y^{i_\gamma}) \mid x^8 \leq \phi_x^{i_\gamma}\}$
NS5$^{\downarrow}_{\gamma}$:	$\mathbb{R}^2$	$\times$	$\{x = x_\gamma^{\downarrow}\}$	$\times$	E	$\times$	$\{0\}$	$\times$	$\{(x^8, \phi_y^{i_\gamma}) \mid x^8 \geq \phi_x^{i_\gamma}\}$

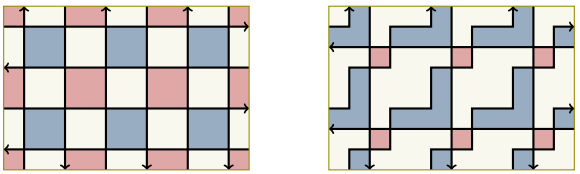
Table 2: A brane tiling configuration for an integrable lattice model. The product between $\mathbb{R}^2 \times \mathbb{R}^2_{67}$ and E is twisted by rotations of $\mathbb{R}^2$ and $\mathbb{R}^2_{67}$ in opposite directions.

called an $(N, 1)$ 5-brane. In this terminology, a stack of N D5-branes may be referred to as an $(N, 0)$ 5-brane. Our choice of the names for various backgrounds was motivated by this 5-brane interpretation.

The 5-brane system realizes a five-dimensional $\mathcal{N} = 1$ supersymmetric gauge theory on $\mathbb{R}^2 \times S^1 \times E$ [96, 97], with the product between $\mathbb{R}^2$ and E being a twisted one. The D3-branes create three-dimensional defects supported on $\{0\} \times S^1 \times E$ in this theory. Thus, the partition function of the lattice model translates to the correlation function of these defects in this theory, also known as the *supersymmetric index* of the theory on $\mathbb{R}^2 \times S^1 \times E$ in the presence of the defects.

If we wish, we can introduce additional 't Hooft lines in the horizontal direction and make a tricolor checkerboard pattern on T^2, as in Figure 14(a); or for that matter, we can consider entirely different patterns of $(N, 0)$ and $(N, \pm 1)$ background regions, such as the one shown in Figure 14(b). Such configurations of 5-branes, called *brane tilings*, realize four-dimensional $\mathcal{N} = 1$ supersymmetric gauge theories on $\mathbb{R}^2 \times E$ [38, 39]. These theories have multiple SU(N) gauge (and flavor) groups, one for each $(N, 0)$ background region, and chiral multiplets in the bifundamental representations under two SU(N) gauge groups associated to $(N, 0)$ background regions sharing a vertex.

The theories realized by the brane tilings in Figure 14 are also examples of $\mathcal{N} = 1$ supersymmetric theories of *class* $\mathcal{S}_k$ [14, 40, 41]. Theories of class $\mathcal{S}_k$ describe the dynamics of M5-branes probing a transverse $\mathbb{C}^2/\mathbb{Z}_k$ singularity, compactified on punctured Riemann surfaces which in our case are tori. This brane setup is obtained by T-duality in the horizontal direction of T^2, followed by a lift to M-theory. The D3-branes are lifted to M2- and M5-branes, producing surface operators in the class-$\mathcal{S}_k$ theories.



(a) (b)

Figure 14: (a) A tricolor checkerboard brane tiling. (b) Another brane tiling.

Spacetime:	$\mathbb{R}^2$	$\times$	T^2	$\times$	E	$\times$	$\mathbb{R}^2_{67}$	$\times$	$\mathbb{R}^2_{89}$
$\mathrm{NS5}_i$:	$\mathbb{R}^2$	$\times$	T^2	$\times$	E	$\times$	$\{0\}$	$\times$	$\{(\phi_x^i, \phi_y^i)\}$
$\mathrm{D0}_\alpha^{\dashrightarrow}$:	$\{0\}$	$\times$	$\{(x_\alpha, y_\alpha)\}$	$\times$	$\{z_\alpha\}$	$\times$	$\{0\}$	$\times$	$\{(x^8, \phi_y^{i_\alpha}) \mid x^8 \le \phi_x^{i_\alpha}\}$
$\mathrm{D2}_\beta^{\dashuparrow}$:	$\{0\}$	$\times$	T^2	$\times$	$\{z_\beta\}$	$\times$	$\{0\}$	$\times$	$\{(\phi_x^{i_\beta}, x^9) \mid x^9 \le \phi_y^{i_\beta}\}$
$\mathrm{D4}_\gamma^{\uparrow}$:	$\mathbb{R}^2$	$\times$	T^2	$\times$	$\{z_\gamma^{\uparrow}\}$	$\times$	$\{0\}$	$\times$	$\{(x^8, \phi_y^{i_\gamma}) \mid x^8 \le \phi_x^{i_\gamma}\}$
$\mathrm{D4}_\gamma^{\downarrow}$:	$\mathbb{R}^2$	$\times$	T^2	$\times$	$\{z_\gamma^{\downarrow}\}$	$\times$	$\{0\}$	$\times$	$\{(x^8, \phi_y^{i_\gamma}) \mid x^8 \ge \phi_x^{i_\gamma}\}$

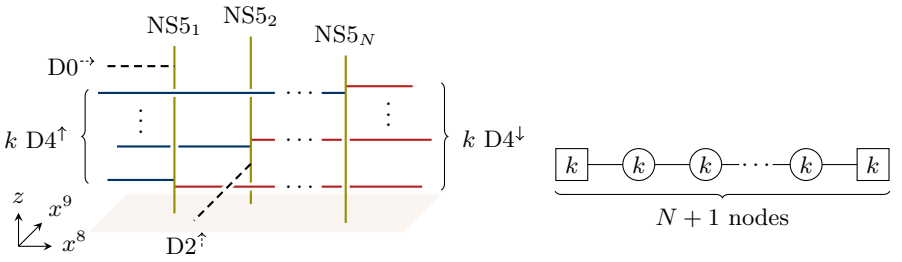
Table 3: A brane configuration of Hanany–Witten type for an integrable lattice model.

It is known that surface operators act on the supersymmetric indices of brane tiling and class-$\mathcal{S}_k$ theories as difference operators [11–17]. Our construction shows that these difference operators are nothing but transfer matrices of L-operators. This result, obtained in [15, 16] from the perspective of brane tilings, was a primary motivation for us to study surface operators in four-dimensional Chern–Simons theory.

