Aspects Of Supergroup Chern-Simons Theories

Victor Mikhaylov

A Dissertation
Presented to the Faculty
of Princeton University
in Candidacy for the Degree
of Doctor of Philosophy

Recommended for Acceptance
by the Department of Physics
Advisor: Edward Witten

September 2015

Abstract

The three-dimensional Chern-Simons gauge theory is a topological quantum field theory, whose correlation functions give metric-independent invariants of knots and three-manifolds. In this thesis, we consider a version of this theory, in which the gauge group is taken to be a Lie supergroup. We show that the analytically-continued version of the supergroup Chern-Simons theory can be obtained by topological twisting from the low energy effective theory of the intersection of D3- and NS5-branes in the type IIB string theory. By S-duality, we deduce a dual magnetic description; and a slightly different duality, in the case of orthosymplectic gauge group, leads to a strong-weak coupling duality between certain supergroup Chern-Simons theories on $\mathbb{R}^3$. Some cases of these statements are known in the literature. We analyze how these dualities act on line and surface operators.

We also consider the purely three-dimensional version of the $\mathfrak{psl}(1|1)$ and the $\mathrm{U}(1|1)$ supergroup Chern-Simons, coupled to a background complex flat gauge field. These theories compute the Reidemeister-Milnor-Turaev torsion in three dimensions. We use the 3d mirror symmetry to derive the Meng-Taubes theorem, which relates the torsion and the Seiberg-Witten invariants, for a three-manifold with arbitrary first Betti number. We also present the Hamiltonian quantization of our theories, find the modular transformations of states, and various properties of loop operators. Our results for the $\mathrm{U}(1|1)$ theory are in general consistent with the results, found for the $\mathrm{GL}(1|1)$ WZW model. We expect our findings to be useful for the construction of Chern-Simons invariants of knots and three-manifolds for more general Lie supergroups.

Acknowledgments

I am deeply grateful to my advisor Edward Witten for his guidance, encouragement and constant support. Working with Edward was a great honor and a great opportunity to learn.

I would like to thank Igor Klebanov and Daniel Marlow for serving on my thesis committee, and Herman Verlinde for agreeing to be a reader of this thesis.

I am much obliged to Alexander Gorsky for educating me in quantum field theory and for organizing a wonderful seminar for students at ITEP, Moscow. I would like to thank M. Dedushenko, I. Denisenko, P. Koroteev, S. Pufu, A. Zhiboedov and all my colleagues for many helpful discussions. Finally, I would like to express gratitude to my family for their love and support.

The work presented in this thesis was partially supported by the Charlotte Elizabeth Procter Fellowship.

Contents

Chapter 1

Introduction

1.1 Topological Quantum Field Theory

A quantum field theory is called topological, if its observables do not depend on the distances. For example, the partition function of such a theory in curved space does not depend on the metric and produces a topological invariant[1]. Such theories are almost trivial in the sense that they do not contain any propagating particles. Nevertheless, they have important applications both in physics and in mathematics. In the real world, these theories describe low-energy limits of gapped systems, and therefore are relevant for the classification of quantum phases of matter. In mathematics, the topological quantum field theory (TQFT) methods have by now become a standard part of topology.

Presumably the first example of a TQFT was considered by A. Schwarz [3] in the late seventies. It is a free, non-interacting theory with the action

$$I = \int \mathrm{d}^3x \, \epsilon^{\mu\nu\rho} A_\mu \partial_\nu A_\rho \,, \tag{1.1}$$

where A_μ is a gauge field in three dimensions and $\epsilon^{\mu\nu\rho}$ is the antisymmetric tensor. (To be

[1] We will not try to make our terminology very precise. What we call "topological invariant" can depend on the choice of a smooth structure, as well as on some other choices.

precise, [3] also considers slightly more general theories.) The action (1.1) is independent of the space-time metric, and is invariant under general coordinate transformations. The theory can be defined on an arbitrary three-manifold W. Since the functional integral with this action is Gaussian, the partition function is simply the inverse of the square root of the determinant of the kinetic operator. To define this determinant, one first needs to fix the usual gauge invariance $A_\mu \to A_\mu + \partial_\mu \alpha$. The gauge-fixing condition necessarily depends on the metric of the manifold, but, as one would expect on physical grounds, this dependence drops out of the partition function. As found in [3], the topological invariant that one gets in this way is what is known as the Ray-Singer torsion, or, equivalently, the combinatorial Reidemeister torsion. A close analog of this theory will be the subject of chapter 3 of this thesis.

Here we would like to briefly discuss some other classical examples of topological theories, which will be important for the present thesis. The three-dimensional Chern-Simons theory is a gauge theory with the action

$$I = \frac{k}{2\pi} \int \mathrm{d}^3 x \; \epsilon^{\mu\nu\rho} \left(A^a_\mu \partial_\nu A^a_\rho + \frac{1}{3} f_{abc} A^a_\mu A^b_\nu A^c_\rho \right) , \tag{1.2}$$

where A^a_μ is a gauge field for some gauge group G, f_{abc} are the structure constants, and k is a coupling constant, which in general has to satisfy some quantization condition, for the path-integral to be gauge-invariant. If the gauge group is U(1), the action clearly reduces to (1.1). Again, the action is metric-independent, and therefore one expects the theory to be topological, provided that it can be regularized in an invariant way. The interesting observables in this theory are Wilson lines for external particles charged under the gauge group G. For example, one can consider a closed Wilson loop, located along some knot in $\mathbb{R}^3$ or in the three-sphere S^3. It has been shown in the foundational paper [4] that the expectation value of such a Wilson loop is the knot invariant, which is known as the Jones polynomial. (This statement applies to the case of gauge group SU(2) and Wilson operators

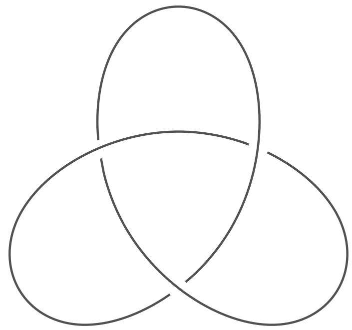

Figure 1.1: The right-handed trefoil knot.

in the two-dimensional representation. For other groups and representations, the polynomials have different names.) It is a Laurent polynomial in the variable $q^{1/2} = \exp(\pi i/k)$, and it can be computed for any given knot by a simple algorithm. It is a topological invariant, in the sense that two knots with different Jones polynomials cannot be continuously deformed into one another without cutting the line. (The opposite, unfortunately, is not true: two knots with the same Jones polynomial need not be identical.) To give an example, for the trefoil knot, shown on fig. 1.1, the polynomial is $P_{\text{trefoil}} = q^{1/2}(-q^4+q^2+q+1)$, while for the unknot it is $P_{\text{unknot}} = q^{1/2} + q^{-1/2}$. These two are different, and the trefoil, indeed, cannot be deformed into the unknot.

The polynomials above have integer coefficients, and the same is true for all Chern-Simons knot polynomials. This, definitely, is a very unusual structure for Wilson loop expectation values in a quantum field theory. Mathematically, the integrality of the coefficients can be explained by the existence of another knot invariant, the Khovanov homology [5]. To a given knot K it associates a vector space $\mathcal{H}_K$, which is bigraded, that is, it has a decomposition into a sum of eigenspaces of two operators F and N. The Jones polynomial can then be obtained as a trace,

$$P_K(q) = \mathrm{Tr}_{\mathcal{H}_K}(-1)^F q^N \,. \tag{1.3}$$

The coefficients of the polynomial are dimensions of subspaces inside $\mathcal{H}_K$, and therefore are integers. Note that the Khovanov homology in general contains more information than the

Jones polynomial, since in taking the trace in (1.3), the eigenvalues of F are relevant only modulo two.

To find a physical interpretation for the Khovanov homology, one needs to construct a four-dimensional TQFT, in which one can define surface operators. (A surface operator is an operator, which is supported on a two-dimensional subspace, like a Wilson line is supported on a one-dimensional subspace.) Suppose that this TQFT is considered on a four-manifold $\mathbb{R}_t \times W$, where $\mathbb{R}_t$ is understood as the time direction, and suppose we add a surface operator, supported on $\mathbb{R}_t \times K$, that is, stretched along the time direction and along the knot $K \subset W$. The Hilbert space of such a topological theory is a vector space, which is naturally a topological invariant of W and K, and, assuming the existence of two conserved charges F and N, has a chance to coincide with the Khovanov homology. Suppose that such a theory is put on $S^1 \times W$. The partition function on this manifold is a trace over the Hilbert space, and, with an insertion of operators $(-1)^F q^N$, would coincide with (1.3). (The trace in the partition function should normally contain the operator $\exp(iTH)$, but the Hamiltonian H of a topological theory is zero.) Therefore, the topological theory in question, upon compactification on a circle, should reduce to the Chern-Simons theory. The TQFT with these properties has indeed been constructed[2] in [6].

Besides Chern-Simons theory, another extremely important example of a TQFT is the Donaldson theory in four dimensions. It can be obtained from the $\mathcal{N} = 2$ supersymmetric Yang-Mills theory by putting it on a curved four-manifold V in a suitable way [7]. Although the Yang-Mills theory has an explicit dependence on the space-time metric, it contains a subsector, singled out by the condition of invariance under a particular fermionic charge Q, in which the correlation functions are metric-independent and define a topological theory. The path-integral in such a theory can be reduced to an integral over the subspace of Q-invariant field configurations. In the case of the Donaldson theory, these are instantons, that is, gauge fields with self-dual field strength. Modulo gauge transformations, this space

[2]For other physical approaches to the Khovanov homology, see [8] and references in [6].

is finite-dimensional. The mathematicians formulate the Donaldson theory in terms of the intersection theory on this finite-dimensional space.

The $\mathcal{N} = 2$ super Yang-Mills theory has a moduli space of vacua, where the gauge group is partially spontaneously broken by an expectation value of an adjoint-valued Higgs field. The theory is asymptotically-free, and therefore at long distances flows to strong coupling, if the expectation value of the Higgs field is not large compared to the Yang-Mills dynamical scale. Usually, it is hard to produce any analytical results for a strongly coupled theory, nevertheless, the exact action for the low-energy effective description of the $\mathcal{N} = 2$ super Yang-Mills theory has been found [9], [10]. This allowed to construct an alternative description of the Donaldson invariants [11]. Indeed, the observables of the topological theory do not depend on the metric, and in particular do not change, if we rescale the metric by a large factor, so that the large-distance effective description of the theory becomes valid. The formulation of this alternative description of the Donaldson theory, known as the Seiberg-Witten invariants, was a major success of topological quantum field theory.

1.2 Overview Of The Thesis

The main subject of the present thesis is the Chern-Simons theory in three dimensions, but with the unusual feature that the gauge group is taken to be a Lie supergroup, rather than an ordinary Lie group.

Chapter 2 of this thesis is based on the paper [1], written in collaboration with Edward Witten. We define and study the analytically-continued version of the supergroup Chern-Simons theory. In the context of ordinary Chern-Simons, the analytical continuation was developed in [12], [13], [6]. It allows to continue the theory to non-integer, and in general even complex values of the level k. This is achieved by defining the path-integral with unusual middle-dimensional integration cycle in the space of complexified fields. To ensure convergence of the integral, the integration cycle is constructed as a Lefschetz thimble for the

Morse function, which is taken to be the real part of the action of the theory. The coordinate parameterizing the Morse flow becomes a new direction in the space, so that the topological theory for the analytically-continued Chern-Simons is essentially four-dimensional. It turns out to be equivalent [6] to the $\mathcal{N} = 4$ super Yang-Mills theory in a half-space, with the Kapustin-Witten twist [14]. In this thesis, we generalize these results to the case of the supergroup Chern-Simons theory. The topological theory in question is obtained by twisting the theory of the D3-NS5 brane intersection. After explaining this construction, we apply various string theory dualities to obtain alternative descriptions of the theory. In particular, we show that the supergroup Chern-Simons invariants can be computed by solving the Kapustin-Witten partial differential equations in the four-dimensional space with a particular three-dimensional defect. We study line and surface operators and their transformations under the S-duality.

An interesting application of our construction arises for the case when the gauge group is taken to be the orthosymplectic supergroup. We point out that the transformation $S^{-1}TS$ of the $\mathrm{SL}(2,\mathbb{Z})$ S-duality group relates the analytically-continued Chern-Simons theories with gauge groups $\mathrm{OSp}(2m+1|2n)$ and $\mathrm{OSp}(2n+1|2m)$. The variable in the knot polynomials is changed as $q \to -q$ under the duality. Since the weak coupling limit corresponds to $q \sim 1$, the duality that we find relates the weak and the strong coupling regimes. We find the transformations of line and surface operators under this duality, and in particular obtain a natural correspondence between non-spinorial representations of the two Lie supergroups. Our results provide a conceptual physical explanation to some known mathematical relations between quantum orthosymplectic supergroups [16], corresponding knot invariants [15] and supergroup conformal field theories [17].

Chapter 3 of this thesis is based on the paper [2]. We consider Chern-Simons theories based on Lie superalgebras $\mathfrak{psl}(1|1)$ and $\mathfrak{u}(1|1)$. We show that they can be coupled to background flat complex gauge fields. With this coupling, these theories compute the invariant of three-manifolds, which is known as the Reidemeister-Milnor-Turaev torsion. We point out that the U(1|1) theory at level k can be obtained by an RG flow from the twisted version of the $\mathcal{N} = 4$ QED with one flavor of charge k. The background flat gauge field comes from a background twisted vector multiplet, whose scalar component defines the FI parameter of the theory in the flat space. The supersymmetric partition function of the QED can be localized on the solutions of the three-dimensional Seiberg-Witten equations. This gives a physical explanation to the theorem of Meng and Taubes [20], which relates the Milnor torsion and the Seiberg-Witten invariants in three dimensions. Our story is in a sense a toy version of the relation between the Donaldson and the Seiberg-Witten invariants in four dimensions, except that here the Seiberg-Witten equations arise in the UV, and not in the IR. For manifolds with small first Betti number, we discuss the matching of the wall-crossing phenomena in the UV and in the IR theories.

We also construct the Hamiltonian quantization of the $\mathfrak{psl}(1|1)$ and the U(1|1) Chern-Simons theories. In particular, the skein relations for the multivariable Alexander polynomial are derived. We illustrate some subtleties that are expected to be important in the quantization of more general supergroup Chern-Simons theories. Our findings are in general agreement with the results, obtained from conformal field theory, however, in this thesis we do not attempt to derive a relation of the supergroup Chern-Simons theories and the WZW models. Finally, we present some brane constructions, realizing the supergroup Chern-Simons theories for general unitary and orthosymplectic gauge groups, and look at possible dualities for those theories.

Chapter 2

Branes And Supergroups

2.1 Introduction

In this Chapter, we consider the analytically-continued version of the Chern-Simons theory with a supergroup. We take an approach, which has been developed in [6] for the case of the ordinary Chern-Simons theory. Let us first give a brief overview of that paper.

2.1.1 Overview Of Previous Work

In the paper [6], the Chern-Simons theory was engineered by a brane construction in type IIB string theory. Consider a stack of n D3-branes, ending on an NS5-brane. The theory on the worldvolume of the D3-branes is the $\mathcal{N} = 4$ super Yang-Mills with gauge group $\mathrm{U}(n)$. One can construct a cohomological TQFT out of it, by making the Kapustin-Witten topological twist [14]. The boundary conditions along the end of the D3-branes on the NS5-branes preserve the topological supercharge Q. The topological theory can then be put on an arbitrary four-manifold M with a three-dimensional boundary W with these boundary conditions. The action turns out to be

$$I = \int_M \{Q, V\} + \frac{i\mathcal{K}}{4\pi} \int_W \mathrm{Tr} \left(\mathcal{A} \wedge \mathrm{d}\mathcal{A} + \frac{2}{3} \mathcal{A} \wedge \mathcal{A} \wedge \mathcal{A} \right). \qquad (2.1)$$

Here $\mathcal{K}$ is a certain complex-valued function of $g_{\rm YM}$ and $\theta_{\rm YM}$ that will be described later.[1] Also, $\mathcal{A}$ is a complexified version of the gauge field, roughly $\mathcal{A}_\mu = A_\mu + i\phi_\mu$, where A_μ is the ordinary gauge field and ϕ_μ denotes some of the scalar fields of $\mathcal{N}=4$ super Yang-Mills theory (which scalar fields enter this formula depends on the choice of Q). The details of the functional V are inessential. Forgetting the scalar field ϕ for a moment, what is written as the Chern-Simons term in (2.1) is really the topological term of the Yang-Mills gauge field in the bulk. Writing it as a Chern-Simons term is correct only as long as one considers small variations of the gauge field. The fact that it is really the bulk topological term means that $\mathcal{K}$ need not be an integer for the path-integral to be gauge-invariant.

If we restrict to Q-invariant observables, localized on the three-dimensional boundary W, the theory with the action (2.1) will actually reproduce the Chern-Simons theory. One important subtlety is that the gauge field is complexified. In fact, as explained in much detail in [12], the four-dimensional topological theory in question is in general equivalent to the Chern-Simons theory with an unusual integration cycle in the path-integral. The middle-dimensional integration cycle in the space of complexified gauge fields can be found by solving the Kapustin-Witten equations along the bulk coordinate, normal to the boundary. Those equations in fact define a gradient flow, with the Morse function being the real part of the action of the theory. The integration cycle then is a Lefschetz thimble. This guarantees that the real part of the action is bounded from below, and the path-integral is convergent. In this thesis, we will mostly try to stay away from the subtleties, related to the choice of the integration contour.

The realization of the (analytically-continued) Chern-Simons theory by a simple brane construction in [6] allowed to apply various string theory dualities, and thus to obtain alternative descriptions of the theory. For example, applying the S-duality, one finds the theory of D3-branes ending on a D5-brane. The corresponding boundary condition in the $\mathcal{N}=4$ super Yang-Mills is known to be the Nahm pole [21]. For this reason, the S-dual, "magnetic"

[1] This function is denoted Ψ in [6, 14]. In the present chapter, we call it $\mathcal{K}$ because of the analogy with the usual Chern-Simons level k.

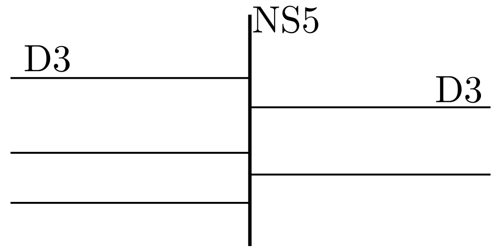

Figure 2.1: An NS5-brane with m D3-branes ending on it from the left and n from the right – sketched here for $m = 3$, $n = 2$. The D3-branes but not the NS5-brane extend in the x_3 direction, which is plotted horizontally, and the NS5-brane but not the D3-branes extend in the $x_{4,5,6}$ directions, which are represented symbolically by the vertical direction in this figure.

description of the Chern-Simons theory is inherently four-dimensional. The path-integral of that theory can be localized on the solutions of the Kapustin-Witten equations with the Nahm pole boundary condition. The space of these solutions is in general discrete. As checked explicitly in [74], counting the solutions reproduces correctly the knot polynomials, with signed counts of the solutions as coefficients. This can be considered as a vast generalization of the theorem of Meng and Taubes [20], which relates the U(1|1) knot polynomials with the three-dimensional Seiberg-Witten invariants. (This special case will be the subject of Chapter 3 of this thesis.)

By applying further a T-duality, one obtains a D4-D5 configuration, and thus a five-dimensional topological field theory in a half-space. It has been conjectured in [6] that the space of supersymmetric ground states in this theory gives a physical realization of the Khovanov homology.

2.1.2 The Two-Sided Problem And Supergroups

In this Chapter, we extend the construction of [6] to the case of Chern-Simons theory with a supergroup. We mainly focus on the U($m|n$) and the OSp($m|2n$) supergroups, for which there exist explicit brane constructions, but our arguments work for other supergroups as well.

We consider the brane configuration of fig. 2.1, with m and n D3-branes on the two sides of the NS5-brane. In field theory, this corresponds to $\mathrm{U}(n)$ and $\mathrm{U}(m)$ maximally-supersymmetric Yang-Mills theories in two half-spaces, joined along a three-dimensional defect. We prove that the action of the theory is given by the same formula (2.1), with an important difference that the gauge field $\mathcal{A}$ is now superalgebra-valued. Namely, it is a sum of a $\mathfrak{u}(n) \oplus \mathfrak{u}(m)$-valued bosonic gauge field, which is obtained by restriction of the bulk gauge fields on the defect W, and a Grassmann one-form field, valued in the bifundamental representation of $\mathrm{U}(m) \times \mathrm{U}(n)$. Therefore, in this two-sided brane configuration, the topological field theory living on the defect is the $\mathrm{U}(m|n)$ supergroup Chern-Simons theory.

This supergroup theory has some peculiarities, which one does not find in the ordinary, bosonic Chern-Simons. We give a brief review of Lie superalgebras and their representations, and then discuss line and surface operators that can be used to define knot invariants in the theory, and some of their properties.

After that we consider some applications. First, as in [6], we apply S-duality and get a description of the theory in terms of the $\mathcal{N} = 4$ Yang-Mills with a D5-type three-dimensional defect. We call this theory "magnetic", while the theory before S-duality is called "electric". The path-integral here can be computed by counting solutions of the Kapustin-Witten equations. This, in principle, gives a way to compute supergroup knot polynomials, though many details remain unclear. We also identify the duals of line and surface operators, found in the electric theory.

Our most interesting application arises for the gauge supergroup $\mathrm{OSp}(2m+1|2n)$. This theory can be realized by essentially the same brane construction, but with an addition of an orientifold three-plane. We find that the element $S^{-1}TS$ of the $\mathrm{SL}(2,\mathbb{Z})$ S-duality group transforms this theory into supergroup Chern-Simons with gauge group $\mathrm{OSp}(2n+1|2m)$. In the special case of $m = 0$, this is a duality of supergroup $\mathrm{OSp}(1|2n)$ Chern-Simons and ordinary, bosonic $\mathrm{O}(2n+1)$ Chern-Simons theory. The variable q in the knot polynomials is mapped under the duality to $-q$. Note that the weak coupling limit is $q \to 1$, so, our duality exchanges the weak and the strong coupling regimes. Again, we describe the mapping of line and surface operators under the duality. In particular, this mapping involves an interesting correspondence between representations of the two supergroups. For the case $m = 0$, this mapping was known in the literature [71].

2.2 Electric Theory

2.2.1 Gauge Theory With An NS-Type Defect

As explained in the introduction, our starting point will be four-dimensional $\mathcal{N} = 4$ super Yang-Mills theory with a three-dimensional half-BPS defect. This theory can be defined in purely gauge-theoretic terms, but it will be useful to consider a brane construction, which gives a realization of the theory for unitary and orthosymplectic gauge groups. We consider a familiar Type IIB setting [36] of D3-branes interacting with an NS5-brane. As sketched in fig. 2.1 of the introduction, where we consider the horizontal direction to be parametrized by[2] $y = x_3$, we assume that there are m D3-branes and thus $\mathrm{U}(m)$ gauge symmetry for $y < 0$ and n D3-branes and thus $\mathrm{U}(n)$ gauge symmetry for $y > 0$. We take the NS5-brane to be at $x_3 = x_7 = x_8 = x_9 = 0$ and hence to be parametrized by x_0, x_1, x_2 and x_4, x_5, x_6, while the semi-infinite D3-branes are parametrized by x_0, x_1, x_2, x_3. With an orientifold projection, which we will introduce in section 2.5, the gauge groups become orthogonal and symplectic. Purely from the point of view of four-dimensional field theory, there are other possibilities.

The theory in the bulk is $\mathcal{N} = 4$ super Yang-Mills, and it is coupled to some three-dimensional bifundamental hypermultiplets, which live on the defect at $y = 0$ and come from the strings that join the two groups of D3-branes. The bosonic fields of the theory are the gauge fields A_i, the scalars $\vec{X}$ that describe motion of the D3-branes along the NS5-brane (that is, in the x_4, x_5, x_6 directions), and scalars $\vec{Y}$ that describe the motion of the D3-branes normal to the NS5-brane (that is, in the x_7, x_8, x_9 directions).

The relevant gauge theory action, including the effects of the defect at $y = 0$, has been

[2] Throughout the chapter, notations y and x_3 are used interchangeably for the same coordinate.

constructed in the paper [23]. In this section we recall some facts about this theory, mostly without derivation. More detailed explanations can be found in the original paper [23] or in the more technical Appendix B below, which is, however, not necessary for understanding the main ideas of the present chapter.

The half-BPS defect preserves $\mathcal{N} = 4$ superconformal supersymmetry in the three-dimensional sense; the corresponding superconformal group is $\mathrm{OSp}(4|4)$. It is important that there exists a one-parameter family of inequivalent embeddings of this supergroup into the superconformal group $\mathrm{PSU}(2,2|4)$ of the bulk four-dimensional theory. For our purposes, it will suffice to describe the different embeddings just from the point of view of global supersymmetry (rather than the full superconformal symmetry). The embeddings differ by which global supersymmetries are preserved by the defect. The four-dimensional bulk theory is invariant under the product $\mathrm{U}_0 = \mathrm{SO}(1,3) \times \mathrm{SO}(6)_R$ of the Lorentz group $\mathrm{SO}(1,3)$ and the R-symmetry group $SO(6)_R$ (or more precisely, a double cover of this associated with spin); this is a subgroup of $\mathrm{PSU}(2,2|4)$. The three-dimensional half-BPS defect breaks U_0 down to a subgroup $\mathrm{U} = \mathrm{SO}(1,2) \times \mathrm{SO}(3)_X \times \mathrm{SO}(3)_Y$; this is a subgroup of $\mathrm{OSp}(4|4)$. Here in ten-dimensional terms, the two factors $\mathrm{SO}(3)_X$ and $\mathrm{SO}(3)_Y$ of the unbroken R-symmetry subgroup act by rotations in the 456 and 789 subspaces, respectively. ($\mathrm{SO}(6)_R$ is broken to $\mathrm{SO}(3)_X \times \mathrm{SO}(3)_Y$ because the NS5-brane spans the 456 directions.) Under U_0, the global supersymmetries transform in a real representation $(\mathbf{2},\mathbf{1},\mathbf{4}) \oplus (\mathbf{1},\mathbf{2},\overline{\mathbf{4}})$. Under U this becomes $V_8 \otimes V_2$, where V_8 is a real eight-dimensional representation $(\mathbf{2},\mathbf{2},\mathbf{2})$ and V_2 is a two-dimensional real vector space with trivial action of U. An embedding of $\mathrm{OSp}(4|4)$ in $\mathrm{PSU}(2,2|4)$ can be fixed by specifying which linear combination of the two copies of V_8 is left unbroken by the defect; these unbroken supersymmetries are of the form $V_8 \otimes \varepsilon_0$, where ε_0 is a fixed vector in V_2. Up to an irrelevant scaling, the choice of ε_0 is parametrized by an angle that we will call ϑ. This angle in turn is determined by the string theory coupling parameter $\tau = i/g_{\mathrm{st}} + \theta/2\pi$, which in field theory terms is $\tau = \frac{4\pi i}{g^2_{\mathrm{YM}}} + \frac{\theta_{\mathrm{YM}}}{2\pi}$. The relation can be found in the brane description, as follows. Let ε_1 and ε_2 be the two ten-dimensional spinors that parametrize supersymmetry transformations in the underlying Type IIB theory. They transform in the **16** of the ten-dimensional Lorentz group Spin$(1,9)$, so

$$\Gamma_{012\ldots9}\varepsilon_i = \varepsilon_i, \quad i = 1, 2, \tag{2.2}$$

where $\Gamma_{012\ldots9}$ is the product of the SO$(1,9)$ gamma-matrices Γ_I, I=0,. . . , 9. The supersymmetry that is preserved by the D3-branes is defined by the condition

$$\varepsilon_2 = \Gamma_{0123}\varepsilon_1\,, \tag{2.3}$$

while the NS5-brane preserves supersymmetries that satisfy

$$\varepsilon_1 = -\Gamma_{012456}(\sin\vartheta\,\varepsilon_1 - \cos\vartheta\,\varepsilon_2)\,, \tag{2.4}$$

where the angle ϑ is related to the coupling parameter τ by

$$\vartheta = \arg(\tau). \tag{2.5}$$

(When $\cos\vartheta = 0$, (2.4) must be supplemented by an additional condition on ε_2.) Altogether the above conditions imply

$$(B_2\sin\vartheta + B_1\cos\vartheta)\varepsilon_1 = \varepsilon_1\,, \tag{2.6}$$

where $B_1 = \Gamma_{3456}$ and $B_2 = \Gamma_{3789}$ are operators that commute with the group U and thus act naturally in the two-dimensional space V_2. The solutions of this condition are of the form $\varepsilon_1 = \varepsilon \otimes \varepsilon_0$, where ε is any vector in V_8, and ε_0 is a fixed, ϑ-dependent vector in V_2. These are the generators of the unbroken supersymmetries.

It will be useful to introduce a new real parameter $\mathcal{K}$ and to rewrite (2.5) as

$$\tau = \mathcal{K}\cos\vartheta\, \mathrm{e}^{i\vartheta}. \tag{2.7}$$

The motivation for the notation is that $\mathcal{K}$ generalizes the level k of purely three-dimensional Chern-Simons theory. For physical values of the coupling τ, one has $\mathrm{Im}\,\tau > 0$; this places a constraint on the variables $\mathcal{K}$ and ϑ. In the twisted topological field theory, $\mathcal{K}$ will turn out to be what was called the canonical parameter Ψ in [14].

In general, let us write G_ℓ and G_r for the gauge groups to the left or right of the defect. From a purely field theory point of view, G_ℓ and G_r are completely arbitrary and moreover arbitrary hypermultiplets may be present at $x_3 = 0$ as long as $\mathrm{Re}\,\tau = \theta_{\mathrm{YM}}/2\pi$ vanishes.[3] However, as soon as $\theta_{\mathrm{YM}} \neq 0$, G_ℓ and G_r and the hypermultiplet representation are severely constrained; to maintain supersymmetry, the product $G_\ell \times G_r$ must be a maximal bosonic subgroup of a supergroup whose odd part defines the hypermultiplet representation and whose Lie algebra admits an invariant quadratic form with suitable properties. These rather mysterious conditions [23] have been given a more natural explanation in a closely related three-dimensional problem [24]; as explained in the introduction, our initial task is to generalize that explanation to four dimensions. We denote the Lie algebras of G_ℓ and G_r as $\mathfrak{g}_\ell$ and $\mathfrak{g}_r$, and denote the Killing forms on these Lie algebras as κ_ℓ and κ_r; precise normalizations will be specified later. We will loosely write $-\mathrm{tr}(\dots)$ for κ_ℓ or κ_r. We also need a form $\kappa = -\kappa_\ell + \kappa_r$ on the direct sum of the two Lie algebras. This will be denoted by $-\mathrm{Tr}(\dots)$. The gauge indices for $\mathfrak{g}_\ell \oplus \mathfrak{g}_r$ will be denoted by Latin letters m, n, p.

As already remarked, from a field theory point of view, as long as $\theta_{\mathrm{YM}} = 0$, the defect at $y = 0$ might support a system of N hypermultiplets transforming in an arbitrary real symplectic representation of $G_\ell \times G_r$. A real symplectic representation of $G_\ell \times G_r$ is a $4N$-dimensional real representation of $G_\ell \times G_r$, equipped with an action of SU(2) that commutes with $G_\ell \times G_r$. (In the context of the supersymmetric gauge theory, this SU(2) will become part of the R-symmetry group, as specified below.) This representation can be conveniently described as follows. Let $\mathcal{R}$ be a complex $2N$-dimensional symplectic representation of $G_\ell \times G_r$, with an invariant two-form ω_{IJ}. We take the sum of two copies of this representation,

[3] The gauge couplings $\tau_{\ell,r}$ and the angles $\vartheta_{\ell,r}$ can also be different at $y < 0$ and $y > 0$, as long as the canonical parameter $\mathcal{K}$ in eqn. (2.7) is the same [23]. For our purposes, this generalization is not important.

with an SU(2) group acting on the two-dimensional multiplicity space, and impose a $G_\ell \times G_r \times \mathrm{SU}(2)$-invariant reality condition. This gives the desired $4N$-dimensional real representation. We denote indices valued in $\mathcal{R}$ as I, J, K, we write T^I_{mJ} for the m^{th} generator of $G_\ell \times G_r$ acting in this representation, and we set $\tau_{mIJ} = T^S_{mI}\omega_{SJ}$, which is symmetric in I, J (and is related to the moment map for the action of $G_\ell \times G_r$ on the hypermultiplets). As remarked above, for $\theta_{\mathrm{YM}} \neq 0$, the representation $\mathcal{R}$ is highly constrained. It turns out that a supersymmetric action for our system with $\theta_{\mathrm{YM}} \neq 0$ can be constructed if and only if

$$\tau_{m(IJ}\tau_{K)Sn}\kappa^{mn} = 0. \tag{2.8}$$

This condition is equivalent [23] to the fermionic Jacobi indentity for a superalgebra $\mathfrak{sg}$, which has bosonic part $\mathfrak{g}_\ell \oplus \mathfrak{g}_r$, with fermionic generators transforming in the representation $\mathcal{R}$ and with $\kappa \oplus \omega$ being an invariant and nondegenerate graded-symmetric bilinear form on $\mathfrak{sg}$; we will sometimes write this form as $-\mathrm{Str}(\dots)$. Concretely, if we denote the fermionic generators of $\mathfrak{sg}$ as f_I, then the commutation relations of the superalgebra are

$$\begin{aligned} [T_m, T_n] &= f^s_{mn} T_s \,, \\ [T_m, f_I] &= T^K_{mI} f_K \,, \\ \{f_I, f_J\} &= \tau_{mIJ}\kappa^{mn} T_n . \end{aligned} \tag{2.9}$$

A short though admittedly mysterious calculation shows that the Jacobi identity for this algebra is precisely (2.8). As already remarked, the closest to an intuitive explanation of this result has been provided in [24], in a related three-dimensional problem. We will write SG for the supergroup with superalgebra $\mathfrak{sg}$.

In more detail, the $\mathcal{R}$-valued hypermultiplet that lives on the defect consists of scalar fields $Q^{I\dot{A}}$ and fermions λ^{IA}_α that transform in the representation $\mathcal{R}$ of the gauge group, and transform respectively as $(\mathbf{1}, \mathbf{1}, \mathbf{2})$ and $(\mathbf{2}, \mathbf{2}, \mathbf{1})$ under $\mathrm{U} = \mathrm{SO}(2,1) \times \mathrm{SO}(3)_X \times \mathrm{SO}(3)_Y$. (Here $A, B = 1, 2$ are indices for the double cover $\mathrm{SU}(2)_X$ of $\mathrm{SO}(3)_X$, and $\dot{A}, \dot{B}$ are similarly related to $\mathrm{SO}(3)_Y$.) They are subject to a reality condition, which e.g. for the scalars reads $\left(Q^I_{\dot{A}}\right)^\dagger = \epsilon^{\dot{A}\dot{B}}\omega_{IJ}Q^J_{\dot{B}}$. To describe the coupling of the bulk fields to the defect theory, it is convenient to rewrite the bulk super Yang-Mills fields in three-dimensional language. The scalars X^a and $Y^{\dot{a}}$, $a, \dot{a} = 1, \ldots, 3$, transform in the vector representations of $\mathrm{SO}(3)_X$ and $\mathrm{SO}(3)_Y$, respectively, and of course the gauge field A_i is $\mathrm{SO}(3)_X \times \mathrm{SO}(3)_Y$ singlet. The super Yang-Mills gaugino field Ψ transforms in the representation $(\mathbf{2}, \mathbf{1}, \mathbf{4}) \oplus (\mathbf{1}, \mathbf{2}, \overline{\mathbf{4}})$ of U_0. Under the subgroup U, it splits into two spinors $\Psi^{A\dot{B}}_{1\alpha}$ and $\Psi^{A\dot{B}}_{2\alpha}$, which transform in the representation $(\mathbf{2}, \mathbf{2}, \mathbf{2})$, like the supersymmetry generator $\varepsilon^{A\dot{B}}_{\alpha}$. More precisely, we define

$$\Psi = -\Psi_2 \otimes B_1\varepsilon_0 + \Psi_1 \otimes B_2\varepsilon_0. \tag{2.10}$$

With this definition, it is straightforward to decompose the supersymmetry transformations of the four-dimensional super Yang-Mills to find the transformations that correspond to $\varepsilon \otimes \varepsilon_0$. In particular, the bosons transform as

$$\begin{aligned}
\delta A_i &= \frac{1}{\sqrt{2}}\varepsilon_{\alpha A\dot{B}}\sigma^{\alpha}_{i\beta}\left(\Psi^{A\dot{B}\beta}_1 \sin\vartheta + \Psi^{A\dot{B}\beta}_2 \cos\vartheta\right) , \\
\delta X^a &= -\frac{i}{\sqrt{2}}\varepsilon^{A\alpha}_{\dot{B}}\Psi^{B\dot{B}}_{1\alpha}\sigma^a_{AB} , \\
\delta Y^{\dot{a}} &= \frac{i}{\sqrt{2}}\varepsilon^{\dot{A}\alpha}_{A}\Psi^{A\dot{B}}_{2\alpha}\sigma^{\dot{a}}_{\dot{A}\dot{B}}.
\end{aligned} \tag{2.11}$$

Here i, j, k and α, β are respectively vector and spinor indices of the three-dimensional Lorentz group $\mathrm{SO}(2,1)$, and σ_i are the Pauli matrices. See Appendix A for some details on our conventions.

The action of the theory has the following form:

$$I_{\text{electric}} = I_{\text{SYM}} - \frac{\theta_{\text{YM}}}{2\pi}\mathrm{CS}(A) + \mathcal{K} I_{\text{hyp}}. \tag{2.12}$$

The terms on the right are as follows. I_{SYM} is the usual action of the $\mathcal{N} = 4$ super Yang-Mills in the bulk. The term proportional to θ_{YM} reflects the bulk "topological" term of four-dimensional Yang-Mills theory

$$I_{\theta_{\rm YM}} = -\frac{\theta_{\rm YM}}{8\pi^2}\int_{x_3<0} \mathrm{tr}\, F\wedge F - \frac{\theta_{\rm YM}}{8\pi^2}\int_{x_3>0}\mathrm{tr}\, F\wedge F, \tag{2.13}$$

which we have split into two contributions at $y<0$ and $y>0$ because in the present context the gauge field (and even the gauge group) jumps discontinuously at $y=0$. Because of this discontinuity, even if we restrict ourselves to variations that are trivial at infinity, $I_{\theta_{\rm YM}}$ has a nontrivial variation supported on the locus W defined by $y=0$. This variation is the same as that of $(\theta_{\rm YM}/2\pi)\mathrm{CS}(A)$, where $\mathrm{CS}(A)$ is the Chern-Simons interaction of $G_\ell\times G_r$:

$$\mathrm{CS}(A) = \frac{1}{4\pi}\int_W \mathrm{Tr}\left(A\wedge dA + \frac{2}{3}A\wedge A\wedge A\right). \tag{2.14}$$

(Recall that the symbol Tr includes the contributions of both G_ℓ and G_r, but with opposite signs.) We lose some information when we replace $I_{\theta_{\rm YM}}$ by $(\theta_{\rm YM}/2\pi)\mathrm{CS}(A)$, since $I_{\theta_{\rm YM}}$ is gauge-invariant as a real number, but $\mathrm{CS}(A)$ is only gauge-invariant modulo an integer. However, the replacement of $I_{\theta_{\rm YM}}$ by $(\theta_{\rm YM}/2\pi)\mathrm{CS}(A)$ is a convenient shorthand. Finally, $I_{\rm hyp}$ is the part of the action that involves the hypermultiplets. More details concerning the action are given in the Appendix B.

We also need some facts about the boundary conditions and supersymmetry transformations in this theory. The bulk scalars $Y_{\dot a}$ obey a Dirichlet type boundary condition. In terms of $Y^m_{\dot A\dot B} = \sigma^{\dot a}_{\dot A\dot B}Y^{\dot a m}$, this boundary condition is

$$Y^m_{\dot A\dot B} = -\frac{1}{2\cos\vartheta}\tau^m_{IJ}Q^I_{\dot A}Q^J_{\dot B}. \tag{2.15}$$

In the brane picture, this boundary condition reflects the fact that the fields $Y^{\dot a}$ describe displacement of the D3-branes from the NS5-brane in the 789 directions, and so vanish at $y=0$ if the hypermultiplets vanish. Notice that, depending on whether m labels a generator of G_ℓ or G_r, the field $Y^m_{\dot A\dot B}$ is defined for $y\leq 0$ or for $y\geq 0$; but the boundary condition (2.15) is valid in both cases. A similar remark applies for other formulas below. Boundary conditions for other fields can be obtained from (2.15) by $\mathcal{N}=4$ supersymmetry transformations, or by ensuring the vanishing of boundary contributions in the variation of the action. For the gauge fields, the relevant part of the action is

$$\frac{1}{2g_{\mathrm{YM}}^2}\int \mathrm{d}^4x\,\mathrm{tr}\,F_{\mu\nu}^2-\frac{\theta_{\mathrm{YM}}}{8\pi^2}\int \mathrm{tr}\,F\wedge F+\mathcal{K}I_{\mathrm{hyp}}. \tag{2.16}$$

Taking the variation and reexpressing the coupling constant using (2.7), one gets on the boundary

$$\sin\vartheta\,F_{k3}^m-\frac{1}{2}\cos\vartheta\,\epsilon_{kij}F_{ij}^m=\frac{2\pi}{\cos\vartheta}J_k^m\,, \tag{2.17}$$

where $J_{mk}=\delta I_{\mathrm{hyp}}/\delta A_k^m$ is the hypermultiplet current, and gauge indices are raised and lowered by the form κ. There is a similar boundary condition for the X^a scalar which we shall not write explicitly here. By making supersymmetry transformations (2.11) of the equation (2.15), one can also find the boundary condition for the bulk fermions,

$$\sqrt{2}\Psi_{2\alpha A\dot{B}}^m=\frac{i}{\cos\vartheta}\tau_{IJ}^m\lambda_{\alpha A}^I Q_{\dot{B}}^J. \tag{2.18}$$

It was shown in [23] that this four-dimensional problem with a half-BPS defect is closely related to a purely three-dimensional Chern-Simons theory with three-dimensional $\mathcal{N}=4$ supersymmetry. A three-dimensional Chern-Simons theory with $\mathcal{N}=3$ supersymmetry exists with arbitrary gauge group and hypermultiplet representation, but with $\mathcal{N}=4$ supersymmetry, one needs precisely the constraints stated above: the gauge group G is the bosonic part of a supergroup SG, and the hypermultiplet representation corresponds to the odd part of the Lie algebra of SG. To compare the action of the four-dimensional model with the defect to the action of the purely three-dimensional model, we first decompose the hypermultiplet action in (2.12) as

$$I_{\mathrm{hyp}}=I_Q(A)+I'_{\mathrm{hyp}}, \tag{2.19}$$

where $I_Q(A)$ is the part of the hypermultiplet action that contains couplings to no bulk fields except A, and $I'_{\rm hyp}$ contains the couplings of hypermultipets to the bulk scalars and fermions. (For details, see Appendix B.) In these terms, the action of the purely three-dimensional theory is

$$-\mathcal{K}\left(\mathrm{CS}(A)+I_Q(A)\right) \tag{2.20}$$

while the contribution to the four-dimensional action at $y=0$ is

$$-\frac{\theta_{\rm YM}}{2\pi}\mathrm{CS}(A)-\mathcal{K}\left(I_Q(A)+I'_{\rm hyp}\right). \tag{2.21}$$

Thus, there are several differences: the defect part (2.21) of the four-dimensional action contains the extra couplings in $I'_{\rm hyp}$, and it has a different coefficient of the Chern-Simons term than that which appears in the purely three-dimensional action (2.20); also, in (2.20), A is a purely three-dimensional gauge field while in (2.21), it is the restriction of a four-dimensional gauge field to $y=0$. There also are differences in the supersymmetry transformations. The supersymmetry transformations in the purely three-dimensional Chern-Simons theory are schematically

$$\delta A\sim\varepsilon\lambda Q\,, \tag{2.22}$$

In the four-dimensional theory with the defect, the transformation for the gauge field in (2.11) is schematically

$$\delta A\sim\varepsilon(\Psi_1+\Psi_2). \tag{2.23}$$

Clearly, the two formulas (2.22) and (2.23) do not coincide. With the help of the boundary condition (2.18), we see that the Ψ_2 term in (2.23), when restricted to $y=0$, has the same form as the purely three-dimensional transformation law (2.22). The term involving Ψ_1 cannot be interpreted in that way; rather, before comparing the four-dimensional theory with a defect to a purely three-dimensional theory, one must redefine the connection A in a way that will eliminate the Ψ_1 term. In section 2.2.2, generalizing the ideas in [6] and in

[24], we will explain how to reconcile the different formulas.

2.2.2 Topological Twisting

After making a Wick rotation to Euclidean signature on $\mathbb{R}^4$, we want to select a scalar supercharge Q, obeying $Q^2 = 0$, in such a way that if we restrict to the cohomology of Q, we get a topological field theory. As part of the mechanism to achieve topological invariance, we require Q to be invariant under a twisted action of the rotations of $\mathbb{R}^4$, that is, under rotations combined with suitable R-symmetries. In Euclidean signature, the rotation and R-symmetry groups are the two factors of $\mathrm{U}_0^E = \mathrm{SO}(4) \times \mathrm{SO}(6)_R$, and the symmetries preserved by the defect are $\mathrm{U}^E = \mathrm{SO}(3) \times \mathrm{SO}(3)_X \times \mathrm{SO}(3)_Y$. The twisting relevant to our problem is the same procedure used in studying the geometric Langlands correspondence via gauge theory [14]. We pick a subgroup $\mathrm{SO}(4)_R \subset \mathrm{SO}(6)_R$, and define $\mathrm{SO}'(4) \subset U_0'$ to be a diagonal subgroup of $\mathrm{SO}(4) \times \mathrm{SO}(4)_R$, such that from the ten-dimensional point of view, $\mathrm{SO}'(4)$ acts by simultaneous rotations in the 0123 and 4567 directions. The space of ten-dimensional supersymmetries transforms as $(\mathbf{2}, \mathbf{1}, \mathbf{4}) \oplus (\mathbf{1}, \mathbf{2}, \overline{\mathbf{4}})$ under $\mathrm{U}_0^E = \mathrm{SO}(4) \times \mathrm{SO}(6)_R \cong \mathrm{SU}(2) \times \mathrm{SU}(2) \times \mathrm{SO}(6)_R$. Each summand has a one-dimensional $\mathrm{SO}'(4)$-invariant subspace; this follows from the fact that the representations $\mathbf{4}$ and $\overline{\mathbf{4}}$ of $\mathrm{SO}(6)_R$ both decompose as $(\mathbf{2}, \mathbf{1}) \oplus (\mathbf{1}, \mathbf{2})$ under $\mathrm{SO}'(4)$. The two invariant vectors coming from $(\mathbf{2}, \mathbf{1}, \mathbf{4})$ and $(\mathbf{1}, \mathbf{2}, \overline{\mathbf{4}})$ give two supersymmetry parameters ε_ℓ and ε_r with definite $\mathrm{SO}(4)$ chiralities. Although there is no natural way to normalize ε_ℓ, there is a natural way[4] to define ε_r in terms of ε_ℓ and one can take Q to be any linear combination $b\varepsilon_\ell + a\varepsilon_r$. We only care about Q up to scaling, so the relevant parameter is $t = a/b$.

In the bulk theory, we can make any choice of t, but in the presence of the half-BPS defect, we must choose a supercharge that is preserved by the defect. As in section 2.2.1, the space of supersymmetries decomposes under U^E as $V_8 \otimes V_2$, where V_8 transforms as $(\mathbf{2}, \mathbf{2}, \mathbf{2})$, and U^E acts trivially on V_2. (In Euclidean signature, the vector spaces V_8 and V_2 are not

[4] One sets $\varepsilon_r = \sum_{\mu=0}^{3} \Gamma_{4+\mu,\mu} \varepsilon_\ell / 4$, as in eqn. (3.8) of [14].

real.) The defect preserves supersymmetry generators of the form $\varepsilon \otimes \varepsilon_0$ with any $\varepsilon \in V_8$ and with a fixed $\varepsilon_0 \in V_2$. Invariance under $\mathrm{SO}'(4)$ restricts to a 1-dimensional subspace of V_8, as explained in the next paragraph. So up to scaling, only one linear combination of ε_ℓ and ε_r is preserved by the defect, and t is uniquely determined.

To find the scalar supersymmetry generator in three-dimensional notation, we note that at $y = 0$, $\mathrm{SO}'(4)$ can be naturally restricted to $\mathrm{SO}'(3)$, which is a diagonal subgroup of $\mathrm{SO}(3) \times \mathrm{SO}(3)_X \subset \mathrm{SO}(4) \times \mathrm{SO}(4)_R$. An $\mathrm{SO}'(3)$-invariant vector in V_8 must have the form

$$\varepsilon_{\mathrm{top}}^{\alpha A \dot{A}} = \epsilon^{\alpha A} v^{\dot{A}}, \tag{2.24}$$

where $\alpha, A, \dot{A} = 1, 2$ label bases of the three factors of $V_8 \sim \mathbf{2} \otimes \mathbf{2} \otimes \mathbf{2}$; $\epsilon^{\alpha A}$ is the antisymmetric symbol; and $v^{\dot{A}}$, which takes values in the $\mathbf{2}$ of $\mathrm{SO}(3)_Y$, is not constrained by $\mathrm{SO}'(3)$ invariance. However, $v^{\dot{A}}$ is determined up to scaling by $\mathrm{SO}'(4)$ invariance. In fact, for any particular $v^{\dot{A}}$, the supersymmetry parameter defined in eqn. (2.24) is invariant under a twisted rotation group that pairs the 0123 directions with 456$\mathbf{v}$, where $\mathbf{v}^{\dot{a}} \sim \overline{v}\sigma^{\dot{a}} v$ is some direction in the subspace 789 (here $\sigma^{\dot{a}}$ are the Pauli matrices). For $\mathrm{SO}'(4)$ invariance, we want to choose $v^{\dot{A}}$ such that $\mathbf{v}$ is the direction x_7. A simple way to do that is to look at the $\mathrm{U}(1)_F$ symmetry subgroup of $\mathrm{SO}(3)_Y$ that rotates the 89 plane and commutes with $\mathrm{SO}'(4)$; thus, $\mathrm{U}(1)_F$ rotates the last two components of $\vec{Y} = (Y_1, Y_2, Y_3)$. We normalize the generator F of $\mathrm{U}(1)_F$ so that the field $\sigma = \frac{Y_2 - iY_3}{\sqrt{2}}$ has charge 2. Then using a standard representation of the $\sigma^{\dot{a}}$, one has

$$Y^{\dot{A}}_{\dot{B}} \equiv Y^{\dot{a}} \sigma^{\dot{A}}_{\dot{a}\dot{B}} = i \begin{pmatrix} Y_1 & \sqrt{2}\,\sigma \\ \sqrt{2}\,\overline{\sigma} & -Y_1 \end{pmatrix}, \tag{2.25}$$

and in this basis, the generator F is

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.26}$$

SO$'$(4) invariance implies that the supersymmetry parameter ε has charge -1 under F (see eqn. (3.11) in [14]), so we can take

$$v^{\dot{A}} = 2^{1/4}\, \mathrm{e}^{-i\vartheta/2} \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{2.27}$$

The normalization factor here is to match the conventions of [14]. For future reference, we also define

$$u^{\dot{A}} = 2^{3/4}\, \mathrm{e}^{i\vartheta/2} \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \tag{2.28}$$

We also will need the relation between the parameter t and the angle ϑ. For that, we use equation (2.26) from [6] for the topological parameter $\varepsilon_\ell + t\varepsilon_r$. Comparing it to our eqn. (2.6), we find that

$$t = \mathrm{e}^{i(\pi-\vartheta)}. \tag{2.29}$$

In the twisted theory, the fields $\vec{X}$ and Y_1 join into a one-form $\phi = \sum_{\mu=0}^{3} \phi_\mu\, dx^\mu$, with components $\phi_i = X_{i+1}$, $i = 0, 1, 2$, and $\phi_3 = Y_1$. Q-invariance (or more precisely the condition $\{Q, \zeta\} = 0$ for any fermionic field ζ) gives a system of equations for A_μ and ϕ_μ. These equations, which have been extensively discussed in [6], take the form $\mathcal{V}^+ = \mathcal{V}^- = \mathcal{V}^0 = 0$, with

$$\begin{aligned} \mathcal{V}^+ &= (F - \phi\wedge\phi + t\,\mathrm{d}_A\phi)^+\,, \\ \mathcal{V}^- &= \left(F - \phi\wedge\phi - t^{-1}\mathrm{d}_A\phi\right)^-\,, \\ \mathcal{V}^0 &= D_\mu\phi^\mu. \end{aligned} \tag{2.30}$$

Here if Λ is a two-form, we denote its selfdual and anti-selfdual projections as Λ^+ and Λ^-, respectively.

2.2.3 Fields And Transformations

If a four-dimensional gauge theory with a defect is related to a purely three-dimensional theory on the defect, then what are the fields in the effective three-dimensional theory? The hypermultiplets supported at $y = 0$ give one obvious source of three-dimensional fields. So let us first discuss these fields from the standpoint of the twisted theory.

The hypermultiplet contains scalar fields $Q^{I\dot{A}}$ that transform as a doublet under $\mathrm{SU}(2)_Y$. In the twisted theory, $\mathrm{SU}(2)_Y$ is reduced to $\mathrm{U}(1)_F$, and accordingly we decompose the $Q^{I\dot{A}}$ in multiplets C^I and $\overline{C}^I$ with charges ± 1 under $\mathrm{U}(1)_F$. (These are upper and lower components in the basis used in (2.26).) The fermionic part of the hypermultiplet λ_α^{AI} has a more interesting decomposition in the twisted theory. Under $\mathrm{SO}'(3)$, both the spinor index α and the $\mathrm{SO}(3)_X$ index A carry spin 1/2, so λ_α^{AI} is a sum of pieces of spin 1 and spin 0. In other words, the fermionic part of the hypermultiplet decomposes into a vector ${\mathcal{A}_f}_i^I$ and a scalar B^I.

The supercharge Q generates the following transformations of these fields:

$$\begin{aligned}
\delta\mathcal{A}_f &= -\mathcal{D}_b C\,,\\
\delta C &= 0\,,\\
\delta\overline{C} &= B\,,\\
\delta B &= \frac{1}{2}[\{C,C\},\overline{C}].
\end{aligned} \tag{2.31}$$

Here for any field Φ, we define $\delta\Phi = [Q,\Phi\}$, where $[\ ,\ \}$ is a commutator or anticommutator for Φ bosonic or fermionic; also, $\mathcal{D}_b$ is the coveriant derivative with a connection $\mathcal{A}_b$ that we define momentarily.

The vector $\mathcal{A}_f$ will become the fermionic part of the $\mathfrak{sg}$-valued gauge field, which we will call $\mathcal{A}$. But where will we find $\mathcal{A}_b$, the bosonic part of $\mathcal{A}$? There is no candidate among the fields that are supported on the defect. Rather, $\mathcal{A}_b$ will be the restriction to the defect worldvolume of a linear combination of bulk fields:

$$\mathcal{A}_b = A + i(\sin\vartheta)\phi. \tag{2.32}$$

This formula defines both a $\mathfrak{g}_\ell$-valued part of $\mathcal{A}_b$ – obtained by restricting $A+i(\sin\vartheta)\phi$ from $y \leq 0$ to $y = 0$ – and a $\mathfrak{g}_r$-valued part – obtained by restricting $A + i(\sin\vartheta)\phi$ from $y \geq 0$ to $y = 0$. (Here $\mathfrak{g}_\ell$ and $\mathfrak{g}_r$ are the Lie algebras of G_ℓ and G_r.) The shift from A to $\mathcal{A}_b$ removes the unwanted term with Ψ_1 in the topological supersymmetry variation (2.23), so that – after restricting to $y = 0$ and using the boundary condition (2.18) – one gets

$$\delta\mathcal{A}_b = \{C, \mathcal{A}_f\}. \tag{2.33}$$

Obviously, since $\mathcal{A}_f$ is only defined at $y = 0$, $\delta\mathcal{A}_b$ can only be put in this form at $y = 0$.

The interpretation of the formulas (2.31) and (2.33) was explained in [24] (where they arose in a purely three-dimensional context): one can interpret C as the ghost field for a partial gauge-fixing of the supergroup SG down to its maximal bosonic subgroup G, and the supercharge Q as the BRST operator for this partial gauge-fixing. Since C has $\mathrm{U}(1)_F$ charge of 1, we should interpret $\mathrm{U}(1)_F$ as the ghost number. Once we interpret C as a ghost field, the transformation laws for $\mathcal{A}_b$ and $\mathcal{A}_f$ simply combine to say that acting on $\mathcal{A} = \mathcal{A}_b + \mathcal{A}_f$, Q generates the BRST transformation $\delta\mathcal{A} = -d_\mathcal{A}C$ with gauge parameter C. The gauge parameter C has opposite statistics from an ordinary gauge generator (it is a bosonic field but takes values in the odd part of the super Lie algebra $\mathfrak{sg}$); this is standard in BRST gauge-fixing of a gauge theory. In such BRST gauge-fixing, one often introduces BRST-trivial multiplets $(\overline{C}, B)$, where $\delta\overline{C} = B$ and δB is whatever it must be to close the algebra. In the most classical case, $\overline{C}$ is an antighost field, with $\mathrm{U}(1)_F$ charge -1, and B is called a Lautrup-Nakanishi auxiliary field. The multiplet $(\overline{C}, B)$ in (2.31) has precisely this form.

If one finds a gauge transformation in which the gauge parameter has reversed statistics to be confusing, one may wish to introduce a formal Grassman parameter η and write $\delta' = \eta\delta$, so that for any field Φ, $\delta'\Phi = [\eta Q, \Phi]$; ηQ is bosonic, so there is an ordinary commutator here. Then

$$\delta'\mathcal{A} = -\mathcal{D}(\eta C), \tag{2.34}$$

showing that the symmetry generated by ηQ transforms the supergroup connection $\mathcal{A}$ by a gauge transformation with the infinitesimal gauge parameter ηC, which has normal statistics.

2.2.4 The Action

After twisting, one can define the $\mathcal{N} = 4$ super Yang-Mills theory on an arbitrary[5] four-manifold M, with the defect supported on a three-dimensional oriented submanifold W. Generically, in this generality, one preserves only the unique supercharge Q.

What is the form of the Q-invariant action of this twisted theory? Any gauge-invariant expression $\{Q, \cdot\}$ is Q-invariant, of course – and also largely irrelevant as long as we calculate only Q-invariant observables, which are the natural observables in the twisted theory. But in addition, any gauge-invariant function of the complex connection $\mathcal{A}$ is Q-invariant, since Q acts on $\mathcal{A}$ as the generator of a gauge transformation. $\mathcal{A}$ is defined only on the oriented three-manifold W, and as we are expecting to make a topological field theory, the natural gauge-invariant function of $\mathcal{A}$ is the Chern-Simons function.

Given this and previous results (concerning the case that there is no defect [14], an analogous purely three-dimensional problem [24], and the case that the fields are nonzero only on one side of W [6]), it is natural to suspect that the action of the twisted theory on M may have the form

$$I = i\mathcal{K}\,\mathrm{CS}(\mathcal{A}) + \{Q, \dots\} = \frac{i\mathcal{K}}{4\pi}\int_W \mathrm{Str}\left(\mathcal{A}d\mathcal{A} + \frac{2}{3}\mathcal{A}^3\right) + \{Q, \dots\}\,, \tag{2.35}$$

[5] If M is not orientable, one must interpret ϕ not as an ordinary 1-form but as a 1-form twisted by the orientation bundle of M.

where if there is a formula of this type, then the coefficient of $\mathrm{CS}(\mathcal{A})$ must be precisely $i\mathcal{K}$, in view of what is already known about the one-sided case.

This is indeed so. Leaving some technical details for Appendix B, we simply make a few remarks here. In the absence of a defect, and assuming that M has no boundary, it was shown in [14] that the action of the twisted super Yang-Mills theory is Q-exact modulo a topological term:

$$I_{\mathrm{SYM}} + \frac{i\theta_{\mathrm{YM}}}{8\pi^2}\int_M \mathrm{tr}\,(F\wedge F) = \frac{i\mathcal{K}}{4\pi}\int_M \mathrm{tr}\,(F\wedge F) + \{Q,\dots\}\,. \tag{2.36}$$

(On the left, I_{SYM} is the part of the twisted super Yang-Mills action that is proportional to $1/g_{\mathrm{YM}}^2$; the part proportional to θ_{YM} is written out explicitly.) In [6], the case that M has a boundary W (and the D3-branes supported on M end on an NS5-brane wrapping T^*W) was analyzed. It was shown that (2.36) remains valid, except that the topological term $\int_M \mathrm{tr}\,F\wedge F$ must be replaced with a Chern-Simons function on $W=\partial M$, not of the real gauge field A but of its complexification $\mathcal{A}_b$. From the point of view of the present thesis, this case means that M intersects the NS5-brane worldvolume in a defect W, and there are gauge fields only on one side of W. Part of the derivation of eqn. (2.35) is simply to use the identity (2.36) on both M_ℓ and M_r, thinking of the integral of $\mathrm{tr}\,F\wedge F$ over M_ℓ or M_r as a Chern-Simons coupling on the boundary.

To get the full desired result, we must include also the hypermultiplets Q that are supported on W. The full action of the theory was described in formulas (2.12) and (2.19). In Euclidean signature it reads

$$I_{\mathrm{electric}} = I_{\mathrm{SYM}} + \frac{i\theta_{\mathrm{YM}}}{2\pi}\mathrm{CS}(A) + \mathcal{K}(I_Q(A) + I'_{\mathrm{hyp}}). \tag{2.37}$$

The identity (2.36) has a generalization that includes the boundary terms:

$$I_{\mathrm{electric}} = i\mathcal{K}\,\left(\mathrm{CS}_{G_r}(\mathcal{A}_b) - \mathrm{CS}_{G_\ell}(\mathcal{A}_b)\right) + \mathcal{K}I_Q(\mathcal{A}_b) + \{Q,\dots\}. \tag{2.38}$$

Since the first three terms are defined purely on the three-manifold W, we can now invoke the result of [24]: this part of the action is $i\mathcal{K}\mathrm{CS}(\mathcal{A}) + \{Q, \dots\}$, where now $\mathrm{CS}_{SG}(\mathcal{A})$ is the Chern-Simons function for the full supergroup gauge field $\mathcal{A} = \mathcal{A}_b + \mathcal{A}_f$, and the Q-exact terms describe partial gauge-fixing from SG to G. This confirms the validity of eqn. (2.35).

We conclude by clarifying the meaning of the supergroup Chern-Simons function $\mathrm{CS}(\mathcal{A})$. With $\mathcal{A} = \mathcal{A}_b + \mathcal{A}_f$, we have

$$\mathrm{CS}(\mathcal{A}) = \mathrm{CS}(\mathcal{A}_b) + \frac{1}{4\pi}\int_W \mathrm{Str}\, \mathcal{A}_f d_{\mathcal{A}_b}\mathcal{A}_f. \tag{2.39}$$

The term involving $\mathcal{A}_f$ is the integral over W of a function with manifest gauge symmetry under $G_\ell \times G_r$ (and even its complexification). It is not affected by the usual subtleties of the Chern-Simons function involving gauge transformations that are not homotopic to the identity. The reason for this is that the supergroup SG is contractible to its maximal bosonic subgroup G; the topology is entirely contained in G. Similarly, with $\mathcal{A}_b = A + i(\sin\vartheta)\phi$, we can expand the complex Chern-Simons function,

$$\mathrm{CS}(\mathcal{A}_b) = \mathrm{CS}(A) + \frac{1}{4\pi}\int_W \mathrm{Tr}\,\left(i(\sin\vartheta)\phi\wedge F - (\sin^2\vartheta)\phi\wedge \mathrm{d}_A\phi - i(\sin^3\vartheta)\phi\wedge\phi\wedge\phi\right), \tag{2.40}$$

and the topological subtleties affect only the first term $\mathrm{CS}(A)$. Here, as in eqns. (2.14) and (2.13), to resolve the topological subtleties and put the action in a form that is well-defined for generic $\mathcal{K}$, we should replace $\mathrm{CS}(A)$ with the corresponding volume integral $(1/4\pi)\int_M \mathrm{tr} F\wedge F$. There is no need for such a substitution in any of the other terms, since they are all integrals over W of gauge-invariant functions. All this reflects the fact that a complex Lie group is contractible to a maximal compact subgroup, so the topological subtlety in $\mathrm{CS}(\mathcal{A}_b)$ is entirely contained in $\mathrm{CS}(A)$.

It is convenient to simply write the action as $i\mathcal{K}\,\mathrm{CS}(\mathcal{A}) + \{Q, \dots\}$, as we have done in eqn. (2.35), rather than always explicitly replacing the term $\mathrm{CS}(A)$ in this action with a bulk integral.

2.2.5 Analytic Continuation

To get the formula (2.33) along W, we have had to replace A by $A+i(\sin\vartheta)\phi$, with the result that the bosonic part of $\mathcal{A}$ is complex-valued. This is related to an essential subtlety [12, 13] in the relation of the four-dimensional theory with a defect to a Chern-Simons theory supported purely on the defect. In general, four-dimensional $\mathcal{N}=4$ super Yang-Mills theory on a four-manifold M, with a half-BPS defect of the type analyzed here on a three-manifold $W \subset M$, is not equivalent to standard Chern-Simons theory on W with gauge supergroup SG, but to an analytic continuation of this theory. The basic idea of this analytic continuation is that localization on the space of solutions of the equations (2.30) defines an integration cycle in the complexified path integral of the Chern-Simons theory. This localization is justified using the fact that the Q-exact terms in (2.38) can be scaled up without affecting Q-invariant observables, so the path integral can be evaluated just on the locus where those terms vanish. The condition for these terms to vanish is the localization equations (2.30), which define the integration cycle. (Thus, the integration cycle is characterized by the fact that $\mathcal{A}_b$ is the restriction to $y=0$ of fields A, ϕ which obey the localization equations and have prescribed behavior for $y \to \pm\infty$.)

For generic W and M, the integration cycle derived from $\mathcal{N}=4$ super Yang-Mills theory differs from the standard one of three-dimensional Chern-Simons theory. For the important case that $W=\mathbb{R}^3$, there is essentially only one possible integration cycle and therefore the two constructions are equivalent. Thus, after including Wilson loop operators (as we do in section 2.3), the four-dimensional construction can be used to study the usual knot invariants associated to three-dimensional Chern-Simons theory.

Unfortunately, it turns out that for supergroups all the observables which can be defined using only closed Wilson loops in $\mathbb{R}^3$ reduce to observables of an ordinary bosonic Chern-Simons theory. This is explained in section 2.3 of the present thesis, and in section 6 of [1]. To find novel observables, one needs to do something more complicated. All of the options seem to introduce some complications in the relation to four dimensions. For example, one can replace $\mathbb{R}^3$ by S^3 and define observables that appear to be genuinely new by considering the path integral with insertion of a Wilson loop in a typical representation (see section 2.3.2.2). But the compactness of S^3 means that one encounters infrared questions in comparing to four dimensions. Because of such complications, our results for supergroup Chern-Simons theory are less complete then in the case of a bosonic Lie group.

A feature of the localization that is special to supergroups is that $\mathcal{A}_b$ is the boundary value of a four-dimensional field (which in the localization procedure is constrained by the equations (2.30)), but $\mathcal{A}_f$ is purely three-dimensional. The reason that this happens is essentially that the topology of the supergroup SG is contained entirely in its maximal bosonic subgroup G. Being fermionic, $\mathcal{A}_f$ is by nature infinitesimal; the Berezin integral for fermions is an algebraic operation (a Gaussian integral in the case of Chern-Simons theory of a supergroup) with no room for choosing different integration cycles. By contrast, in the integration over the bosonic fields, it is possible to pick different integration cycles and the relation to four-dimensional $\mathcal{N}=4$ super Yang-Mills theory does give a very particular one.

One important qualitative difference between purely three-dimensional Chern-Simons theory and what one gets by extension to four dimensions is as follows. In the three-dimensional theory, the "level" k must be an integer, but in the analytically continued version given by the relation to four-dimensional $\mathcal{N}=4$ super Yang-Mills, k is generalized to a complex parameter $\mathcal{K}$. Part of the mechanism for this is that although the Chern-Simons function $CS(A)$ is only gauge-invariant modulo 1, in the four-dimensional context it can be replaced by a volume integral $\int_M \mathrm{Tr}\, F\wedge F$, which is entirely gauge-invariant, so there is no need to quantize the parameter.

2.2.6 Relation Among Parameters

At first sight, eqn. (2.35) seems to tell us that the relation between the parameter $\mathcal{K}$ in four dimensions and the usual parameter k of Chern-Simons theory, which appears in the purely three-dimensional action

$$i\frac{k}{4\pi}\int_W \mathrm{Str}\,\left(A\wedge \mathrm{d}A+\frac{2}{3}A\wedge A\wedge A\right), \tag{2.41}$$

would be $\mathcal{K}=k$. However, for the purely one-sided case, the relation, according to [6], is really[6]

$$\mathcal{K}=k+h\ \mathrm{sign}\,(k). \tag{2.42}$$

An improved explanation of this is as follows.

The purely three-dimensional Chern-Simons theory for a compact gauge group G involves a path integral over the space of real connections A. This is an oscillatory integral and in particular, at one-loop level, in expanding around a classical solution, one has to perform an oscillatory Gaussian integral.[7] After diagonalizing the matrix that governs the fluctuations, the oscillatory Gaussian integral is a product of one-dimensional integrals

$$\int_{-\infty}^{\infty}\frac{\mathrm{d}x}{\sqrt{\pi}}\exp(i\lambda x^2)=\frac{\exp(i(\pi/4)\mathrm{sign}\lambda)}{|\lambda|}, \tag{2.43}$$

where the phase comes from rotating the contour by $x=\exp(i(\pi/4)\mathrm{sign}\lambda)x'$ to get a real convergent Gaussian integral for x'. In Chern-Simons gauge theory, the product of these phase factors over all modes of the gauge field and the ghosts gives (after suitable regularization) a factor $\exp(i\pi\eta/4)$, where η is the Atiyah-Patodi-Singer η-invariant. This factor has the effect of shifting the effective value of k in many observables to $k+h\ \mathrm{sign}\,k$, where h is the dual Coxeter number of G (this formula is often written as $k\to k+h$, with k assumed

[6] A careful reader will ask what precisely we mean by k in the following formula. In defining k precisely, we will assume that it is positive; if it is negative, one makes the same definitions after reversing orientations. One precise definition is that k is the level of a two-dimensional current algebra theory that is related to the given Chern-Simons theory in three dimensions. (The level is defined as the coefficient of a c-number term appearing in the product of two currents.) Another precise definition is that, for integer k, the space of physical states of the Chern-Simons theory on a Riemann surface Σ is $H^0(\mathcal{M},\mathcal{L}^k)$, where $\mathcal{M}$ is the moduli space of holomorphic G-bundles over Σ and $\mathcal{L}$ generates the Picard group of $\mathcal{M}$. (For simplicity, in this statement, we assume G to be simply-connected.)

[7] The following is explained more fully on pp. 358-9 of [4], where however a nonstandard normalization is used for η. See also [87].

to be positive).

One can think of the shift $k \to k+h\ \mathrm{sign}\, k$ as arising in a Wick rotation in field space from the standard integration cycle of Chern-Simons theory (real A) to an integration cycle on which the integral is convergent rather than oscillatory. But this is precisely the integration cycle that is used in the four-dimensional description (see [12, 13]). Accordingly, in the four-dimensional description, there is no one-loop shift in the effective value of $\mathcal{K}$ and instead the shift must be absorbed in the relation between parameters in the four- and three-dimensional descriptions by $\mathcal{K} = k + h\ \mathrm{sign}\, k$.

Up to a point, the same logic applies in our two-sided problem. The four-dimensional path integral has no oscillatory phases and hence no one-loop shift in the effective value of the Chern-Simons coupling. So any such shift that would arise in a purely three-dimensional description must be absorbed in the relationship between $\mathcal{K}$ and a three-dimensional parameter k. We are therefore tempted to guess that the relationship between $\mathcal{K}$ and the parameter k of a purely three-dimensional Chern-Simons theory of the supergroup SG is

$$\mathcal{K} = k + h_{\mathfrak{sg}}\, \mathrm{sign}\, k, \tag{2.44}$$

where $h_{\mathfrak{sg}}$ is the dual Coxeter number of the supergroup. The trouble with this formula is that it assumes that the effective Chern-Simons level for a supergroup has the same one-loop renormalization as for a bosonic group. The validity of this claim is unclear for reasons explored in Appendix E of [1]. (In brief, the fact that the invariant quadratic form on the bosonic part of the Lie superalgebra $\mathfrak{sg}$ is typically not positive-definite means it is not clear what should be meant by $\mathrm{sign}\, k$, and also means that a simple imitation of the standard one-loop computation of bosonic Chern-Simons theory does not give the obvious shift $k \to k + h_{\mathfrak{sg}}\, \mathrm{sign}\, k$.) We actually do not know the proper treatment of purely three-dimensional Chern-Simons theory of a supergroup. In this chapter, we concentrate on the four-dimensional description, in which the bosonic part of the path integral is convergent, not oscillatory, and accordingly there is no one-loop shift in the effective value of $\mathcal{K}$. Thus we should just think of $\mathcal{K}$ as the effective parameter of the Chern-Simons theory.

Let us go back to the purely bosonic or one-sided case. For G simple and simply-laced, Chern-Simons theory is usually parametrized in terms of

$$q = \exp(2\pi i/(k + h \operatorname{sign} k)) = \exp(2\pi i/\mathcal{K}). \tag{2.45}$$

If G is not simply-laced, it is convenient to take $q = \exp(2\pi i/n_{\mathfrak{g}}\mathcal{K})$, where $n_{\mathfrak{g}}$ is the ratio of length squared of long and short roots of $\mathfrak{g}$. Including the factor of $1/n_{\mathfrak{g}}$ in the exponent ensures that q is the instanton-counting parameter in a magnetic dual description. Similarly, for a supergroup SG, we naturally parametrize the theory in terms of

$$q = \exp(2\pi i/n_{\mathfrak{sg}}\mathcal{K}), \tag{2.46}$$

where $n_{\mathfrak{sg}}$ is the ratio of length squared of the longest and shortest roots of a maximal bosonic subgroup of SG, computed using an invariant bilinear form on $\mathfrak{sg}$ (for the supergroups we study in this thesis, $n_{\mathfrak{sg}}$ can be 1, 2, or 4). To write this formula in terms of a purely three-dimensional parameter k, we would have to commit ourselves to a precise definition of such a parameter. Each of the definitions given for bosonic groups in footnote 6 may generalize to supergroups, but in neither case is the proper generalization immediately clear.

2.3 Observables In The Electric Theory

The most important observables in ordinary Chern-Simons gauge theory are Wilson line operators, labeled by representations of the gauge group. To understand their analogs in supergroup Chern-Simons theory, we need to know something about representations of supergroups. The theory of Lie supergroups has some distinctive features, compared to the ordinary Lie group case, and these special features have implications for Chern-Simons theory and its line observables. Accordingly, we devote section 2.3.1 to a brief review of Lie supergroups and superalgebras. Then in section 2.3.2, we discuss the peculiarities of line observables in three-dimensional supergroup Chern-Simons theory. In sections 2.3.3 and 2.3.4, we return to the four-dimensional construction, and explain, in fairly close parallel with [6], how line operators of supergroup Chern-Simons theory are realized as line or surface operators in $\mathcal{N} = 4$ super Yang-Mills theory. Finally, in section 2.3.5 we summarize some unclear points.

In the four-dimensional construction, in addition to the line and surface operators considered here, it is possible to construct Q-invariant local operators. They are described in Appendix D.

2.3.1 A Brief Review Of Lie Superalgebras

We begin with the basics of Lie superalgebras, Lie supergroups, and their representations. For a much more complete exposition see e.g. [37, 38, 39].

A Lie superalgebra decomposes into its bosonic and fermionic parts, $\mathfrak{sg} = \mathfrak{g}_{\bar{0}} + \mathfrak{g}_{\bar{1}}$. We will assume that $\mathfrak{g}_{\bar{0}}$ is a reductive Lie algebra (the sum of a semi-simple Lie algebra and an abelian one). Moreover, to define the supergroup gauge theory action, we need the superalgebra $\mathfrak{sg}$ to possess a non-degenerate invariant bilinear form. (This also determines a superinvariant volume form on the SG supergroup manifold.) Finite-dimensional Lie superalgebras with these properties are direct sums of some basic examples. These include the unitary and the orthosymplectic superalgebras, as well as a one-parameter family of deformations of $\mathfrak{osp}(4|2)$, and two exceptional superalgebras, as specified in Table 2.1. For the unitary Lie superalgebras, one can also restrict to the supertraceless matrices $\mathfrak{su}(m|n)$, and for $m = n$ further factor by the one-dimensional center down to $\mathfrak{psu}(n|n)$. In what follows, by a Lie superalgebra we mean a superalgebra from this list.[8]

[8]We avoid here using the term "simple superalgebra," since, e.g., $\mathfrak{u}(1|1)$ is not simple (it is solvable), but is perfectly suitable for supergroup Chern-Simons theory. Let us mention that Lie superalgebras with the properties we have required which in addition are simple are called basic classical superalgebras.

Though we use real notation in denoting superalgebras, for instance in writing $\mathfrak{u}(m|n)$ and not $\mathfrak{gl}(m|n)$, we never really are interested in choosing a real form on the full superalgebra. One reason for this is that we will actually be studying analytically-continued versions of supergroup Chern-Simons theories. If one considers all possible integration cycles, then the real form is irrelevant. More fundamentally, as we have already explained in section 2.2.5, to define a path integral for supergroup Chern-Simons theory, one needs to pick a real integration cycle for the bosonic fields, but one does not need anything like this for the fermions. Correspondingly, we might need a real structure on $\mathfrak{g}_{\bar{0}}$ (and this will generally be the compact form) but not on the full supergroup or the superalgebra. So for our purposes, a three-dimensional Chern-Simons theory is naturally associated to a so-called *cs*-supergroup, which is a complex Lie supergroup together with a choice of real form for its bosonic subgroup.

If we choose the compact form of a maximal bosonic subgroup of a supergroup SG, then one can calculate the volume of SG with respect to its superinvariant measure. This volume has the following significance in Chern-Simons theory. The starting point in Chern-Simons perturbation theory on a compact three-manifold is to expand around the trivial flat connection; in doing so one has to divide by the volume of the gauge group. But this volume is actually[9] 0 for any Lie supergroup whose maximal bosonic subgroup is compact, with the exception of $\mathrm{B}(0,n) = \mathrm{OSp}(1|2n)$. This fact is certainly one reason that one cannot expect to develop supergroup Chern-Simons theory by naively imitating the bosonic theory.