6.4. Linear quiver theories

Another interesting chain of dualities we can apply to the brane configuration in Table 1 is S-duality and T-duality in the horizontal direction of T^2. This turns $\mathrm{D5}_i$ into NS5-branes $\mathrm{NS5}_i$, $\mathrm{F1}_\alpha^{\dashrightarrow}$ into D0-branes $\mathrm{D0}_\alpha^{\dashrightarrow}$, $\mathrm{F1}_\beta^{\dashuparrow}$ into D2-branes $\mathrm{D2}_\beta^{\dashuparrow}$, and $\mathrm{D3}_\gamma^{\uparrow}$ and $\mathrm{D3}_\gamma^{\downarrow}$ into D4-branes $\mathrm{D4}_\gamma^{\uparrow}$ and $\mathrm{D4}_\gamma^{\downarrow}$, as summarized in Table 3. A schematic picture of this brane setup is shown in Figure 15(a). These branes are placed in a background with a nonzero B-field.

Let us decompactify the holomorphic surface $C = E$ to $\mathbb{C}$. Then, the part of the system consisting of the D4- and NS5-branes is a well-known



(a) (b)

Figure 15: (a) A D4–NS5 brane configuration of Hanany–Witten type with additional D0- and D2-branes. (b) The linear quiver for the theory realized by the D4–NS5 brane configuration.

brane configuration studied in Witten's classic paper [42], which builds on his earlier work [98] with Hanany.

The D4–NS5 brane configuration realizes a four-dimensional $\mathcal{N} = 2$ supersymmetric gauge theory on $\mathbb{R}^2 \times T^2$. This theory is described by a linear quiver shown in Figure 15(b). A circle node represents a vector multiplet for an SU(k) gauge group, a square node an SU(k) flavor group, and an edge a bifundamental hypermultiplet.

The value $\phi_x^{i+1} - \phi_x^i$ determines the gauge coupling of the ith SU(k) gauge group, while the difference of the periodic scalars on NS5$_i$ and NS5$_{i+1}$ gives the θ-angle for this group; together they form a complexified gauge coupling. The positions $z_\gamma^\uparrow$ and $z_\gamma^\downarrow$ of the D4-branes in C determine the masses of the hypermultiplets charged under the left and right SU(k) flavor groups, respectively. For generic values of ϕ_y^i, the theory is in the Higgs phase in which the gauge symmetry is completely broken.

The topological twist used in the construction of the six-dimensional topological–holomorphic theory becomes the Donaldson–Witten twist of the linear quiver theory, as can be seen as follows. If there are only the NS5- and D4-branes, the dualities used above can be applied to a more general setup where M is the product of a three-manifold W and S^1, instead of $\mathbb{R}^2 \times T^2$. By dimensional reduction on S^1, the linear quiver theory reduces to a three-dimensional $\mathcal{N} = 4$ supersymmetric gauge theory on W. There are two topological twists for a general $\mathcal{N} = 4$ supersymmetric gauge theory [99], and what we get here is the one using the SU(2) R-symmetry coming from the rotation symmetry of $\mathbb{R}^3_{679}$. This is known to be the dimensional reduction of the Donaldson–Witten twist.

The presence of the B-field and other ϵ-dependent part of the background has the effect of introducing the standard Ω-deformation. A quick way to see this is to note that if we apply S-duality, T-duality in the horizontal direction of T^2, and T-duality on E to the brane configuration in Table 2, we arrive at an almost identical Hanany–Witten configuration, in which E is replaced with the dual elliptic curve $E^\vee$. The linear quiver theory realized by this brane configuration is clearly subjected to the standard Ω-deformation because the last T-duality is applied to a twisted product of $\mathbb{R}^2$ and E and, as discussed earlier, this is how the standard Ω-deformation is constructed. The theories realized by the two Hanany–Witten configurations are related by a diffeomorphism between the elliptic curves, so the deformations they receive are the same.

The D0-branes insert local operators in the linear quiver theory, while the D2-branes create surface operators supported on $\{0\} \times T^2$. In particular, each D0-brane acts on the partition function of the Ω-deformed linear quiver theory as a transfer matrix.

Let us consider the situation where all $\mathrm{D4}^\uparrow_\gamma$ and $\mathrm{D4}^\downarrow_\gamma$ end on the same NS5-brane, say $\mathrm{NS5}_1$. In this case, this transfer matrix is constructed from k copies of a rational version of $L^{(N,1)}$ corresponding to the decompactification of E to $\mathbb{C}$.[10]

If we further specialize to the case $N = 2$, these L-operators are R-matrices for the rational six-vertex model (the rational limit of the eight-vertex model) whose vertical lines carry Verma modules of $\mathfrak{sl}_2$. The module structure comes from dynamical creation and annihilation of D2-branes stretched between $\mathrm{D4}^\downarrow_\gamma$ and $\mathrm{NS5}_2$ [5].

A transfer matrix of the rational six-vertex model is a generating function of the conserved charges of the XXX spin chain. Thus, our brane construction naturally explains the appearance of the "noncompact" XXX spin chain of length k, whose spins take values in Verma modules of $\mathfrak{sl}_2$, from the Ω-deformed $\mathcal{N} = 2$ supersymmetric gauge theory with a single $\mathrm{SU}(k)$ gauge group and two fundamental hypermultiplets [4, 5]. This phenomenon generalizes to any $N \geq 2$, for which an $\mathfrak{sl}_N$ spin chain arises [6–8].

Now let us make C compact again, taking $C = E$. Then, the D4–NS5 brane configuration realizes a six-dimensional lift of the linear quiver theory

[10]The dynamical parameter is absent for $C = \mathbb{C}$ as we explain in Section 6.5, so the decompactification acts as if wrapping T^2 with a surface operator and then taking the rational limit. The transfer matrix still consists of $L^{(N,1)}$ if the positions of the 't Hooft lines are pairwise interchanged. To be precise, the L-operator that enters the transfer matrix is not equal but gauge equivalent to $L^{(N,1)}$ because we have defined $L^{(N,1)}$ as the L-operator in the background with $\mathcal{A}_x = \mathcal{A}_y = 0$.