Another difference between ordinary groups and supergroups is that in the supergroup case, we have to distinguish between irreducible representations and indecomposable ones. A representation R of $\mathfrak{sg}$ is called irreducible if it does not contain a non-trivial $\mathfrak{sg}$-invariant

[9] A quick proof is as follows. Let SG be a Lie supergroup whose maximal bosonic subgroup is compact (this assumption ensures that there are no infrared subtleties in defining and computing the volume of SG). Suppose that there is a fermionic generator $\mathcal{C}$ of $\mathfrak{sg}$ with the property that $\{\mathcal{C},\mathcal{C}\} = 0$. Such a $\mathcal{C}$ exists for every Lie supergroup except $\mathrm{OSp}(1|2n)$. We view $\mathcal{C}$ as generating a supergroup F of dimension $0|1$, which we consider to act on SG on (say) the left. This gives a fibration $SG \to SG/\mathrm{F}$ with fibers F. The volume of SG can be computed by first integrating over the fibers of the fibration. But the volume of the fibers is 0, so (given the existence of $\mathcal{C}$) the volume of SG is 0. The volume of the fibers is 0 because, since $\{\mathcal{C},\mathcal{C}\} = 0$, there are local coordinates in which the fibers are parametrized by an odd variable ψ and $\mathcal{C} = \partial/\partial\psi$. $\mathcal{C}$-invariance of the volume then implies that the measure for integration over ψ is invariant under adding a constant to ψ; the volume of the fiber is therefore $\int \mathrm{d}\psi \cdot 1 = 0$.

superalgebra	bosonic part	fermionic part	type
$\mathfrak{u}(m\vert n)$	$\mathfrak{u}(m) \oplus \mathfrak{u}(n)$	$(m, \overline{n}) \oplus (\overline{m}, n)$	I
$B(m,n) \simeq \mathfrak{osp}(2m+1\vert 2n)$	$\mathfrak{so}(2m+1) \oplus \mathfrak{sp}(2n)$	$(2m+1, 2n)$	II
$C(n+1) \simeq \mathfrak{osp}(2\vert 2n)$	$\mathfrak{u}(1) \oplus \mathfrak{sp}(2n)$	$(1, 2n) \oplus (\overline{1}, 2n)$	I
$D(m,n) \simeq \mathfrak{osp}(2m\vert 2n), m > 1$	$\mathfrak{so}(2m) \oplus \mathfrak{sp}(2n)$	$(2m, 2n)$	II
$D(2,1;\alpha), \alpha \in \mathbb{C} \setminus \{0, -1\}$	$\mathfrak{su}(2) \oplus \mathfrak{su}(2) \oplus \mathfrak{su}(2)$	$(2,2,2)$	II
$G(3)$	$\mathfrak{su}(2) \oplus \mathfrak{g}_2$	$(2,7)$	II
$F(4)$	$\mathfrak{su}(2) \oplus \mathfrak{so}(7)$	$(2,8)$	II

Table 2.1: Lie superalgebras suitable for the supergroup Chern-Simons theory. (We do not list explicitly the subquotients of the unitary superalgebra, which are mentioned in the text.)

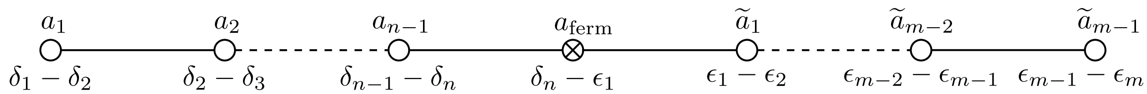

Figure 2.2: Dynkin diagram for the $\mathfrak{su}(m|n)$ superalgebra. The subscripts are expressions for the roots in terms of the orthogonal basis $\delta_\bullet$, $\epsilon_\bullet$. The superscripts represent the Dynkin labels of a weight. The middle root denoted by a cross is fermionic.

subspace R_0, and it is called indecomposable if it cannot be decomposed as $R_0 \oplus R_1$ where R_0 and R_1 are non-trivial $\mathfrak{sg}$-invariant subspaces. In a reducible representation, the representation matrices are block triangular $\begin{pmatrix} * & * \\ 0 & * \end{pmatrix}$, while in a decomposable representation, they are block diagonal. For ordinary reductive Lie algebras, these notions coincide (if the matrices are block triangular, there is a basis in which they are block diagonal), but for Lie superalgebras as defined above, they do not coincide, with the sole exception of $\mathrm{B}(0,n)$. It is not a coincidence that $\mathrm{B}(0,n)$ is an exception to both statements; a standard way to prove that a reducible representation of a compact Lie group is also decomposable involves averaging over the group, and this averaging only makes sense because the volume is nonzero. For $\mathrm{B}(0,n)$, taken with the compact form of its maximal bosonic subgroup, the same proof works, since the volume is not zero. A physicist's explanation of the "bosonic" behavior of $\mathrm{B}(0,n)$ might be that, as we argue later, the Chern-Simons theory with this gauge supergroup is dual to an ordinary bosonic Chern-Simons theory with the gauge group $\mathrm{SO}(2n+1)$. This forces $\mathrm{B}(0,n)$ to behave somewhat like an ordinary bosonic group.

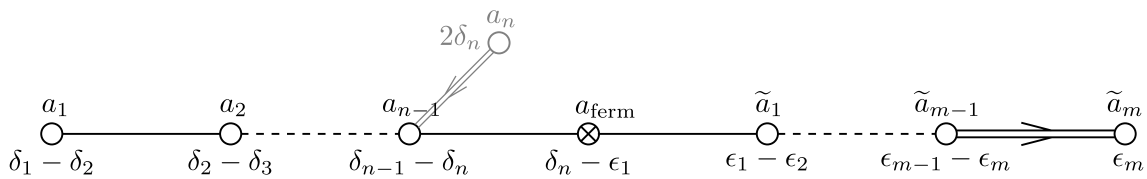

Figure 2.3: Dynkin diagram for the $\mathfrak{osp}(2m+1|2n)$ superalgebra, $m \geq 1$. The subscripts are expressions for the roots in terms of the orthogonal basis $\delta_\bullet$, $\epsilon_\bullet$. The superscripts represent the Dynkin labels of a weight. The arrows point in the direction of a shorter root. The middle root denoted by a cross is fermionic. Roots of the $\mathfrak{sp}(2n)$ and $\mathfrak{so}(2m+1)$ subalgebras are on the left and on the right of the fermionic root. The site shown in grey and labeled a_n is the long simple root of the $\mathfrak{sp}(2n)$ subalgebra, which does not belong to the set of simple roots of the superalgebra.

The structure theory for a simple Lie superalgebra $\mathfrak{sg}$ can be described similarly to the case of an ordinary Lie algebra. One starts by picking a Cartan subalgebra $\mathfrak{t}$, which for our superalgebras is just a Cartan subalgebra of the bosonic part. Then one decomposes $\mathfrak{sg}$ into root subspaces. These subspaces lie either in $\mathfrak{g}_{\bar{0}}$ or in $\mathfrak{g}_{\bar{1}}$, and the roots are correspondingly called bosonic or fermionic. Then one makes a choice of positive roots, or, equivalently, of a Borel subalgebra $\mathfrak{b} \supset \mathfrak{t}$. Unlike in the bosonic case, different Borel subalgebras can be non-isomorphic. However, there is a distinguished Borel subalgebra – the one which contains precisely one simple fermionic root. This is the choice that we shall make. For each choice of Borel subalgebra, one can construct a Dynkin diagram. The distinguished Dynkin diagrams for the unitary and the odd orthosymplectic superalgebras are shown in fig. 2.2 and fig. 2.3.

The fermionic $\mathbb{Z}_2$-grading of a Lie superalgebra can be lifted (in a way that is canonical up to conjugacy) to a $\mathbb{Z}$-grading, which can be defined as follows. The subalgebra of degree zero is generated by the Cartan subalgebra together with the bosonic simple roots of the superalgebra. The fermionic simple root of the distinguished Dynkin diagram is assigned degree one. The grading for the other elements of the superalgebra is then determined by the commutation relations. This $\mathbb{Z}$-grading is defined by a generator of $\mathfrak{g}_{\bar{0}}$.

For example, for the unitary superalgebra this element can be taken to be the central generator of $\mathfrak{u}(n)$. The degree zero subalgebra in this case is just the bosonic subalgebra, while the fermions decompose as $\mathfrak{g}_{\bar{1}} \simeq \mathfrak{g}_{-1} \oplus \mathfrak{g}_1$. Another example would be the odd orthosymplectic superalgebra $\mathfrak{osp}(2m+1|2n)$, for which the situation is slightly different. There exists a simple root of the bosonic subalgebra, which is not a simple root of the superalgebra, but rather is a multiple of a fermionic simple root, and therefore will not have degree zero. It is shown in grey in fig.2.3. The degree zero subalgebra consists of a semisimple Lie algebra $\mathfrak{sl}(n) \oplus \mathfrak{o}(2m+1)$ with the Dynkin diagram obtained from fig.2.3 by deleting the fermionic node, plus a central $\mathfrak{u}(1)$. This central element is the generator of the $\mathbb{Z}$-grading. The bosonic subalgebra decomposes into degrees ± 2 and 0, while the fermions again live in degrees ± 1.

More generally, for any superalgebra, the distinguished $\mathbb{Z}$-grading takes values from -1 to 1 or from -2 to 2, and the superalgebras are classified accordingly as type I or type II. In a type I superalgebra, the bosonic subalgebra lies completely in degree 0. The representation of $\mathfrak{g}_{\bar{0}}$ on the fermionic subalgebra $\mathfrak{g}_{\bar{1}}$ is reducible, and $\mathfrak{g}_{\bar{1}}$ decomposes into subspaces of degree -1 and 1. The unitary superalgebra is an example of a type I superalgebra. For the type II superalgebras, the action of $\mathfrak{g}_{\bar{0}}$ on $\mathfrak{g}_{\bar{1}}$ is irreducible. Under the $\mathbb{Z}$-grading, the bosonic subalgebra decomposes as $\mathfrak{g}_{\bar{0}} \simeq \mathfrak{g}_{-2} \oplus \mathfrak{g}_0 \oplus \mathfrak{g}_2$, and the fermions decompose as $\mathfrak{g}_{\bar{1}} \simeq \mathfrak{g}_{-1} \oplus \mathfrak{g}_1$. The $\mathfrak{osp}(2m+1|2n)$ superalgebra is an example of the type II case. The type of a superalgebra is important for representation theory, and we indicate it in Table 2.1.

We need to introduce some further notation. Let $\Delta_{\bar{0}}^+$ and $\Delta_{\bar{1}}^+$ be the sets of positive bosonic and fermionic roots, respectively, and let $\overline{\Delta}_{\bar{1}}^+$ be the set of positive fermionic roots with zero length. The length is defined using the invariant quadratic form on $\mathfrak{sg}$, which we normalize in a standard way so that the length squared of the longest root is 2. A root of zero length is called isotropic; isotropic roots are always fermionic. It is convenient to expand the roots and the weights in terms of a vector basis $\delta_\bullet$ and $\epsilon_\bullet$, orthogonal with respect to the invariant scalar product, with $\langle \delta_i, \delta_i \rangle = -\langle \epsilon_j, \epsilon_j \rangle > 0$. For example, the positive roots for the unitary superalgebra $\mathfrak{su}(m|n)$ are

$$\begin{aligned} \Delta_{\bar{0}}^+ &= \left\{ \delta_i - \delta_{i+p}, \epsilon_j - \epsilon_{j+p} \right\}, \quad i = 1 \ldots n, \quad j = 1 \ldots m, \quad p > 0, \\ \Delta_{\bar{1}}^+ &= \overline{\Delta}_{\bar{1}}^+ = \left\{ \delta_i - \epsilon_j \right\}. \end{aligned} \tag{2.47}$$

	$\mathfrak{su}(m\|n),\ n,m \geq 0$	$\mathfrak{osp}(m\|2n),\ m \geq 0,\ n \geq 1$	$\mathfrak{so}(n)$
h	$n-m$	$n-m/2+1$	$n-2$
(s)dim	$(n-m)^2-1$	$(2n-m)(2n-m+1)/2$	$n(n-1)/2$

Table 2.2: (Super)dimensions and dual Coxeter numbers.

The quadratic Casimir operator is defined using the invariant form on $\mathfrak{sg}$ (normalized in the standard way). In this thesis, by the dual Coxeter number h we mean one-half of the quadratic Casimir in the adjoint representation.[10] For future reference, in Table 2.2 we collect the superdimension (the difference between the dimension of $\mathfrak{g}_{\bar{0}}$ and that of $\mathfrak{g}_{\bar{1}}$) and the dual Coxeter number for the unitary and orthosymplectic superalgebras.

For a given Borel subalgebra, one defines the bosonic and fermionic Weyl vectors as

$$\rho_{\bar{0}} = \frac{1}{2} \sum_{\alpha \in \Delta^+_{\bar{0}}} \alpha \,, \quad \rho_{\bar{1}} = \frac{1}{2} \sum_{\alpha \in \Delta^+_{\bar{1}}} \alpha \,, \tag{2.48}$$

and the superalgebra Weyl vector as $\rho = \rho_{\bar{0}} - \rho_{\bar{1}}$. The Weyl group of a superalgebra, by definition, is generated by reflections with respect to the even (that is, bosonic) roots.

2.3.1.1 Representations

The finite-dimensional irreducible representations are labeled by their highest weights. The weights can be parametrized in terms of Dynkin labels. For a weight Λ, the Dynkin label associated to a simple root α_i is defined as $a_i = \dfrac{2\langle \Lambda, \alpha_i \rangle}{\langle \alpha_i, \alpha_i \rangle}$, if the length of the root α_i is non-zero, and $a_i = \langle \Lambda, \alpha_i \rangle$, if the length of the root is zero.

For a type I superalgebra, the Dynkin diagram coincides with the diagram for the semisimple part of the bosonic subalgebra $\mathfrak{g}_{\bar{0}}$, if one deletes the fermionic root. The finite-dimensional superalgebra representations are labeled by the same data as the representations of the bosonic subalgebra. For example, for the dominant weights of $\mathfrak{su}(m|n)$ all the Dynkin labels, except a_{ferm}, must be non-negative integers. The fermionic label can be an arbitrary

[10] This definition is different from the definition of [43].

complex number, if we consider representations of the superalgebra, or an arbitrary integer, if we want the representation to be integrable to a representation of the compact form of the bosonic subgroup.

For a type II superalgebra, if one deletes the fermionic node of the Dynkin diagram (and the links connecting to it), one gets a diagram for the semisimple part of the degree-zero subalgebra $\mathfrak{g}_0 \subset \mathfrak{g}_{\bar{0}}$. The long simple root of the bosonic subalgebra $\mathfrak{g}_{\bar{0}}$ is "hidden" behind the fermionic simple root, and is no longer a simple root of the superalgebra. This is illustrated in fig. 2.3 for the $\mathrm{B}(m,n)$ case. For us it will be convenient to parametrize the dominant weights in terms of the Dynkin labels of the bosonic subalgebra, so, for type II, instead of a_{ferm} we will use the Dynkin label with respect to the long simple root of $\mathfrak{g}_{\bar{0}}$. For example, for $\mathrm{B}(m,n)$ this label is[11] a_n, as shown on the figure, and the weights will be parametrized by $(a_1,\dots,a_n,\widetilde{a}_1,\dots,\widetilde{a}_m)$. Clearly, in this case for the superalgebra representation to be finite-dimensional, it is necessary for these Dynkin labels to be non-negative integers. It turns out that there is an additional supplementary condition. For example, for $\mathrm{B}(m,n)$ this condition says that if $a_n < m$, then only the first a_n of the labels $(\widetilde{a}_1,\dots,\widetilde{a}_m)$ can be non-zero. For the other type II superalgebras the supplementary conditions can be found e.g. in Table 2 of [37]. The finite-dimensional irreducible representations are in one-to-one correspondence with integral dominant weights that satisfy these extra conditions.

For a generic highest weight, the irreducible superalgebra representation can be constructed rather explicitly. For a type I superalgebra, one takes an arbitrary representation R^0_Λ of the bosonic part $\mathfrak{g}_{\bar{0}}$, with highest weight Λ. A representation of the superalgebra can be induced from R^0_Λ by setting the raising fermionic generators $\mathfrak{g}_1$ to act trivially on R^0_Λ, and the lowering fermionic generators $\mathfrak{g}_{-1}$ to act freely. The resulting representation in the vector space

$$\mathcal{H}_\Lambda = \wedge^\bullet \mathfrak{g}_{-1} \times \mathrm{R}^0_\Lambda \tag{2.49}$$

is called the Kac module. For a generic highest weight, this gives the desired finite-dimensional

[11] Our notation here is slightly unconventional: notation a_n is usually used for what we call a_{ferm}.

irreducible representation. For a type II superalgebra, the representation can be similarly induced from a representation of the degree-zero subalgebra $\mathfrak{g}_0 \subset \mathfrak{g}_{\bar{0}}$, but the answer is slightly more complicated than (2.49), since the fermionic creation or annihilation operators do not anticommute among themselves.

The Kac module, which one gets in this way, is irreducible only for a sufficiently generic highest weight. In this case, the highest weight Λ and the representation are called typical. Typical representations share many properties of representations of bosonic Lie algebras, e.g., a reducible representation with a typical highest weight is always decomposable, and there exist simple analogs of the classical Weyl character formula for their characters and supercharacters.

However, if Λ satisfies the equation

$$\langle \Lambda + \rho, \alpha \rangle = 0\,, \tag{2.50}$$

for some isotropic root $\alpha \in \overline{\Delta}_{\bar{1}}^{+}$, then the Kac module acquires a null vector. The irreducible representation then is a quotient of the Kac module by a maximal submodule. Such weights and representations are called atypical. Let $\Delta(\Lambda)$ be the subset of $\overline{\Delta}_{\bar{1}}^{+}$ for which (2.50) is satisfied. The number of roots in $\Delta(\Lambda)$ is called the degree of atypicality of the weight and of the corresponding representation.

The maximal possible degree of atypicality of a dominant weight is called the defect of the superalgebra. For $\mathfrak{u}(m|n)$, for a dominant Λ all the roots in $\Delta(\Lambda)$ are mutually orthogonal, and therefore the maximal number of such isotropic roots is $\min(m,n)$. In the corresponding IIB brane configuration, this is the number of D3-branes which can be recombined and removed from the NS5-brane. (This symmetry breaking process is analyzed in section 6 of [1].)

A Kac-Wakimoto conjecture [43, 44] states that the superdimension of a finite-dimensional irreducible representation is non-zero if and only if it has maximal atypicality. (For ordinary

Lie algebras and for $\mathrm{B}(0,n)$, the maximal atypicality is zero, and all representations should be considered as both typical and maximally atypical.)

2.3.1.2 The Casimir Operators And The Atypical Blocks

The Casimir operators, by definition, are invariant polynomials in the generators of $\mathfrak{sg}$; in a fancier language, they generate the center $\mathfrak{Z}$ of the universal enveloping algebra $\mathcal{U}(\mathfrak{sg})$. We introduce some facts about them, which will be useful for the discussion of Wilson lines.

There is a well-known formula for the value of the quadratic Casimir in a representation with highest weight Λ,

$$c_2(\Lambda) = \langle \Lambda + \rho, \Lambda + \rho \rangle - \langle \rho, \rho \rangle \,, \tag{2.51}$$

which continues to hold in the superalgebra case. A remote analog of this formula for the higher Casimirs is known as the Harish-Chandra isomorphism (see e.g. [45]), which we now briefly review.

By the Poincaré-Birkhoff-Witt theorem, a Casimir element $c \in \mathfrak{Z}$ can be brought to the normal-ordered form, where in the Chevalley basis, schematically, $c = \sum (E^-)^{k_1} H^{k_2} (E^+)^{k_1}$. When acting on the highest weight vector of some representation, the only non-zero contribution comes from the purely Cartan part. This gives a homomorphism $\widehat{\xi} : \mathfrak{Z} \to S(\mathfrak{t})$, where $S(\mathfrak{t})$ are the symmetric polynomials in elements of $\mathfrak{t}$, and the value of the Casimir in a representation R_Λ with highest weight Λ is evaluated as $c(\Lambda) = (\widehat{\xi}(c))[\Lambda]$. Here the square brackets mean the evaluation of a polynomial from $S(\mathfrak{t})$ on an element of $\mathfrak{t}^*$. By making appropriate shifts of the Lie algebra generators in the polynomial $\widehat{\xi}(c)$, one can define a different polynomial $\xi(c)$, such that the formula becomes

$$c(\Lambda) = (\xi(c))[\Lambda + \rho] \,. \tag{2.52}$$

This is a minor technical redefinition, which will be convenient.

For ordinary Lie algebras, the Harish-Chandra theorem states that the image of the homomorphism ξ consists of the Weyl-invariant polynomials $S_W(\mathfrak{t}) \subset S(\mathfrak{t})$, and ξ is actually an isomorphism of commutative algebras $\mathfrak{Z} \simeq S_W(\mathfrak{t})$. To summarize, the Casimirs can be represented by Weyl-invariant Cartan polynomials, and their values in a representation R_Λ are obtained by evaluating these polynomials on $\Lambda + \rho$.

In the superalgebra case, the Harish-Chandra isomorphism [46] identifies $\mathfrak{Z}$ with a subalgebra $S^0_W(\mathfrak{t}) \subset S_W(\mathfrak{t})$, consisting of Weyl-invariant polynomials p with the following invariance property,

$$p[\Lambda + \rho + \mathrm{x}\alpha] = p[\Lambda + \rho] \tag{2.53}$$

for any $\mathrm{x} \in \mathbb{C}$ and $\alpha \in \Delta(\Lambda)$.

For a highest weight representation R_Λ, the corresponding set of eigenvalues of the Casimir operators (equivalently, a homomorphism from $\mathfrak{Z}$ into the complex numbers) is called the central character, denoted χ_Λ. The Harish-Chandra isomorphism allows one to describe the sets of weights which share the same central character. If the weight is typical, then the other weights with the same central character can be obtained by the shifted Weyl action $\Lambda \to w(\Lambda + \rho) - \rho$. The orbit of this transformation can contain no more than one dominant weight; therefore, two different typical finite-dimensional representations have different central characters. This is no longer the case for the atypical weights. Given an atypical dominant weight Λ, we can shift it by a linear combination of elements of $\Delta(\Lambda)$ to obtain new dominant weights with the same central character. More generally, we can apply a sequence of shifts and Weyl transformations without changing the central character. All the representations that are obtained in this way will have the same degree of atypicality, and they will share the same eigenvalues of the Casimir operators. The set of atypical finite-dimensional representations which have a common central character is called an atypical block. In this chapter, we are interested mostly in the irreducible representations, and, somewhat imprecisely,[12] by an atypical block we will usually mean a set of irreducible

[12]This phrasing is imprecise because it does not take account the difference between reducibility of a representation and decomposability.

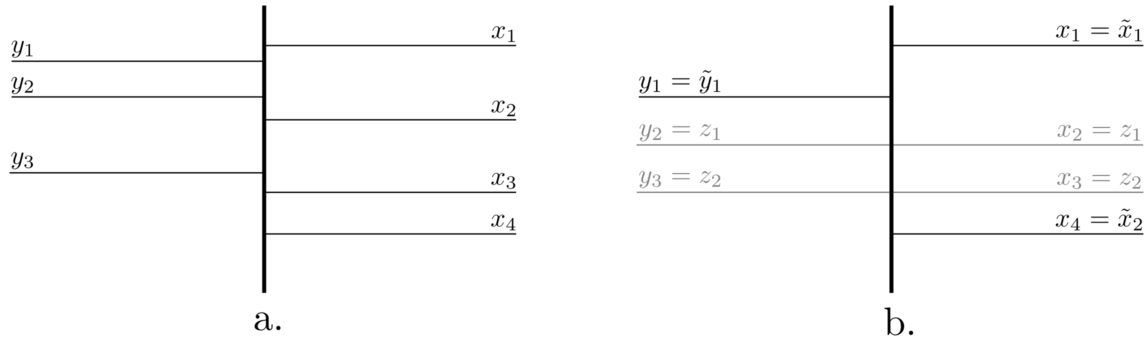

Figure 2.4: Examples of dominant weights for $\mathfrak{u}(3|4)$. a. A typical weight. b. A weight of atypicality two, which is part of a block of atypical weights. The block is labeled by $\widetilde{x}_1$, $\widetilde{x}_2$, and $\widetilde{y}_1$, which correspond to a dominant weight of $\mathfrak{u}(1|2)$. The weights that make up this block are parametrized by z_1 and z_2, which can be thought of as labels of a maximally atypical weight of $\mathfrak{u}(2|2)$.

representations (or, equivalently, dominant weights) with the same central character.

As an example, let us describe the atypical blocks for the $\mathfrak{u}(m|n)$ superalgebra. It is convenient to parametrize a weight Λ as

$$\Lambda + \rho = \sum_{i=1}^{n} x_i \delta_i - \sum_{j=1}^{m} y_j \epsilon_{m+1-j}. \tag{2.54}$$

For Λ to be dominant, the two sequences $\{x_i\}$ and $\{y_j\}$ must be strictly increasing, and satisfy an appropriate integrality condition. A dominant weight can be represented graphically, as shown in fig. (2.4a). This is essentially the weight diagram of [47]. The picture shows an obvious analogy between a dominant weight of $\mathfrak{u}(m|n)$ and a vacuum of a brane system; we will develop this analogy in section 2.4.4.4. This description also confirms that dominant weights of $\mathfrak{u}(m|n)$ correspond to dominant weights of the purely bosonic subalgebra $\mathfrak{u}(m) \times \mathfrak{u}(n)$. In this correspondence, of the two central generators of $\mathfrak{u}(m) \times \mathfrak{u}(n)$, one linear combination corresponds to the fermionic root a_{ferm} of $\mathfrak{su}(m|n)$ and the other to the center of $\mathfrak{u}(m|n)$.

For atypicality r, the set $\Delta(\Lambda)$ consists of r isotropic roots $\delta_{i_l} - \epsilon_{j_l}$, $l = 1 \ldots r$, which are mutually orthogonal, that is, each basis vector $\delta_\bullet$ or $\epsilon_\bullet$ can appear no more than once.[13] The

[13]Suppose that in $\{\delta_1 - \epsilon_1, \delta_1 - \epsilon_2\}$ the vector δ_1 appears more than once. Then, by taking a difference,

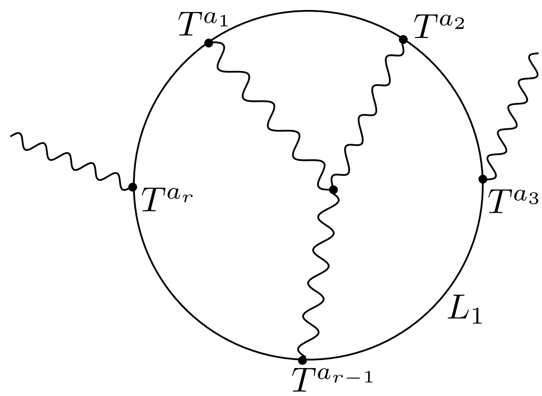

Figure 2.5: A diagram contributing to the expectation value of a link. A component L_1 of the link is shown. The propagators running from T^{a_3} and T^{a_r} connect to the other components of the link.

atypicality condition (2.50) then says that r of the x-labels are "aligned" with (equal to) the y-labels. Let these labels be $x_{i_l} = y_{m+1-j_l} \equiv z_l$, $l = 1 \ldots r$, and the others be $\widetilde{x}_1, \ldots, \widetilde{x}_{n-r}$, $\widetilde{y}_1, \ldots, \widetilde{y}_{m-r}$. Then the atypical blocks of atypicality r are labeled by the numbers $\widetilde{x}_\bullet$ and $\widetilde{y}_\bullet$, which can be thought of as labels for a dominant weight of $\mathfrak{u}(m-r|n-r)$, and the weights inside the same atypical block are parametrized by a sequence $z_\bullet$, which can be thought of as a dominant maximally atypical weight of $\mathfrak{u}(r|r)$. An example is shown in fig. (2.4b). An atypical block is described by the following statement: the category of finite-dimensional representations (not necessarily irreducible) from the same atypical block of atypicality r is equivalent to the category of maximally atypical representations of $\mathfrak{u}(r|r)$ from the atypical block, which contains the trivial representation [47]. A similar statement holds for the orthosymplectic superalgebras; the role of $\mathfrak{u}(r|r)$ is played by $\mathfrak{osp}(2r|2r)$, $\mathfrak{osp}(2r+1|2r)$ or $\mathfrak{osp}(2r+2|2r)$.

2.3.2 Line Observables In Three Dimensions

We begin the discussion of line operators by considering purely three-dimensional Chern-Simons theory of a supergroup. As it is explained in Appendix E of [1], there are some puzzles about this theory, but they do not really affect the following remarks. In any event, these

we would get that $\langle \Lambda + \rho, \epsilon_1 - \epsilon_2 \rangle = 0$, which contradicts the assumption that Λ is dominant.

remarks are applicable to the analytically-continued theory as defined in four dimensions, to which we return in section 2.3.3.

Consider a supergroup Chern-Simons theory on $\mathbb{R}^3$ with a link L which consists of p closed Wilson loops $L_1, \ldots, L_p$, labeled by representations $R_{\Lambda_1}, \ldots, R_{\Lambda_p}$ of the supergroup. Let us look at the perturbative expansion of this observable. On $\mathbb{R}^3$, the trivial flat connection is the only one, up to gauge transformation, and perturbation theory is an expansion about it. The trivial flat connection is invariant under constant gauge transformations, but as the generators of constant gauge transformations on $\mathbb{R}^3$ are not normalizable, we do not need to divide by the volume of the group of constant gauge transformations. This is just as well, as this volume is typically zero in the case of a supergroup.

A portion of a diagram that contributes to the expectation value is shown in fig. 2.5. We focus on a single component of the link, say L_1, and sketch only the gluon lines that are attached to this component. Let r be the number of such lines. Then in evaluating this diagram, we have to evaluate a trace

$$\mathrm{Str}_{R_{\Lambda_1}}(T^{a_1} \ldots T^{a_r})\, d_{a_1 \ldots a_r} \,, \tag{2.55}$$

where T^{a_i} are bosonic or fermionic generators of the superalgebra, and everything except the group factor for the component L_1 is hidden inside the invariant tensor $d_{a_1 \ldots a_r}$ (whose construction depends on the rest of the diagram). By gauge invariance, the operator $T^{a_1} \ldots T^{a_r}\, d_{a_1 \ldots a_r}$ is a Casimir operator $c_{L_1,p} \in \mathfrak{Z}$, acting in the representation R_{Λ_1}. The Casimir can be replaced simply by a number, and what then remains of the group factor is the supertrace of the identity. So this contribution to the expectation value can be written as $c_{L_1,p}(\Lambda_1)\,\mathrm{sdim} R_{\Lambda_1}$. From this we learn two things. First of all, up to a constant factor, the expectation value for the link L will not change if we replace any of the representations R_{Λ_i} by a representation with the same values of the Casimirs. That is, for an irreducible atypical representation, the expectation value depends only on the atypical block, and not on the specific representative. Second, if the supertrace over any of the representations R_{Λ_i} vanishes, the expectation value of the link in $\mathbb{R}^3$ vanishes. Recall from the previous section that the superdimension can be non-zero only for a representation of maximal atypicality. We conclude that in $\mathbb{R}^3$ for a non-trivial link observable, the components of the link should be labeled by maximally atypical blocks or else the expectation value will be zero. For example, for the unitary supergroup $\mathrm{U}(m|n)$, maximally atypical blocks correspond to irreducible representations of the ordinary Lie group $\mathrm{U}(|n-m|)$.

In section 6 of [1], it is argued that for knots on $\mathbb{R}^3$ (and more generally on any space with enough non-compact directions) one can give expectation values to the superghost fields C, without changing the expectation value of a product of loop operators. For instance, in this way, the $\mathrm{U}(m|n)$ theory can be Higgsed down to $\mathrm{U}(|n-m|)$. Therefore, on $\mathbb{R}^3$ the supergroup theory does not give any new Wilson loop observables, beyond those that are familiar from $\mathrm{U}(|n-m|)$. The symmetry breaking procedure shows that the expectation value of a Wilson loop labeled by a maximally atypical representation of $\mathrm{U}(m|n)$ is equal to the expectation value of the corresponding $\mathrm{U}(|n-m|)$ Wilson loop.

Yet it is known from the point of view of quantum supergroups [30, 31] that knot invariants can be associated to arbitrary highest weights of $\mathrm{U}(m|n)$, not necessarily maximally atypical. It is believed that generically these invariants are new, that is, they cannot be trivially reduced to invariants constructed using bosonic Lie groups. To make such a construction from the gauge theory point of view, one needs to remove the supertrace which in the case of a representation that is not maximally atypical multiplies the expectation value by $\mathrm{sdim} R_\Lambda = 0$. One strategy is to consider a Wilson operator supported not on a compact knot but on a non-compact 1-manifold that is asymptotic at infinity to a straight line in $\mathbb{R}^3$ (but which may be knotted in the interior). The invariant of such a non-compact knot would be an operator acting on the representation R_Λ, rather than a number. This approach may give invariants associated to arbitrary supergroup representations. This strategy seems plausible to us because it appears to make sense at least in perturbation theory, but we will not investigate it here.

The Higgsing argument suggests another approach that turns out to work well for typical representations. (For representations that are neither typical not maximally atypical, the only strategy we see is the one mentioned in the last paragraph.) In this approach, we look at the loop observables on a manifold with less then three non-compact directions. We will focus on the case of S^3. Again perturbation theory is an expansion around the trivial flat connection. But now, unlike the $\mathbb{R}^3$ case, the generators of constant gauge transformations are normalizable and we do have to divide by the volume of the gauge group. As was mentioned in our superalgebra review, this volume is zero for any supergroup except $\mathrm{OSp}(1|2n)$. Therefore, for almost all supergroups the partition function $Z(S^3)$ on S^3 is divergent,

$$Z(S^3) = \infty. \tag{2.56}$$

If we try to include a Wilson loop in a non-maximally atypical representation, we get an indeterminacy $0 \cdot \infty$.

There is a natural way to resolve this indeterminacy in the case of typical representations, but it involves an additional tool. In three-dimensional Chern-Simons theory with a compact simple gauge group G, Wilson line operators and line operators defined by a monodromy singularity are equivalent [4, 48, 49]. The proof involves using the Borel-Weil-Bott (BWB) theorem to "de-quantize" an irreducible representation of G, interpreting it as arising by quantizing some auxiliary space (the flag manifold of G), in what we will call BWB quantum mechanics. To resolve the indeterminacy that was just noted, we need the analog of this for supergroups.

2.3.2.1 BWB Quantum Mechanics

We first recall this story in the case of an ordinary bosonic group. Let G be a compact reductive Lie group, $T \subset G$ a maximal torus, and $\lambda \in \mathfrak{t}^*$ an integral weight. Assume in addition, that λ is regular, that is $\langle\lambda,\alpha\rangle \neq 0$ for any root α, or equivalently the coadjoint orbit of λ is G/T. (If this is not so, there is a similar story to what we will explain with G/T replaced by G/L, where L is a subgroup of G that contains T. L is called a Levi subgroup of G. Its Lie algebra is obtained by adjoining to $\mathfrak{t}$ the roots α that obey $\langle\lambda,\alpha\rangle = 0$.) One can consider a quantum mechanics in phase space G/T with the Kirillov-Kostant symplectic form corresponding to λ. The functional integral for this theory can be written as

$$\int \mathrm{D}h \exp\left(-i\int \lambda\big(h^{-1}\partial_s h\big)\mathrm{d}s\right) , \tag{2.57}$$

where we integrate over maps of a line (or a circle) to G/T. The action here is defined using an arbitrary lift of the map $h(s)$ valued in G/T into a map valued in G. The functional integral is independent of this lift, as long as the weight is integral.

Let V_λ be a one-dimensional T-module, where T acts with weight λ. The prequantization line bundle over the phase space is defined as $\mathcal{L}_\lambda \simeq G\times_T V_\lambda$; thus, it is a line bundle associated to the principal T-bundle $G\to G/T$. To define an actual quantization, one needs to make a choice of polarization. For that we need a complex structure. To that end, pick a Borel subgroup $B\supset T$ in the complexified gauge group $G_\mathbb{C}$. The complex Kähler manifold $\mathcal{M}\simeq G_\mathbb{C}/B$ is isomorphic, as a real manifold, to our phase space, and this gives it a complex structure. The prequantum line bundle is likewise endowed with a holomorphic structure, $\mathcal{L}_\lambda\simeq G_\mathbb{C}\times_B V_\lambda$.

An accurate description of geometric quantization also involves the metaplectic correction. Instead of being just sections of the prequantum line bundle, the wave-functions are usually taken to be half-forms valued in this line bundle. For example, this is the source of the 1/2 shift in the Bohr-Sommerfeld quantization condition. The metaplectic correction is important for showing the independence of the Hilbert space on the choice of polarization. In a holomorphic polarization, the bundle of half-densities is a square root of the canonical line bundle K. For the flag manifolds that we consider, K is simply $\mathcal{L}_{-2\rho}$, where ρ is the Weyl vector for the chosen Borel subgroup. So our wave-functions will live, roughly speaking, in $\mathcal{L}_\lambda \otimes K^{1/2} \simeq \mathcal{L}_{\lambda-\rho}$.

The precise characterization of the Hilbert space is given by the Borel-Weil-Bott theorem. Let $w \in \mathcal{W}$ be the element of the Weyl group that conjugates λ into a weight that is dominant with respect to the chosen B. Since λ was assumed to be regular, the weight

$$\Lambda = w(\lambda) - \rho\,, \tag{2.58}$$

is also dominant. The BWB theorem states that the cohomology $H^\bullet(\mathcal{M}, \mathcal{L}_{\lambda-\rho})$ is non-vanishing precisely in one degree $\ell(w)$, which is the length of the element w in terms of the simple reflections. The group $G_\mathbb{C}$ acts naturally on the cohomology, and $H^{\ell(w)}(\mathcal{M}, \mathcal{L}_{\lambda-\rho}) \simeq R_\Lambda$. This can be taken naturally as the Hilbert space $\mathcal{H}$ of our system. Clearly, R_Λ depends only on λ, and not on the choice of the Borel subgroup, that is, the polarization. If B is taken such that λ is dominant, then this is the usual Kähler quantization, since $H^0(\mathcal{M}, \mathcal{L}_{\lambda-\rho})$ is the space of holomorphic sections.

The fact that the resulting Hilbert space $H^{\ell(w)}(\mathcal{M}, \mathcal{L}_{\lambda-\rho})$ is independent of the choice of complex structure (or equivalently the choice of B) has a direct explanation. On a Kähler manifold, the bundle $\Omega^{0,\bullet}(\mathcal{M}) \otimes K^{1/2}$ is isomorphic to the Dirac bundle $S \simeq S^+ \oplus S^-$, where S^+ and S^- are spinors of positive or negative chirality. The Dirac operator is $\not{D} = \overline{\partial} + \overline{\partial}^*$, and the cohomology of $\overline{\partial}$ acting in $\Omega^{0,\bullet}(\mathcal{M}) \otimes K^{1/2}$ is isomorphic to the space of zero-modes of the Dirac operator, by a standard Hodge argument. Therefore, the Hilbert space that we defined is simply the kernel of the Dirac operator acting on $S \otimes V_\lambda$.

For the application to Wilson operators, we need to decide if the particle running in the loop is bosonic or fermionic. If the Hilbert space lies in the ℓ-th cohomology group, it is natural to define the operator $(-1)^F$ that distinguishes bosons from fermions as $(-1)^F = (-1)^\ell$. In the Dirac operator terminology, the particle is a boson or a fermion depending on whether the zero-modes lie in S^+ or in S^-. In particular, the amplitude of propagation of the particle along a loop (with zero Hamiltonian) is naturally defined as the index of the Dirac operator ind $\not{D} = \pm \dim R_\Lambda$, to account for the -1 factor for a fermion loop. Note that an elementary Weyl reflection of λ along a simple root reverses the orientation of $\mathcal{M}$, and therefore exchanges S^+ with S^- and exchanges bosons and fermions.

In what follows, we will always work in the Borel subalgebra in which λ is dominant, and therefore $\Lambda = \lambda - \rho$.

Now we return to the supergroup case. We would like to write the same functional integral (2.57), with matrices replaced by supermatrices. A technical detail is as follows. In the bosonic case, the integral goes over G/T, where G is the real compact form of the group. In the supergroup case, we choose the compact real form of the bosonic subgroup $G_{\overline{0}}$, since this is the only choice that will lead to finite-dimensional representations of SG. The compact form of $G_{\overline{0}}$ may not extend to a real form of SG (for $\mathrm{OSp}(n|2m)$ it does not), so one has to develop the theory without assuming a real form of SG. Similarly to what we have said in the beginning of section 2.3.1 for the Chern-Simons case, to make sense of the BWB path integral, a real form is needed only in the bosonic directions. The path integral of the supergroup BWB model goes over a sub-supermanifold in $SG_{\mathbb{C}}/T_{\mathbb{C}}$ whose reduced manifold is the bosonic phase space $G_{\overline{0}}/T$. For instance, in our analysis shortly of a type I supergroup, $h_0 \in G_{\overline{0}}/T$, and θ and $\widetilde{\theta}$ are independent variables valued in $\mathfrak{g}_1$ and $\mathfrak{g}_{-1}$, with no reality condition.

We claim that a simple supergroup version of the BWB model produces an irreducible representation of SG as the Hilbert space. To exclude zero-modes, we assume that λ is regular, so that $\langle \lambda, \alpha \rangle \neq 0$ for any $\alpha \in \Delta^+$, bosonic or fermionic. It means in particular that the weight $\Lambda = \lambda - \rho$ is typical. In this case, a direct analog of the BWB theorem exists [50], and the same logic as for the bosonic group leads to the conclusion that the Hilbert space of the system is indeed the irreducible representation R_Λ.

For a type I superalgebra, this statement can be heuristically explained as follows. Take a parametrization of the supergroup element as $h = \mathrm{e}^{\theta} h_0 \mathrm{e}^{\widetilde{\theta}}$, with h_0 an element of the bosonic subgroup, and $\widetilde{\theta}$ and θ belonging to $\mathfrak{g}_{-1}$ and $\mathfrak{g}_1$, respectively. The action of the theory is

$$-\int \lambda(h^{-1}\mathrm{d}h) = \int \mathrm{Str}(\lambda^\circ h^{-1}\mathrm{d}h)\,, \tag{2.59}$$

where $\lambda^\circ = \kappa^{mn}\lambda_m T_n$ is the dual of λ, defined using the superinvariant bilinear form.[14] Using the fact that $\{\mathfrak{g}_{-1},\mathfrak{g}_{-1}\} = 0$, and the fact that the invariant bilinear form is even, one can rewrite this as

$$\int \mathrm{Str}\left(\lambda^\circ h^{-1}\mathrm{d}h\right) = \int \mathrm{Str}\left(\lambda^\circ h_0^{-1}\mathrm{d}h_0\right) + \int \mathrm{Str}\left(h_0[\widetilde{\theta},\lambda^\circ]h_0^{-1}\mathrm{d}\theta\right)\,, \tag{2.60}$$

If λ is regular, the commutation with it in $[\widetilde{\theta},\lambda^\circ]$ simply multiplies the different components of the fermion $\widetilde{\theta}$ by non-zero numbers. Then we can set $\theta' = h_0[\widetilde{\theta},\lambda^\circ]h_0^{-1}$, with θ' a new fermionic variable. The resulting theory is a BWB quantum mechanics for the bosonic field h_0, together with the free fermions θ' and θ. The Hilbert space is a tensor product (2.49), as expected for a typical representation of a type I superalgebra.[15]

What if λ is atypical, so that there exist isotropic fermionic roots α for which $\langle\lambda,\alpha\rangle = 0$? The usual BWB action (2.57) is degenerate, as it is independent of some modes of θ and $\widetilde{\theta}$. This is analogous to the problem that one has in the bosonic case if λ is non-regular, and one can proceed in a similar way. We replace SG/T with SG/L, where L is a subgroup of G whose Lie algebra includes the roots with $\langle\lambda,\alpha\rangle = 0$. ($L$ is a superanalog of a Levi subgroup of a simple bosonic Lie group.) Then we quantize SG/L instead of SG/T. This seems to give a well-defined quantum mechanics, but we will not try to analyze it. The BWB theory for atypical representations is more complicated than a naive generalization from the

[14] The circle denotes the dual with respect to the bosonic part of the superinvariant bilinear form $\kappa = \kappa_r - \kappa_\ell$. The dual with respect to the positive definite form $\kappa_r + \kappa_\ell$ will be denoted by a star.

[15] There is a small caveat in this discussion. By our logic, the theory (2.60) gives $\mathcal{H} \simeq \wedge^\bullet \mathfrak{g}_{-1} \times R^0_{\lambda-\rho_{\overline{0}}}$, which is the superalgebra representation with the highest weight $\lambda - \rho_{\overline{0}}$, whereas the supergroup BWB predicts the highest weight to be $\lambda - \rho$. Presumably, the discrepancy can be cured if one takes into account the Jacobian of the transformation from the superinvariant measure in the full set of variables to the free measure in the (θ',θ) variables. In other words, that Jacobian gives the difference between the one-loop shift for SG and for its maximal bosonic subgroup G.

bosonic case [47]. One expects the Hilbert space of the SG/L model to be a finite-dimensional representation with highest weight Λ. However, rather than the irreducible representation, it might be the Kac module, or some quotient of it, or some more complicated indecomposable representation.

2.3.2.2 Monodromy Operators In The Three Dimensional Theory

By coupling the gauge field of Chern-Simons theory to the currents of BWB quantum mechanics, supported on a knot K, we can write a path integral representation of a Wilson operator supported on K:

$$\mathrm{Str}_{R_\Lambda} P \exp\left(-\oint_K A\right) = \int \mathrm{D}h \exp\left(-i\oint_K \lambda(h^{-1}\mathrm{d}_A h)\right). \tag{2.61}$$

Here K is an arbitrary knot – that is, a closed oriented 1-manifold in W. As we have explained, this replacement is justified at least for typical representations. In the atypical case, we expect this replacement to be valid if R_Λ is chosen correctly within its block.

To establish the relation between Wilson lines and monodromy operators, we remove the BWB degrees of freedom by a gauge transformation. This is possible because G acts transitively on G/T; thus, we can pick a gauge transformation along K that maps h to a constant element of G/T. For a regular weight λ, the choice of this constant element breaks the G gauge symmetry along K down to T. What remains of the functional integral (2.61) is an insertion of an abelian Wilson line

$$\exp\left(-i\oint_K \lambda(A)\right). \tag{2.62}$$

With this insertion, the classical equations derived from the Chern-Simons functional integral require the gauge field strength to have a delta-function singularity along the knot,

$$F = \frac{2\pi\lambda^\circ}{\mathcal{K}}\delta_K. \tag{2.63}$$

For example, if r, θ are polar coordinates in the normal plane to the knot, then this equation can be obeyed with

$$A = \frac{\lambda^\circ}{\mathcal{K}}\, \mathrm{d}\theta. \tag{2.64}$$

We note that $\mathrm{d}\theta$ is singular at $r = 0$, that is, along K. In quantum theory, the classical equations do not always hold. However, to develop a sensible quantum theory, it is necessary to work in a space of fields in which it is possible to obey the classical equations. One accomplishes this in the present case by quantizing the theory in a space of fields characterized by

$$A = \frac{\lambda^\circ}{\mathcal{K}}\, \mathrm{d}\theta + \dots \tag{2.65}$$

where the ellipses refer to terms less singular than $\mathrm{d}\theta$ at $r = 0$. This gives the definition of a monodromy operator.

Note that in (2.61), to rewrite a Wilson line for a dominant weight Λ, we used a weight $\lambda = \Lambda + \rho$. The motivation for this shift was given in our review of the coadjoint orbit quantum mechanics, but this point requires more explanation. In the ordinary three-dimensional formulation of Chern-Simons theory, it is known that such shifts of the weights should not be included in the definition of the monodromy operators, but rather they appear in the final answers as quantum corrections [48]. This is analogous to the shift in the level $k \to k + h\ \mathrm{sign}(k)$. However, in the analytically-continued theory, we have to put the shift of k by hand into the classical action, and one expects that the same should be done with the shifts of the weights.[16] For example, let us look at the expectation value for the unknot, labeled with the spin j representation, in the SU(2) Chern-Simons theory on $\mathbb{R}^3$. This expectation value is

$$Z_j(\mathbb{R}^3) = \frac{\sin\left(2\pi\left(j + 1/2\right)/\mathcal{K}\right)}{\sin\left(\pi/\mathcal{K}\right)}. \tag{2.66}$$

This formula is derived from the relation with conformal field theory, so j here is a non-

[16] Both of these shifts arise from the phase of an oscillatory Gaussian integral, as was explained in the case of k in section 2.2.6. In the 4d analytically-continued version of the theory, the Gaussian integrals are real and will not generate shifts.

negative half-integer. In the analytically-continued theory, we want to replace the Wilson line of the spin j representation with a monodromy operator, and assume that the answer is given by the same simple formula (2.66). The prescribed monodromy around the knot is defined by

$$F = 2\pi i \frac{j'\,\sigma_3}{\mathcal{K}} \delta_K \,, \tag{2.67}$$

where $\sigma_3 \in \mathfrak{su}(2)$ is the Pauli matrix. We need to choose between taking $j' = j$ or $j' = j+1/2$. Note that the Weyl transformation of the field in (2.67) brings j' to $-j'$. It should leave the expectation value invariant, up to sign.[17] The symmetry of the formula (2.66) is consistent with this, if we take $j' = j + 1/2$.

So we will assume that the monodromy operator in the analytically-continued theory, which corresponds to a representation with weight Λ, should be defined using the shifted weight $\lambda = \Lambda + \rho$. However, let us comment on some possible issues related to these shifts. For a type I superalgebra, the Weyl vector ρ has integral Dynkin labels, so, if Λ is an integral weight, then λ is also an integral weight of the superalgebra. But for the $\mathfrak{u}(m|n)$ case, it might not be an integral weight of the supergroup. This can be illustrated even in the purely bosonic case. For U(2), the quantum correction $\Lambda \to \Lambda + \rho$ shifts the SU(2) spin by one-half, and does not change the eigenvalue of the central generator $\mathfrak{u}(1) \subset \mathfrak{u}(2)$. The resulting weight is a well-defined weight of $\mathrm{SU}(2) \times \mathrm{U}(1)$, but not of $\mathrm{U}(2) \simeq (\mathrm{SU}(2) \times \mathrm{U}(1))/\mathbb{Z}_2$. For a type II superalgebra, the problem can be worse. If Λ is an integral weight of the superalgebra, λ might not be an integral weight of the superalgebra itself, because the Weyl vector ρ can have non-integer Dynkin labels.[18]

We will not try to resolve these puzzles, but will just note that in one approach to the line observables of the analytically-continued theory, one replaces a Chern-Simons monodromy

[17]For a knot in $\mathbb{R}^3$ or S^3, there is essentially only one integration cycle, so the Weyl reflection maps the integration cycle to an equivalent one. But it may reverse the orientation of the integration cycle, and that is the reason for the sign. A related explanation of the sign was given in section 2.3.2.1.

[18]For any simple root α of the superalgebra, it is true that $2\langle\rho, \alpha\rangle = \langle\alpha, \alpha\rangle$. From this one infers that the Dynkin label of the Weyl vector is either one or zero. However, for a type II superalgebra there exists a simple root of the bosonic subalgebra, which is not a simple root of the superalgebra, and for that root the Dynkin label of ρ need not be integral.

operator by a surface operator in four dimensions. In that case, the fact that λ° in eqn. (2.65) is defined using a non-integral weight presents no problem with gauge-invariance for much the same reason that the non-integrality of $\mathcal{K}$ presents no problem: the "big" gauge transformations that lead to integrality of the parameters in purely three-dimensional Chern-Simons theory do not have analogs[19] in the four-dimensional setting.

Finally, we can return to the question of making sense of a path integral for a knot $K \subset S^3$ labeled by a typical representation. As remarked following eqn. (2.56), a direct attempt to do this in the language of Wilson operators leads to a $0 \cdot \infty$ degeneracy. This degeneracy is naturally resolved by replacing the Wilson operator by a monodromy operator with weight λ. In perturbation theory in the presence of a monodromy operator supported on a knot K, the functional integral is evaluated by expanding near classical flat connections on the complement of K whose monodromy around K has a prescribed conjugacy class. The group H of unbroken gauge symmetries of any such flat connection, for λ typical, is a purely bosonic subgroup of SG, because the fermionic gauge symmetries have been explicitly broken by the reduction of the gauge symmetry along K from SG to a bosonic subgroup (this subgroup is T if λ is regular as well as typical).[20] To compute the functional integral expanded around a classical flat connection, one has to divide by the volume of H, but this presents no problem: as H is purely bosonic, its volume is not zero. So in the monodromy operator approach, there is no problem to define a path integral on S^3 with insertion of a knot labeled by a typical representation.

Now let us consider loop operators in $\mathbb{R}^3$ rather than S^3. We have claimed that a path integral on $\mathbb{R}^3$ with a Wilson operator labeled by a representation of non-maximal atypicality

[19]In going from three to four dimensions, the support of a monodromy operator is promoted from a knot K to a two-manifold C with boundary K. If C is compact, a homotopically non-trivial map from K to the maximal torus $T \subset G$ does not extend over C. If $C = K \times \mathbb{R}_+$, such a gauge transformation can be extended over C, but the extension does not approach 1 at infinity. In a noncompact setting, one only requires invariance under gauge transformations that are 1 at infinity.

[20]For any K, there is an abelian flat connection on the complement of K with the prescribed monodromy around K, unique up to gauge transformation. Its automorphism group is T if λ is regular as well as typical (and otherwise is a Levi subgroup L that lies between T and G). In general, there may be nonabelian flat connections with the required monodromies; their automorphism groups are smaller.

is 0. This must remain true if the Wilson operator is replaced by a corresponding monodromy operator. Let us see how this happens. The difference between $\mathbb{R}^3$ and S^3 is that in defining the path integral on $\mathbb{R}^3$, we only divide by gauge transformations that are trivial at infinity. If on S^3 a flat connection has an automorphism subgroup H, then on $\mathbb{R}^3$ it will give rise to a family of irreducible connections, with moduli space SG/H. The volume of this moduli space will appear as a factor (in the numerator!) in evaluating the path integral. If H is purely bosonic, then the quotient SG/H has fermionic directions, and its volume generally vanishes.[21] Therefore, the expectation value of a closed monodromy operator in $\mathbb{R}^3$, for λ typical, vanishes (except for $\mathrm{B}(0,n)$), in agreement with the corresponding statement for the Wilson loop. To analyze the case that the weight λ is not typical, we need to extend the BWB quantum mechanics for atypical weights, and presumably we will then need to compute the invariant volume of a homogeneous space SG/H, where now H will be a subsupergroup. It is plausible that for λ of sufficient atypicality, this volume can be non-zero,[22] so that the monodromy operator can have a non-trivial expectation value. But we have not performed this computation.

2.3.3 Line Observables In Four Dimensions

Our next goal is to interpret the line operators that we have discussed in the full four-dimensional construction. First we consider Wilson lines, and explore their symmetries in the physical 4d super Yang-Mills theory associated to the D3-NS5 system.

2.3.3.1 Wilson Operators

For generic values of t, $\mathcal{N}=4$ super Yang-Mills theory in bulk does not admit Q-invariant Wilson operators. (They exist precisely if $t^2=-1$, a fact that is important in the geometric

[21] As always, the exception is a supergroup from the series $\mathrm{B}(0,n)$.

[22] In view of an argument explained in footnote 9, a necessary condition is that any fermionic generator $\mathcal{C}$ of SG that obeys $\{\mathcal{C},\mathcal{C}\}=0$ must be conjugate to a generator of H. This ensures that the group F generated by $\mathcal{C}$ does not act freely on SG/H, so that the argument of footnote 9 cannot be used to show that the volume of SG/H is 0. For $\mathrm{U}(m|n)$, it follows from this criterion that SG/H has zero volume except possibly if λ is maximally atypical.

Langlands correspondence [14].) However, on the defect W there always exist supersymmetric Wilson operators

$$\mathcal{W}_R(K) = \mathrm{Str}_R P\exp\left(-\oint_K \mathcal{A}\right) , \tag{2.68}$$

labeled by an arbitrary representation R of the supergroup SG. Here $\mathcal{A}$ is the supergroup gauge field and Q-invariance is clear since Q acts on $\mathcal{A}$ by gauge transformations.

These are the most obvious Q-invariant line operators, but they have a drawback that makes them harder to study: as operators in the physical $\mathcal{N}=4$ super Yang-Mills theory, they have less symmetry than one might expect. We will analyze the symmetry of these operators in different situations.

The procedure by which we constructed a topological field theory involved twisting four of the six scalars of $\mathcal{N}=4$ super Yang-Mills theory, leaving two untwisted scalars and hence an unbroken R-symmetry group $\mathrm{U}(1)_F = \mathrm{SO}(2) \subset \mathrm{SO}(6)_R$. In the special case that $M = \mathbb{R}\times W$ with a product metric, there is no need for twisting in the $\mathbb{R}$ direction to maintain supersymmetry, so three scalars remain untwisted and $\mathrm{U}(1)_F$ is enhanced to $\mathrm{SU}(2)_Y$. The supercharge Q that we chose in constructing a topological field theory was one component of an $\mathrm{SU}(2)_Y$ doublet. For $M = \mathbb{R}\times W$, the twisted action is invariant under $\mathrm{SU}(2)_Y$ as well as Q, so it inevitably preserves two supercharges – both components of the doublet containing Q. Likewise, the Wilson loop operators (2.68) are invariant under $\mathrm{SU}(2)_Y$ as well as Q, so on $M = \mathbb{R}\times W$, they really preserve two supersymmetries.

Now let us specialize further to the case that $W = \mathbb{R}^3$ is flat, with $M = \mathbb{R}\times W = \mathbb{R}^4$. In this case, no topological twisting is necessary, but the half-BPS defect supported on W breaks the R-symmetry group to $\mathrm{SO}(3)_X\times\mathrm{SO}(3)_Y$. In addition, there is an unbroken rotation group $\mathrm{SO}(3)$, and, as explained in section 2.2.1, the unbroken supersymmetries transform as $(\mathbf{2},\mathbf{2},\mathbf{2})$ under $\mathrm{SO}(3)\times\mathrm{SO}(3)_X\times\mathrm{SO}(3)_Y$. Let us consider a Wilson operator $\mathcal{W}_R(K)$ where K is a straight line $\mathbb{R}\subset W$, say the line $x_1 = x_2 = 0$, parametrized by x_0. K is invariant under a subgroup $\mathrm{SO}(2)\subset\mathrm{SO}(3)$ of rotations of x_1 and x_2. To identify the global symmetry of $\mathcal{W}_R(K)$ involves a crucial subtlety. First let us consider the one-sided case studied in [6], in other words the case of an ordinary gauge group G rather than a supergroup SG. In this case, the supergroup connection reduces to $\mathcal{A}_b = A + i(\sin\vartheta)\phi$, and the Wilson operator for a straight Wilson line depends on one component ϕ_0 of a triplet (ϕ_0, ϕ_1, ϕ_2) of $\mathrm{SO}(3)_X$. This field is invariant under a subgroup $\mathrm{SO}(2)_X \subset \mathrm{SO}(3)_X$, and hence a straight Wilson line in the case of an ordinary gauge group has global (bosonic) symmetry $\mathrm{SO}(2)\times\mathrm{SO}(2)_X\times\mathrm{SO}(3)_Y$. In the supergroup case, we must remember that the supergroup connection also has a fermionic part $\mathcal{A}_f$ which began life as part of a field that transforms as $(\mathbf{2},\mathbf{2})$ under $\mathrm{SO}(3)\times\mathrm{SO}(3)_X$. As a result, the component of $\mathcal{A}_f$ in the x_0 direction is not separately invariant under $\mathrm{SO}(2)$ and $\mathrm{SO}(2)_X$ but only under a diagonal combination $\mathrm{SO}'(2) \subset \mathrm{SO}(2)\times\mathrm{SO}(2)_X$. Hence the bosonic global symmetry of a straight Wilson line in the supergroup case is $\mathrm{SO}'(2)\times\mathrm{SO}(3)_Y$, reduced from the corresponding symmetry in the case of an ordinary Lie group.

The supersymmetry of a straight Wilson line $\mathcal{W}_R(K)$ is likewise reduced in the supergroup case from what it is in the case of an ordinary Lie group. A supersymmetry has no chance to preserve the straight Wilson line if its commutator with the complexified bosonic gauge field $\mathcal{A}_b$ has a contribution proportional to Ψ_1. Indeed, the boundary conditions do not tell us anything about the behaviour of Ψ_1 at $x_3 = 0$, so there would be no way to cancel such a term. Inspection of the supersymmetry transformations (2.11) reveals that, apart from the $\mathrm{SO}'(3)$-invariant supersymmetries with generators

$$\varepsilon^{\alpha A\dot{A}} = \epsilon^{\alpha A} w^{\dot{A}} \tag{2.69}$$

(familiar from eqn. (2.24)), with arbitrary two-component spinor $w^{\dot{A}}$, the only supersymmetries that do not produce variations of $\mathcal{A}_b$ proportional to Ψ_1 are those with generators

$$\varepsilon^{\alpha A\dot{A}} = \sigma_0^{\alpha A} \widetilde{w}^{\dot{A}}\,, \tag{2.70}$$

where again $\widetilde{w}^{\dot{A}}$ is an arbitrary spinor. Since $\widetilde{w}^{\dot{A}}$ transforms as a spinor of $\mathrm{SU}(2)_Y$, an $\mathrm{SU}(2)_Y$-invariant Wilson operator is invariant under this transformation for all $\widetilde{w}^{\dot{A}}$ if and only if it is invariant for some particular nonzero $\widetilde{w}^{\dot{A}}$. A choice that is convenient because it enables us to write simple formulas in the language of the twisted theory is to set $\widetilde{w}^{\dot{A}} = v^{\dot{A}}$ (where $v^{\dot{A}}$ was defined in (2.28)). Writing $\widetilde{\delta}$ for the transformation generated by the corresponding supersymmetry, one computes that

$$\widetilde{\delta}\mathcal{A}_0 = -i[\overline{C}, \mathcal{B}\}\,, \tag{2.71}$$

where we define

$$\mathcal{B} = \{C, \overline{C}\} + B. \tag{2.72}$$

Since (2.71) is non-zero, our Wilson lines do not preserve supersymmetries (2.70) for a generic representation. Therefore, they preserve only the two supersymmetries (2.69). They are 1/4-BPS objects from the standpoint of the defect theory (or 1/8 BPS relative to the underlying $\mathcal{N} = 4$ super Yang-Mills theory). This is an important difference from the case of a purely bosonic gauge group, in which Wilson lines preserve four supersymmetries (a fact that greatly simplifies the analysis of the dual 't Hooft operators [6, 51]). In fact, if the representation R that labels the Wilson line $\mathcal{W}_R(K)$ is such that the fermionic generators act trivially, then (2.71) vanishes, and $\mathcal{W}_R(K)$ becomes 1/2-BPS (in the defect theory), as in the bosonic or one-sided case. More generally, for (2.71) to vanish it is enough that the anticommutators of the fermionic generators vanish in the representation R. Of course, in the case of a supergroup such as $\mathrm{U}(m|n)$, this is a very restrictive condition.

One can also construct other Q-invariant Wilson operators in the electric theory, by adding a polynomial of the Higgs field $\mathcal{B}$ to the connection in the Wilson line. The resulting operators preserve $1/4$ or $1/8$ of the three-dimensional supersymmetry. In the Q-cohomology, such operators are equivalent to the ordinary Wilson lines (2.68), and for this reason we will not say more about them.

Why do we care about the reduced supersymmetry of the supergroup Wilson loop operators? One of our goals will be to understand what happens to line operators under nonperturbative dualities. For this purpose, the fact that the supergroup Wilson operators are only 1/4 and not 1/2 BPS is rather inconvenient. Possible constructions of a dual operator that preserve 4 supercharges are much more restrictive than possible constructions that preserve only 2 supercharges. We will obtain a reasonable duality picture for certain 1/2 BPS Wilson-'t Hooft line operators that will be introduced in section 2.3.3.2. These Wilson-'t Hooft operators are labeled by weights of SG and the way they are constructed suggests that from the point of view of the twisted topological field theory – the supergroup Chern-Simons theory – they are equivalent to Wilson operators. But because of their enhanced supersymmetry, it is much easier to find their duals.

About the Wilson operators, we make the following remarks. We were not able to find a construction of 't Hooft-like disorder operators – characterized by a singularity of some kind – with precisely the right global symmetries so that they might be dual to the Wilson operators constructed above. It may be that one has to supplement an 't Hooft-like construction by adding some quantum mechanical variables that live along the line operators (analogous to the BWB variables that we discussed in section 2.3.2.1). With only 2 supersymmetries to be preserved, there are many possibilities and we do not know a good approach. Also, the fact that the two-dimensional space of supersymmetries preserved by a Wilson operator is not real suggests that it is difficult to realize such an object in string theory. A string theory realization would probably have helped in understanding the action of duality.

2.3.3.2 Wilson-'t Hooft Operators

For all these reasons, we now move on to consider Wilson-'t Hooft operators.

$\mathcal{N} = 4$ super Yang-Mills theory supports BPS Wilson-'t Hooft line operators in the bulk [52]. Though they preserve 8 supersymmetries, generically these do not include the specific supersymmetry Q. The condition for a Wilson-'t Hooft operator in bulk to be Q-invariant is that its electric and magnetic charges must be proportional with a ratio $\mathcal{K}$ [14]. Since both charges have to be integral, Q-invariant Wilson-'t Hooft operators exist in the bulk only for

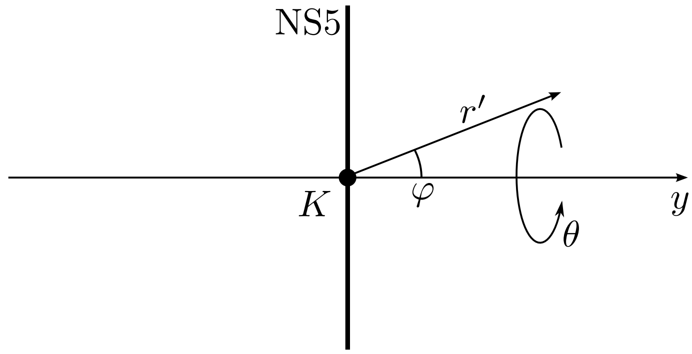

Figure 2.6: The hyperplane $x^0 = 0$ showing our notation for the coordinates. $y = x^3$ runs horizontally; the $x^0 = 0$ section of the knot K is shown as a dot in the center.

rational values of the canonical parameter $\mathcal{K}$. In the present chapter, we generally assume $\mathcal{K}$ to be generic.

However, we are interested in operators that are supported not in the bulk but along the defect at $x^3 = 0$. The gauge theories with gauge groups G_ℓ and G_r live in half-spaces, and the magnetic flux for each gauge group can escape through the boundary of the half-space and so is not quantized. So a Wilson-'t Hooft operator that lives only at $y = 0$ is no longer constrained to have an integral magnetic charge. Such operators can exist for any (integer) electric charge and arbitrary $\mathcal{K}$. To define them precisely, we work in the weak coupling regime, where g_{YM} is small, and therefore, according to (2.7), $\mathcal{K}$ is large. The weight of the representation is taken to scale with $\mathcal{K}$, so that the monodromy of the gauge field, which is proportional to $\lambda/\mathcal{K}$, is fixed.

Consider a line operator located at $y = 0$ along the x^0 axis. (See fig. 2.6 for the notation.) We want to find a model solution of the BPS equations (2.30) that will define the singular asymptotics of the fields near the operator. For definiteness, consider the Yang-Mills theory on the right of the three-dimensional defect. We make a conformally-invariant abelian ansatz which preserves the $\mathrm{SO}(2) \times \mathrm{SO}(2)_X \times \mathrm{SO}(3)_Y$ symmetry:

$$
\begin{aligned}
A &= c_a \frac{\mathrm{d}x^0}{r'} + \mathfrak{m}_r (1 - \cos\varphi)\mathrm{d}\theta\,, \\
\phi &= c_\phi \frac{\mathrm{d}x^0}{r'}\,.
\end{aligned}
\tag{2.73}
$$

Here $\mathfrak{m}_r$ is the magnetic charge (which as noted above will not be quantized). The ray $\varphi = 0$ points in the $y > 0$ direction, and the signs were chosen such that there is no Dirac string along this ray. The localization equations (2.30) are satisfied if

$$c_a = i\mathfrak{m}_r \tan\vartheta\,, \quad c_\phi = -\frac{\mathfrak{m}_r}{\cos\vartheta}\,, \tag{2.74}$$

where ϑ is the angle related to the twisting parameter t, as introduced in section 2.2. The hypermultiplet fields are taken to vanish. The factor of i in the Coulomb singularity of the gauge field A is an artifact of the Euclidean continuation; in Lorentz signature, the solution would be real. Eqn. (2.73) fixes the behaviour of the bulk fields near a line operator. For a generic magnetic charge, the fields of the hypermultiplet do not commute with the singularity in (2.73), and thus are required to vanish along the operator.

Let us check that our model solution satisfies also the boundary conditions at $y = 0$. The boundary conditions can be derived from (2.17) and an analogous expression for the scalar X^a. This is done in Appendix B. However, in the topological theory one can understand the relevant features by a more simple argument. The boundary condition should require vanishing of the boundary part of the variation of the action of the theory. Suppose that we consider a configuration in which all the fermions vanish, and the bosonic fields satisfy the localization equations. The variation of the non-$\mathcal{Q}$-exact Chern-Simons term (equivalently, the topological term) gives the gauge field strength $\mathcal{F}_b$. The $\mathcal{Q}$-exact terms in the action come in two different sorts. There is a bulk contribution, whose bosonic part is proportional to the sum of squares of the localization equations (2.30). The variation of these terms vanishes when the equations are satisfied. There are also $\mathcal{Q}$-exact terms supported on the defect; they furnish gauge fixing of the fermionic gauge symmetry of the supergroup Chern-Simons. Their variation is proportional to the hypermultiplet fields. Therefore, we conclude that if the fields satisfy the localization equations, and the three-dimensional hypermultiplet vanishes, the boundary condition reduces to

$$\imath^* \left(\mathcal{F}_b\right) = 0\,, \tag{2.75}$$

where $\imath : W \hookrightarrow M$ is the natural embedding of the three-manifold into the bulk manifold. This boundary condition is indeed satisfied by the model solution (2.73), (2.74), because in the complexified gauge field $\mathcal{A}_b = A + i(\sin\vartheta)\phi$, the Coulomb parts of A and ϕ cancel. (The magnetic part is annihilated by $\imath^*$.) In fact, at $y = 0$, the complexified field $\mathcal{A}_b$ reduces to the field of a Chern-Simons monodromy operator (2.63), if we identify $\mathfrak{m} = \lambda^\circ/\mathcal{K}$, where now $\mathfrak{m}$ includes both the part in $\mathfrak{g}_\ell$ and in $\mathfrak{g}_r$.

In Chern-Simons theory, in the presence of a monodromy defect, the bulk action is supplemented with an abelian Wilson line (2.62) along the defect; in our derivation in section (2.3.2.2), this is what remained after gauge-fixing the BWB action. The Chern-Simons action with an insertion of an abelian Wilson line is characterized by the fact that its variation near the background singular field (2.63) does not have a delta function term supported on the knot (a delta function term that would come from the variation of Chern-Simons in the presence of the monodromy singularity is canceled by the variation of the abelian Wilson operator). In four dimensions, in the presence of a singularity along a knot K, the topological action (2.35) should be integrated over the four-manifold with a neighborhood of K cut out, and taking into account the singularity along K of the Wilson-'t Hooft operator, this produces a term in the variation with delta-function support along K:

$$\delta\left(\frac{i\mathcal{K}}{4\pi}\int_{M\setminus K}\mathrm{tr}\left(\mathcal{F}\wedge\mathcal{F}\right)\right) = i\oint_K \mathrm{Str}\left(\lambda^\circ\delta\mathcal{A}\right). \tag{2.76}$$

To cancel this variation, just like in three dimensions, one inserts an abelian Wilson line (2.62).

But now we learn something fundamental. Although the Wilson-'t Hooft operators that we have constructed do not have a quantized magnetic charge, they have a quantized electric charge. The abelian Wilson line is only gauge-invariant if λ is an integral weight of $G_\ell \times G_r$. For a type I superalgebra such as $\mathfrak{u}(m|n)$, an integral weight of $G_\ell \times G_r$ corresponds to an integral weight of the supergroup $\mathrm{U}(m|n)$ and therefore, these Wilson operators are classified by integral weights of the supergroup. The Weyl group of $\mathrm{U}(m|n)$ is the same as that of its bosonic subgroup $\mathrm{U}(m) \times \mathrm{U}(n)$, so an equivalent statement is that Wilson operators of the supergroup (for irreducible typical representations, or some particular atypical representations) are in correspondence with this class of Wilson-'t Hooft operators. The advantage of the Wilson-'t Hooft operators is that they have more symmetry: in addition to Q-invariance, they are half-BPS operators with the full $\mathrm{SO}(2) \times \mathrm{SO}(2)_X \times \mathrm{SO}(3)_Y$ symmetry, just like a Wilson line in the one-sided problem.

For a type II superalgebra, such as $\mathfrak{osp}(2m+1|2n)$, there is a slight complication. For such algebras, some "small" dominant weights do not correspond to representations. (These are the weights that do not satisfy the "supplementary condition," as defined in section 2.3.1.1. See also section 2.5.6.2 for details in the case of $\mathrm{OSp}(2m+1|2n)$.) Our construction gives a half-BPS line operator for every dominant weight whether or not this weight corresponds to a representation. It is hard to study explicitly why some Wilson-'t Hooft operators with small weights do not correspond to representations, because the semiclassical description of a Wilson-'t Hooft operator is valid for large weights.[23]

2.3.3.3 Twisted Line Operators

In section 2.5, we will discuss a non-perturbative duality for Chern-Simons theory with orthosymplectic supergroup $\mathrm{OSp}(r|2n)$. It will turn out that line operators labeled by dominant weights of the supergroup are not a closed set of operators under that duality. To get a duality-invariant picture, one needs to include what we will call twisted line operators.

The clearest explanation seems to be also the most naive one. We consider 4d super

[23]Given this, one may wonder if the half-BPS property is lost when the weights are too small. We doubt that this is the right interpretation because the construction of half-BPS line operators on the magnetic side, discussed in section 2.4.4, appears to be valid for all weights.

Yang-Mills theory on $W \times \mathbb{R}_y$, where $\mathbb{R}_y$ is parametrized by y. For $y < 0$, the gauge group is $\mathrm{SO}(r)$; for $y > 0$, it is $\mathrm{Sp}(2n)$. Along $W \times \{y = 0\}$ is a bifundamental hypermultiplet.

Now we pick a knot $K \subset W$, and define a line operator supported on K by saying that the hypermultiplet fermions change sign under monodromy around K. Locally, this makes perfect sense. Globally, to make sense of it, we have to say essentially that the hypermultiplets are not just bifundamentals, but are twisted by a $\mathbb{Z}_2$ bundle defined on $W \times \{y = 0\}$ that has monodromy around K. If such a flat bundle does not exist, we say that the path integral with insertion of the given line operator is 0. If there are inequivalent choices for this flat bundle, we sum over the choices.

This procedure actually defines not just a single new line operator, but a whole class of them, which we will call twisted line operators. The reason is that the monodromy around K forces the hypermultiplets to vanish along K, and therefore there is no problem to include arbitrary $\mathrm{SO}(r)\times\mathrm{Sp}(2n)$ Wilson operators along K. This class of operators will be important in section 2.5.

Can we do something similar for $\mathrm{U}(m|n)$? In this case, we can pick an arbitrary nonzero complex number e^{ic}, embedded as an element of the center of $\mathrm{U}(n)$ (or of its complexification if c is not real), and twist the hypermultiplet fields by e^{ic} under monodromy around K. Then we proceed as just explained, and get a family of line operators that depend on the parameter c. Again, from a global point of view, this means the hypermultiplets are bifundamentals twisted by a flat line bundle with monodromy e^{ic} around K, and we define the path integral by summing over the possible flat bundles that obey this condition. And again, we can generalize this definition by including Wilson operators of $\mathrm{U}(m) \times \mathrm{U}(n)$.

2.3.4 Surface Operators

In the relation of 3d Chern-Simons theory to 4d gauge theory, there are two possible strategies for finding a 4d construction related to a line operator in the Chern-Simons theory.

In one picture, the 3d line operator is promoted to a 4d line operator that lives on the defect that supports the Chern-Simons gauge fields. In the second picture, a line operator in 3d is considered to have its support in codimension 2, and it is promoted to a surface operator in 4d, whose support is in codimension 2.

So if Chern-Simons theory on a three-manifold W is related to 4d super Yang-Mills on $W \times \mathbb{R}_y$, where $\mathbb{R}_y$ is a copy of $\mathbb{R}$ parametrized by y with a defect at $y = 0$, then in the first approach, a 3d line operator supported on $K \subset W$ is promoted in 4d to a line operator supported on $K \times \{y = 0\}$. In the second approach, a 3d line operator supported on K is promoted to a 4d surface operator supported on a two-manifold C such that $C \cap \{y = 0\} = K$. For example, C might be simply $K \times \mathbb{R}_y$.

Both of these viewpoints were explored in [6] for the one-sided problem, although the first one based on 4d line operators was developed in more detail. In the two-sided case, we have followed the first viewpoint so far but now we turn to the second one and consider surface operators.

We focus on the simplest half-BPS surface operators, which were described in the bulk in [53]. Our problem is to understand what happens when one of these operators intersects a fivebrane. In the present section, we answer this question on the electric side (that is, for an NS5-brane). In section 2.4, we answer the question on the magnetic side (that is, for a D5-brane).

One advantage to the formulation via surface operators in four dimensions rather than line operators is that the behavior under S-duality is simple to understand. That is because, in the 4d bulk, one already knows the behavior under S-duality of the surface operators we will be studying. Given a surface operator intersecting an NS5-brane, the S-dual of this configuration will have to consist of the S-dual surface operator intersecting a D5-brane. So all we have to do is to determine what happens when a surface operator intersects an NS5-brane or a D5-brane. S-duality will then take care of itself.

2.3.4.1 Surface Operators In The Bulk

The simplest half-BPS surface operators in $\mathcal{N}=4$ super Yang-Mills theory are labeled by a set of four parameters $(\alpha,\beta,\gamma,\eta)$. The first three define the singular behavior of the fields near the support of the operator, which will be a two-manifold C. If r and θ are polar coordinates in the normal plane to C, we require the fields near C to behave like

$$\begin{aligned} A &= \alpha\,\mathrm{d}\theta + \dots\,, \\ \phi &= \beta\,\frac{\mathrm{d}r}{r} - \gamma\,\mathrm{d}\theta + \dots\,, \end{aligned} \tag{2.77}$$

where the ellipses represent less singular terms. The parameters α, β and γ take values in a Cartan subalgebra $\mathfrak{t} \subset \mathfrak{g}$. More precisely, one can make big gauge transformations on the complement of C that shift α by an arbitrary cocharacter; therefore, α should be considered as an element of the maximal torus[24] $T \simeq \mathfrak{t}/\Gamma_{\mathrm{cochar}}$.

The meaning of the fourth parameter η is the following. Assume that the triple (α,β,γ) is regular, that is, it commutes only with $\mathfrak{t}$. In this case the singular behavior (2.77) reduces the gauge group along the surface operator to its maximal torus T, and it makes sense to speak of the first Chern class of the resulting T-bundle on C. One can define the $\mathfrak{t}^*$-valued theta-angle η coupled to this Chern class, and introduce a factor

$$\exp\left(i\int_C \eta(F)\right) \tag{2.78}$$

in the functional integral. By integrality of the first Chern class, this expression is invariant under a shift of the theta-angle by an element of the character lattice $\Gamma_{\mathrm{char}} \subset \mathfrak{t}^*$, so η really takes values in the maximal torus of the Langlands-dual group, $\eta \in T^\vee \simeq \mathfrak{t}^*/\Gamma_{\mathrm{char}}$. Dividing by the action of the Weyl group $\mathcal{W}$, which is a remnant of the non-abelian gauge symmetry, we get that the parameters $(\alpha,\beta,\gamma,\eta)$ take values in $(T,\mathfrak{t},\mathfrak{t},T^\vee)/\mathcal{W}$.