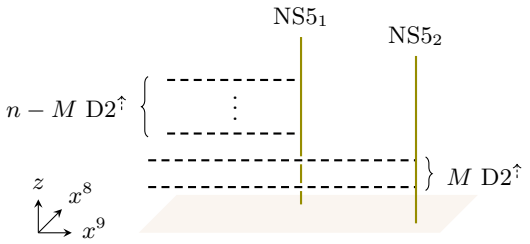


Figure 16: A brane configuration for a two-dimensional $\mathcal{N}=(4,4)$ supersymmetric gauge theory.

compactified on E, as one can see by applying T-duality on E. Correspondingly, the six-vertex model is promoted to the eight-vertex model, whose transfer matrix generates the conserved charges of the XYZ spin chain. If we compactify only one direction so that $C=\mathbb{C}^\times$, the brane configuration produces a five-dimensional gauge theory and the XXZ spin chain.

6.5. Nekrasov–Shatashvili realization of compact spin chains

In the same brane configuration, the crossings of the D0- and D2-branes create transfer matrices constructed from R-matrices in the vector representation of $\mathfrak{sl}_N$. Therefore, the $\mathfrak{sl}_N$ spin chains with spins in the vector representation also appear in this setup. It is interesting to look at these spin chains from the point of view of the D2-branes.

For the moment let us take $N=2$, so there are two NS5-branes. The possible configurations of n D2-branes ending on either NS5-brane are classified by an integer M such that $0\le M\le n$, namely the number of D2-branes ending on NS5$_2$. This is the magnon number of the spin chain, counting the total number of "up" spins in the chain. A case with $M=2$ is illustrated in Figure 16.

In the case when $C=\mathbb{C}$ and the Ω-deformation is absent, the D2–NS5 brane configuration with fixed M realizes an $\mathcal{N}=(4,4)$ supersymmetric gauge theory on T^2. This theory has a U(M) gauge group and a hypermultiplet in the bifundamental representation of the gauge group and a U(n) flavor symmetry.

The separation $\phi_y^2-\phi_y^1$ of the NS5-branes in the x^9-direction determines the gauge coupling. The separation $\phi_x^2-\phi_x^1$ in the x^8-direction is proportional to the Fayet–Iliopoulos (FI) parameter r for the U(1) part of the gauge group, and it combines with the two-dimensional θ-angle ϑ to form

the complexified FI parameter

$$t = \frac{\vartheta}{2\pi} + \mathrm{i}r\,. \tag{321}$$

The positions z_β of $\mathrm{D2}^{\hat{\imath}}_{\beta}$ in C are twisted masses of the hypermultiplet. These complex mass parameters may be thought of as the eigenvalues of the scalar field in the nondynamical vector multiplet for the $\mathrm{U}(n)$ flavor symmetry.

To this theory the Ω-deformation is applied. This makes use of the $\mathrm{U}(1)$ isometry of a plane transverse to the D2-branes, and breaks the $\mathcal{N} = (4,4)$ supersymmetry to $\mathcal{N} = (2,2)$ supersymmetry. In the language of $\mathcal{N} = (2,2)$ supersymmetry, the $\mathcal{N} = (4,4)$ vector multiplet consists of a vector multiplet and a chiral multiplet in the adjoint representation, whereas the $\mathcal{N} = (4,4)$ fundamental hypermultiplet splits into a pair of fundamental and antifundamental chiral multiplets. The Ω-deformation gives the adjoint chiral multiplet a twisted mass u proportional to ϵ, and the fundamental and antifundamental chiral multiplets twisted masses $-u/2$ and $u/2$, respectively [86].

The topological twist is the A-twist here. We can see this from the fact that the scalar field σ of the vector multiplet for the gauge symmetry, whose eigenvalues parametrize the positions of the D2-branes on C, is unaffected by the twist. Alternatively, we may note that the D2-branes are surface operators in the theory on the D4-branes, and the Donaldson–Witten twist reduces to the A-twist in two dimensions.

Suppose that $r \neq 0$ and the twisted masses are vanishing, including those induced by the Ω-deformation. Then, the theory is in the Higgs phase and flows in the infrared to a topological sigma model whose target space is the cotangent bundle $T^*\mathrm{Gr}(M,n)$ of the Grassmannian $\mathrm{Gr}(M,n)$, endowed with a hyperkähler metric. This is the A-model [100], and its algebra of local operators is given by the quantum cohomology ring $QH^\bullet(T^*\mathrm{Gr}(M,n))$. By the state–operator correspondence this is isomorphic as a vector space to the Hilbert space of states.

Now we turn on all the twisted masses. As the supercharge of an A-twisted gauge theory squares to a gauge transformation generated by the adjoint scalar in the vector multiplet, this amounts to working equivariantly with respect to the $\mathrm{U}(n)$ flavor symmetry as well as the $\mathrm{U}(1)$ isometry used in the Ω-deformation. The algebra of local operators is therefore deformed to the equivariant quantum cohomology $QH^\bullet_{(\mathbb{C}^\times)^n \times \mathbb{C}^\times}(T^*\mathrm{Gr}(M,n))$, where $(\mathbb{C}^\times)^n$ is the diagonal torus of the complexification of the $\mathrm{U}(n)$ flavor symmetry and the last $\mathbb{C}^\times$ is the complexification of the $\mathrm{U}(1)$ isometry.

The D0-branes create local operators in the theory. According to our brane construction, these operators can be understand as transfer matrices, constructed from a rational version of $R^{(N,0)}$. Although there is a rational solution of the dynamical Yang–Baxter equation [56], what we get here is a *nondynamical* one: flat connections on $C = \mathbb{C}$ are all gauge equivalent to zero, so the relevant moduli space has no directions that would correspond to a dynamical parameter. (In our brane setup ϕ goes to a constant value at the infinity of C, and with this boundary condition the argument in Section 4 applies.) The transfer matrices of this rational R-matrix are those of the XXX spin chain whose spins are in the vector representation.

By integrability, these transfer matrices generate a commutative algebra of operators, called a *Bethe algebra*, which has the same dimension as the Hilbert space of the spin chain. Since the total spin is a conserved quantity in the XXX spin chain, this is the direct sum of $n+1$ commutative algebras, each acting on a subspace of a fixed magnon number. In the present setup, the local operators created by the D0-branes generate the summand corresponding to the M-magnon sector. The dimension of this summand is actually equal to the dimension of $QH^\bullet_{(\mathbb{C}^\times)^n\times\mathbb{C}^\times}(T^*\mathrm{Gr}(M,n))$, so the D0-branes generate the whole algebra of local operators of the A-model. Hence, the Bethe algebra for the M-magnon sector of the XXX spin chain of length n is isomorphic to $QH^\bullet_{(\mathbb{C}^\times)^n\times\mathbb{C}^\times}(T^*\mathrm{Gr}(M,n))$.