[24] In this section we discuss only the bulk $\mathcal{N}=4$ Yang-Mills theory, and all our notation refers to its bosonic gauge group, and not to a supergroup.

The singular asymptotics of the fields (2.77) satisfy the localization equations (2.30) for any value of t, if supplemented with appropriate sources,

$$\begin{aligned} F - \phi \wedge \phi &= 2\pi\alpha\, \delta_C\,, \\ \mathrm{d}_A \phi &= -2\pi\gamma\, \delta_C \\ \mathrm{d}_A \star \phi &= 2\pi\beta\, \mathrm{d}x^0 \wedge \mathrm{d}y \wedge \delta_C \end{aligned} \tag{2.79}$$

where $\delta_C = \mathrm{d}(\mathrm{d}\theta)/2\pi$ is the δ-function 2-form that is Poincaré dual to the surface C, and x^0 and y are coordinates along the surface.

The prescribed singularities (2.77) define the space of fields over which one integrates to define $\mathcal{N} = 4$ super Yang-Mills theory in the presence of the surface operator. Let us also define more precisely what functional we are integrating over this space. The action of the bulk topological theory consists of the topological term and some $\mathcal{Q}$-exact terms (2.36). In the presence of the surface operator, the topological term is defined as

$$\frac{i\mathcal{K}}{4\pi} \int_M' \mathrm{tr}(F \wedge F)\,, \tag{2.80}$$

where the symbol $\int_M'$ denotes an integral over $M \setminus C$, not including a delta function contribution along C. Alternatively, we can write this as an integral over the whole four-manifold, and explicitly subtract the contribution which comes from the delta-function singularity of the curvature:

$$\frac{i\mathcal{K}}{4\pi} \int_M' \mathrm{tr}(F \wedge F) = \frac{i\mathcal{K}}{4\pi} \int_M \mathrm{tr}(F \wedge F) - i\mathcal{K} \int_C \mathrm{tr}(\alpha F) - i\pi\mathcal{K}\, \mathrm{tr}(\alpha^2)\, C \cap C. \tag{2.81}$$

The c-number contribution proportional to the self-intersection number $C \cap C$ appears here from the square of the delta-function.

In the absence of the surface operator, the $\mathcal{Q}$-exact part of the action has the form

$$-\frac{1}{g_{\text{YM}}^2}\int \operatorname{tr}\left(\frac{2t^{-1}}{t+t^{-1}}\,\mathcal{V}^+\wedge\mathcal{V}^+ - \frac{2t}{t+t^{-1}}\,\mathcal{V}^-\wedge\mathcal{V}^- + \mathcal{V}^0\wedge\star\mathcal{V}^0\right)\,, \tag{2.82}$$

where $\mathcal{V}^+$, $\mathcal{V}^-$ and $\mathcal{V}^0$ are the left hand sides of the supersymmetric localization equations, as defined in (2.30). In the presence of the surface operator, the localization equations acquire delta-function sources, as in (2.79). The action (2.82) is modified accordingly, *e.g.*, the first term becomes

$$-\frac{1}{g_{\text{YM}}^2}\int \operatorname{tr}\left(\frac{2t^{-1}}{t+t^{-1}}\left(\mathcal{V}^+ - 2\pi(\alpha - t\gamma)\delta_C^+\right)\wedge\left(\mathcal{V}^+ - 2\pi(\alpha - t\gamma)\delta_C^+\right)\right). \tag{2.83}$$

Because it contains the square of a delta function, this expression is at risk of being divergent. To make the action finite, one works in a class of fields in which the localization equations (2.79) are satisfied, modulo smooth terms. In other words, the left hand side of the localization equations must contain the same delta functions as the right hand side.

In the definition of the surface operator, it was assumed that the singularity defined by (α, β, γ) is regular, that is, the gauge group along the operator is broken down to the maximal torus. This is the case for which the theta-angles η can be defined classically. But it can be argued that the surface operator is actually well-defined quantum mechanically as long as the full collection of couplings $(\alpha, \beta, \gamma, \eta)$ is regular. One approach to showing this involves a different construction of these surface operators with additional degrees of freedom along the surface as described in section 3 of [54]. In this chapter, we will try to avoid these issues.

2.3.4.2 Surface Operators In The Electric Theory

Let us specialize to a four-manifold $M = W\times\mathbb{R}_y$, with an NS5-type defect along $W\times\{y=0\}$. To incorporate a loop operator along the knot K in the Chern-Simons theory, we insert surface operators in the left and right Yang-Mills theories along a two-surface $C = C_\ell \cup C_r$ that intersects the $y = 0$ hyperplane along K. We could simply take C to be an infinite cylinder $K \times \mathbb{R}_y$, or we could take an arbitrary finite 2-surface. The orientations are taken to be such that $\partial C_r = -\partial C_\ell = K$. The parameters of the surface operators on the right and on the left will be denoted by letters with a subscript r or ℓ. Sometimes we will use notation without subscript to denote the combined set of parameters on the right and on the left (e.g., $\alpha = (\alpha_r, \alpha_\ell)$).

We would like to understand the meaning of the parameters of a surface operator in the Chern-Simons theory. It is clear that a surface operator with $\beta = \gamma = \eta = 0$ and non-zero α is equivalent to a monodromy operator in Chern-Simons, with weight $\lambda^\circ = \mathcal{K}\alpha$. Such a surface operator can be obtained e.g. as a Dirac string, which is produced by moving a Wilson-'t Hooft line operator in the four-dimensional theory into the bulk.

The parameter β has no direct interpretation in Chern-Simons, and defines just a deformation of the integration contour, without changing the path integral. As noted in [6], sometimes it might not be possible to turn on β. For example, let the bosonic gauge group be abelian, and let the three-manifold W be compact (e.g., $W \simeq S^3$). If we have a link with components labeled by $\beta_1, \dots, \beta_p$, then, integrating the third equation in (2.79) over W, we get that $\sum \beta_i l_i = 0$, where l_i is the length of the i-th component of the link. We see that if there is only one component, then β has to be zero.

The case of a surface operator with non-zero γ is a little subtle. It is not clear to us whether such an operator in the physical theory[25] can intersect (or end on) the three-dimensional defect in a Q-invariant way, and if it can, then to what line operator in Chern-Simons theory it would correspond. In topological theory, when one takes the parameter t to be real, such an operator makes perfect sense and has a natural Morse theory interpretation [6, 12]. In that case, the bosonic part of the action, modulo Q-exact terms, is defined in presence of a surface operator by an integral of the local density $\mathrm{tr}(\mathcal{F}_b \wedge \mathcal{F}_b)$ over the four-

[25]By the "physical theory" we mean the theory that in flat space describes the D3-NS5 intersection. In this theory, t is given by (2.29) and lies on the unit circle, and $\mathcal{K}$ is real. By the "topological theory," we mean the theory which arises naturally from the Morse theory construction [12, 13], with t being real, and $\mathcal{K}$ in general complex. In this chapter, we focus on the physical theory.

manifold $M \setminus C$. Up to some field-independent constants, we have, analogously to (2.81),

$$\frac{i\mathcal{K}}{4\pi}\int_{M_r}' \mathrm{tr}(\mathcal{F}_b \wedge \mathcal{F}_b) = \frac{i\mathcal{K}}{4\pi}\int_{M_r} \mathrm{tr}(\mathcal{F}_b \wedge \mathcal{F}_b) - i\mathcal{K}\int_{C_r} \mathrm{tr}((\alpha_r - w\gamma_r)\mathcal{F}_b). \tag{2.84}$$

(Here we focus on the integral on the right hand side of the defect.) The combination $\alpha_r - w\gamma_r$ under the trace came from the monodromy of the complexified gauge field $\mathcal{A}_b = A + w\phi$, where w is some complex number with non-zero imaginary part. (In physical theory, $w = i\sin\vartheta$.) Such an operator clearly corresponds to a Chern-Simons monodromy operator of weight $\lambda^\circ = \mathcal{K}(\alpha - w\gamma)$, which generically is complex. Now, the problem with such an operator in the physical theory is that the right hand side of (2.84) contains an integral of $i\mathcal{K}w\mathrm{tr}(\gamma F)$ over C, which cannot have any interpretation in the bulk physical theory, since w is not real. (Comparing e.g. to (2.78), we could say that this insertion corresponds to $\eta = w\mathcal{K}\gamma$, which is not an element of the real Lie algebra.) What one should really do in the physical theory is to write the action as a four-dimensional integral of the density $\mathrm{tr}(F \wedge F)$, with gauge field non-complexified, plus the three-dimensional integral of a three-form which can be found on the right hand side of equation (2.40). In the presence of a surface operator, one should omit C from the four-dimensional integral of $\mathrm{tr}(F \wedge F)$, and the knot K from the boundary integral of the just-mentioned three-form. In the bulk, this gives an ordinary surface operator of the sort reviewed in section 2.3.4.1. However, it is not completely clear whether with this definition the intersection of the operator with the defect at $y = 0$ can be made Q-invariant, and to what Chern-Simons weight it would correspond. In the S-dual description of the theory in section 2.4, we will find natural half-BPS surface operators with non-zero $\gamma^\vee$, and the Chern-Simons weight will not depend on this parameter. So we would expect that in the physical theory, Q-invariant surface operators with $\gamma \neq 0$, intersecting the boundary, do exist, and that γ plays much the same role as β – that is, it only deforms the integration contour. But this point is not completely clear.

Finally, turning on the parameter η of the surface operator corresponds to adding an abelian Wilson insertion along the line K, where the surface operator crosses the $y = 0$ hyperplane. Naively, this happens because of the "identity" $\exp(i\eta \int_{C\cap M_r} F) = \exp(i\eta \oint_K A)$ where A is an abelian gauge field with curvature F. We cannot take this formula literally, since $\oint_K A$ is only gauge-invariant mod $2\pi\mathbb{Z}$. But the "identity" is correct for computing classical equations of motion, and thus shifting $\eta_{\ell,r}$ has the same effect on the equations of motion as shifting the electric charges that live on $K = C \cap W$. Note that in presence of the three-dimensional defect the parameter η is lifted from the maximal torus $T^\vee$, and takes values in the dual Cartan subalgebra $\mathfrak{t}^*$.

Let us briefly summarize. A surface operator with parameters $(\alpha, \beta, 0, \eta)$, supported on a surface $C = C_\ell \cup C_r$, corresponds in the analytically-continued three-dimensional Chern-Simons theory to a monodromy operator with weight $\lambda^\circ = \mathcal{K}\alpha - \eta^*$. (Recall that a circle denotes the dual with respect to the superinvariant bilinear form $\kappa = \kappa_r - \kappa_\ell$, and a star represents the dual with respect to the positive definite form $\kappa_r + \kappa_\ell$.) Let λ_ℓ and λ_r be the parts of the weight, lying in the Cartan of the left and right bosonic gauge groups, respectively. Then, more explicitly,

$$\begin{aligned} \lambda_\ell &= -\mathcal{K}\alpha_\ell^* + \eta_\ell \, , \\ \lambda_r &= \mathcal{K}\alpha_r^* - \eta_r . \end{aligned} \tag{2.85}$$

We have set γ to zero, since its role is not completely clear. For a given weight λ, we have a freedom to change α and η, while preserving $\lambda_{\ell,r}$. So a given line operator in the Chern-Simons theory can be represented by a family of surface operators in the four-dimensional theory.

Now let us specialize for a moment to the operators of type $(\alpha, 0, 0, 0)$. The action of the Weyl group on α, together with the large gauge transformations which shift α by an element of the coroot lattice[26] Γ_w^* of the bosonic subalgebra, generate the action of

[26]For simplicity, here we restrict to a simply-connected gauge group, where the cocharacter lattice is the coroot lattice.

the affine Weyl group $\widehat{\mathcal{W}}_1 = \mathcal{W} \ltimes \Gamma_w^*$ at level 1. Equivalently, on the quantum-corrected weights λ these transformations act as the affine Weyl group $\widehat{\mathcal{W}}_{\mathcal{K}} = \mathcal{W} \ltimes \mathcal{K}\Gamma_w^*$ at level[27] $\mathcal{K}$. Though the description by surface operators makes sense for arbitrary λ, let us look specifically at the integral weights $\lambda \in \Gamma_w$. For generic $\mathcal{K}$, the subgroup of $\widehat{\mathcal{W}}_{\mathcal{K}}$ which maps the weight lattice to itself consists only of the ordinary Weyl transformations. Therefore, the space of integral weights modulo the action of $\widehat{\mathcal{W}}_{\mathcal{K}}$ in this case is the space $\Gamma_w/\mathcal{W}$ of dominant weights of the superalgebra, and the Chern-Simons observables corresponding to these weights are generically all inequivalent. Of course, this is a statement about the analytically-continued theory, which is the only theory that makes sense for generic $\mathcal{K}$. If however $\mathcal{K}$ is a rational number $\mathfrak{p}/\mathfrak{q}$, then there are infinitely many elements of the affine Weyl group, which preserve the integral weight lattice Γ_w. (For example, such are all the transformations from $\widehat{\mathcal{W}}_{\mathfrak{p}} \subset \widehat{\mathcal{W}}_{\mathcal{K}}$.) Modulo these transformations, there is only a finite set of inequivalent integral weights.

For an ordinary bosonic Chern-Simons theory and integer leve,l this can be compared to the well-known three-dimensional result according to which the inequivalent Chern-Simons line operators are labeled by the integrable weights $\Lambda \in \Gamma_w/\widehat{\mathcal{W}}_k$. The connection between the two descriptions is that the weight Λ is integrable at level k if and only if the corresponding quantum corrected weight $\lambda = \Lambda + \rho$ belongs to the *interior* of the fundamental Weyl chamber $\Gamma_w/\widehat{\mathcal{W}}_{\mathcal{K}}$, while the operators with λ belonging to the boundary of the fundamental Weyl chamber decouple in the Chern-Simons. This explains how the four-dimensional description by codimension-two operators with quantum-corrected level $\mathcal{K}$ and weight λ can be equivalent (for integer $\mathcal{K}$ and if the four-dimensional theory is specialized to an appropriate class of observables) to the analogous three-dimensional description by operators defined with ordinary k and Λ. For the case of a supergroup, where the purely three-dimensional description is not completely clear, this discussion supports the view that, similarly to the

[27]By the affine Weyl group at some level $\mathfrak{p}$ we mean the group which acts on the Cartan subalgebra by ordinary Weyl transformations together with shifts by $\mathfrak{p}$ times a coroot. Our terminology is slightly imprecise, since as an abstract group, the affine Weyl group does not depend on the level.

bosonic case, at integer level there is a distinguished theory with only a finite set of inequivalent line operators. One detail to mention is that in the four-dimensional construction, we did not show that the operators with λ lying on the boundary of the affine Weyl chamber decouple from the theory. We do not know for sure if this is true for supergroups in the context of a hypothetical theory with only the distinguished set of line operators. Another caveat is that we worked with the half-BPS surface operators, and therefore our conclusion might not hold for the atypical supergroup representations.

2.3.5 Various Problems

We conclude by emphasizing a few unclear points.

In the four-dimensional construction, we have separately defined Wilson line operators and Wilson-'t Hooft line operators in the 3d defect $W \subset M$. They are parametrized by the same data – at least in the case of typical weights. The Wilson line operators generically have less symmetry. Is it conceivable that they flow in the infrared to Wilson-'t Hooft line operators with enhanced symmetry?

For an atypical weight, there are many possible Wilson operators but only one half-BPS Wilson-'t Hooft operator. This in itself is no contradiction. But in the S-dual description of section 2.4 (see in particular section 2.4.4.5), we will find several half-BPS line operators for a given atypical weight. The counterparts of this on the electric side seem to be missing.

One more technical puzzle arises for type II superalgebras. The half-BPS Wilson-'t Hooft operators seem to be well-defined for an arbitrary integral weight λ, at least if it is typical, even though in some cases there is no corresponding representation. (For a weight to correspond to a finite-dimensional representation, the weight should satisfy an extra constraint, as was recalled in section 2.3.1.1.) There is no contradiction, but it is perhaps a surprise to apparently find half-BPS Wilson-'t Hooft line operators that do not correspond to representations.

Additional line operators can presumably be constructed by coupling the bulk fields to some quantum mechanical degrees of freedom that live only along the line operator. This may help in constructing additional half-BPS line operators. Perhaps it is important to understand better the BWB quantum mechanics for atypical weights.

2.4 Magnetic Theory

2.4.1 Preliminaries

In this section we explore the S-dual description of our theory. Throughout this section the reader may assume that the theory considered corresponds to the supergroup $SG = \mathrm{U}(m|n)$. This means in particular that the maximal bosonic subgroup $SG_{\overline{0}} = \mathrm{U}(m) \times \mathrm{U}(n)$ is simply-laced. Some minor modifications that arise for other supergroups will be discussed in section 2.5.

We would like to recall how the supersymmetries and various parameters transform under S-duality. It is convenient to look again on the Type IIB picture. Under the element

$$\mathcal{M} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{2.86}$$

of the S-duality group $SL(2, \mathbb{Z})$, the coupling constant of the theory transforms as

$$\tau \to \frac{a\tau + b}{c\tau + d}. \tag{2.87}$$

The supersymmetries of the Type IIB theory transform according to

$$\varepsilon_1 + i\varepsilon_2 \to \mathrm{e}^{i\alpha/2}(\varepsilon_1 + i\varepsilon_2)\,, \tag{2.88}$$

where $\alpha = -\arg(c\tau + d)$. In particular, for the supersymmetries that are preserved by the

D3-brane we can use the relation (2.3) to rewrite this as

$$\varepsilon_1 \to \exp\left(-\frac{1}{2}\alpha\,\Gamma_{0123}\right)\varepsilon_1\,, \tag{2.89}$$

in Lorentz signature. In [14] this relation was derived from the field theory point of view.

Under the duality transformation $\mathcal{M}$, the charges of the fivebranes transform as

$$(p\;\;q) \to (p\;\;q)\mathcal{M}^{-1}\,, \tag{2.90}$$

where $(p,q) = (1,0)$ for the NS5-brane and $(p,q) = (0,\pm 1)$ for the D5- or $\overline{\text{D5}}$-brane. For future reference we describe the supersymmetries that are preserved by a defect consisting of a general (p,q)-fivebrane. The supersymmetries preserved by such a brane, stretched in the 012456 directions, are given by the same formula as in (2.4), where now

$$\vartheta = \arg(p\tau + q). \tag{2.91}$$

Equation (2.4) can be rewritten in a more convenient form

$$\varepsilon_1 + i\varepsilon_2 = i\mathrm{e}^{i\vartheta}\,\Gamma_{012456}(\varepsilon_1 - i\varepsilon_2). \tag{2.92}$$

Under the S-duality, ϑ is shifted by angle $\alpha = -\arg(c\tau + d)$, so one can see that equation (2.92) indeed transforms covariantly, if the supersymmetries are mapped as in equation (2.88). The twisting parameter $t = -\mathrm{e}^{-i\vartheta}$ is multiplied by $\mathrm{e}^{-i\alpha}$, that is,

$$t \to t\frac{c\tau + d}{|c\tau + d|}. \tag{2.93}$$

The canonical parameter $\mathcal{K}$ of the bulk theory was defined in equation (2.36). In terms of the gauge coupling and the parameter t,

$$\mathcal{K} = \frac{\theta_{\mathrm{YM}}}{2\pi} + \frac{4\pi i}{g_{\mathrm{YM}}^2}\frac{t - t^{-1}}{t + t^{-1}}. \tag{2.94}$$

For the special case that t corresponds to the supersymmetry preserved by the D3-NS5 system, this reduces to eqn. (2.7). Under S-duality, the canonical parameter transforms [14] in the same way as the gauge coupling,

$$\mathcal{K} \to \frac{a\,\mathcal{K} + b}{c\,\mathcal{K} + d}. \tag{2.95}$$

Let us specialize to the case of interest. The basic S-duality transformation that exchanges electric and magnetic fields is usually described (for simply-laced groups) as $\tau \to -1/\tau$, but this does not specify it uniquely, since it does not determine the sign of the matrix $\mathcal{M}$ of eqn. (2.86). We fix the sign by taking

$$\mathcal{M} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{2.96}$$

This means, according to eqn. (2.90) that an NS5-brane, with $(p, q) = (1, 0)$, transforms to a $\overline{\mathrm{D5}}$-brane, with $(p, q) = (0, -1)$, so that according to eqn. (2.91), $\vartheta^\vee = \pi$ and $t^\vee = 1$. Then from the definition (2.94) of the canonical parameter, it follows that $\mathcal{K}^\vee = \frac{\theta^\vee_{\mathrm{YM}}}{2\pi}$.

Unlike in the electric theory, the twisted action is very simple on the dual magnetic side. As in the purely bosonic case [6], the action is Q-exact except for a multiple of the instanton number (see Appendix C for a detailed explanation). In Euclidean signature, we have

$$I_{\mathrm{magnetic}} = \frac{i\theta^\vee_{\mathrm{YM}}}{8\pi^2}\int \mathrm{tr}\,(F \wedge F) + \{Q, \dots\}. \tag{2.97}$$

If we set

$$q = \exp(-i\theta^\vee_{\mathrm{YM}})\,, \tag{2.98}$$

then the dependence of the theory on q is easily described: a solution of the localization equations of instanton number n makes a contribution $\pm q^n$ to the path integral. (The sign is given by the sign of the fermion determinant.) This simple result arises in the usual way because of cancellation between bosonic and fermionic fluctuations around a solution of the localization equations. If therefore the instanton number is integer-valued and is bounded above and below in all solutions of the localization equations,[28] then the path integral is a Laurent polynomial in q with integer coefficients, namely

$$Z = \sum_n a_n q^n, \tag{2.99}$$

where a_n is the number of solutions (weighted by sign) of instanton number n.

It is straightforward to express q in terms of the parameters of the electric theory. As explained above, in the magnetic theory $\mathcal{K}^\vee = \theta_{\mathrm{YM}}^\vee/2\pi$; also, according to (2.95), $\mathcal{K}^\vee = -1/\mathcal{K}$. So

$$\theta_{\mathrm{YM}}^\vee = -2\pi/\mathcal{K}\,, \tag{2.100}$$

and hence

$$q = \exp\left(\frac{2\pi i}{\mathcal{K}}\right). \tag{2.101}$$

For an ordinary (simple, compact, and simply-laced) bosonic group, this is the standard variable in which the quantum knot invariants are conveniently expressed, and for a supergroup it is the closest analog. These matters were described in section 2.2.6.

We now proceed to describing the localization equations and the boundary conditions in the magnetic theory, leaving many technical details for Appendix C. Some relevant aspects of the gauge theory have been studied in [21]. The details depend on the difference of the numbers of D3-branes on the two sides of the D5-brane. We describe different cases in the

[28] One expects the instanton number to be bounded in any solution, though this has not been proved. However, the claim that the instanton number is integer-valued is oversimplified; for example, if the gauge group is simply-connected or M is contractible, the instanton number takes values in $\mathbb{Z} + c$ where c is a constant determined by the boundary conditions. In such a situation, the partition function is q^c times a Laurent polynomial in q.

subsequent sections.

2.4.2 Gauge Groups Of Equal Rank

In the case of an equal number of D3-branes on the two sides, the effective theory is a $\mathrm{U}(n)$ super Yang-Mills theory in the whole four-dimensional space, with an additional three-dimensional matter hypermultiplet localized on the defect, at $x_3 = 0$. This hypermultiplet comes from the strings that join the D5-brane and the D3-branes, and therefore it transforms in the fundamental of the $\mathrm{U}(n)$ gauge group. Under the global bosonic symmetries $\mathrm{U} = \mathrm{SO}(1,2) \times \mathrm{SO}(3)_X \times \mathrm{SO}(3)_Y$, the scalars Z^A of the hypermultiplet transform as a doublet $(\mathbf{1},\mathbf{2},\mathbf{1})$, and the fermions $\zeta^{\alpha\dot{A}}$ transform as $(\mathbf{2},\mathbf{1},\mathbf{2})$. The bulk fields have discontinuities at $x_3 = 0$ as a result of their interaction with the defect. For example, the equations of motion of the gauge field, in Euclidean signature, can be deduced from the action

$$-\frac{1}{2(g^\vee_{\mathrm{YM}})^2}\int \mathrm{d}^4x \,\mathrm{tr}\, F^2_{\mu\nu} + \frac{1}{(g^\vee_{\mathrm{YM}})^2} I^\vee_{\mathrm{hyp}}. \tag{2.102}$$

(In the magnetic description, the topological term $\int \mathrm{tr}\, F\wedge F$ is integrated over all of $\mathbb{R}^4$ and so does not affect the equations of motion.) The equations of motion that come from the variation of this action have a delta-term supported on the defect,

$$D_3 F^m_{3i} - \frac{1}{2}\delta(x_3) J^m_i = 0\,, \tag{2.103}$$

where $J^i_m = \delta I^\vee_{\mathrm{hyp}}/\delta A^m_i$ is the current.[29] The delta-term in this equation means that the gauge field has a cusp at $x_3 = 0$, so that F_{3i} has a discontinuity:

$$F^m_{3i}|^{\pm} = \frac{1}{2} J^m_i. \tag{2.104}$$

[29] Indices m, n continue to denote gauge indices, although now the gauge group is just one copy of $\mathrm{U}(n)$ throughout $\mathbb{R}^4$. Gauge indices are raised and lowered with the positive-definite Killing form $\delta_{mn} = -\mathrm{tr}(T_m T_n)$.

Here and in what follows we use the notation $\varphi|^{\pm} = \varphi(x_3 + 0) - \varphi(x_3 - 0)$ for the jump of a field across the defect. By supersymmetry, this discontinuity equation can be extended to a full three-dimensional current supermultiplet. The most important for us will be the lowest component of the current multiplet, which governs the discontinuity of the bulk scalar fields X^a:

$$X^{am}\big|_{-}^{+} = \frac{1}{2}\mu^{am}\,, \tag{2.105}$$

where the hyperkahler moment map for the defect hypermultiplets is

$$\mu^a_m = \overline{Z}_A \sigma^{aA}_B T_m Z^B. \tag{2.106}$$

(The other bulk scalar fields $Y^{\dot a}$ are continuous at $x_3 = 0$.)

Now we turn to the twisted theory. Recall, that for twisting we use an SO(4) subgroup of the R-symmetry, which on the defect naturally reduces to $\mathrm{SO}(3)_X$. Thus, the hypermultiplet scalars Z^A become spinors Z^α under the twisted Lorentz group. They are invariant under $\mathrm{SU}(2)_Y$, and therefore have ghost number zero. The hypermultiplet fermions $\zeta^{\alpha\dot A}$ remain spinors. Since they also transform as a doublet of $\mathrm{SU}(2)_Y$, we can expand them in the basis given by the vectors u and v of eqns. (2.28) and (2.27) (where now we take $\vartheta^\vee = \pi$):

$$\begin{aligned}\zeta^{\dot A} &= iu^{\dot A}\zeta_u + iv^{\dot A}\zeta_v\,,\\ \overline{\zeta}^{\dot A} &= iu^{\dot A}\overline{\zeta}_u + iv^{\dot A}\overline{\zeta}_v.\end{aligned} \tag{2.107}$$

The u- and v-components of ζ and $\overline{\zeta}$ have ghost number plus or minus one, respectively.

As usual, the path integral can be localized on the solutions of the BPS equations $\{Q, \xi\} = 0$, where ξ is any fermionic field. The resulting equations for the bulk fermions were partly described in eqn. (2.30). At $t^\vee = 1$, they have a particularly simple form,

$$\begin{aligned} &F - \phi \wedge \phi + \star \mathrm{d}_A \phi = \frac{1}{2} \star (\delta_W \wedge \mu) \ , \\ &D_\mu \phi^\mu = 0. \end{aligned} \tag{2.108}$$

Here $\delta_W = \delta(x_3)\mathrm{d}x_3$ is Poincaré dual to the three-manifold W on which the defect is supported. The delta function term on the right hand side of the first equation in (2.108) is related to the discontinuity (2.105) of the 1-form field ϕ. There is no such term in the second equation, because the only field whose x_3 derivative appears in this equation is ϕ_3; this field originates as a component of $Y^{\dot{a}}$, and is continuous at $x_3 = 0$. The condition that $\{Q, \xi\} = 0$ for all ξ also leads to conditions on the ghost field σ:

$$D_\mu \sigma = [\phi_\mu, \sigma] = [\overline{\sigma}, \sigma] = 0. \tag{2.109}$$

These equations say that the infinitesimal gauge transformation generated by σ is a symmetry of the solution. In this chapter we generally do not consider reducible solutions, so we generally can set σ to 0.

We also should consider the condition $\{Q, \xi\} = 0$ where ξ is one of the defect fermions. For the u-component of the fermions that are defined in eqn. (2.107), $\{Q, \xi\}$ equals the variation of the defect fields under the gauge transformation generated by σ, so the condition for it to vanish, when combined with (2.109) says that the full configuration including the fields on the defect is σ-invariant. More important for us will be the condition $\{Q, \xi\} = 0$ for the v-components:

$$\not{D} Z + \phi_3 Z = 0 \, , \quad \not{D} \overline{Z} - \overline{Z} \phi_3 = 0. \tag{2.110}$$

Eqns. (2.108) and (2.110) together give the condition for a supersymmetric configuration.

2.4.3 Gauge Groups Of Unequal Rank

Now consider the case that the number of D3 branes jumps from n to $n+r$, $r>0$, upon crossing the D5-brane. The gauge groups on the left and on the right are $\mathrm{U}(n)$ and $\mathrm{U}(n+r)$, and will be denoted by G_ℓ and G_r, respectively. The behavior along the defect has been described in [21]. In contrast to the case $r=0$, there are no hypermultiplets supported along the defect at $y=0$. What does happen is different according to whether $r=1$ or $r>1$. We first describe the behavior for $r>1$.

The main feature of this problem is that some of the bulk fields have a singular behavior (known as a Nahm pole singularity) near $y=0$. Assuming that r is positive, the singular behavior arises as one approaches $y=0$ from above. To describe the singularity, we first pick a subgroup $H=\mathrm{U}(n)\times\mathrm{U}(r)\subset\mathrm{U}(n+r)$, and we set $H'=\mathrm{U}(n)\times\mathrm{U}(1)$, where $\mathrm{U}(1)$ is the center of the second factor in H. The singularity will break $G_r=\mathrm{U}(n+r)$ to H'. The fields with a singular behavior are the scalar fields that we have called X^a in the untwisted theory or as ϕ_i in the twisted theory. The behavior of ϕ as y approaches 0 from above is

$$\phi_i=\frac{t_i}{y}+\dots\,, \tag{2.111}$$

where the ellipsis represent less singular terms, and the matrices t_i represent an irreducible embedding of $\mathfrak{su}(2)$ into the Lie algebra $\mathfrak{u}(r)$ of the second factor of $H=\mathrm{U}(n)\times\mathrm{U}(r)$. Thus the matrices t_i are $(n+r)\times(n+r)$ matrices that vanish except for a single $r\times r$ block, as shown here for $n=2$, $r=3$:

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & * & * & * \\ 0 & 0 & * & * & * \\ 0 & 0 & * & * & * \end{pmatrix}. \tag{2.112}$$

These matrices are traceless, so their nonzero blocks are actually valued in $\mathfrak{su}(r)\subset\mathfrak{u}(r)$.

The Nahm pole singularity breaks the gauge symmetry for $y > 0$ from $\mathrm{U}(n+r)$ to $H' = \mathrm{U}(n) \times \mathrm{U}(1)$, and there is to begin with a $G_\ell = \mathrm{U}(n)$ gauge symmetry for $y < 0$. There is therefore a $\mathrm{U}(n)$ gauge symmetry on both sides of the defect, and the condition obeyed by the $\mathrm{U}(n)$ gauge fields is just that they are continuous at $y = 0$, making a $\mathrm{U}(n)$ gauge symmetry throughout the whole spacetime. On the other hand, the fields supported at $y > 0$ that do not commute with the Nahm pole singularity acquire very large masses near $y = 0$, and they vanish for $y \to 0$. (This statement applies to fields in the adjoint representation of $\mathfrak{su}(r)$ and also to fields in the bifundamental of $\mathrm{U}(n) \times \mathrm{U}(r)$.) To finish describing the gauge theory of the defect, we must explain the behavior at $y = 0$ of the fields in the second factor of $H' = \mathrm{U}(n) \times \mathrm{U}(1)$. These fields make up a single vector multiplet, which obeys what we might call Dirichlet boundary conditions (the gauge fields A_i and scalars $Y^{\dot{a}}$ in this multiplet obey Dirichlet boundary conditions, while the scalars X^a obey Neumann boundary conditions; these conditions are extended to the fermions in a fashion determined by supersymmetry).

For $r = 1$, this description requires some modification, because $\mathfrak{su}(1) = 0$ and accordingly the matrices t_i vanish. Still, the defect breaks the $G_r = \mathrm{U}(n+1)$ gauge symmetry for $y > 0$ to a subgroup $H' = \mathrm{U}(n) \times \mathrm{U}(1) \subset \mathrm{U}(n+1)$. Just as at $r > 1$, the $\mathfrak{u}(n)$-valued gauge fields on the two sides of the defect fit smoothly into continuous $\mathfrak{u}(n)$-valued fields throughout the whole spacetime. For $y > 0$, the gauge fields valued in the orthocomplement of $\mathfrak{u}(n)$ obey the same Dirichlet boundary conditions described at the end of the last paragraph.

So far, we have described this construction as if the matrices t_i in eqn. (2.111) are just constant matrices. This makes sense if $W = \mathbb{R}^3$, but in general, we must recall that in the twisted theory on $M = W \times \mathbb{R}$, $\phi = \sum_i \phi_i \mathrm{d}x^i$ transforms as a 1-form along W. The proper interpretation of the Nahm pole singularity in this general setting is as follows (see section 3.4 of [6]; the considerations there carry over to the present case without essential change). The $\mathfrak{u}(r)$ bundle along W must be derived from a spin bundle S_W via a homomorphism $\varrho : \mathfrak{su}(2) \to \mathfrak{u}(k)$ defined by the t_i. The restriction to $W \times \{y = 0\}$ of the $\mathfrak{u}(r)$-valued part of the gauge field is the Levi-Civita connection ω of S_W, embedded in $\mathfrak{su}(r)$ via ϱ. We describe this by saying that when restricted to $y = 0$, the $\mathfrak{u}(r)$-valued part of the gauge field A is $A_{\mathfrak{u}(r)} = \varrho(\omega)$.

2.4.3.1 The Framing Anomaly

It is now possible to make an interesting check of the relationship between Chern-Simons theory of $\mathrm{U}(n|n+r)$ and the defect theory just described. Here we will be rather brief, assuming that the reader is familiar with the description of the one-sided case in section 3.5.3 of [6]. Recall that in general the partition function of Chern-Simons theory on a three-manifold W is not quite a topological invariant of W; W must be endowed with a framing (or more precisely a two-framing [55]) to define this partition function. A framing is a trivialization of the tangent bundle of W. Under a unit change of framing, the partition function acquires a factor [4]

$$\exp(2\pi i c \,\mathrm{sign}(k))/24)\,, \tag{2.113}$$

where c is the central charge of the relevant current algebra at level k. For a compact simple gauge group G this is $c = k \dim G/(k + h\ \mathrm{sign}(k))$, where h is the dual Coxeter number of G. We will assume that the same formula for c applies, at least modulo an integer, for a simple supergroup SG, which in our case will be $\mathrm{SU}(n|n+r)$:

$$c = \frac{k\,\mathrm{sdim}\, SG}{k + h_{\mathfrak{sg}}\mathrm{sign}(k)} \mod \mathbb{Z}. \tag{2.114}$$

This is a non-trivial assumption, since some of the standard arguments do not apply for supergroups, as it is described in Appendix E of [1]. (Replacing $\mathrm{SU}(n|n+r)$ by $\mathrm{U}(n|n+r)$, which is isomorphic locally to $\mathrm{SU}(n|n+r) \times \mathrm{U}(1)$, shifts c by 1, which will not be important as we will only study c mod $\mathbb{Z}$. So the following discussion will be phrased for the simple supergroup $\mathrm{SU}(n|n+r)$, rather than $\mathrm{U}(n|n+r)$.) It is useful to factor (2.113) as follows:

$$\exp\left(2\pi i\,\mathrm{sign}(k)\,\mathrm{sdim}\,SG/24\right)\cdot q^{-h_{\mathfrak{sg}}\,\mathrm{sdimSG}/24}. \tag{2.115}$$

Perturbation theory is an expansion in powers of $1/\mathcal{K}$, with an ℓ-loop diagram making a contribution of order $\mathcal{K}^{1-\ell}$. Accordingly, the exponent $2\pi i\,\mathrm{sign}(k)\,\mathrm{sdim}\,SG/24$ in the first factor in (2.115), being invariant under scaling of k, is a 1-loop effect. Since it is not analytic in $\mathcal{K}$, we cannot hope to reproduce it from four dimensions. If this factor – or a similar one that arises if c is shifted by an integer – appears in a purely three-dimensional construction, then it must appear in a comparison between the relevant measures in three and four dimensions, as discussed in section 2.2.6 above and in section 3.5.3 of [6]. However, the second factor in (2.115), which is a simple power of q, comes from diagrams with ≥ 2 loops and can be reproduced from four dimensions.

As in [6], this factor arises from a subtlety in the definition of instanton number in the presence of the Nahm pole. The condition that along $W\times\{y=0\}$, $A_{\mathfrak{u}(r)}=\varrho(\omega)$ means that the instanton number, defined in the obvious way from the integral $\int_{M_\ell}\mathrm{Tr}\,F\wedge F+\int_{M_r}\mathrm{Tr}\,F\wedge F$, is not a topological invariant. If one varies the metric of W, the second term picks up a variation from the change in ω. To compensate for this, one must add to the instanton number a multiple of the Chern-Simons invariant of ω, but this is only gauge-invariant (as a real number) once we pick a framing on W. From the viewpoint of the dual magnetic description, that is why Chern-Simons theory on W requires a framing of W. To adapt the analysis of [6] to the present problem, we simply proceed as follows. In the $\mathrm{U}(n|n+r)$ case, the Nahm pole is embedded in a $\mathfrak{u}(r)$ subalgebra, and therefore the framing-dependence that is introduced when we define the instanton number for this problem is independent of n and is the same as it is for the one-sided problem with $n=0$ and gauge group $\mathrm{U}(r)$. Hence, to obtain in the magnetic description the expected factor $q^{-h_{\mathfrak{sg}}\mathrm{sdim}\,G/24}$ in the framing dependence, we need the identity

$$h_{\mathfrak{su}(n|n+r)} \operatorname{sdim} SU(n|n+r) = h_{\mathfrak{su}(r)} \dim SU(r). \tag{2.116}$$

This is true because $\operatorname{sdim} SU(n|n+r)$ is independent of n and likewise $h_{\mathfrak{su}(n|n+r)}$ is independent of n. See Table 2.2.

2.4.4 Line And Surface Operators In The Magnetic Theory

Our next goal is to identify the S-duals of the line and surface operators that we have found on the electric side. We use the fact that we know how S-duality acts on the bulk surface operators. For an "electric" surface operator, the magnetic dual [53] has parameters $(\alpha^\vee, \beta^\vee, \gamma^\vee, \eta^\vee) = (\eta, |\tau|\beta^*, |\tau|\gamma^*, -\alpha)$, where τ is the gauge coupling constant. This determines the singularity of the fields along the operator in the bulk, away from the three-dimensional defect. We still have to find the model solution which describes the behavior of the fields near the end of the surface operator at $y = 0$. This will be the main subject of the present section.

In bulk, for a surface operator with parameters $(\alpha, \beta, \gamma, \eta)$, the parameters α and η are both periodic. In the presence of a defect, this is no longer the case. In the electric description, η is not a periodic variable on a D3-brane that ends on (or intersects) an NS5-brane. Shifting η by an integral character would add a unit of charge along the defect. Dually to this, for the D3-D5 system, in the case of a surface operator with parameters $(\alpha^\vee, \beta^\vee, \gamma^\vee, \eta^\vee)$, $\alpha^\vee$ is not a periodic variable. In the model solutions that we construct below, if $\alpha^\vee$ is shifted by an integral cocharacter (of $G^\vee$), then the solution is unchanged in the bulk up to a gauge transformation, but is modified along the defect.

It follows from this that once we construct model solutions for surface operators with parameters $(\alpha^\vee, \beta^\vee, \gamma^\vee)$, we can trivially construct magnetic line operators. We return to this in section 2.4.5.

2.4.4.1 Reduction Of The Equations

We focus first on the case of gauge groups of equal rank, as described in section 2.4.2. The discussion can be transferred to the unequal rank case in a straightforward way, and we shall comment on this later.

To give a definition of a surface operator whose support intersects the three-dimensional defect, we have to find a model solution of the localization equations (2.108) and (2.110) for the fields near the surface C and near the hyperplane $y = 0$. The classical solution does not depend on the two-dimensional theta-angles $\eta^\vee$, so we label it by three parameters $(\alpha^\vee, \beta^\vee, \gamma^\vee)$. We consider a surface operator stretched along $C = \mathbb{R}_{x^0} \times \mathbb{R}_y$ in $\mathbb{R}^4$,and look for a time-independent, scale-invariant solution. We aim to construct a model solution that is 1/2-BPS, that is, it preserves the four supersymmetries (2.69) and (2.70). It should also be invariant under the $\mathrm{SO}(3)_Y$ subgroup of the R-symmetry groups. The symmetries allow us to considerably reduce the localization equations. An analogous problem in the one-sided theory was considered in section 3.6 of [6], where the reader can find many details which we do not repeat here.

First of all, for an irreducible solution the field σ is zero, and therefore, by $\mathrm{SO}(3)_Y$ symmetry, ϕ_3 should also vanish. The Q-invariance together with $\mathrm{SO}(3)_Y$ symmetry makes the solution invariant under the first pair of supersymmetries (2.69). Using the explicit formulas for the transformations (2.229), one can also impose invariance under the second pair of supersymmetries (2.70). For $t^\vee = 1$, which is the case in the magnetic theory, this fixes A_0 to be zero. The reduced localization equations can be written in a concise form, after introducing some convenient notation. Following [6], we define three operators

$$
\begin{aligned}
\mathcal{D}_1 &= 2D_{\overline{z}}\,, \\
\mathcal{D}_2 &= D_3 - i\phi_0\,, \\
\mathcal{D}_3 &= 2\phi_z\,,
\end{aligned}
\tag{2.117}
$$

where $z = x_1 + ix_2$ is a complex coordinate, $\phi_z = (\phi_1 - i\phi_2)/2$ is the z-component of ϕ, and $D_{\overline{z}}$ and D_3 are covariant derivatives. We also denote the components of the bosonic spinor field Z^α as $Z \equiv Z^1$ and $\widetilde{Z} \equiv (Z^2)^\dagger$. For simplicity, we assume the gauge group $G^\vee$ to be $\mathrm{U}(n)$. Then the components of the moment map (2.106) can be written as

$$\mu_0 = i(\widetilde{Z}^\dagger \otimes \widetilde{Z} - Z \otimes Z^\dagger)\,, \quad \mu_z = -iZ \otimes \widetilde{Z}. \tag{2.118}$$

With this notation, the reduced localization equations are

$$\begin{aligned} &[\mathcal{D}_1, \mathcal{D}_2] = 0\,, \quad [\mathcal{D}_3, \mathcal{D}_1] = 0\,, \quad [\mathcal{D}_2, \mathcal{D}_3] - \mu_z \delta(y) = 0\,, \\ &\mathcal{D}_1 Z = \mathcal{D}_1 \widetilde{Z} = 0\,, \end{aligned} \tag{2.119}$$

together with

$$\sum_i [\mathcal{D}_i, \mathcal{D}_i^\dagger] + i\mu_0 \delta(y) = 0. \tag{2.120}$$

The space of fields in which we look for the solution is the space of continuous connections on $\mathbb{R}^4 \setminus C$, and Higgs fields with an arbitrary discontinuity across the hyperplane $y = 0$. (The fields should also be vanishing at infinity.) The correct discontinuity (2.105) is enforced by the delta-terms in the localization equations. To put the real and imaginary parts A_3 and ϕ_0 of the connection in $\mathcal{D}_2$ on equal footing, let us also allow A_3 to have an arbitrary discontinuity across $y = 0$, and to compensate for this, we divide the space of solutions by the gauge transformations, which are allowed to have a cusp across the defect hyperplane.

The analysis of these equations in the one-sided case in [6] was based on the fact that the equations (2.119) are actually invariant under complex-valued gauge transformations, not just real-valued ones. One can try to solve the equations in a two-step procedure in which one first solves eqn. (2.119) and then tries to find a complex-valued gauge transformation to a set of fields that obeys (2.120) as well.

Though we could follow that strategy here as well, we will instead follow a more direct approach. We are motivated by the fact that the basic surface operator in the absence of any defect or boundary is described by a trivial abelian solution. In the one-sided problem, one requires a Nahm pole along the boundary and therefore the full solution is always irreducible. However, in the two-sided case with equal ranks, there is no Nahm pole. Is it too much to hope that we can find something interesting by taking simple abelian solutions for $y < 0$ and $y > 0$, somehow glued together along $y = 0$?

2.4.4.2 Some "Abelian" Solutions

We look for a model solution for a surface operator with parameters $(\alpha^\vee, 0, 0)$, and initially we assume $\alpha^\vee$ regular. Since we take $\beta^\vee = \gamma^\vee = 0$, we look for a model solution invariant under the SO(2) group of rotations in the 12-plane, and under the $\mathrm{SO}(2)_X$ subgroup of the R-symmetry. Accordingly, the field ϕ_z should vanish. Indeed, the $\mathrm{SO}(2)_X$ acts on ϕ_z by multiplication by a phase. In a fully non-abelian solution, this phase could possibly be undone by a gauge transformation, but in a solution that is abelian away from $y = 0$ – as we will assume here – that is not possible and ϕ_z must vanish. Therefore, from the discontinuity equation for ϕ_z it follows that either Z or $\widetilde{Z}$ should vanish. So for definiteness, assume that $\widetilde{Z} = 0$ and $Z \neq 0$.

For now we focus on irreducible solutions, for which the gauge group along K is broken completely. We postpone the discussion of reducible solutions.

A simple abelian solution of the localization equations would be $A = \alpha^\vee \cos\varphi \mathrm{d}\theta$, $\phi = \alpha^\vee \mathrm{d}x^0/r'$. For $y \to \infty$, ϕ vanishes, and A approaches the simple surface operator solution $\alpha^\vee \mathrm{d}\theta$ for $y \to +\infty$ ($\theta = 0$) or $-\alpha^\vee \mathrm{d}\theta$ for $y \to -\infty$ ($\theta = \pi$). However, we want a solution in which A will approach independent limits $\alpha^\vee_\ell \mathrm{d}\theta$ and $\alpha^\vee_r \mathrm{d}\theta$ for $y \to -\infty$ and $y \to +\infty$. Also we want to allow for the possibility that a gauge transformation by a constant matrix g has to be made to match the solutions for $y<0$ and $y>0$. So we try

$$
\begin{aligned}
y>0: \quad & A=\alpha_r^\vee \cos\varphi\, \mathrm{d}\theta\,, & \phi &= \alpha_r^\vee \frac{\mathrm{d}x^0}{r'}\,,\\
y<0: \quad & A=-g\alpha_\ell^\vee g^{-1} \cos\varphi\, \mathrm{d}\theta\,, & \phi &= -g\alpha_\ell^\vee g^{-1} \frac{\mathrm{d}x^0}{r'}\,.
\end{aligned} \tag{2.121}
$$

We also have to impose the discontinuity equation $\phi_0\big|^{\pm}=\frac{i}{2}(\widetilde{Z}^\dagger\otimes\widetilde{Z}-Z\otimes Z^\dagger)$. Note first of all that taking the trace of this gives $i(\mathrm{tr}(\alpha_r^\vee)+\mathrm{tr}(\alpha_\ell^\vee))=r'(|Z|^2-|\widetilde{Z}|^2)/2$. Therefore, the choice of whether Z or $\widetilde{Z}$ is non-zero is determined by the sign of the combination of parameters on the left hand side of this equation. We assume this combination to be positive, and take

$$
Z=\frac{\mathtt{v}}{\sqrt{z}}, \tag{2.122}
$$

where $\mathtt{v}$ is some constant vector. We have taken Z to be holomorphic to satisfy $\mathcal{D}_1 Z=0$ (this is one of the localization equations, eqn. (2.119)). Note that A does not appear in this equation, since it vanishes at $y=0$, so the formula for Z does not depend on $\alpha_\ell^\vee$ or $\alpha_r^\vee$. Also, (2.122) means that Z has a monodromy -1 around the knot, which in this description is located at $z=0$. So we have to assume that this monodromy of Z is part of the definition of the surface operator in this magnetic description.

The discontinuity equation now becomes

$$
i\alpha_r^\vee+ig\alpha_\ell^\vee g^{-1}=\frac{1}{2}\mathtt{v}\otimes\mathtt{v}^\dagger. \tag{2.123}
$$

This is a set of n^2 equations for a unitary matrix g and a vector $\mathtt{v}$, which are together n^2+n variables. The equations are invariant under the diagonal unitary gauge transformations, which remove n parameters. Therefore, generically one expects to have a finite number of solutions.

The equations can be conveniently formulated as follows. For a given hermitian matrix $N=i\alpha_r^\vee$, find a vector $\mathtt{v}$, such that the hermitian matrix $N'=N-\frac{1}{2}\mathtt{v}\otimes\mathtt{v}^\dagger$ has the same eigenvalues as $M = -i\alpha_\ell^\vee$. Using the identity $\det(X + \mathtt{v} \otimes \mathtt{v}^\dagger) = (1 + \mathtt{v}^\dagger X^{-1}\mathtt{v}) \det(X)$, the characteristic polynomial for N' can be written as

$$\det\left(\mathbb{1} \cdot \lambda - N + \frac{1}{2}\mathtt{v} \otimes \mathtt{v}^\dagger\right) = \det(\mathbb{1} \cdot \lambda - N)\left(1 + \frac{1}{2}\sum_{i=1}^{n} \frac{|\mathtt{u}_i^\dagger \mathtt{v}|^2}{\lambda - \lambda_i}\right), \tag{2.124}$$

where $\mathtt{u}_i$ are the eigenvectors of N with eigenvalues λ_i. First let us assume that $\mathtt{u}_i^\dagger \mathtt{v} \neq 0$ for all i. Then the eigenvalues of N' are solutions of the equation

$$1 + \frac{1}{2}\sum_{i=1}^{n} \frac{|\mathtt{u}_i^\dagger \mathtt{v}|^2}{\lambda - \lambda_i} = 0. \tag{2.125}$$

Note that all the eigenvalues of N are distinct – this is the regularity condition for the weight, which says that $\langle \lambda, \alpha_{\bar{0}} \rangle \equiv \langle \Lambda + \rho, \alpha_{\bar{0}} \rangle \neq 0$ for all the superalgebra bosonic roots $\alpha_{\bar{0}}$. By sketching a plot of the function in the left hand side of (2.125), it is easy to observe that the equation has n solutions $\lambda = \lambda_i'$, $i = 1, \ldots, n$. These solutions interlace the eigenvalues λ_i, in the sense that if the λ_i and λ_i' are arranged in increasing order then $\lambda_1' < \lambda_1 < \lambda_2' < \cdots < \lambda_n$. Had we assumed $\widetilde{Z}$ rather than Z to be non-zero, we would have obtained the opposed interlacing condition $\lambda_1 < \lambda_1' < \lambda_2 < \cdots < \lambda_n'$. Moreover, by tuning the n coefficients $|\mathtt{u}_i^\dagger \mathtt{v}|^2$ of the equation, one can in a unique way put these solutions to arbitrary points inside the intervals $(-\infty, \lambda_1)$, (λ_1, λ_2), $\ldots$, $(\lambda_{n-1}, \lambda_n)$, to which they belong. To do this, we simply view eqn. (2.125) as a system of linear equations for the constants $|\mathtt{u}_i^\dagger \mathtt{v}|^2$. The interlacing condition ensures that there is no problem with the positivity of those constants. An important special case is that $|\mathtt{u}_i^\dagger \mathtt{v}|^2 \to 0$ precisely when λ_j' (for $j = i$ or $i \pm 1$) approaches λ_i. The facts we have just stated are used in some applications of random matrix theory; for example, see p. 16 of [56].

We conclude that the equation (2.123) has a solution, which moreover is unique (modulo diagonal gauge transformations), if and only if the eigenvalues of $i\alpha_r^\vee$ and $-i\alpha_\ell^\vee$ are interlaced. Since the eigenvalues of $i\alpha_r^\vee$ and $i\alpha_\ell^\vee$ should be the weights of a dual Wilson-'t Hooft operator on the electric side, we have a reasonable candidate for the dual of such operators when certain inequalities are satisfied. If some of the eigenvalues of $i\alpha_r^\vee$ coincide with eigenvalues of $-i\alpha_\ell^\vee$, then the corresponding components of $Z = \mathtt{v}/\sqrt{z}$ vanish. (We return to this point in section 2.4.4.5.)

If the eigenvalues are not interlaced, the abelian ansatz fails. As a motivation to understand what to do in this case, we will describe a possibly more familiar problem that leads to the same equations and conditions that we have just encountered. We look at the system of N D3-branes intersecting a D5-brane from a different point of view. Instead of studying a surface operator, we look for a supersymmetric vacuum state in which the fields $\vec{X}$ have one asymptotic limit $\vec{X}_\ell$ for $y \to -\infty$ and another limit $\vec{X}_r$ for $y \to +\infty$. Such a vacuum exists for any choice of $\vec{X}_\ell$, $\vec{X}_r$, and is unique up to a gauge transformation. Macroscopically, this vacuum is often just understood by saying that a D3-brane can end on a D5-brane so the value of $\vec{X}$ can jump from $\vec{X}_\ell$ to $\vec{X}_r$ in crossing the D5-brane. Thus, one represents the vacuum by the simple picture of fig. 2.1 of section 2.1.2, but now with the fivebrane in the picture understood as a D5-brane.

Although this picture is correct macroscopically, from a more microscopic point of view, the vacuum of the D3-D5 system is found by solving Nahm's equations for the D3 system, with the D3-D5 intersection contributing a hypermultiplet that appears as an impurity. This has been analyzed in detail in [21]. Let us just consider the case that the branes are separated at $y \to \pm\infty$ only in the X_4 direction, where X_4 corresponds to ϕ_0 in our notation here. A natural ansatz would then be to assume that $X_5 = X_6 = 0$ everywhere. That leads to simple equations. Nahm's equations with $X_5 = X_6 = 0$ just reduce to $\mathrm{d}X_4/\mathrm{d}y = 0$ (for $y \neq 0$), so X_4 is one constant matrix for $y > 0$ and a second constant matrix for $y < 0$. After diagonalizing X_4 for $y > 0$, we can write $X_4 = \alpha_r^\vee$ for $y > 0$, $X_4 = -g\alpha_\ell^\vee g^{-1}$ for $y < 0$, with $\alpha_\ell^\vee, \alpha_r^\vee \in \mathfrak{t}$, $g \in \mathrm{U}(n)$. Finally, in the construction of the vacuum, the jump condition at the location of the hypermultiplet is precisely (2.105).

So in constructing the vacuum assuming that $X_5 = X_6 = 0$ identically, the solution exists if and only if the eigenvalues of X_4 are interlaced, so that the branes are placed as shown

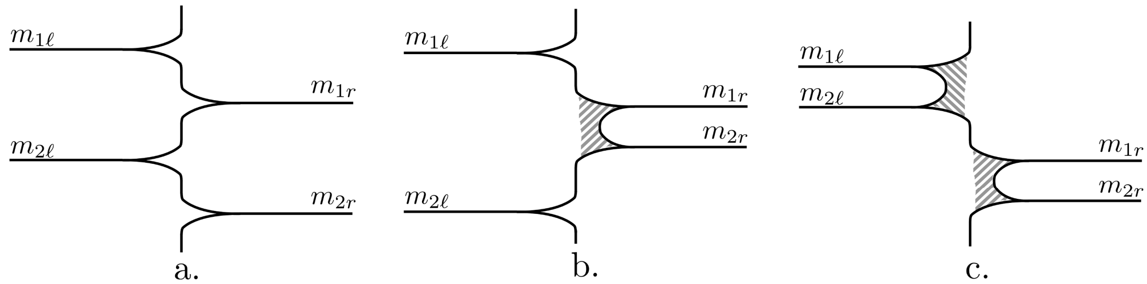

Figure 2.7: D3-branes ending on the two sides of a D5-brane. If the branes are not interlaced, they can form a fuzzy funnel.

in fig. 2.7(a). What if they are not interlaced? A unique vacuum solution still exists, but the assumption that X_5 and X_6 are identically 0 is no longer valid. For example, if two of the eigenvalues of X_4 for $y \to -\infty$ or for $y \to +\infty$ are very close – in other words if two of the λ_i or two of the λ'_i are very close – then the neighboring branes form a fuzzy funnel, as in fig. 2.7(b,c). The fuzzy funnel is described [57] by a nonabelian solution of Nahm's equations, with $X_4, X_5 \neq 0$. If $X_4, X_5 \to 0$ for $y \to \pm\infty$, then in the appropriate solution of Nahm's equations, $X_4 \pm iX_5$ is nilpotent, but not zero [21]. This suggests that we should try a new ansatz with ϕ_z nilpotent but not zero in order to find the missing solutions when the weights are not interlaced. For now, we present this as heuristic motivation for a more general ansatz, but later we will explain a precise map between the problem of finding half-BPS surface operators and Nahm's equations for a D3-D5 vacuum.

2.4.4.3 General Solution For U(2)

We consider the first non-trivial example of this problem, which is for gauge group U(2), corresponding to U(2|2) on the electric side. We focus on the configuration shown in fig. 2.7(b). The positions of the branes in that figure should be interpreted as the eigenvalues of the matrices which appear in the $1/r'$ singularity of the field ϕ_0. If the weights are $\alpha_r^\vee = i\,\mathrm{diag}(m_{1r}, m_{2r})$ and $\alpha_\ell^\vee = -i\,\mathrm{diag}(m_{1\ell}, m_{2\ell})$, then $m_{1r,\ell}$ and $m_{2r,\ell}$ label the positions of the horizontal lines in fig. 2.7. We assume that, by a Weyl conjugation, $\alpha^\vee$ was brought to the form with $m_{1r} > m_{2r}$ and $m_{1\ell} > m_{2\ell}$.

We introduce a convenient variable ς defined as $\sinh\varsigma = \cot\varphi$ (or $\tanh\varsigma = \cos\varphi$). It runs from $-\infty$ to 0 on the left of the defect, and from 0 to $+\infty$ on the right. For the fields on the left of the defect, we use the same abelian ansatz (2.121). For the fields on the right, we want to find a conformally- and $\mathrm{SO}(2)_X$-invariant solution with ϕ_z belonging to the non-trivial nilpotent conjugacy class. A family of such solutions, which actually contains all the solutions with these symmetries, was found in [6], and has the following form,

$$\begin{aligned} A &= \frac{i}{2}\begin{pmatrix} m_{1r}+m_{2r}+\partial_\varsigma V_r & 0 \\ 0 & m_{1r}+m_{2r}-\partial_\varsigma V_r \end{pmatrix}\cos\varphi\,\mathrm{d}\theta\,, \\ \phi_0 &= \frac{i}{2r'}\begin{pmatrix} m_{1r}+m_{2r}+\partial_\varsigma V_r & 0 \\ 0 & m_{1r}+m_{2r}-\partial_\varsigma V_r \end{pmatrix}, \\ \phi_z &= \frac{1}{2z}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\exp(-V_r)\,, \end{aligned} \tag{2.126}$$

where the function $V_r(\varsigma)$ is found from the localization equations to be

$$V_r = \log\left(\frac{\sinh(a_r\varsigma+b_r)}{a_r}\right). \tag{2.127}$$

The ansatz is $\mathrm{SO}(2)_X$-invariant up to a diagonal gauge transformation. In (2.126), a_r and b_r are some unknown constants. We choose a_r to be positive. Then b_r should also be positive, so that no singularity appears[30] in the interval $\varsigma\in(0,\infty)$. The requirement that the behavior of the gauge field at $\varsigma\to\infty$ should agree with the surface operator $A=\alpha_r^\vee\mathrm{d}\theta$ fixes $a = m_{1r}-m_{2r}$. (Had we chosen the opposite Weyl chamber for $\alpha^\vee$, we would have to make a Weyl transformation on the ansatz (2.126), making ϕ_z lower-triangular.) Note that, due to the $\cos\varphi$ factor, the gauge field at $y=0$ vanishes; this agrees with our requirement that $Z^\alpha\sim 1/\sqrt{z}$ should have monodromy -1. The next step is to impose the discontinuity

[30]The singularity that the solution has at $a_r\varsigma+b_r=0$ is the Nahm pole. In the one-sided problem, one chooses $b_r=0$ to have this pole precisely at $\varsigma=0$.

equations at $\varsigma = 0$, and to hope that they will have a solution for some positive real b_r. The z-component of the discontinuity equations tells us that the hypermultiplet fields should have the form

$$Z = \frac{1}{\sqrt{z}} \begin{pmatrix} \mathtt{s} \\ 0 \end{pmatrix}, \quad \widetilde{Z} = \frac{1}{\sqrt{z}} (0 \ i\mathtt{w}). \tag{2.128}$$

Unlike the interlaced case, here there is no freedom to include a general non-abelian gauge transformation in gluing the left and the right side. Such a gauge transformation would not be consistent with the symmetry, since generically it would not commute with the U(1) subgroup of the gauge group which is used to undo the $\mathrm{SO}(2)_X$ rotations. The only possible non-abelian gluing gauge transformation is the Weyl conjugation. The equations will tell us that in this case it is not needed. The ϕ_0 and ϕ_z discontinuity conditions give

$$\frac{a_r}{\sinh b_r} = \mathtt{sw}\,,$$

$$m_{1r} + m_{2r} - 2m_{1\ell} + a_r \coth b_r = -|\mathtt{s}|^2\,,$$

$$m_{1r} + m_{2r} - 2m_{2\ell} - a_r \coth b_r = |\mathtt{w}|^2. \tag{2.129}$$

Subtracting the last two equations, we see that a solution with positive b cannot exist unless $m_{1\ell} - m_{2\ell} > 0$. This is consistent with our choice of the Weyl chamber, so no gluing gauge transformation is needed. Eliminating $\mathtt{s}$ and $\mathtt{w}$ from (2.129), we obtain

$$\frac{m_{1\ell} - m_{2\ell}}{m_{1r} - m_{2r}} = \coth b_r + \sqrt{\left(\frac{m_{1r} + m_{2r} - m_{1\ell} - m_{2\ell}}{m_{1r} - m_{2r}}\right)^2 + \frac{1}{\sinh^2 b_r}}. \tag{2.130}$$

The function on the right is monotonically decreasing. It is easy to see that the equation has a solution $b_r > 0$ if and only if the eigenvalues are arranged as in fig. 2.7b.

The last case to consider for the U(2) group is that of fig. 2.7c. Here fields on both sides of the defect should have a non-zero nilpotent ϕ_z. The fields on the right are given by the same ansatz (2.126). The fields on the left are given by the same ansatz, but with V_r replaced by

$$V_\ell = \log\left(\frac{\sinh(-a_\ell\varsigma + b_\ell)}{a_\ell}\right). \tag{2.131}$$

Again, we assume a_ℓ to be positive, and then b_ℓ should also be positive to avoid the singularity on the interval $\varsigma \in (-\infty, 0)$. We fix a_ℓ from the asymptotics at $\varsigma \to -\infty$ to be $a_\ell = m_{1\ell} - m_{2\ell}$, though in this case the gauge field A asymptotically is proportional to $\mathrm{diag}(m_{2\ell}, m_{1\ell})$. We could make a Weyl gauge transformation to bring it to the other Weyl chamber.

In gluing left and right, we cannot make any non-diagonal gauge transformations, as follows again from the $\mathrm{SO}(2)_X$ symmetry. There are two separate cases to consider. First assume that ϕ_z has a non-trivial jump at $y = 0$. This forces the hypermultiplet fields Z and $\widetilde{Z}$ to have the form (2.128). The discontinuity equations give

$$\begin{aligned}
\frac{a_r}{\sinh b_r} \pm \frac{a_\ell}{\sinh b_\ell} &= \mathtt{sw}\,, \\
a_r \coth b_r + a_\ell \coth b_\ell &= -\frac{|\mathtt{w}|^2 + |\mathtt{s}|^2}{2}\,, \\
m_{1r} + m_{2r} - m_{1\ell} - m_{2\ell} &= \frac{|\mathtt{w}|^2 - |\mathtt{s}|^2}{2}.
\end{aligned} \tag{2.132}$$

(The sign in the first equation can be exchanged by an abelian gluing gauge transformation.) The second equation clearly has no positive solutions for $b_{r,\ell}$.

Therefore, the field ϕ_z has to be continuous at $y = 0$. In this case, either Z or $\widetilde{Z}$ should be zero. Assume that it is $\widetilde{Z}$, and

$$Z = \frac{1}{\sqrt{z}}\begin{pmatrix} \mathtt{s} \\ \mathtt{w} \end{pmatrix}. \tag{2.133}$$

Since the field ϕ_0 is diagonal, the matrix $Z \otimes Z^\dagger$ should be also diagonal, so either $\mathtt{s}$ or $\mathtt{w}$ is zero. We have to choose $\mathtt{s} = 0$ to avoid the same sign problem which caused trouble in the second equation in (2.132). The discontinuity equations become

$$\frac{a_r}{\sinh b_r} - \frac{a_\ell}{\sinh b_\ell} = 0\,,$$

$$a_r \coth b_r + a_\ell \coth b_\ell = |\mathtt{w}|^2/2\,,$$

$$m_{1r} + m_{2r} - m_{1\ell} - m_{2\ell} = -|\mathtt{w}|^2/2. \tag{2.134}$$

The last equation here implies that $m_{1r} + m_{2r} < m_{1\ell} + m_{2\ell}$. In the opposite case, we would have to take Z and not $\widetilde{Z}$ to be zero. Eliminating $|\mathtt{w}|$ and b_ℓ, we get

$$\frac{m_{1\ell} + m_{2\ell} - m_{1r} - m_{2r}}{m_{1r} - m_{2r}} = \coth b_r + \sqrt{\left(\frac{m_{1\ell} - m_{2\ell}}{m_{1r} - m_{2r}}\right)^2 + \frac{1}{\sinh^2 b_r}}. \tag{2.135}$$

This equation has a solution precisely when the eigenvalues are arranged as in fig. 2.7c.

2.4.4.4 General Surface Operators

We have described the abelian solutions for the $\mathrm{U}(n|n)$ case, and some more general solutions for $\mathrm{U}(2|2)$ for surface operators of type $(\alpha^\vee, 0, 0)$. In this section we look at the general singularities of type $(\alpha^\vee, \beta^\vee, \gamma^\vee)$, aiming to make a precise statement about the correspondence between surface operators and supersymmetric vacua of the theory.

Let us go from the coordinates (t, x_1, x_2, y) to $(t, \varsigma, \theta, r')$, in which the rotational and scaling symmetries act in the most simple way. The flat metric in these coordinates is conformally equivalent to $\cosh^2\varsigma(\mathrm{d}t^2 + \mathrm{d}r'^2)/r'^2 + \mathrm{d}\varsigma^2 + \mathrm{d}\theta^2$, which is $AdS_2 \times \mathbb{R}_\varsigma \times S^1_\theta$, up to a warping factor $\cosh^2\varsigma$. In conformal field theory, finding a model solution for a surface operator is equivalent to finding a vacuum configuration in this space, with the asymptotics of the scalar fields at $\varsigma \to \pm\infty$ defined by the charges of the surface operator. To make this intuition precise, let us rewrite our localization equations (2.119), (2.120) in terms of these coordinates. We make a general scale-invariant and rotationally-invariant ansatz for the fields,

$$\phi_0 = \frac{1}{r'} M(\varsigma)\,, \quad \phi_z = \frac{1}{z} N(\varsigma)\,, \quad A = M_1(\varsigma)\cos\varphi\,\mathrm{d}\theta. \tag{2.136}$$

(We could have absorbed $\cos\varphi = \tanh\varsigma$ into M_1, but it is more convenient to write it this way.) The equations reduce to

$$[\partial_\varsigma - iM, N] = 0\,, \quad [\partial_\varsigma - iM_1, N] = 0\,, \quad [\partial_\varsigma - iM, \partial_\varsigma - iM_1] + \frac{2i}{\sinh 2\varsigma}(M - M_1) = 0\,, \tag{2.137}$$

together with

$$\frac{\sinh^2\varsigma\,\partial_\varsigma M_1 + \partial_\varsigma M}{\cosh^2\varsigma} + 2i[N, N^\dagger] + \frac{\sinh\varsigma}{\cosh^3\varsigma}(M_1 - M) = 0. \tag{2.138}$$

The first set of equations almost implies that $M_1 = M$. In fact, there is a class of reducible solutions for which this equality is not true. They will be described in the next section, but for now we take $M_1 = M$ as an ansatz. Then the equations reduce simply to Nahm's equations $[\partial_\varsigma - iM, N] = 0$ and $\partial_\varsigma M + 2i[N, N^\dagger] = 0$ for the scalar fields M, $\mathrm{Re}(N)$ and $\mathrm{Im}(N)$. At $\varsigma \to \pm\infty$, these fields should approach limiting values given by the parameters of the surface operator $\alpha^\vee$, $\beta^\vee$ and $\gamma^\vee$. At $\varsigma = 0$, assuming the regularity of $M(\varsigma)$, the conformally invariant solution for Z and $\widetilde{Z}$ is given by $1/\sqrt{z}$ times some constant vectors, which should be found from the discontinuity equations.

In this way, the problem of finding the model solution for a surface operator is indeed reduced to the problem of finding the supersymmetric vacuum of the D3-D5 system for given asymptotic values of the scalar fields. To actually find the solutions, one needs to find the solutions of the Nahm's equations on a half-line, with asymptotics of the fields given by the regular triple $(\alpha^\vee, \beta^\vee, \gamma^\vee)$, and then glue them at $y = 0$, according to the discontinuity equations. The relevant solutions of the Nahm's equations can be found e.g. in [51]. The problem reduces to solving a set of algebraic equations for the integration constants of the solutions and the components of the hypermultiplet field Z^α. Solving these equations seems like a tedious problem even for the U(2) case, and we will not attempt to do it here. The relation to the supersymmetric vacua guarantees that for any values of the parameters a model solution exists, unique up to gauge invariance.

The reduction that we have just described works for the unequal rank case as well. The gluing conditions of section 2.4.3 for the conformally-invariant solution (2.136) at $y = 0$ reduce to the gluing conditions for the scalar fields M and N. In particular, a $1/y$ Nahm pole boundary condition translates into a $1/\varsigma$ Nahm pole for the vacuum scalar fields.

2.4.4.5 Reducible Solutions

So far we have concentrated on irreducible solutions, but there are reducible solutions as well.

Returning to eqn. (2.137), instead of setting $M_1 = M$, we write $M_1 = M + S$. We find that the equations are obeyed if M and N obey the same conditions as before, while

$$[S, N] = [S, M] = \partial_\varsigma S + \frac{2}{\sinh 2\varsigma} S = 0. \tag{2.139}$$

The last equation means that

$$S = \coth \zeta S_0 = \frac{1}{\cos\varphi} S_0 \tag{2.140}$$

with a constant matrix S_0.

The interpretation is very simple. First we describe the equal rank case. In $\mathrm{U}(n)$, we pick a subgroup $\mathrm{U}(n-m)\times\mathrm{U}(m)$. In $\mathrm{U}(n-m)$, we pick matrices M, N and defect fields $Z, \widetilde{Z}$ that satisfy Nahm's equations and the jump conditions at $y = 0$, giving an irreducible solution (in $\mathrm{U}(n-m)$) as described in section 2.4.4.4. In $\mathrm{U}(m)$, we embed a trivial abelian solution with $A = \alpha^\vee d\theta$, $\phi_z = (\beta^\vee + i\gamma^\vee)/(2z)$, $\phi_0 = 0$. (This trivial solution is obtained by taking $S = \alpha^\vee$, and taking the $\mathfrak{u}(m)$-valued part of N to be the constant matrix $(\beta^\vee + i\gamma^\vee)/2$.) This describes a solution that can exist if m eigenvalues of $\vec{\zeta}_\ell^\vee = (\alpha_\ell^\vee, \beta_\ell^\vee, \gamma_\ell^\vee)$ coincide with m eigenvalues of $\vec{\zeta}_r^\vee = (\alpha_r^\vee, \beta_r^\vee, \gamma_r^\vee)$. For left and right eigenvalues to coincide is the condition for an atypical weight, so these solutions govern atypical weights.

For the same atypical weight, however, we could have simply used the irreducible $\mathrm{U}(n)$-valued solution with $S = 0$ constructed in section 2.4.4.4. After all, this solution exists for any weights. More generally, consider an atypical weight of $\mathrm{U}(n|n)$ with s eigenvalues of $\vec{\zeta}_\ell^\vee$ equal to corresponding eigenvalues of $\vec{\zeta}_r^\vee$. For any $m \leq s$, we can obtain a surface operator solution with this weight, based on a subgroup $\mathrm{U}(n-m) \times \mathrm{U}(m) \subset \mathrm{U}(n)$. We simply take a trivial abelian solution in $\mathrm{U}(m)$ based on m of the s common weights, and combine this with an irreducible solution in $\mathrm{U}(n-m)$ for all the other weights. For each m, there are $\binom{s}{m}$ such solutions, since we had to pick m of the s common weights. Considering all values of m from 0 to s, this gives 2^s surface operator solutions for a weight of $\mathrm{U}(n|n)$ of atypicality s. Qualitatively, this is in agreement with what one finds on the electric side, where a finite-dimensional representation with a given highest weight is unique only if the weight is typical. In the case that the weights $\alpha_\ell^\vee$ and $\alpha_r^\vee$ are integral and $\beta_\ell^\vee, \gamma_\ell^\vee$ and $\beta_r^\vee, \gamma_r^\vee$ all vanish, so that the model solutions that we have constructed are related to line operators (see section 2.4.5), this leads to 2^s line operators associated to a weight of atypicality s; we suspect that they are dual to 2^s distinguished representations with the given highest weight.

The story is similar for unequal ranks. The gauge group is $\mathrm{U}(n)$ for $y < 0$ and $\mathrm{U}(n+r)$ for $y > 0$. We pick subgroups $\mathrm{U}(n-m) \times \mathrm{U}(m) \subset \mathrm{U}(n)$ and $\mathrm{U}(n+r-m) \times \mathrm{U}(m) \subset \mathrm{U}(n+r)$. We combine a trivial abelian $\mathrm{U}(m)$-valued solution on the whole y line with an irreducible solution based on $\mathrm{U}(n-m)$ for $y < 0$ and $\mathrm{U}(n+r-m)$ for $y > 0$. Just as in the last paragraph, we get 2^s solutions for a weight of $\mathrm{U}(n|n+r)$ of atypicality s.

Another type of reducible solution was found in section 2.4.4.2. If one of the eigenvalues of $\alpha_r^\vee$ is equal to an eigenvalue of $-\alpha_\ell^\vee$, then the corresponding matrix elements of Z and $\widetilde{Z}$

vanish and a U(1) subgroup of the gauge group is unbroken. The basic phenomenon occurs actually for the gauge group U(1), corresponding to the supergroup U(1|1). There is a surface operator described by a trivial abelian solution with $A = \alpha^\vee \cos\varphi\,\mathrm{d}\theta$ and $\phi = \alpha^\vee\,\mathrm{d}x^0/r'$ everywhere and $Z = \widetilde{Z} = 0$. (This solution has $\alpha_r^\vee = \alpha = -\alpha_\ell^\vee$ because $\cos\varphi = 1$ on the positive y axis and -1 on the negative y axis.) Clearly since Z and $\widetilde{Z}$ vanish, the U(1) gauge symmetry is unbroken. This is a reducible solution that can occur for a typical weight, since $\alpha_r^\vee = -\alpha_\ell^\vee$ is not a condition for atypicality. Such a surface operator does not seem to be well-defined. Since the gauge symmetry remains unbroken along the knot K, the gauge field near K is free to fluctuate. In particular, it follows that the variation of the topological term in the presence of this model singularity is not zero, but is proportional to $\int_K \alpha\delta A$, and therefore, the action is not Q-invariant. We do not know how to interpret the singularity that seems to arise when an eigenvalue of $\alpha_\ell^\vee$ approaches one of $-\alpha_r^\vee$, or how to describe a half-BPS surface operator in this case. A possibly similar problem arises in the bulk in $\mathcal{N} = 4$ super Yang-Mills theory with any nonabelian gauge group if one tries to define a surface operator with parameters $(0, 0, 0, \eta^\vee)$. Classically, it is hard to see how to do this, since the definition of $\eta^\vee$ requires a reduction of the gauge symmetry to the maximal torus along the support of the surface operator, and this is lacking classically if $\alpha^\vee = \beta^\vee = \gamma^\vee = 0$. Yet the surface operator in question certainly exists; it is S-dual to a surface operator with parameters $(\alpha, 0, 0, 0)$ that can be constructed semiclassically. One approach to defining it involves adding additional variables along the surface (see section 3 of [54]).

2.4.5 Line Operators And Their Dualities

We have constructed surface operators, but there is an easy way to construct line operators from them. We simply observe that if we set $\beta^\vee = \gamma^\vee = 0$, and also take $\alpha^\vee$ to be integral, then the bulk solution $A = \alpha^\vee\,\mathrm{d}\theta$ defining a surface operator in the absence of any D5-brane can be gauged away. So for those parameters, the surface operators that we have constructed are trivial far away from the D5-brane defect. That means that those surface operators reduce macroscopically to line operators supported on the defect.

Saying that $\alpha^\vee$ is "integral" means that it is a cocharacter of the maximal torus of the dual group $G^\vee$, or in other words a character of the maximal torus of G. Up to the action of the Weyl group, this character corresponds to a dominant weight of G. In other words, we have found line operators of the magnetic description by $G^\vee$ gauge theory that are classified by dominant weights (or representations) of the electric group G.

In all these statements, G is either G_ℓ or G_r, the gauge group to the left or right of the D5-brane defect. Taking account of the behavior on both sides, these line operators are really classified by dominant weights of $G_\ell \times G_r$. (In our main example of $\mathrm{U}(m|n)$, G is $\mathrm{U}(m)$ or $\mathrm{U}(n)$ and the distinction between G and its dual group $G^\vee$ is not important. However, this part of the analysis is more general and carries over also to the orthosymplectic case that we discuss in section 2.5.)

Wilson-'t Hooft operators of the "electric" description involving an NS5-brane are also classified by dominant weights of $G_\ell \times G_r$ (or equivalently by dominant weights of the supergroup SG), as we learned in section 2.3.3.2. Thus an obvious duality conjecture presents itself: the line operator associated to a given weight of $G_\ell \times G_r$ in one description is dual to the line operator associated to the same weight in the other description.

This statement is a natural analog of the usual duality between Wilson and 't Hooft operators, adapted to the present situation. But a detail remains to be clarified. In the standard mapping between Wilson operators of G and 't Hooft operators of $G^\vee$, there is a minus sign that to some extent is a matter of convention. That is because electric-magnetic duality could be composed with charge conjugation for either G or $G^\vee$. Charge conjugation acts by reversing the sign of a weight, up to a Weyl transformation.

In the supergroup case, let $(\lambda_\ell, \lambda_r)$ be a weight of $G_\ell \times G_r$, and let $(\alpha_\ell^\vee, \alpha_r^\vee)$ be a magnetic weight of $G_\ell^\vee \times G_r^\vee$. If we specify that we want a duality transformation that maps λ_ℓ to $+\alpha_\ell^\vee$, then it becomes a well-defined question whether λ_r maps to $+\alpha_r^\vee$ or to $-\alpha_r^\vee$. The correct answer is the one with a minus sign:

$$(\lambda_\ell, \lambda_r) \leftrightarrow (\alpha_\ell^\vee, -\alpha_r^\vee). \tag{2.141}$$

To see this, we observe that there is a symmetry of the problem that exchanges the left and right of the defect and exchanges λ_ℓ with λ_r but $\alpha_\ell^\vee$ with $-\alpha_r^\vee$. For a defect at $x^3 = 0$ and a line operator supported on the line $L : x^1 = x^2 = x^3 = 0$, we can take this symmetry to be $x^2 \to -x^2$, $x^3 \to -x^3$, with x^0, x^1 fixed. This has been chosen to exchange the left and right sides of the defect, while mapping the line L to itself and preserving the orientation of spacetime, so as to leave $\mathcal{K}$ fixed. It does not affect electric charge, but it reverses the sign of $\alpha^\vee$ because it reverses the orientation of the x^1x^2 plane.

As was already remarked in section 2.3.5, in the case of an atypical weight, our pictures on the magnetic and electric sides do not quite match. On the magnetic side, for a given atypical weight, we have found multiple possible 1/2 BPS surface and line operators, as explained in section 2.4.4.5. On the electric side, for any weight, even atypical, we found only a single 1/2 BPS surface or Wilson-'t Hooft line operator.

2.4.6 A Magnetic Formula For Knot And Link Invariants

The Q-invariant line and surface operators that we have constructed can be used to get magnetic formulas for knot and link invariants. In the case of line operators, we have little to add to what was stated in eqn. (2.99). Here we will elaborate on the construction of knot and link invariants using surface operators. After some general observations, we will comment on what happens for atypical weights.

We start on the electric side with a knot invariant defined by including a surface operator with parameters $(\alpha, \beta, \gamma, \eta)$ supported on a two-surface C that intersects the hyperplane $y = 0$ along a knot K. One can take simply $C = K \times \mathbb{R}_y$ (where $\mathbb{R}_y$ is parametrized by y) or one can choose C to be compact. The dual magnetic description involves a surface operator wrapped on C with parameters $(\alpha^\vee, \beta^\vee, \gamma^\vee, \eta^\vee) = (\eta, |\tau|\beta^*, |\tau|\gamma^*, -\alpha)$.

The parameters of the surface operator in the magnetic case define the singularities of the fields near C, but also they determine some insertions that must be made in the functional integral along C. The action of the theory in the presence of the surface operator is

$$\frac{i\mathcal{K}^\vee}{4\pi}\int_M \operatorname{tr}(F\wedge F) - i\int_C \operatorname{tr}\left((\mathcal{K}^\vee\alpha^\vee - \eta^{\vee *})F\right)\,, \tag{2.142}$$

modulo Q-exact terms. We have used eqns. (2.81) and (2.78) for the terms proportional to $\alpha^\vee$ and $\eta^{\vee *}$. The integral in the four-dimensional topological term is taken over M, but alternatively, we could take it over $M\setminus C$, and that would absorb the term proportional to $\alpha^\vee$. Note that the objects which appear in this formula are topological invariants, because the bundle is naturally trivialized both at infinity and in the vicinity of K, where the fields Z^α become large. (For now we consider the generic irreducible case, when the gauge group is completely broken along K; we do not consider the problem mentioned at the end of section 2.4.4.5.) Using the relation (2.85) between weights and parameters of the surface operator, the action can be alternatively written as

$$\frac{i\mathcal{K}^\vee}{4\pi}\int_M \operatorname{tr}(F\wedge F) + i\mathcal{K}^\vee\int_{C_r}\operatorname{tr}(\lambda_r F) - i\mathcal{K}^\vee\int_{C_\ell}\operatorname{tr}(\lambda_\ell F). \tag{2.143}$$

The insertion of the two-dimensional observable in this formula is essentially the S-dual of the analogous insertion in the electric theory. This statement can be justified explicitly if the gauge group is abelian. In that case, the two-observable $\int F$ is the second descendant of the Q-closed field σ. Under S-duality, both the gauge-invariant polynomials of σ and their descendants are mapped to each other. (See Appendix D for details on the descent procedure in the presence of the three-dimensional defect.)

The functional integral in the magnetic theory can be localized on the space of solutions to the localization equations (2.108), (2.110). The knot polynomial can be obtained by counting the solutions of the localization equations in the presence of a singularity of type

type $(\alpha^\vee, \beta^\vee, \gamma^\vee)$, weighted by the combination (2.143) of topological numbers of the solution, as well as the sign of the fermion determinant. (These statements hold for both the equal-rank and unequal rank cases, though one uses different equations and model solutions in the two cases.) For a given weight, there are different possible choices of surface operator. We can vary $\alpha^\vee$ and $\eta^\vee$, as long as their appropriate combination is equal to the weight. We can also turn on arbitrary $\gamma^\vee$ and $\beta^\vee$, as long as it is not forbidden for topological reasons. All this simply reflects the fact that the problem of counting solutions of elliptic equations is formally invariant under continuous deformations of parameters. Note that, in particular, the operators with $\gamma^\vee \neq 0$ are well-defined and 1/2-BPS, and changing $\gamma^\vee$ does not change the weight in (2.143), with which the solutions of the localization equations are counted. This supports the view, proposed in section 2.3.4.2, that in the physical theory γ plays much the same role, as β: it deforms the contour of integration in the functional integral, without changing the Chern-Simons observables.[31]

It is conceivable that the counting of the solutions of the localization equations is only generically independent of the parameters $(\alpha^\vee, \beta^\vee, \gamma^\vee)$, and that wall-crossing phenomena can occur. (A prototype of what might happen has been seen for the three-dimensional Seiberg-Witten equations [20].) We will not attempt to analyze this possibility here, and will simply assume that for any regular triple $(\alpha^\vee, \beta^\vee, \gamma^\vee)$, the counting of solutions is the same. Let $\mathcal{S}_0$ be the space of these solutions. It is convenient to introduce variables $\mathbf{t}_r = q^{-\lambda_r^*}$ and $\mathbf{t}_\ell = q^{\lambda_\ell^*}$, valued in the complexification of the maximal tori of the left and the right bosonic gauge groups of the electric theory. The knot polynomial is then given by

$$\sum_{s\in\mathcal{S}_0} (-1)^f q^{N^\vee}\, \mathbf{t}_\ell^{c_{1\ell}}\, \mathbf{t}_r^{c_{1r}}. \tag{2.144}$$

[31] All this is true for the physical theory, where both $\mathcal{K}$ and the weights are real. We expect the situation to be different in the topological theory, where on the electric side the surface operators with $\gamma \neq 0$ are defined according to eq. (2.84). In that case, γ is related to the imaginary part of the weight. In particular, the insertion of $i\mathcal{K}w\int \mathrm{tr}(\gamma\mathcal{F})$ in (2.84) will lead on the magnetic side to a similar insertion, which will complexify the weight in eq. (2.143).

Here $(-1)^f$ is the sign of the fermion determinant, evaluated in the background of the classical solution s, $N^\vee = \frac{1}{8\pi^2}\int_{M\backslash K} \mathrm{tr}(F\wedge F)$ is the instanton number, and $c_{1\ell,r} = \frac{1}{2\pi}\int_{C_{r,\ell}} F$ are the $\mathfrak{t}^\vee$-valued relative first Chern classes for the abelian bundles on C_r and C_ℓ. One can consider (2.144) as a polynomial in q, after expressing $\mathbf{t}_{\ell,r}$ in terms of q for a particular weight λ, but one can also treat $\mathbf{t}_{\ell,r}$ as independent formal variables.

What happens if the weight λ is atypical? By varying $\alpha^\vee$ and $\eta^\vee$, while preserving λ, we can still make the model solution irreducible. So we can use the solutions from $\mathcal{S}_0$ to obtain the knot polynomial, and simply substitute our λ in eqn.(2.144). We expect that this polynomial will correspond to the Kac module of highest weight λ. This expectation follows from the fact that a typical representation can be continuously deformed into an atypical one by varying the fermionic Dynkin label a_{ferm}. Since this label need not be integral, this variation makes sense, and the limit of this typical representation, when the weight becomes atypical, is the Kac module. In the magnetic theory, to take the limit of a knot invariant, we simply substitute the atypical weight into the universal polynomial (2.144), evaluated on $\mathcal{S}_0$. So this type of polynomial indeed corresponds to the Kac module.

For an atypical weight, rather than an irreducible model solution, we can also use surface operators defined by reducible solutions. For any weight of atypicality at least p, we can consider a surface operator whose irreducible part is associated to a surface operator of $\mathrm{U}(m-p|n-p)$. This surface operator breaks the bosonic group $\mathrm{U}(m)\times\mathrm{U}(n)$ to an subgroup H that generically is $\mathrm{U}(1)^p$ (it can be a nonabelian group containing $\mathrm{U}(1)^p$ if the reducible part of the solution is non-regular). Let $T_H \cong \mathrm{U}(1)^p$ be the maximal torus of H. The group H acts on the space of solutions of the localization equations. In such a situation, by standard localization arguments,[32] the invariants can be computed by just counting the T_H-invariant

[32] Generically, one expects that the solutions consist of a finite set of points, and if so, these points are all invariant under the continuous group T_H. However, suppose that some of the solutions make up a manifold U that has a non-trivial action of T_H. Then by standard arguments of cohomological field theory [58], the contribution of the manifold U to the counting of solutions is $(-1)^f\chi(U;V)$, where $(-1)^f$ is the sign of the fermion determinant, $V\to U$ is a certain "obstruction bundle" (a real vector bundle of rank equal to the dimension of U), and $\chi(U;V)$ is the Euler characteristic of $V\to U$. Let U' be the fixed point set of the action of T_H on U and let $V'\to U'$ be the T_H-invariant subbundle of $V|_{U'}$. A standard topological argument shows that $(-1)^f\chi(U;V) = (-1)^{f'}\chi(U';V')$ (if U' is not connected, one must write a sum over components

solutions. The T_H-invariant subgroup of $\mathrm{U}(m) \times \mathrm{U}(n)$ is $T_H \times \mathrm{U}(m-p) \times \mathrm{U}(n-p)$. There are no interesting solutions valued in the abelian group T_H, so in fact, the $\mathrm{U}(m|n)$ invariants with a surface operator of this type can be computed by counting solutions for $\mathrm{U}(m-p|n-p)$. Some simple group theory shows that the signs of the two fermion determinants are the same and hence the $\mathrm{U}(m|n)$ invariants for a weight of atypicality $\geq p$ coincide with $\mathrm{U}(m-p|n-p)$ invariants. In particular, $\mathrm{U}(m|n)$ invariants of maximal atypicality coincide with invariants of the bosonic group $\mathrm{U}(|n-m|)$. (This reasoning also makes it clear that the knot and link invariants constructed using a reducible model solution do not depend on the weights in the abelian part of the solution.)

For a weight of atypicality r, we can take any $p \leq r$ in this construction. We have argued that for $p = 0$, we expect to get invariants associated to the Kac module, while $p = r$ presumably corresponds to the irreducible atypical representation. The intermediate values of p plausibly correspond to the reducible indecomposables, which are obtained by taking non-minimal subquotients of the Kac module.

In section 6 of [1], an alternative approach to comparing $\mathrm{U}(m|n)$ with $\mathrm{U}(m-p|n-p)$ is given. The key idea there is gauge symmetry breaking. This approach is very natural on the electric side.

In the rather formal discussion that we have given here, we have not taken into account some of the insight from section 2.3.2.2. From that analysis, we know that for the knot invariants to be nonzero, we can consider a typical weight for a knot in S^3 or a maximally atypical weight for a knot in $\mathbb{R}^3$. For other weights, a slightly different approach is needed. We have not understood the analogs of these statements on the magnetic side.

2.4.7 A Possible Application

Here we will briefly indicate a possible application of this work, for gauge group $\mathrm{U}(1|1)$. This direction will be explored in more detail in Chapter 3 of this thesis.

---

on the right hand side). In our problem, this means that we can consider only the $\mathrm{U}(m-p|n-p)$ solutions and count them just as we would for $\mathrm{U}(m-p|n-p)$, ignoring the embedding in $\mathrm{U}(m|n)$.

Using the fact that the supergroup U(1|1) is solvable, the invariant for a knot $K \subset S^3$ labeled by a typical representation of U(1|1) can be explicitly computed by repeated Gaussian integrals. It turns out to equal the Alexander polynomial [59, 60, 61]. The usual variable q on which the Alexander polynomial depends is a certain function of the Chern-Simons coupling and the typical weight.

The Alexander polynomial of K can also be computed [20] by counting solutions of a 3d version of the Seiberg-Witten equations with a prescribed singularity along K. Such solutions can be labeled by an integer-valued invariant Θ (a certain relative first Chern class), and if b_n is the number of solutions with $\Theta = n$ (weighted as usual with the sign of a certain fermion determinant), then the Alexander polynomial is $Z(q) = \sum_n b_n q^n$. The proof that $Z(q)$ equals the Alexander polynomial is made by showing that the two functions obey the same "skein relations."

The question arises of whether one could find a more direct explanation of this result, or perhaps a more direct link between U(1|1) Chern-Simons theory and the Seiberg-Witten equations. From the point of view of the present chapter, U(1|1) Chern-Simons theory can be represented in terms of $\mathcal{N} = 4$ super Yang-Mills theory with gauge group $\mathrm{U}(1)_\ell \times \mathrm{U}(1)_r$ on $S^3 \times \mathbb{R}$, interacting with a bifundamental hypermultiplet that is supported on $S^3 \times \{0\}$. However, we can just as well replace $\mathbb{R}$ here by S^1. If we do that, we get U(1|1) Chern-Simons theory with a different integration cycle. However, as long as one considers only Wilson operators on $\mathbb{R}^3$ or S^3, all integration cycles are equivalent and so $\mathcal{N} = 4$ super Yang-Mills theory on $S^3 \times S^1$ with a bifundamental hypermultiplet on $S^3 \times \{0\}$ should give another way to study the Alexander polynomial.[33]

S-duality converts this to a "magnetic" problem on $S^3 \times S^1$, now with U(1) gauge fields in bulk and a twisted hypermultiplet supported on $S^3 \times S^1$. If one takes the radius of S^1 to be small compared to that of S^3, the four-dimensional localization equations can be expected

[33] Once we replace $S^3 \times \mathbb{R}$ with $S^3 \times S^1$, the left and right of the defect are connected. So we now have a single U(1) vector multiplet on $S^3 \times S^1$, with the fields allowed to have different limits as $S^3 \times \{0\}$ is approached from the left or right. The two limits give two different sets of 3d fields to which the "bifundamental" hypermultiplet is coupled.

to reduce to three-dimensional effective equations. These will be equations in which U(1) gauge fields are coupled to a hypermultiplet, and one can argue that the relevant equations are the Seiberg-Witten equations.

Thus one can hope that, as in [20], it will be possible to compute the Alexander polynomial by counting solutions of the Seiberg-Witten equations. Unfortunately, in working on $S^3 \times S^1$, one encounters a number of technical difficulties. In Chapter 3, we instead consider the three-dimensional theory, which is obtained by compactifying on an interval with particular boundary conditions, instead of S^1.

2.5 Orthosymplectic Chern-Simons Theory

In this section, we return to the D3-NS5 system of fig. 2.1, but now we add an O3-plane parallel to the D3-branes. A D3-O3 system can have orthogonal or symplectic gauge symmetry, depending on which type of O3-plane is chosen. The gauge symmetry jumps from orthogonal to symplectic in crossing an NS5-brane. Accordingly, the construction of section 2.2, with an O3-plane added, is related to Chern-Simons theory of an orthosymplectic gauge group $\mathrm{OSp}(r|2n)$, where the integers r and n depend on the numbers of D3-branes on the two sides of the NS5-brane. As in section 2.4, an S-duality transformation that converts the D3-O3-NS5 system to a D3-O3-D5 system gives a magnetic dual of three-dimensional $\mathrm{OSp}(r|2n)$ Chern-Simons theory. This is a close analog of what we have already seen for unitary groups.

However, something novel happens if $r = 2m+1$ is odd. In this case, a slightly different procedure yields a duality between two "electric" descriptions. In three-dimensional terms, we will learn that Chern-Simons theory of $\mathrm{OSp}(2m+1|2n)$, with coupling parameter q, is equivalent to Chern-Simons theory of $\mathrm{OSp}(2n+1|2m)$, with coupling parameter $-q$. (The Chern-Simons theories that appear in this statement are defined via the brane constructions which as usual allow analytic continuation away from integer levels.) Since weak coupling

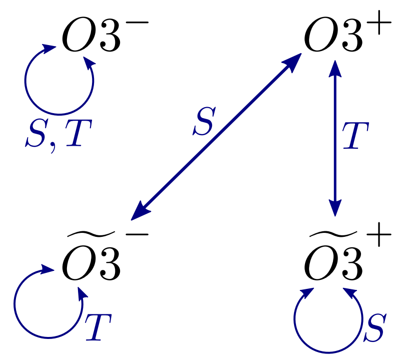

Figure 2.8: Action of the S-duality group on the orientifold planes.

in Chern-Simons theory is $q \to 1$, while $q \to -1$ is a strongly-coupled limit, this duality exchanges strong and weak coupling.

2.5.1 Review Of Orientifold Planes

We start with a brief review of the orientifold 3-planes of Type IIB superstring theory [25, 26] (see also section 7 of [29]).

There are four kinds of O3-plane, distinguished by $\mathbb{Z}_2$-valued discrete fluxes of the NS and RR two-form fields of Type IIB supergravity. An O3-plane in which both fluxes vanish is denoted $\mathrm{O3}^-$; in the presence of m parallel D3-branes (and their images) it gives $\mathrm{O}(2m)$ gauge symmetry (for some purposes, we consider only the connected component $\mathrm{SO}(2m)$). Adding discrete RR flux gives an $\widetilde{\mathrm{O3}}^-$-plane, which with the addition of m parallel D3-branes gives $\mathrm{O}(2m+1)$ gauge symmetry. An orientifold 3-plane with only NS flux is denoted $\mathrm{O3}^+$ and gives $\mathrm{Sp}(2m)$ gauge symmetry. Finally, the orientifold $\widetilde{\mathrm{O3}}^+$ with both kinds of flux gives again $\mathrm{Sp}(2m)$ gauge symmetry, but (as we recall shortly) with a shift in the value of the theta-angle θ_{YM}, a fact that we abbreviate by saying that the gauge group is $\mathrm{Sp}'(2m)$. The transformation properties of the orientifold 3-planes under the $\mathrm{SL}(2,\mathbb{Z})$ S-duality group are summarized in fig. 2.8.

When an O3-plane crosses an NS5-brane, its NS flux jumps; when it crosses a D5-brane, its RR flux jumps. More generally, when an O3-plane crosses a (p, q)-fivebrane its $(\mathrm{NS}, \mathrm{RR})$

fluxes jump by (p,q) mod 2.

Regardless of the type of O3-plane, a D3-O3 system has the same supersymmetry as a system of D3-branes only. In particular, this supersymmetry is parametrized by the angle ϑ, which is related to the string coupling in the usual way, as in eqn. (2.91). To find the classical effective action for the gauge theory that describes a D3-O3 system at low energies, we simply take the effective action of a D3-brane system, restrict the fields to be invariant under the orientifold projection, and divide by 2. The restriction reduces a $\mathrm{U}(n)$ gauge symmetry to $\mathrm{O}(n)$ or $\mathrm{Sp}(n)$, depending on the type of O3-plane. We divide by 2 because the orientifolding operation is a sort of discrete gauge symmetry in string theory. (As we explain shortly, there is a subtlety in dividing θ_{YM} by 2.) The same procedure of restricting to the invariant subspace and dividing by 2 enables us to deduce the effective action of a D3-O3-NS5 or D3-O3-D5 system from those of a D3-NS5 or D3-D5 system.

For the $\mathrm{U}(n)$ gauge fields along a system of n parallel D3-branes, we write the gauge theory action as

$$\frac{1}{2g_{\mathrm{YM}}^2}\int \mathrm{d}^4x\,\mathrm{tr}\,F_{\mu\nu}^2-\frac{\theta_{\mathrm{YM}}}{8\pi^2}\int \mathrm{tr}\,F\wedge F\,, \tag{2.145}$$

where tr is the trace in the fundamental representation of $\mathrm{U}(n)$, and the Yang-Mills parameters g_{YM} and θ_{YM} are related to the τ parameter of the underlying Type IIB superstring theory by the standard formula

$$\tau=\frac{\theta_{\mathrm{YM}}}{2\pi}+\frac{2\pi i}{g_{\mathrm{YM}}^2}. \tag{2.146}$$

The action (2.145) is defined so that θ_{YM} couples precisely to the instanton number

$$N=\frac{1}{8\pi^2}\int \mathrm{tr}\,F\wedge F, \tag{2.147}$$

normalized to be an integer on a four-manifold without boundary. This ensures that the theory is invariant under $\tau\to\tau+1$, which corresponds to $\theta_{\mathrm{YM}}\to\theta_{\mathrm{YM}}+2\pi$.

If we include an O3 plane that reduces the gauge symmetry from $\mathrm{U}(n)$ to $\mathrm{O}(n)$, then we write the action in the same way, with tr now representing a trace in the fundamental representation of $\mathrm{O}(n)$. But since we have to divide the action by 2, we express the gauge theory parameters in terms of τ not by (2.146) but by

$$\frac{\tau}{2} = \frac{\theta_{\mathrm{YM}}}{2\pi} + \frac{2\pi i}{g_{\mathrm{YM}}^2}. \tag{2.148}$$

We write

$$\frac{\tau}{2} = \tau_{\mathrm{YM}}, \tag{2.149}$$

where τ_{YM} is expressed in terms of g_{YM} and θ_{YM} in the usual way. An important detail now is that the quantity N, which is $\mathbb{Z}$-valued in $\mathrm{U}(n)$ gauge theory, takes values in[34] $2\mathbb{Z}$ in $\mathrm{O}(n)$ gauge theory for $n \geq 4$. Because of this, the $\mathrm{O}(n)$ gauge theory is invariant under $\tau \to \tau + 1$, even though θ_{YM} couples to $N/2$.

Next consider the orientifold plane to be $\mathrm{O3}^+$, reducing the gauge symmetry from $\mathrm{U}(n)$ to $\mathrm{Sp}(n)$ (here n must be even). The action is still defined as in eqn. (2.145), now with tr representing the trace in the fundamental representation of $\mathrm{Sp}(n)$. Furthermore, the coupling parameter τ of Type IIB superstring theory is still related to the gauge theory parameters as in (2.148). Now, however, the quantity N is integer-valued (a minimal $\mathrm{Sp}(n)$ instanton is an $\mathrm{SU}(2)$ instanton of instanton number 1 embedded in $\mathrm{Sp}(2) \cong \mathrm{SU}(2)$), so the operation $\tau \to \tau + 1$ of the underlying string theory is not a symmetry of the gauge theory. Instead, this operation maps an $\mathrm{O3}^+$ orientifold plane to a $\widetilde{\mathrm{O3}}^+$-plane, in which the gauge group is still $\mathrm{Sp}(n)$ but the relation between string theory and gauge theory parameters is shifted from (2.148) to

$$\frac{\tau + 1}{2} = \frac{\theta_{\mathrm{YM}}}{2\pi} + \frac{2\pi i}{g_{\mathrm{YM}}^2}. \tag{2.150}$$

The term $\mathrm{Sp}'(n)$ gauge theory is an abbreviation for $\mathrm{Sp}(n)$ gauge theory with coupling

[34] For $n \geq 4$, an $\mathrm{O}(n)$ instanton of minimal instanton number can be embedded in an $\mathrm{SO}(4)$ subgroup. An $\mathrm{SO}(4)$ instanton of minimal instanton number (on $\mathbb{R}^4$; we do not consider here effects associated to the second Stieffel-Whitney class) is simply an $\mathrm{SU}(2)$ instanton of instanton number 1, embedded in one of the two factors of $\mathrm{Spin}(4) \cong \mathrm{SU}(2) \times \mathrm{SU}(2)$. Upon embedding $\mathrm{O}(n)$ in $\mathrm{U}(n)$, the $\mathrm{O}(n)$ instanton constructed this way is a $\mathrm{U}(n)$ instanton of instanton number 2, explaining why the instanton number normalized as in (2.147) is an even integer in $\mathrm{O}(n)$. In the case of $\mathrm{O}(3)$, there is not room for the construction just described, and the minimal instanton has $N = 4$.

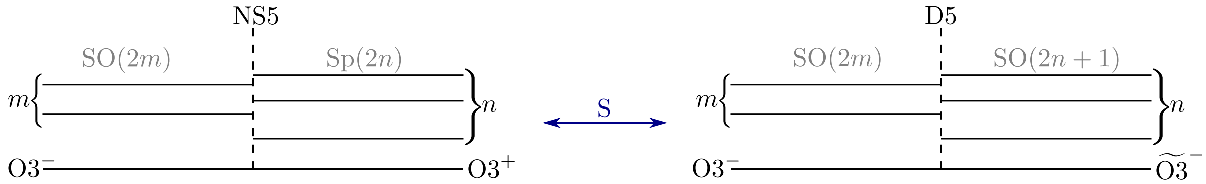

Figure 2.9: The brane configurations that realize the electric and magnetic theory for the four-dimensional construction of the $\mathrm{OSp}(2m|2n)$ Chern-Simons theory.

parameters related in this way to the underlying string theory parameters.

2.5.2 The Even Orthosymplectic Theory

Now we begin our study of the D3-O3 system interacting with a fivebrane. On the left of fig. 2.9, we sketch an $\mathrm{O3}^-$-plane that converts to an $\mathrm{O3}^+$-plane in crossing an NS5-brane. The gauge group is therefore $\mathrm{SO}(2m)$ on the left and $\mathrm{Sp}(2n)$ on the right, where m and n are the relevant numbers of D3-branes. In the topologically twisted version of the theory, along the defect, one sees a Chern-Simons theory of the supergroup $\mathrm{OSp}(2m|2n)$. After the orientifold projection, the action can be written just as in eqn. (2.35):

$$I = \frac{i\mathcal{K}_{\mathfrak{osp}}}{4\pi}\int_W \mathrm{Str}\left(\mathcal{A}\mathrm{d}\mathcal{A} + \frac{2}{3}\mathcal{A}^3\right) + \{Q, \dots\}\,, \tag{2.151}$$

Now Str denotes the supertrace in the fundamental representation of the orthosymplectic group. This follows by simply projecting the effective action described in section 2.2 onto the part that is invariant under the orientifold projection. The expression for $\mathcal{K}_{\mathfrak{osp}}$ in terms of string theory parameters τ, ϑ is the same as in equation (2.7) except for a factor of 2 associated to the orientifolding:

$$\frac{\tau}{2} = \tau_{\mathrm{YM}} = \mathcal{K}_{\mathfrak{osp}} \cos\vartheta\, \mathrm{e}^{i\vartheta}. \tag{2.152}$$

Note that the bosonic part of the Chern-Simons action in (2.151) can be also expressed as

$$\frac{i\mathcal{K}_{\mathfrak{osp}}}{4\pi}\int_W \mathrm{Tr}\left(\mathcal{A}_b \mathrm{d}\mathcal{A}_b + \frac{2}{3}{\mathcal{A}_b}^3\right) = i\mathcal{K}_{\mathfrak{osp}}\Big(\mathrm{CS}(A_{\mathfrak{sp}}) - 2\mathrm{CS}(A_{\mathfrak{so}})\Big)\,, \tag{2.153}$$

where the Chern-Simons functionals $\mathrm{CS}(A_{\mathfrak{sp}})$ and $\mathrm{CS}(A_{\mathfrak{so}})$ are normalized to take values in $\mathbb{R}/2\pi\mathbb{Z}$ for simply connected gauge groups and $m > 1$.

Now we apply the usual S-duality transformation $\tau \to \tau^\vee = -1/\tau$. As indicated in the figure, this leaves the $\mathrm{O3}^-$-plane invariant but converts the $\mathrm{O3}^+$-plane to an $\widetilde{\mathrm{O3}}^-$-plane; now the gauge group is $\mathrm{SO}(2m)$ on the left and $\mathrm{SO}(2n+1)$ on the right. What we get this way is a magnetic dual of Chern-Simons theory of $\mathrm{OSp}(2m|2n)$.

The appropriate effective action to describe this situation is found by simply projecting the effective action described in section 2.4.3 onto the part invariant under the orientifold projection. There is no analog of the case $m = n$ that was important in section 2.4.3, since $2m$ never coincides with $2n+1$. The condition analogous to $|n-m| \geq 2$ is $|2m-(2n+1)| \geq 3$. If this is the case, the appropriate description involves a Nahm pole associated to an irreducible embedding $\mathfrak{su}(2) \to \mathfrak{so}(|2m-(2n+1)|)$. The Nahm pole appears on the left or the right of the defect depending on the sign of $2m-(2n+1)$. What commutes with the Nahm pole is an $\mathrm{SO}(w)$ gauge theory theory that fills all space; here w is the smaller of $2m$ and $2n+1$. If $|2m-(2n+1)| = 1$, then as in section 2.4.3, there is no Nahm pole and the vector multiplets that transform in the fundamental representation of $\mathrm{SO}(w)$ obey Dirichlet boundary conditions along the defect.

The action can still be expressed as in (2.97)

$$I_{\text{magnetic}} = \frac{i\theta^\vee_{\mathrm{YM}}}{8\pi^2}\int \mathrm{tr}\,(F\wedge F) + \{Q,\dots\}, \tag{2.154}$$

where now tr is the trace in the fundamental representation of the orthogonal group, and

$\tau^\vee_{\rm YM} = \theta^\vee_{\rm YM}/2\pi + 4\pi i/(g^\vee_{\rm YM})^2$ is related to the underlying string theory parameters by

$$\tau^\vee_{\rm YM} = \frac{1}{2}\tau^\vee = -\frac{1}{2\tau}. \tag{2.155}$$

We recall from section 2.5.1 that the instanton number $N^\vee = (1/8\pi^2)\int \mathrm{tr}\, F \wedge F$ takes even integer values in the case of an orthogonal gauge group. Hence the natural instanton-counting parameter is

$$q = \exp(-2i\theta^\vee_{\rm YM}), \tag{2.156}$$

in the sense that a field of $N^\vee = 2r$ contributes $\pm q^r$ to the path integral (as usual the sign depends on the sign of the fermion determinant).

The variable q can be expressed in terms of the canonical parameter $\mathcal{K}_{\mathfrak{osp}}$ of the electric description. In (2.100), we have obtained $\mathrm{Re}\,(\tau^\vee) = -1/\mathcal{K}$, where $\mathcal{K}$ is the canonical parameter for the theory with no orientifolds. In the orientifolded theory, the canonical parameter $\mathcal{K}_{\mathfrak{osp}}$ that appears in the action (2.151) is one-half of that. Hence, using equation (2.155), we find that

$$\frac{\theta^\vee_{\rm YM}}{2\pi} = \mathrm{Re}\,\tau^\vee_{\rm YM} = \frac{1}{2}\mathrm{Re}\,\tau^\vee = -\frac{1}{2\mathcal{K}} = -\frac{1}{4\mathcal{K}_{\mathfrak{osp}}}\,, \tag{2.157}$$

and therefore the definition (2.156) gives

$$q = \exp\left(\frac{\pi i}{\mathcal{K}_{\mathfrak{osp}}}\right). \tag{2.158}$$

By contrast, Chern-Simons theory or two-dimensional current algebra for a purely bosonic group G with Lie algebra $\mathfrak{g}$ is naturally parametrized by

$$q_{\mathfrak{g}} = \exp\left(\frac{2\pi i}{n_{\mathfrak{g}}\mathcal{K}_{\mathfrak{g}}}\right), \tag{2.159}$$

where $n_{\mathfrak{g}}$ is the ratio of length squared of long and short roots of $\mathfrak{g}$. (This is also the natural instanton-counting parameter in the magnetic dual description of this theory [6].)

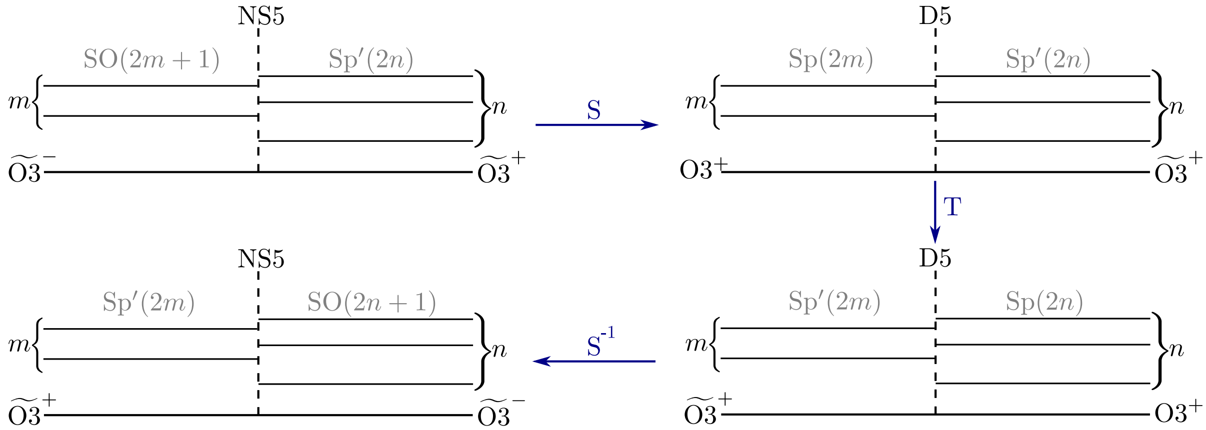

Figure 2.10: The figure in the upper left corner shows the brane configuration, which gives the four-dimensional construction for the $\mathrm{OSp}(2m+1|2n)$ Chern-Simons theory. The other figures are obtained by acting with various elements of the $\mathrm{SL}(2,\mathbb{Z})$ S-duality group. In particular, the transformation $S^{-1}TS$ maps the configuration in the upper left to the one in the lower left.

The parameter q defined in eqn. (2.158) is an analog of this, with $n_{\mathfrak{g}}$ replaced by the ratio of length squared of the longest and shortest bosonic roots; for $\mathfrak{osp}(2m|2n)$, this ratio is equal to 2.

2.5.3 The Odd Orthosymplectic Theory

2.5.3.1 Preliminaries

Now we will repeat the analysis of the D3-O3-NS5 system, with just one important change: we give the O3-planes a discrete RR flux. As depicted in the upper left of fig. 2.10, we take the O3-plane to be of type $\widetilde{\mathrm{O3}}^-$ to the left of the NS5-brane and (therefore) of type $\widetilde{\mathrm{O3}}^+$ to the right. The gauge groups realized on the D3-O3 system on the two sides of the defect are $\mathrm{SO}(2m+1)$ and $\mathrm{Sp}'(2n)$, so this configuration describes an analytically-continued version of $\mathrm{OSp}(2m+1|2n)$ Chern-Simons theory. Up to a point, the four-dimensional gauge theory description of this system can be found just as in section 2.5.2: we restrict the fields of the familiar $\mathrm{U}(2m+1|2n)$ system to be invariant under the orientifold projection, and divide the action by 2.

However, there are some crucial subtleties that do not have a close analog in the previous case:

(1) The gauge theory theta-angle jumps by π in crossing the defect, because the gauge theory on the right is of type $\mathrm{Sp}'(2m)$. By itself, this would spoil the supersymmetry of the defect system, since when one verifies supersymmetry at the classical level, one assumes that τ_{YM} is continuous in crossing the defect.[35]

(2) This suggests that a quantum anomaly may be relevant, and in fact there is one: in three dimensions, the bifundamental hypermultiplet of $\mathrm{SO}(2m+1)\times\mathrm{Sp}(2n)$ that is supported on the three-dimensional defect suffers from a global anomaly.

These two problems, in fact, compensate each other. Indeed, the anomalous fermionic path-integral can be made well-defined by adding a half-integer Chern-Simons term, that is, by considering the combination

$$\mathrm{Pf}(\not{D})\exp\left(\frac{i}{2}\mathrm{CS}(A)\right). \tag{2.160}$$

Here $\mathrm{Pf}(\not{D})$ is the Pfaffian of the fermionic kinetic operator, which changes sign under large $\mathrm{Sp}(2n)$ gauge transformations. The half-integral Chern-Simons term, sitting at the defect, has the same local variation, as a bulk theta-term with theta-angle equal to π. Thus, adding the half-integral Chern-Simons term simultaneously restores the invariance under the gauge symmetry and under the supersymmetry.

The combination (2.160) is what is typically used in physical literature. However, the overall sign of this expression is not well-defined. It is better to use the APS index theorem to replace this combination by the eta-invariant, which is gauge-invariant and well-defined. So, we write instead

$$|\mathrm{Pf}(\not{D})|\exp(i\pi\eta'/2) \tag{2.161}$$

where

$$\exp(i\pi\eta'/2)=\exp(i\pi\widehat{\eta}/2-im\mathrm{CS}(A_{\mathfrak{sp}})-2in\mathrm{CS}(A_{\mathfrak{so}}))\,, \tag{2.162}$$

[35]Supersymmetry actually allows certain discontinuities [23], but not a jump in θ_{YM} at fixed ϑ.

and $\widehat{\eta}$ is one-half of the eta-invariant of the kinetic operator of the 3d fermions. Under local variations of the gauge field, the expression (2.161) changes in the same way as (2.160).

2.5.3.2 The Dual Theory

We can find now a magnetic dual of $\mathrm{OSp}(2m+1|2n)$ Chern-Simons theory by applying the S-duality transformation $\tau \to -1/\tau$. Its action on the brane configuration is shown in the upper part of fig. 2.10. The new string coupling is $\tau^\vee = -1/\tau$. The gauge groups are now $\mathrm{Sp}(2m)$ in M_ℓ and $\mathrm{Sp}'(2n)$ in M_r. We continue to use the notation $\tau_{\mathrm{YM}}^\vee = \frac{1}{2}\tau^\vee$ for the gauge coupling. The minimal instanton number for the symplectic group is 1, so the natural instanton-counting parameter analogous to (2.156) is $q = \exp(-i\theta_{\mathrm{YM}}^\vee)$. Using (2.157), this can be presented as

$$q = \exp\left(\frac{\pi i}{2\mathcal{K}_{\mathfrak{osp}}}\right). \tag{2.163}$$

This agrees with the general definition (2.159), since the ratio of length squared of the longest and shortest bosonic roots for the odd orthosymplectic algebras is $n_{\mathfrak{g}} = 4$.

In the "magnetic" description, one of the orientifold planes is again of type $\widetilde{\mathrm{O3}}^+$, which means that the θ_{YM} jumps by π upon crossing the defect. As in the electric description, this jump appears to violate supersymmetry. The resolution is similar to what it was in the electric description. First we consider the case that $m = n$. For this case, the gauge group is simply $\mathrm{Sp}(2n)$ filling all of spacetime. There is a fundamental hypermultiplet supported on the defect. Its Pfaffian has the sign anomaly, similarly to the one mentioned in the previous section. The anomaly is canceled roughly speaking via a half-integral Chern-Simons term supported on the defect, or more accurately via an η-invariant. The combined path integral involving the fermion Pfaffian, the η-invariant, and the jump in θ_{YM} (as well as other factors) is gauge-invariant and supersymmetric. The factors involved in the anomaly cancellation are the familiar ones from eqn. (2.161):

$$|\mathrm{Pf}(\not{D})| \exp(i\pi\eta'/2) \exp\left(-\frac{i}{8\pi}\int_{M_r} \mathrm{tr}_{\mathfrak{sp}}\, F \wedge F\right). \tag{2.164}$$

(Here we have written explicitly the term, corresponding to the jump of the theta-angle.) According to the APS index theorem, the product of the last two factors equals ± 1 (possibly multiplied by a factor that only depends on $\int_{M_r} R^2$). This factor of ± 1 must be incorporated in the sum over instanton solutions. We denote it as

$$\mathrm{sign}_{y\geq 0} = \exp(i\pi\eta'/2)\exp\left(-\frac{i}{8\pi}\int_{M_r}\mathrm{tr}_{\mathfrak{sp}}\, F\wedge F\right). \tag{2.165}$$

What happens if $n \neq m$? In this case, there are no hypermultiplets supported on the defect. Instead, there is a jump in the gauge group in crossing the defect. Along the defect there is a Nahm pole, associated to an irreducible embedding of $\mathfrak{su}(2)$ in $\mathfrak{sp}(|2n-2m|)$. As usual, the pole is on the side on which the gauge group is larger. The gauge group that is unbroken throughout all space is $\mathrm{Sp}(2s)$, where s is the smaller of n and m.

At first sight, it is not clear how to generalize (2.164) to $n \neq m$. If there are no fermions supported on the defect, how can we possibly use an anomaly in a fermion determinant as part of a mechanism to compensate for a jump in θ_{YM} by π? To understand what must happen, recall that we can deform from $n = m$ to $n \neq m$ by Higgsing – by moving some of the D3-branes (on one side or the other of the defect) away from the rest of the system. When we do this, the bifundamental hypermultiplet which is responsible for some of the interesting factors in (2.164) does not simply vanish in a puff of smoke. It mixes with some of the bulk degrees of freedom and gains a large mass. When this happens, whatever bulk degrees of freedom remain will carry whatever anomaly existed before the Higgsing process.

So the resolution of the puzzle must involve a subtlety in the fermion path integral for $n \neq m$. Going back to (2.164), naively $\not{D}$ is the Dirac operator just of the defect fermions and η' is one-half their η-invariant. There are also bulk fermions, but they have no anomaly and vanishing η-invariant, so it does not seem interesting to include them in (2.164). However, precisely because they have no anomaly and vanishing η-invariant, we could include them in (2.164) (and their coupling to the defect fermions) without changing anything. This is a better starting point to study the Higgsing process, since Higgsing disturbs the decoupling.

Upon Higgsing, the first two factors in eqn. (2.164) keep their form, but some modes become massive and – in the limit that $|2n-2m|$ D3-branes are removed on one side or the other – the defect fermions disappear and we are left with an expression of the same form as (2.164), but now the Pfaffian and the η-invariant are those of the bulk fermions in the presence of the Nahm pole. The Dirac operator of the bulk fermions in the presence of the Nahm pole can be properly defined, with some subtlety, as an elliptic differential operator [70]. This gives a framework in which one could investigate its Pfaffian and η-invariant. For the theory that we are discussing here to make sense, there must be an anomaly in the sign of the Pfaffian of this operator, and it must also have a nontrivial η-invariant that compensates in the familiar way for the jump in θ_{YM}. These points have not yet been investigated, but there do not seem to be any general principles that exclude the required behavior.

2.5.4 The Framing Anomalies

In section 2.4.3.1 we have verified that our constructions predict the correct value for the global framing anomaly for the Chern-Simons theory of the unitary supergroup. Here we repeat the same analysis for the orthosymplectic gauge group.

In the non-simply-laced case, the analog of the formula (2.115) for the framing factor is

$$\exp\left(2\pi i\,\mathrm{sign}(k)\,\mathrm{sdim}\,SG/24\right)\cdot q^{-n_{\mathfrak{g}}h_{\mathfrak{sg}}\,\mathrm{sdim}\,\mathrm{SG}/24}. \tag{2.166}$$

The difference with the simply-laced case is the factor of $n_{\mathfrak{g}}$ in the exponent, which compensates for the analogous factor in the definition (2.159) of the q variable. As usual in this chapter, we will ignore the one-loop contribution to the anomaly, and focus only on the power of q. To compare the anomalies for different groups, it is convenient to express them in terms of the theta-angle of the magnetic theory. What we need to know is that for a theory with a bosonic gauge group the variable q is defined as $q=\exp(-2i\theta^{\vee}_{\mathrm{YM}})$, if the gauge group in the magnetic description is orthogonal, and as $q = \exp(-i\theta^\vee_{\mathrm{YM}})$, if this group is symplectic. We have explained the reason behind this definition, when we discussed the magnetic theories for the orthosymplectic supergroups.

Consider first the even orthosymplectic algebra $\mathfrak{osp}(2m|2n)$. As we recalled in section 2.4.3.1, the framing anomaly in the magnetic description comes from the peculiarities of the definition of the instanton number in the presence of the Nahm pole. We set $r = n - m$. For $r > 0$, the Nahm pole in the magnetic theory is embedded into an $\mathfrak{so}(2r+1)$ subalgebra of $\mathfrak{so}(2n+1)$. This means that the framing anomaly depends only on r and not on m; setting $m = 0$, we reduce to the magnetic dual of $\mathrm{Sp}(2r)$ Chern-Simons theory and we should get the same framing anomaly. The anomaly factor for the orthosymplectic case is expected to be

$$q_{\mathfrak{osp}}^{-n_{\mathfrak{osp}(2m|2n)} h_{\mathfrak{osp}} \mathrm{sdimOSp}/24} = \exp\left(-4i\theta^\vee_{\mathrm{YM}} h_{\mathfrak{osp}} \mathrm{sdim}\,\mathrm{OSp}/24\right). \tag{2.167}$$

For the symplectic gauge group this factor is

$$q_{\mathfrak{sp}}^{-n_{\mathfrak{sp}} h_{\mathfrak{sp}} \mathrm{dimSp}/24} = \exp\left(-4i\theta^\vee_{\mathrm{YM}} h_{\mathfrak{sp}} \mathrm{dim}\,\mathrm{Sp}/24\right). \tag{2.168}$$

The two expressions agree, since

$$h_{\mathfrak{osp}(2m|2n)}\, \mathrm{sdim}\,\mathrm{OSp}(2m|2n) = h_{\mathfrak{sp}(2r)}\, \mathrm{dim}\,\mathrm{Sp}(2r) = 2r(r+1/2)(r+1). \tag{2.169}$$

This identity is the analog of (2.116); see Table 2.2 for the numerical values.

If $r < 0$, the Nahm pole lives in the $\mathfrak{so}(-2r-1)$ subalgebra on the other side of the defect. This is the same Nahm pole that would arise in the magnetic dual of $\mathrm{SO}(-2r)$ Chern-Simons theory, so the framing anomaly should agree with that theory. For the bosonic theory with the even orthogonal gauge group we have

$$q_{\mathfrak{so}}^{-h_{\mathfrak{so}} \dim \mathrm{SO}} = \exp\left(-2i\theta^\vee_{\mathrm{YM}} h_{\mathfrak{so}} \mathrm{dim}\,\mathrm{SO}/24\right). \tag{2.170}$$

This agrees with (2.167), since

$$h_{\mathfrak{osp}(2m|2n)}\,\mathrm{sdim}\,\mathrm{OSp}(2m|2n) = -\frac{1}{2}h_{so(-2r)}\mathrm{dim}\,\mathrm{SO}(-2\mathrm{r}) = 2r(r+1/2)(r+1). \qquad (2.171)$$

The minus sign appears here, because the Nahm pole for the orthosymplectic theory with $r<0$ is on the left side of the defect.

Alternatively, we could think of the $\mathfrak{so}(-2r-1)$ Nahm pole as corresponding to the $\mathrm{Sp}(-2r-2)$ electric theory. This would give the same result.

Let us repeat the same story for the odd orthosymplectic superalgebra $\mathfrak{osp}(2m+1|2n)$. Again, we set $r=n-m$. The Nahm pole is embedded in the $\mathfrak{sp}(2|r|)$ subalgebra. In the purely bosonic case, the same embedding would arise for the $SO(2|n-m|+1)$ electric theory. Therefore, we would expect that the global framing anomaly for the superalgebra case is the same as for this purely bosonic Lie algebra, at least above one loop. The framing factor for the odd orthosymplectic case should be

$$q_{\mathfrak{osp}}^{-n_{\mathfrak{osp}(2m+1|2n)}h_{\mathfrak{osp}}\mathrm{sdim OSp}/24} = \exp\left(-4i\theta_{\mathrm{YM}}^{\vee}h_{\mathfrak{osp}}\mathrm{sdim}\,\mathrm{OSp}/24\right). \qquad (2.172)$$

In the $\mathrm{SO}(2|r|+1)$ the answer is

$$q_{\mathfrak{so}}^{-n_{\mathfrak{so}}h_{\mathfrak{so}}\mathrm{dim}\,\mathrm{SO}} = \exp\left(-2i\theta_{\mathrm{YM}}^{\vee}h_{\mathfrak{so}}\mathrm{dim}\,\mathrm{SO}/24\right). \qquad (2.173)$$

The two expressions (2.172) and (2.173) agree, since from Table 2.2 we have

$$h_{\mathfrak{osp}(2m+1|2n)}\,\mathrm{sdim}\,\mathrm{OSp}(2m+1|2n) = \frac{1}{2}h_{so(2|r|+1)}\mathrm{dim}\,\mathrm{SO}(2|\mathrm{r}|+1) = 2r(r^2-1/4). \qquad (2.174)$$

The sign in the right hand side changes, depending on the sign of r, in accord with the fact that the Nahm pole is on the right or on the left of the defect. Note also, that up to this change of sign the formula is symmetric under the exchange of m and n. This reason for this symmetry will become clear in section 2.5.5.

2.5.5 Another Duality

So far in this chapter, we have just exploited the duality $S : \tau \to -1/\tau$, exchanging NS5-branes with D5-branes. The full S-duality group $\mathrm{SL}(2, \mathbb{Z})$ of Type IIB superstring theory contains much more. In particular, it has a non-trivial subgroup that maps an NS5-brane to itself. This subgroup is generated by the element

$$S^{-1}TS = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}. \tag{2.175}$$

That this element maps an NS5-brane to itself follows from the action of duality on fivebrane charges given in eqn. (2.90). (Concretely, S converts an NS5-brane to a D5-brane, T leaves fixed the D5-brane, and S^{-1} maps back to an NS5-brane.) This transformation will map a D3-NS5 system, possibly with an O3-plane, to a system of the same type. In the approach to Chern-Simons theories followed in the present chapter, this transformation will map an "electric" description to another "electric" description, and thus it will give a duality of Chern-Simons theories (analytically continued away from integer levels).

Let us first see what this duality does to a D3-NS5 system, associated to the supergroup $\mathrm{U}(m|n)$. The operation $S^{-1}TS$ maps D3-branes and NS5-branes to themselves, so it maps the Chern-Simons theory of $\mathrm{U}(m|n)$ to itself, while transforming the canonical parameter according to (2.95), which in this case gives

$$\frac{1}{\mathcal{K}} \to \frac{1}{\mathcal{K}} - 1 = \frac{1}{\mathcal{K}'}. \tag{2.176}$$

This transformation leaves fixed the variable $q = \exp(2\pi i/\mathcal{K})$ in terms of which the knot invariants are usually expressed. (In fact, the symmetry (2.176) can be viewed as the reason that the knot invariants can be expressed in terms of q rather than being more general functions of $\mathcal{K}$.) This duality acts trivially on line operators of $\mathrm{U}(m|n)$. To argue this, we just observe that T can be understood classically – as a 2π shift in θ_{YM} – and does not affect the model solution that is used to define a line operator.

The action of STS^{-1} on a surface operator can be determined by looking at the behavior far away from the defect. We have

$$\begin{pmatrix} \alpha \\ \eta \end{pmatrix} \overset{S}{\to} \begin{pmatrix} \eta \\ -\alpha \end{pmatrix} \overset{T}{\to} \begin{pmatrix} \eta \\ \eta - \alpha \end{pmatrix} \xrightarrow{S^{-1}} \begin{pmatrix} \alpha - \eta \\ \eta \end{pmatrix}. \tag{2.177}$$

Using the relation (2.85), the action on the weight λ can be conveniently written

$$\frac{\lambda'}{\mathcal{K}'} = \frac{\lambda}{\mathcal{K}}. \tag{2.178}$$

Since knot invariants computed using surface operators by the procedure explained in section 2.4.6 only depend on the ratio $\lambda/\mathcal{K}$, this shows that they are invariant under $S^{-1}TS$. Using the relation (2.176) between $\mathcal{K}'$ and $\mathcal{K}$, eqn. (2.178) is equivalent to

$$\lambda' = \lambda + \mathcal{K}'\lambda. \tag{2.179}$$

Let us check whether these formulas are consistent with the idea that if λ is integral, the same knot and link invariants can be computed using either line operators or surface operators. $S^{-1}TS$ acts trivially on the weight of a line operator, but acts on the weight of a surface operator as in (2.179). However, knot invariants computed from surface operators are unchanged in shifting λ by $\mathcal{K}$ times an integral cocharacter. Since the groups $\mathrm{U}(n)$ and $\mathrm{U}(m)$ are selfdual, if λ is an integral character, it is also an integral cocharacter.

Now let us apply this duality to the configuration of fig. 2.9, which corresponds to an even orthosymplectic group $\mathrm{OSp}(2m|2n)$. The transformation $S^{-1}TS$ maps the O3-planes that appear in this configuration to themselves, so again it maps Chern-Simons theory of

OSp$(2m|2n)$ to itself. The canonical parameter $\mathcal{K}_{\mathfrak{osp}}$ of the orthosymplectic theory was defined as one-half of the object $\mathcal{K}$ defined in section 2.4, so the transformation rule (2.176) can be written

$$\frac{1}{\mathcal{K}_{\mathfrak{osp}}} \to \frac{1}{\mathcal{K}_{\mathfrak{osp}}} - 2 = \frac{1}{\mathcal{K}'_{\mathfrak{osp}}}\,, \tag{2.180}$$

Therefore, the natural Chern-Simons parameter $q = \exp(\pi i/\mathcal{K}_{\mathfrak{osp}})$, defined in eqn. (2.158), is invariant, just as for the unitary case. The Chern-Simons theory again is simply mapped to itself. It takes a little more effort to understand the duality action on line and surface operators. For this reason, the discussion of the operator mapping will be presented in a separate section 2.5.6. There we will find that, unlike for the unitary superalgebra, the duality acts on the set of line operators by a non-trivial involution.

For the odd orthosymplectic group OSp$(2m+1|2n)$, matters are more interesting. The action of $S^{-1}TS$ on the brane configuration associated to OSp$(2m+1|2n)$ is described in fig. 2.10. Chasing clockwise around the figure from upper left to lower left, we see that the duality maps a brane configuration associated to OSp$(2m+1|2n)$ to one associated to OSp$(2n+1|2m)$. Since the gauge group changes, this is definitely a non-trivial duality of (analytically-continued) Chern-Simons theories. For example, setting $n = 0$, we get a duality between Chern-Simons theory of the ordinary bosonic group O$(2m+1)$ and Chern-Simons theory of the supergroup OSp$(1|2m)$. How does this duality act on the natural variable q that parametrizes the knot invariants? For the odd orthosymplectic group, the natural variable in terms of which the knot invariants are expressed is $q = \exp(\pi i/2\mathcal{K}_{\mathfrak{osp}})$, introduced in eqn. (2.163). The transformation (2.180) acts on this variable by[36]

$$q \to -q. \tag{2.181}$$

[36]There is a subtlety here. The Killing form for a superalgebra can be defined with either sign. Since the duality maps theories with, say, Sp group at $y > 0$ to Sp group at $y < 0$, it exchanges the two choices. If we want to define the sign of the Killing form to be always positive, say, for the $\mathfrak{sp}$ subalgebra, we should rather say that q maps to $-q^{-1}$. What is written in the text assumes that the sign of the Killing form in M_ℓ or M_r is unchanged in the duality.

The minus sign means that the duality we have found exchanges weak and strong coupling. Indeed, in three-dimensional Chern-Simons theory, the weak coupling limit is $q \to 1$, and $q \to -1$ is a point of strong coupling.

It is inevitable that the duality must map weak coupling to strong coupling, since the classical representation theories of $\mathrm{OSp}(2m+1|2n)$ and $\mathrm{OSp}(2n+1|2m)$ are not equivalent. A duality mapping weak coupling to weak coupling would imply an equivalence between the two classical limits, but this does not hold.

Some instances of the duality predicted by the brane construction have been discovered previously. For $n = 0$ and $m = 1$, the relation between knot invariants has been discussed in [17]; for $n = 0$ and any m, this subject has been discussed in [15] in a different language. For related discussion from the standpoint of quantum groups see [16], and see [71] for associated representation theory. We will say more on some of these results in section 2.5.6.

Now let us look at the same duality in the magnetic dual language. Our two electric theories are sketched in the upper and lower left of fig. 2.10, and the corresponding magnetic duals, obtained by acting with S, are shown in the upper and lower right of the same figure. One involves an $\mathrm{Sp}(2m) \times \mathrm{Sp}'(2n)$ gauge theory, and the other involves an $\mathrm{Sp}'(2m) \times \mathrm{Sp}(2n)$ gauge theory. There is no change in the gauge groups, the localization equations, or in the hypermultiplet fermions if $n = m$ or in the Nahm pole singularity if $n \neq m$. The only difference is that in one case θ_{YM} differs on the right by π from the underlying Type IIB theta-angle, and in the other case, it differs on the left by π from the underlying Type IIB theta-angle. In the upper right of fig. 2.10, a solution of the localization equations with instanton number $N^\vee$ is weighted by the product of $q^{N^\vee}$ with the sign factor of eqn. (2.165). There is an additional sign that we will call $(-1)^f$; this is the sign of the determinant of the operator obtained by linearizing around a solution of the localization equations. This factor is not affected by the duality. The combination is

$$(-1)^f q^{N^\vee} \mathrm{sign}_{y \geq 0} = (-1)^f q^{N^\vee} \exp(i\pi\eta'/2) \exp\left(-\frac{i}{8\pi}\int_{M_r} \mathrm{tr}_{\mathfrak{sp}}\, F \wedge F\right). \tag{2.182}$$

On the lower left of the figure, the sign factor $\mathrm{sign}_{y\geq 0}$ is replaced with

$$\mathrm{sign}_{y\leq 0} = \exp(i\pi\eta'/2)\exp\left(+\frac{i}{8\pi}\int_{M_\ell}\mathrm{tr}_{\mathfrak{sp}}\, F\wedge F\right). \tag{2.183}$$

We also have to replace q with $-q$. So (2.182) is replaced with

$$(-1)^f(-q)^{N^\vee}\exp(i\pi\eta'/2)\exp\left(+\frac{i}{8\pi}\int_{M_\ell}\mathrm{tr}_{\mathfrak{sp}}\, F\wedge F\right). \tag{2.184}$$

The two expressions (2.182) and (2.184) are equal, since

$$N^\vee = N^\vee_\ell + N^\vee_r, \tag{2.185}$$

with

$$N^\vee_\ell = \frac{1}{8\pi^2}\int_{M_\ell}\mathrm{tr}_{\mathfrak{sp}}\, F\wedge F,\qquad N^\vee_r = \frac{1}{8\pi^2}\int_{M_r}\mathrm{tr}_{\mathfrak{sp}}\, F\wedge F. \tag{2.186}$$

The above formulas can be written more elegantly by using the Atiyah-Patodi-Singer (APS) index theorem [69] for the Dirac operator on a manifold with boundary. This will also be useful later. We let ν_ℓ (or ν_r) be the index of the Dirac operator on M_ℓ (or M_r), acting on spinors with values in the fundamental representation of $\mathrm{Sp}(2n)$ (or $\mathrm{Sp}(2m)$). This index is defined by counting zero-modes of spinor fields that are required to be square-integrable at infinite ends of M_ℓ or M_r, and to obey APS global boundary conditions along the finite boundary W. The APS index theorem gives

$$\begin{aligned}(-1)^{\nu_\ell} &= \exp(i\pi\eta'/2)\exp\left(+\frac{i}{8\pi}\int_{M_\ell}\mathrm{tr}_{\mathfrak{sp}}\, F\wedge F\right)\\ (-1)^{\nu_r} &= \exp(i\pi\eta'/2)\exp\left(-\frac{i}{8\pi}\int_{M_r}\mathrm{tr}_{\mathfrak{sp}}\, F\wedge F\right).\end{aligned} \tag{2.187}$$

Thus the factors weighting a given solution in the dual constructions of fig. 2.10 are respectively

$$(-1)^f(-q)^{N^\vee}(-1)^{\nu_\ell} \tag{2.188}$$

and

$$(-1)^f q^{N^\vee}(-1)^{\nu_r}. \tag{2.189}$$

The most convenient way to compare these two formulas is as follows. Let ν be the index of the Dirac operator on the whole four-manifold $M = M_\ell \cup M_r$. Additivity of the index under gluing gives

$$\nu = \nu_\ell + \nu_r. \tag{2.190}$$

But we also have

$$\nu = N^\vee. \tag{2.191}$$

To obtain this formula, one can first deform the gauge field into an $\mathrm{Sp}(2s)$ subgroup, where $s = \min(n, m)$, so as not to have to consider the jump from n to m (which is not present in standard formulations of index problems). Then (2.191) is a consequence of the ordinary Atiyah-Singer index theorem, or of the APS theorem on the noncompact four-manifold $M = W \times \mathbb{R}$ (with the contributions of the ends at infinity canceling). It follows from these statements that

$$(-1)^{N^\vee}(-1)^{\nu_r} = (-1)^{\nu_\ell}, \tag{2.192}$$

showing that the two descriptions do give the same result.

We now proceed to describe the action of the duality on line and surface operators of the orthosymplectic theory.

2.5.6 Duality Transformation Of Orthosymplectic Line And Surface Operators

2.5.6.1 Magnetic Duals Of Twisted Line Operators

Before we can describe the action of the duality on line operators, we need some preparation. In section 2.3.3.3, we have introduced the twisted line operators in the electric description.

One needs to include them in the story to get a consistent picture for the $S^{-1}TS$ duality of line operators in the orthosymplectic theory. For this reason, here we make a digression to describe their magnetic duals.

This question arises already for $\mathrm{U}(m|n)$, so we start there. Consider a knot K in a three-manifold W. W is embedded in a four-manifold M, for example $W \times \mathbb{R}$. The definition of twisted line operators on the electric side depended on the existence of a flat line bundle with some twist c around the knot K. For a generic twist, such a bundle can only exist if the cycle K is trivial in $H_1(M)$. In addition to the twist, the line operator also supports a Wilson operator of the bosonic subgroup with some weight Λ. In the magnetic theory, we propose the following definition for the dual of a twisted operator. Let $\lambda = \Lambda + \rho_{\bar{0}}$ be the quantum-corrected weight. Note that here we use the bosonic Weyl vector for the quantum correction, since Λ was the highest weight of a representation of the bosonic subgroup. For a twisted operator of quantum-corrected weight λ, we define the dual magnetic operator, using the irreducible model solution of section 2.4.4, corresponding to the weight λ, but also make the following modification. For definiteness, let $n \geq m$. Then the $\mathrm{U}(m)$-part of the gauge field is continuous across the three-dimensional defect. Pick a surface Σ bounded by K, or, more precisely, a class[37] in the relative homology $\mathrm{H}_2(M, K)$. The $\mathrm{U}(m)$ bundle is trivialized along the knot K, so it makes sense to evaluate its first Chern class on the class Σ, and to include a factor

$$\exp\left(ic\int_\Sigma \mathrm{tr}\, F/2\pi\right) \tag{2.193}$$

in the functional integral. Here c is an angular variable, which we conjecture to equal the twist of the line operator on the electric side.[38] This proposal can be justified by noting that the insertion (2.193) is essentially an abelian surface operator of type $(0, 0, 0, \eta^\vee)$, with $\eta^\vee$ valued in the center of the Lie algebra of the magnetic gauge group. After doing the

[37] Since K is trivial in the homology, Σ exists, but it might not be unique. If it is not unique, we should probably sum over possible choices. For simple manifolds like $\mathbb{R}^4$ and $\mathbb{R} \times S^3$ that we mostly consider in this chapter, this question does not arise.

[38] Note that one cannot define such twisted operators in the one-sided, purely bosonic theory, because there the gauge bundle is trivialized completely along $y = 0$, and not only along the knot.

S-duality transformation, this becomes an operator of type $(\alpha, 0, 0, 0)$ in the electric theory. The singularity $\alpha d\theta$ in the abelian gauge field can be removed by making a gauge transformation around this surface operator. Such a gauge transformation closes only up to the element $\exp(ic)$ of the center, and therefore introduces a twist by $\exp(ic)$ to the boundary hypermultiplets.

Now let us turn to the orthosymplectic Chern-Simons theory. For the $\mathrm{OSp}(2m|2n)$ case the magnetic gauge group is $\mathrm{SO}(2m) \times \mathrm{SO}(2\mathrm{n}+1)$, and its subgroup which is not broken by the three-dimensional defect is $\mathrm{SO}(N)$, where $N = 2m$ or $N = 2n+1$, depending on m, n. As is clear from the electric description of section 2.3.3.3, for the twisted operator to have a non-zero matrix element, the knot K should be trivial in $H_1(M; \mathbb{Z}_2)$, that is, we should have $K = \partial\Sigma + 2K'$, where Σ is a two-cycle in $H_2(M, K)$, and K' is an integral cycle. In the magnetic description we define a twisted operator of quantum-corrected weight $\lambda = \Lambda + \rho_{\overline{0}}$ by the same irreducible model solution that we would use for an untwisted operator, but we also make an insertion in the functional integral. Namely, when we sum over different bundles, we add an extra minus sign if the $\mathrm{SO}(N)$-bundle, restricted to Σ, cannot be lifted to a $\mathrm{Spin}(N)$-bundle. In other words, we add a factor

$$(-1)^{\int_\Sigma w_2}\,, \tag{2.194}$$

where w_2 is the second Stiefel-Whitney class.[39]

There is no analog of this for an odd orthosymplectic group $\mathrm{OSp}(2m+1|2n)$. For example, for $m = n$, the magnetic dual is simply an $\mathrm{Sp}(2n)$ gauge theory with a fundamental hypermultiplet along the defect. The existence of this hypermultiplet means that the gauge bundle restricted to Σ must be an $\mathrm{Sp}(2n)$ bundle, not a bundle with structure group $\mathrm{PSp}(2n) = \mathrm{Sp}(2n)/\mathbb{Z}_2$. For $m \neq n$, the model solution has a Nahm pole valued in $\mathrm{Sp}(|2m-2n|)$, and this is incompatible with a twist defined using the center of $\mathrm{Sp}(2n)$.

[39]What we have described about the S-duality of twisted line operators is rather similar to the result of [72]: choosing a topological type of bundle on one side of the duality translates on the other side to choosing a fugacity in the sum over bundles.

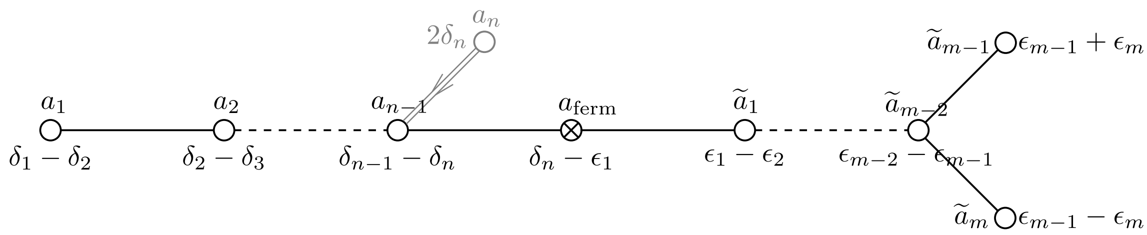

Figure 2.11: Dynkin diagram for the $\mathfrak{osp}(2m|2n)$ superalgebra, $m \geq 2$. The subscripts are expressions for the roots in terms of the orthogonal basis $\delta_\bullet$, $\epsilon_\bullet$. The superscripts represent the Dynkin labels of a weight. The middle root denoted by a cross is fermionic. Roots of the $\mathfrak{sp}(2n)$ and $\mathfrak{so}(2m)$ subalgebras are on the left and on the right of the fermionic root. The site shown in grey and labeled a_n is the long simple root of the $\mathfrak{sp}(2n)$ subalgebra, which does not belong to the set of simple roots of the superalgebra.

The magnetic duals of twisted and untwisted operators are nonetheless different, but that is because the model solutions used to define them are different, as explained in section 2.5.6.6.

2.5.6.2 More On The Orthosymplectic Lie Superalgebras

We also need to review some facts about the orthosymplectic Lie superalgebras. We start with the even orthosymplectic superalgebra $\mathrm{D}(m,n) \simeq \mathfrak{osp}(2m|2n)$. Here we assume that $m > 1$, since $m = 1$ corresponds to the type I superalgebra $\mathrm{C}(n) \simeq \mathfrak{osp}(2|2n)$ (the analysis of its line and surface operators is analogous to the $\mathfrak{u}(m|n)$ case, which we have discussed in section 2.5.5). We also assume that $n > 1$; the case $n = 1$ can be treated with minor modifications.

The Dynkin diagram for $\mathrm{D}(m,n)$ is shown on fig. 2.11. The positive bosonic and fermionic roots of $\mathfrak{osp}(2m|2n)$ are

$$\begin{aligned} &\Delta_{\bar{0}}^{+} = \left\{\delta_i \pm \delta_{i+p},\, 2\delta_i,\, \epsilon_j \pm \epsilon_{j+p}\right\}, \\ &\Delta_{\bar{1}}^{+} = \left\{\delta_i \pm \epsilon_j\right\}, \quad i = 1\ldots n,\, j = 1\ldots m,\, p > 0\,, \end{aligned} \tag{2.195}$$

where the mutually orthogonal basis vectors are normalized as

$$\langle \delta_i, \delta_i \rangle = \frac{1}{2}\,, \qquad \langle \epsilon_i, \epsilon_i \rangle = -\frac{1}{2}, \tag{2.196}$$

to ensure that the longest root has length squared 2. The bosonic and fermionic Weyl vectors are

$$\rho_{\bar{0}} = \sum_{i=1}^{n} (n+1-i)\,\delta_i + \sum_{j=1}^{m} (m-j)\,\epsilon_j\,, \quad \rho_{\bar{1}} = m \sum_{i=1}^{n} \delta_i\,, \tag{2.197}$$

and the superalgebra Weyl vector is $\rho = \rho_{\bar{0}} - \rho_{\bar{1}}$.

A weight with Dynkin labels[40] $a_\bullet$, $\widetilde{a}_\bullet$ is decomposed in terms of the basis vectors as

$$\begin{aligned} \Lambda &= a_1\delta_1 + \cdots + a_n(\delta_1 + \cdots + \delta_n) + \widetilde{a}_1\epsilon_1 + \cdots + \widetilde{a}_{m-2}(\epsilon_1 + \cdots + \epsilon_{m-2}) \\ &+ \frac{1}{2}\widetilde{a}_{m-1}(\epsilon_1 + \cdots + \epsilon_{m-1} + \epsilon_m) + \frac{1}{2}\widetilde{a}_m(\epsilon_1 + \cdots + \epsilon_{m-1} - \epsilon_m)\,. \end{aligned} \tag{2.198}$$

It is a dominant weight of a finite-dimensional representation, if the Dynkin labels are non-negative integers, and also satisfy the following supplementary condition: if $a_n \leq m-2$, then no more than the first a_n of the labels $\widetilde{a}_\bullet$ can be non-zero; if $a_n = m-1$, then $\widetilde{a}_{m-1} = \widetilde{a}_{m-2}$; if $a_n \geq m$, there is no constraint. We will call a weight (and the corresponding representation) spinorial if the number $\widetilde{a}_{m-1} + \widetilde{a}_m$ is odd. Clearly, a spinorial dominant weight must have $a_n \geq m$. Also, such a weight is always typical.

Now let us turn to the odd orthosymplectic superalgebra $\mathrm{B}(m,n) \simeq \mathfrak{osp}(2m+1|2n)$. The distinguished Dynkin diagram and the simple roots for $\mathfrak{osp}(2m+1|2n)$ and for its bosonic subalgebra $\mathfrak{so}(2m+1)\times\mathfrak{sp}(2n)$ can be found in fig. 2.3 of section 2.3.1. The positive bosonic and fermionic roots of this superalgebra are

$$\begin{aligned} \Delta^+_{\bar{0}} &= \left\{\delta_i - \delta_{i+p},\, \delta_i + \delta_{i+p},\, 2\delta_i,\, \epsilon_j - \epsilon_{j+p},\, \epsilon_j + \epsilon_{j+p},\, \epsilon_j\right\}\,, \\ \Delta^+_{\bar{1}} &= \left\{\delta_i - \epsilon_j,\, \delta_i + \epsilon_j,\, \delta_i\right\}\,, \quad i = 1\ldots n,\, j = 1\ldots m,\, p > 0\,, \end{aligned} \tag{2.199}$$

where the mutually orthogonal basis vectors are normalized as in (2.196). The bosonic and

[40] The Dynkin label of a weight Λ for a simple bosonic root α is defined as usual as $a = 2\langle\Lambda,\alpha\rangle/\langle\alpha,\alpha\rangle$. However, the Dynkin labels used in (2.198) are for the simple roots of $\mathfrak{so}(2m)\times\mathfrak{sp}(2n)$, not for the superalgebra $\mathfrak{osp}(2m|2n)$. In practice, this means that a_n is the weight for the long root $2\delta_n$ of $\mathfrak{sp}(2n)$, and we do not use the label a_{ferm} associated to the fermionic root of the superalgebra.

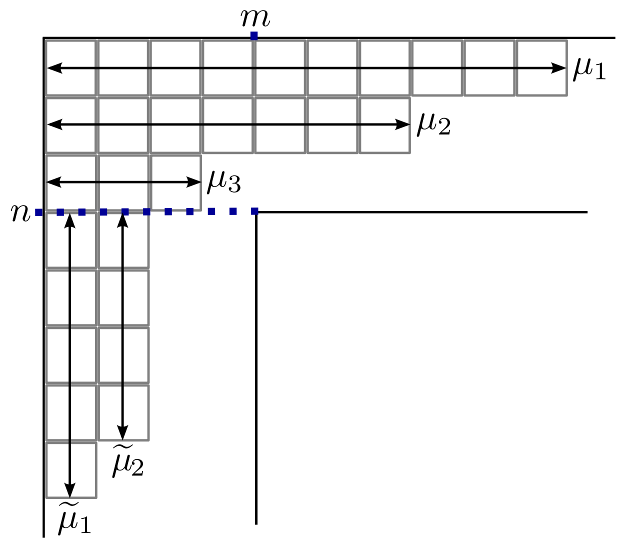

Figure 2.12: Example of a hook partition for $\mathfrak{osp}(9|6)$. The labels μ_i, $i = 1, \dots n$ and $\widetilde{\mu}_j$, $j = 1, \dots m$ were defined in (2.202). Here $\mu_3 = 3$, and, clearly, no more than the first three $\widetilde{\mu}$'s can be non-zero.

fermionic Weyl vectors are

$$\rho_{\overline{0}} = \sum_{i=1}^{n} (n + 1 - i)\, \delta_i + \sum_{j=1}^{m} \left(m + \frac{1}{2} - j \right) \epsilon_j \,, \quad \rho_{\overline{1}} = \left(m + \frac{1}{2} \right) \sum_{i=1}^{n} \delta_i \,, \tag{2.200}$$

and as usual the superalgebra Weyl vector is $\rho = \rho_{\overline{0}} - \rho_{\overline{1}}$.

If we parametrize a weight as

$$\Lambda = \sum_{i=1}^{n} \mu_i \delta_i + \sum_{i=1}^{m} \widetilde{\mu}_i \epsilon_i \,, \tag{2.201}$$

then, in terms of its Dynkin labels, one has

$$\mu_i = \sum_{j=i}^{n} a_j \,, \quad \widetilde{\mu}_i = \sum_{j=i}^{m-1} \widetilde{a}_j + \frac{1}{2} \widetilde{a}_m . \tag{2.202}$$

A weight Λ is a highest weight of a finite-dimensional representation of $\mathfrak{osp}(2m + 1|2n)$, if its Dynkin labels are non-negative integers, and no more than the first a_n of the $\mathfrak{so}(2m + 1)$ labels $(\widetilde{a}_1, \dots, \widetilde{a}_m)$ are non-zero. The last condition is trivial if $a_n \geq m$. We will call an irreducible representation "large" if $a_n \geq m$, and "small" in the opposite case. An irreducible representation is spinorial if the Dynkin label $\widetilde{a}_m$ is odd, and non-spinorial in the opposite case. Clearly, any spinorial representation is "large." It is also easy to see that all the "small" representations are atypical, and all the spinorial representations are typical.

Non-spin highest weights can be conveniently encoded in terms of hook partitions [40, 41, 42]. These are simply Young diagrams which are constrained to fit inside a hook with sides of width n and m, as shown in fig. 2.12 for $n = 3$ and $m = 4$. The figure shows how the labels $\mu_\bullet$ and $\widetilde{\mu}_\bullet$ parametrizing the weight are read from the diagram. This presentation implements automatically the constraint that only the first a_n of the $\mathfrak{so}(2m+1)$ Dynkin labels can be non-zero. In this notation, the "small" representations are those for which the Young diagram does not fill the upper left $n \times m$ rectangle.

Finally, let us note that for typical representations of any superalgebra there exist simple analogs of the Weyl formula to compute characters and supercharacters. For the character of a representation with highest weight Λ, the formula reads

$$\mathrm{ch}\,(R_\Lambda) = L^{-1} \sum_{w \in \mathcal{W}} (-1)^{\ell(w)} \exp\left(w(\Lambda + \rho)\right). \tag{2.203}$$

Here the sum goes over the elements of the Weyl group $\mathcal{W}$, which, by definition, is generated by reflections along the bosonic roots. The number $\ell(w)$ is the length of the reduced expression for the Weyl group element w. The Weyl denominator L is

$$L = \frac{\prod_{\alpha \in \Delta_{\bar{0}}^+} (\mathrm{e}^{\alpha/2} - \mathrm{e}^{-\alpha/2})}{\prod_{\alpha \in \Delta_{\bar{1}}^+} (\mathrm{e}^{\alpha/2} + \mathrm{e}^{-\alpha/2})}. \tag{2.204}$$

2.5.6.3 OSp$(2m|2n)$: The Mapping Of Line Operators

To understand the action of the $S^{-1}TS$ duality on the line operators of the D(m,n) Chern-Simons theory, we need to understand the action of the T-transformation on their magnetic duals. Since T is just a shift of the theta-angle, it does not change the model solution that is used to define the operator. Therefore one might conclude, as we did for the unitary superalgebra, that line operators are invariant under this transformation. As we now explain, this is indeed true for a subclass of line operators, but not for all of them.

In section 2.5.1 we have defined the instanton number $\mathfrak{N}^\vee$ for the orthogonal group. The action contained a term $i\theta_s\mathfrak{N}^\vee/2$, where θ_s is the string theory theta-angle. The 2π-periodicity of θ_s relied on the fact that $\mathfrak{N}^\vee$ takes values in $2\mathbb{Z}$. While this assertion is true on $\mathbb{R}^4$ or $\mathbb{R}\times S^3$, it is not always true on more general manifolds. We now want to show that it is not true even on simple manifolds like $\mathbb{R}^4$ in the presence of some line operators, and therefore such line operators transform non-trivially under the T-transformation.

Before explaining the details, let us state clearly the result. Consider a Wilson-'t Hooft operator (untwisted or twisted) in the electric theory, located along a knot K. We claim that in the presence of its S-dual, the instanton number $\mathfrak{N}^\vee$ of the magnetic theory takes values in $2\mathbb{Z}$, if the quantum-corrected weight λ of the operator is non-spin, and it takes values[41] in $\mathbb{Z}$, if this weight is spin. Therefore, T acts trivially on the non-spinorial line operators, but not on the spinorial ones. We will show that for spinorial weights the transformation T exchanges twisted and untwisted operators of a given quantum-corrected weight λ. In terms of the electric theory, we say that the knot invariants that are obtained from an untwisted spinorial operator in the theory with level $\mathcal{K}_{\mathfrak{osp}}$ are equal to the invariants obtained from a twisted spinorial operator in the theory with level $\mathcal{K}'_{\mathfrak{osp}}$, where $\mathcal{K}'_{\mathfrak{osp}}$ is given by (2.180). The mapping of non-spinorial line operators (whether untwisted or twisted) between the Chern-Simons theories with levels $\mathcal{K}_{\mathfrak{osp}}$ and $\mathcal{K}'_{\mathfrak{osp}}$ is trivial: the weight is unchanged and twisted or untwisted operators map to themselves.