Our brane construction thus explains the correspondence between the XXX spin chain and the equivariant cohomology of the cotangent bundles of Grassmannians, discovered by Nekrasov and Shatashvili [2, 3] and mathematically developed in [101–103]. The above brane configuration has been studied in this context in [104].

If we take $C = \mathbb{C}^\times$, the brane configuration realizes a three-dimensional lift of the above theory, and the rational R-matrix is replaced with a trigonometric one. This is again a nondynamical R-matrix for the following reason. Physically, we expect that the trigonometric case is equivalent to the limit $\tau \to \mathrm{i}\infty$ of the elliptic case where E degenerates to a cylinder. If the dynamical parameter $\bar{\lambda}$ is fixed in this limit, the dynamical elliptic R-matrix becomes a dynamical trigonometric R-matrix. In our case, however, $\bar{\lambda}$ is determined by the background gauge field according to formula (161). Provided that the holonomy of A around the one-cycle $\mathcal{C}_a$ is generic and fixed, taking $\tau \to \mathrm{i}\infty$ entails the limit $|\lambda_i| \to \infty$. In this limit of infinite dynamical parameter, the dynamical trigonometric R-matrix reduces to a nondynamical one.

Hence, the three-dimensional theory corresponds to the M-magnon sector of the XXZ spin chain, whose transfer matrices coincide with those of the nondynamical trigonometric R-matrix. The trigonometric case of

the Nekrasov–Shatashvili correspondence has been mathematically established [105].

For the elliptic case $C = E$, one may be tempted to say that the four-dimensional lift of the theory should correspond to the "M-magnon sector" of the XYZ spin chain. However, such a statement does not make sense since the total spin is not a conserved quantity in the XYZ spin chain. This is not a contradiction. The point is that the R-matrix for the XYZ spin chain is Baxter's nondynamical elliptic R-matrix, while what the theory gives is the dynamical elliptic R-matrix $R^{(N,0)}$. The transfer matrices of $R^{(N,0)}$ do preserve the total spin. The correct statement is therefore that the Higgs branch of the four-dimensional theory corresponds to the Bethe algebra for the M-magnon sector of the spin chain defined by $R^{(N,0)}$.

For general $N \geq 2$, the configurations of n D2-branes ending on N NS5-branes are classified by integers $\mathbf{M} = (M_0, \dots, M_N)$ such that

$$0 = M_0 \leq M_1 \leq \cdots \leq M_N = n\,. \tag{322}$$

The gauge theory on the D2-branes has gauge group $\mathrm{U}(M_1) \times \cdots \times \mathrm{U}(M_{N-1})$ and flavor group $\mathrm{U}(M_N)$, and a bifundamental hypermultiplet of $\mathrm{U}(M_i) \times \mathrm{U}(M_{i+1})$ for each $i = 1, \dots, N-1$. For generic values of the FI parameters, it flows to the A-model (or its three- or four-dimensional lift) whose target space is the cotangent bundle $T^*\mathcal{F}_{\mathbf{M}}$ of the partial flag manifold $\mathcal{F}_{\mathbf{M}}$ parametrizing chains of subspaces

$$0 = F_0 \subset F_1 \subset \cdots \subset F_N = \mathbb{C}^n\,, \qquad \dim F_i = M_i\,. \tag{323}$$

Everything we have said about $T^*\mathrm{Gr}(M,n)$ generalizes straightforwardly to $T^*\mathcal{F}_{\mathbf{M}}$, and we find that the subsector of an $\mathfrak{sl}_N$ spin chain with total $\mathfrak{sl}_N$ weight $\sum_{i=1}^{N}(M_i - M_{i-1})\bar{\omega}_{N-i+1}$ arises from this theory.

6.6. Q-operators

In all of these spin chains there are important operators called *Q-operators*, which are of great help in solving the spectra. One of the main results of [2, 3] is that the Q-operator $Q(z)$ for the XXX spin chain is identified with the local operator

$$\det(z - \sigma) \tag{324}$$

in the gauge theory. (Similar results have been obtained for the trigonometric case in [106, 107].) We can understand this identification as follows.

Let us enrich the system by introducing an additional 't Hooft line along the horizontal direction of the lattice, with the Dirac string extending along $C = \mathbb{C}$ and going off to ∞. By following the chain of dualities we see that this is another kind of D2-brane, which covers the Ω-deformation plane $\mathbb{R}^2$ and ends on one of the two NS5-branes, say $\mathrm{NS5}_2$, from the positive x^9-direction. (Which NS5-brane it ends on is immaterial due to the symmetry under flipping of all spins, or the isomorphism $\mathrm{Gr}(M,n) \cong \mathrm{Gr}(n-M,n)$, or Hanany–Witten transitions involving D6-branes.) This D2-brane is supported at a point on T^2, hence creates a local operator in the theory.

From the point of view of the linear quiver theory on $\mathbb{R}^2 \times T^2$, the original D2-branes and the additional one represent two kinds of surface operators, one extending along T^2 and the other along $\mathbb{R}^2$. Such intersecting surface operators were studied in [108] by means of the correspondence to Liouville theory [109]. There it was found that open strings stretched between intersecting D2-branes give rise to a zero-dimensional $\mathcal{N} = (0,2)$ Fermi multiplet at the intersection. In the present case, this multiplet takes values in the bifundamental representation of $\mathrm{U}(M) \times \mathrm{U}(1)$, where $\mathrm{U}(M)$ is the flavor symmetry on the original M D2-branes attached to $\mathrm{NS5}_2$ (which is the gauge symmetry of the two-dimensional gauge theory) and $\mathrm{U}(1)$ is the global symmetry on the additional D2-brane. The partition function of this multiplet turns out to be given precisely by the operator (324), with z being the value of the scalar field in the $\mathrm{U}(1)$ vector multiplet, or the position of the additional D2-brane on C.

Thus, we identify $Q(z)$ with a horizontal 't Hooft line with spectral parameter z crossing the vertical Wilson lines. In our forthcoming paper, we will present a more explicit derivation of this identification using a description of surface operators in four-dimensional Chern–Simons theory in terms of two-dimensional degrees of freedom.