Now let us prove our assertions about the instanton number. Assume for simplicity that the four-manifold M is 2-connected (that is, $\pi_1(M)=\pi_2(M)=0$). Our goal is to evaluate the instanton number $\mathfrak{N}^\vee$ for an $\mathrm{SO}(2m)\times\mathrm{SO}(2n+1)$ bundle on the knot complement $M\backslash K$ with a fixed trivialization along K, which is defined by a model solution of weight λ. For now

[41] As we have already explained in footnote 28, a more precise statement is that the instanton number takes values in $2\mathbb{Z}+c$ or $\mathbb{Z}+c$ for some constant c. Here we are interested only in the difference of instanton numbers for different bundles, so we will ignore the constant shift.

let us assume that $m \leq n$, so that the SO($2m$) subgroup of the gauge group is left unbroken by the three-dimensional defect at $y = 0$. Let Σ' be a two-sphere in M that encircles some point of the knot (this means that the linking number of Σ' with K is 1; for instance, Σ' can be the sphere $x^0 = 0$, $r' =$const in the language of fig. 2.6 of section 3.3.1), and Σ be a surface, bounded by the knot. Σ represents the non-trivial cycle in the relative homology $H_2(M, K)$.

We will focus on SO($2m$) bundles $\mathcal{V}$ on the knot complement, and ignore what happens in the SO($2(n-m)+1$)-part of the gauge group, which is broken everywhere at $y = 0$ by the boundary condition of the 3d defect. The reason we can do so is that all interesting things will come from different extensions of the SO($2m$) bundle from the knot neighborhood $K \times \Sigma'$ to the cycle Σ, while for the SO($2(n-m)+1$) subgroup this extension is uniquely fixed by the boundary condition. This is also the reason that there is no non-trivial analog of this story for the one-sided problem [6].

So far we have not been precise about the global form of the structure group of our bundle $\mathcal{V} \to M$. In the most general case, the structure group is the projective orthogonal group PSO($2m$) (the quotient of SO($2m$) by its center $\{\pm 1\}$), and this structure group might or might not lift to SO($2m$) or Spin($2m$). If it does lift to SO($2m$) or Spin($2m$), we say that $\mathcal{V}$ has a vector or a spin structure, respectively. To study obstructions to the existence of a vector or a spin structure (and more generally, obstructions related to $\pi_1(G)$ for G-bundles), it is enough to look at the restriction of the bundle to the two-skeleton of the manifold. Let Σ_0 be a two-manifold with G-bundle $\mathcal{V} \to \Sigma_0$; we assume that G is a connected group, and that Σ_0 is closed or that $\mathcal{V}$ is trivialized on its boundary. Such a $\mathcal{V} \to \Sigma_0$ is classified topologically by a characteristic class x valued in $H^2(\Sigma_0, \pi_1(G))$. Concretely, x is captured by an element of $\pi_1(G)$ that is used as a gluing function to construct the bundle $\mathcal{V} \to \Sigma_0$. Thus, x associates to Σ_0 an element $\widehat{x}$ of the center of the universal cover $\widehat{G}$ of G. A bundle $\mathcal{V}_R$ associated to $\mathcal{V}$ in a representation R exists if and only if $\widehat{x}$ acts trivially on R.

In our application, Σ_0 is either Σ or Σ', and $G = \mathrm{PSO}(2m)$. We note that the surface Σ

can be deformed to lie entirely in the region $y > 0$, where the gauge group is $\mathrm{SO}(2n+1)$. Since $\mathrm{SO}(2m)$ and not $\mathrm{PSO}(2m)$ is a subgroup of $\mathrm{SO}(2n+1)$, the restriction of $\mathcal{V}$ to Σ always has vector structure.

Let λ be a non-spinorial weight of the gauge group of the electric theory. This means that λ belongs to the character lattice of $\mathrm{SO}(2m) \times \mathrm{Sp}(2n)$, and therefore the parameter of the S-dual magnetic operator belongs to the cocharacter lattice of the dual group, which is $\mathrm{SO}(2m) \times \mathrm{SO}(2n+1)$. Therefore, the model solution for the line operator defines on Σ' a bundle with vector structure. Together with the facts that we explained a few lines above, this means that $\mathcal{V}$ has vector structure, *i.e.* it is an $\mathrm{SO}(2m)$ bundle. For its instanton number we can use the formula

$$\mathfrak{N}^\vee = \int_M w_2 \wedge w_2 \bmod 2\,, \tag{2.205}$$

where w_2 is the second Stiefel-Whitney class, or more precisely an arbitrary lift of it to the integral cohomology. (For a derivation of this formula, see *e.g.* [73].) On our manifold we can rewrite[42] this as

$$\mathfrak{N}^\vee = 2\left(\int_\Sigma w_2\right)\left(\int_{\Sigma'} w_2\right) \bmod 2\,, \tag{2.206}$$

which means that whatever w_2 is, the instanton number is even. Therefore, a shift of the theta-angle by 2π in presence of a non-spinorial line operator is still a symmetry, and such operators are mapped trivially under the T-transformation.

Now let the weight λ be spinorial. Then it belongs to the character lattice of $\mathrm{Spin}(2m) \times \mathrm{Sp}(2n)$ (and not to its sublattice corresponding to $\mathrm{SO}(2m) \times \mathrm{Sp}(2n)$), and therefore the parameter of the dual magnetic operator belongs to the cocharacter lattice of $\mathrm{PSO}(2m) \times \mathrm{SO}(2n+1)$ (and not to the cocharacter lattice of $\mathrm{SO}(2m) \times \mathrm{SO}(2n+1)$). The bundle that is defined on Σ' by such a model solution is a $\mathrm{PSO}(2m)$ bundle with no vector structure. What

[42]For a quick explanation, think of w_2 in this geometry as a sum $a+b$, where a is possibly non-trivial on Σ but trivial on Σ', and b is trivial on Σ but possibly non-trivial on Σ'. Then $w_2^2 = 2ab = 0 \bmod 2$, accounting for the factor of 2 in eqn. (2.206).

we then expect to get is roughly speaking that the factor $\int_{\Sigma'} w_2$ in (2.206) now becomes $1/2$, which would give us $\mathfrak{N}^\vee = \int_\Sigma w_2 \bmod 2$ for the instanton number. Let us prove this in a more rigorous way.

For that we adapt arguments used in [73], where more detail can be found. The topology of two PSO-bundles that coincide on the two-skeleton can differ only by the embedding of some number of bulk instantons. Therefore the instanton numbers of such bundles can only differ by an even integer. To find $\mathfrak{N}^\vee \bmod 2$, it is enough to study any convenient bundle with a given behavior on Σ and Σ'. Consider first the case of the group $\mathrm{PSO}(6) = \mathrm{SU}(4)/\mathbb{Z}_4$. Its fundamental group is $\mathbb{Z}_4$. Let x be the $\mathbb{Z}_4$-valued characteristic class which defines the topology of the restriction of the bundle to the two-skeleton (i.e., to Σ and Σ'). Let $\mathcal{L}$ be a line bundle with first Chern class $c_1 = x \bmod 4$. Let $\mathcal{O}$ be the trivial line bundle, and consider the bundle

$$\mathcal{V}_4 = \mathcal{L}^{1/4} \otimes (\mathcal{L}^{-1} \oplus \mathcal{O} \oplus \mathcal{O} \oplus \mathcal{O})\,. \tag{2.207}$$

It does not exist as an SU(4) bundle, unless $x = 0$, but its associated adjoint bundle $3\mathcal{L} \oplus 3\mathcal{L}^{-1} \oplus 9\mathcal{O}$ does exist; this bundle has structure group PSO(6). The associated bundle in the vector representation of SO(6) is the antisymmetric part of $\mathcal{V}_4 \otimes \mathcal{V}_4$; it exists precisely when $x = 0 \bmod 2$, since it contains $\mathcal{L}^{1/2}$. Though $\mathcal{V}_4$ might not exist, we can use the standard formulas to compute its Chern number

$$\int_M c_2(\mathcal{V}_4) = -\frac{3}{4} \int_\Sigma c_1(\mathcal{L}) \int_{\Sigma'} c_1(\mathcal{L}) = \frac{1}{4} \int_\Sigma x \int_{\Sigma'} x \bmod 1\,. \tag{2.208}$$

This Chern number is the instanton number normalized to be $\mathbb{Z}$-valued for an SU(4) bundle, so it is $\mathfrak{N}^\vee/2$. Note that, since the bundle on Σ' has no vector structure, we have $\int_{\Sigma'} x = \pm 1$. On the contrary, on Σ there is vector structure, and we can write $\int_\Sigma x = 2\int_\Sigma w_2 \bmod 4$. We finally get

$$\mathfrak{N}^\vee = \int_\Sigma w_2 \bmod 2\,. \tag{2.209}$$

Comparing to the definition of the magnetic duals of the twisted operators in section 2.5.6.1, we conclude that the T-transformation, besides shifting the theta-angle by 2π, ialso interchanges the twisted and untwisted spinorial line operators. One can easily extend these arguments to the even orthogonal groups other than SO(6). The relevant facts are explained in [73] in a similar context, and will not be repeated here.

In our discussion, we have assumed that the ranks of the two gauge groups satisfy $m \leq n$. One can extend the arguments to the case $n > m$ with some technical modifications. Rather than explaining this, we will now give an alternative argument, which uses the language of surface operators, and does not depend on the rank difference $n - m$.

2.5.6.4 OSp($2m|2n$): The Mapping Of Surface Operators

Our discussion will be analogous to what we have said about the case of the unitary superalgebra in section 2.5.5. The $S^{-1}TS$ duality transformation acts on the half-BPS surface operators in the following way,

$$\begin{pmatrix} \alpha \\ \eta \end{pmatrix} \xrightarrow{S} \begin{pmatrix} \eta \\ -\alpha \end{pmatrix} \xrightarrow{T} \begin{pmatrix} \eta \\ \eta^{*so} - \alpha \end{pmatrix} \xrightarrow{S^{-1}} \begin{pmatrix} \alpha - \eta^{*so} \\ \eta \end{pmatrix}. \tag{2.210}$$

Here the T-transformation acts in the magnetic description of the theory. Therefore, its definition involves taking the dual of η with respect to the canonically-normalized Killing form of the orthogonal Lie group, which is the gauge group in the magnetic description. To emphasize this fact, we have denoted this dual by η^{*so}.

Recall that the action in the electric theory was defined using the canonically-normalized Killing form of the superalgebra, whose bosonic part, according to (2.153), is $\kappa_{\mathfrak{osp}} = \kappa_{\mathfrak{sp}} - 2\kappa_{\mathfrak{so}}$, where $\kappa_{\mathfrak{so}}$ and $\kappa_{\mathfrak{sp}}$ are the canonically-normalized Killing forms for the corresponding bosonic Lie algebras. Let us consider the positive-definite form $\kappa_{\mathfrak{sp}} + 2\kappa_{\mathfrak{so}}$, and denote the dual with respect to this form by a star. (In fact, this notation has already been defined in footnote

14.) The equation (2.210) in this notation is equivalent to

$$\begin{pmatrix} \alpha \\ \\ \eta \end{pmatrix} \xrightarrow{S^{-1}TS} \begin{pmatrix} \alpha - 2\eta^* \\ \\ \eta \end{pmatrix}. \tag{2.211}$$

For the $\mathfrak{so}(2m)$ part of the parameters, the factor of two in this formula simply follows from the analogous factor in front of $\kappa_{\mathfrak{so}}$ in $\kappa_{\mathfrak{osp}}$. For the $\mathfrak{sp}(2n)$ part of the parameters, one needs to compare the canonically-normalized Killing forms of $\mathfrak{sp}(2n)$ and $\mathfrak{so}(2n+1)$ on $\mathfrak{t}^*_{\mathfrak{sp}} \simeq \mathfrak{t}_{\mathfrak{so}}$. The S-duality maps the root lattice in $\mathfrak{t}^*_{\mathfrak{sp}}$ to the coroot lattice in $\mathfrak{t}_{\mathfrak{so}}$. Comparing these lattices, one finds that in $\mathfrak{t}^*_{\mathfrak{sp}} \simeq \mathfrak{t}_{\mathfrak{so}}$ the S-duality identifies δ_i with ϵ_i, in the notations of section 2.5.6.2. The canonically-normalized forms for $\mathfrak{sp}(2n)$ and $\mathfrak{so}(2n+1)$ give respectively[43] $\langle\delta_i, \delta_j\rangle_{\mathfrak{sp}} = \delta_{ij}/2$ and $\langle\epsilon_i, \epsilon_j\rangle_{\mathfrak{so}} = \delta_{ij}$, and their ratio gives the factor of two in (2.211).

The equation (2.85), which defines the relation between the weight and the parameters of a surface operator in the electric theory, continues to hold for the orthosymplectic Chern-Simons theory, if one replaces the level $\mathcal{K}$ in that equation by $\mathcal{K}_{\mathfrak{osp}}$. Using this, and also the transformation laws (2.180) and (2.211), we conclude that the $S^{-1}TS$ duality transforms the weights according to

$$\frac{\lambda'}{\mathcal{K}'_{\mathfrak{osp}}} = \frac{\lambda}{\mathcal{K}_{\mathfrak{osp}}}. \tag{2.212}$$

Again, the procedure of section 2.4.6 for computing knot invariants using surface operators is obviously invariant under this transformation.

Let us compare the surface operator and the line operator approaches in the case that the weight λ is integral. The equation (2.212) can alternatively be written as

$$\lambda' = \lambda + 2\mathcal{K}'_{\mathfrak{osp}}\lambda. \tag{2.213}$$

First let us look at the part λ_r of the weight, which corresponds to the symplectic Lie

[43] Note that the canonical normalization of the Killing form for $\mathfrak{so}(2n+1)$ is different from the superalgebra normalization (2.196).

subalgebra. In the action (2.153), the level $\mathcal{K}_{\mathfrak{osp}}$ multiplies the Chern-Simons term for the $\mathfrak{sp}(2n)$ subalgebra, which is defined using the canonically-normalized $\mathfrak{sp}(2n)$ Killing form. Therefore the knot invariants computed using the surface operators are unchanged when the weight λ_r is shifted by $\mathcal{K}_{\mathfrak{osp}}$ times an integral coroot of the $\mathfrak{sp}(2n)$ subalgebra. If λ_r is an integral weight, then $2\lambda_r$ is an integral coroot, and therefore the difference between λ_r' and λ_r in (2.213) is inessential for computing the knot invariants.

For the part λ_ℓ of the weight, which corresponds to the orthogonal subalgebra, the situation is more complicated. The canonically-normalized Chern-Simons term for the orthogonal subalgebra in the action (2.153) is multiplied by $2\mathcal{K}_{\mathfrak{osp}}$. For this reason, the knot invariants computed using the surface operators are invariant under the shift of λ_ℓ by $2\mathcal{K}_{\mathfrak{osp}}$ times an integral coroot of the $\mathfrak{so}(2m)$ subalgebra. Therefore, the shift of λ_ℓ in the equation (2.213) is trivial from the point of view of the knot observables if and only if the integral weight λ_ℓ is actually a coroot. What if it is not? Since the $\mathfrak{so}(2m)$ Lie algebra is simply-laced, any integral weight is also an element of the dual root lattice Γ_r^*. Therefore the group element $\exp(2\pi\lambda_\ell)$ actually belongs to the center of the orthogonal group. Let us make a singular gauge transformation in the electric theory around the surface operator on the left side of the three-dimensional defect, using the group element $\exp(\theta\lambda_\ell)$, where θ is the azimuthal angle in the plane normal to the surface operator. This transformation maps a surface operator corresponding to the weight λ_ℓ' back to a surface operator with weight λ_ℓ. Since our gauge transformation is closed only up to the central element $\exp(2\pi\lambda_\ell)$, it also introduces a twist of the boundary hypermultiplets by this group element. In the fundamental representation of $\mathrm{SO}(2m)$, to which the hypermultiplets belong, the element $\exp(2\pi\lambda_\ell)$ acts trivially if the weight λ is non-spinorial, and it acts by -1 if λ is spinorial. We have reproduced the result that was derived in the previous section in the language of line operators: $S^{-1}TS$ acts trivially on Chern-Simons line observables labeled by non-spinorial representations, but exchanges the twisted and the untwisted operators for a spinorial weight.

2.5.6.5 $\mathrm{OSp}(2m|2n)$: Comparing The Representations

We would like to look closer at the mapping of spinorial line operators. Consider a line operator, labeled by a supergroup representation of spinorial highest weight $\Lambda = \lambda - \rho$, and an $S^{-1}TS$-dual twisted operator, which is labeled by a representation of the bosonic subgroup with highest weight $\Lambda' = \lambda - \rho_{\overline{0}}$. Note that the Weyl vectors ρ and $\rho_{\overline{0}}$, which can be found from (2.197), are non-spinorial integral weights, and therefore the property of being spinorial/non-spinorial is the same for the weights and for the quantum-corrected weights of $\mathrm{OSp}(2m|2n)$.

We would like to see more explicitly how the duality mapping acts in terms of representations. We have $\lambda = \Lambda + \rho = \Lambda' + \rho_{\overline{0}}$, or equivalently, $\Lambda' = \Lambda - \rho_{\overline{1}}$. Using the formulas (2.197) and (2.198), this can be translated into a mapping of Dynkin labels,

$$\begin{aligned} \widetilde{a}'_j &= \widetilde{a}_j, \quad j = 1, \ldots, m\,, \\ a'_i &= a_i, \quad i = 1, \ldots, n-1\,, \\ a'_n &= a_n - m\,. \end{aligned} \tag{2.214}$$

As was noted in section 2.5.6.2, for a spinorial superalgebra representation one has[44] $a_n \geq m$. Therefore, the mapping of Dynkin labels written above is a one-to-one correspondence between the irreducible spinorial representations of the $D(m,n)$ superalgebra and its bosonic subalgebra.

We can make an additional test of the duality by comparing the local framing anomalies of the line operators. Recall that the knot polynomials in Chern-Simons theory are invariants of framed knots. If the framing of a knot is shifted by one unit via a 2π twist, the knot

[44] As we have mentioned in a similar context in section 2.3.5, we do not really know why the supplementary condition should be imposed in the present discussion, since it is not a general condition on 1/2-BPS line operators. Nonetheless, imposing this condition works nicely, as we have just seen. This shows once again that our understanding of line operators in the theory is incomplete. We will find something similar for odd OSp supergroups.

polynomial is multiplied by a factor

$$\exp(2\pi i \Delta_O)\,, \tag{2.215}$$

where Δ_O is the dimension of the conformal primary O that corresponds in the WZW model[45] to the given Wilson line. For a Wilson line in representation R, this framing factor is

$$\exp\left(i\pi\,\frac{c_2(R)}{k+h\ \mathrm{sign}(k)}\right) = q^{c_2(R)}\,, \tag{2.216}$$

where $c_2(R) = \langle\lambda,\lambda\rangle - \langle\rho,\rho\rangle$ is the value of the quadratic Casimir in the representation R. The variable q was defined for the $\mathrm{D}(m,n)$ superalgebra in (2.158). In the bosonic, one-sided case these formulas have been derived in [6] from the magnetic description of the theory. It would be desirable to give such a derivation for the two-sided case, but we will not attempt to do it here.

To compare the framing factors for our dual operators, we need to derive a formula for the framing anomaly of a twisted operator. The energy-momentum tensor of the conformal field theory is given by the Sugawara construction

$$T(z) = \frac{\widehat{\kappa}_{nm}\,{:}J^m(z)J^n(z){:}}{2(k+h)}\,, \tag{2.217}$$

where $\widehat{\kappa} = \kappa \oplus \omega$ is the superinvariant bilinear form[46] on the superalgebra, and $J^m(z)$ is the holomorphic current with the usual OPE

$$J^m(z)J^n(w) \sim \frac{k\,\widehat{\kappa}^{mn}}{(z-w)^2} + \frac{f_p^{mn}J^p(w)}{z-w}. \tag{2.218}$$

[45] As it is explained in Appendix E of [1], there actually is not a straightforward relation between 3d Chern-Simons theory and 2d current algebra in the case of a supergroup. Nonetheless, some results work nicely and the one we are stating here seems to be one.

[46] Here we slightly depart from our usual notation, and use indices $m, n, \ldots$ both for bosonic and fermionic generators of the superalgebra. Also, note that the inverse tensor is defined by $\widehat{\kappa}^{mn}\widehat{\kappa}_{pn} = \delta_p^m$, hence the unusual order of indices in the Sugawara formula.

One can easily verify that for a simple superalgebra the formula (2.217) gives the energy-momentum tensor with a correct OPE.

Normally, the current $J^m(z)$ is expanded in integer modes. The eigenvalue of the Virasoro generator L_0, acting on a primary field, is determined by the action of the zero-modes of the current, which give the quadratic Casimir, as stated in eqn. (2.216). However, for a primary field corresponding to a twisted operator in Chern-Simons, one naturally expects the fermionic components of the current $J^m(z)$ to be antiperiodic. In that case, the bosonic part of the current gives the usual contribution to the conformal dimension, which for a weight Λ is proportional to the bosonic quadratic Casimir $\langle \Lambda + 2\rho_{\bar{0}}, \Lambda \rangle$. The fermionic part of the current in the twisted sector has no zero-modes, and its contribution to the L_0 eigenvalue is just a normal-ordering constant, independent of the weight Λ. One can evaluate this constant from (2.217), (2.218), and get for the dimension of the operator

$$\Delta_O^{tw} = \frac{\langle \Lambda + 2\rho_{\bar{0}}, \Lambda \rangle - k \dim(\mathfrak{g}_{\bar{1}})/8}{2(k+h)}. \tag{2.219}$$

Using the identity $\langle \rho_{\bar{0}}, \rho_{\bar{0}} \rangle = \langle \rho, \rho \rangle + h \dim(\mathfrak{g}_{\bar{1}})/8$, which actually is valid for any of our superalgebras, one obtains an expression for the framing factor

$$\exp\left(i\pi \frac{\langle \lambda, \lambda \rangle - \langle \rho, \rho \rangle}{k + h \ \mathrm{sign}(k)} \right) \exp\left(-i\pi \dim(\mathfrak{g}_{\bar{1}}) \mathrm{sign}(k)/8\right). \tag{2.220}$$

Here we have restored the dependence on the sign of the level k, and used our definition of λ for the twisted operators. The second factor in this formula does not map correctly under the duality, but that is what one could have expected, since this factor is non-analytic in $\mathcal{K} = k + h \,\mathrm{sign}(k)$ (compare to the discussion of the global framing anomalies in sections 2.4.3.1 and 2.5.4). The first factor is analytic in $\mathcal{K}$, and it is clear from comparison to eq. (2.216) that it does map correctly under the duality.

2.5.6.6 Duality For The Odd Orthosymplectic Superalgebra

Let us turn to the case of the odd orthosymplectic superalgebra. As was already noted in section 2.3.2.2, the definition of line operators in this theory has some peculiarities. As follows from the equation (2.200), for $\mathrm{B}(m,n)$ the bosonic Weyl vector $\rho_{\overline{0}}$ is an integral spinorial weight, while the superalgebra Weyl vector ρ is not an integral weight: it has a half-integral Dynkin label with respect to the short coroot of the $\mathfrak{sp}(2n)$ subalgebra. This means that the quantum-corrected weight $\lambda = \Lambda + \rho$ for an untwisted operator is not an integral weight, and therefore a Wilson-'t Hooft operator, as defined in section (2.3.3.2), is not gauge-invariant classically. The resolution of this puzzle should come from another peculiarity of the $\mathrm{B}(m,n)$ Chern-Simons theory. The definition of the path-integral of this theory includes an η-invariant (2.161), which comes from the one-loop determinant (or rather the Pfaffian) of the hypermultiplet fermions. In the presence of a monodromy operator, one should carefully define this fermionic determinant, and we expect an anomaly that will cancel the problem that exists at the classical level. We will not attempt to explain the details of this in the present chapter.

Unlike the case $\mathrm{OSp}(2m|2n)$, a magnetic line operator of $\mathrm{OSp}(2m+1|2n)$ is completely determined[47] by its weight λ, as explained at the end of section 2.5.6.1. However, the quantum-corrected weights for twisted and untwisted operators belong to different lattices, due to the different properties of ρ and $\rho_{\overline{0}}$, mentioned above. So the magnetic duals of twisted and untwisted electric line operators are simply described by different model solutions. Since the T-transformation preserves the model solution, the $S^{-1}TS$ duality should preserve the quantum-corrected weight.

We need to introduce some further notation. In the orientifold construction, we took the Killing form to be positive on the $\mathfrak{sp}$ part of $\mathrm{B}(m,n)$. In the dual theory, it will be positive on the $\mathfrak{so}$ part, and for this reason we denote the superalgebra of the dual theory by $\mathrm{B}'(n,m)$.

[47] Here we ignore the issues related to the atypical representations. We will say a little more on this later in this section.

The basis vectors in the dual $\mathfrak{t}^{*\prime}$ of the Cartan subalgebra of $\mathrm{B}'(n,m)$ will be denoted by δ'_j, $j = 1,\ldots,m$, and ϵ'_i, $i = 1,\ldots,n$, and their scalar products are defined to have opposite sign relative to (2.196). The Dynkin labels for the representations of $\mathrm{B}'(n,m)$ will be denoted as a'_j, $j = 1,\ldots,m$, and $\widetilde{a}'_i$, $i = 1,\ldots,n$. To make precise sense of the statement that the $S^{-1}TS$ duality preserves the quantum-corrected weight, it is necessary to specify how one identifies $\mathfrak{t}^*$ and $\mathfrak{t}^{*\prime}$. We use the mapping which identifies ϵ'_i with δ_i and δ'_j with ϵ_j. This linear map preserves the scalar product. In principle, one could derive this identification from the S-duality transformations of surface operators, but we will simply take it as a conjecture and show that it passes some non-trivial tests.

We can make one such test before we go into the details of the operator mapping. According to the equations (2.216), (2.220) and the definition (2.163) of the variable q, the framing anomaly factor in the $\mathrm{B}(m,n)$ theory for an operator of quantum-corrected weight λ is equal to q^{2c_2}, where $c_2 = \langle\lambda,\lambda\rangle - \langle\rho,\rho\rangle$. (This formula is true for both twisted and untwisted operators, modulo non-analytic terms.) From this we can see that our map does preserve the framing anomaly.[48] Indeed, it preserves λ and the scalar product, and although the Weyl vectors ρ and ρ' for the two dual superalgebras $\mathrm{B}(m,n)$ and $\mathrm{B}'(n,m)$ are different, their lengths happen to coincide, as one can verify from the explicit formula (2.200).

In the rest of this section we will examine the mapping

$$\lambda = \lambda' \tag{2.221}$$

in more detail. We will see that it gives a correspondence between the untwisted non-spinorial operators of the two theories, maps the twisted non-spinorial operators to the untwisted spinorial operators, and finally indentifies the twisted spinorial operators of one theory with the twisted spinorial operators of the other one. To put it shortly, it exchanges the spin and the twist. It is important to note that one might need to refine the mapping

[48]To be precise, there is actually a little mismatch for the spinorial operators. In that case the quadratic Casimir c_2 can be non-integral, and therefore there is a difference by a root of unity due to the fact that q is mapped to $-q$. Hopefully, this discrepancy can be cured in a more accurate treatment.

(2.221) for atypical weights. We will indeed encounter an ambiguity in interpreting (2.221) for the "small" atypical weights.

First let us focus on the non-spinorial untwisted line operators, for which the duality should give a correspondence between the non-spinorial representations of the two superalgebras. The map (2.221) of the dominant weights is already known in the literature for the special case of $m = 0$. In fact, a remarkable correspondence between finite-dimensional representations of $\mathfrak{osp}(1|2n)$ and non-spinorial finite-dimensional representations of $\mathfrak{so}(2n+1)$ was established in [71]. It preserves the full set of Casimirs, including the quadratic one. For $n = 1$, the map is so elementary that one can describe it by hand. This will make our later discussion more concrete. The spin s representation of $\mathfrak{so}(3)$, for non-negative integer s, is mapped to the trivial representation of $\mathfrak{osp}(1|2)$ for $s = 0$, and otherwise to the representation of $\mathfrak{osp}(1|2)$ that is a direct sum of bosonic states of spin $s/2$ (under $\mathfrak{sp}(2) \cong \mathfrak{su}(2)$) and fermionic states of spin $(s-1)/2$. Note that if we ignore the statistics of the states, the given $\mathfrak{so}(3)$ and $\mathfrak{osp}(1|2)$ representations both have dimension $2s+1$. This is a special case of a correspondence between characters found in [71].

An equivalent explanation is that a representation of $\mathfrak{so}(3)$ whose highest weight is s is mapped, if s is an integer, to a representation of $\mathfrak{osp}(1|2)$ whose highest weight is s times the smallest strictly positive weight of this algebra. The spinorial representations of $\mathfrak{so}(3)$ – the representations with half-integral s – do not participate in this correspondence, since there is no representation of $\mathfrak{osp}(1|2)$ whose highest weight is a half-integral multiple of the smallest positive weight. The spinorial representations of $\mathfrak{so}(3)$ have a dual in terms of twisted line operators, but not in terms of representations.

This correspondence between $\mathfrak{so}(s)$ and $\mathfrak{osp}(1|2)$ maps tensor products of $\mathfrak{so}(3)$ representations to tensor products of $\mathfrak{osp}(1|2)$ representations if one ignores whether the highest weight of an $\mathfrak{osp}(1|2)$ representation is bosonic or fermionic. To illustrate this correspondence, let $\mathbf{s}$ denote an irreducible $\mathfrak{so}(3)$ representation of spin s. Let $\mathbf{s}'$ and $\widetilde{\mathbf{s}}'$ denote irreducible $\mathfrak{osp}(1|2)$ representations whose highest weight is s times the smallest positive weight, with the highest weight vector being bosonic or fermionic, respectively. Then one has, for example,

$$\begin{cases} \mathbf{1}\otimes\mathbf{1}\cong\mathbf{2}\oplus\mathbf{1}\oplus\mathbf{0} & \text{for } \mathfrak{so}(3) \\ \mathbf{1}'\otimes\mathbf{1}'\cong\mathbf{2}'\oplus\widetilde{\mathbf{1}}'\oplus\mathbf{0}' & \text{for } \mathfrak{osp}(1|2). \end{cases} \tag{2.222}$$

There is an obvious matching, if we ignore the reversed statistics of $\widetilde{\mathbf{1}}'$ on the $\mathfrak{osp}(1|2)$ side. We interpret this matching to reflect the fact that the duality between $\mathfrak{so}(3)$ and $\mathfrak{osp}(1|2)$ preserves the operator production expansion for Wilson line operators. (In Chern-Simons theory, for generic q the OPE of line operators is given by the classical tensor product, so we can compare such OPE's by comparing classical tensor products.) However, we do not know the interpretation of the reversed statistics of $\widetilde{\mathbf{1}}'$. Perhaps it somehow involves the fact that the quantum duality changes the sign of q. In [71], it is shown that an analogous matching of tensor products holds in general.

Additional relevant results are in [16]. Let $\mathcal{U}_q(\mathfrak{osp}(1|2n))$ and $\mathcal{U}_{q'}(\mathfrak{so}(2n+1))$ be the quantum deformations of the universal enveloping algebras of the corresponding Lie (super)algebras. It has been shown in [16] that there exists a natural map between the representations of these two quantum groups if one takes $q'=-q$, and restricts to non-spinorial representations of the latter. One would expect such a result from our duality, assuming that Chern-Simons theory of a supergroup is related to a corresponding quantum group in the manner that is familiar in the bosonic world.

Now we return to our mapping $\lambda=\lambda'$ (eqn. (2.221)), which extends the known results described above to general m and n. It has several nice properties. As follows from our discussion of the framing anomaly, it preserves the quadratic Casimir. From the Harish-Chandra isomorphism, it follows that, for non-spinorial weights, (2.221) gives a natural mapping not only of the quadratic Casimir, but of the higher Casimirs as well. It would be interesting to find an explanation of this directly from the quantum field theory. The map also preserves the atypicality conditions (2.50). Next, let us look at the Weyl character formula (2.203), assuming that the weights are typical. The Weyl groups for the two superalgebras are equivalent and act in the same way on $\mathfrak{t}^* \simeq \mathfrak{t}^{*\prime}$; therefore, with the mapping (2.221), the numerators of the character formula coincide for the dual representations. The denominators are also equal, as one can easily check, using the list of simple roots (2.199). However, the supercharacters are not mapped in any simple way. In particular, the duality preserves the dimensions of typical representations, but not the superdimensions.[49]

Let us actually see what the equation (2.221) says about the map of representations. Writing it as $\Lambda' = \Lambda + \rho - \rho'$ and using equations (2.200), (2.201), one gets that the labels $\mu_\bullet$ and $\widetilde{\mu}_\bullet$, defined in those equations, transform into $\mu'_j = \widetilde{\mu}_j + n$, $\widetilde{\mu}'_i = \mu_i - m$. According to the equation (2.202), this gives a mapping for the Dynkin labels,

$$\begin{aligned} \widetilde{a}'_i &= a_i, \quad i = 1, \ldots, n-1\,, \\ a'_j &= \widetilde{a}_j, \quad j = 1, \ldots, m-1\,, \\ \widetilde{a}'_n &= 2(a_n - m)\,, \\ a'_m &= \frac{1}{2}\widetilde{a}_m + n. \end{aligned} \tag{2.223}$$

If we restrict to "large" non-spin dominant weights ($a_n \geq m$), then this formula gives a one-to-one correspondence. The non-spin condition means that $\widetilde{a}_m$ is even, so that the mapping (2.223) is well-defined, and the "large" condition $a_n \geq m$ ensures that $\widetilde{a}'_n \geq 0$.

It is not immediately obvious what to say for "small" representations, since for them the dual Dynkin label $\widetilde{a}'_n$ comes out negative. Note that all the "small" representations are atypical, and in general we have less control over them by methods of this chapter. There can be different possible conjectures as to how to make sense of our map for them. First of all, we can still treat (2.221) as a correspondence between monodromy operators. Then to understand to which representation a given operator corresponds, we should make a Weyl transformation on λ', to bring it to a positive Weyl chamber. This is one possible way to understand the map (2.221) for the "small" representations. (For an atypical weight, there

[49] Of course, for $m, n \neq 0$ the superdimensions of typical representations on both sides of the duality are simply zero. But for m or n equal to 0, they are non-zero and do not agree.

can be several different ways to conjugate it to the positive Weyl chamber; these give different weights, though belonging to the same atypical block.)

There is another very elegant possibility. If we simply transpose the hook diagram for a $\mathrm{B}(m,n)$ weight, we will get some weight of $\mathrm{B}'(n,m)$. It is a curious observation that for the "large" representations, this operation reproduces our duality (2.223). Moreover, one can prove that even for the "small" representations this flip preserves the quadratic Casimir operator and therefore the framing anomaly, and can be a candidate for the generalization of our map to the "small" highest weights. Unfortunately, this is merely a possible guess.

In short, we have found a natural 1-1 mapping between non-spinorial representations of $\mathrm{OSp}(2m+1|2n)$ and $\mathrm{OSp}(2n+1|2m)$. Now let us turn to spinorial ones. The mapping (2.221) sends spinorial line operators to twisted operators. Here is a simple consistency check of this statement. In the electric theory, consider a Wilson-'t Hooft operator in a spinorial representation R that is supported on a knot K in a three-manifold W. If the class of K in $H_1(W;\mathbb{Z}_2)$ is nonzero, then the expectation of the operator vanishes because it is odd under a certain "large" gauge transformation that is single-valued in $\mathrm{SO}(2m+1)$ but not if lifted to $\mathrm{Spin}(2m+1)$. (The gauge transformations along a Wilson-'t Hooft operator are constrained to lie in the maximal torus, but there is no problem in choosing such an abelian "large" gauge transformation.) The dual of such a Wilson-'t Hooft operator under the $S^{-1}TS$ duality should have the same property. Indeed, a twisted operator, as described in section 2.3.3.3, does have this property (in this case because the definition of the twisted operator involves picking a $\mathbb{Z}_2$ bundle with monodromy around K).

Let Λ be a spinorial dominant weight of the $\mathrm{B}(m,n)$ superalgebra, and let Λ' be a non-spinorial weight of the bosonic algebra $\mathfrak{so}(2n+1)\times\mathfrak{sp}(2m)$ that we use in defining a twisted line operator. The mapping (2.221) would then be $\Lambda'+\rho'_{\overline{0}}=\Lambda+\rho$. The bosonic Weyl vector that is used here can be obtained from (2.200) by exchanging $\epsilon_\bullet$ with $\delta_\bullet$ and m with n. From this one finds that the coefficients in the expansion of the weights in the $\delta_\bullet$, $\epsilon_\bullet$ basis transform as $\widetilde{\mu}'_i=\mu_i-m$, $\mu'_j=\widetilde{\mu}_j-1/2$. Therefore, according to (2.202), the Dynkin labels of the weights are related as

$$\begin{aligned}
\widetilde{a}'_i &= a_i, \quad i = 1, \dots, n-1\,, \\
a'_j &= \widetilde{a}_j, \quad j = 1, \dots, m-1\,, \\
\widetilde{a}'_n &= 2(a_n - m)\,, \\
a'_m &= \frac{1}{2}(\widetilde{a}_m - 1).
\end{aligned} \tag{2.224}$$

This gives a one-to-one correspondence between the spinorial supergroup representations and the non-spinorial weights of the bosonic algebra $\mathfrak{so}(2n+1) \times \mathfrak{sp}(2m)$. In fact, for a spinorial representation of $\mathfrak{osp}(2m+1|2n)$, $\widetilde{a}_m$ is odd, ensuring that a'_m is an integer. On the other hand, $\widetilde{a}'_n$ is always even, so the twisted line operator with Dynkin labels $a'_i, \widetilde{a}'_j$ is always associated to a non-spinorial representation of the bosonic subalgebra of $\mathrm{OSp}(2n+1|2m)$. Moreover, the supplementary condition guarantees that $a_n - m$ is non-negative for a spinorial superalgebra representation.

The twisted operators for spinorial representations of the bosonic subgroup should be mapped into similar twisted spinorial operators. The mapping (2.221) reduces in this case to $\Lambda + \rho_{\overline{0}} = \Lambda' + \rho'_{\overline{0}}$. This gives $\widetilde{\mu}'_i = \mu_i + 1/2$, $\mu'_j = \widetilde{\mu}_j - 1/2$, or, in terms of the Dynkin labels,

$$\begin{aligned}
\widetilde{a}'_i &= a_i, \quad i = 1, \dots, n-1\,, \\
a'_j &= \widetilde{a}_j, \quad j = 1, \dots, m-1\,, \\
\widetilde{a}'_n &= 2a_n + 1\,, \\
a'_m &= \frac{1}{2}(\widetilde{a}_m - 1)\,,
\end{aligned} \tag{2.225}$$

which is indeed a one-to-one correspondence between the spinorial representations of the bosonic subgroups. In other words, the weights a'_i and $\widetilde{a}'_j$ are integers if the a_i and $\widetilde{a}_j$ are integers and $\widetilde{a}_m$ is odd, and moreover in that case $\widetilde{a}'_n$ is odd.

2.6 Appendix A: Conventions And Supersymmetry Transformations

We mostly follow the notation of [23, 6], with some minor differences. Euclidean signature is used, except in section 2.2.1 and the beginning of section 2.4. The Lorentz vector indices are denoted by Greek letters μ, $\nu, \ldots$ in four dimensions and by Latin i, j, k in three dimensions. The defect is at $x^3 = 0$, and x^3 is assumed to be the normal coordinate such that ∂_3 is the unit normal vector at the defect. The 3d spinor indices are denoted by α, β, ... When the indices are not shown explicitly, they are contracted as $v^\alpha w_\alpha$. They are raised and lowered with epsilon symbols,

$$\begin{aligned} &\epsilon^{12} = \epsilon_{12} = 1\,, \\ &v^\alpha = \epsilon^{\alpha\beta} v_\beta. \end{aligned} \tag{2.226}$$

Vector and spinor notation are related by sigma-matrices,

$$V_{\alpha\beta} = \sigma^i_{\alpha\beta} V_i = \begin{pmatrix} -iV_2 + V_3 & iV_1 \\ iV_1 & iV_2 + V_3 \end{pmatrix}. \tag{2.227}$$

With this definition, the product of the sigma-matrices is

$$\sigma^{i\alpha\beta}\sigma^j_{\beta\gamma} = \delta^{ij}\delta^\alpha_\gamma + \epsilon^{ijk}\sigma^\alpha_{k\gamma}. \tag{2.228}$$

The boundary conditions are invariant under 3d supersymmetry, with R-symmetry group $\mathrm{SU}(2)_X \times \mathrm{SU}(2)_Y$. The spinor indices for these two groups are denoted by $A, B, \ldots$ and $\dot{A}, \dot{B}, \ldots$, respectively, and the vector indices are denoted by a, b, c and $\dot{a}, \dot{b}, \dot{c}$. Conventions for the R-symmetry indices are the same as for the Lorentz indices. In particular, the R-symmetry sigma-matrices are as in 2.227.

Fields that take values in the adjoint representation are understood as anti-hermitian matrices.

The three-dimensional $\mathcal{N}=4$ supersymmetry acts on the fields in the following way:

$$\begin{aligned}
\delta A_i &= -\frac{1}{\sqrt{2}}\varepsilon^{\alpha}_{A\dot{B}}\left(\Psi_1^{A\dot{B}\beta}\sin\vartheta+\Psi_2^{A\dot{B}\beta}\cos\vartheta\right)\sigma_{i\alpha\beta}\,,\\
\delta A_3 &= -\frac{i}{\sqrt{2}}\varepsilon^{\alpha}_{A\dot{B}}\left(-\Psi^{A\dot{B}}_{1\alpha}\cos\vartheta+\Psi^{A\dot{B}}_{2\alpha}\sin\vartheta\right)\,,\\
\delta X^a &= -\frac{i}{\sqrt{2}}\varepsilon^{A}_{\dot{B}}\Psi_1^{B\dot{B}}\sigma^a_{AB}\,,\\
\delta Y^a &= \frac{i}{\sqrt{2}}\varepsilon^{\dot{A}}_{A}\Psi_2^{A\dot{B}}\sigma^a_{\dot{A}\dot{B}}\,,
\end{aligned}$$

$$\begin{aligned}
\sqrt{2}\delta\Psi^{A\dot{B}}_{1\alpha} &= \varepsilon^{\beta B\dot{B}}\left(-\not{D}_{\alpha\beta}X^A_B-\frac{i}{2}\epsilon_{\alpha\beta}\sin\vartheta[X^{AC},X_{BC}]\right)\\
&\quad -\varepsilon^{A\dot{A}}_{\alpha}\left(iD_3Y^{\dot{B}}_{\dot{A}}+\frac{i}{2}\sin\vartheta[Y_{\dot{A}\dot{C}},Y^{\dot{B}\dot{C}}]\right)\\
&\quad +i\cos\vartheta\varepsilon^{C\dot{C}}_{\alpha}[X^A_C,Y^{\dot{B}}_{\dot{C}}]+\varepsilon^{\beta A\dot{B}}\left(\frac{i}{2}\sin\vartheta\epsilon_{ijk}F^{ij}+\cos\vartheta F_{k3}\right)\sigma_{k\alpha\beta}\,,\\
\sqrt{2}\delta\Psi^{A\dot{B}}_{2\alpha} &= \varepsilon^{\beta A\dot{A}}\left(\not{D}_{\alpha\beta}Y^{\dot{B}}_{\dot{A}}-\frac{i}{2}\epsilon_{\alpha\beta}\cos\vartheta[Y^{\dot{B}\dot{C}},Y_{\dot{A}\dot{C}}]\right)\\
&\quad -\varepsilon^{B\dot{B}}_{\alpha}\left(iD_3X^A_B-\left[\frac{i}{2}\mu^A_B\delta(x^3)\right]-\frac{i}{2}\cos\vartheta[X^{AC},X_{BC}]\right)\\
&\quad -i\sin\vartheta\varepsilon^{C\dot{C}}_{\alpha}[X^A_C,Y^{\dot{B}}_{\dot{C}}]-\varepsilon^{\beta A\dot{B}}\left(-\frac{i}{2}\cos\vartheta\epsilon_{ijk}F^{ij}+\sin\vartheta F_{k3}\right)\sigma_{k\alpha\beta}\,,
\end{aligned}$$

$$\delta Q^I_{\dot A} = -\varepsilon^A_{\dot A}\lambda^I_A\,,$$
$$\delta\lambda^I_{\alpha A} = \varepsilon^{\beta\dot A}_A i \not{D}_{\alpha\beta} Q^I_{\dot A} - \varepsilon_{\alpha A\dot A}\omega^{IJ}\frac{\partial W_4}{\partial Q^J_{\dot A}} + \varepsilon^{B\dot B}_\alpha \sin\vartheta X^m_{AB} T^I_{mJ} Q^J_{\dot B}\,,$$
$$\delta Z^A = -\varepsilon^A_{\dot A}\zeta^{\dot A}\,,$$
$$\delta\overline{Z}^A = -\varepsilon^A_{\dot A}\overline{\zeta}^{\dot A}\,,$$
$$\delta\zeta^{\dot A} = \varepsilon^{\dot A}_A i\not{D} Z^A - \varepsilon_{B\dot B} Y^{\dot A\dot B} Z^B\,,$$
$$\delta\overline{\zeta}^{\dot A} = \varepsilon^{\dot A}_A i\not{D}\overline{Z}^A + \varepsilon_{B\dot B}\overline{Z}^B Y^{\dot A\dot B}. \tag{2.229}$$

The term with the moment map μ^A_B in the transformation of the Ψ_2 fermion is present only for the magnetic theory. In the language of $\mathcal{N}=1$ three-dimensional superfields, it comes from the $\delta(x_3)$ term in the auxiliary field F_Y (see eqn. (2.264) for more details). This term propagates in all equations in combination with D_3X^a, canceling the delta-contribution from the discontinuity of the field X^a.

2.7 Appendix B: Details On The Action And The Twisting

2.7.1 Constructing The Action From $\mathcal{N}=1$ Superfields

In this section, we review the construction [23] of the action for the D3-NS5 system. One of the reasons for discussing this in some detail is that we will need parts of it to write out the action for the magnetic theory.

Here we work in Euclidean signature. In [23], the D3-NS5 action was constructed by writing an $\mathcal{N}=1$ 3d supersymmetric action with a global SU(2) symmetry, and then adjusting the couplings to extend this symmetry to a product $\mathrm{SU}(2)_X\times\mathrm{SU}(2)_Y$. This group does not commute with the supersymmetry generators, and therefore extends the $\mathcal{N}=1$ supersymmetry to $\mathcal{N}=4$. The $\mathcal{N}=1$ multiplets in the bulk are a vector multiplet[50] (A_i,ξ_A)

[50] The subscript A in ξ_A is not an R-symmetry index.

and three chiral multiplets (X^a, ρ_1^a, F_X^a), (Y^a, ρ_2^a, F_Y^a) and (A_3, ξ_3, F_3), where X^a and Y^a are the six scalars of the $\mathcal{N} = 4$ SYM[51], and A_3 is a component of the gauge field. The fermionic fields can be packed into two $\mathcal{N} = 4$ SUSY covariant combinations

$$
\begin{aligned}
\sqrt{2}\Psi_1^{A\dot{B}} &= -i\rho_1^{(A\dot{B})} + \epsilon^{A\dot{B}}(-\sin\vartheta\,\xi_A + \cos\vartheta\,\xi_3)\,, \\
\sqrt{2}\Psi_2^{A\dot{B}} &= -i\rho_2^{(A\dot{B})} + \epsilon^{A\dot{B}}(-\cos\vartheta\,\xi_A - \sin\vartheta\,\xi_3).
\end{aligned}
\tag{2.230}
$$

The action of the bulk $\mathcal{N} = 4$ super Yang-Mills, rephrased in three-dimensional notation, has the following form,

$$
\begin{aligned}
&-\frac{1}{g_{\mathrm{YM}}^2}\int \mathrm{d}^4x\,\mathrm{tr}\left(\frac{1}{2}F_{\mu\nu}^2 + (D_iX^a)^2 + (D_iY^a)^2\right.\\
&+i\Psi_{1\alpha}^{A\dot{B}}D_\beta^\alpha\Psi_{1A\dot{B}}^\beta + i\Psi_{2\alpha}^{A\dot{B}}D_\beta^\alpha\Psi_{2A\dot{B}}^\beta + 2\Psi_{2\alpha}^{A\dot{B}}D_3\Psi_{1A\dot{B}}^\alpha\\
&+X_B^A\left(-\sin\vartheta([\Psi_1^{B\dot{C}\alpha},\Psi_{1A\dot{C}\alpha}] - [\Psi_2^{B\dot{C}\alpha},\Psi_{2A\dot{C}\alpha}]) - 2\cos\vartheta[\Psi_2^{B\dot{C}\alpha},\Psi_{1A\dot{C}\alpha}]\right)\\
&+Y_{\dot{D}}^{\dot{C}}\left(-\cos\vartheta([\Psi_1^{A\dot{D}\alpha},\Psi_{1A\dot{C}\alpha}] - [\Psi_2^{A\dot{D}\alpha},\Psi_{2A\dot{C}\alpha}]) + 2\sin\vartheta[\Psi_2^{A\dot{D}\alpha},\Psi_{1A\dot{C}\alpha}]\right)\\
&-F_X^2 - F_Y^2 - F_3^2 + 2D_3\left(F_XY\right) - 2F_3[X,Y]\\
&+F_X^a\left(-2D_3Y_a - \sin\vartheta\epsilon_{abc}([X^b,X^c] - [Y^b,Y^c]) - 2\cos\vartheta\epsilon_{abc}[X^b,Y^c]\right)\\
&\left.+F_Y^a\left(2D_3X_a - \cos\vartheta\epsilon_{abc}([X^b,X^c] - [Y^b,Y^c]) + 2\sin\vartheta\epsilon_{abc}[X^b,Y^c]\right)\right)\\
&+\frac{i\theta_{\mathrm{YM}}}{8\pi^2}\int\mathrm{tr}\left(F\wedge F\right)\\
&+\int\mathrm{d}^4x\,\mathrm{tr}\left(\frac{\theta_{\mathrm{YM}}}{8\pi^2}\partial_3\left(\xi_A^2\right) - \frac{1}{g_{\mathrm{YM}}^2}\partial_3\left((\xi_A^2 - \xi_3^2)\sin\vartheta\cos\vartheta - 2\xi_3\xi_A\cos^2\vartheta\right)\right).
\end{aligned}
\tag{2.231}
$$

Here the first four lines are the usual kinetic and Yukawa terms. The next three lines contain the auxiliary fields, after eliminating which these terms will give the usual quartic $\mathcal{N} = 4$ super Yang-Mills potential, but they will also give some total ∂_3 derivatives, which we cannot drop if we want to couple the theory to the defect in a supersymmetric way. Next, there is also a theta-term, and finally in the last line there are some total derivatives of the non-R-

[51] In non-R-symmetrized expressions, where only the diagonal subgroup of the $\mathrm{SU}(2)_X\times\mathrm{SU}(2)_Y$ is explicitly visible, it does not make sense to distinguish $\mathrm{SU}(2)_X$ and $\mathrm{SU}(2)_Y$ indices.

symmetric combinations of fermions, which appear from rearranging the fermionic kinetic terms and from the theta-term.

For the NS5-type defect we can use (2.7) to reduce the last line in (2.231) to

$$\frac{\cot\vartheta}{g_{\mathrm{YM}}^2}\int \mathrm{d}^4x\,\partial_3\mathrm{tr}\left(\xi_A\cos\vartheta+\xi_3\sin\vartheta\right)^2. \tag{2.232}$$

This term is important for R-symmetrizing the fermionic couplings on the boundary.

On the three-dimensional defect live chiral multiplets (Q^A,λ^A,F_Q^A). In $\mathcal{N}=1$ notation, the action on the defect includes a usual kinetic term for the Q-multiplet, a quartic superpotential $\frac{\mathcal{K}}{4\pi}\mathcal{W}_4(\mathcal{Q})$ with

$$\begin{aligned}\mathcal{W}_4 &= \frac{1}{12}t_{IJ;KS}\epsilon^{AB}\epsilon^{CD}\mathcal{Q}_A^I\mathcal{Q}_B^J\mathcal{Q}_C^K\mathcal{Q}_D^S\,,\\ t_{IJ;KS} &= \frac{1}{4}\kappa^{mn}\left(\tau_{mIK}\tau_{nJS}-\tau_{mIS}\tau_{nJK}\right)\,,\end{aligned} \tag{2.233}$$

and a superpotential that couples the four-dimensional scalar X^a to the defect theory,

$$\mathcal{W}_{QXQ} = -\frac{\mathcal{K}}{4\pi}\sin\vartheta\,\mathcal{Q}^{IA}\mathcal{X}_{AB}^m\tau_{mIJ}\mathcal{Q}^{JB}. \tag{2.234}$$

This choice of the superpotential corresponds to the case when the $NS5$-brane is stretched in directions 456. Indeed, the bifundamental fields will have a mass term proportional to X^2, *i.e.* their mass is proportional to the displacement in these directions.

The boundary conditions of the theory form a current multiplet of three-dimensional

$\mathcal{N}=4$ supersymmetry,

$$\begin{aligned}
&Y^m_{\dot{A}\dot{B}} = -\frac{1}{2\cos\vartheta}\tau^m_{IJ}Q^I_{\dot{A}}Q^J_{\dot{B}}\,,\\
&\sqrt{2}\Psi^m_{2\alpha A\dot{B}} = \frac{i}{\cos\vartheta}\tau^m_{IJ}\lambda^I_{\alpha A}Q^J_{\dot{B}}\,,\\
&\sin\vartheta F^m_{k3} - \frac{i}{2}\cos\vartheta\epsilon_{ijk}F^m_{ij} = -\frac{2\pi}{\cos\vartheta}\kappa^{mn}J_{nk}\,,\\
&D_3X^m_a - \frac{1}{2}\cos\vartheta\epsilon_{abc}f^m_{np}X^{bn}X^{cp} = \frac{1}{2}\tan\vartheta\,\omega_{IJ}\epsilon_{\dot{A}\dot{B}}X^n_aT^{mI}_KT^J_{nS}Q^{K\dot{A}}Q^{S\dot{B}}\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad -\frac{1}{4\cos\vartheta}\lambda^I_A\sigma^{AB}_a\tau^m_{IJ}\lambda^J_B\,,
\end{aligned} \tag{2.235}$$

where J_{mk} is the current

$$J_{mi} = \frac{\delta I_Q}{\delta A^{im}} = \frac{1}{4\pi}\tau_{mIJ}\left(\epsilon^{\dot{A}\dot{B}}Q^I_{\dot{A}}D_iQ^J_{\dot{B}} + \epsilon^{AB}\frac{i}{2}\lambda^I_A\sigma_i\lambda^J_B\right). \tag{2.236}$$

The first of the boundary conditions has the following origin. At stationary points of the action the auxiliary field F^a_X has a contribution from the boundary, proportional to the delta function. Then the term F^2_X would produce a square of the delta function. To avoid this and to make sense of the action, the boundary contribution to F^a_X should be set to zero, and this gives the boundary condition for the field $Y^{\dot{a}}$. The other three boundary conditions can be obtained in a usual way from the variation of the action, after eliminating the auxiliary fields.

The complete action after eliminating the auxiliary fields is

$$\begin{aligned}
I_{\text{electric}} \;=\;& I_{\text{SYM}} + \frac{i\theta_{\text{YM}}}{2\pi}\text{CS}(A) + \mathcal{K}I_Q(A)\\
&+\frac{\mathcal{K}}{4\pi}\int \mathrm{d}^3x\left(\frac{1}{2}\sin^2\vartheta\omega_{IJ}\epsilon_{\dot{A}\dot{B}}X^{ma}X^{na}T^I_{mK}T^J_{nS}Q^{K\dot{A}}Q^{S\dot{B}} - \frac{1}{2}\sin\vartheta\lambda^I_AX^{mAB}\tau_{mIJ}\lambda^J_B\right)\\
&+\frac{1}{g^2_{\text{YM}}}\int \mathrm{d}^3x\,\text{Tr}\left(-\frac{2}{3}\epsilon_{abc}\cos\vartheta X^aX^bX^c - \frac{2}{3}\epsilon_{abc}\sin\vartheta Y^aY^bY^c + 2\Psi^{A\dot{B}}_1\Psi_{2A\dot{B}}\right)
\end{aligned} \tag{2.237}$$

where

$$I_Q(A) = \frac{1}{4\pi}\int \mathrm{d}^3x\,\bigg(\frac{1}{2}\epsilon^{\dot{A}\dot{B}}\omega_{IJ}D_iQ^I_{\dot{A}}D^iQ^J_{\dot{B}} - \frac{i}{2}\epsilon^{AB}\omega_{IJ}\lambda^I_A \not{D}\lambda^J_B$$
$$+\frac{1}{4}\kappa^{mn}\tau_{mIJ}\tau_{nKS}Q^{I\dot{A}}Q^K_{\dot{A}}\lambda^{JC}\lambda^S_C + \frac{1}{2}\epsilon_{\dot{A}\dot{B}}\omega^{IJ}\frac{\partial W_4}{\partial Q^I_{\dot{A}}}\frac{\partial W_4}{\partial Q^J_{\dot{B}}}\bigg). \tag{2.238}$$

is the $\mathcal{N}=4$ super Chern-Simons action with the CS term omitted.

Before proceeding to twisting, it is useful to remove the term $\lambda X \lambda$ in the action, using the last line in the boundary conditions[52] (2.235). Then the action is

$$\begin{aligned}
I_{\text{electric}} \;=\;& I_{\text{SYM}} + \frac{i\theta_{\text{YM}}}{2\pi}\mathrm{CS}(A) + \mathcal{K}I_Q(A)\\
&+\frac{\mathcal{K}}{4\pi}\int \mathrm{d}^3x\left(-\frac{1}{2}\sin^2\vartheta\,\omega_{IJ}\epsilon_{\dot{A}\dot{B}}X^{ma}X^{na}T^I_{mK}T^J_{nS}Q^{K\dot{A}}Q^{S\dot{B}}\right)\\
&+\frac{1}{g^2_{\text{YM}}}\int \mathrm{d}^3x\,\mathrm{Tr}\left(-\frac{2}{3}\epsilon_{abc}\cos\vartheta X^aX^bX^c - \frac{2}{3}\epsilon_{abc}\sin\vartheta Y^aY^bY^c + 2\Psi_1^{A\dot{B}}\Psi_{2A\dot{B}}\right)\\
&-\frac{2}{g^2_{\text{YM}}}\int \mathrm{d}^3x\,\mathrm{Tr}\left(X^aD_3X_a - \cos\vartheta\epsilon_{abc}X^aX^bX^c\right).
\end{aligned} \tag{2.239}$$

The supersymmetry transformations for this theory can be found by R-symmetrization of the $\mathcal{N}=1$ supersymmetry transformations, or, for the bulk super Yang-Mills fields, by reduction from the $\mathcal{N}=4$ formulas in four dimensions. The result can be found in Appendix 2.6.

2.7.2 Twisted Action

Now we would like to twist the theory and to couple it to the metric. Let us recall, what is the set of fields of our topological theory. The four scalars X^a and Y^1 of the bulk super Yang-Mills become components of a 1-form ϕ, and the other two scalars are combined as $\sigma = \frac{Y_2 - iY_3}{\sqrt{2}}$ and $\overline{\sigma} = \frac{Y_2 + iY_3}{\sqrt{2}}$. The fermions of the twisted bulk theory are [14] two scalars η and $\widetilde{\eta}$, two one-forms ψ and $\widetilde{\psi}$, and a 2-form χ. The selfdual and anti-selfdual parts of the two forms are denoted by $\pm$ superscripts. These fermions are related to the fields of the

[52]One might be worried that after this transformation the action no longer gives the same boundary conditions from the boundary variation. In section 2.7.3 we will make our argument more accurate.

physical theory as follows,

$$
\begin{aligned}
2\sqrt{2}\Psi_1^{\alpha A\dot{A}} &= (\widetilde{\eta}-t^{-1}\eta)\epsilon^{\alpha A}v^{\dot{A}}+(-\widetilde{\psi}-t\psi)_3\epsilon^{\alpha A}u^{\dot{A}}+ \\
&\quad +2(t^{-1}\chi^{+}+\chi^{-})_{i3}\sigma_i^{\alpha A}v^{\dot{A}}+(\widetilde{\psi}-t\psi)_i\sigma_i^{\alpha A}u^{\dot{A}}\,, \\
-2\sqrt{2}i\Psi_2^{\alpha A\dot{A}} &= (-\widetilde{\eta}-t^{-1}\eta)\epsilon^{\alpha A}v^{\dot{A}}+(-\widetilde{\psi}+t\psi)_3\epsilon^{\alpha A}u^{\dot{A}}+ \\
&\quad +2(t^{-1}\chi^{+}-\chi^{-})_{i3}\sigma_i^{\alpha A}v^{\dot{A}}+(\widetilde{\psi}+t\psi)_i\sigma_i^{\alpha A}u^{\dot{A}}.
\end{aligned} \tag{2.240}
$$

Here is a summary of Q-transformations of the bulk fields, as derived in [14],

$$
\begin{aligned}
&\delta A = it\widetilde{\psi}+i\psi\,, &&\delta\phi = -i\widetilde{\psi}+it\psi\,, \\
&\delta\sigma = 0\,, &&\delta\overline{\sigma} = it\widetilde{\eta}+i\eta\,, \\
&\delta\eta = tP+[\overline{\sigma},\sigma]\,, &&\delta\widetilde{\eta} = -P+t[\overline{\sigma},\sigma]\,, \\
&\delta\psi = D\sigma+t[\phi,\sigma]\,, &&\delta\widetilde{\psi} = tD\sigma-[\phi,\sigma]\,, \\
&\delta\chi = H\,,
\end{aligned} \tag{2.241}
$$

where on-shell

$$
P = D^{\mu}\phi_{\mu}\,, \quad H^{+} = \mathcal{V}^{+}(t)\,, \quad H^{-} = t\mathcal{V}^{-}(t) \tag{2.242}
$$

and

$$
\begin{aligned}
\mathcal{V}^{+}(t) &= (F-\phi\wedge\phi+tD\phi)^{+}\;, \\
\mathcal{V}^{-}(t) &= \left(F-\phi\wedge\phi-t^{-1}D\phi\right)^{-}.
\end{aligned} \tag{2.243}
$$

As it was described in [14], the manifestly Q-invariant topological action for the bulk super Yang-Mills theory contains a topological term and a Q-variation of a fermionic expression (see section 3.4 of that paper). In our case the theory is defined on the two half-spaces with the defect W between them, and therefore the equations have to be completed with some boundary terms:

$$
\begin{aligned}
I_{\mathrm{SYM}} &= \{Q, \dots\} - \frac{t - t^{-1}}{t + t^{-1}} \frac{4\pi}{g_{\mathrm{YM}}^2} \, \mathrm{CS}(A) \\
&+ \frac{1}{g_{\mathrm{YM}}^2} \int_W \mathrm{Tr} \left(\frac{4}{t + t^{-1}} \left(F \wedge \phi - \frac{1}{3} \phi \wedge \phi \wedge \phi \right) + \frac{t - t^{-1}}{t + t^{-1}} \phi \wedge D\phi \right) \\
&+ \frac{1}{g_{\mathrm{YM}}^2} \int_W \mathrm{d}^3 x \sqrt{\gamma} \, \mathrm{Tr} \left(2\overline{\sigma} D_3 \sigma + \gamma^{ij} \phi_i D_j \phi_3 - \gamma^{ij} \phi_3 D_j \phi_i \right) . \qquad (2.244)
\end{aligned}
$$

Let us give some explanations on this formula. Recall that in our notation, I_{SYM} is the part of the bulk super Yang-Mills action, which is proportional to $1/g_{\mathrm{SYM}}^2$, – that is, with the θ_{YM}-part omitted. Here and in what follows we ignore expressions on W bilinear in the bulk fermions, because in the end they have to cancel by supersymmetry, anyway. As usual, the Chern-Simons form $CS(A)$ is just a notation for the bulk topological term. By γ we denote the induced metric on W. The third component of various bulk tensors on the boundary is defined as a contraction of these tensors with a unit vector field n^μ, normal to the defect. For example, $D_j \phi_3$ means a pullback to W of a one-form $n^\nu D_\mu \phi_\nu$.

The first line in the expression above is the formula that was used in [14]. The coefficient of the topological term in this expression adds with the usual theta parameter θ_{YM} to become the canonical parameter, which we called $\mathcal{K}$. The second line in this formula is what appeared in the purely bosonic Chern-Simons case [6]. Finally, the last line was dropped in that paper as a consequence of the boundary conditions, but in our case it is non-zero.

A useful transformation is to integrate by parts in the last line of (2.244) to change $-\phi_3 D^i \phi_i$ into another $\phi^i D_i \phi_3$, but in doing so we have to remember that the metric connection in the covariant derivatives is four-dimensional. Because of this, the integration by parts produces a curvature term

$$
\frac{1}{g_{\mathrm{YM}}^2} \int_W \mathrm{d}^3 x \sqrt{\gamma} \, \mathrm{Tr} \left(-s_{ij} \phi^i \phi^j + s^i_i \phi_3 \phi_3 \right) , \qquad (2.245)
$$

where s_{ij} is the second fundamental form of the hypersurface W. This curvature term should be canceled by adding a curvature coupling to the last line in (2.239).

We will substitute what we have just learned about I_{SYM} into the action (2.239) of the theory, but first let us make some transformations of the action (2.239). We would like to complexify the gauge field in the hypermultiplet action $I_Q(A)$. The seagull term for $(DQ)^2$ comes from $XXQQ$ in the second line of (2.239). To change the terms linear in the gauge field we need to add and subtract $i\sin\vartheta X$ times the boundary current (2.236). Using the third of the boundary conditions (2.235), the current can be rewritten as a combination of gauge field strengths. After these manipulations, a twisted version of (2.239) will look like

$$\begin{aligned} I_{\text{electric}} \;=\;& I_{\text{SYM}} + \frac{i\theta_{\text{YM}}}{2\pi}\text{CS}(A) + \mathcal{K}I_Q(\mathcal{A}_b) + \frac{1}{g^2_{\text{YM}}}\int \mathrm{d}^3x\,\text{Tr}\left(-\frac{2}{3}\cos\vartheta\phi\wedge\phi\wedge\phi\right) \\ &+\frac{1}{g^2_{\text{YM}}}\int \mathrm{d}^3x\,\sqrt{\gamma}\,\text{Tr}\left(-i\sin\vartheta\phi_3[\overline{\sigma},\sigma] - 2\phi^i D_3\phi_i - 2i\sin\vartheta\,\phi^i F_{i3} + s_{ij}\phi^i\phi^j - s^i_i\phi_3\phi_3\right) \\ &+\frac{2\cos\vartheta}{g^2_{\text{YM}}}\int \mathrm{d}^3x\,\text{Tr}\left(\phi\wedge\phi\wedge\phi - \phi\wedge F\right). \end{aligned} \tag{2.246}$$

Now we substitute here the expression (2.244) for the super Yang-Mills action. The Chern-Simons term in (2.244) changes the coefficient in front of the Chern-Simons term in (2.246) from $\theta_{\text{YM}}/2\pi$ to $\mathcal{K}$. Expression in the second line in (2.244) and the term with $\phi\wedge\phi\wedge\phi$ in the first line of (2.246) combine with the Chern-Simons term, changing the gauge field in it from A into complexified gauge field $\mathcal{A}_b$, as shown in [6]. We are left with the following action,

$$\begin{aligned} &I_{\text{electric}} = \{Q,\dots\} + i\mathcal{K}\text{CS}(\mathcal{A}_b) + \mathcal{K}I_Q(\mathcal{A}_b) \\ &+\frac{1}{g^2_{\text{YM}}}\int \mathrm{d}^3x\,\sqrt{\gamma}\,\text{Tr}\left(-i\sin\vartheta\phi_3[\overline{\sigma},\sigma] + 2\overline{\sigma}D_3\sigma\right) \\ &+\frac{1}{g^2_{\text{YM}}}\int \mathrm{d}^3x\,\sqrt{\gamma}\,\text{Tr}\left(-2\phi^i D_3\phi_i + 2\phi^i D_i\phi_3 - 2i\sin\vartheta\,\phi^i F_{i3}\right) \\ &+\frac{2\cos\vartheta}{g^2_{\text{YM}}}\int \text{Tr}\left(-\phi\wedge F + \phi\wedge\phi\wedge\phi\right). \end{aligned} \tag{2.247}$$

We are almost done. All we need to show is that the last three lines here are Q-exact. This is indeed so (again, we ignore the fermion bilinears):

$$
\begin{aligned}
&\int \mathrm{d}^3x\,\sqrt{\gamma}\,\mathrm{Tr}\,(\overline{\sigma}D_3\sigma) = -\frac{1}{2\cos\vartheta}\left\{Q, \int \mathrm{d}^3x\,\sqrt{\gamma}\,\mathrm{Tr}\left(\overline{\sigma}(t^{-1}\psi_3 + \widetilde{\psi}_3)\right)\right\}\,,\\
&\int \mathrm{d}^3x\,\sqrt{\gamma}\,\mathrm{Tr}\,(\phi_3[\overline{\sigma},\sigma]) = -\frac{1}{2\cos\vartheta}\left\{Q, \int \mathrm{d}^3x\,\sqrt{\gamma}\,\mathrm{Tr}\left(\overline{\sigma}(t^{-1}\widetilde{\psi}_3 - \psi_3)\right)\right\}\,, \qquad (2.248)\\
&\int \mathrm{Tr}\,(\phi\wedge(\star D\phi - i\sin\vartheta\,\star F - \cos\vartheta(F-\phi\wedge\phi))) = \left\{Q, \int \mathrm{Tr}\left(\phi\wedge(t^{-1}\chi^+ + \chi^-)\right)\right\}.
\end{aligned}
$$

Up to Q-exact terms, our action is the sum of the Chern-Simons term and the twisted action $I_Q(\mathcal{A}_b)$. This combination is just the (twisted) action of the $\mathcal{N}=4$ Chern-Simons theory. Let us see, how it is related [24] to the Chern-Simons theory with a supergroup. We define the fields of the twisted theory as

$$
\begin{aligned}
Q^{\dot{A}} &= iv^{\dot{A}}\overline{C} + \frac{1}{2}u^{\dot{A}}C\,,\\
\lambda^{\alpha A} &= -\frac{i}{2}\epsilon^{\alpha A}B + i\sigma^{i\alpha A}\mathcal{A}_{f_i}.
\end{aligned} \qquad (2.249)
$$

Substituting this into the action and using the explicit form (2.238) of $I_Q(A)$, one finds,

$$
i\mathcal{K}\,\mathrm{CS}(\mathcal{A}_b) + \mathcal{K}\,I_Q(\mathcal{A}_b) = i\mathcal{K}\,\mathrm{CS}(\mathcal{A}) + i\mathcal{K}I_{\text{g.f.}}\,, \qquad (2.250)
$$

where $\mathcal{A} = \mathcal{A}_b + \mathcal{A}_f$ is the complexified superconnection. The Q-exact gauge fixing term $I_{\text{g.f.}} = \{Q, V_{\text{g.f.}}\}$ for the fermionic part of the superalgebra is

$$
\begin{aligned}
I_{\text{g.f.}} &= \int \mathrm{d}^3x\,\sqrt{\gamma}\,\mathrm{Str}\left(-\mathcal{D}_b^iB\mathcal{A}_{f_i} + \mathcal{D}_b^i\overline{C}\mathcal{D}_{bi}C + \{\mathcal{A}_f,\overline{C}\}\{\mathcal{A}_f,C\}\right.\\
&+ \left.\frac{1}{4}\{\overline{C},B\}\{C,B\} + \frac{1}{16}[C,\{\overline{C},\overline{C}\}][\overline{C},\{C,C\}]\right)\,, \qquad (2.251)\\
V_{\text{g.f.}} &= \int \mathrm{d}^3x\,\sqrt{\gamma}\,\mathrm{Str}\left(-\mathcal{D}_b^i\overline{C}\mathcal{A}_{f_i} + \frac{1}{8}\{\overline{C},\overline{C}\}\{C,B\}\right).
\end{aligned}
$$

2.7.3 Boundary Conditions

Let us rewrite the boundary conditions (2.235) in terms of fields of the twisted theory. The first line of that formula gives

$$\sigma = \frac{i}{2}\frac{1}{1+t^2}\{C,C\}\,, \quad \overline{\sigma} = \frac{i}{1+t^{-2}}\{\overline{C},\overline{C}\}\,, \quad \phi_3 = -\frac{1}{t+t^{-1}}\{C,\overline{C}\}. \tag{2.252}$$

These three formulas are related to one another by $\mathrm{SU}(2)_Y$ rotations. The boundary condition for the fermion in (2.235) gives one new relation

$$t^{-1}\chi^+_{i3} - \chi^-_{i3} = \frac{2}{t+t^{-1}}\{\mathcal{A}_{f_i},\overline{C}\}\,, \tag{2.253}$$

two relations, that can be obtained from (2.252) by Q-transformations

$$\widetilde{\eta} + t^{-1}\eta = \frac{2}{t+t^{-1}}\{B,\overline{C}\}\,, \quad -\widetilde{\psi}_3 + t\psi_3 = \frac{i}{t+t^{-1}}\{B,C\}\,, \tag{2.254}$$

and one relation which comes from the bulk and boundary Q-variation of the gauge field $\mathcal{A}_b$, which we have already discussed,

$$\widetilde{\psi}_i + t\psi_i = -\frac{2i}{t+t^{-1}}\{\mathcal{A}_{f_i},C\}. \tag{2.255}$$

The third line in (2.235) gives boundary condition for the gauge field,

$$\cos\vartheta\,\imath^*\left(i\sin\vartheta\star F + \cos\vartheta F\right) = -\mathcal{A}_f\wedge\mathcal{A}_f + \frac{1}{2}\star_3\left(\{C,D\overline{C}\} - \{\overline{C},DC\} + [B,\mathcal{A}_f]\right). \tag{2.256}$$

The twisted version of the last line in (2.235) is a long expression with a contribution from the curvature coupling. It can be somewhat simplified by subtracting a D_i derivative of the boundary condition (2.252) for ϕ_3. The result is the following,

$$\cos\vartheta\imath^*\left(\star D\phi+\cos\vartheta\,\phi\wedge\phi\right)=-\mathcal{A}_f\wedge\mathcal{A}_f+\frac{1}{2}\star_3\bigl(D\{C,\overline{C}\}+i\sin\vartheta\left(\{\overline{C},[\phi,C]\}-\{C,[\phi,\overline{C}]\}\right)-[B,\mathcal{A}_f]\bigr)\,. \tag{2.257}$$

If we subtract (2.257) and (2.256), we get just a Q-variation of the fermionic boundary condition (2.253). A new relation results, if we add these two:

$$\mathcal{F}_b+\mathcal{A}_f\wedge\mathcal{A}_f=\star_3\{C,D\overline{C}-i\sin\vartheta[\phi,\overline{C}]\}-\{Q,\chi_t\}\ , \tag{2.258}$$

where we defined

$$\chi_t=\frac{t^{-2}-3}{4}\chi^+ +\frac{t^2-3}{4}\chi^-. \tag{2.259}$$

Q^2 acts as a gauge transformation with parameter $-i(1+t^2)\sigma$ in the bulk and with parameter $\{C,C\}/2$ on the defect (2.33), (2.31). This agrees with the boundary conditions.

The Q-transformations of the set of boundary ghosts $\overline{C}$, C and B were given in (2.31). To fix the residual gauge symmetry in perturbation theory, we introduce the usual ghosts c, $\overline{c}$ and the Lagrange multiplier field b, and the BRST-differential $Q_{\rm bos}$, associated to this gauge fixing. This differential acts on all fields in the usual fashion. The topological differential Q acts trivially on b and $\overline{c}$, but generates the following transformation, when acting on c:

$$\delta c=i(1+t^2)\sigma. \tag{2.260}$$

On the boundary, this corresponds to [24]

$$\delta c=-\frac{1}{2}\{C,C\}. \tag{2.261}$$

The full BRST differential in the gauge fixed theory is the sum $Q+Q_{\rm bos}$. This operator squares to zero, and in the boundary theory it corresponds to the usual gauge fixing for the full supergroup gauge symmetry.

Finally, let us comment on the fact that we used the boundary conditions to transform the action (to pass from (2.238) to (2.239), and then to get (2.246)). We did it to exploit more directly the relation to the $\mathcal{N}=4$ Chern-Simons theory, but that transformation was not really necessary. Indeed, the terms that came from using the boundary conditions gave essentially the last line in the list (2.248) of Q-exact expressions. The combination of the boundary conditions that we used was just a Q-variation of the boundary condition for the χ fermion (2.253). (More precisely, this combination differs by a derivative of (2.252), but this is fine, since the boundary condition (2.252) is Dirichlet.) So we could equally well keep the expressions that involved the hypermultiplet fields, instead of transforming them into the bulk fields, and this would give Q-exact expressions as well.

2.8 Appendix C: Details On The Magnetic Theory

2.8.1 Action Of The Physical Theory

Here we would like to give some details on the derivation of the action and the boundary conditions for the D3-D5 system, with equal numbers if the D3-branes in the two sides of the D5-brane. This action has been constructed in [83], but our treatment of the boundary conditions is slightly different.

As in the electric theory, we write the action in the three-dimensional $\mathcal{N}=1$ formalism. The bulk super Yang-Mills part of the action has been given in (2.231) (one should set ϑ to π in that formula). On the defect there is a fundamental hypermultiplet (Z^A, ζ^A, F^A), where the first two fields have already appeared in our story, and F^A is the auxiliary field. Besides the usual kinetic term, the boundary action contains a superpotenial that couples the bulk and the boundary fields,

$$\mathcal{W}_{\overline{Z}YZ} = -\overline{\mathcal{Z}}_A \mathcal{Y}^A_B \mathcal{Z}^B. \qquad (2.262)$$

This superpotential has been chosen in such a way as to make the boundary interactions invariant[53] under the full $SO(3)_X \times SO(3)_Y$ R-symmetry group. Specifically, the boundary action contains Yukawa couplings $-i\overline{\zeta}_A \xi_A Z^A + i\overline{Z}_A \xi_A \zeta^A$ coming from the kinetic term, and $\overline{Z}_A \rho_2^{AB} \zeta^B + \overline{\zeta}_A \rho_2^{AB} Z_B$ from the superpotential. They can be packed into R-symmetric couplings

$$i\sqrt{2}\left(\overline{Z}_A \Psi_2^{A\dot{B}} \zeta_{\dot{B}} + \overline{\zeta}_{\dot{B}} \Psi_2^{A\dot{B}} Z_A\right) , \tag{2.263}$$

where the $\mathcal{N}=4$ fermion $\Psi_2^{A\dot{B}}$ was defined in (2.230).

The superpotential contains a coupling of the auxiliary field F_Y to the moment map μ_m^a, which was defined in (2.106). This coupling will add a delta-function contribution to the equation for the auxiliary field,

$$F_Y^{am} = D_3 X^{am} + \frac{1}{2}\epsilon^{abc}([X_b, X_c] - [Y_b, Y_c])^m - \frac{1}{2}\mu^{am}\delta(x^3). \tag{2.264}$$

The square of the auxiliary field in the Yang-Mills action would produce a term with a square of this delta-function. To make this contribution finite, we require the scalars X^a to have a discontinuity across the defect. This discontinuity equation extends via the supersymmetry to a set of equations for two three-dimensional current multiplets,

$$\begin{aligned}
&X^{am}\big|_-^+ = \frac{1}{2}\mu^{am} , \\
&\sqrt{2}\Psi_{1m}^{A\dot{B}}\big|_-^+ = i\left(\overline{\zeta}^{\dot{B}} T_m Z^A + \overline{Z}^A T_m \zeta^{\dot{B}}\right) , \\
&F_{3i}^m\big|_-^+ = \frac{1}{2}J_i^m , \\
&D_3 Y_m^a\big|_-^+ = \frac{1}{2}\left(\overline{Z}_A\{Y^a, T_m\}Z^A - \overline{\zeta}^{\dot{A}} T_m \zeta^{\dot{B}} \sigma^a_{\dot{A}\dot{B}}\right) ,
\end{aligned} \tag{2.265}$$

where the current is

$$J_{mi} = \frac{\delta I_{\text{hyp}}^{\vee}}{\delta A^{mi}} = -\overline{Z}_A T_m D_i Z^A + D_i \overline{Z}_A T_m Z^A - i\overline{\zeta}^{\alpha}_{\dot{A}} T_m \sigma^{\beta}_{i\alpha} \zeta^{\dot{A}}_{\beta}. \tag{2.266}$$

[53] As we have said, we choose $t^{\vee} = 1$. For $t^{\vee} = -1$ the sign of the superpotential would be the opposite.

Next we have to substitute expressions for all the auxiliary fields into the Lagrangian, and make it manifestly R-symmetric. Also, we would like to rearrange the action in such a way that the squares of the delta-function would not appear. In the Yang-Mills action (2.231) there is a potentially dangerous term F_Y^2, but with the gluing conditions (2.265) it is non-singular and produces no finite contribution at $x_3 = 0$. Then for this term we can replace the x_3-integral over $\mathbb{R}$ by an integral over $x_3 < 0$ and $x_3 > 0$. The term F_X^2 is also non-singular, so we delete the plane $x^3 = 0$ in the same way. There is a singular term $D_3(F_X Y)$, but in can be dropped as a total derivative. The total ∂_3 derivative of the non-R-symmetric fermion combination in (2.231) can be dropped in the same way. There is also a delta-function contribution from the D_3 part of the fermionic kinetic term. Collecting all the boundary terms in the integrals with $x_3 = 0$ deleted, we get a simple action

$$\begin{aligned} I_{\text{magnetic}} &= I_{\text{SYM}} + \frac{i\theta^\vee_{\text{YM}}}{8\pi^2}\int \operatorname{tr}(F\wedge F) \\ &+ \frac{1}{(g^\vee_{\text{YM}})^2}\int \mathrm{d}^3x \left(D_i\overline{Z}_A D^i Z^A - i\overline{\zeta}_{\dot{A}}\not{D}\zeta^{\dot{A}} - \overline{\zeta}_{\dot{A}}Y^{\dot{A}}_{\dot{B}}\zeta^{\dot{B}} - \overline{Z}_A Y^a Y_a Z^A\right) \\ &+ \frac{1}{(g^\vee_{\text{YM}})^2}\int \mathrm{d}^3x\, \frac{2}{3}\operatorname{tr}\left(\epsilon_{abc}(X^aX^bX^c)\big|^+_-\right). \end{aligned} \tag{2.267}$$

Here in I_{SYM} the usual super Yang-Mills Lagrangian in the bulk is integrated over the two half-spaces $x_3 < 0$ and $x_3 > 0$, with the hyperplane $x_3 = 0$ deleted. On the defect the $\overline{Z}YYZ$ terms from the superpotential combined with the XYY term from the bulk action into an R-symmetric coupling. The Yukawa terms $\overline{\zeta}\Psi_2 Z + \overline{Z}\Psi_2\zeta$ canceled with the delta-contribution from the bulk fermionic kinetic energy.

2.8.2 Action Of The Twisted Theory

From the action of supersymmetry (2.229) one finds the following Q-transformations for the boundary fields of the twisted theory,

$$\begin{aligned}
\delta Z &= -2i\zeta_u\,, \\
\delta \overline{Z} &= -2i\overline{\zeta}_u\,, \\
\delta \zeta_u &= \sigma Z\,, \\
\delta \overline{\zeta}_u &= -\overline{Z}\sigma.
\end{aligned} \tag{2.268}$$

The two other fermions transform as $\delta\zeta_v = f$ and $\delta\overline{\zeta}_v = \overline{f}$, where

$$\begin{aligned}
f &= \not{D} Z + \phi_3 Z\,, \\
\overline{f} &= \not{D}\overline{Z} - \overline{Z}\phi_3\,,
\end{aligned} \tag{2.269}$$

but with these transformation rules the algebra does not close off-shell. For this reason we introduce two auxiliary bosonic spinor fields F and $\overline{F}$, for which the equations of motion should impose $F = f$ and $\overline{F} = \overline{f}$. The topological transformations are then

$$\begin{aligned}
\delta \zeta_v &= F\,, \\
\delta \overline{\zeta}_v &= \overline{F}\,, \\
\delta F &= -2i\sigma\zeta_v\,, \\
\delta \overline{F} &= 2i\overline{\zeta}_v\sigma.
\end{aligned} \tag{2.270}$$

The transformation rules for the auxiliary fields were chosen in a way to ensure that the square of the topological supercharge acts by the same gauge transformation, by which it acts on the other fields.

Now we would like to prove our claim that the action of the magnetic theory is Q-exact

(2.97), up to the topological term. The first step is to notice that the following identity holds, up to terms bilinear in the bulk fermions,

$$
\begin{aligned}
&\int \mathrm{d}^3x\sqrt{\gamma}\left(D_i\overline{Z}_\alpha D^i Z^\alpha - i\overline{\zeta}_{\dot{A}}\not{D}\zeta^{\dot{A}} + \overline{\zeta}_{\dot{A}}Y^{\dot{A}}_{\dot{B}}\zeta^{\dot{B}} + \overline{Z}_\alpha\left(-\phi_3^2 - \{\overline{\sigma},\sigma\} + \frac{1}{4}R\right)Z^\alpha\right)\\
&=\left\{Q, \int \mathrm{d}^3x\sqrt{\gamma}\left(\left(\frac{1}{2}\overline{F}-\overline{f}\right)\zeta_v + \overline{\zeta}_v\left(\frac{1}{2}F-f\right) + \overline{Z}\overline{\sigma}\zeta_u - \overline{\zeta}_u\overline{\sigma}Z\right)\right\}\\
&+\int \mathrm{d}^3x\sqrt{\gamma}\,\mathrm{tr}\left(\phi_3 D_i\mu^i\right) - \int \mathrm{d}^3x\,\mathrm{tr}\left(F\wedge\mu\right).
\end{aligned}
\tag{2.271}
$$

In the first line R is the scalar curvature of the three-dimensional metric γ_{ij}, which appears in this equation from the Lichnerowicz identity.

We can apply this formula to the action (2.267) of the theory, after adding appropriate curvature couplings. We see that there are several unwanted terms, which are not Q-exact. They come from the last line in the identity (2.271), from the boundary terms in the Yang-Mills action (2.244), and, finally, there is a cubic XXX term in (2.267). Using the Dirichlet boundary condition (2.105), we see that most of these terms cancel. What is left is the $\mathrm{tr}(\overline{\sigma}D_3\sigma|^{\pm})$ term from the super Yang-Mills action (2.244), but this term is Q-exact (after adding appropriate fermion bilinear), as we noted in (2.248). So the only non-trivial term in the action of the magnetic theory is the topological term. This is, of course, what one would expect, since in the electric theory we are integrating the fourth descendant of the scalar BRST-closed observable $\mathrm{tr}\sigma^2$. In the S-dual picture this should map to the fourth descendant of the analogous scalar operator, which gives precisely the topological term.

Let us comment on the role of the discontinuity equations (2.265) in the localization computations. In fact, only the first condition in (2.265) should be explicitly imposed on the solutions of the localization equations. Indeed, one can show with some algebra that the last two conditions in that formula follow from the first one automatically, if the localization equations $\{Q,\lambda\}=0$ for every fermion are satisfied.

2.9 Appendix D: Local Observables

In a topological theory of cohomological type (see [58] for an introduction), there generally are interesting local observables. In fact, typically there are Q-invariant zero-form observables (local operators that are inserted at points) and also p-form observables which must be integrated over p-cycles to achieve Q invariance. They are derived from the local observables by a "descent" procedure.

We will describe here the local observables in our problem and the descent procedure. In the magnetic description, everything proceeds in a rather standard way, so we have little to say. The action of electric-magnetic duality on local observables is also straightforward. The zero-form operators of the electric theory are gauge-invariant polynomials in σ, as we discuss below, and duality maps them to the corresponding gauge-invariant polynomials in $\sigma^\vee$; the duality mapping of k-form observables is then determined by applying the descent procedure on both sides of the duality. We focus here on the peculiarities of the electric description that reflect the fact that there are two different gauge groups on the two sides of a defect.

First we recall what happens in bulk, away from the defect. The theory has a complex adjoint-valued scalar σ (defined in eqn. (2.25)) that has ghost number 2 (that is, charge 2 under $\mathrm{U}(1)_F$). This ensures that $\{Q,\sigma\}=0$, as super Yang-Mills theory has no field of dimension 3/2 and ghost number 3 (the elementary fermions have ghost number ± 1). The gauge-invariant and Q-invariant local operators are simply the gauge-invariant polynomials in σ. For a semisimple Lie group of rank r, it is a polynomial ring with r generators. To be concrete, we consider gauge group $\mathrm{U}(n)$, in which the generators are $\mathrm{O}_k=\frac{1}{k}\mathrm{tr}\,\sigma^k$, $k=1,\dots,n$. These are the basic Q-invariant local observables.

In a topological field theory, one would expect that it does not matter at what point in spacetime the operator O_k is inserted. This follows from the identity

$$\mathrm{dO}_p = \left\{ Q, \frac{1}{2}\mathrm{tr}\,\left(\sigma^{p-1}(t^{-1}\widetilde{\psi} + \psi)\right) \right\}, \tag{2.272}$$

where $\mathrm{d} = \sum \mathrm{d}x^\mu \partial_\mu$ is the exterior derivative, and ψ and $\widetilde{\psi}$ are fermionic one-forms. (See Appendix 2.7.2 for a list of fields of the bulk theory and their Q-transformations.) This identity, which says that the derivative of O_k is Q-exact, is actually the first in a hierarchy [7]. If we rename O_k as $\mathrm{O}_k^{(0)}$ to emphasize the fact that it is a zero-form valued operator, then for each k, there is a hierarchy of s-form valued operators $\mathrm{O}_k^{(s)}$, $s = 0, \ldots, 4$, obeying

$$\mathrm{dO}_k^{(s)} = [Q, \mathrm{O}_k^{(s+1)}\}. \tag{2.273}$$

Construction of this hierarchy is sometimes called the descent procedure. This formula can be read in two ways. If Σ_s is a closed, oriented s-manifold in W, then $I_{k,\Sigma_s} = \int_{\Sigma_s} \mathrm{O}_k^{(s)}$ is a Q-invariant observable, since

$$\left[Q, \int_{\Sigma_s} \mathrm{O}_k^{(s)} \right\} = \int_{\Sigma_s} \mathrm{dO}_k^{(s-1)} = 0. \tag{2.274}$$

And I_{k,Σ_s} only depends, modulo $[Q, \ldots\}$, on the homology class of Σ_s, since if Σ_s is the boundary of some Σ_{s+1}, then

$$\int_{\Sigma_s} \mathrm{O}_k^{(s)} = \int_{\Sigma_{s+1}} \mathrm{dO}_k^{(s)} = \left[Q, \int_{\Sigma_{s+1}} \mathrm{O}_k^{(s+1)} \right\}. \tag{2.275}$$

For $s = 0$, Σ_s is just a point p , and $\int_p \mathrm{O}_k^{(0)}$ is just the evaluation of $\mathrm{O}_k = \mathrm{O}_k^{(0)}$ at p; the statement that $\int_{\Sigma_s} \mathrm{O}_k^{(s)}$ only depends on the homology class of Σ_s means that it is independent of p, as we explained already above via eqn. (2.272).

In the magnetic description, we simply carry out this procedure as just described. However, in the electric theory, it is not immediately obvious how much of this standard picture survives when a four-manifold M is divided into two halves M_ℓ and M_r by a defect W. Starting with zero-forms, to begin with we can define separate observables $\mathrm{O}_{k,\ell} = \frac{1}{k}\mathrm{tr}_\ell\,\sigma^k$ and $\mathrm{O}_{k,r} = \frac{1}{k}\mathrm{tr}_r\,\sigma^k$ in M_ℓ and M_r respectively. $\mathrm{O}_{k,\ell}$ is constant mod $\{Q,\dots\}$ in M_ℓ, and similarly $\mathrm{O}_{k,r}$ is constant mod $\{Q,\dots\}$ in M_r. But is there any relation between these two observables? Such a relation follows from boundary condition (2.15), which tells us that on the boundary

$$\sigma = \frac{i}{2}\frac{1}{1+t^2}\{C,C\}. \tag{2.276}$$

(This concise formula, when restricted to the Lie algebras of G_ℓ or of G_r, expresses the boundary value of σ on M_ℓ or on M_r in terms of the same boundary field C.) Hence the invariance of the supertrace implies that $\mathrm{Str}\,\sigma^k = 0$ along W, or in other words that

$$\mathrm{tr}_\ell\,\sigma^k = \mathrm{tr}_r\sigma^k \tag{2.277}$$

when restricted to the boundary between M_ℓ and M_r, where it makes sense to compare these two operators.

Now let us reconsider the descent procedure in this context. We will try to construct an observable by integration on a closed one-cycle $\Sigma_1 = \Sigma_{1\ell}\cup\Sigma_{1r}$, which lies partly in M_ℓ and partly in M_r,

$$\int_{\Sigma_1}\mathrm{O}^{(1)}_k \equiv \int_{\Sigma_{1\ell}}\mathrm{O}^{(1)}_{k,\ell} + \int_{\Sigma_{1r}}\mathrm{O}^{(1)}_{k,r}. \tag{2.278}$$

Given that $\mathrm{O}^{(0)}_{k,\ell} = \mathrm{O}^{(0)}_{k,r}$ along $M_\ell\cap M_r = W$, and in particular on $C_0 = \Sigma_1\cap W$, our observable is Q-closed,

$$\left[Q,\int_{\Sigma_1}\mathrm{O}^{(1)}_k\right\} = \int_{C_0}\left(\mathrm{O}^{(0)}_{k,r} - \mathrm{O}^{(0)}_{k,\ell}\right) = 0. \tag{2.279}$$

The relative minus sign comes in here, because $\Sigma_{1\ell}$ and Σ_{1r} end on C_1 with opposite orientations.

Next we would like to go one step further and define an analogous 2-observable. To check Q-invariance of such an observable, analogously to the case just considered, we would need a relation between $\mathrm{O}^{(1)}_{k,\ell}$ and $\mathrm{O}^{(1)}_{k,r}$. From the relations $\mathrm{dO}^{(0)}_{k,\ell} = [Q, \mathrm{O}^{(1)}_{k,\ell}\}$ in M_ℓ, $\mathrm{dO}^{(0)}_{k,r} = [Q, \mathrm{O}^{(1)}_{k,r}\}$ in M_r, it follows that, if $\imath : W \hookrightarrow M$ is the natural embedding, then

$$\left[Q, \imath^*(\mathrm{O}^{(1)}_{k,\ell} - \mathrm{O}^{(1)}_{k,r})\right\} = 0. \tag{2.280}$$

In topological theory, a Q-closed unintegrated one-form should be Q-exact, so there should exist some operator $\widetilde{\mathrm{O}}^{(1)}_k$, such that

$$\imath^*(\mathrm{O}^{(1)}_{k,\ell} - \mathrm{O}^{(1)}_{k,r}) = \left[Q, \widetilde{\mathrm{O}}^{(1)}_k\right\}, . \tag{2.281}$$

Then for a closed 2-cycle $\Sigma_2 = \Sigma_{2\ell} \cup \Sigma_{2r}$ that intersects W along some C_1 we can define an observable

$$\int_{\Sigma_2} \mathrm{O}^{(2)}_k \equiv \int_{\Sigma_{2\ell}} \mathrm{O}^{(2)}_{k,\ell} + \int_{\Sigma_{2r}} \mathrm{O}^{(2)}_{k,r} + \int_{C_1} \widetilde{\mathrm{O}}^{(1)}_k. \tag{2.282}$$

This observable is Q-closed.

Let us see how to define the next descendant. From the definition of $\mathrm{O}^{(2)}$ and from (2.281) we have

$$\left[Q, \imath^*(\mathrm{O}^{(2)}_{k,\ell} - \mathrm{O}^{(2)}_{k,r})\right\} = \left[Q, \mathrm{d}\widetilde{O}^{(1)}_k\right\}\,, \tag{2.283}$$

therefore, there exists $\widetilde{\mathrm{O}}^{(2)}_k$ such that

$$\imath^*(\mathrm{O}^{(2)}_{k,\ell} - \mathrm{O}^{(2)}_{k,r}) = \mathrm{d}\widetilde{O}^{(1)}_k + \left[Q, \widetilde{\mathrm{O}}^{(2)}_k\right\}. \tag{2.284}$$

Continuing in the same way, we find $\widetilde{\mathrm{O}}^{(n)}_k$ such that

$$\imath^*(\mathrm{O}^{(n)}_{k,\ell} - \mathrm{O}^{(n)}_{k,r}) = \mathrm{d}\widetilde{O}^{(n-1)}_k + \left[Q, \widetilde{\mathrm{O}}^{(n)}_k\right\}\,, \tag{2.285}$$

and the Q-invariant (n)-observable can be defined as

$$\int_{\Sigma_n} \mathrm{O}_k^{(n)} \equiv \int_{\Sigma_{n\ell}} \mathrm{O}_{k,\ell}^{(n)} + \int_{\Sigma_{nr}} \mathrm{O}_{k,r}^{(n)} + \int_{C_{n-1}} \widetilde{\mathrm{O}}_k^{(n-1)}. \tag{2.286}$$

Let us find explicit representatives for all these operators in our case. A formula for $\mathrm{O}_k^{(1)}$ was already given in the right hand side of (2.272):

$$\mathrm{O}_{k,\ell,r}^{(1)} = \mathrm{tr}_{\ell,r}\left(\sigma^{k-1}\psi_t\right) , \tag{2.287}$$

where we now defined

$$\psi_t = \frac{1}{2}(t^{-1}\widetilde{\psi} + \psi). \tag{2.288}$$

This field has useful properties

$$\{Q, \psi_t\} = \mathcal{D}_b\sigma\,, \quad [Q, \mathcal{F}_b] = i(1+t^2)\mathcal{D}_b\psi_t\,, \tag{2.289}$$

and satisfies the boundary condition

$$\imath^*(\psi_t) = -\frac{i}{1+t^2}\{\mathcal{A}_f, C\}. \tag{2.290}$$

Therefore on the defect

$$\imath^*\left(\mathrm{O}_{k,\ell}^{(1)} - \mathrm{O}_{k,r}^{(1)}\right) \sim \mathrm{Str}\left(\{C, C\}^{k-1}\{C, \mathcal{A}_f\}\right) = 0. \tag{2.291}$$

Since this is zero, $\widetilde{\mathrm{O}}_k^{(1)}$ vanishes, and the 2-observable can be defined without a boundary contribution. A representative for the 2-observable is

$$\mathrm{O}_{k,\ell,r}^{(2)} = \mathrm{tr}_{\ell,r}\left(\frac{1}{2}\sum_{k-2}\sigma^{\mathrm{j}_1}\,\psi_t \wedge \sigma^{\mathrm{j}_2}\,\psi_t - \frac{i}{1+t^2}\,\sigma^{k-1}\mathcal{F}_b\right) , \tag{2.292}$$

where $\mathcal{F}_b$ is the field strength for the complexified gauge field $\mathcal{A}_b$. Here and in what follows we use the notation $\sum_m$ for a sum where the set of indices $\mathrm{j}_1, \mathrm{j}_2, \ldots$ runs over partitions of m.

Using the boundary condition (2.290) and invariance of the supertrace, one finds on the boundary,

$$\imath^* \left(\mathrm{O}^{(2)}_{k,\ell} - \mathrm{O}^{(2)}_{k,r} \right) = \frac{i}{1+t^2} \,\mathrm{Str}\left(\sigma^{k-1} \mathcal{F}' \right) , \tag{2.293}$$

where $\mathcal{F}' = \mathcal{F}_b + \mathcal{A}_f \wedge \mathcal{A}_f$ is the part of the super gauge field strength that lies in the bosonic subalgebra. The expression under the supertrace is non-zero, but we know that it should be Q-exact. Indeed, one finds that this is a Q-variation of

$$\widetilde{\mathcal{O}}^{(2)}_k = \frac{1}{2} \left(\frac{i}{1+t^2} \right)^k \mathrm{Str}\left(C^{2k-3} \mathcal{D}_b \mathcal{A}_f \right) . \tag{2.294}$$

Proceeding further with the descent procedure, we can find the 3-descendant,

$$\mathcal{O}^{(3)}_{k,\ell,r} = \mathrm{tr}_{\ell,r} \left(\frac{1}{3} \sum_{k-3} \sigma^{\mathrm{j}_1} \psi_t \wedge \sigma^{\mathrm{j}_2} \psi_t \wedge \sigma^{\mathrm{j}_3} \psi_t - \frac{i}{1+t^2} \sum_{k-2} \sigma^{\mathrm{j}_1} \mathcal{F}_b \wedge \sigma^{\mathrm{j}_2} \psi_t \right) . \tag{2.295}$$

On the boundary after some computation we find

$$\widetilde{\mathcal{O}}^{(3)}_k = \frac{1}{2} \left(\frac{i}{1+t^2} \right)^k \mathrm{Str} \left(\sum_{2k-4} C^{\mathrm{j}_1} \mathcal{A}_f C^{\mathrm{j}_2} \mathcal{D}_b \mathcal{A}_f \right) . \tag{2.296}$$

The bulk part of the four-observable has a representative

$$\begin{aligned} \mathrm{O}^{(4)}_{k,\ell,r} \;&=\; \mathrm{tr}_{\ell,r} \left(\frac{1}{4} \sum_{k-4} \sigma^{\mathrm{j}_1} \psi_t \wedge \sigma^{\mathrm{j}_2} \psi_t \wedge \sigma^{\mathrm{j}_3} \psi_t \wedge \sigma^{\mathrm{j}_4} \psi_t - \frac{i}{1+t^2} \sum_{k-3} \sigma^{\mathrm{j}_1} \mathcal{F}_b \wedge \sigma^{\mathrm{j}_2} \psi_t \wedge \sigma^{\mathrm{j}_3} \psi_t \right. \\ &- \left. \frac{1}{2(1+t^2)^2} \sum_{k-2} \sigma^{\mathrm{j}_1} \mathcal{F}_b \wedge \sigma^{\mathrm{j}_2} \mathcal{F}_b \right) . \end{aligned} \tag{2.297}$$

The four-observable, which is formed from (2.296) and (2.297), has ghost number zero for $k = 2$. In this case, of course, it reduces just to our super Chern-Simons action.

One might wonder how unique this procedure is. Clearly, for the n^{th} descendant of $\mathrm{O}^{(0)}_k$, we can try to modify it by adding a suitable $(n-1)$-observable with ghost number $(2k-n)$, integrated over $C_{n-1} = \Sigma_n \cap W$. Since C_{n-1} is a boundary in the bulk (it is the boundary of $\Sigma_n \mathcal{M}_\ell$, for example), such a modification would be non-trivial only if the observable that we add cannot be extended into the bulk. One possible example is adding a Wilson loop to $\widetilde{\mathrm{O}}_1^{(1)}$ in the 2-descendant of the operator $\operatorname{tr} \sigma$. What other boundary observables might one consider? If we denote the bosonic subgroup of the supergroup by $SG_{\bar{0}} \cong G_\ell \times G_r$, the Q-invariant scalar observables on the defect correspond to the $SG_{\bar{0}}$-invariant polynomials of the ghost field C. However, one can check that for the basic classical Lie superalgebras all such polynomials come[54] from the invariant polynomials in $\sigma \sim \{C, C\}$, and therefore the corresponding observables are bulk observables.

[54]See, e.g., a list of these polynomials in [84].

Chapter 3

Analytic Torsion, 3d Mirror Symmetry and Supergroup Chern-Simons Theories

3.1 Introduction

In this chapter, we study the topological quantum field theory that computes the Reidemeister-Milnor-Turaev torsion [88], [89] in three dimensions. This is a Gaussian theory of a number of bosonic and fermionic fields in a background flat complex GL(1) gauge field. It can be obtained by topological twisting from a free hypermultiplet with $\mathcal{N} = 4$ supersymmetry. This theory is very simple and can be given different names – the one-loop Chern-Simons path-integral [90], or the Rozansky-Witten model [91] with target space $\mathbb{C}^2$, or the U(1|1) supergroup Chern-Simons theory [61] at level equal to one, but we prefer to call it $\mathfrak{psl}(1|1)$ supergroup Chern-Simons theory.

Let us give a brief summary of the chapter. In section 3.2, we describe the definition of the theory. We explain that its functional integral computes a ratio of determinants, which depends holomorphically on a background flat GL(1) bundle $\mathcal{L}$. We also define various line operators, the most important of which lead to the Alexander polynomial for knots and links.

In section 3.3, we use mirror symmetry in three dimensions to represent the $\mathfrak{psl}(1|1)$

theory as the endpoint of an RG flow, that starts from the twisted version of the $\mathcal{N} = 4$ QED with one fundamental flavor. The computation of the partition function of the QED can be localized on the set of solutions to the three-dimensional version of the Seiberg-Witten equations [11]. This provides a physicist's derivation of the relation between the Reidemeister-Turaev torsion and the Seiberg-Witten invariants, which is known as the Meng-Taubes theorem [20], [92]. We consider, in particular, the subtle case of three-manifolds with first Betti number $b_1 \leq 1$ and show, how the quantum field theory manages to reproduce the details of the Meng-Taubes theorem in this case. Previously, the same RG flow has been used in [93] to derive a special case of the Meng-Taubes theorem for the trivial background bundle, when the torsion degenerates to the Casson-Walker invariant. (We elaborate a little more on this in the end of section 3.3.) In comparison to [93], the new ingredient in our thesis is the coupling of the QED to the background flat bundle $\mathcal{L}$, so let us explain, how this works. In flat space and before twisting, the QED has a triplet of FI terms ϕ^a, which transform as a vector under the $\mathrm{SU}(2)_X$-subgroup of the $\mathrm{SU}(2)_X \times \mathrm{SU}(2)_Y$ R-symmetry. (In our notations, the scalars of the vector multiplet of the QED transform in the vector representation of $\mathrm{SU}(2)_Y$.) These FI terms can be thought of as a vev of the scalars of a background twisted vector multiplet. The vector field B_i of the same multiplet can be coupled in a supersymmetric way to the current of the topological U(1)-symmetry of the QED. Upon twisting the theory by $\mathrm{SU}(2)_X$, the scalar and the vector fields of the twisted vector multiplet combine into a complex gauge field $B+i\phi$. Invariance under the topological supercharge Q requires this background field to be flat. One can easily see that the partition function depends on it holomorphically. In the $\mathfrak{psl}(1|1)$ theory, which emerges in the IR, the field $B + i\phi$ gives rise to the complex flat connection that is used in the definition of the Reidemeister-Turaev torsion.

In section 3.4, we consider the $\mathrm{U}(1|1)$ supergroup Chern-Simons theory. It is obtained from the $\mathfrak{psl}(1|1)$ theory by coupling it to $\mathrm{U}(1)_k \times \mathrm{U}(1)_{-k}$ Chern-Simons gauge fields. It has been argued previously [59], [61] that this theory computes the torsion that we study. We show that, in fact, the U(1|1) theory for the compact form of the gauge group is a $\mathbb{Z}_k$-orbifold of the $\mathfrak{psl}(1|1)$ theory, and thus, indeed, computes essentially the same invariant. To be more precise, there exist different versions of the U(1|1) theory, which differ by the global form of the gauge group, but they all are related to the $\mathfrak{psl}(1|1)$ theory. Mirror symmetry maps the U(1|1) Chern-Simons theory at level k to an orbifold of the same twisted $\mathcal{N} = 4$ QED, or equivalently, to an $\mathcal{N} = 4$ QED with one electron of charge k.

In section 3.5, we present the Hamiltonian quantization of the theory. This section does not depend on the results of section 3.3, and can be read separately. By considering braiding transformations of the states on a punctured sphere, we recover the skein relations for the multivariable Alexander polynomial. We consider in some detail the Hilbert space of the $\mathfrak{psl}(1|1)$ theory on a torus, and the correspondence between the states and the loop operators. We find the OPEs of line operators and the action of the modular group. In fact, as long as the background bundle $\mathcal{L}$ has non-trivial holonomies along the cycles of the Riemann surface, on which the theory is quantized, the Hilbert space is one-dimensional, and our analysis is very straightforward. We also discuss the canonical quantization of the U(1|1) Chern-Simons theory. We consider modular transformations of the states on the torus, and find results very similar to those obtained from the GL(1|1) WZW model [60]. To our knowledge, this is the first example of the canonical quantization of a supergroup Chern-Simons theory, that does not assume an *a priori* relation to the WZW model.

In section 3.6, we discuss possible generalizations to other supergroup Chern-Simons theories. We make a summary of properties of such theories. (Some of these were briefly discussed in Chapter 2.) We also present some brane constructions, and consider possible dualities.

Besides the papers that we have already mentioned, previous work on the topological field theory interpretation of the Meng-Taubes theorem includes [94], where the subject was approached from the four-dimensional Donaldson theory, and [95], where a mathematically rigorous proof of the Meng-Taubes theorem using TQFT was presented. All the mathematical facts about the Reidemeister-Turaev torsion, the Seiberg-Witten invariants and the Meng-Taubes theory, that we touch upon in this chapter, can be found in a comprehensive review [88].

Finally, let us mention that there exists yet another approach [96] to the Reidemeister-Turaev torsion, which presumably can be given a physical interpretation, – in this case, in terms of the first-quantized theory of Seiberg-Witten monopoles. Unfortunately, this will not be considered in the present thesis.

3.2 Electric Theory

In this section, we describe the theory, which computes an analytic analog of the Reidemeister-Turaev torsion. Up to some details, it is simply the theory of the degenerate quadratic functional [3]. One important difference, however, is that we introduce a coupling to a *complex* background flat bundle, and consider the torsion as a holomorphic function of it. Our definition is similar but not quite identical to the definition of the analytic torsion, known in the mathematical literature [97]. The discussion will be phrased in the language of supergroup Chern-Simons theory. Though this might seem like an unnecessary over-complication, it will make our formulas a little more compact, and will also help, when we discuss generalizations in later sections. Throughout the chapter, the theory of this section will be called "electric", while its mirror, considered in section 3.3, will be called "magnetic".

3.2.1 The Simplest Supergroup Chern-Simons Theory

In this section we introduce the $\mathfrak{psl}(1|1)$ Chern-Simons theory. We work on a closed oriented three-manifold W.

The superalgebra $\mathfrak{g} \simeq \mathfrak{psl}(1|1)$ is simply the supercommutative Grassmann algebra $\mathbb{C}^{0|2}$. The Chern-Simons gauge field will be a $\mathbb{C}^2$-valued fermionic one-form $A = A^I \widehat{f}_I$, where $\widehat{f}_+$ and $\widehat{f}_-$ are the superalgebra generators. To make the theory interesting, we want to couple it to a background flat bundle. It could possibly be a GL(2)-bundle, where GL(2) acts on $\mathfrak{g}$ in the obvious way. However, the definition of the Chern-Simons action requires a choice of an invariant bilinear form. This reduces the symmetry to SL(2), so we couple the theory to a flat SL(2)-bundle $\mathcal{B}$. The Chern-Simons action can be written as[1]

$$I_{\mathfrak{psl}(1|1)} = \frac{i}{4\pi}\int_W \mathrm{Str}\, A\, \mathrm{d}_{\mathcal{B}} A\,, \tag{3.1}$$

where the supertrace denotes an invariant two-form, $\mathrm{Str}(ab) = \epsilon_{IJ} a^I b^J$, and $\mathrm{d}_{\mathcal{B}}$ is the covariant differential, acting on the forms valued in $\mathcal{B}$. One could eliminate the flat gauge field from $\mathrm{d}_{\mathcal{B}}$ by a suitable choice of trivialization of $\mathcal{B}$, but we prefer not to do so.

The supergroup gauge transformations act by $A \to A - \mathrm{d}_{\mathcal{B}}\alpha$. To fix the gauge, we introduce a $\mathfrak{g}$-valued ghost field C. Since our gauge symmetry is fermionic, this field has to be bosonic: its two components are complex scalars C^+ and C^-. We also introduce a bosonic $\mathfrak{g}$-valued antighost field $\overline{C}$ and a $\mathfrak{g}$-valued fermionic Lagrange multiplier λ. The BRST generator Q is defined to act as

$$\delta A = -\mathrm{d}_{\mathcal{B}} C\,, \quad \delta C = 0\,, \quad \delta\lambda = 0\,, \quad \delta\overline{C} = \lambda\,. \tag{3.2}$$

Next we have to choose an appropriate gauge-fixing action. It will contain in particular the kinetic term for the bosonic fields C and $\overline{C}$, and we want to make sure that this term is positive-definite. To that end, we pick a hermitian structure on our flat bundle and restrict to unitary gauges. We impose a reality condition $\overline{C}^I = -\epsilon^{IJ}(C^J)^\dagger$. The complex flat connection in $\mathcal{B}$ can be decomposed as $B + i\phi$, where B is a hermitian connection and ϕ is a section of the adjoint bundle. We introduce a covariant derivative $D_i = \partial_i + iB_i$, and also introduce notations $\mathcal{D}_i = D_i - \phi_i$ for the covariant derivative in the flat bundle $\mathcal{B}$ and $\overline{\mathcal{D}}_i = D_i + \phi_i$ for the covariant derivative with the conjugate gauge field. We pick a metric γ on W and take

[1] Throughout the chapter we use Euclidean conventions, in which the functional under the path-integral is $\exp(-I)$.

the gauge-fixing action to be

$$I_{\text{g.f.}} = \left\{ Q, \int \mathrm{d}^3x \sqrt{\gamma}\gamma^{ij}\, \mathrm{Str}\left(\overline{\mathcal{D}}_i \overline{C} A_j\right) \right\} = -\int \mathrm{d}^3x \sqrt{\gamma}\gamma^{ij}\, \mathrm{Str}\left(\overline{\mathcal{D}}_i \overline{C} \mathcal{D}_j C - A_i \overline{\mathcal{D}}_j \lambda\right) . \quad (3.3)$$

The bosonic part of this action is manifestly positive-definite. The gauge-fixing condition is $\overline{\mathcal{D}}_i A^i = 0$. The action has a ghost number symmetry $\mathrm{U}(1)_{\mathrm{F}}$, under which the ghost and the antighost fields have charges ± 1. If the background field satisfies $[\overline{\mathcal{D}}^i, \mathcal{D}_i] = 0$, or equivalently $D^i \phi_i = 0$, this symmetry is enhanced to $\mathrm{SU}(2)$, which rotates C and $\overline{C}$ as a doublet and which we will call $\mathrm{SU}(2)_Y$. If we turn off the background gauge field completely, we also recover the "flavor" $\mathrm{SU}(2)_{\mathrm{fl}}$ symmetry, which is the unitary subgroup of the $\mathrm{SL}(2)$ automorphism group of the superalgebra. The groups $\mathrm{SU}(2)_Y$ and $\mathrm{SU}(2)_{\mathrm{fl}}$ commute. Together they generate an action of $\mathrm{SO}(4)$ on the real four-dimensional space parameterized by C and $\overline{C}$.

In this chapter, we will not consider the general $\mathrm{SL}(2)$ analytic torsion[2]. From now on, we restrict our attention to the case that the background flat bundle is abelian, $\mathcal{B} = \mathcal{L} \oplus \mathcal{L}^{-1}$, where[3] $\mathcal{L} \in \mathrm{Hom}(H_1(W), \mathbb{C}^*)$. By abuse of notation, we will denote the connection in $\mathcal{L}$ by the same letters $B + i\phi$, where now B is understood to be a connection in a flat unitary line bundle, and ϕ is a closed one-form, whose cohomology class determines the absolute values of the holonomies in $\mathcal{L}$.

The abelian background field preserves a $\mathrm{U}(1)_{\mathrm{fl}}$-subgroup of the flavor symmetry group $\mathrm{SU}(2)_{\mathrm{fl}}$. We will furthermore assume that ϕ is chosen to be the harmonic representative in its cohomology class, so that the $\mathrm{SU}(2)_Y$-symmetry is present.

[2]The reason is that the Meng-Taubes theorem, which will be the subject of section 3.3, does not seem to generalize to $\mathrm{SL}(2)$ torsion, since only the abelian part of the symmetry is visible in the UV. However, what could be generalized to the $\mathrm{SL}(2)$ torsion (and, in fact, to $\mathrm{Sp}(2n, \mathbb{C})$ torsion) is the Hamiltonian quantization that we consider in section 3.5. This generalization will be discussed elsewhere.

[3]Throughout the chapter, the coefficients in homology and cohomology are assumed to be $\mathbb{Z}$, unless explicitly specified otherwise.

3.2.2 Relation To A Free Hypermultiplet

Our theory can be obtained by making a topological twist of the theory of a free $\mathcal{N} = 4$ hypermultiplet. This is a trivial special case of the general relation between supergroup Chern-Simons and $\mathcal{N} = 4$ Chern-Simons-matter theories, found in [24]. For completeness, we provide some details.

The R-symmetry group of $\mathcal{N} = 4$ supersymmetry in three dimensions is $\mathrm{SU}(2)_X \times \mathrm{SU}(2)_Y$. The supercharges transform in the $(\mathbf{2}, \mathbf{2}, \mathbf{2})$-representation of $\mathrm{SU}(2)_{\mathrm{Lorentz}} \times \mathrm{SU}(2)_X \times \mathrm{SU}(2)_Y$. A supersymmetric theory can be twisted by taking the Lorentz spin-connection to act by elements of the diagonal subgroup of $\mathrm{SU}(2)_{\mathrm{Lorentz}} \times \mathrm{SU}(2)_X$. This leaves an $\mathrm{SU}(2)_Y$ doublet of invariant supercharges. We pick one of them, to be called Q, and use it to define a cohomological topological theory. The ghost number symmetry $\mathrm{U}(1)_{\mathrm{F}}$ is the subgroup of $\mathrm{SU}(2)_Y$, for which Q is an eigenvector.

The scalars of the free hyper give rise to the ghost fields C and $\overline{C}$. They parameterize a copy of the quaternionic line $\mathbb{H}$, which has a natural action of two commuting $\mathrm{SU}(2)$ groups. One of them is identified with the R-symmetry group $\mathrm{SU}(2)_Y$, and the other is the flavor symmetry $\mathrm{SU}(2)_{\mathrm{fl}}$. The hypermultiplet fermions, which transform in the $(\mathbf{2}, \mathbf{2}, \mathbf{1})$ representation of the Lorentz and R-symmetry groups, upon twisting give rise to a vector field and a scalar, which we identify with the fermionic gauge field A_i and the Lagrange multiplier field λ.

Finally, the imaginary part of the flat connection ϕ_i originates from the $\mathrm{SU}(2)_X$-triplet of hypermultiplet masses. While they are constant parameters in the untwisted theory, they are promoted to a closed one-form in the topological theory, still preserving the Q-invariance. Different terms in the action (3.1), (3.3) can be easily seen to originate from the kinetic and the mass terms for the hypermultiplet scalars and fermions.

3.2.3 A Closer Look At The Analytic Torsion

Here we would like to take a closer look at the invariant that our theory computes. We discuss its properties and relation to other known definitions of the torsion. For simplicity, the manifold W is assumed to be closed, unless indicated otherwise.

3.2.3.1 Definition And Properties

The partition function of the theory can be written as a ratio of determinants,

$$\tau(\mathcal{L}) = \frac{\det L_-}{\det^2 \Delta_0}\,. \tag{3.4}$$

Here the operator $L_- = \star(d_B - \phi) + (d_B + \phi)\star$ is acting in $\Omega^1_{\mathcal{L}}(W) \oplus \Omega^3_{\mathcal{L}}(W)$, where $\Omega^p_{\mathcal{L}}(W)$ is the space of p-forms valued in $\mathcal{L}$. The twisted Laplacian $\Delta_0 = -D_i D^i + \phi_i \phi^i$ is acting in $\Omega^0_{\mathcal{L}}(W)$. Note that the operator Δ_0 is hermitian, while L_- is hermitian only when $\phi = 0$.

The ratio $\tau(\mathcal{L})$, by construction, is a holomorphic function of the flat bundle, even though the determinants in (3.4) are not. We can understand the analytic properties of $\tau(\mathcal{L})$ rather explicitly. The absolute value of the torsion can be written in the usual Ray-Singer form as

$$|\tau(\mathcal{L})| = \frac{(\det \Delta_1)^{1/2}}{(\det \Delta_0)^{3/2}}\,, \tag{3.5}$$

where Δ_1 is the twisted Laplacian, acting on one-forms. The numerator in this formula vanishes, whenever the twisted cohomology $H^1(W, \mathcal{L})$ is non-empty. This subspace, possibly with the exception of the trivial flat bundle, is the locus of zeros of $\tau(\mathcal{L})$. The denominator vanishes, when the twisted cohomology $H^0(W, \mathcal{L})$ is non-empty, which is precisely when the flat bundle $\mathcal{L}$ is trivial. At this point the function $\tau(\mathcal{L})$ can potentially have a singularity. In fact, if the first Betti number b_1 of W is greater than one, the singularity would be of codimension at least two, which is not possible for a holomorphic function. For $b_1 = 1$, let the holonomies of $\mathcal{L}$ around the torsion[4] one-cycles be trivial, and let $\mathbf{t}$ be the holonomy around the non-torsion one-cycle. At $\mathbf{t} = 1$, the operators Δ_0 and Δ_1 have one zero mode each. At small $\mathbf{t} - 1$, these eigenfunctions become quasi-zero modes with eigenvalues of order $(\mathbf{t} - 1)^2$, according to the non-degenerate perturbation theory. Plugging this into (3.5), we see that the ratio $\tau(\mathcal{L})$ near the trivial flat bundle is proportional to $1/(\mathbf{t} - 1)^2$, that is, has a second-order pole. Finally, for $b_1 = 0$ the torsion is a function on the discrete set of flat bundles. For the trivial flat bundle and $b_1 = 0$, it is natural to set τ to be equal to infinity[5].

Another important property of the torsion is the relation

$$\tau(\mathcal{L}) = \tau(\mathcal{L}^{-1}) \,, \tag{3.6}$$

which follows from the charge conjugation symmetry $\mathcal{C}$ that maps the superalgebra generators as $\widehat{f}_\pm \to \pm\widehat{f}_\mp$, and the line bundle $\mathcal{L}$ to its dual $\mathcal{L}^{-1}$.

3.2.3.2 Details Of The Definition

We would like to make a more precise statement about what we mean by the formal definition (3.4). Let us assume for now that the flat bundle $\mathcal{L}$ is unitary. If we eliminate the ambiguities in the definition of $\tau(\mathcal{L})$ for such bundles, the definition for complex flat bundles will also be unambiguous, by analyticity.

The absolute value (3.5) of $\tau(\mathcal{L})$ is (the inverse of) the Ray-Singer torsion, which is a well-defined and metric-independent object. However, as is well-known in the context of Chern-Simons theory [4], the definition of the phase of $\tau(\mathcal{L})$ requires more care. With our assumption that $\mathcal{L}$ is unitary, the operator L_- is hermitian and has real eigenvalues.

[4] A cycle is called "torsion" if it lies in the torsion part of $H_1(W)$, that is, if some multiple of it is trivial. This use of the word "torsion" is totally unrelated to "torsion" as an invariant of the manifold. Hopefully, this will not cause confusion.

[5] One could say that for the trivial bundle the path-integral is undefined, since it has both bosonic and fermionic zero modes. But it is natural to set it equal to infinity for $b_1 = 0$, because, thinking in terms of gauge-fixing, the path-integral has a factor of inverse volume of the gauge supergroup, which is infinity, since this volume is zero. Taking $Z(S^3) = \infty$ also makes the factorization formulas of the ordinary Chern-Simons valid in the supergroup case.

Since the determinant of L_- comes from a fermionic path-integral, it is natural to choose a regularization, in which it is real. The only possible ambiguity then is in its sign. Note that this is mainly interesting in the case when there is torsion in $H_1(W)$, so that the space of flat bundles is not connected, and signs can potentially be changed for different connected components.

Let us suggest a way to define the sign of L_-. What we are about to say might not seem particularly natural at first sight, but, as we show later, matches well with known definitions of the analytic and combinatorial torsion. Let us pick a spin structure s on the three-manifold W, and take some oriented spin four-manifold V, of which W with a given spin structure is a boundary. The line bundle $\mathcal{L}$ can be extended onto V, though the extension might not be flat. On V we consider the Donaldson operator $L_4 : \Omega^1_{\mathcal{L}}(V) \to \Omega^0_{\mathcal{L}}(V) \oplus \Omega^{2,-}_{\mathcal{L}}(V)$ that arises from the linearization of the self-duality equations, twisted by the line bundle $\mathcal{L}$. Here $\Omega^{2,-}$ is the bundle of anti-selfdual two-forms. We define the sign of the determinant of L_-, and therefore of the torsion $\tau(\mathcal{L})$, using the index of the elliptic operator L_4,

$$\operatorname{sign} \tau(\mathcal{L}) = (-1)^{\operatorname{ind}(L_4) - \operatorname{ind}(L_{4,\mathrm{triv}})}\,, \tag{3.7}$$

where $L_{4,\mathrm{triv}}$ is the Donaldson operator coupled to the trivial line bundle. The motivation behind this definition is that, if we were to compute the change of sign of $\det L_-$ under a continuous change of $\mathcal{L}$, we could naturally do it by using the formula (3.7) with the four-manifold taken to be the cylinder $W \times I$, since the index of L_4 on such a cylinder computes the spectral flow of L_-.

We started with a choice of a spin structure, but so far it did not explicitly enter the discussion. Its role is the following. For two different choices of the four-manifold, the change in the sign of $\det L_-$ is governed by the index of L_4 on a closed four-manifold V', which, according to the index theorem, is

$$\mathrm{ind}(L_4) - \mathrm{ind}(L_{4,\mathrm{triv}}) = \int_{V'} c_1(\mathcal{L})^2 \,. \tag{3.8}$$

However, since the spin structure on W can be extended to V', the four-manifold V' is spin, and therefore its intersection form is even, and so is the right hand side of (3.8). We conclude that the sign of $\tau(\mathcal{L})$ depends on a spin structure on W, but not on the choice of the four-manifold. (This is equivalent to the well-known fact [98] that a choice of a spin-structure allows to define a half-integral Chern-Simons term for an abelian gauge field.)

It is not hard to calculate the dependence on the spin structure explicitly. Let s_1 and s_2 be two spin structures on W, which differ by some $x \in H^1(W, \mathbb{Z}_2)$. Let V_1 and V_2 be four-manifolds with boundary W, onto which s_1 and s_2 extend. Now the closed four-manifold V', glued from V_1 and V_2 along their boundary W, need not be spin, and its Stiefel-Whitney class $w_2 \in H^2(V', \mathbb{Z}_2)$ can be non-zero. The intersection form is not even, but its odd part is governed by the Wu's formula, which tells us that $\bar{c}_1^2 = \bar{c}_1 \smile w_2$, where $\bar{c}_1$ is the mod 2 reduction of $c_1(\mathcal{L})$. (This is true for any $H^2(V', \mathbb{Z}_2)$ class, of course.) The Stiefel-Whitney class of V' is determined by x. For a given good covering of V', the two spin structures s_1 and s_2 define a lift of the transition functions of the tangent bundle of V' from SO(4) to Spin(4), and this lift is consistent everywhere, except for a codimension-two chain, lying in W. This chain defines the Stiefel-Whitney class of V', but it is also the Poincaré dual of the class x in W. These arguments allow us to write

$$\int_{V'} c_1(\mathcal{L})^2 = \int_{V'} c_1(\mathcal{L}) \smile w_2 = \int_{\mathrm{PD}(w_2)} c_1(\mathcal{L}) = \int_{\mathrm{PD}(x)} c_1(\mathcal{L}) = \int_W c_1(\mathcal{L}) \smile x \mod 2 \,, \tag{3.9}$$

where PD stands for Poincaré dual. We conclude that under a change of the spin structure by x, the sign of $\tau(\mathcal{L})$ changes by the factor

$$(-1)^{\int_W c_1(\mathcal{L}) \smile x} \,. \tag{3.10}$$

It will be useful to rearrange this formula a little. For that we need to recall a couple of topological facts. The topology of a flat line bundle is completely defined by its holonomies around the torsion one-cycles. This is formalized by the following exact sequence,

$$H^1(W) \to H^1(W,\mathbb{R}) \xrightarrow{\alpha} H^1(W,\mathrm{U}(1)) \xrightarrow{\beta} \operatorname{tor} H^2(W) \to 0\,, \tag{3.11}$$

which is associated to the short exact sequence of coefficients $0 \to \mathbb{Z} \to \mathbb{R} \to \mathrm{U}(1) \to 0$. By Pontryagin duality, $H^1(W,\mathrm{U}(1)) \simeq \mathrm{Hom}(H_1(W),\mathrm{U}(1))$ is the abelian group of (unitary) flat line bundles on W. The morphism α gives a flat bundle with trivial holonomy around the torsion cycles and given holonomy around the non-torsion cycles[6]. The morphism β maps a given flat bundle to its first Chern class, which depends only on the holonomies around the torsion cycles, by exactness of the sequence. Pick a pair of classes y_1 and y_2 from $\operatorname{tor} H^2(W)$. Let $\mathcal{L}_1$ be some flat bundle with Chern class y_1. Its holonomies around the torsion cycles are completely defined by y_1. We can take a holonomy of $\mathcal{L}_1$ around the one-cycle, Poincaré-dual to y_2. The logarithm of this number gives a pairing $\operatorname{tor} H^2(W) \times \operatorname{tor} H^2(W) \to \mathbb{Q}/\mathbb{Z}$, which is known as the linking form. An important fact is that it is bilinear and symmetric. (Actually, this pairing is just the $\mathrm{U}(1)\times\mathrm{U}(1)$ Chern-Simons term for flat bundles.)

Returning to the formula (3.10), we note that $x \in H^1(W,\mathbb{Z}_2)$ defines a $\mathbb{Z}_2$-bundle, and (3.10) is the holonomy of this bundle around the one-cycle, Poincaré dual to $c_1(\mathcal{L})$. Since the linking form is symmetric, this holonomy is equal to the holonomy of $\mathcal{L}$ around the one-cycle, Poincaré dual to $c_1(x)$, where, to construct $c_1(x)$, we think of the $\mathbb{Z}_2$-bundle defined by x as of a $\mathrm{U}(1)$-bundle. This holonomy will be denoted by $\mathcal{L}(c_1(x))$. We conclude that it defines the change of the sign of $\tau(\mathcal{L})$, when the spin structure on W is changed by x. To indicate the dependence on the spin structure explicitly, we will sometimes write the torsion as $\tau_s(\mathcal{L})$, so that

$$\tau_{x\cdot s}(\mathcal{L}) = \mathcal{L}(c_1(x))\,\tau_s(\mathcal{L})\,. \tag{3.12}$$

[6] What one means by non-torsion cycles is not canonically defined, but this does not matter, when the holonomies around the torsion cycles are trivial.

It is noteworthy that if the line bundle $\mathcal{L}$ has trivial holonomies around 2-torsion cycles, the definition of $\tau(\mathcal{L})$ is independent of any choices at all.

In fact, even for a generic flat bundle, $\tau_s(\mathcal{L})$ depends on something less than a spin structure. There is a natural map from the set of spin structures to the set of spin-$\mathbb{C}$ structures with trivial determinant, which is given by tensoring with a trivial line bundle,. This map is not an isomorphism, because in general two different spin structures can map to the same spin-$\mathbb{C}$ structure. Since the change of the sign of $\tau_s(\mathcal{L})$ under a change of s by an element x of $H^1(W, \mathbb{Z}_2)$ depends only on the first Chern class of the line bundle obtained from x, the sign of $\tau_s(\mathcal{L})$ really depends only on a spin-$\mathbb{C}$ structure with trivial determinant, and not on the spin structure itself.

One could consider some trivial generalizations of our definition of the torsion. For example, τ_s can be naturally defined for an arbitrary spin-$\mathbb{C}$ structure s, not necessarily with trivial determinant. Let s_0 be some arbitrary spin-$\mathbb{C}$ structure with trivial determinant, s be an arbitrary spin-$\mathbb{C}$ structure, and let $y \in H^2(W)$ be such that $y \cdot s = s_0$. We can set $\tau_s(\mathcal{L}) = \mathcal{L}(y)\tau_{s_0}(\mathcal{L})$. Clearly, (3.12) implies that τ_s depends only on s, and not on the choice of s_0. In quantum field theory terms, this modification amounts to adding to the action a local topologically-invariant functional of the background gauge field – the Wilson loop of $\mathcal{L}$ around the cycle, Poincaré-dual to $y \in H^2(W)$. Another possible modification of the definition would be to add a Chern-Simons term for the background field B. Note that, if we choose the coefficient of this term to be a half-integer, this would eliminate the dependence of τ_s on the spin structure. In what follows, we will mostly restrict to our most basic definition of τ_s, unless indicated otherwise.

3.2.3.3 Comparison To Known Definitions

Let us comment on the relation of our torsion to some known definitions from the mathematical literature. A rigorous definition of the complex analytic torsion was given in [97].

The authors consider essentially[7] the same ratio of determinants (3.4) and use the ζ-function regularization to define it as a holomorphic function of the flat bundle $\mathcal{L}$. An important difference, however, is that for a unitary flat bundle their torsion is not real, but has a phase, proportional to the eta-invariant of L_-. In the language of functional integral, such definition is perhaps more natural [4], when the determinant of L_- comes from a bosonic, rather than a fermionic functional integral. The relation to our definition is given by the APS index theorem: to transform the eta-invariant into the index, one needs to subtract what might be called a half-integral Chern-Simons term of the flat connection in the line bundle $\mathcal{L}$. This is why the dependence on a spin structure appeares in our story, but not in [97].

In fact, there is a combinatorial definition of torsion, which, as we conjecture, is precisely equal to our $\tau_s(\mathcal{L})$. This is the Turaev's refinement of Reidemeister torsion[8]. We briefly summarize some facts about it. For a detailed review, as well as references, the reader can consult [88].

Let W' be a compact three-manifold, which is closed or is a complement of a link neighborhood in a closed three-manifold, so that it has a boundary consisting of a number of tori. (In our language, non-empty boundary will correspond to adding line operators, to which we turn in the next section.) In either case, the Euler characteristic of W' is zero. Reidemeister torsion of W' is defined as the determinant of a particular acyclic complex, twisted by a vector representation of the fundamental group of the manifold. The determinant of this combinatorially defined complex can be viewed as a discretisation of the functional integral, which computes the analytic torsion. We will assume that the representation of the fundamental group is given by the flat line bundle $\mathcal{L}$. Reidemeister torsion is defined only up to

[7] There are some differences. The discussion in [97] is more general: the authors consider a manifold of arbitrary odd dimension, and not necessarily one-dimensional flat vector bundles. Another difference from our approach, if phrased in path-integral language, is that in [97] the gauge-fixing term in the analog of (3.3) is defined using the derivative $\mathcal{D}$, rather than its conjugate. This eliminates the need to pick a hermitian structure on the flat bundle, but makes the functional integral representation of the determinant more formal. Finally, the ratio of determinants in [97] is actually the inverse of ours.

[8] Note that sometimes Reidemeister torsion is defined to be the inverse of what we consider here. With the definition that we use, the absolute value of the combinatorial torsion is equal to the *inverse* of Ray-Singer torsion, defined in the usual way.

a sign and up to multiplication by a holonomy of $\mathcal{L}$ around an arbitrary cycle in W'. This happens because the determinant depends on the basis in the complex, of which there is no canonical choice. Turaev has shown [99] that this ambiguity can be eliminated, once one makes a choice of what he called an Euler structure[9]. In analytical terms, it is a choice of a nowhere vanishing vector field, up to homotopy and up to an arbitrary modification inside a three-ball. Such vector fields always exist on W', since $\chi(W') = 0$. In three dimensions, it is not hard to see that Euler structures are in a canonical one-to-one correspondence with spin-$\mathbb{C}$ structures. For a spin-$\mathbb{C}$ structure s, let us denote the Reidemeister-Turaev combinatorial torsion by $\tau_s^{\mathrm{RT}}(\mathcal{L})$. Under a change of the spin-$\mathbb{C}$ structure by an element $y \in H^2(W')$, the torsion changes as

$$\tau_{y\cdot s}^{\mathrm{RT}}(\mathcal{L}) = \mathcal{L}(y)\, \tau_s^{\mathrm{RT}}(\mathcal{L})\,, \tag{3.13}$$

where, as usual in our notations, $\mathcal{L}(y)$ is the holonomy of $\mathcal{L}$ around the cycle Poincaré dual to y. The combinatorial torsion also has a charge conjugation symmetry $\mathcal{C}$

$$\tau_s^{\mathrm{RT}}(\mathcal{L}^{-1}) = (-1)^{\ell} \tau_{\overline{s}}^{\mathrm{RT}}(\mathcal{L}) = (-1)^{\ell}\, \mathcal{L}^{-1}(c_1(\det s))\tau_s^{\mathrm{RT}}(\mathcal{L})\,, \tag{3.14}$$

where $\overline{s}$ is the conjugate of the spin-$\mathbb{C}$ structure s, and ℓ is the number of connected components of the boundary of W'. The second equality here follows from (3.13).

If the three-manifold W' is closed and the spin-$\mathbb{C}$ structure s has trivial determinant, we claim that τ_s^{RT} coincides with our analytic torsion τ_s. (Modulo signs, that is, ignoring the dependence on the spin structure, this statement would follow from the results of [97] and [100].) For a spin-$\mathbb{C}$ structure with trivial determinant, the properties (3.13) and (3.14) reduce to our formulas (3.12) and (3.6), respectively. When the three-manifold W' is not closed but is a complement of a link, the relation between τ_s^{RT} and our τ_s should still hold, with an appropriate definition of the analytic torsion in presence of line operators. This will

[9] More precisely, the choice of an Euler structure eliminates the freedom to multiply the torsion by a holonomy of $\mathcal{L}$, while the overall sign can be fixed by choosing an orientation in the homology $H_\bullet(W')$. At least for a closed three-manifold, there exists a canonical homology orientation, defined by the Poincaré duality, and we assume that our theory automatically picks this orientation.

be discussed in the next section.

An important special case is when the flat bundle $\mathcal{L}$ has trivial holonomies around the torsion one-cycles. Then $\tau_s^{\mathrm{RT}}(\mathcal{L})$ is a holomorphic function of $b_1(W')$ variables $\mathbf{t}_1, \ldots, \mathbf{t}_{b_1}$. Let us also ignore the dependence on s, so that we consider τ modulo sign and modulo multiplication by powers of $\mathbf{t}_\bullet$. This variant of the combinatorial torsion is known as the Milnor torsion. A theorem due to Milnor [101] and Turaev [102] describes its relation to the Alexander polynomial Δ of W', which is a function of the same variables $\mathbf{t}_1, \ldots \mathbf{t}_{b_1}$. If $b_1(W') > 1$, then $\tau = \Delta$. If $b_1(W') = 1$, then $\tau = \mathbf{t}\Delta/(\mathbf{t}-1)^2$, if W' is a closed three-manifold, and $\tau = \Delta/(\mathbf{t}-1)$, if W' is a complement of a knot in a closed three-manifold. For a closed W', these statements are in agreement with the analytical properties of our τ, described in section 3.2.3.1.

3.2.4 Line Operators

We would like to define some line operators in our theory, in order to study knot invariants. First thing that comes to mind is to use Wilson lines. For these to be invariant under the transformations (3.2), they should be labeled by representations of $\mathfrak{pl}(1|1)$. This superalgebra contains $\mathfrak{psl}(1|1)$ as well as one bosonic generator, which acts on the fermionic generators with charges ± 1. The Wilson lines should be defined with the $\mathfrak{pl}(1|1)$ connection $A + B + i\phi$. In fact, the only irreducible representations of $\mathfrak{pl}(1|1)$ are one-dimensional representations, to be denoted (m), in which the bosonic generator acts with some charge m, and the fermionic generators act trivially. Inserting a Wilson loop in representation (m) along a knot K is equivalent to multiplying the path-integral by the m-th power of the holonomy of the background bundle $\mathcal{L}$ around the cycle K. Though this operator is of a rather trivial sort, it will be convenient to consider it as a line operator. It will be denoted by L_m, $m \in \mathbb{Z}$. According to the remarks at the end of section 3.2.3.2, inserting operators L_m around various cycles is equivalent to changing the spin-$\mathbb{C}$ structure, with which the torsion is defined.

All the other representations of $\mathfrak{pl}(1|1)$ are reducible, but, in general, can be indecompos-

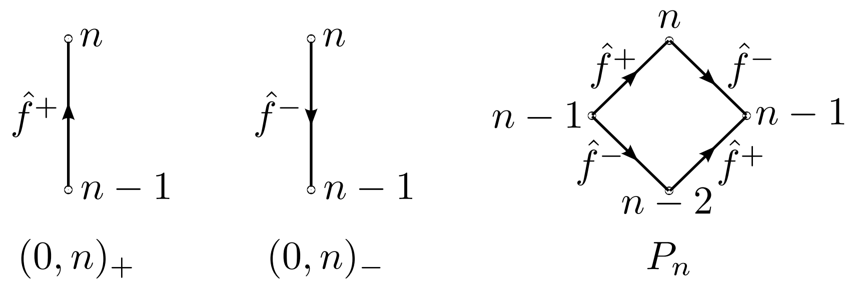

Figure 3.1: Examples of reducible indecomposable representations of $\mathfrak{pl}(1|1)$. The dots are the basis vectors, and the arrows show the action of the fermionic generators $\widehat{f}^{\pm}$. The numbers $n, n-1, \ldots$ are the eigenvalues of the bosonic generator of $\mathfrak{pl}(1|1)$, that is, the $\mathrm{U}(1)_{\mathrm{fl}}$-charges. The representations $(0,n)_-$ and $(0,n)_+$ are known as the (anti-)Kac modules.

able[10]. Some examples are shown on fig. 3.1. (There are more such representations – they are listed *e.g.* in [103], – but we will not need them.) In this chapter, we are mostly interested in closed loop operators. Naively, due to the presence of the supertrace, a closed Wilson loop labeled by a reducible indecomposable representation splits into a sum of Wilson loops for the invariant subspaces and quotients by them. If this were true, the indecomposable representations would not need to be considered separately. We will later find that, due to regularization issues, at least for some indecomposable representations the Wilson loops do not actually reduce to sums of Wilson loops L_m. This will be discussed in section 3.5, but till then we will not consider indecomposable representations.

In the case that the holonomy of the background field along some loop K is trivial, one can construct a line operator by inserting an integral $\oint_K A^{\pm}$ into the path-integral. Note that these operators transform as a doublet under the $\mathrm{SU}(2)_{\mathrm{fl}}$ flavor symmetry. These will play the role of creation/annihilation operators in the Hamiltonian picture in section 3.5, but, again, will not be important till there.

The most useful line operator can be obtained by cutting a knot (or a link) K out of W, and requiring the background gauge field to have a singularity near K with some prescribed holonomy $\mathbf{t}$ around the meridian of the knot complement[11]. It is this type of line operators

[10] That is, they have invariant subspaces, but need not split into direct sums.

[11] The meridian is the cycle that can be represented by a small circle, linking around the knot. A longitude is a cycle that goes parallel to the knot. The longitude, unlike the meridian, is not canonically defined. Its

that will give rise to the Alexander knot polynomial.

One has to be careful in defining the determinants (3.4) in presence of such a singularity. In this chapter, our understanding of the determinants in this case will be much less complete than in the case of closed three-manifolds. We will not attempt to give a rigorous definition, but will simply state some results that are consistent with other approaches to line operators, which are discussed later in the thesis, and with known properties of the Alexander polynomial. Let $\mathbf{t}$ be the holonomy around the meridian of the knot K, and $\mathbf{t}_{\|}$ be the holonomy around the longitude. While $\mathbf{t}$ is a part of the definition of the line operator along K, $\mathbf{t}_{\|}$ depends on the flat connection and, in particular, on other line operators, linked with K. The problem with the determinants (3.4) in presence of line operators is that in general they can be anomalous, that is, they can change sign under large gauge transformations of the background gauge field. Equivalently, one will in general encounter half-integral powers of $\mathbf{t}$ and $\mathbf{t}_{\|}$ in the expectation values. One possible resolution is to choose a square root of the holonomies, or, equivalently, to take $\mathcal{L} \simeq \mathcal{L}'^2$, and to consider the knot polynomial as a function of the holonomies of $\mathcal{L}'$. One expects this to produce a version of the Alexander polynomial known as the Conway function. (See section 4 of [102] for a review.) Alternative approach, which we will assume in most of the chapter, is to add along the longitude of the knot a Wilson line for the background gauge field. So, we will in general consider combined line operators, labeled by two parameters $\mathbf{t}$ and m, with m being the charge for the Wilson line for the background field $B + i\phi$. It will be clear from the discussion of the U$(1|1)$ theory in section 3.4 that for gauge invariance, the charge m should be taken valued in $1/2 + \mathbb{Z}$. It is more convenient to work with an integer parameter $n = m + 1/2$, and we will accordingly label our line operators as $L_{\mathbf{t},n}$. Note that, since the longitude cycle is not canonically defined, the definition of these line operators depends on the knot framing. Under a unit change of framing, the Wilson line for the background gauge field will produce a factor of $\mathbf{t}^{n-1/2}$. With suitable choices of framing, half-integral powers of $\mathbf{t}$ will not appear in the

choice is equivalent to choosing a framing of the knot.

expectation values.

The operators $L_{\mathbf{t},n}$ will sometimes be called typical, while L_n and Wilson lines for the indecomposable representations will be called atypical. This terminology originates from the classification of superalgebra representations, as we briefly recall in section 3.4.1.

3.3 Magnetic Theory And The Meng-Taubes Theorem

As was explained in section 3.2.2, our Chern-Simons theory can be obtained from the theory of a free $\mathcal{N} = 4$ hypermultiplet by twisting. An alternative description of the same topological theory can be obtained, if we recall that the free hypermultiplet describes the infrared limit of the $\mathcal{N} = 4$ QED with one electron. This is the basic example of mirror symmetry [104] in three-dimensional abelian theories, which was understood in [105] as a functional Fourier transform. By metric independence of the topological observables, they can be equally well computed in the UV or in the IR description. We now consider the topologically-twisted version of the UV gauge theory, which we will call the "magnetic" description.

(On a compact manifold, the claim that the RG flow from the UV theory leads to a free hypermultiplet depends on the presence of the non-trivial background flat bundle, which forces the theory to sit near its conformally-invariant vacuum. When the background gauge field is turned off, *e.g.* as is necessarily the case for a manifold with trivial H_1, the situation is more subtle. This and some other details will be discussed in part 3.3.3 of the present section.)

3.3.1 The $\mathcal{N} = 4$ QED With One Electron

We now describe the bosonic fields of the theory. The fermionic fields, as well as the details on the action, are discussed in the Appendix A. Bosonic fields of the vector multiplet are a gauge field A_i and an $\mathrm{SU}(2)_Y$-triplet of scalars $Y^{\dot{a}}$. (Bosonic gauge field A_i here is completely unrelated to the fermionic gauge field of the electric gauge theory. In fact, the fields of the electric description emerge from the monopole operators of the UV theory.) In the twisting construction we use the $\mathrm{SU}(2)_X$-subgroup of the R-symmetry, so the scalars of the vector multiplet will remain scalars. It is convenient to introduce a combination $\sigma = (Y_2 - iY_3)/\sqrt{2}$, which has charge 2 under the ghost number symmetry $\mathrm{U}(1)_F$. The remaining component Y_1 has ghost number zero. The hypermultiplet contains an $\mathrm{SU}(2)_X$-doublet of complex scalars, which upon twisting become a spinor Z^α. They have charge one under the gauge group. The imaginary part ϕ of the background flat connection originated from the masses in the electric description. Under the mirror symmetry, it is mapped to a Fayet-Iliopoulos parameter.

The flavor symmetry $\mathrm{SU}(2)_{\mathrm{fl}}$ is emergent in the infrared limit. In the UV, only its Cartan part is visible – it is identified with the shift symmetry of the dual photon. The current for this symmetry is $\frac{-i}{2\pi} \star F$, where $F = \mathrm{d}A$. The real part of the background gauge field couples to this symmetry, so, it should enter the action in the interaction $-\frac{i}{2\pi}\int B \wedge F$. In fact, the whole action of the topological theory has the form

$$I_{\mathrm{QED}} = \{Q, \dots\} + I_{\mathrm{top}}\,, \tag{3.15}$$

where

$$I_{\mathrm{top}} = -\frac{i}{2\pi}\int (B + i\phi) \wedge F\,. \tag{3.16}$$

(More details are given in Appendix A.) This can be more accurately written as

$$\exp(-I_{\mathrm{top}}) = \mathcal{L}^{-1}\left(c_1(\mathcal{A})\right)\,, \tag{3.17}$$

where $\mathcal{A}$ is a line bundle, in which A is the connection. The fields Z^α take values in a spin-$\mathbb{C}$ bundle, and correspondingly, the path-integral includes a sum over spin-$\mathbb{C}$ structures s'. We view this spin-$\mathbb{C}$ bundle as a spin bundle S for some fixed spin structure s, tensored with the line bundle $\mathcal{A}$. We identify the reference spin structure s with the spin structure, which was used in the definition of torsion on the electric side. A change of the spin structure by an element $x \in H^1(W, \mathbb{Z}_2)$ is equivalent to twisting the bundle $\mathcal{A}$ by the $\mathbb{Z}_2$-bundle, corresponding to x. The formula (3.17) then changes in the same way (3.12) as the torsion $\tau_s(\mathcal{L})$, in agreement with the mirror symmetry[12]. The theory also has a charge conjugation symmetry, which, as on the electric side, implies that the invariants for $\mathcal{L}$ and $\mathcal{L}^{-1}$ are the same.

Note that, instead of (3.16), we could try to use

$$\exp(-I_{\mathrm{top}}) \stackrel{?}{=} \mathcal{L}^{-1}\left(\frac{1}{2}c_1(\det s')\right) . \tag{3.18}$$

Here $\det s'$ is the determinant line bundle of the spin-$\mathbb{C}$ bundle, in which the fields Z^α live. However, the factor of $1/2$ inside the brackets means that one has to take a square root of the holonomy of $\mathcal{L}$, and therefore the sign of this quantity is not well defined. This is the same ambiguity that we encountered in section 3.2.3.2, and it is resolved, again, by picking a reference spin structure s.

The functional integral of the magnetic theory can be localized on the solutions of BPS equations $\{Q, \psi\} = 0$, where ψ is any fermion of the theory. One group of these equations actually tells us that the solution should be invariant under the gauge transformation with parameter, equal to the field σ. We will only consider irreducible solutions, and therefore σ must be zero. We also only consider the case that the background field satisfies $\mathrm{d}\star\phi = 0$, so that the twisted theory has the full $\mathrm{SU}(2)_Y$-symmetry. (We have seen on the electric side that $\mathrm{d}\star\phi = 0$ is the condition for this symmetry to be present. On the magnetic side, one can also explicitly check this, as shown in the Appendix A.) This symmetry, together with vanishing of σ, implies that Y_1 is also zero. With this vanishing assumed, the remaining BPS

[12] Again, s should be more appropriately viewed as a spin-$\mathbb{C}$ structure with trivial determinant. Of course, we could equally well take an arbitrary reference spin-$\mathbb{C}$ structure. That would give the trivial generalization of τ_s to arbitrary spin-$\mathbb{C}$ structures, as described at the end of section 3.2.3.2.

equations take the form of the three-dimensional Seiberg-Witten equations,

$$F + \frac{1}{2} \star \left(\mu - e^2\phi\right) = 0\,,$$
$$\not{D} Z = 0\,, \tag{3.19}$$

where $\mu = i\sigma_{j\,\alpha}^{\beta}\, Z^{\alpha}\overline{Z}_{\beta}\, \mathrm{d}x^j$ is the moment map, with $\sigma_{i\,\alpha}^{\beta}$ being the Pauli matrices contracted with the vielbein, e^2 is the gauge coupling, and $\not{D}$ is the Dirac operator, acting on the sections of $S \otimes \mathcal{A}$. Generically, the localization equations have a discrete set of solutions, and the partition function of the theory can be written as

$$\tau_s(\mathcal{L}) = \sum_{\mathfrak{S}} (-1)^f\, \mathcal{L}^{-1}(c_1(\mathcal{A}))\,, \tag{3.20}$$

where the sum goes over the set $\mathfrak{S}$ of solutions of the Seiberg-Witten equations, $\mathcal{A}$ is a line bundle, corresponding to the given solution, and $(-1)^f$ is the sign of the fermionic determinant.

The relation between the Reidemeister-Turaev torsion and the Seiberg-Witten invariant in three-dimensions is the content of the Meng-Taubes theorem [20] and its refinement due to Turaev [92]. We have presented a physicist's derivation of this theorem. Some subtleties that arise for three-manifolds with $b_1 \leq 1$ are discussed later in this section.

3.3.2 Adding Line Operators

Let us describe the magnetic duals of line operators, which were introduced in section 3.2.4. The first type of line operators were the integrals of the fermionic gauge field $\int_K A^{\pm}$. On the magnetic side, their duals will be the integrals of monopole operators, which we will not discuss. The second type were the Wilson lines for the background gauge field. Obviously, their definition will be the same on the magnetic side.

Non-trivial and interesting line operators were defined by singularities of the background flat connection. We denoted them by $L_{\mathbf{t},n}$ in section 3.2.4. Since the one-form ϕ enters the BPS equations (3.19) on the magnetic side, the singularity of ϕ implies that those equations will have solutions with a singularity along the knot K. The line operator is then defined by requiring the fields to diverge near K as in a particular singular model solution. We use notation W for the closed three-manifold, and W' for the manifold, obtained from W by cutting out a small toric neighborhood of the singular line operator. Let r and θ be the polar coordinates in the plane, perpendicular to K. Near the knot, the singularity of the imaginary part of the background gauge field has the form

$$\phi = -\gamma \, \mathrm{d}\theta + \beta \, \frac{\mathrm{d}r}{r} \, . \tag{3.21}$$

(We follow the notations of [53].) Note that whenever the parameter β is non-zero, the closed two-form $\star\phi$ has a non-vanishing integral around the boundary of the toric neighborhood of the link. This might be forbidden for topological reasons – *e.g.*, if K is a one-component link in S^3. In such cases, β cannot be turned on. Even when the parameter β can be non-zero, we expect the invariants to be independent of it.

To find the model solution, consider the case that W is the flat space, and K is a straight line. Let Z^1 and Z^2 be the two components of Z^α, and $z = r\exp(i\theta)$ be the complex coordinate in the plane, perpendicular to K. We are looking for a time-independent, scale-invariant solution of the Seiberg-Witten equations. The gauge field in such a solution can be set to zero, so that the remaining equations give

$$Z^1(Z^2)^\dagger = \frac{e^2(\beta + i\gamma)}{2z} \, , \quad Z^1(Z^1)^\dagger - Z^2(Z^2)^\dagger = 0 \, , \quad \partial_{\bar{z}} Z^1 = \partial_z Z^2 = 0 \, , \tag{3.22}$$

and the scale-invariant solution is simply $Z^1 = a/\sqrt{z}$, $Z^2 = b/\sqrt{\bar{z}}$, where $ab^\dagger = e^2(\beta + i\gamma)/2$ and $|a|^2 = |b|^2$. Note that the field Z^α here is antiperiodic around K. Since we view Z^α as a spinor field on the closed three-manifold W, it should rather be periodic, so, we make a gauge transformation to bring the model solution to the form

$$Z^1 = \frac{a}{\sqrt{r}}\,, \quad Z^2 = \frac{b}{\sqrt{r}}\exp(i\theta)\,, \quad A = -\frac{1}{2}\mathrm{d}\theta\,. \tag{3.23}$$

To complete the definition of the line operator, we also need to explain, how the topological action (3.17) is defined in presence of the singularity. The flat bundle $\mathcal{L}$ is naturally an element of $\mathrm{Hom}(H_1(W'), \mathbb{C}^*)$. By Poincaré duality, it can be paired with an element of the relative cohomology $H^2(W', \partial W')$, and this pairing will define the action. If we forget for a moment about possible torsion, the relative cohomology class that we need is naturally the cohomology class $[F/2\pi]$ of the gauge field strength for a given solution. However, here we encounter the mirror of the problem that we had on the electric side: this class in general is not integral. The reason, roughly speaking, is the antiperiodicity of the field Z^α around K, or equivalently, the half-integral term $-\frac{1}{2}\mathrm{d}\theta$ in the gauge field (3.23) near the line operator. (Depending on the topology, there can also appear a similar term with θ replaced by the angle along K.) This will in general cause half-integral powers of the holonomies $\mathbf{t}$ and $\mathbf{t}_{\|}$ to appear in the torsion invariant. To remove them, just as in section 3.2.4, one introduces along K a Wilson line for the background gauge field with a half integral charge $n - 1/2$. With a suitable choice of framing, this is enough to remove the half-integral powers of holonomies.

Here we viewed the field Z^α as a section of the spin bundle on W, tensored with a line bundle $\mathcal{A}$ with connection A. A more systematic way to define these line operators is to allow spin (or spin-$\mathbb{C}$) structures on W' that do not necessarily extend to W. The antiperiodicity of Z in the model solution (3.23) can then be absorbed into the definition of this spin structure. The field Z^α then provides an honest cross-section of the line bundle $\mathcal{A}$ in the neighborhood of the link, and this allows to canonically define an integer-valued relative Chern class $c_1(\mathcal{A}) \in H^2(W', \partial W')$. The charge n of the background field Wilson line and the choice of the framing are then absorbed into the choice of the spin-$\mathbb{C}$ structure.

This is the approach taken in the mathematical literature[13], see *e.g.* [88]. This point of view is consistent with the picture that inserting line operators of type L_m, or changing the parameter n for operators $L_{\mathbf{t},n}$, is equivalent to changing the spin-$\mathbb{C}$ structure.

We only considered the case that the holonomy of the background flat connection around the meridian of the knot is not unimodular. In the opposite case, we have $\gamma = 0$ in eq. (3.21), and, assuming that the parameter β is also zero, the singular model solution seems to disappear. This makes it unclear, how to define the magnetic duals of line operators with unimodular holonomy, except by the analytic continuation from $\gamma \neq 0$. This problem looks analogous to the one described in the end of section 2.4.4.5 of Chapter 2.

3.3.3 More Details On The Invariant

In our derivation of the relation between the Seiberg-Witten invariant and the Reidemeister-Turaev torsion we ignored some subtleties [20], [106], which occur for three-manifolds with $b_1 \leq 1$. Here we would like to close this gap. First we look at the UV theory, and then we describe the RG flow to the IR theory in more detail. We will see that the claim that the IR theory is the $\mathfrak{psl}(1|1)$ Chern-Simons model sometimes has to be corrected.