6.7. Theories for open spin chains

In the above discussions on the appearances of spin chains from linear quiver theories and theories related to the cotangent bundles of partial flag manifolds, it is crucial that the horizontal direction of T^2 is periodic because we need to use T-duality in this direction to arrive at the relevant brane configurations. Consequently, the spin chains appearing in these theories are closed ones with periodic boundary conditions.

In four-dimensional Chern–Simons theory, however, there is nothing that stops us from considering lattices on a noncompact surface such as $\mathbb{R}^2$. Even though we can no longer apply the T-duality then, S-duality still leads to

an interesting configuration consisting of NS5-branes $\mathrm{NS5}_i$, D1-branes $\mathrm{D1}^{\dashrightarrow}_{\alpha}$ and $\mathrm{D1}^{\dashuparrow}_{\beta}$, and D3-branes $\mathrm{D3}^{\uparrow}_{\gamma}$ and $\mathrm{D3}^{\downarrow}_{\gamma}$.

For $C = \mathbb{C}$, the part of the system comprised of $\mathrm{NS5}_i$ and $\mathrm{D1}^{\dashuparrow}_{\beta}$ realizes a one-dimensional gauge theory which is the dimensional reduction of the two-dimensional $\mathcal{N} = (2,2)$ supersymmetric gauge theory discussed above. Our construction therefore implies that an *open* $\mathfrak{sl}_N$ spin chain arises from this one-dimensional theory.

Since the horizontal Wilson lines now extend indefinitely, they do not represent transfer matrices anymore. Rather, these Wilson lines crossing the vertical ones are monodromy matrices T_α constructed from the rational version of $R^{(N,0)}$, which is a nondynamical R-matrix. The monodromy matrices satisfy the RLL relation (254), with T_α taking the place of $L^{(N,0)}_\alpha$. As such, they provide a representation of the corresponding quantum algebra, namely the Yangian of $\mathfrak{sl}_N$.

Again, the theory flows to a sigma model on $T^*\mathcal{F}_{\mathbf{M}}$ for generic values of the FI parameters. Due to the topological twist this is topological quantum mechanics on $T^*\mathcal{F}_{\mathbf{M}}$, whose algebra of local operators is the equivariant cohomology $H^\bullet_{(\mathbb{C}^\times)^n \times \mathbb{C}^\times}(T^*\mathcal{F}_{\mathbf{M}})$. Thus, the action of monodromy matrices on the Hilbert space defines an action of the Yangian on $H^\bullet_{(\mathbb{C}^\times)^n \times \mathbb{C}^\times}(T^*\mathcal{F}_{\mathbf{M}})$. This statement was proved in [101–103].

Similarly, we obtain in the trigonometric case an action of the quantum loop algebra of $\mathfrak{sl}_N$ on the equivariant K-theory of $T^*\mathcal{F}_{\mathbf{M}}$ [105], and in the elliptic case an action of the elliptic quantum group of $\mathfrak{sl}_N$ on the equivariant elliptic cohomology of $T^*\mathcal{F}_{\mathbf{M}}$ [110–112].

6.8. Yangians in three-dimensional linear quiver theories

The D3–NS5 part of the above system, for $C = \mathbb{C}$, is the original Hanany–Witten configuration for a three-dimensional $\mathcal{N} = 4$ supersymmetric gauge theory, described by a linear quiver with $\mathrm{U}(k)$ gauge and flavor groups [98]. In this theory $\mathrm{D1}^{\dashrightarrow}_{\alpha}$ create local operators which involve monopoles, and they represent monodromy matrices with k vertical lines carrying infinite-dimensional representations of $\mathfrak{sl}_N$.

For generic values of the FI parameters, the theory is in the Higgs phase, and the topological twist and the Ω-deformation reduce it to topological quantum mechanics on the moduli space of vortices [113]. Therefore, $\mathrm{D1}^{\dashrightarrow}_{\alpha}$ act as the Yangian on the equivariant cohomology of this space. This conclusion fits nicely with results obtained in [9, 10], where it was found that the algebra

of local operators of the topologically twisted and Ω-deformed linear quiver theory is a certain quotient of the Yangian.

Acknowledgments

We would like to thank Davide Gaiotto for valuable discussions. This work is supported by the NSERC Discovery Grant program and by the Perimeter Institute for Theoretical Physics. Part of the research was conducted while JY was supported by the ERC Starting Grant No. 335739 "Quantum fields and knot homologies" funded by the European Research Council under the European Union's Seventh Framework Programme and by the National Key Research and Development Program of China (No. 2020YFA0713000). Research at Perimeter Institute is supported by the Government of Canada through Industry Canada and by the Province of Ontario through the Ministry of Research and Innovation.

References

[1] R. J. Baxter, Exactly Solved Models in Statistical Mechanics, Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], London, (1989).

[2] N. A. Nekrasov and S. L. Shatashvili, *Supersymmetric vacua and Bethe ansatz*, Nucl. Phys. B Proc. Suppl. **192/193** (2009), 91–112.

[3] N. A. Nekrasov and S. L. Shatashvili, *Quantum integrability and supersymmetric vacua*, Prog. Theor. Phys. Suppl. **177** (2009), 105–119.

[4] N. A. Nekrasov and S. L. Shatashvili, *Quantization of integrable systems and four dimensional gauge theories*, in: XVIth International Congress on Mathematical Physics, p. 265, World Sci. Publ., Hackensack, NJ, (2010).

[5] N. Dorey, S. Lee, and T. J. Hollowood, *Quantization of integrable systems and a 2d/4d duality*, JHEP **10** (2011) 077.

[6] H.-Y. Chen, N. Dorey, T. J. Hollowood and S. Lee, *A new 2d/4d duality via integrability*, JHEP **09** (2011) 040.

[7] N. Nekrasov and V. Pestun, *Seiberg–Witten geometry of four dimensional $N = 2$ quiver gauge theories*, `arXiv:1211.2240`.

[8] N. Nekrasov, V. Pestun, and S. Shatashvili, *Quantum geometry and quiver gauge theories*, Comm. Math. Phys. **357** (2018), 519–567.

[9] M. Bullimore, T. Dimofte, and D. Gaiotto, *The Coulomb branch of 3d $\mathcal{N} = 4$ theories*, Comm. Math. Phys. **354** (2017) 671–751.