3.3.3.1 Seiberg-Witten Equations For $b_1 \leq 1$

Let us look closer at the Seiberg-Witten counting problem. Our goal here is not to derive something new, but merely to understand, how gauge theory takes care of some subtleties in the formulation of the Meng-Taubes theorem.

Note that in the analogous problem in the context of Donaldson theory in four dimensions, the gauge theory gives the first of the Seiberg-Witten equations roughly in the form $F^+ + \overline{Z}Z = 0$. To avoid dealing with reducible solutions with $F^+ = 0$, one introduces by hand a deformation two-form in the equation [11]. In our case, the situation is different: the

[13] There is also another difference of our treatment of line operators from mathematical literature. There, the analogs of line operators are typically introduced by gluing in an infinite cylindrical end to the manifold W', rather than by considering solutions on W with singularities.

deformation two-form $e^2 \star \phi/2$ is already there. In nice situations, the counting of solutions does not depend on the choice of this deformation, so any two-form could be taken. But sometimes it is not true, and then it will be important, what deformation two-form is chosen for us by the gauge theory.

The properties of the counting problem depend on the first Betti number $b_1(W)$, whose role here is analogous to b_2^+ in four dimensions. For $b_1 > 0$, a reducible solution has $Z = 0$ and $F = e^2 \star \phi/2$. For such a solution to occur, the cohomology class of $e^2 \star \phi/2$ has to be integral. When in the parameter space we pass through such a point, so that reducible solutions are possible, the counting of solutions can in principle jump. This makes the Seiberg-Witten invariant dependent on the deformation two-form, or the metric and e^2, if we prefer to keep the deformation two-form equal to $e^2 \star \phi/2$ with fixed ϕ. Actually, for $b_1 > 1$ no jumping is possible, since in the space of deformation two-forms we can always bypass the point, where reducible solutions can occur. But for $b_1 = 1$, non-trivial wall-crossing phenomena do happen. As we change the two-form $e^2 \star \phi/2$, and its cohomology class passes through an integer point, the number of solutions with first Chern class $[F/2\pi]$ equal to this integer does change in a known way [107]. (For the particular case of $S^1 \times S^2$, the Seiberg-Witten counting problem is worked out in detail in the Example 4.1 in [88].)

There is another related issue. As we explained in section 3.2.3.1, the torsion, to which the Seiberg-Witten invariant is supposed to be equal to, for $b_1 = 1$ has a second order pole. Just for concreteness, consider the manifold $S^1 \times S^2$, for which the torsion is[14] $\tau(\mathbf{t}) = \mathbf{t}/(\mathbf{t}-1)^2$, where $\mathbf{t}$ is the holonomy around the non-trivial cycle. If we expand this, say, near $\mathbf{t} = 0$, we get a semi-infinite Laurent series $\mathbf{t} + 2\mathbf{t}^2 + \dots$. However, it is known that for any given deformation two-form the Seiberg-Witten equations have only a finite number of solutions.

The resolution of these puzzles is that we need to take the infrared limit of the theory, *e.g.* by taking e^2 to infinity. This means that we have to take the deformation two-form

[14] The function $\tau(\mathbf{t})$ should have a second order pole at $\mathbf{t} = 1$. Also, it cannot have any zeros for $\mathbf{t} \in \mathbb{C}^* \backslash \{1\}$, since the twisted cohomology $H^1(S^1 \times S^2, \mathcal{L})$ for such $\mathbf{t}$ is empty. Imposing also invariance under the charge conjugation $\mathcal{C}$, we recover the stated result, up to a constant numerical factor.

to be $+\infty \cdot \star\phi$. That is, it should be proportional to $\star\phi$ with a positive coefficient, and, to count solutions with a given Chern class $[F/2\pi]$, we should use a deformation two-form with Chern class much larger than $[F/2\pi]$ in absolute value. This is equivalent to the prescription of Meng and Taubes. Depending on the sign of ϕ, the two expansions that we get in this way for $S^1 \times S^2$ would be $\mathbf{t} + 2\mathbf{t}^2 + \ldots$ and $\mathbf{t}^{-1} + 2\mathbf{t}^{-2} + \ldots$. One can check that the sign of ϕ is such that $|\mathbf{t}| < 1$ in the first case and $|\mathbf{t}| > 1$ in the second, so that in either case the expansion is absolutely convergent.

Just like for closed three-manifolds, for manifolds with links in them, the Seiberg-Witten counting problem for $b_1(W') = 1$ is special. (This case arises *e.g.* when one cuts out a one-component knot from a simply-connected manifold.) As we reviewed in the end of section 3.2.3.3, the Reidemeister torsion for such a manifold has a first order pole. Therefore, it has two different Laurent expansions near $\mathbf{t} = 0$ and $\mathbf{t} = \infty$. The Seiberg-Witten equations in this case have an infinite number of solutions, with Chern class unbounded from above or from below, depending on the sign of the deformation two-form $e^2 \star \phi$. The sign of $e^2 \star \phi$ is such that these expansions are absolutely convergent. Unlike the case of a compact three-manifold, here we do not need to explicitly take e^2 to infinity, since the deformation two-form $e^2 \star \phi$ already diverges near the knot.

When W is a rational homology sphere, that is $b_1 = 0$, there is no way to avoid reducible solutions in working with the Seiberg-Witten equations. Because of this, a simple signed count of solutions is no longer a topological invariant. Still, one can define a topological invariant by adding an appropriate correction term [107]. We will not attempt to derive it from the quantum field theory, but will in what follows use the fact that the definition of the invariant for $b_1 = 0$ does exist.

3.3.3.2 Massive RW Model And The Casson-Walker Invariant

Let us now turn to the IR theory, which is a valid description, when the size of the three-manifold W is scaled to be large. The topological theory reduces in this case to the Rozansky-

Witten (RW) sigma-model [91] with the target space being the Coulomb branch manifold, which for the $\mathcal{N}=4$ QED is [108] the smooth Taub-NUT space X_{TN}. The U(1) graviphoton translation symmetry is generated on X_{TN} by a Killing vector field V. The coupling of the UV theory to the flat gauge field $B+i\phi$ translates into a coupling of the RW model to the same flat gauge field via the isometry V. In the untwisted language, the imaginary part ϕ of the gauge field would be a hyperkäler triplet of mass terms. For this reason, we call our IR theory the massive Rozansky-Witten model. An explicit Lagrangian and more detailed treatment of this theory will appear elsewhere. The coupling of the Rozansky-Witten model to a *dynamical* Chern-Simons gauge field has been previously considered in [24]. We will now see, how and when the massive RW model reduces to the Gaussian $\mathfrak{psl}(1|1)$ theory.

First, let us turn off the background flat gauge field. What we get then is the ordinary RW model for X_{TN}. The path-integral of that theory has the following structure [91]. The kinetic terms have both bosonic and fermionic zero modes. The bosonic ones correspond to constant maps to the target space. The integral over the bosonic zero modes thus is an integral over X_{TN}. The one-loop path-integral produces a simple measure factor, while most of the higher-loop diagrams vanish. The reason is that all the interactions (which involve the curvature of X_{TN}) are irrelevant in the RG sense, and can be dropped, when the worldsheet metric is scaled to infinity. However, some diagrams do survive due to the presence of the zero modes. Overall, the path-integral for each b_1 is given by a simple Feynman diagram, which captures the topological information about W, times the integral of the Euler density of X_{TN}. It is important that the path-integral depends on the target space only by this curvature integral. The Euler numbers happen to be the same for X_{TN} and for the Atiyah-Hitchin manifold X_{AH}. This was used in [93] to derive a special case of the Meng-Taubes theorem by the following argument. The RW model for X_{AH} can be obtained from the IR limit of the topologically-twisted $\mathcal{N}=4$ SU(2) Yang-Mills theory [108], which computes the Casson-Walker invariant [109], [110], [111]. Since the Rozansky-Witten invariants computed using X_{TN} and X_{AH} are the same, the Casson-Walker invariant is equal to the Seiberg-Witten invariant, when the background bundle is trivial.

Now let us turn on the background bundle back again. In its presence, the kinetic terms of the RW model have no zero modes. The classical solution, around which one expands, is the map to the fixed point of the vector field V, that is, to the conformally-invariant vacuum. In the absence of the zero modes, all the irrelevant curvature couplings can be thrown away. In this way, the RW path-integral reduces to the Gaussian integral of the $\mathfrak{psl}(1|1)$ model. It is natural to expect the path-integral to be continuous in $\mathcal{L}$. To the extent that this is true, the torsion $\tau(\mathcal{L})$ evaluated for trivial $\mathcal{L}$ should thus coincide with the Casson-Walker invariant. Note that on the level of Feynman diagrams this is not completely trivial, since for $B+i\phi = 0$ the interaction vertices come from the curvature terms, while for $B+i\phi \neq 0$ they come from expanding the Gaussian path-integral in powers of the background gauge field. Still, the actual Feynman integrals should coincide. We will not explicitly analyze the diagrams here (most of them were analyzed in [91]), but will just use the known relation between τ and the Casson-Walker invariant to check the continuity of the massive RW path-integral in $\mathcal{L}$.

Let $\tau(1)$ denote the torsion evaluated for the trivial background flat bundle[15], and CW be the Casson-Walker invariant. For $b_1 \geq 2$, it is indeed true that $\tau(1) = \mathrm{CW}$. For $b_1 = 1$, the torsion has the form

$$\tau(\mathbf{t}) = \frac{\mathbf{t}\Delta(\mathbf{t})}{(\mathbf{t}-1)^2}\,, \tag{3.24}$$

where $\Delta(\mathbf{t})$ is the Alexander polynomial. Setting $\mathbf{t} = \exp(m)$ and expanding this in m, we get

$$\tau(\mathbf{t}) = \frac{\Delta(1)}{m^2} + \frac{1}{2}\Delta''(1) - \frac{1}{12}\Delta(1) + O(m^2)\,. \tag{3.25}$$

Dropping the $1/m^2$ term, we define the regularized torsion $\tau_{\mathrm{reg}}(1) = \frac{1}{2}\Delta''(1) - \frac{1}{12}\Delta(1)$. This combination, again, is equal to the Casson-Walker invariant for $b_1 = 1$. However, the presence of the extra divergent piece $\Delta(1)/m^2$ means that the path-integral of the massive RW model in this case is not continuous in its dependence on the background gauge field:

[15] Note that for $\tau(1)$ the dependence on the spin-$\mathbb{C}$ structure drops out.

for $\mathcal{L}$ approaching the trivial flat bundle, the torsion tends to infinity, while for $\mathcal{L}$ taken to be exactly the trivial flat bundle, the invariant is finite. One can trace the origin of this discontinuity to the wall-crossing in the UV theory. Indeed, for non-zero ϕ, the Seiberg-Witten invariant is evaluated using the deformation two-form $e^2 \star \phi/2$, which in the infrared limit $e^2 \to \infty$ lands us in the infinite wall-crossing chamber. The $1/m^2$ singularity of the torsion for $m \to 0$ arises from the infinite number of solutions of the Seiberg-Witten equations in this chamber. On the other hand, for trivial $\mathcal{L}$ we have $\phi = 0$, and the deformation two-form vanishes for all e^2. To evaluate the invariant, one should properly deal with reducible solutions. Instead, we will simply assume that the deformation two-form is non-zero, but infinitesimally small. It is known [88] that in such chamber the Seiberg-Witten invariant is equal to $\frac{1}{2}\Delta''(1)$, which, again, is the Casson-Walker invariant, up to a correction $-\frac{1}{12}\Delta(1) = -\frac{1}{12}|\mathrm{tor}\, H_1(W)|$, which, presumably, would be recovered with an appropriate treatment of the reducible solutions. Thus, one can say that the discontinuity at trivial $\mathcal{L}$ in the massive RW model for $b_1 = 1$ is a "squeezed version" of the wall-crossing in the UV theory[16]

Finally, for $b_1 = 0$, assuming that the torsion subgroup tor $H_1(W)$ is non-empty, the Reidemeister-Turaev torsion is a function on the discrete set of flat bundles. For non-trivial $\mathcal{L}$, the Seiberg-Witten counting problem computes the torsion, while for trivial $\mathcal{L}$ it computes the Casson-Walker invariant, which now is not related to the torsion, since there is no way to continuously interpolate to the trivial $\mathcal{L}$, starting from a non-trivial $\mathcal{L}$. In fact, for $b_1 = 0$ the Casson invariant is computed by a two-loop Feynman integral [91], and it is clearly not possible to obtain it from the one-loop torsion.

Let us summarize. The UV topological theory, and thus the Seiberg-Witten counting problem, is equivalent to the massive RW theory. For non-trivial $\mathcal{L}$, this theory reduces to the $\mathfrak{psl}(1|1)$ Chern-Simons theory and computes the Reidemeister-Turaev torsion. For trivial $\mathcal{L}$, it computes the Casson-Walker invariant, which for $b_1 > 0$ can be obtained from a limit

[16] It is a "squeezed version", because the wall-crossing condition is not conformally-invariant, and thus we cannot see all the walls in the IR theory, but only see a discontinuity at $\phi = 0$. This can be contrasted with the situation in the Donaldson theory in four dimensions, where the wall-crossing condition is conformally-invariant, and the walls can be seen both in the UV and in the IR descriptions [112].

of the $\mathfrak{psl}(1|1)$ invariant, while for $b_1 = 0$ is not related to it. Our results agree with the mathematical literature [88], [106].

3.4 U(1|1) Chern-Simons Theory

In a series of papers [59, 60, 61], it has been shown that the Alexander polynomial and the Milnor torsion can be computed from the U(1|1) Chern-Simons theory. We would like to revisit this subject and to show, how it fits together with our discussion in previous sections. We point out that for the compact form of the bosonic gauge group, the U(1|1) Chern-Simons theory is simply an orbifold of the $\mathfrak{psl}(1|1)$ theory. (A direct analog of this statement is well-known in the ABJM context.) In particular, it contains no new information compared to the $\mathfrak{psl}(1|1)$ Chern-Simons with a coupling to a general background flat bundle $\mathcal{L}$, and computes, indeed, essentially the same invariant.

3.4.1 Lie Superalgebra $\mathfrak{u}(1|1)$

We start with a brief review of the superalgebra $\mathfrak{u}(1|1)$. A more complete discussion can be found *e.g.* in [103]. Let $\widehat{f}_+$ and $\widehat{f}_-$ be the fermionic generators, and $\widehat{t}_\ell$ and $\widehat{t}_r$ the generators of the left and right bosonic $\mathfrak{u}(1)$ factors. It will also be convenient to use a different basis in the bosonic subalgebra, which is $\widehat{E} = \widehat{t}_r + \widehat{t}_\ell$ and $\widehat{N} = (\widehat{t}_r - \widehat{t}_\ell)/2$. The element $\widehat{N}$ acts on the fermionic subalgebra by the $\mathrm{U}(1)_{\mathrm{fl}}$ transformations, and the element $\widehat{E}$ is central. Explicitly, the non-trivial commutation relations are

$$[\widehat{N}, \widehat{f}_\pm] = \pm \widehat{f}_\pm\,, \quad \{\widehat{f}_+, \widehat{f}_-\} = \widehat{E}\,. \tag{3.26}$$

The group of even automorphisms of $\mathfrak{u}(1|1)$ is generated by the charge conjugation $\widehat{E} \to -\widehat{E}$, $\widehat{N} \to -\widehat{N}$, $\widehat{f}_\pm \to \pm\widehat{f}_\pm$, rescalings $\widehat{f}_\pm \to a_\pm \widehat{f}_\pm$, $\widehat{E} \to a_+ a_- \widehat{E}$ with $a_\pm \in \mathbb{R} \setminus 0$, and shifts $\widehat{N} \to \widehat{N} + b\widehat{E}$, $b \in \mathbb{R}$.

As for any Lie superalgebra, the representations of $\mathfrak{u}(1|1)$ can be usefully divided into two classes – the typical and the atypical ones. (For a brief review of superalgebra representations, the reader can consult section 2.3.1.) The typicals are precisely the ones, in which the central generator $\widehat{E}$ acts non-trivially. They are two-dimensional, and the generators, in some basis, act by matrices

$$\widehat{E} = w \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} , \quad \widehat{N} = \begin{pmatrix} n & 0 \\ 0 & n-1 \end{pmatrix} , \quad \widehat{f}_+ = \begin{pmatrix} 0 & w \\ 0 & 0 \end{pmatrix} , \quad \widehat{f}_- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} , \tag{3.27}$$

with $w \neq 0$. These will be called representations of type (w, n). To be precise, one has to make a choice, whether to assign a bosonic or a fermionic parity to the highest weight vector. This effectively doubles the number of representations. In our applications, the representations will be labeling closed Wilson loops, which come with a supertrace. Therefore, different parity assignments will be just a matter of overall sign, and we will mostly ignore this.

In the atypical representations the generator $\widehat{E}$ acts trivially, and therefore they can be equivalently thought of as representations of $\mathfrak{pl}(1|1)$. These have already been described in section 3.2.4. Note that the indecomposable representation $(0, n)_-$ of fig. 3.1 can be obtained as a degeneration of the typical representation (w, n) for $w \to 0$. With a suitable rescaling of the generators $\widehat{f}^{\pm}$ before taking the limit, one can similarly obtain the representation $(0, n)_+$ of fig. 3.1. The representations $(0, n)_-$ and $(0, n)_+$ are known as the atypical Kac module and anti-Kac module.

Let us also write out some tensor products. Tensoring any representation with the one-dimensional atypical (n) simply shifts the $\widehat{N}$-charges. The other tensor products are

$$(w_1, n_1) \otimes (w_2, n_2) = (w_1 + w_2, n_1 + n_2) \oplus (w_1 + w_2, n_1 + n_2 - 1)' , \quad w_1 + w_2 \neq 0 ; \tag{3.28}$$

$$(w, n_1) \otimes (-w, n_2) = P_{n_1+n_2} , \tag{3.29}$$

where the indecomposables P_n were defined on fig. 3.1. The prime on the second representation in the r.h.s. of (3.28) means that the highest weight vector in it has reversed parity. The set of representations (n), (w, n), P_n is closed under tensor products.

The superalgebra $\mathfrak{u}(1|1)$ possesses a two-dimensional family of non-degenerate invariant bilinear two-forms, which can be obtained by taking a supertrace over a (w, n) representation with $w \neq 0$. Note that all the representations (w, n) for different values of $w \neq 0$ and n, and therefore also the corresponding invariant forms, are related by the superalgebra automorphisms.

3.4.2 Global Forms

There exist different versions of Chern-Simons theory based on the superalgebra $\mathfrak{u}(1|1)$, and here we would like to classify them. To define such a theory, one needs to pick a global form of the gauge group, and also to choose an invariant bilinear form, with which to define the action. These data should be consistent, in the sense that the action should be invariant under the large gauge transformations. Theories related by the superalgebra automorphisms are equivalent. We can use this symmetry to bring either the invariant bilinear form, or the lattice, which defines the global form of the group, to some simple canonical form. To classify the theories, it is convenient to take the first approach.

Let $\mathfrak{g}_{\overline{0}} \simeq \mathbb{R}^2$ be the bosonic subalgebra of $\mathfrak{u}(1|1)$. The $\mathfrak{u}(1|1)$ gauge field, in components, is $A = A^N \widehat{N} + A^E \widehat{E} + A^+ \widehat{f}_+ + A^- \widehat{f}_-$. For the bosonic part of the gauge field, we will also use expansion in a different basis, $A^N \widehat{N} + A^E \widehat{E} \equiv A^\ell \,\widehat{t}_\ell + A^r \,\widehat{t}_r$. The action of the theory can be written as

$$I_{\mathfrak{u}(1|1)} = I_{\mathrm{bos}} + I_{\mathfrak{psl}(1|1)}(\mathcal{L}_{A^N} \otimes \mathcal{L}) + I_{\mathrm{g.f.}}\,, \tag{3.30}$$

where I_{bos} is the Chern-Simons term for the bosonic gauge field, $I_{\mathfrak{psl}(1|1)}$ is the action (3.1), coupled to the line bundle $\mathcal{L}_{A^N}$ with connection A^N, and to some background flat bundle $\mathcal{L}$. Finally, $I_{\mathrm{g.f.}}$ is the gauge-fixing action (3.3) for the fermionic part of the gauge symmetry.

By using the superalgebra automorphisms, we bring the bosonic Chern-Simons term to the form

$$I_{\text{bos}} = \frac{i}{4\pi} \int_W A^r \mathrm{d}A^r - A^\ell \mathrm{d}A^\ell \,. \tag{3.31}$$

(As usual, this formula is literally true only for topologically-trivial bundles. More generally, it is implicitly understood that the action is defined by integrating Chern classes of a continuation of the bundle to some four-manifold.)

Different versions of the theory will correspond to different choices of the global form of the bosonic subgroup $\mathrm{G}_{\bar{0}}$ of $\mathrm{U}(1|1)$. A global form is fixed, once we choose a cocharacter lattice $\Gamma_{\text{coch}} \subset \mathfrak{g}_{\bar{0}}$, that is, the lattice by which to factorize the vector space $\mathfrak{g}_{\bar{0}}$ to get the torus $\mathrm{G}_{\bar{0}}$. The first constraint on possible choices of the lattice Γ_{coch} comes from the fact that the fermionic generators of $\mathfrak{u}(1|1)$ should transform in a well-defined representation of $\mathrm{G}_{\bar{0}}$. In the basis dual to $(\widehat{t}^\ell, \widehat{t}^r)$, the corresponding weight has coordinates $(-1, 1)$, and we require that this vector be contained in the dual lattice $\Gamma_{\text{ch}} \simeq \Gamma^*_{\text{coch}}$.

We also need to make sure that the action (3.31) is invariant under the large gauge transformations. This will be true, if the number

$$\frac{1}{2} \int_V c_1^r \wedge c_1^r - c_1^\ell \wedge c_1^\ell \tag{3.32}$$

is integer on any closed spin four-manifold V. (We restrict to spin four-manifolds, because we already have a choice of a spin structure on W.) Here $c_1^{r,\ell} = [\mathrm{d}A^{r,\ell}/2\pi]$ are the $H^2(V, \mathbb{R})$-valued Chern classes for some extension of the $\mathrm{G}_{\bar{0}}$-bundle onto V.

The classes c^r and c^ℓ for different $\mathrm{G}_{\bar{0}}$-bundles form a lattice in $H^2(V, \mathbb{R}) \oplus H^2(V, \mathbb{R})$, which is naturally isomorphic to $\Gamma_{\text{coch}} \otimes H^2(V)$ (modulo torsion). Any element of this lattice can be expanded as $v_1\omega_1 + v_2\omega_2$, where ω_1 and ω_2 are arbitrary classes in $H^2(V)$, and v_1 and v_2 are the generators of the lattice Γ_{coch}. The quadratic form (3.32) can be explicitly written as

$$a_{11} \int_V \frac{1}{2}\omega_1 \wedge \omega_1 + a_{12} \int_V \omega_1 \wedge \omega_2 + a_{22} \int_V \frac{1}{2}\omega_2 \wedge \omega_2 \,, \tag{3.33}$$

with $a_{11} = (v_1^r)^2 - (v_1^\ell)^2$, $a_{12} = v_1^r v_2^r - v_1^\ell v_2^\ell$ and $a_{22} = (v_2^r)^2 - (v_2^\ell)^2$. For (3.33) to be an integer for arbitrary ω_1 and ω_2, the three coefficients a_{ij} should be integers. (We used again the fact that the intersection form on a spin four-manifold is even.) This condition is precisely equivalent to the requirement for Γ_{coch} to be an integral lattice in $\mathbb{R}^{1,1}$. We conclude that U(1|1) Chern-Simons theories are labeled by integral lattices in $\mathbb{R}^{1,1}$, whose dual contains the vector $(-1,1)$.

3.4.3 The Orbifold

To show that the theory is an orbifold of $\mathfrak{psl}(1|1)$ Chern-Simons, it is convenient to rewrite it in a different way. Let us use the basis $(\widehat{E},\widehat{N})$ in $\mathfrak{g}_{\bar{0}}$, in which the $\mathbb{R}^{1,1}$ scalar product is $(u,v) = u^N v^E + u^E v^N$. Let k and ν be some positive integers, and ξ be an integer or a half-integer, defined modulo k. By taking $v_1 = (k/\nu, 0)$ and $v_2 = (\xi/\nu, \nu)$ as the generators, for any such set we define a lattice, which actually has the right properties to serve as Γ_{coch}. The opposite is also true: any lattice Γ_{coch} has a basis of this form, and it is unique modulo shifting ξ by a multiple of k. (The parameter k is actually the area of the fundamental domain of Γ_{coch}.) This can be seen as follows. Let $v_1 = (a,b)$ and $v_2 = (c,d)$ be some generators of Γ_{coch}. The condition that the weight of the fermionic part of the superalgebra is a well-defined weight of $\mathrm{G}_{\bar{0}}$ means that b and d are integers. Let ν be their greatest common divisor. Then, by Euclidean algorithm, there exists an $\mathrm{SL}(2,\mathbb{Z})$-matrix of the form

$$\begin{pmatrix} d/\nu & -b/\nu \\ p & q \end{pmatrix}, \tag{3.34}$$

with some p and q. Transforming the basis of the lattice with this matrix, we find a basis of the form $v_1 = (a',0)$, $v_2 = (b',\nu)$. (We choose a' to be positive.) The integrality of the lattice means that $a'\nu \in \mathbb{Z}$ and $2b'\nu \in \mathbb{Z}$, so we can indeed parameterize the basis vectors in terms of k, ξ and ν. Residual $\mathrm{SL}(2,\mathbb{Z})$-transformations of the basis shift ξ by multiples of k.

Now we can make a superalgebra automorphism $\widehat{E}' = \frac{k}{\nu}\widehat{E}$, $\widehat{N}' = \widehat{N} + \frac{\xi}{\nu^2}\widehat{E}$ to transform this basis into $v_1' = (1,0)$, $v_2' = (0,\nu)$, at the expense of changing the action from its canonical form (3.31) to

$$I_{\rm bos} = \frac{i}{2\pi}\int_W \frac{k}{\nu}A^N \mathrm{d}A^E + \frac{\xi}{\nu^2}A^N\mathrm{d}A^N\,. \tag{3.35}$$

The path-integral involves a sum over topological classes of bundles, which are parameterized by the first Chern classes of the A^E and A^N bundles, which take values in $H^2(W)$ and $\nu H^2(W)$, respectively. For every topological type, let us write the gauge field A^E as a sum of some fixed connection $A^E_{(0)}$ and a one-form a^E. Integrating over a^E produces a delta-function, which localizes the integral to those connections A^N, which are flat. The $\mathfrak{psl}(1|1)$ part of the path-integral can then be taken explicitly, and we get for the U(1|1) partition function,

$$\tau_s^{\mathrm{U}(1|1)}(\mathcal{L}) = \int \mathrm{D}A^N \sum_{c_1^E} \delta(k\mathrm{d}A^N/2\pi)\,\mathcal{L}_{A^N}(kc_1^E)\,\exp(\xi \mathrm{CS}(\mathcal{L}_{A^N}))\,\tau_s(\mathcal{L}^\nu_{A^N}\otimes\mathcal{L})\,. \tag{3.36}$$

Here for convenience we changed the integration variable $A^N\to\nu A^N$. The origin of different terms here is as follows. The sum over the (integral) Chern classes c_1^E is what remained from the functional integral over A^E. The delta-function came from the integration over a^E. The holonomy of the flat bundle $\mathcal{L}_{A^N}$ around the Poincaré dual of kc_1^E is just a rewriting of the exponential of the Chern-Simons term $kA^N\mathrm{d}A^E/2\pi$. The Chern-Simons term for $\mathcal{L}_{A^N}$ with coefficient ξ came from the $A^N\mathrm{d}A^N/2\pi$ term in the action (3.35). Finally, τ_s is the $\mathfrak{psl}(1|1)$ torsion evaluated for a flat bundle, which is the ν-th power of $\mathcal{L}_{A^N}$, tensored with some background flat bundle $\mathcal{L}$.

Essentially the same path-integral as (3.36) was considered in section 2.2 of [113]. It was noted that the sum over c_1^E is proportional to the delta-function, supported on flat bundles with $\mathbb{Z}_k$-valued holonomy, since the pairing between $H^2(W)$ and the group of flat bundles is perfect. (That paper actually considered $k=1$.) Using this, we finally get

$$\tau_s^{\mathrm{U}(1|1)}(\mathcal{L}) = \frac{1}{k}\sum_{\mathcal{L}_k}\exp(\xi\mathrm{CS}(\mathcal{L}_k))\,\tau_s(\mathcal{L}_k^\nu\otimes\mathcal{L})\,, \tag{3.37}$$

where the sum goes over all $\mathbb{Z}_k$-bundles $\mathcal{L}_k$. The factor of k appeared from the delta-function in (3.36). To be precise, the explanations that we gave are sufficient to fix this formula only up to a prefactor. For manifolds with $b_1 = 0$, the normalization (3.37) can be recovered from the considerations in section 2.2 of [113]. We expect that it is correct in general. The factor of $1/k$ has a natural interpretation in terms of the orbifold – it is the volume of the isotropy subgroup, which is $\mathbb{Z}_k$.

An important special case is the U(1|1) Chern-Simons defined with the most natural global form of the group, where we simply set $\exp(2\pi i\widehat{t_\ell}) = \exp(2\pi i\widehat{t_r}) = 1$. The action is (3.31) with an integer factor k in front of it. By making an automorphism transformation, this theory can be mapped to the form (3.35) with $\xi = k/2$ and $\nu = 1$. Interestingly, it becomes independent of the spin structure, if k is odd. This is because the sign of the fermionic determinant is changing in the same way as the half-integral Chern-Simons term for A^N. For the general version of the theory, the dependence on the spin structure drops out when $\nu/2 + \xi \in \mathbb{Z}$. In what follows, we restrict to the version of the theory with $\xi = 0$ and $\nu = 1$.

Let us make some terminological comments. We call the theory U(1|1) Chern-Simons, and not $\mathfrak{gl}(1|1)$ or $\mathfrak{u}(1|1)$, because we need to choose a reality condition and a global form for the bosonic subgroup – and we take it to be $\mathrm{U}(1) \times \mathrm{U}(1)$. One could in principle consider other real and global forms. Those theories, if well-defined, would not need to be related to the $\mathfrak{psl}(1|1)$ theory by orbifolding. For the $\mathfrak{psl}(1|1)$ theory, we do not use the name PSU(1|1), because there is no bosonic subgroup, and therefore no choice of the real form or the global form. This theory is naturally associated to the complex Lie superalgebra.

In this thesis, we will not attempt to derive a relation between the supergroup Chern-Simons theory and the WZW models. However, if such a relation does exist, then what we have explained in this section would imply some correspondence between the U(1|1) and the $\mathfrak{psl}(1|1)$ WZW models. A duality of this kind is indeed known [114], although its derivation does not look similar to ours.

3.4.4 Magnetic Dual

The dual magnetic description of the theory is, of course, simply the orbifold of the QED of section 3.3. (This fact can also be independently derived from brane constructions, as we review later in section 3.6.3.) For the polynomial (3.20), summing over flat bundles has simply the effect of picking only powers of holonomies, which are multiples of k. Equivalently, note that the action of the magnetic theory will have the form analogous to (3.30), but with $I_{\mathfrak{psl}(1|1)} + I_{\text{g.f.}}$ replaced by the QED action. The field A^N couples to the QED topological current $iF/2\pi$. Integrating over A^N, we simply get that the Chern class of the QED gauge field is the k-th multiple of the Chern class of the A^E bundle. Since this bundle is arbitrary, we conclude that the orbifold of the magnetic theory is just the same QED, but with a constraint that the Chern class of the gauge field takes values in $kH^2(W)$. This can be equivalently viewed as[17] an $\mathcal{N}=4$ QED with one electron of charge k.

The $\mathfrak{u}(1|1)$ partition function $\tau_s^{\mathrm{U}(1|1)}(\mathcal{L})$ inherits from the torsion τ_s the dependence on the spin-$\mathbb{C}$ structure with trivial determinant. As we noted in the end of section 3.2.3.2, the definition of $\tau_s(\mathcal{L})$ can be easily extended to construct a torsion, which depends on an arbitrary spin-$\mathbb{C}$ structure, with no constraint. The same applies to $\tau_s^{\mathrm{U}(1|1)}(\mathcal{L})$. Now, consider the limit $k \to \infty$. Since now we sum essentially over all flat bundles, the U(1|1) partition function cannot depend on the unitary part of the flat connection in $\mathcal{L}$. Therefore, by holomorphicity, it will not depend on $\mathcal{L}$ at all. We denote this version of the torsion by $\tau_{s,\infty}$. This is a number, which depends only on W and on the choice of a spin-$\mathbb{C}$ structure. Looking at the magnetic side, it is clear that this number is precisely the signed count of solutions to the Seiberg-Witten equations, with the fields Z^α valued in a given spin-$\mathbb{C}$ bundle s. We conclude that the version of the electric theory with $k=\infty$ has these integers as its partition function. We note that this version of the torsion invariant has been defined and studied in [115] and [92]. The fact that it is an integer was demonstrated by purely combinatorial methods. One pedantic comment that we have to make is that $\tau_{s,\infty}$ is

[17] I thank N. Seiberg for pointing this out.

completely independent of $\mathcal{L}$ only for a manifold with $b_1 > 1$. For $b_1 = 1$, it does depend on the orientation in $H^1(W, \mathbb{R})$, induced by the absolute value of the holonomy of $\mathcal{L}$, since we need to choose the chamber, in which the Seiberg-Witten invariant is computed.

3.4.5 Line Operators

In the U(1|1) theory, we can define some Wilson loops. For the atypical representations, these are essentially the operators that were already defined earlier in section 3.2.4 for the $\mathfrak{psl}(1|1)$ theory. These are the operators L_n, labeled by one-dimensional atypicals (n), as well as Wilson lines for the indecomposable representations, whose role we still have to clarify.

For the typical representations (w, n), we want to claim that the Wilson lines are actually equivalent to the twist line operators of type $L_{\mathbf{t},n}$ with $\mathbf{t} = \exp(2\pi i w/k)$. This relation is the usual statement of equivalence of Wilson lines and monodromy operators in Chern-Simons theory. (For U(1|1), this relation was first suggested in [61].) The argument adapted to the supergroup case is given[18] in section 2.3.2. One consistency check can be made by looking at the transformation of these operators under the charge conjugation symmetry $\mathcal{C}$. As can be seen from (3.27), the representation changes as $(w, n) \to (-w, 1-n)$, while the twist operator changes as $L_{\mathbf{t},n} \to L_{\mathbf{t}^{-1},1-n}$, as follows from its definition in section 3.2.4. This is consistent with the identification of the operators. Note also that the boson-fermion parity of the highest weight vector of the representation (w, n) is changed under the charge conjugation. A Wilson loop with a supertrace will consequently change its sign. This can be taken as an explanation of the factor $(-1)^\ell$ in the formula (3.14) for the charge conjugation transformation of torsion in presence of the boundary. For $\mathbf{t} = \exp(2\pi i w/k)$, we will also denote the operators $L_{\mathbf{t},n}$ by $L_{w,n}$. Hopefully, this will not cause confusion.

[18]In fact, for U(1|1) the statement is quite obvious. The two-dimensional representation (w, n) can be obtained by quantizing a pair of fermions, living on the Wilson line. After gauging these fermions away, one is left with a singularity in the gauge field, which is equivalent to the monodromy $\mathbf{t}$. The ubiquitous shift of n by $1/2$ can be understood as a shift of the weight by the Weyl vector of the superalgebra $\mathfrak{u}(1|1)$. The combination $m = n - 1/2$, which appeared in section 3.2.4, is the "quantum-corrected" weight.

3.5 Hamiltonian Quantization

It is a well-established fact that the quantization of the Chern-Simons theory with an ordinary compact gauge group leads to conformal blocks of a WZW model [4, 48, 86, 116]. For the supergroup case, it is often assumed that a similar relation holds [59, 61, 34], however, to our knowledge, no derivation of this statement is available in the literature, and the properties of the supergroup theories in the Hamiltonian picture are fundamentally unclear. In this section, we take an opportunity to bring some clarity to the subject by explicitly quantizing the Chern-Simons theories, which were considered in previous sections. Since these theories are essentially Gaussian, the quantization is straightforward. In this thesis, we do not attempt to derive a relation to the conformal field theory.

3.5.1 Generalities

In the quantization of an ordinary, bosonic Chern-Simons theory on a Riemann surface Σ, the classical phase space to be quantized is the moduli space of flat connections on Σ. Dividing by the gauge group typically introduces singularities, which, however, do not play much role – the correct thing to do is to throw them away by replacing the moduli space of flat connections by the moduli space of stable holomorphic bundles. In the supergroup case, this approach does not seem to lead to consistent results. Reducible connections here can lead to infinite partition functions (as in the case of the theory on S^3), and that should somehow be reflected in the canonical quantization. The correct approach, we believe, is to consider the theory with gauge-fixed fermionic part of the gauge symmetry. The Hilbert space of the supergroup Chern-Simons should then be constructed by taking the cohomology of the BRST supercharge in the joint Hilbert space of gauge fields and superghosts. Due to "non-compactness" of the fermionic directions, even in the ghost number zero sector this cohomology is not equivalent to throwing the ghosts away.

First we consider the quantization of the $\mathfrak{psl}(1|1)$ Chern-Simons theory. We take the three-manifold to be a product $\mathbb{R}_t \times \Sigma$, where $\mathbb{R}_t$ is the time direction, and Σ is a connected oriented Riemann surface. Non-zero modes of the fields along Σ do not contribute to the cohomology of Q, and can be dropped. Zero-modes are present, when the cohomology $H^\bullet(\Sigma, \mathcal{L})$ of the de Rham differential on Σ, twisted by the connection in the flat bundle $\mathcal{L}$, is non-trivial. When $H^1(\Sigma, \mathcal{L})$ is non-empty, there is a moduli space of fermionic flat connections on Σ. This gives a number of fermionic creation and annihilation operators, and a finite-dimensional factor for the Hilbert space, – in complete analogy with the ordinary, bosonic Chern-Simons. This will be illustrated in examples later in this section. The zeroth cohomology $H^0(\Sigma, \mathcal{L})$ is non-empty, if and only if the flat bundle $\mathcal{L}$ is trivial on Σ. In this case, the cohomology is one-dimensional, since we have assumed Σ to be connected. The ghosts and the time component A_0 of the fermionic gauge field now have zero modes, which organize themselves into the quantum mechanics of a free superparticle in $\mathbb{R}^{4|4}$, with the action

$$-\int \mathrm{d}t \, \mathrm{Str}(-A_0\dot{\lambda} + \dot{\overline{C}}\dot{C}) \,. \tag{3.38}$$

(Here for simplicity we did not write the coupling to the external gauge field.) The Hilbert space[19], before we reduce to the cohomology of Q, is the space of functions on $\mathbb{C}^2$ (with holomorphic coordinates given by the components $C^\pm$ of the scalar superghost), tensored with the four-dimensional Hilbert space of the fermions $\lambda^\pm$ and $A_0^\pm$. We can write the states as

$$\psi_0|0\rangle + \psi_+\lambda^+|0\rangle + \psi_-\lambda^-|0\rangle + \psi_{+-}\lambda^+\lambda^-|0\rangle \,, \tag{3.39}$$

where $|0\rangle$ is annihilated by $A_0^\pm$, and $\psi_\bullet$ are functions of C and $\overline{C}$. We recall from eq. (3.2) that the BRST differential transforms $\overline{C}$ into λ. If we treat $\lambda^\pm$ as the differentials $\mathrm{d}\overline{C}^\pm$ and identify the wavefunctions (3.39) with differential forms on $\mathbb{C}^2$ with antiholomorphic indices, then Q acts as the Dolbeault operator. Thus, formally, the Hilbert space of the ghost system

[19]Here and in what follows, by "Hilbert space" we really mean the space of states. It does not, in general, have an everywhere-defined non-degenerate scalar product.

$\mathcal{H}_{\text{gh}}$ is the Dolbeault cohomology[20] of $\mathbb{C}^2$ with antiholomorphic indices.

Since $\mathbb{C}^2$ is non-compact, it is not obvious, how to make precise sense of this statement. Certainly, the path-integral of the theory on some three-manifold with a boundary produces a Q-closed state on the boundary. But to divide by Q-exact wavefunctions, we need to specify, what class of states is considered. For example, one could consider differential forms with no constraints on the behavior at infinity. This would lead to the ordinary Dolbeault complex. By the $\overline{\partial}$-Poincaré lemma, the cohomology is supported in degree zero, and consists simply of holomorphic functions on $\mathbb{C}^2$. This space will be denoted by $H^{0,0}_{\overline{\partial}}$, and the states will be called non-compact. In our applications, we can usually restrict to states, which are invariant under the $\mathrm{U}(1)_F$ ghost number symmetry. In $H^{0,0}_{\overline{\partial}}$, such states are multiples of $v_0 = |0\rangle$, the constant holomorphic function. Another possibility is to look at the cohomology with compact support[21]. By Serre duality, it is the dual of the space of holomorphic functions, and lives in degree $(0,2)$. We will denote this space by $H^{0,2}_{\overline{\partial},\text{comp}}$, and call the corresponding states compact. The $\mathrm{U}(1)_F$-invariant states here are multiples of $v_1 = \delta^{(4)}(C,\overline{C})\lambda^+\lambda^-|0\rangle$.

To understand the interpretation of these states in our theory, we need to recall some properties of the torsion. Let W' be a three-manifold with boundary Σ, together with some choice of the flat bundle $\mathcal{L}$ and, possibly, line operators inside. Let the holonomies of $\mathcal{L}$ be trivial on Σ, so that $H^0(\Sigma,\mathcal{L})$ is non-empty. If the flat bundle $\mathcal{L}$ is completely trivial even inside W', and, in particular, W' contains no line operators $L_{\mathbf{t},n}$, we call the manifold with this choice of the flat bundle unstable. In the opposite case, we call it stable. Let W be a connected sum of two three-manifolds W_1 and W_2 along their common boundary Σ, with no holonomies of $\mathcal{L}$ along the cycles of Σ. There are three possibilities. If both W_1 and W_2 are stable, the path-integral on W vanishes, because of the fermionic zero modes, – this property of the torsion is known as "unstability". If both W_1 and W_2 are unstable, the path-integral is not well-defined, because of the presence of both fermionic and bosonic zero modes. Finally,

[20] In the context of general Rozansky-Witten theories this statement – with $\mathbb{C}^2$ replaced by a compact hyper-Kähler manifold – appeares already in the original paper [91].

[21] For our purposes, the cohomology with compact support and the integrable cohomology will be considered as identical.

if one of W_1, W_2 is stable, and the other is unstable, the functional integral generically has no zero modes, and the torsion is a finite number.

We claim that our functional integral for an unstable three-manifold W' with boundary Σ naturally yields a state for the ghosts in the non-compact cohomology $H^{0,0}_{\overline{\partial}}$. Indeed, the zero modes of C, $\overline{C}$ and λ are completely free to fluctuate inside W', and therefore the wavefunction as a function of C is constant and should not contain insertions of λ, – so, it is a multiple of v_0. On the other hand, if the manifold W' is stable, we get a state in the compact cohomology $H^{0,2}_{\overline{\partial},\mathrm{comp}}$. The holonomies of the flat bundle inside W' do not allow the zero modes of the ghosts and λ to freely go to infinity. Modulo Q, the wavefunction in this case is a multiple of the state v_1. The natural pairing between the compact and the non-compact cohomology yields a finite answer for a closed three-manifold, glued from a stable and an unstable piece. If, on the other hand, we try to pair two stable manifolds, we get zero, since we have too many insertions of the operators $\lambda^{\pm}$ in the product of the wavefunctions. If we try to pair two non-compact, unstable states, the result is not well-defined, because one encounters both bosonic and fermionic zero modes[22]. This is consistent with the properties of the torsion, described above.

In the special case that Σ is a two-sphere with no punctures, the ghost Hilbert space $\mathcal{H}_{\mathrm{gh}}$ is all of the Hilbert space. Since it is not one-dimensional, the topological theory contains non-trivial local operators. They are in correspondence with $\overline{\partial}$-closed $(0,p)$-forms on $\mathbb{C}^2$. Again, one might think that all of these, except for the holomorphic functions, are Q-exact, and therefore decouple, but this is not in general true due to the non-compactness of the field space. Let us introduce a special notation $\mathcal{O}_1$ for the operator $\lambda^+\lambda^-\delta^{(4)}(C,\overline{C})$, which we will need in what follows.

[22]For a manifold W glued from two unstable pieces, depending on the situation, it can be natural to define the torsion to be infinity, or zero, or some finite number, by perturbing $\mathcal{L}$ away from the singular case. However, it does not seem to be possible to give any universal meaning to the pairing of non-compact wavefunctions in the ghost Hilbert space.

3.5.2 The Theory On $S^1 \times \Sigma$

Let us illustrate in some examples, how this machinery works. First we compute the invariants for the theory on $S^1 \times \Sigma$, with Σ a closed Riemann surface with no punctures. Then we add punctures and derive the skein relations for the Alexander polynomial. In the whole section 3.5, we typically ignore the overall sign of the torsion, and its dependence on the spin structure.

3.5.2.1 No Punctures

Consider a three-manifold $S^1 \times \Sigma$, where Σ is a Riemann surface of genus g. Let the flat bundle $\mathcal{L}$ have a holonomy $\mathbf{t}$ along the S^1, and no holonomies along the cycles of Σ. We would like to compute the torsion $\tau(\mathbf{t})$ of this manifold. For simplicity, we take $|\mathbf{t}| = 1$.

The topological theory on this manifold reduces to the quantum mechanics of zero modes of the fields on Σ. The components of the gauge field $A^{\pm}$, tangential to Σ, will produce $4g$ fermionic zero modes, which can be grouped into $2g$ pairs of fermions, corresponding to some choice of a- and b-cycles on Σ. For each pair of the fermions, the action is defined with the kinetic operator $i\partial_t + B_0$, where B_0 is the background gauge field in the time direction. If we denote the determinant of this operator by $d(\mathbf{t})$, the gauge fields contribute a factor of $d^{2g}(\mathbf{t})$ to the torsion. The time component of the gauge field $A_0^{\pm}$ together with the Lagrange multiplier λ give two more pairs of fermions with the same action, and hence a factor of $d^2(\mathbf{t})$. Finally, the zero-modes of the superghosts $C^{\pm}$ and $\overline{C}^{\pm}$ give two complex scalars, which contribute a factor of $d^{-4}(\mathbf{t})$. The torsion altogether is $\tau(\mathbf{t}) = d^{2g-2}(\mathbf{t})$. Using the zeta-regularization,[23] one readily computes $d(\mathbf{t}) = \mathbf{t}^{1/2} - \mathbf{t}^{-1/2}$. For the torsion of $S^1 \times \Sigma$, we get

$$\tau(\mathbf{t}) = (\mathbf{t}^{1/2} - \mathbf{t}^{-1/2})^{2g-2} \,. \tag{3.40}$$

[23] One needs to use the identity $\exp\left(-\zeta'(0,a) - \zeta'(0,1-a)\right) = 2\sin(\pi a)$ for the derivative $\partial_s\zeta(s,a)$ of the Hurwitz zeta-function. In the text we ignored the factor of $-i$, which results from this computation, since we are not interested in the overall sign of $\tau(\mathbf{t})$.

Let us derive the same result by a Hilbert space computation. The torsion can be computed by taking the supertrace $\mathrm{Str}_{\mathcal{H}}\mathbf{t}^{\widehat{J}}$ over the Hilbert space, where $\widehat{J}$ is the generator of the $\mathrm{U}(1)_{\mathrm{fl}}$-symmetry. In this formalism, it is obvious that the contribution of a single pair of fermions is indeed $d(\mathbf{t}) = \mathbf{t}^{1/2} - \mathbf{t}^{-1/2}$. The contribution of the superghosts C and $\overline{C}$ can also be easily computed. We set $\mathbf{t} = \exp(i\alpha)$. The quantum mechanics of the complex field C^+ is the theory of a free particle in $\mathbb{R}^2$, and we need to find the trace of the rotation operator $\exp(i\alpha\widehat{J})$ over its Hilbert space,

$$\mathrm{tr}\,\exp\left(i\alpha\widehat{J}\right) = \int \frac{\mathrm{d}^2\vec{p}\,\mathrm{d}^2\vec{x}}{(2\pi)^2}\exp(i\vec{p}\vec{x})\exp(-i\vec{p}\,'\vec{x}) = \frac{1}{4\sin^2(\alpha/2)}\,, \tag{3.41}$$

where $\vec{p}\,'$ is the vector obtained from $\vec{p}$ by a rotation by the angle α. This is equal to $-d^{-2}(\mathbf{t})$, and together with a similar contribution from C^- leads to the correct result $d^{-4}(\mathbf{t})$.

In the computation above, the trace was taken over the whole Hilbert space of the ghost system, and not over the cohomology of Q, since it is not clear in general, what one should mean by this cohomology. However, it is curious to observe that one can obtain the same results by tracing over the non-compact (or over the compact) Dolbeault cohomology. Indeed, $H^{0,0}_{\overline{\partial}}$ is the space of holomorphic functions on $\mathbb{C}^2$, which can be expanded in the basis generated by the monomials 1, C^+, C^-, $(C^+)^2$, *etc.* The trace of $\mathbf{t}^{\widehat{J}}$ over this space can be written as

$$\mathbf{t}^0 + \mathrm{e}^{-\epsilon}(\mathbf{t}^{-1} + \mathbf{t}) + \mathrm{e}^{-2\epsilon}(\mathbf{t}^{-2} + \mathbf{t}^0 + \mathbf{t}^2) + \dots\,, \tag{3.42}$$

where we introduced a regulator $\epsilon > 0$. The sum of this convergent series for $\epsilon \to 0$ is equal to $-d^{-2}(\mathbf{t})$, which is the correct contribution of the ghost system to the torsion. I do not know, if this computation should be taken seriously.

3.5.2.2 Surfaces With Punctures

Next, let us incorporate some line operators. Consider a Riemann surface Σ of genus g with $p \geq 2$ punctures, corresponding to p parallel line operators $L_{\mathbf{t}_1,n_1}\dots L_{\mathbf{t}_p,n_p}$, stretched along the S^1. For consistency, we assume $\mathbf{t}_1\mathbf{t}_2\ldots\mathbf{t}_p = 1$. Let there also be a background holonomy $\mathbf{t}$ around the S^1. We introduce the number $N = \sum_i (n_i - 1/2)$, which measures the total $\mathrm{U}(1)_{\text{fl}}$-charge. For $N = 0$, the configuration is symmetric under the charge conjugation (up to the substitution $\mathbf{t} \to \mathbf{t}^{-1}$ for all the holonomies.)

Due to the presence of line operators, the cohomology $H^0(\Sigma, \mathcal{L})$ is empty, and the Hilbert space does not contain the ghost factor $\mathcal{H}_{\mathrm{gh}}$. However, the cohomology $H^1(\Sigma, \mathcal{L}) \equiv H^1$ is in general non-empty, so there will be $h = \dim H^1$ zero modes of the fermionic gauge field A^+ and h zero modes of the field A^-. Our Lie superalgebra is a direct sum, and correspondingly it is convenient to choose a polarization, in which the modes of A^+ are the creation operators, and the modes of A^- are the annihilation operators. The Hilbert space is

$$(\det H^1)^{-1/2+N/h} \otimes \wedge^\bullet H^1 \,. \tag{3.43}$$

It contains states with charges ranging from $-h/2 + N$ to $h/2 + N$, with

$$\mathsf{N}(q) = \binom{h}{q + h/2 - N} \tag{3.44}$$

states of charge q. (The overall power of $\det H^1$ was chosen so as to ensure that for $N = 0$ the spectrum of $\mathrm{U}(1)_{\text{fl}}$-charges is symmetric.) Taking the supertrace of $\mathbf{t}^{\widehat{J}}$ over this Hilbert space, we find the invariant for $S^1 \times \Sigma$,

$$\tau(\mathbf{t}) = \mathbf{t}^N (\mathbf{t}^{1/2} - \mathbf{t}^{-1/2})^h \,. \tag{3.45}$$

As we will see, $h = -\chi = 2g - 2 + p$. An important special case is that Σ is S^2 with two marked points. Then $h = 0$, the Hilbert space is one-dimensional, and the invariant τ is equal to one, up to an overall power of $\mathbf{t}$.

Let us give a more explicit description of the twisted cohomology for the simple case of $\Sigma \simeq S^2$. In the presence of a singular background field, corresponding to an insertion of a line operator $L_{\mathbf{t},n}$ along some knot K, the behavior of the dynamical fields of the $\mathfrak{psl}(1|1)$ theory near K is determined by a boundary condition, which is described in Appendix B. It says that the superghost fields $C^{\pm}$ should vanish near K, while the components of the fermionic gauge field $A^{\pm}$, perpendicular to K, are allowed to have a singularity, which however has to be better than a pole. This boundary condition is elliptic. The cohomology H^1, therefore, can be represented by $\mathcal{L}$-twisted one-forms, which lie in the kernel of the operator $d + d^*$ on Σ and which near the marked points are less singular than $1/r$. Just for illustration, we can write an explicit formula for these one-forms. For that, pick a complex structure on Σ, and let the marked points be $z_1, \dots, z_p$. The cohomology will be represented by holomorphic $(1,0)$- and antiholomorphic $(0,1)$-forms. Let us write $\mathbf{t}_i = \exp(2\pi i\, a_i)$, with $a_i \in (0,1)$, for the holonomies. (We assume that the bundle $\mathcal{L}$ is unitary.) Note that the sum $\sum a_i$ is a positive integer. Any twisted holomorphic one-form can be written as

$$\omega = \prod_{i=1}^{p} (z - z_i)^{a_i} P(z) \mathrm{d}z \,, \tag{3.46}$$

with some rational function $P(z)$, which is allowed to have simple poles at points z_i, according to our boundary condition. Assuming that infinity is not among the marked points, we should have $\omega \sim \mathrm{d}z/z^2 + o(1/z^2)$ at large z. Writing $P(z)$ as $\sum P_i/(z - z_i)$, the condition at infinity gives $1 + \sum a_i$ linear equations on the coefficients P_i, so the space of twisted holomorphic forms is of dimension $p - 1 - \sum a_i$. Similarly, the space of twisted antiholomorphic forms has dimension $\sum a_i - 1$, and the total dimension of H^1 is $p - 2$, in agreement with the formula $h = -\chi$.

Instead of working with cohomology, it is more convenient to look at the dual homology, which for S^2 with marked points is generated by contours, connecting different punctures[24]. (The differential forms, which behave better than $1/r$ near the punctures, can be integrated over such contours, and the integrals do not change, when the forms are shifted by differentials of functions that vanish at the punctures. Moreover, the pairing between this version

[24] I am grateful to E. Witten for the suggestion to look at the homology and for helpful explanations.

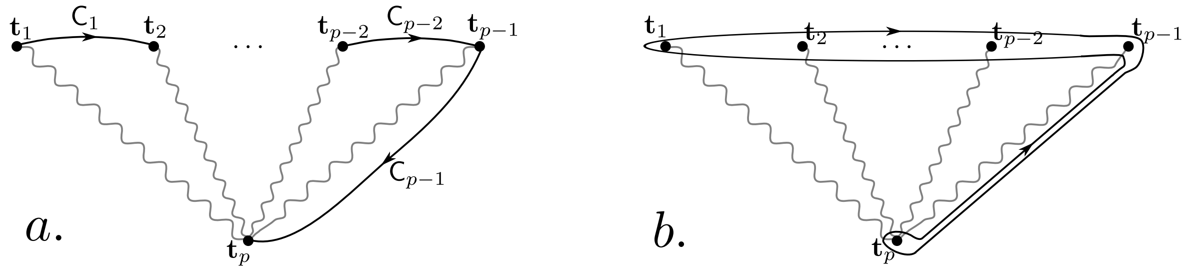

Figure 3.2: a. Marked points, basis contours, and a particular choice of cuts on the p-punctured sphere. Locally-constant sections of $\mathcal{L}$ pick a factor of $\mathbf{t}_i$ upon going counter-clockwise around the i-th puncture. b. This contour is trivial, since it can be pulled off to infinity. This gives a relation $(1-\mathbf{t}_1)\mathsf{C}_1 + \cdots + (1-\mathbf{t}_1 \ldots \mathbf{t}_{p-1})\mathsf{C}_{p-1} = 0$.

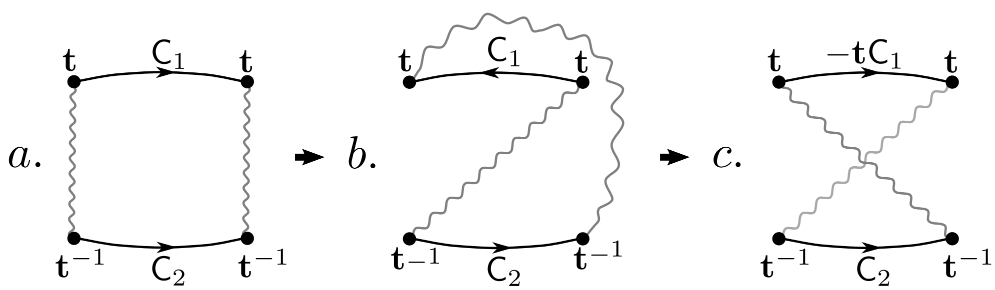

Figure 3.3: a. Two cuts and two basis contours for a four-punctured sphere. b. The result of the braiding transformation. c. We moved the contour C_1 across the cut and reversed its orientation, which produced a factor of $-\mathbf{t}$.

of homology and the twisted cohomology is non-degenerate.) The basis in the homology consists of $p-2$ contours $\mathsf{C}_1, \ldots, \mathsf{C}_{p-2}$, shown on fig. 3.2a. One might think that the contour C_{p-1} should also be included in the basis, but actually it can be expressed in terms of $\mathsf{C}_1, \ldots, \mathsf{C}_{p-2}$, using the relation of fig. 3.2b. On a general Riemann surface, one obtains in the same way that the dimension of the homology is $h = -\chi = p - 2 + 2g$.

It is possible to find modular transformations of states in the Hilbert space. For that, one needs to find the action of large diffeomorphisms on the twisted cohomology H^1, or, equivalently, on the basis contours in the dual homology. To give an example of such argument, we derive the skein relations for the Alexander polynomial[25]. Consider a Riemann

[25]Rather similar contour manipulations are used in [59] to obtain braiding transformations of the states from the CFT free-field representation. The contours in question are then integration contours for the screening fields. In fact, the two computations seem to be directly related, since the screening fields are the CFT currents, which in the Chern-Simons theory correspond to the gauge fields $A^{\pm}$, whose modes are

sphere with four punctures, two of which are labeled by holonomies $\mathbf{t}$, and two by $\mathbf{t}^{-1}$. This configuration arises on the boundary of a solid three-ball with two line operators $L_{\mathbf{t},n}$ inside. We set the parameters n equal to one-half, so that the line operators are expected to have trivial framing transformations and to give rise to the Conway function. (We have to mention once again that our understanding of these line operators is incomplete. This will lead to some uncontrollable minus signs in their expectation values.) The twisted cohomology H^1 on the four-punctured sphere is two-dimensional. A pair of basis contours C_1 and C_2 for the dual homology is shown on fig. 3.3a. We make a large diffeomorphism, which exchanges the two punctures labeled by $\mathbf{t}$. This leads to the configuration of fig. 3.3b. We move the upper cut through the contour C_1. This multiplies C_1 by a factor of $\mathbf{t}$. This brings us to the configuration of fig. 3.3c, where we have also reversed the orientation of the upper contour. The cuts can now be deformed back to the configuration of fig. 3.3a, and we find that the braiding transformation acts on the contours as

$$\begin{pmatrix} \mathsf{C}_1' \\ \mathsf{C}_2' \end{pmatrix} = \begin{pmatrix} -\mathbf{t} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mathsf{C}_1 \\ \mathsf{C}_2 \end{pmatrix} . \tag{3.47}$$

The Hilbert space of the four-punctured sphere, according to eq. (3.43), consists of four states – one of $\mathrm{U}(1)_{\mathrm{fl}}$-charge -1, one of charge $+1$, and two of charge 0. The neutral states are the ones that arise on the boundary of a three-ball with a pair of line operators inside. The state of charge -1 transforms under the braiding by some phase. From eq. (3.43), we would expect this phase to be the inverse square root of the determinant of the matrix in (3.47). The two $\mathrm{U}(1)_{\mathrm{fl}}$-invariant states then transform with the matrix

$$\begin{pmatrix} i\mathbf{t}^{1/2} & 0 \\ 0 & -i\mathbf{t}^{-1/2} \end{pmatrix} . \tag{3.48}$$

the cohomology that we are considering. To make the connection more precise, one needs to switch to the holomorphic polarization.

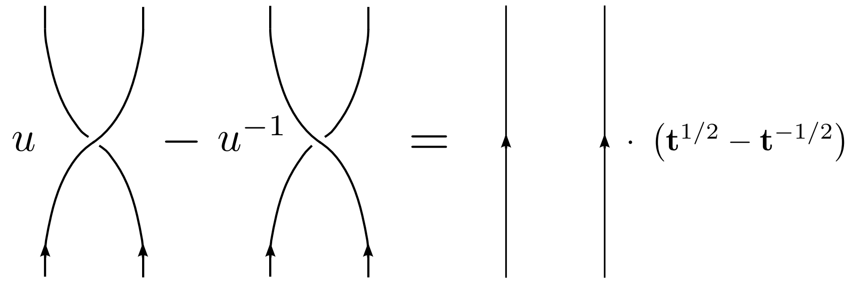

Figure 3.4: A skein relation for the Alexander polynomial. In the canonical framing, $u = 1$.

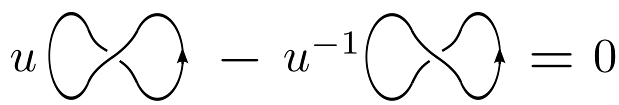

Figure 3.5: The result of closing the strands in the skein relation and using the fact that the Alexander polynomial for a disjoint link is zero. The relation is consistent, if it is written in the vertical framing, and the invariant transforms by a factor of u under a unit change of framing.

Note that the braiding action (3.47) is defined only up to an overall phase, since we could make a constant $\mathrm{U}(1)_{\mathrm{fl}}$ gauge transformation, or, equivalently, could move the cuts on fig. 3.3 around the sphere any number of times. Such a phase, however, would cancel out in 3.48, since the two states of interest are $\mathrm{U}(1)_{\mathrm{fl}}$-invariant.

From (3.48) it follows [4] that the knot invariant satisfies the skein relation of fig.3.4, with $u = i$. (On the way, we made an arbitrary choice of the square root of the determinant of the matrix (3.47). With an opposite choice, we would get $u = -i$.) Initially, we assumed that our line operators have no framing dependence. But now we can see that this would be inconsistent with fig. 3.5, which is obtained from the skein relation by closing the braids and using the fact that the Alexander polynomial of a disjoint link is zero. We are seemingly forced to conclude that our invariant does have a framing dependence, with a framing factor $u = i$. On S^3, there exists a canonical choice of framing, in which the self-linking number of all components of the link is zero. If we bring all the links to this choice of framing, the polynomial would satisfy the skein relation of fig. 3.4, but with $u = 1$. This skein relation, together with a normalization condition, which we derive later in this section, defines the single-variable Alexander polynomial (or the Conway function), as expected.

But the fact that we found a non-trivial framing dependence is rather unsatisfactory. In the dual Seiberg-Witten description, the knot invariant is clearly a polynomial with real (and integral) coefficients, and there can be no factors of i. To get rid of the problem, we have to put an extra factor of i in the braiding transformation of the highest weight state of $\mathrm{U}(1)_{\mathrm{fl}}$-charge -1. This will multiply the matrix (3.48) by i, and make $u = 1$ in the skein relation. It would be desirable to understand the physical origin of this factor.

To be able to compute the multivariable Alexander polynomial, that is, the invariant for multicomponent links, with different components labeled by arbitrary holonomies, one needs two more skein relations [117]. We derive them in Appendix C.

3.5.3 T^2 And Line Operators

In this section, we look more closely on the Hamiltonian quantization of the theory on a two-torus T^2. First we describe the Hilbert space abstractly, and then relate different states to line operators of the theory.

3.5.3.1 The Torus Hilbert Space

Let us fix a basis of cycles a and b on T^2, and denote the corresponding holonomies of the background bundle by $\mathbf{t}_a$ and $\mathbf{t}_b$. Assume first that at least one of the holonomies is non-trivial. In this case, the twisted cohomology $H^\bullet(T^2,\mathcal{L})$ is empty, and the torus Hilbert space $\mathcal{H}_{\mathbf{t}_a,\mathbf{t}_b}$ is one-dimensional. Let us choose some basis vector $|\mathbf{t}_a,\mathbf{t}_b\rangle$ for each of these Hilbert spaces. We pick a normalization such that under any $\mathrm{SL}(2,\mathbb{Z})$ modular transformation $\mathcal{M}$ the vectors map as

$$\mathcal{M}|\mathbf{t}_a,\mathbf{t}_b\rangle = |\mathbf{t}_{\mathcal{M}^{-1}(a)},\mathbf{t}_{\mathcal{M}^{-1}(b)}\rangle\,, \tag{3.49}$$

without any extra factors. Note that the charge conjugation symmetry $\mathcal{C}$ is equivalent to the modular transformation $\mathcal{S}^2$, which flips the signs of both cycles.

A slightly more complicated case is $\mathbf{t}_a = \mathbf{t}_b = 1$. The Hilbert space $\mathcal{H}_{1,1}$ is a product, with one factor being the vector space $\mathcal{H}_{\rm gh}$ of states of the ghosts, which was described before. Another factor comes from the fact that the fermionic gauge fields now have zero modes A_a^+, A_b^+, A_a^- and A_b^-, arising from components of the one-forms $A^\pm$ along the a- or the b-cycle. With a natural choice of polarization, the modes of A^- are the annihilation operators, and the modes of A^+ are the creation operators. The four states in the Hilbert space of the vector fields are

$$|-1\rangle, \quad |0_a\rangle \equiv A_a^+|-1\rangle, \quad |0_b\rangle \equiv A_b^+|-1\rangle, \quad |+1\rangle \equiv A_a^+A_b^+|-1\rangle\,. \tag{3.50}$$

The states $|\pm1\rangle$ are of charge ±1, and are invariant under the modular group $\mathrm{SL}(2,\mathbb{Z})$, since they have nowhere to transform. The two states $|0_a\rangle$ and $|0_b\rangle$ are neutral, and transform under $\mathrm{SL}(2,\mathbb{Z})$ as a doublet.

3.5.3.2 Line Operators $L_{\mathbf{t}\,,n}$

Consider a solid torus with boundary T^2, with cycle a contractible, and put a line operator of type $L_{\mathbf{t}_a,n}$ along the b-cycle inside. Here it is assumed that $\mathbf{t}_a \neq 1$. The operator is taken with the natural framing for loops in the solid torus. We can also turn on a background holonomy $\mathbf{t}_b$. The resulting state lives in $\mathcal{H}_{\mathbf{t}_a,\mathbf{t}_b}$, and we claim that it is

$$|L_{\mathbf{t}_a\,,n},\mathbf{t}_b\rangle = \mathbf{t}_b^{n-1/2}|\mathbf{t}_a,\mathbf{t}_b\rangle\,, \tag{3.51}$$

with a suitable normalization of $|\mathbf{t}_a,\mathbf{t}_b\rangle$. (Note that $\mathbf{t}_b$ is not a parameter of the line operator itself, but is defined by the background bundle, and in particular by the other line operators, linked with the given one.) It is easy to see that both sides of (3.51) depend on n in the same way. (In taking a half-integer power of $\mathbf{t}_b$, we ignored the sign ambiguity, since we generically do not try to fix the overall signs in this section. A more accurate treatment of signs would require keeping track of spin structures.) The non-trivial content of this equation is the statement that $|\mathbf{t}_a,\mathbf{t}_b\rangle$, defined in this way, transforms under the modular group as in (3.49), without any extra factors. For the charge conjugation symmetry $\mathcal{C}$, this is easy to see from the transformation properties of the line operators $L_{\mathbf{t}_a,n}$. For the element $\mathcal{T}$ of the modular group $\mathrm{SL}(2,\mathbb{Z})$, the l.h.s. changes into $\mathbf{t}_a^{-n+1/2}|L_{\mathbf{t}_a,n},\mathbf{t}_b\rangle$, where the factor of $\mathbf{t}_a$ is due to the change of framing. This is again consistent with (3.49). It requires a little more work to see that $|\mathbf{t}_a,\mathbf{t}_b\rangle$ transforms as in (3.49) also for the element $\mathcal{S}$ of $\mathrm{SL}(2,\mathbb{Z})$. Note that a pair of solid tori can be glued together to produce $S^1\times S^2$ with two parallel line operators along the S^1. The gluing identifies the b-cycles of the two tori, and maps the a-cycle of one torus to the $-a$ of the other. This gives a bilinear pairing between the Hilbert spaces $\mathcal{H}_{\mathbf{t}_a^{-1},\mathbf{t}_b}$ and $\mathcal{H}_{\mathbf{t}_a,\mathbf{t}_b}$. In the section 3.5.2.2, we learned that the dimension of the Hilbert space on S^2 with two marked points is equal to one. It follows that, under the bilinear pairing,

$$\big(|\mathbf{t}_a^{-1},\mathbf{t}_b\rangle,|\mathbf{t}_a,\mathbf{t}_b\rangle\big)=\big(|L_{\mathbf{t}_a^{-1},1-n},\mathbf{t}_b\rangle,|L_{\mathbf{t}_a,n},\mathbf{t}_b\rangle\big)=1\,. \tag{3.52}$$

Note that we can apply the elements $\mathcal{CS}$ and $\mathcal{S}$ to the two vectors in this equation, and get the same gluing of the tori. Suppose that the $\mathcal{S}$-transformation of the state $|\mathbf{t}_a,\mathbf{t}_b\rangle$ gives the state $|\mathbf{t}_b^{-1},\mathbf{t}_a\rangle$ with some factor $f(\mathbf{t}_a,\mathbf{t}_b)$. It then follows that $f(\mathbf{t}_a,\mathbf{t}_b)f(\mathbf{t}_a^{-1},\mathbf{t}_b)=1$. The function f should be holomorphic, and can only have zeros or singularities at $\mathbf{t}_a$ or $\mathbf{t}_b$ equal to 0, 1 or infinity. However, 1 is excluded by the equation above. Then, f can only be a monomial in powers of $\mathbf{t}_a$, but this possibility is excluded by the charge conjugation symmetry. We conclude that the vectors $|\mathbf{t}_a,\mathbf{t}_b\rangle$, defined as in (3.51), transform under the modular group according to (3.49). (We did not exclude the possibility of non-trivial $\mathbf{t}$-independent phases in (3.49), but there seem to be no possible candidates for such phases.)

As a check of the modular transformations that we have described, consider a Hopf link, formed by two unknots with some operators $L_{\mathbf{t}_a,n}$ and $L_{\mathbf{t}_b,m}$ in S^3. Up to powers of $\mathbf{t}_\bullet$, which depend on the framings, the invariant for this configuration is equal to the same scalar product (3.52), that is, to one. This is the correct result for the Alexander polynomial of the Hopf link. In the discussion of the Hilbert space of empty S^2, we have defined a local operator $\mathcal{O}_1$. Now we can give it a geometric interpretation[26]: it can be obtained by inserting a small Hopf link of loop operators of type $L_{\mathbf{t},n}$.

3.5.3.3 Other Line Operators

Consider again the same solid torus, and put a line operator L_n along the b-cycle[27]. We first assume that $\mathbf{t}_b \neq 1$, so that the resulting state is $|L_n, \mathbf{t}_b\rangle = \mathbf{t}_b^n\, g(\mathbf{t}_b)|1, \mathbf{t}_b\rangle$, for some holomorphic function $g(\mathbf{t})$. To fix it, note that the invariant for $S^1 \times S^2$ with holonomy $\mathbf{t}_b$ around S^1 can be represented by

$$\tau(S^1 \times S^2, \mathbf{t}_b) = (|L_0, \mathbf{t}_b\rangle, |L_0, \mathbf{t}_b\rangle) = g^2(\mathbf{t}_b)\,. \tag{3.53}$$

On the other hand, it is equal to $(\mathbf{t}_b^{1/2} - \mathbf{t}_b^{-1/2})^{-2}$, so we find that $g(\mathbf{t}_b) = 1/(\mathbf{t}_b^{1/2} - \mathbf{t}_b^{-1/2})$, and therefore

$$|L_n, \mathbf{t}_b\rangle = \frac{1}{1 - \mathbf{t}_b^{-1}}\, \mathbf{t}_b^{n-1/2}\, |1, \mathbf{t}_b\rangle\,. \tag{3.54}$$

Using this, we can find the Milnor torsion for an unknot in S^3. This invariant is equal to

$$(|L_{\mathbf{t},n}, 1\rangle, \mathcal{S}|L_0, \mathbf{t}\rangle) = \frac{1}{\mathbf{t}^{1/2} - \mathbf{t}^{-1/2}}\,, \tag{3.55}$$

which is the correct result. (One can get rid of the half-integer power of $\mathbf{t}$ by choosing a different framing.) Another application is to find the degeneration of the operator $L_{\mathbf{t},n}$ in the limit $\mathbf{t} \to 1$. From (3.51) and (3.54) we find

$$\lim_{\mathbf{t}\to 1} L_{\mathbf{t},n} = L_n - L_{n-1}\,, \quad \mathbf{t}_b \neq 1\,. \tag{3.56}$$

[26]This operator can be given yet another interpretation. Consider cutting out a small three-ball, and gluing in a non-compact space, which is the complement of the three-ball in $\mathbb{R}^3$. The zero-modes of the ghosts cannot freely fluctuate in such geometry, so, this construction produces the desired operator. We can also give arbitrary non-zero vevs $C_0 \in \mathbb{C}^2$ to the fields C in the asymptotic region. This would produce the operator $\lambda^+\lambda^-\delta^{(4)}(C - C_0, \overline{C} - \overline{C}_0)$.

[27]These operators differ from the vacuum just by a factor of $\mathbf{t}^n$, so, we would loose nothing by considering only $n = 0$. But we prefer to keep general n, because it will be helpful, when we come to the U(1|1) theory.

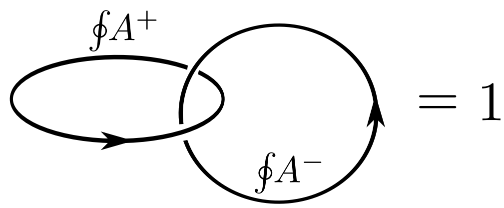

Figure 3.6: The anticommutation relation for the modes of A^+ and A^-, written geometrically.

(This formula is valid only in the sector $\mathbf{t}_b \neq 1$, that is, in presence of a non-trivial holonomy along the line operator.) This relation, when applied to invariants of links in the three-sphere, is known as the Torres formula [118].

Now, consider the case that $\mathbf{t}_b = 1$, so that L_n is inserted inside a solid torus with no background holonomy. The parameter n then does nothing, and the resulting state corresponds just to the empty torus. We want to identify the corresponding state $|\mathrm{vac}\rangle$ in $\mathcal{H}_{1,1}$. In the ghost Hilbert space, it is the vector v_0, as defined in section 3.5.1. In the gauge fields Hilbert space, it is some vector from (3.50), which should have zero charge and should be invariant under the $\mathcal{T}$-transformation. The vector with these properties is $|0_a\rangle$, so we find

$$|L_n, 1\rangle = |\mathrm{vac}\rangle = v_0 \otimes |0_a\rangle \,. \tag{3.57}$$

Let us also give a geometrical interpretation to some other states in $\mathcal{H}_{1,1}$. For that, we simply need to write the modes $A^{\pm}_{a,b}$, used as the creation and annihilation operators in (3.50), as integrals of $A^{\pm}$ over different cycles. The anticommutation relation for these operators is equivalent to a geometrical identity, shown on fig. 3.6. To obtain the state $v_0 \otimes |-1\rangle$, one inserts into the empty solid torus the operator $\oint_b A^-$, effectively undoing the action of A^+_a in (3.50). Similarly, the states $v_0 \otimes |0_b\rangle$ and $v_0 \otimes |+1\rangle$ can be obtained by inserting operators $\oint_b A^+ \oint_b A^-$ and $\oint_a A^+$, respectively. On fig. 3.7, we show the operators needed to create the neutral states, which are obtained by applying transformations $\mathcal{T}^p$ to the $\mathcal{S}$-transform of the vacuum.

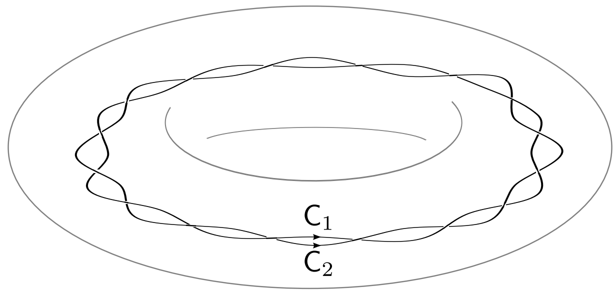

Figure 3.7: By inserting the operator $\oint_{\mathsf{C}_1} A^+ \oint_{\mathsf{C}_2} A^-$, with cycles C_1 and C_2 shown on the figure, one obtains the state $\mathcal{T}^p \mathcal{S}|\mathrm{vac}\rangle = (\mathcal{S}+p)|\mathrm{vac}\rangle$, with p equal to the number of times the cycles wind around each other.

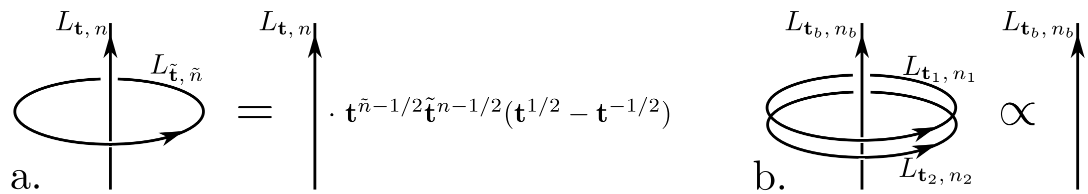

Figure 3.8: a. A relation that follows from one-dimensionality of the Hilbert space of a sphere with two marked points. b. The configurations on the left and on the right are proportional with some coefficient.

3.5.3.4 OPEs of Line Operators

We would like to find the OPEs of our line operators. For products involving the atypical operator L_n, the OPE is trivial: such an operator simply shifts the value of n for the other operators, with which it is multiplied. More interesting are the products of the typical operators $L_{\mathbf{t},\,n}$. To find their OPE, we will need the relation of fig. 3.8a. It can be derived from the fact that the Hilbert space of the two-punctured sphere is one dimensional, and from comparison of the invariants for two linked unknots and for a single unknot in S^3.

To derive the expansion for the product $L_{\mathbf{t}_1,\,n_1} \times L_{\mathbf{t}_2,\,n_2}$, we place two parallel operators along the b-cycle inside a solid torus, and look at the resulting state on the boundary T^2. Assuming that $\mathbf{t}_1\mathbf{t}_2 \neq 1$, the Hilbert space for the torus with this insertion is one-dimensional, and the state created by the insertion of the two operators is proportional to the state created by $L_{\mathbf{t}_1\mathbf{t}_2,\,n_1+n_2}$, with some proportionality coefficient f, which in general can be a holomorphic function of $\mathbf{t}_1$, $\mathbf{t}_2$, and also of the holonomy $\mathbf{t}_b$ of the background bundle along the b-cycle of the torus. To fix this coefficient, consider the configuration on fig. 3.8b. To get from the l.h.s. to the r.h.s., one can apply the relation of fig. 3.8a twice, or one can first fuse $L_{\mathbf{t}_1,\,n_1}$ and $L_{\mathbf{t}_2,\,n_2}$, and then apply the relation once. The two ways of reducing the picture should be equivalent, and this fixes the proportionality factor f, mentioned above, to be equal to $1-\mathbf{t}_b^{-1}$. This leads to the following OPE,

$$L_{\mathbf{t}_1,\,n_1}\times L_{\mathbf{t}_2,\,n_2}=L_{\mathbf{t}_1\mathbf{t}_2,\,n_1+n_2}-L_{\mathbf{t}_1\mathbf{t}_2,\,n_1+n_2-1}\,. \tag{3.58}$$

(Here we absorbed a factor of $\mathbf{t}_b^{-1}$ into the shift $n_1+n_2\to n_1+n_2-1$.)