[10] A. Braverman, M. Finkelberg, and H. Nakajima, *Coulomb branches of 3d $\mathcal{N} = 4$ quiver gauge theories and slices in the affine Grassmannian (with appendices by Alexander Braverman, Michael Finkelberg, Joel Kamnitzer, Ryosuke Kodera, Hiraku Nakajima, Ben Webster, and Alex Weekes)*, `arXiv:1604.03625`.

[11] D. Gaiotto, L. Rastelli and S. S. Razamat, *Bootstrapping the superconformal index with surface defects*, JHEP **01** (2013) 022.

[12] A. Gadde and S. Gukov, *2d index and surface operators*, JHEP **03** (2014) 080.

[13] L. F. Alday, M. Bullimore, M. Fluder, and L. Hollands, *Surface defects, the superconformal index and q-deformed Yang–Mills*, JHEP **10** (2013) 018.

[14] D. Gaiotto and S. S. Razamat, *$\mathcal{N} = 1$ theories of class $\mathcal{S}_k$*, JHEP **07** (2015) 073.

[15] K. Maruyoshi and J. Yagi, *Surface defects as transfer matrices*, PTEP. Prog. Theor. Exp. Phys. (2016) 113B01.

[16] J. Yagi, *Surface defects and elliptic quantum groups*, JHEP **06** (2017) 013.

[17] Y. Ito and Y. Yoshida, *Superconformal index with surface defects for class $\mathcal{S}_k$*, Nuclear Phys. B **962** (2021) 115277.

[18] K. Costello, *Supersymmetric gauge theory and the Yangian*, `arXiv:1303.2632`.

[19] K. Costello, *Integrable lattice models from four-dimensional field theories*, in: String-Math 2013, vol. 88 of Proc. Sympos. Pure Math., pp. 3–23, Amer. Math. Soc., Providence, RI, (2014).

[20] K. Costello, E. Witten, and M. Yamazaki, *Gauge theory and integrability, I*, ICCM Notices **6** (2018) 46–119.

[21] K. Costello, E. Witten, and M. Yamazaki, *Gauge theory and integrability, II*, ICCM Notices **6** (2018), 120–146.

[22] N. A. Nekrasov, *Seiberg–Witten prepotential from instanton counting*, Adv. Theor. Math. Phys. **7** (2003), 831–864.

[23] N. A. Nekrasov and A. Okounkov, *Seiberg–Witten theory and random partitions*, in: The Unity of Mathematics, vol. 244 of Progr. Math., p. 525, Birkhäuser Boston, Boston, MA, (2006).

[24] N. Nekrasov and E. Witten, *The Omega deformation, branes, integrability and Liouville theory*, JHEP **09** (2010) 092.

[25] E. Witten, *Topological quantum field theory*, Comm. Math. Phys. **117** (1988), 353–386.

[26] E. Witten, *Mirror manifolds and topological field theory*, in: Essays on Mirror Manifolds, pp. 120–158, Int. Press, Hong Kong, (1992).

[27] K. Hori, S. Katz, A. Klemm, R. Pandharipande, R. Thomas, C. Vafa, et al., Mirror Symmetry, Vol. 1 of Clay Mathematics Monographs, American Mathematical Society, Providence, RI, (2003).

[28] J. Yagi, Ω-*deformation and quantization*, JHEP **08** (2014) 112.

[29] Y. Luo, M.-C. Tan, J. Yagi, and Q. Zhao, Ω-*deformation of B-twisted gauge theories and the 3d-3d correspondence*, JHEP **02** (2015) 047.

[30] N. Nekrasov, *Tying up instantons with anti-instantons*, pp. 351–388. World Sci. Publ., 2018.

[31] M. Jimbo, T. Miwa, and M. Okado, *Solvable lattice models whose states are dominant integral weights of* $A^{(1)}_{n-1}$, Lett. Math. Phys. **14** (1987), 123–131.

[32] M. Jimbo, T. Miwa, and M. Okado, *Local state probabilities of solvable lattice models: an* $A^{(1)}_{n-1}$ *family*, Nucl. Phys. B **300** (1988), 74–108.

[33] G. Felder, *Conformal field theory and integrable systems associated to elliptic curves*, in: Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Zürich, 1994), pp. 1247–1255, Birkhäuser, Basel, (1995).

[34] G. Felder, *Elliptic quantum groups*, in: XIth International Congress of Mathematical Physics (Paris, 1994), pp. 211–218, Int. Press, Cambridge, MA, (1995).

[35] R. J. Baxter, *Eight-vertex model in lattice statistics*, Phys. Rev. Lett. **26** (1971), 832–833.

[36] R. J. Baxter, *Partition function of the eight-vertex lattice model*, Ann. Phys. **70** (1972), 193–228.

[37] A. A. Belavin, *Dynamical symmetry of integrable quantum systems*, Nucl. Phys. B **180** (1981), 189–200.

[38] A. Hanany and K. D. Kennaway, *Dimer models and toric diagrams*, `arXiv:hep-th/0503149`.

[39] S. Franco, A. Hanany, D. Vegh, B. Wecht, and K. D. Kennaway, *Brane dimers and quiver gauge theories*, JHEP (2006) 096.

[40] D. Gaiotto, $N = 2$ *dualities*, JHEP **08** (2012) 034.

[41] D. Gaiotto, G. W. Moore, and A. Neitzke, *Wall-crossing, Hitchin systems, and the WKB approximation*, Adv. Math. **234** (2013), 239–403.

[42] E. Witten, *Solutions of four-dimensional field theories via M-theory*, Nucl. Phys. B **500** (1997), 3–42.

[43] M. Atiyah, *Magnetic monopoles and the Yang–Baxter equations*, Int. J. Mod. Phys. A **6** (1991), 2761–2774.

[44] V. P. Spiridonov, *Elliptic beta integrals and solvable models of statistical mechanics*, in: Algebraic Aspects of Darboux Transformations, Quantum Integrable Systems and Supersymmetric Quantum Mechanics, Vol. 563 of Contemp. Math., p. 181, Amer. Math. Soc., Providence, RI, (2012).

[45] M. Yamazaki, *Quivers, YBE and 3-manifolds*, JHEP **05** (2012) 147.

[46] J. Yagi, *Quiver gauge theories and integrable lattice models*, JHEP **10** (2015) 065.

[47] V. V. Bazhanov and S. M. Sergeev, *A master solution of the quantum Yang-Baxter equation and classical discrete integrable equations*, Adv. Theor. Math. Phys. **16** (2012), 65–95.

[48] V. V. Bazhanov and S. M. Sergeev, *Elliptic gamma-function and multi-spin solutions of the Yang-Baxter equation*, Nucl. Phys. B **856** (2012), 475–496.

[49] M. Ashwinkumar, M.-C. Tan, and Q. Zhao, *Branes and categorifying integrable lattice models*, Adv. Theor. Math. Phys. **24** (2020), 1–24.

[50] A. Kapustin and E. Witten, *Electric-magnetic duality and the geometric Langlands program*, Commun. Number Theory Phys. **1** (2007), 1–236.

[51] K. K. Uhlenbeck, *Removable singularities in Yang–Mills fields*, Comm. Math. Phys. **83** (1982), 11–29.

[52] K. Hori, A. Iqbal, and C. Vafa, *D-branes and mirror symmetry*, `arXiv:hep-th/0005247`.

[53] E. Witten, *A new look at the path integral of quantum mechanics*, in: Surveys in Differential Geometry. Volume XV. Perspectives in Mathematics and Physics, vol. 15 of Surv. Differ. Geom., p. 345, Int. Press, Somerville, MA, (2011).

[54] M. S. Narasimhan and C. S. Seshadri, *Stable and unitary vector bundles on a compact Riemann surface*, Ann. of Math. (2) **82** (1965), 540–567.

[55] J. Yagi, *Branes and integrable lattice models*, Mod. Phys. Lett. A **32** (2017), 1730003, 16.

[56] P. Etingof and A. Varchenko, *Solutions of the quantum dynamical Yang-Baxter equation and dynamical quantum groups*, Comm. Math. Phys. **196** (1998), 591–640.

[57] R. J. Baxter, *Eight-vertex model in lattice statistics and one-dimensional anisotropic Heisenberg chain. II. Equivalence to a generalized ice-type lattice model*, Ann. Phys. **76** (1973), 25–47.

[58] P. Etingof, *On the dynamical Yang–Baxter equation*, in: Proceedings of the International Congress of Mathematicians, Vol. II (Beijing, 2002), pp. 555–570, Higher Ed. Press, Beijing, (2002).

[59] A. Kapustin, *Wilson–'t Hooft operators in four-dimensional gauge theories and S-duality*, Phys. Rev. D (3) **74** (2006) 025005.

[60] S. Gukov and E. Witten, *Gauge theory, ramification, and the geometric Langlands program*, in: Current Developments in Mathematics, 2006, p. 35, Int. Press, Somerville, MA, (2008).

[61] M. P. Richey and C. A. Tracy, $\mathbb{Z}_n$ *Baxter model: symmetries and the Belavin parametrization*, J. Stat. Phys. **42** (1986), 311–348.

[62] A. A. Belavin and V. G. Drinfeld, *Solutions of the classical Yang–Baxter equation for simple Lie algebras*, Funktsional. Anal. i Prilozhen. **16** (1982), 1–29, 96.

[63] A. M. Levin, M. A. Olshanetsky, and A. V. Zotov, *Monopoles and modifications of bundles over elliptic curves*, SIGMA **5** (2009) 065.

[64] A. Levin, M. Olshanetsky, A. Smirnov, and A. Zotov, *Characteristic classes of* $\mathrm{SL}(N,\mathbb{C})$*-bundles and quantum dynamical elliptic R-matrices*, J. Phys. A **46** (2013) 035201.

[65] K. Hasegawa, *Ruijsenaars' commuting difference operators as commuting transfer matrices*, Comm. Math. Phys. **187** (1997), 289–325.

[66] S. N. M. Ruijsenaars, *Complete integrability of relativistic Calogero-Moser systems and elliptic function identities*, Comm. Math. Phys. **110** (1987), 191–213.

[67] O. Foda, K. Iohara, M. Jimbo, R. Kedem, T. Miwa, and H. Yan, *An elliptic quantum algebra for* $\widehat{\mathrm{sl}}_2$, Lett. Math. Phys. **32** (1994), 259–268.

[68] H. Fan, B.-y. Hou, K.-j. Shi, and W.-l. Yang, *The elliptic quantum algebra* $A_{q,p}(\widehat{sl}_n)$ *and its bosonization at level one*, J. Math. Phys. **39** (1998), 4356–4368.

[69] J. Avan, L. Frappat, M. Rossi, and P. Sorba, *Deformed* $\mathcal{W}_N$ *algebras from elliptic* $sl(N)$ *algebras*, Comm. Math. Phys. **199** (1999), 697–728.

[70] M. Jimbo, H. Konno, S. Odake, and J. Shiraishi, *Quasi-Hopf twistors for elliptic quantum groups*, Transform. Groups **4** (1999), 303–327.

[71] E. K. Sklyanin, *Some algebraic structures connected with the Yang–Baxter equation*, Funktsional. Anal. i Prilozhen. **16** (1983), 27–34.

[72] I. V. Cherednik, *Some finite-dimensional representations of generalized Sklyanin algebras*, Funktsional. Anal. i Prilozhen. **19** (1985), 89–90.

[73] B. Y. Hou and H. Wei, *Algebras connected with the* Z_n *elliptic solution of the Yang–Baxter equation*, J. Math. Phys. **30** (1989), 2750–2755.

[74] Y.-H. Quano and A. Fujii, *Generalized Sklyanin algebra*, Mod. Phys. Lett. A **6** (1991), 3635–3640.

[75] B. Y. Hou, K. J. Shi, and Z. X. Yang, $\mathrm{sl}_q(n)$ *quantum algebra as a limit of the elliptic case*, in: Infinite Analysis, Part A, B (Kyoto, 1991), Vol. 16 of Adv. Ser. Math. Phys., pp. 391–403, World Sci. Publ., River Edge, NJ, (1992).

[76] E. K. Sklyanin, *Some algebraic structures connected with the Yang–Baxter equation. Representations of a quantum algebra*, Funktsional. Anal. i Prilozhen. **17** (1983), 34–48.

[77] B. Y. Hou, K. J. Shi, and Z. X. Yang, *Cyclic representation and function difference representation of the* Z_n *Sklyanin algebra*, J. Phys. A **26** (1993), 4951–4965.

[78] B. Y. Hou, K. J. Shi, W. L. Yang, and Z. X. Yang, *Function space representation of the Z_n Sklyanin algebra*, Phys. Lett. A **178** (1993), 73–80.

[79] H. Konno, *An elliptic algebra $U_{q,p}(\widehat{sl}_2)$ and the fusion RSOS model*, Comm. Math. Phys. **195** (1998), 373–403.

[80] M. Jimbo, H. Konno, S. Odake, and J. Shiraishi, *Elliptic algebra $U_{q,p}(\widehat{\mathfrak{sl}}_2)$: Drinfeld currents and vertex operators*, Comm. Math. Phys. **199** (1999), 605–647.

[81] H. Konno, *Elliptic quantum group $U_{q,p}(\widehat{\mathfrak{sl}}_2)$, Hopf algebroid structure and elliptic hypergeometric series*, J. Geom. Phys. **59** (2009), 1485–1511.

[82] K. Hasegawa, *Crossing symmetry in elliptic solutions of the Yang–Baxter equation and a new L-operator for Belavin's solution*, J. Phys. A **26** (1993), 3211–3228.

[83] M. Bershadsky, C. Vafa, and V. Sadov, *D-branes and topological field theories*, Nucl. Phys. B **463** (1996), 420–434.

[84] M. Bershadsky, A. Johansen, V. Sadov, and C. Vafa, *Topological reduction of 4D SYM to 2D σ-models*, Nucl. Phys. B **448** (1995), 166–186.

[85] J. A. Harvey, G. Moore and A. Strominger, *Reducing S duality to T duality*,Phys. Rev. D (3) **52** (1995), 7161–7167.

[86] S. Hellerman, D. Orlando, and S. Reffert, *String theory of the Omega deformation*, JHEP **01** (2012) 148.

[87] S. Reffert, *General omega deformations from closed string backgrounds*, JHEP **04** (2012) 059.

[88] S. Hellerman, D. Orlando, and S. Reffert, *The omega deformation from string and M-theory*, JHEP **07** (2012) 061.

[89] D. Orlando and S. Reffert, *Deformed supersymmetric gauge theories from the fluxtrap background*, Int. J. Mod. Phys. A **28** (2013) 1330044.

[90] A. Giveon, M. Porrati, and E. Rabinovici, *Target space duality in string theory*, Phys. Rep. **244** (1994), 77–202.

[91] C. V. Johnson, D-Branes, Cambridge Monographs on Mathematical Physics, Cambridge University Press, Cambridge, (2003).

[92] S. Yamaguchi, *Wilson loops of anti-symmetric representation and D5-branes*, JHEP **05** (2006) 037.

[93] J. Gomis and F. Passerini, *Holographic Wilson loops*, JHEP **08** (2006) 074.

[94] J. Gomis and F. Passerini, *Wilson loops as D3-branes*, JHEP **01** (2007) 097.

[95] S. A. Cherkis and A. Kapustin, *Singular monopoles and supersymmetric gauge theories in three dimensions*, Nucl. Phys. B **525** (1998), 215–234.

[96] O. Aharony and A. Hanany, *Branes, superpotentials and superconformal fixed points*, Nucl. Phys. B **504** (1997), 239–271.

[97] O. Aharony, A. Hanany, and B. Kol, *Webs of (p, q) five-branes, five dimensional field theories and grid diagrams*, JHEP **01** (1998) 002.

[98] A. Hanany and E. Witten, *Type IIB superstrings, BPS monopoles, and three-dimensional gauge dynamics*, Nucl. Phys. B **492** (1997), 152–190.

[99] M. Blau and G. Thompson, *Aspects of $N_T \geq 2$ topological gauge theories and D-branes*, Nucl. Phys. B **492** (1997), 545–590.

[100] E. Witten, *Topological sigma models*, Comm. Math. Phys. **118** (1988), 411–449.

[101] V. Gorbounov, R. Rimányi, V. Tarasov, and A. Varchenko, *Quantum cohomology of the cotangent bundle of a flag variety as a Yangian Bethe algebra*, J. Geom. Phys. **74** (2013), 56–86.

[102] D. Maulik and A. Okounkov, *Quantum groups and quantum cohomology*, Astérisque **408** (2019), ix+209.

[103] R. Rimányi, V. Tarasov, and A. Varchenko, *Partial flag varieties, stable envelopes, and weight functions*, Quantum Topol. **6** (2015), 333–364.

[104] D. Orlando and S. Reffert, *Twisted masses and enhanced symmetries: the A&D series*, JHEP **02** (2012) 060.

[105] R. Rimányi, V. Tarasov, and A. Varchenko, *Trigonometric weight functions as K-theoretic stable envelope maps for the cotangent bundle of a flag variety*, J. Geom. Phys. **94** (2015), 81–119.

[106] P. P. Pushkar, A. V. Smirnov, and A. M. Zeitlin, *Baxter Q-operator from quantum K-theory*, Adv. Math. **360** (2020) 106919.

[107] P. Koroteev, P. P. Pushkar, A. Smirnov, and A. M. Zeitlin, *Quantum K-theory of quiver varieties and many-body systems*, arXiv: 1705.10419.

[108] J. Gomis, B. Le Floch, Y. Pan, and W. Peelaers, *Intersecting surface defects and two-dimensional CFT*, Phys. Rev. D **96** (2017) 045003.

[109] L. F. Alday, D. Gaiotto, and Y. Tachikawa, *Liouville correlation functions from four-dimensional gauge theories*, Lett. Math. Phys. **91** (2010), 167–197.

[110] M. Aganagic and A. Okounkov, *Elliptic stable envelopes*, J. Amer. Math. Soc. **34** (2021), 79–133.

[111] G. Felder, R. Rimányi, and A. Varchenko, *Elliptic dynamical quantum groups and equivariant elliptic cohomology*, SIGMA **14** (2018) 132.

[112] H. Konno, *Elliptic stable envelopes and finite-dimensional representations of elliptic quantum group*, J. Integrable Syst. **3** (2018) xyy012.

[113] M. Bullimore, T. Dimofte, D. Gaiotto, J. Hilburn, and H.-C. Kim, *Vortices and Vermas*, Adv. Theor. Math. Phys. **22** (2018), 803–917.

Perimeter Institute for Theoretical Physics
Waterloo, ON N2L 2Y5 Canada
E-mail address: kcostello@perimeterinstitute.ca

Yau Mathematical Sciences Center, Tsinghua University
Beijing 100084 China
E-mail address: junyagi@tsinghua.edu.cn