Now let us turn to the more subtle case of $\mathbf{t}_1\mathbf{t}_2=1$. Let us write the OPE as

$$L_{\mathbf{t},\,n_1}\times L_{\mathbf{t}^{-1},\,n_2}=L_{P,\,n_1+n_2}\,, \tag{3.59}$$

where $L_{P,\,n}$ is some new line operator, to be determined. Again, assume that the operators $L_{\mathbf{t},\,n_1}$ and $L_{\mathbf{t}^{-1},\,n_2}$ lie along the b-cycle of a solid torus. In the sector $\mathbf{t}_b\neq 1$, the Hilbert space on T^2 with this insertion is one-dimensional, and one can apply the same arguments that we used above. The result is

$$L_{P,\,n_1+n_2}=L_{n_1+n_2}-2L_{n_1+n_2-1}+L_{n_1+n_2-2}\,,\quad \mathbf{t}_b\neq 1\,, \tag{3.60}$$

where we applied the relation (3.56) to the OPE (3.58). For $\mathbf{t}_b=1$, the product $L_{\mathbf{t},\,n_1}\times L_{\mathbf{t}^{-1},\,n_2}$ creates some state $|L_{P,\,n_1+n_2},1\rangle$ in the Hilbert space $\mathcal{H}_{1,1}$. In the ghost sector, this state is v_1 (in the notations of sec. 3.5.1), since the singularities in $L_{\mathbf{t},\,n_1}$ and $L_{\mathbf{t}^{-1},\,n_2}$ do not allow the ghosts to fluctuate. We also need to find, what linear combination of the states (3.50) of the fermionic gauge fields is created by $L_{P,\,n_1+n_2}$. For that, we note that gluing a solid torus with the operator $L_{P,\,n_1+n_2}$ to an empty solid torus produces $S^1\times S^2$ with two line operators

$L_{\mathbf{t},n_1}$ and $L_{\mathbf{t}^{-1},n_2}$ along S^1. The corresponding invariant is equal to one, so

$$(|L_{P,n_1+n_2},1\rangle, v_0 \otimes |0_a\rangle) = 1\,. \tag{3.61}$$

On the other hand, if we glue the same tori, but with transformation $\mathcal{S}$ sliced in between, we get a three-sphere with two unlinked unknots $L_{\mathbf{t},n_1}$ and $L_{\mathbf{t}^{-1},n_2}$ inside. The invariant for this configuration is zero, so

$$(|L_{P,n_1+n_2},1\rangle, v_0 \otimes |0_b\rangle) = 0\,. \tag{3.62}$$

From the two equations above, we find that

$$|L_{P,n_1+n_2},1\rangle = v_1 \otimes |0_b\rangle\,. \tag{3.63}$$

Thus, the line operator $L_{P,n}$, which can be obtained from the OPE of $L_{\mathbf{t},n_1}$ and $L_{\mathbf{t}^{-1},n_2}$, is defined by (3.60) in the sector $\mathbf{t}_b \neq 1$, and by (3.63) in the sector $\mathbf{t}_b = 1$.

The set of line operators $L_{\mathbf{t},n}$, L_n and $L_{P,n}$ for different values of n and $\mathbf{t} \neq 1$ forms a closed operator algebra. The OPEs of operators $L_{P,n}$ with themselves and with $L_{\mathbf{t},n}$ follow from (3.58) and (3.59) by associativity.

3.5.3.5 A Comment On Indecomposable Representations

It is convenient to think of the operators $L_{\mathbf{t},n}$ as of Wilson lines, coming from the typical representations of the $\mathfrak{u}(1|1)$ superalgebra, though, of course, this will be literally true only in the U(1|1) theory, and not in $\mathfrak{psl}(1|1)$. The OPE (3.58) of these operators agrees with the tensor product decomposition (3.28) of the typical representations. For the second OPE (3.59) to agree with (3.29), we have to assume that the line operator $L_{P,n}$ is actually the Wilson line for the indecomposable representation P_n, defined in fig. 3.1. This statement makes sense already in the $\mathfrak{psl}(1|1)$ theory, since P_n is also a representation of $\mathfrak{pl}(1|1)$. In

(3.60) we found that $L_{P,n}$ reduces in a special case to a sum of atypical line operators L_n. Comparing this statement to fig. 3.1, we see that it agrees with the decomposition that one would expect to happen for the Wilson loop in representation P_n. (Recall that Wilson loops in reducible indecomposable representations are naively expected to decompose into sums of Wilson loops for irreducible representations.) But we also note that this decomposition does not hold always. Indeed, if it were true also in the sector $\mathbf{t}_b = 1$, the r.h.s. of (3.60) would tell us that $L_{P,n}$ is identically zero in that sector, which is not correct, since for $\mathbf{t}_b = 1$ the operator $L_{P,n}$ actually produces a non-zero state $v_1 \otimes |0_b\rangle$. This state can be obtained by inserting the operators $\oint A^+$ and $\oint A^-$, as shown on fig. 3.7, together with the local operator $\mathcal{O}_1$, to produce the ghost wavefunction v_1. It is tempting to speculate that this combination of operators should arise as some point-splitting regularization of the Wilson loop in representation P_n, but we do not know, how to make this statement precise.

If the typical operators $L_{\mathbf{t},n}$ are thought of as Wilson lines in the typical representations (w, n), then their limit for $\mathbf{t} \to 1$ should correspond to Wilson lines in the (anti-)Kac modules $(0, n)_\pm$, introduced on fig. 3.1. The Torres formula (3.56) then says that the Wilson loops in these indecomposable representations actually reduce to sums of Wilson loops L_n for the irreducible building blocks of the indecomposables. This statement, again, is true in the sector $\mathbf{t}_b \neq 1$. For $\mathbf{t}_b = 1$, one should find some independent way to fix the state in $\mathcal{H}_{1,1}$, produced by the operator $L_{1,n}$. More precisely, since there are two different versions $(0, n)_+$ and $(0, n)_-$ of the limit of (w, n) for $w \to 0$, one would expect that there are two versions $L_{1,n,+}$ and $L_{1,n,-}$ of the operator $\lim_{\mathbf{t}\to 1} L_{\mathbf{t},n}$, which produce two different states in $\mathcal{H}_{1,1}$. We are not sure, what these states are.[28].

The general situation with Wilson loops in reducible indecomposable representations is the following. It is consistent to assume that they do split into sums of Wilson loops L_n, if the background monodromy $\mathbf{t}_b$ along the knot is non-trivial. When $\mathbf{t}_b = 1$, one has to

[28]One possible guess would be that $L_{1,n,+}$ for $\mathbf{t}_b = 1$ is equivalent to $\lambda^- \delta^{(2)}(C^-, \overline{C}^+) \oint A^+$, and similarly for $L_{1,n,-}$, with plus and minus indices interchanged. The reason is that this combination is $\mathrm{U}(1)_{\mathrm{fl}}$-invariant, and depends only on A^+, and not on A^-, as the Wilson line in representation $(0, n)_+$ should.

find some independent way to determine, what states in $\mathcal{H}_{1,1}$ they produce. For P_n, we used the OPE of two typical operators, and for the (anti-)Kac modules $(0,n)_\pm$, one could possibly use the relation to the degeneration limit of the typical operators. But for general indecomposable representations, there seems to be no natural way to determine the state in $\mathcal{H}_{1,1}$, and therefore it does not make much sense to consider such Wilson loops as separate operators at all.

3.5.4 U$(1|1)$ Chern-Simons

Since the U$(1|1)$ theory is the $\mathbb{Z}_k$-orbifold of the $\mathfrak{psl}(1|1)$ Chern-Simons, it is completely straightforward to write out its Hamiltonian quantization, once it is known for $\mathfrak{psl}(1|1)$. For that, one simply needs to restrict to states with U$(1)_{\text{fl}}$-charge divisible by k, and to sum over winding sectors.

For illustration, we consider explicitly the torus Hilbert space. The windings around the two cycles will be labeled by integers w and w', which we take to lie in the range $0 \leq w, w' \leq k-1$. The corresponding holonomies are $\mathbf{t}_w = \exp(2\pi i w/k)$ and $\mathbf{t}_{w'} = \exp(2\pi i w'/k)$. Let $\mathcal{H}_{0,0}$ be the $\mathbb{Z}_k$-invariant subspace of the $\mathfrak{psl}(1|1)$ zero-winding Hilbert space $\mathcal{H}_{1,1}$, and let $\mathcal{H}_{w,w'} \equiv \mathcal{H}_{\mathbf{t}_w,\mathbf{t}_{w'}}$ be the one-dimensional Hilbert spaces in the sectors with windings w and w'. The Hilbert space of the U$(1|1)$ theory on T^2 is the direct sum $\mathcal{H}_{T^2} = \oplus_{w,w'}\mathcal{H}_{w,w'}$.

To find the states that are created by loop operators $L_{w,n}$, L_n and $L_{P,n}$, we take corresponding states in the $\mathfrak{psl}(1|1)$ theory, set the longitudinal holonomy $\mathbf{t}_b$ to be equal to $\exp(2\pi i w'/k)$, and sum over the winding sectors $w' = 0, \dots k-1$. Setting $|w, w'\rangle \equiv |\mathbf{t}_w, \mathbf{t}_{w'}\rangle$, from the equations (3.51), (3.54), (3.57), (3.60) and (3.63) we find

$$\begin{aligned} |L_{w,n}\rangle &= \sum_{w'=0}^{k-1} \exp(2\pi i(n-1/2)w'/k)|w,w'\rangle, \quad w\neq 0\,; \\ |L_n\rangle &= v_0\otimes|0_a\rangle + \frac{1}{2i}\sum_{w'=1}^{k-1}\frac{\exp(2\pi i n w'/k)}{\sin(\pi w'/k)}|0,w'\rangle\,; \\ |L_{P,n}\rangle &= v_1\otimes|0_b\rangle + 2i\sum_{w'=1}^{k-1}\sin(\pi w'/k)\exp(2\pi i(n-1)w'/k)|0,w'\rangle\,. \end{aligned} \tag{3.64}$$

The parameter n is periodic with period k, and we take it to belong to the interval $0\leq n\leq k-1$. If we project out the subspace $\mathcal{H}_{0,0}$, the states $|L_n\rangle$ and $|L_{w,n}\rangle$ with $n=0,\dots k-1$, $w=1\dots k-1$, corresponding to a restricted set of irreducible representations, would form a basis in the remaining Hilbert space. This is what one would have in the ordinary, bosonic Chern-Simons theory. In the full Hilbert space $\mathcal{H}_{T^2}$, the states created by the line operators that we have discussed do not form a basis. More precisely, it is not even clear, what one would mean by such a basis, due to the rather weird nature of $\mathcal{H}_{0,0}$.

The bilinear product of states in U(1|1) theory is $1/k$ times the product in the $\mathfrak{psl}(1|1)$ theory, where the factor $1/k$ comes from eq. (3.37). In particular, we have

$$(|w,w'\rangle,|\widetilde{w},\widetilde{w}'\rangle) = \frac{1}{k}\delta_{w+\widetilde{w}\,\mathrm{mod}\,k,\,0}\,\delta_{w'-\widetilde{w}'\,\mathrm{mod}\,k,\,0}\,, \tag{3.65}$$

and therefore

$$(|L_{w,n}\rangle,|L_{\widetilde{w},\widetilde{n}}\rangle) = \delta_{w+\widetilde{w}\,\mathrm{mod}\,k,\,0}\,\delta_{n+\widetilde{n}-1\,\mathrm{mod}\,k,\,0}\,. \tag{3.66}$$

Let us look at the modular properties of the states, created by the line operators. Under the transformation $\mathcal{T}$, the state $|w,w'\rangle$ transforms into $|w,w'-w\rangle$. The operator $L_{w,n}$ thus picks a phase $\exp(2\pi i w(n-1/2)/k)$. The combination $w(n-1/2)$ is the quadratic Casimir for the typical representation (w,n), and the framing factor that we got is what one would expect from the conformal field theory. The operator L_n is invariant under $\mathcal{T}$. The operator

$L_{P,n}$ does not transform with a simple phase, but rather is shifted as

$$\mathcal{T}|L_{P,n}\rangle = |L_{P,n}\rangle + v_1 \otimes |0_a\rangle \,. \tag{3.67}$$

Geometrically, the reason is that the operator, which defines the state $|0_b\rangle$, is given by integration of A^+ and A^- over the contours of fig. 3.7. Under the $\mathcal{T}$-transformation, the winding number of the two contours changes. We note that in the sector $\mathcal{H}_{0,0}$ the operator $\mathcal{T}$ is not diagonalizable. This is the signature of the logarithmic behavior of the CFT, which presumably corresponds to our Chern-Simons theory.

Under the modular transformation $\mathcal{S}$, the state $|L_{w,n}\rangle$ changes into $\sum_{R'} S^{R'}_{w,n}|L_{R'}\rangle$ with

$$S^{w'n'}_{w,n} = \frac{1}{k}\exp(-2\pi i((n-1/2)w' + (n'-1/2)w)/k)\,, \tag{3.68}$$
$$S^{n'}_{w,n} = \frac{2i\sin(\pi w/k)}{k}\exp(-2\pi i n' w/k)\,. \tag{3.69}$$

The other line operators transform as $\mathcal{S}|L_n\rangle = v_0 \otimes |0_b\rangle + \sum_{R'} S^{R'}_n |L_{R'}\rangle$ and $\mathcal{S}|L_{P,n}\rangle = -v_1 \otimes |0_a\rangle + \sum_{R'} S^{R'}_{P,n}|L_{R'}\rangle$, with

$$S^{w',n'}_{n} = -\frac{1}{2ik\sin(\pi w'/k)}\exp(-2\pi i n w'/k)\,, \tag{3.70}$$
$$S^{w',n'}_{P,n} = -\frac{2i\sin(\pi w'/k)}{k}\exp(-2\pi i(n-1)w'/k) \tag{3.71}$$

Modular transformations very similar to (3.68)-(3.71) were previously derived in the U(1|1) WZW model in [60]. There are, however, some differences. The transformations most similar to ours, but with $\mathcal{H}_{0,0}$ part omitted, are called "naive" in that paper. A slightly different version of transformations is derived using a particular regularization, whose role is essentially to avoid dealing with $\mathcal{H}_{0,0}$. (The Chern-Simons interpretation of this regularization is explained on fig. 11-12 of that paper.) We will not attempt to rederive the modular transformations with the regularization of [60], since in our approach a regularization is not needed.

3.6 Some Generalizations

In this section, we make some brief comments on supergroup Chern-Simons theories other than $\mathfrak{psl}(1|1)$ or $\mathrm{U}(1|1)$. Much of what we are going to say here is a summary of results of [1]. The reason we decided to make this summary is that there, the focus was not on the three-dimensional, but on the analytically-continued version of the theory. Here we would also like to emphasize the importance of coupling to a background flat bundle. Our understanding of the supergroup Chern-Simons theories is very limited, and this section will contain more questions than answers.

3.6.1 Definition And Brane Constructions

To define a supergroup Chern-Simons theory, one needs to choose a complex Lie superalgebra $\mathfrak{g}$, which possesses a non-degenerate invariant bilinear form. The bosonic and the fermionic parts of $\mathfrak{g}$ will be denoted by $\mathfrak{g}_{\bar{0}}$ and $\mathfrak{g}_{\bar{1}}$, respectively. One also needs to choose a real form $\mathfrak{g}_{\bar{0}}^{\mathbb{R}}$ for $\mathfrak{g}_{\bar{0}}$, and a global form $G_{\bar{0}}$ for the corresponding ordinary real Lie group[29]. A real form for the whole superalgebra $\mathfrak{g}$ is not needed. The action of the theory is the usual Chern-Simons action, except that the gauge field is a sum of an ordinary $\mathfrak{g}_{\bar{0}}^{\mathbb{R}}$-valued gauge field and a Grassmann $\mathfrak{g}_{\bar{1}}$-valued one-form. The action is multiplied by a level k, whose quantization condition is determined by the global form $G_{\bar{0}}$, as in the usual Chern-Simons theory. More precisely, the fermionic part of the action can have a global anomaly, in which case the quantization condition for k should be shifted by $1/2$, to cancel the anomaly. To state exactly what we mean by k, we have to specify the regularization scheme. In flat space, one can make the path-integral absolutely convergent by adding a Yang-Mills term, at the expense of breaking the supersymmetry from $\mathcal{N}=4$ to $\mathcal{N}=3$. The Chern-Simons level

[29] One could also imagine defining a complex supergroup Chern-Simons theory, in which the bosonic gauge fields would be valued in the complex Lie algebra $\mathfrak{g}_{\bar{0}}$, and the fermions – in two copies of $\mathfrak{g}_{\bar{1}}$. More generally, it should be possible to define quivers of supergroup Chern-Simons theories, as mentioned in section 2.2.6 of [1].

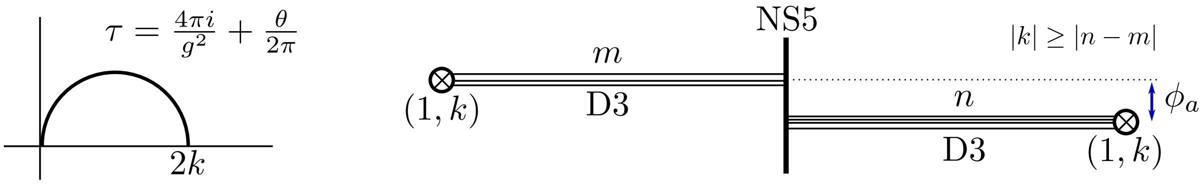

Figure 3.9: Brane construction for an $\mathcal{N} = 4$ Gaiotto-Witten theory. The complexified type IIB string coupling should belong to a semicircle of radius k, as shown on the left. The relative displacement ϕ_a of the two $(1,k)$-branes is the $\mathrm{SU}(2)_X$-triplet of masses. The D3-branes are shown slightly displaced along the direction of the NS5-brane just for clarity of the picture.

then receives no one-loop renormalization. By k we mean this "quantum-corrected" level[30]. An equivalent definition of k is by a brane construction, which is presented below. On a curved space, the correct treatment of the theory at one-loop is not entirely clear (see *e.g.* Appendix E of [1].)

By analogy with the ordinary Chern-Simons theory, one can define an "uncorrected" level k' by

$$k = k' + |h_{\mathfrak{g}}| \operatorname{sign}(k') \,, \tag{3.72}$$

where $h_{\mathfrak{g}}$ is the dual Coxeter number of the superalgebra. One expects that this k' is the level of the current algebra, which one would find in the Hamiltonian quantization of the theory, but that remains to be shown. We note that, while k can be a half-integer, with definition (3.72) k' is always an integer.

Completely analogously to the $\mathfrak{psl}(1|1)$ case, the fermionic part of the gauge symmetry can be globally gauge-fixed. This introduces $\mathfrak{g}_{\bar{1}}$-valued bosonic superghost C and antighost $\overline{C}$, as well as a fermionic $\mathfrak{g}_{\bar{1}}$-valued Lagrangian multiplier λ. Observables of the topological theory are then in the cohomology of a BRST charge Q. This partial gauge-fixing procedure for supergroup Chern-Simons was first described in [24].

As was found in [24], supergroup Chern-Simons theories can be obtained by topological twisting from the $\mathcal{N} = 4$ Chern-Simons-matter theories of [23]. For unitary and orthosym-

[30] Note that we changed notations slightly compared to chapter 2. What we call k here is equal to what was called $\mathcal{K}$ in that chapter.

	0	1	2	3	4	5	6	7	8	9
D3	✓	✓	✓	(✓)	–	–	–	–	–	–
NS5	✓	✓	✓	–	✓	✓	✓	–	–	–
$(1,k)$	✓	✓	✓	–	–	–	–	✓	✓	✓

Table 3.1: Details on the brane configuration of fig. 3.9. The D3-branes span a finite interval in the third direction. The R-symmetry groups $\mathrm{SU}(2)_X$ and $\mathrm{SU}(2)_Y$ act on the directions 456 and 789, respectively.

plectic gauge groups, the latter can be engineered in type IIB string theory by brane constructions [119], [120], [121]. For the $\mathrm{U}(m|n)$ theory, the brane configuration is shown on fig. 3.9. Table 3.1 shows, in which directions the branes are stretched. For eight supersymmetries to be preserved, the complexified type IIB coupling should lie on a semicircle of radius k, as shown on the left of fig. 3.9. The coupling constant can thus be of order $g^2 \sim 1/k$, so, the theory has a well-defined perturbative expansion in $1/k$. The level k for $\mathrm{U}(m|n)$ should satisfy the generalized s-rule condition $|k| \geq |n-m|$, and otherwise the theory breaks supersymmetry [36], [119], [122], [123]. One can also turn on an $\mathrm{SU}(2)_X$-triplet of masses ϕ_a, which correspond in the brane picture to the relative displacement of the $(1,k)$-branes in directions 456, as shown on fig. 3.9. For this deformation to preserve supersymmetry, the generalized s-rule requires $|k| \geq \max(m,n)$.

Let us also discuss brane construction for the orthosymplectic theories. For that, we add an orientifold three-plane to the configuration of fig. 3.6.3. (For a review of orientifold planes, see [25], [26], or section 2.5.1.) Recall that the orientifold three-planes have two $\mathbb{Z}_2$-charges, one of which is usually denoted by plus or minus, and the other by a tilde. Upon crossing a (p,q)-fivebrane, the type of the orientifold changes: if $p \bmod 2 \neq 0$, then plus is exchanged with minus, and if $q \bmod 2 \neq 0$, then the tilde is added or removed. A possible configuration is shown on fig. 3.10. In the interval between the two $(1,k)$-fivebranes, the gauge group is $\mathrm{O}(2m+1)$ on the left and $\mathrm{Sp}(2n)$ on the right. The leftmost and rightmost orientifold planes on the figure have a tilde, if k is even, and do not have it, if k is odd. If the $\widetilde{\mathrm{O}3}^-$-plane would appear on the far right, the theory would have an extra three-dimensional hypermultiplet, coming from the fundamental strings that join the D3-branes and the $\widetilde{\mathrm{O3}}^{-}$-plane. That would give a theory different from what we want. Therefore, we have to take k to be an odd integer. In the $\mathrm{OSp}(2m+1|2n)$ Chern-Simons, we normalize the action to be

$$\frac{k_{\mathfrak{osp}}}{4\pi}\int \mathrm{Str}\left(A\mathrm{d}A+\frac{2}{3}A^3\right)\,, \tag{3.73}$$

where Str is the supertrace in the fundamental representation of the supergroup. Here $k_{\mathfrak{osp}}=k/2$, where the factor of $1/2$ comes from the orientifolding. Let us call a bosonic Chern-Simons term canonically-normalized, if it transforms by arbitrary multiples of 2π under large gauge transformations, assuming that the gauge group is connected and simply-connected. With the normalization (3.73), the level $k_{\mathfrak{osp}}$ multiplies the canonically-normalized Chern-Simons term for the $\mathrm{Sp}(2n)$ subgroup, and twice the canonically-normalized action[31] for $\mathrm{Spin}(2m+1)$. From what we have said about the brane configuration, we see that k is odd, and thus $k_{\mathfrak{osp}}\in 1/2+\mathbb{Z}$. Therefore, the $\mathrm{Sp}(2n)$ part of the bosonic action is anomalous under large gauge transformations. But that precisely compensates for the anomaly for $2m+1$ hypermultiplets in the fundamental of $\mathrm{Sp}(2n)$, so, the theory is well-defined. For any supergroup Chern-Simons theory, one expects the analog of the s-rule to be $|k_{\mathfrak{g}}|\geq|h_{\mathfrak{g}}|$. This is equivalent to the requirement that $k'_{\mathfrak{g}}$, as defined in (3.72), does exist. For $\mathrm{OSp}(2m+1|2n)$, this condition reads as $|k|\geq|2(n-m)+1|$.

For the even orthosymplectic group $\mathrm{OSp}(2m|2n)$, the brane configuration is shown on fig. 3.11. To avoid having an $\widetilde{\mathrm{O3}}^{-}$-plane and an extra hypermultiplet, this time we have to take k to be even, and therefore $k_{\mathfrak{osp}}=k/2$ is an arbitrary integer, consistently with the fact that the fermionic determinant has no global anomaly. The generalized s-rule is $|k|\geq 2|n-m+1|$.

[31] More precisely, this is true for $m>1$. For $m=1$, it is four times the canonically-normalized action.

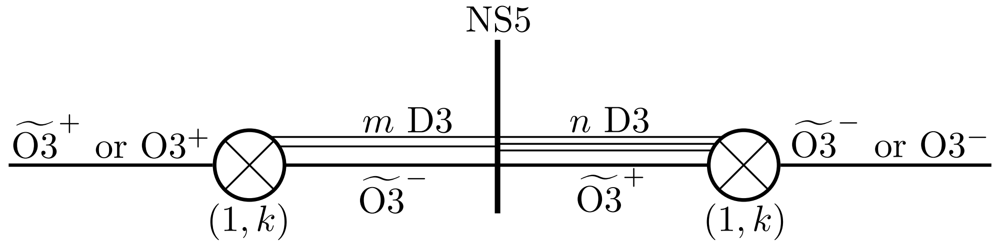

Figure 3.10: The brane construction for the $\mathcal{N}=4$ Gaiotto-Witten theory, which upon twisting would give $\mathrm{OSp}(2m+1|2n)$ Chern-Simons. The leftmost and rightmost orientifold planes are $\widetilde{\mathrm{O3}}^{\pm}$, if k is even, and $O3^{\pm}$, if k is odd.

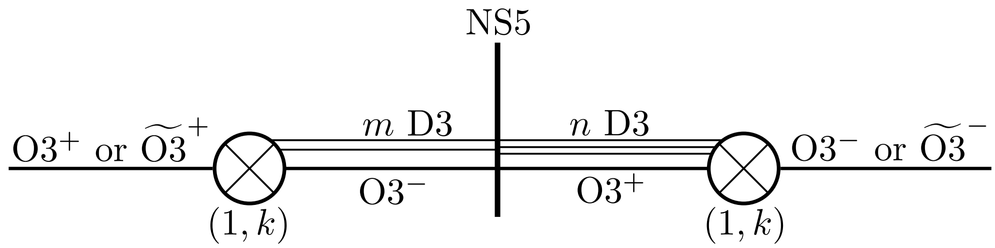

Figure 3.11: The brane construction for the $\mathcal{N}=4$ Gaiotto-Witten theory, which upon twisting would give $\mathrm{OSp}(2m|2n)$ Chern-Simons. The leftmost and rightmost orientifold planes are $O3^{\pm}$, if k is even, and $\widetilde{\mathrm{O3}}^{\pm}$, if k is odd.

3.6.2 Some Properties

Importantly, for Lie superalgebras there exist automorphisms, which commute with the bosonic subalgebra. For the so-called type I superalgebras, the group of these automorphisms is U(1). Type I superalgebras are $\mathfrak{gl}(m|n)$, together with the subquotients $\mathfrak{sl}$ and $\mathfrak{psl}$, and the orthosymplectic superalgebras $\mathfrak{osp}(2|2n)$. The fermionic part $\mathfrak{g}_{\bar{1}}$ for type I decomposes under the action of $\mathfrak{g}_{\bar{0}}$ into a direct sum of two representations. The U(1)-automorphism acts on them with charges ± 1. For superalgebras of type II, which are all the other superalgebras, the relevant group of automorphisms is only $\mathbb{Z}_2$. It acts trivially on $\mathfrak{g}_{\bar{0}}$, and flips the sign of elements of $\mathfrak{g}_{\bar{1}}$. In Chern-Simons theory, one can use these automorphisms to couple the theory to a background flat connection. For type I, this can be a complex flat line bundle, just as we found for $\mathfrak{psl}(1|1)$ and $\mathrm{U}(1|1)$. The partition function of the theory depends on the background complex flat connection holomorphically. In flat space, the imaginary part of the background flat connection can be identified with the $SU(2)_X$-triplet of masses, mentioned above. For a theory with a type II superalgebra, the background bundle can only be a $\mathbb{Z}_2$-bundle. Equivalently, one can assign antiperiodic boundary conditions around various cycles of the three-manifold for the $\mathfrak{g}_{\bar{1}}$-valued fields.

Line observables of the supergroup Chern-Simons theory include Wilson lines in various representations of the supergroup, as well as vortex operators, which are expected to be equivalent to the Wilson lines, at least modulo Q. One can also construct twist line operators by turning on a singular holonomy for the background flat gauge field, as we did in simple examples in the present chapter. For special values of the holonomy, those operators can be equivalent to ordinary vortex operators.

Consider the theory on $\mathbb{R}^3$, or other space with three non-compact directions, and assume that the background flat bundle was turned off. It is then possible to give vevs to the scalar superghost fields C and $\overline{C}$ and to partially Higgs the theory. For example, the $U(m|n)$ gauge supergroup can in this way be reduced down to $U(|n-m|)$. (In the brane picture, this corresponds to recombining a number of D3-branes and taking them away from the NS5-brane in the directions 789.) Since the superghosts appear only in Q-exact terms, this procedure does not change the expectation values of observables in the Q-cohomology. By this Higgsing argument one can see that the expectation values of Wilson loops vanish for almost all representations, except for the maximally-atypical ones. The classes of maximally-atypical representations are in a natural correspondence with representations of $U(|n-m|)$, and the Wilson loops in those representations reduce to Wilson loops of the ordinary, bosonic $U(|n-m|)$ Chern-Simons theory upon Higgsing. Thus, on $\mathbb{R}^3$ the $U(m|n)$ supergroup theory does not produce new knot invariants. (A similar story holds for other supergroups[32].) It is however interesting to turn on a background flat bundle, which in flat space means just a constant $SU(2)_X$-triplet of mass terms. Looking at the brane picture, one would expect that

[32] Almost all supergroup Chern-Simons theories can be reduced in this way to bosonic Chern-Simons. One exception is the series $OSp(2m+1|2n)$, which can be Higgsed only to $OSp(1|2n)$. However, we found in section 2.5.5 that the analytically-continued version of $OSp(1|2n)$ Chern-Simons is dual to the ordinary Chern-Simons with gauge group $O(2n+1)$.

for large ϕ_a the $\mathrm{U}(m|n)$ theory would reduce to $\mathrm{U}(m) \times \mathrm{U}(n)$ Chern-Simons. If this were true, then, in particular, we would have a knot invariant, which interpolates between the $\mathrm{U}(|n-m|)$ and the $\mathrm{U}(m) \times \mathrm{U}(n)$ invariants. This is certainly very puzzling. Unfortunately, we cannot test this in the simple examples considered in this chapter, since the atypical representations of $\mathrm{U}(1|1)$ do not produce non-trivial knot invariants.

On a compact closed three-manifold, the theory has both bosonic and fermionic zero modes. To get a well-defined invariant, one needs to turn on a background flat bundle. The partition function is then a holomorphic function thereof. Alternatively, one can insert loops with vortex operators. As discussed in section 2.3.2, to remove all the zero modes by a single vortex operator, it has to be labeled by a typical weight of the superalgebra.

3.6.3 Dualities

The configuration of fig. 3.9 is clearly similar to the brane contruction for the analytically-continued theory, discussed in Chapter 2. If we moved the $(1,k)$-branes along the third direction away to infinity, we would recover precisely the configuration studied in [6] and in the previous Chapter. In the language of the analytically-continued theory, the role of the $(1,k)$-branes is to choose the real integration contour for the path-integral. Indeed, the fluctuations of the D3-branes in the directions 456 are described in the 4d $\mathcal{N}=4$ Yang-Mills theory by three components of the adjoint-valued scalar field. Upon twisting, those become the imaginary part of the gauge field of the analytically-continued Chern-Simons theory. At the positions of the $(1,k)$-branes these fields are set to zero, which means that we are working with the real integration contour.

Having a brane construction, one can apply various string theory dualities. In the analytically-continued Chern-Simons, it has been shown that the S-dual theory gives a new way to compute the Chern-Simons invariants [6], [74]. One might ask, whether we can obtain anything useful by considering the S-dual of our configuration of fig. 3.9, which is shown on fig. 3.12. Unfortunately, this does not seem to be the case, beyond the duality for the $\mathfrak{psl}(1|1)$

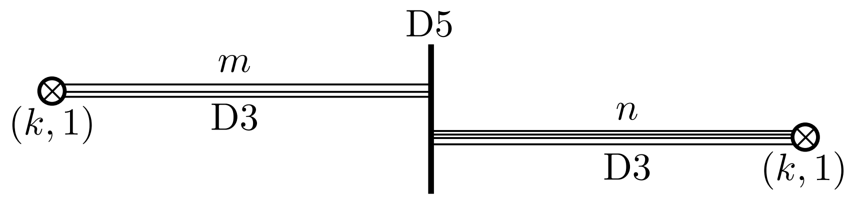

Figure 3.12: The S-dual of the brane configuration, which describes the $\mathrm{U}(m|n)$ Chern-Simons theory.

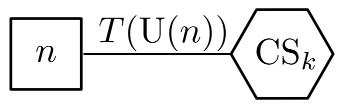

Figure 3.13: The interaction of n D3-branes with a $(k,1)$-brane is described by coupling the D3-brane gauge fields to the $T(\mathrm{U}(n))$ theory via the $\mathrm{U}(n)$ symmetry of the Higgs branch of $T(\mathrm{U}(n))$, and gauging the Coulomb branch of $T(\mathrm{U}(n))$ with a $\mathrm{U}(n)$ Chern-Simons gauge field at level k.

and $\mathrm{U}(1|1)$ theory, which has been considered in previous sections.

The problem is that the S-dual configuration of fig. 3.12 contains D3-branes ending on $(k,1)$-fivebranes. The low energy field theory for such a "tail" has been described in [29], and is shown on fig. 3.13. The $\mathrm{U}(n)$ gauge theory of n D3-branes is coupled to the Higgs branch of the three-dimensional theory $T(\mathrm{U}(n))$, the Coulomb branch of which is gauged by a level k Chern-Simons gauge field. The $T(\mathrm{U}(n))$ theory with non-abelian symmetries of the Coulomb branch gauged does not have a Lagrangian description, and therefore the configuration of fig. 3.12 does not seem to be particularly useful for the purpose of studying supergroup topological invariants.

More precisely, there exists one case, where gauging the Coulomb branch of $T(\mathrm{U}(n))$ is easy [29] — namely, $n=1$. Using the description of this case in [29], one can readily see that the configuration of fig. 3.12 for $m=n=1$ gives the mirror of $\mathrm{U}(1|1)$ Chern-Simons, which was considered in section 3.4.4.

One can alternatively view the mirror transformation of the $\mathrm{U}(m|n)$ theory as follows. We represent the bifundamental hypermultiplet of the $\mathrm{U}(m|n)$ theory as the IR limit of the Coulomb branch of some UV theory, and then couple it to bosonic Chern-Simons gauge fields. The relevant UV theory can be found by replacing the $(1,k)$-fivebranes on fig. 3.9 by

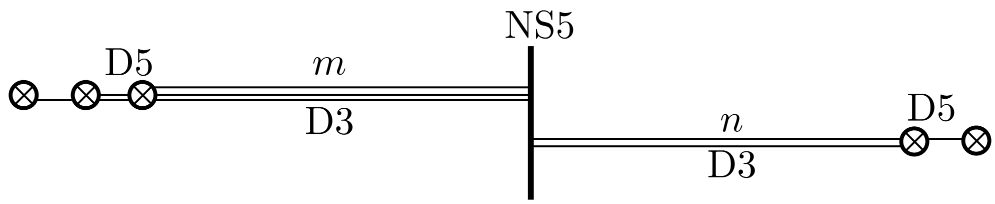

Figure 3.14: A brane configuration, which produces a free $U(n)\times U(m)$ bifundamental hypermultiplet. There are m and n D5-branes on the left and on the right, arranged so as to impose the Dirichlet boundary condition in the 4d $\mathcal{N}=4$ Yang-Mills theory.

a bunch of D5-branes, so as to impose the Dirichlet boundary condition (see fig. 3.14), and then applying the S-duality and making some Hanany-Witten moves. (For $n=m=1$, this procedure would give the $\mathfrak{psl}(1|1)$ theory and its mirror.) The resulting UV theory is given by the quiver of fig. 3.15. It is an "ugly" quiver, in the terminology of [29]. As demonstrated in section 2.4 of that paper, it has nm monopole operators, which in the IR give rise to nm free hypermultiplets, as expected.

Again, this description is not useful for non-abelian supergroup Chern-Simons theories, since the non-abelian symmetry of the Coulomb branch of the quiver emerges only in the IR. We can nevertheless play a game similar to what we did for the single hypermultiplet. We can couple the quiver theory to $n+m-1$ flat GL(1) gauge fields, using the dual photon translation symmetries and FI terms of the UV theory. On the one hand, it is clear from the IR theory that the resulting invariant is a product of nm abelian torsions. On the other hand, it can be computed by solving non-abelian Seiberg-Witten equations[33] for the quiver of fig. 3.15. One expects that the solutions to those equations, in the limit of large FI terms, can be obtained by embedding nm solutions of the abelian equations, so as to reproduce a product of abelian torsions. Since, anyway, this invariant does not produce anything new, we will not consider it in more detail.

There is one last case, where the mirror symmetry can be useful for supergroup Chern-Simons. This is when the level k is equal to one. The reason is that a $(1,1)$-fivebrane can be related by S-duality, say, to a D5-brane, while preserving the NS5-brane in the configuration

[33]Those equations are completely analogous to the abelian ones, and are written out in Appendix3.7.

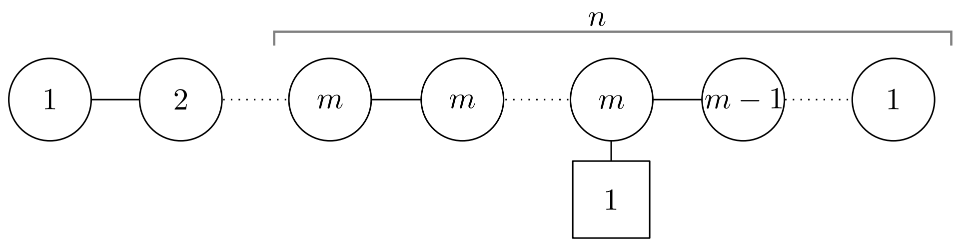

Figure 3.15: A quiver gauge theory, which is obtained by S-duality and a sequence of Hanany-Witten moves from the brane configuration of fig. 3.14.. We follow the notations of [29]: the circles denote unitary gauge groups, the square is the fundamental hypermultiplet, and connecting lines are bifundamental hypermultiplets.

of fig. 3.9. The generalized s-rule requires in this case that $|n-m| \leq 1$. By applying a further S-duality, the theory can be mapped to an $\mathcal{N}=4$ Yang-Mills with no matter or with a single fundamental hypermultiplet. In this way, *e.g.*, the $\mathrm{U}(n|n)$ Chern-Simons theory at level one would be related to the non-abelian $\mathrm{U}(n)$ Seiberg-Witten equations. The problem, however, is that the s-rule in this case does not allow us to turn on a background flat bundle, except for the case of the $\mathrm{U}(1|1)$ theory. Therefore, even if the mirror theory does compute some non-trivial invariant, it will not be computable in the $\mathrm{U}(m|n)$ supergroup Chern-Simons. It is possible that in the orthosymplectic $\mathrm{OSp}(2m+1|2n)$ case the situation is better, and one can turn on a background $\mathbb{Z}_2$-bundle and get a non-trivial duality of invariants, but we will not explore this here.

3.7 Appendix A: Details On The $\mathcal{N}=4$ QCD

Here we describe the fields, the BRST transformations and the Lagrangian for the topologically twisted $\mathcal{N}=4$ SQCD with one fundamental flavor. The bosonic fields of the theory are the gauge field A, the triplet of scalars, which we write as a complex scalar σ and a real field Y_1, and the hypermultiplet scalar fields, which upon twisting become a spinor Z^α. The fermions of the vector multiplet transform in the $(2,2,2)$ representation of the Lorentz and R-symmetry groups, and upon twisting produce fermionic scalars η and $\widetilde{\psi}$ of ghost numbers -1 and $+1$, a one-form ψ of ghost number $+1$, and a two-form χ of ghost number -1. The fermions of the hypermultiplet after twisting remain spinors, and will be denoted by ζ_u (of ghost number $+1$) and ζ_v (of ghost number -1).

The BRST transformations of the fields can be obtained by dimensional reduction[34] from the formulas of Chapter 2,

$$\begin{aligned}
&\delta A = \psi\,, \quad \delta\sigma = 0\,, \quad \delta\overline{\sigma} = \eta\,, \quad \delta Y_1 = \widetilde{\psi}\,, \quad \delta Z = \zeta_u \\
&\delta\eta = i[\sigma,\overline{\sigma}]\,, \quad \delta\psi = -d_A\sigma\,, \quad \delta\widetilde{\psi} = i[\sigma, Y_1]\,, \quad \delta\chi = H\,, \quad \delta\zeta_u = i\sigma Z\,, \quad \delta\zeta_v = f\,.
\end{aligned}$$

Here $\mu = iZ^\alpha \otimes \overline{Z}_\beta \sigma_\alpha^\beta$ is the moment map, and H and f are auxiliary fields. The equations of motion set $H = F + \star\left(d_A Y_1 + \frac{1}{2}\mu - \frac{1}{2}e^2\phi\right)$ and $f = \not{D}Z + iY_1 Z$, and the Seiberg-Witten equations are

$$F + \star\left(d_A Y_1 + \frac{1}{2}\mu - \frac{1}{2}e^2\phi\right) = 0\,, \tag{3.74}$$

$$\not{D}Z + iY_1 Z = 0\,. \tag{3.75}$$

The FI one-form ϕ is valued in the center of $\mathfrak{u}(n)$. Here are a couple of useful identities,

$$\begin{aligned}
&\int \mathrm{d}^3x\sqrt{\gamma}\left(D_i\overline{Z}_\alpha D^i Z^\alpha + \overline{Z}_\alpha\left(Y_1^2 + \frac{1}{4}R\right)Z^\alpha\right) \\
= &\int \mathrm{d}^3x\sqrt{\gamma}\,|f|^2 - \int \mathrm{d}^3x\sqrt{\gamma}\,\mathrm{tr}\left(Y_1 D_i\mu^i\right) + \int \mathrm{d}^3x\,\mathrm{tr}\left(F\wedge\mu\right)\,,
\end{aligned} \tag{3.76}$$

$$\begin{aligned}
&\int \mathrm{d}^3x\sqrt{\gamma}\,\mathrm{tr}\left(\frac{1}{2}F_{ij}^2 + (D_iY)^2 + \frac{1}{4}(\mu_i - e^2\phi_i)^2 - e^2\phi^i D_i Y_1\right) \\
= &\int \mathrm{tr}\left(H\wedge\star H\right) + \int \mathrm{tr}\left(F\wedge e^2\phi - F\wedge\mu\right) + \int \mathrm{d}^3x\sqrt{\gamma}\,\mathrm{tr}\left(Y_1 D_i\mu^i\right).
\end{aligned} \tag{3.77}$$

where R is the scalar curvature. These identities allow to rewrite the SQCD action in the form (3.15)-(3.16). (Our normalization of the coupling constant is such that the gauge field kinetic term is $\int \mathrm{tr}\,F_{ij}^2/4\pi e^2$.)

The action of the twisted theory in general contains the term $Y_1 D^i\phi_i$, which breaks the

[34] Our notations here are slightly different from Chapter 2 in that here the adjoint-valued fields are Hermitian. The covariant differential is $d_A = d + iA$.

$\mathrm{SU}(2)_Y$-symmetry. This, in fact, is the same term that we saw in section 3.2.1 in the electric theory. If $\mathrm{d} \star \phi = 0$, the symmetry is restored. For an irreducible solution of the Seiberg-Witten equations, one then has $Y_1 = \sigma = \overline{\sigma} = 0$, and the equations (3.74)-(3.75) can be simplified to (3.19). For a more general ϕ, the field Y_1 is non-zero and can be found by applying d_A to the equation (3.74).

We focused on the QCD with one fundamental flavor, but this twisting procedure generalizes in an obvious way to an arbitrary quiver theory with vector multiplets and hypermultiplets.

3.8 Appendix B: Boundary Conditions Near A Line Operator

In general, in giving a definition of a disorder operator, one needs to specify the boundary conditions for the fields near the singularity, to ensure that the Hamiltonian in presence of the operator remains self-adjoint. (A closely related condition is that in Euclidean signature the kinetic operator of the fields should remain Fredholm.) For that, the boundary conditions should satisfy two requirements. First, to verify the Hermiticity of the Hamiltonian, one integrates by parts, and the boundary term should vanish. Second, the boundary conditions should set to zero half of the modes near the boundary. Here we would like to sketch these boundary conditions for our disorder operators $L_{\mathbf{t},n}$, since we use them explicitly in section 3.5.2.2. (Note that sometimes in similar problems there exist families of possible boundary conditions, and this leads to important physical consequences [124], [125]. In our case, nothing like this happens.)

We consider an operator $L_{\mathbf{t},n}$, stretched along a straight line in $\mathbb{R}^3$. The coordinate along the operator will be denoted by t and will be treated as time, and the polar coordinates in the transverse plane will be denoted by r and θ. For the background gauge field, we choose the gauge in which B is zero, but fields with positive $\mathrm{U}(1)_{\mathrm{fl}}$-charge are multiplied by $\mathbf{t}$ in going around the operator. We assume $\mathbf{t}$ to be unimodular and write it as $\mathbf{t} = \exp(2\pi i a)$, with $a \in (0,1)$.

For the scalar field C^+, we want to impose a boundary condition with which the two-dimensional Laplacian Δ would be self-adjoint. The field can be expanded in modes of different angular momentum ℓ, valued in $a + \mathbb{Z}$. Near $r = 0$, the modes with angular momentum ℓ behave like $r^{\pm|\ell|}$. We impose the boundary condition $C|_{r\to 0} = 0$. It actually implies that C vanishes at least as $r^{\min(a,1-a)}$. This boundary condition has the required properties.

The Q transformations act as

$$\delta A_0 = -\partial_t C\,, \quad \delta A = -\mathrm{d}C\,, \quad \delta\overline{C} = \lambda\,, \tag{3.78}$$

where we separated the fermionic gauge field into its time component A_0 and components A in the transverse plane. The boundary condition for the fermions, which is compatible with vanishing of C and with Q-invariance, is to require that λ and A_0 vanish at $r = 0$, and that A is less singular than $1/r$, in an orthonormal frame. Then, in fact, the fields λ, A_0 and rA vanish at least as $r^{\min(a,1-a)}$, and are square-integrable. The fermionic Hamiltonian is the operator $d + d^*$ in two dimensions, acting on the field $\mathcal{A} = A_0 + A + \star\lambda$, where $\star$ is the 2d Hodge operator. It is easy to see (on the physical level of rigor) that with our boundary condition the Hamiltonian is self-adjoint. If $z = r\exp(i\theta)$ is the complex coordinate, then the operator reduces to

$$\begin{pmatrix} 0 & -\overline{\partial} \\ \partial & 0 \end{pmatrix}, \tag{3.79}$$

acting on the doublet $((A_0 + i\lambda)/2, A_z)$, plus a similar operator for the other pair of fields $((A_0 - i\lambda)/2, A_{\overline{z}})$. In verifying the Hermiticity of this operator, the boundary term in the integration by parts vanishes. The boundary condition sets to zero a minimal possible number of modes, so one expects that the operator is not only Hermitian, but is self-adjoint.

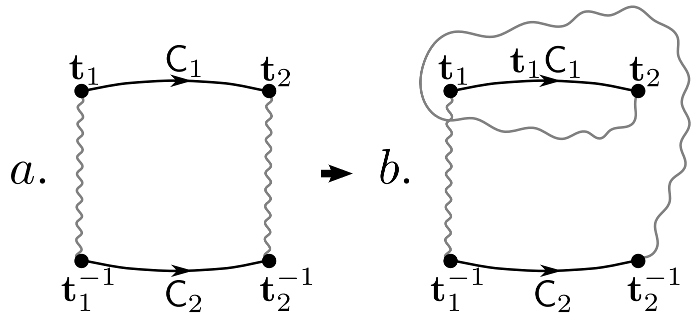

Figure 3.16: a. Two cuts and two basis contours for a four-punctured sphere. b. The result of the braiding transformation. The contour C_1 got multiplied by $\mathbf{t}_1$ in crossing the left cut. It will also get a factor of $\mathbf{t}_2$, when the right cut is moved back to its place.

3.9 Appendix C: Skein Relations For The Multivariable Alexander Polynomial

Here we derive two skein relations for the multivariable Alexander polynomial, which are known [117] to define it completely, together with the skein relation of fig. 3.4, the normalization (3.55), the formula of fig. 3.8a, and the fact that the invariant is zero for a disjoint link.

Consider the case of two strands, labeled by holonomies $\mathbf{t}_1$ and $\mathbf{t}_2$. The sphere with four punctures $\mathbf{t}_1$, $\mathbf{t}_1^{-1}$, $\mathbf{t}_2$, $\mathbf{t}_2^{-1}$ and two basis contours is shown on fig. 3.16a. Upon performing a braiding transformation, which brings the marked point $\mathbf{t}_2$ around the point $\mathbf{t}_1$, we arrive at the picture on fig. 3.16b. The contour C_1 gets a factor of $\mathbf{t}_1$ in crossing the left cut. To compare to fig. 3.16a, we also need to move the right cut back to its place. That will multiply the contour C_1 by a factor of $\mathbf{t}_2$. Overall, the transformation acts on the contours as

$$\begin{pmatrix} \mathsf{C}_1' \\ \mathsf{C}_2' \end{pmatrix} = \begin{pmatrix} \mathbf{t}_1\mathbf{t}_2 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mathsf{C}_1 \\ \mathsf{C}_2 \end{pmatrix}. \tag{3.80}$$

Therefore, the state of $\mathrm{U}(1)_{\mathrm{fl}}$-charge -1 transforms by a factor $(\mathbf{t}_1\mathbf{t}_2)^{-1/2}$, and the two $\mathrm{U}(1)_{\mathrm{fl}}$-neutral states are transformed by a matrix with eigenvalues $(\mathbf{t}_1\mathbf{t}_2)^{1/2}$ and $(\mathbf{t}_1\mathbf{t}_2)^{-1/2}$.

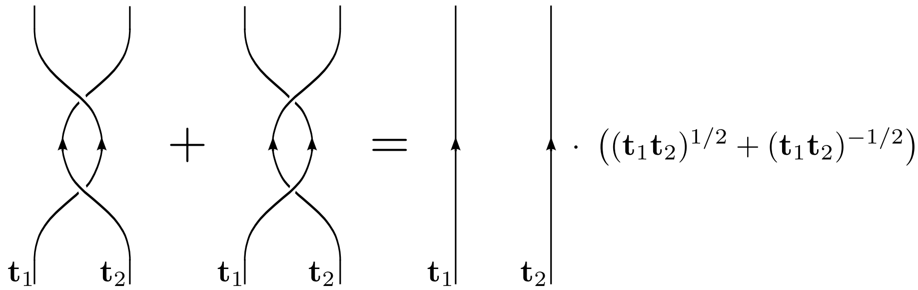

Figure 3.17: A skein relation for the multivariable Alexander polynomial.

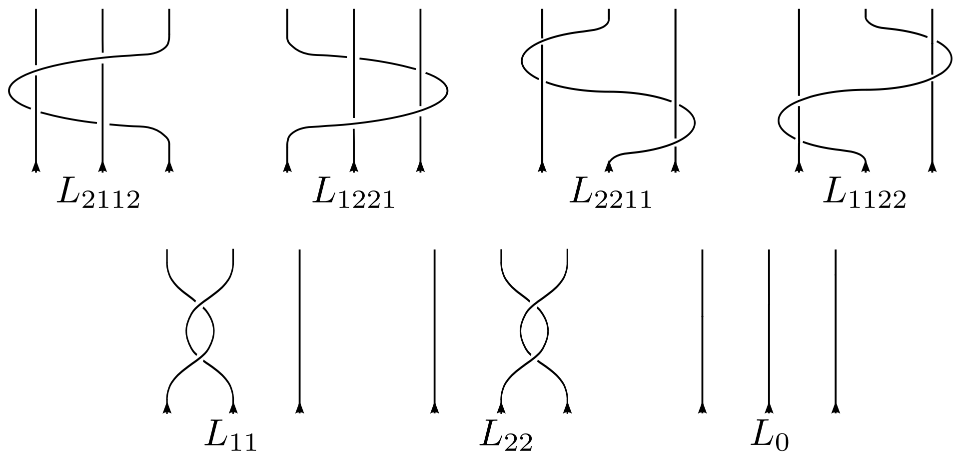

Figure 3.18: Braids for the 3-strand skein relation.

(In taking a square root, we made a choice of sign such that the resulting skein relation for $\mathbf{t}_1 = \mathbf{t}_2$ is consistent with fig. 3.4.) The skein relation that we find is shown on fig. 3.17.

To completely characterize the multivariable Alexander polynomial, one more skein relation is needed [117]. It relates seven three-strand configurations, shown on fig. 3.18. The existence of this skein relation follows from the fact that the dimension of the $\mathrm{U}(1)_{\mathrm{fl}}$-invariant subspace of the Hilbert space of the six-punctured sphere, according to (3.44), is $\binom{4}{2} = 6$.

We need to find the action of the braiding transformations of fig. 3.18 on the four contours that generate the twisted homology of the six-punctured sphere. For example, let us consider the link L_{2211}. The basis contours and the result of the braiding transformation are shown on fig. 3.19. On the contours C_1 and C_2 we put cross-marks at some points, which are not moved in the transformation. At these points the one-form, which is being integrated over

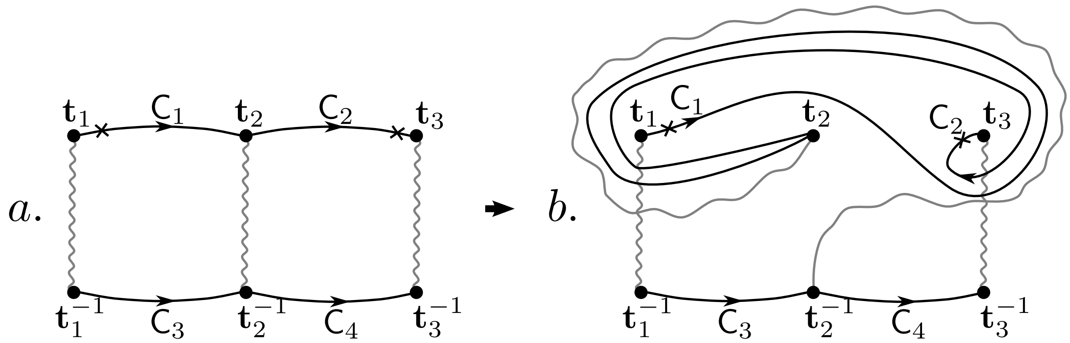

Figure 3.19: a. A particular choice of cuts and basis contours for the six-punctured sphere. b. The result of the braiding transformation, corresponding to the link L_{2211}.

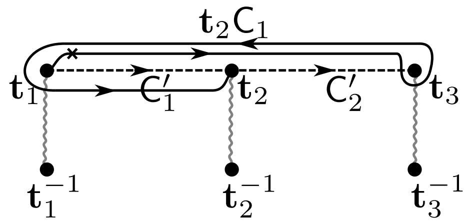

Figure 3.20: Contour $\mathbf{t}_2\mathsf{C}_1$, which comes from C_1 of fig. 3.19 after moving the middle cut back to its place. We show the new basis contours C'_1 and C'_2 by dashed lines.

the contour, is taken on the first sheet, and on the rest of the contour it is defined by analytic continuation. The first step in comparing figures 3.19b and 3.19a is to bring the middle cut back to its place. On the way, it will cross the contours C_1 and C_2, and that will multiply them by $\mathbf{t}_2$. On fig. 3.20, we show the contour $\mathbf{t}_2\mathsf{C}_2$. We need to expand it in the new basis C'_1 and C'_2, which is shown by dashed lines. We start comparing the contours from the cross-mark, and add a factor of $\mathbf{t}^{-1}$ each time we cross a cut counterclockwise around a puncture $\mathbf{t}$. We find

$$\mathbf{t}_2\mathsf{C}_1 = \mathsf{C}'_1 + \mathsf{C}'_2 + \mathbf{t}_3^{-1}(-\mathsf{C}'_2 - \mathsf{C}'_1 + \mathbf{t}_1^{-1}\mathsf{C}'_1)\,. \tag{3.81}$$

Repeating the same steps for C_2, and for each link from fig. 3.18, we find the braiding matrices

$$
B_{2112} = \mathbf{t}_3 \begin{pmatrix} 1 & 0 \\ \mathbf{t}_2(\mathbf{t}_1 - 1) & \mathbf{t}_1\mathbf{t}_2 \end{pmatrix}, \qquad B_{1221} = \mathbf{t}_1 \begin{pmatrix} \mathbf{t}_2\mathbf{t}_3 & \mathbf{t}_3 - 1 \\ 0 & 1 \end{pmatrix},
$$
$$
B_{2211} = \mathbf{t}_2 \begin{pmatrix} \mathbf{t}_1 & \mathbf{t}_1(1-\mathbf{t}_3) \\ 1-\mathbf{t}_1 & 1-\mathbf{t}_1+\mathbf{t}_1\mathbf{t}_3 \end{pmatrix}, \quad B_{1122} = \mathbf{t}_2 \begin{pmatrix} 1-\mathbf{t}_3+\mathbf{t}_1\mathbf{t}_3 & \mathbf{t}_2^{-1}(1-\mathbf{t}_3) \\ \mathbf{t}_2\mathbf{t}_3(1-\mathbf{t}_1) & \mathbf{t}_3 \end{pmatrix},
$$
$$
B_{11} = \begin{pmatrix} \mathbf{t}_1\mathbf{t}_2 & 0 \\ \mathbf{t}_2(1-\mathbf{t}_1) & 1 \end{pmatrix}, \quad B_{22} = \begin{pmatrix} 1 & 1-\mathbf{t}_3 \\ 0 & \mathbf{t}_2\mathbf{t}_3 \end{pmatrix}, \quad B_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{3.82}
$$

Here we defined the matrices by $(\mathsf{C}_1', \mathsf{C}_2')^T = B(\mathsf{C}_1, \mathsf{C}_2)^T$. The contours C_3 and C_4 are transformed trivially.

Let $a^+_{1,2,3,4}$ be the four creation operators, obtained by integrating the fermionic gauge field A^+ over the corresponding contours. The Hilbert space of the six-punctured sphere contains one state of charge -2, from which we build the other states by applying $a^+_\bullet$. The six neutral states, which we are interested in, are

$$
\begin{aligned}
&a_1^+ a_2^+ |{-2}\rangle, \\
&a_1^+ a_3^+ |{-2}\rangle, \quad a_2^+ a_3^+ |{-2}\rangle, \\
&a_1^+ a_4^+ |{-2}\rangle, \quad a_2^+ a_4^+ |{-2}\rangle, \\
&a_3^+ a_4^+ |{-2}\rangle\,.
\end{aligned} \tag{3.83}
$$

The highest weight state $|{-2}\rangle$ transforms under braiding by a factor $\det^{-1/2} B$, and therefore so does the state $a_3^+ a_4^+ |{-2}\rangle$. The state $a_1^+ a_2^+ |{-2}\rangle$ transforms by a factor $\det^{1/2} B$. The states in the second and the third lines of (3.83) transform in doublets by the matrix $B \det^{-1/2} B$. In total, for each braiding transformation, the 6×6 braiding matrix has $1 + 1 + 4 = 6$ independent matrix elements. We can collect them in a 7×6 matrix, in which the rows correspond to the diagrams of fig. 3.18. The null-vector of this matrix will give us the skein relation. Let us set $g_\pm(\mathbf{t}) = \mathbf{t}^{1/2} \pm \mathbf{t}^{-1/2}$. Using the explicit expressions for the braiding matrices (3.82), one finds the skein relation to be

$$\begin{aligned} &g_+(\mathbf{t}_1)g_-(\mathbf{t}_2)L_{2112} - g_-(\mathbf{t}_2)g_+(\mathbf{t}_3)L_{1221} + g_-(\mathbf{t}_1\mathbf{t}_3^{-1})(L_{2211} + L_{1122}) \\ &+ g_-(\mathbf{t}_2\mathbf{t}_3\mathbf{t}_1^{-1})g_+(\mathbf{t}_3)L_{11} - g_-(\mathbf{t}_1\mathbf{t}_2\mathbf{t}_3^{-1})g_+(\mathbf{t}_1)L_{22} + g_-(\mathbf{t}_1^2\mathbf{t}_3^{-2})L_0 = 0\,. \end{aligned} \tag{3.84}$$

This, indeed, is the correct skein relation for the multivariable Alexander polynomial. Together with other relations and normalization conditions that we have found, it fixes the knot invariant completely [117]. We should note, however, that we did not explain, how to properly choose the square root of the determinant of the braiding matrix in the transformation of the highest weight state. Thus, our derivation does not allow to unambiguously fix relative signs of different diagrams in the skein relation.

Bibliography

[1] V. Mikhaylov and E. Witten, “Branes And Supergroups,” arXiv:1410.1175. Accepted for publication in Commun. Math. Phys.

[2] V. Mikhaylov, “Analytic Torsion, 3d Mirror Symmetry And Supergroup Chern-Simons Theories,” arXiv:1505.03130.

[3] A. S. Schwarz, “The partition function of degenerate quadratic functional and Ray-Singer invariants,” Lett. Math. Phys. 2 (1977/78), no. 3, 247–252.

[4] E. Witten, “Quantum Field Theory and the Jones Polynomial,” Commun. Math. Phys. **121**, 351 (1989).

[5] M. Khovanov, “A Categorification Of The Jones Polynomial,” Duke. Math. J. **101** (2000) 359-426.

[6] E. Witten, “Fivebranes and Knots,” Quantum Topology **1** (2012) 1-137, arXiv:1101.3216.

[7] E. Witten, “Topological Quantum Field Theory,” Commun. Math. Phys. **117** (1988) 353.

[8] S. Gukov, A. S. Schwarz, and C. Vafa, “Khovanov-Rozansky Homology And Topological Strings,” Lett. Math. Phys. **74** (2005) 53-74, arXiv:hep-th/0412243.

[9] N. Seiberg and E. Witten, “Electric - magnetic duality, monopole condensation, and confinement in N=2 supersymmetric Yang-Mills theory,” Nucl. Phys. B **426**, 19 (1994), [Nucl. Phys. B **430**, 485 (1994)], [hep-th/9407087].

[10] N. Seiberg and E. Witten, “Monopoles, duality and chiral symmetry breaking in N=2 supersymmetric QCD,” Nucl. Phys. B **431**, 484 (1994) [hep-th/9408099].

[11] E. Witten, “Monopoles and four manifolds,” Math. Res. Lett. **1**, 769 (1994), [hep-th/9411102].

[12] E. Witten, “Analytic Continuation Of Chern-Simons Theory,” arXiv:1001.2933, in *Chern-Simons Gauge Theory: Twenty Years After*, ed. J. E. Andersen et. al. (American Mathematical Society, 2011), pp. 347-446.

[13] E. Witten, “A New Look At The Path Integral Of Quantum Mechanics,” arXiv:1009.6032.

[14] A. Kapustin and E. Witten, “Electric-Magnetic Duality and the Geometric Langlands Program,” Commun. Numb. Th. Phys. **1** (2007) 1-236, arXiv:hep-th/0604151.

[15] S. C. Blumen, “On the $\mathrm{U}_q(osp(1|2n))$ and $\mathrm{U}_{-q}(so(2n+1))$ Uncolored Quantum Link Invariants,” J. Knot Theory Ramifications **19** (2010), no. 3, 335–353.

[16] R. B. Zhang, “Finite-Dimensional Representations of $\mathrm{U}_q(\mathfrak{osp}(1|2n))$ and its Connection with Quantum $\mathfrak{so}(2n+1)$,” Lett. Math. Phys. **25** (1992) 317-325.

[17] I. P. Ennes, P. Ramadevi, A. V. Ramallo and J. M. Sanchez de Santos, “Duality in $\mathfrak{osp}(1|2)$ Conformal Field Theory and Link Invariants,” Int. J. Mod. Phys. A **13**, 2931 (1998), arXiv:hep-th/9709068.

[18] P. Ozsvath, J. Rasmussen, and Z. Szabo, “Odd Khovanov Homology,” arXiv:0710.4300.

[19] A. P. Ellis and A. D. Lauda, “An Odd Categorification Of Quantum $sl(2)$,” arXiv:1307.7816.

[20] G. Meng and C. H. Taubes, "SW=Milnor Torsion," Math. Res. Lett. **3** (1996) 661-674.

[21] D. Gaiotto and E. Witten, "Supersymmetric Boundary Conditions in $\mathcal{N}=4$ Super Yang-Mills Theory," J. Stat. Phys. **135** (2009) 789-855. arXiv:0804.2902.

[22] H. Ooguri and C. Vafa, "Knot Invariants And Topological Strings," Nucl. Phys. **B577** (2000) 419, arXiv:hep-th/9912123.

[23] D. Gaiotto and E. Witten, "Janus Configurations, Chern-Simons Couplings, and the Theta-Angle in $\mathcal{N}=4$ Super Yang-Mills Theory," JHEP **1006**, 097 (2010) arXiv:0804.2907.

[24] A. Kapustin and N. Saulina, "Chern-Simons-Rozansky-Witten Topological Field Theory," Nucl. Phys. B **823**, 403 (2009) arXiv:0904.1447.

[25] E. Witten, "Baryons and Branes in Anti-de Sitter Space," JHEP **9807**, 006 (1998), arXiv:hep-th/9805112.

[26] A. Hanany and B. Kol, "On Orientifolds, Discrete Torsion, Branes and M Theory," JHEP **0006**, 013 (2000), arXiv:hep-th/0003025.

[27] B. Feng and A. Hanany, "Mirror Symmetry by O3 Planes," JHEP **0011**, 033 (2000), arXiv:hep-th/0004092.

[28] N. J. Evans, C. V. Johnson and A. D. Shapere, "Orientifolds, Branes, and Duality of 4-D Gauge Theories," Nucl. Phys. B **505**, 251 (1997), arXiv:hep-th/9703210.

[29] D. Gaiotto and E. Witten, "S-Duality Of Boundary Conditions In $\mathcal{N}=4$ Super Yang-Mills Theory," arXiv:0807.3720.

[30] R. B. Zhang, "Braid Group Representations Arising From Quantum Supergroups With Arbitrary q And Link Polynomials," J. Math. Phys. **33** (1992) 3918-30.

[31] J. R. Links, M. D. Gould, and R. B. Zhang, "Quantum Supergroups, Link Polynomials, and Representations Of The Braid Generator," Rev. Math. Phys. **5** (1993) 345-61.

[32] R. B. Zhang, "Three-Manifold Invariants Arising From $\mathcal{U}_q(\mathfrak{osp}(1|2))$," Modern Phys. Lett. **A9** (1994) 1453-65.

[33] S. Blumen, "Topological Invariants Of Three-Manifolds From $\mathcal{U}_q(\mathfrak{osp}(1|2n))$," arXiv:math/0209218.

[34] J. H. Horne, "Skein Relations and Wilson Loops in Chern-Simons Gauge Theory," Nucl. Phys. B **334**, 669 (1990).

[35] C. Vafa, "Brane / anti-brane systems and U(N|M) supergroup," hep-th/0101218.

[36] A. Hanany and E. Witten, "Type IIB Superstrings, BPS Monopoles, and Three-Dimensional Gauge Dynamics," Nucl. Phys. B **492**, 152 (1997), arXiv:hep-th/9611230.

[37] V.G. Kac, "Representations of Classical Lie Superalgebras," in *Differential Geometrical Methods in Mathematical Physics, II* , (Proc. Conf., Univ. Bonn, Bonn, 1977) 597-626, Lecture Notes in Math., 676, Springer, Berlin, 1978.

[38] T. Quella and V. Schomerus, "Superspace Conformal Field Theory," J. Phys. A **46**, 494010 (2013), arXiv:1307.7724.

[39] L. Frappat, P. Sorba and A. Sciarrino, "Dictionary on Lie Superalgebras," arXiv:hep-th/9607161.

[40] A. Baha Balantekin and I. Bars, "Dimension and Character Formulas for Lie Supergroups," J. Math. Phys. **22**, 1149 (1981).

[41] R.J. Farmer, P.D. Jarvis, "Representations of Orthosymplectic Superalgebras. II. Young Diagrams and Weight Space Techniques." J. Phys. A 17 (1984), no. 12, 2365–2387.

[42] C.J. Cummins, R.C. King, "Young diagrams, Supercharacters of OSp($M|N$) and Modification Rules." J. Phys. A 20 (1987), no. 11, 3103–3120.

[43] V. G. Kac, M. Wakimoto, "Integrable Highest Weight Modules over Affine Superalgebras and Number Theory," *Lie Theory and Geometry*, 415–456, Progr. Math., **123**, Birkhäuser Boston, 1994.

[44] V. Serganova, "On the Superdimension of an Irreducible Representation of a Basic Classical Lie Superalgebra," *Supersymmetry in Mathematics and Physics*, Lecture Notes in Mathematics 2011, pp 253-273.

[45] J. E. Humphreys, "Introduction to Lie Algebras and Representation Theory," Graduate Texts in Mathematics, 9 (Springer-Verlag, New York-Berlin, 1978).

[46] V. G. Kac, "Laplace Operators of Infinite-Dimensional Lie Algebras and Theta Functions," Proc. Nat. Acad. Sci. U.S.A. **81** (1984), no. 2, Phys. Sci., 645–647.

[47] C. Gruson, V. Serganova, "Cohomology of Generalized Supergrassmannians and Character Formulae for Basic Classical Lie Superalgebras," Proc. Lond. Math. Soc. **(3)** 101 (2010), no. 3, 852–892.

[48] S. Elitzur, G. W. Moore, A. Schwimmer and N. Seiberg, "Remarks on the Canonical Quantization of the Chern-Simons-Witten Theory," Nucl. Phys. B **326**, 108 (1989).

[49] C. Beasley, "Localization for Wilson Loops in Chern-Simons Theory," arXiv:0911.2687.

[50] I.B. Penkov, "An Introduction to Geometric Representation Theory for Complex Simple Lie Superalgebras," *Differential Geometric Methods in Theoretical Physics* (Shumen, 1984), 89–106, World Sci. Publishing, Singapore, 1986.

[51] V. Mikhaylov, "On the Solutions of Generalized Bogomolny Equations," JHEP **1205** (2012) 112, arXiv:1202.4848.

[52] A. Kapustin, "Wilson-'t Hooft Operators in Four-Dimensional Gauge Theories and S-Duality," Phys. Rev. D **74**, 025005 (2006), arXiv:hep-th/0501015.

[53] S. Gukov and E. Witten, "Gauge Theory, Ramification, And The Geometric Langlands Program," in J. David et. al. (eds.) *Current Developments In Mathematics* (International Press, Somerville Mass., 2008), arXiv:hep-th/0612073.

[54] S. Gukov and E. Witten, "Rigid Surface Operators," Adv. Theor. Math. Phys. **14** (2010), arXiv:0804.1561.

[55] M. F. Atiyah, "On Framings Of 3-Manifolds," Topology **29** (1990), no. 1, 1–7.

[56] N. Srivastava, "Spectral Sparsification And Restricted Invertibility" (Ph.D. thesis, Yale University, 2010), [http://www.cs.yale.edu/homes/srivastava/dissertation.pdf].

[57] N. R. Constable, R. C. Myers, O. Tafjord, "Fuzzy Funnels: Nonabelian Brane Intersections," arXiv:hep-th/0105035.

[58] E. Witten, "Introduction To Cohomological Field Theories," Int. J. Mod. Phys. **A6** (1991) 2775-92.

[59] L. Rozansky and H. Saleur, "Quantum Field Theory for the Multivariable Alexander-Conway Polynomial," Nucl. Phys. B **376**, 461 (1992).

[60] L. Rozansky and H. Saleur, "S and T Matrices For The Super U(1|1) WZW Model," Nucl. Phys. **B389** (1993) 365-423.

[61] L. Rozansky and H. Saleur, "Reidemeister Torsion, The Alexander Polynomial, and U(1|1) Chern-Simons Theory," arXiv:hep-th/9209073, J.Geom.Phys. **13** (1994) 105-123.

[62] A. N. Redlich, "Parity Violation and Gauge Noninvariance of the Effective Gauge Field Action in Three-Dimensions," Phys. Rev. D **29**, 2366 (1984).

[63] L. Alvarez-Gaume and E. Witten, "Gravitational Anomalies," Nucl. Phys. B **234**, 269 (1984).

[64] E. Witten, "An SU(2) Anomaly," Phys. Lett. **B117** (1982) 324-8.

[65] M. Z. Hasan and C. L. Kane, "Topological Insulators," Rev. Mod. Phys. **82** (2010) 3045.

[66] X.-L. Qi and S.-C. Zhang, "Topological Insulators And Superconductors," Rev. Mod. Phys. **83** (2011) 1057.

[67] M. Z. Hasan and J. E. Moore, "Three-Dimensional Topological Insulators," Ann. Rev. Condens. Matt. Phys. **2** (2011) 55.

[68] A. Bernevig (with T. L. Hughes), *Topological Insulators And Topological Superconductors* (Princeton University Press, 2013).

[69] M. F. Atiyah, V. Patodi, I. Singer, "Spectral Asymmetry And Riemannian Geometry I-III," Math. Proc. Cambridge Philos. Soc. **77** (1975) 43, **78** (1975) 405, **79** (1976) 71.

[70] R. Mazzeo, E. Witten, "The Nahm Pole Boundary Condition," arXiv:1311.3167 [math].

[71] V. Rittenberg, M. Scheunert, "A Remarkable Connection Between the Representations of the Lie Superalgebras $\mathfrak{osp}(1|2n)$ and the Lie Algebras $\mathfrak{o}(2n+1)$," Comm. Math. Phys. **83** (1982) 1-9.

[72] C. Vafa and E. Witten, "A Strong coupling test of S duality," Nucl. Phys. B **431**, 3 (1994), arXiv:hep-th/9408074.

[73] E. Witten, "Supersymmetric Index in Four-Dimensional Gauge Theories," Adv. Theor. Math. Phys. **5**, 841 (2002), arXiv:hep-th/0006010.

[74] D. Gaiotto and E. Witten, "Knot Invariants From Four-Dimensional Gauge Theory," arXiv:1106.4789.

[75] C. Candu, T. Creutzig, V. Mitev, and V. Schomerus, “Cohomological Reduction Of Sigma Models,” arXiv:1001.1344.

[76] E. Witten, “Khovanov Homology And Gauge Theory,” arXiv:1108.3103.

[77] E. Witten, “Two Lectures On The Jones Polynomial And Khovanov Homology,” arXiv:1401.6996 [math.GT].

[78] E. Witten, “Supersymmetry and Morse Theory,” J. Diff. Geom. **17**, 661 (1982).

[79] D. Gaiotto, G. Moore, and E. Witten, to appear.

[80] K. Hori, S. Katz, A. Klemm, R. Pandharipande, R. Thomas, C. Vafa, R. Vakil and E. Zaslow, *Mirror Symmetry*, (Clay Mathematics Monographs, 1).

[81] E. Gorsky, S. Gukov, and M. Stosic, “Quaduply-Graded Colored Homology Of Knots,” arXiv:1304.3481.

[82] C. G. Callan, Jr. and J. A. Harvey, “Anomalies And Fermion Zero-Modes On Strings And Domain Walls,” Nucl. Phys. **B250** (1985) 427-436.

[83] O. DeWolfe, D. Z. Freedman and H. Ooguri, “Holography and Defect Conformal Field Theories,” Phys. Rev. D **66**, 025009 (2002), arXiv:hep-th/0111135.

[84] B.D. Boe, J.R. Kujawa, D.K. Nakano, “Cohomology and Support Varieties for Lie Superalgebras,” Trans. Amer. Math. Soc. 362 (2010), no. 12, 6551-6590.

[85] D. Bar-Natan and E. Witten, “Perturbative Expansion of Chern-Simons Theory with Noncompact Gauge Group,” Commun. Math. Phys. **141**, 423 (1991).

[86] E. Witten, “On Holomorphic Factorization Of WZW And Coset Models,” Comm. Math. Phys. **144** (1992) 189-212.

[87] D. S. Freed, R. E. Gompf, “Computer Calculation of Witten’s 3-Manifold Invariant,” Comm. Math. Phys. **141** (1991), no. 1, 79–117.

[88] L.I. Nicolaescu, "The Reidemeister torsion of 3-manifolds," de Gruyter Studies in Mathematics, 30. Walter de Gruyter, Berlin, 2003.

[89] V. Turaev, "Introduction to combinatorial torsions," Notes taken by Felix Schlenk. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2001.

[90] L. Rozansky, "A Contribution of the Trivial Connection to the Jones Polynomial and Witten's Invariant of 3-d Manifolds. 1.," Commun. Math. Phys. **175**, 275 (1996) [hep-th/9401061].

[91] L. Rozansky and E. Witten, "HyperKahler geometry and invariants of three manifolds," Selecta Math. **3**, 401 (1997) [hep-th/9612216].

[92] V.G. Turaev, "A combinatorial formulation for the Seiberg-Witten invariants of 3-manifolds," Math. Res. Lett. 5 (1998), no. 5, 583–598.

[93] M. Blau and G. Thompson, "On the relationship between the Rozansky-Witten and the three-dimensional Seiberg-Witten invariants," Adv. Theor. Math. Phys. **5**, 483 (2002) [hep-th/0006244].

[94] M. Mariño and G. W. Moore, "Three manifold topology and the Donaldson-Witten partition function," Nucl. Phys. B **547**, 569 (1999) [hep-th/9811214].

[95] S. K. Donaldson, "Topological field theories and formulae of Casson and Meng-Taubes," Geom. Topol. Monographs **2**, 87 (1999) [math/9911248].

[96] M. Hutchings, Y.-J. Lee, "Circle-valued Morse theory and Reidemeister torsion," Geom. Topol. 3 (1999), 369–396.

[97] M. Braverman, T. Kappeler, "Refined analytic torsion," J. Differential Geom. 78 (2008), no. 2, 193–267, arxiv:math/0505537.

[98] R. Dijkgraaf and E. Witten, "Topological Gauge Theories and Group Cohomology," Commun. Math. Phys. **129**, 393 (1990).

[99] V.G. Turaev, "Euler structures, nonsingular vector fields, and Reidemeister-type torsions," (Russian) Izv. Akad. Nauk SSSR Ser. Mat. 53 (1989), no. 3, 607–643, 672; translation in Math. USSR-Izv. 34 (1990), no. 3, 627–662.

[100] Rung-Tzung Huang, "Refined analytic torsion: comparison theorems and examples," Illinois J. Math. 51 (2007), no. 4, 1309–1327.

[101] J.W. Milnor, "A Duality Theorem for Reidemeister Torsion," Ann.Math., 76 (1962), p. 137-147.

[102] V.G. Turaev, "Reidemeister Torsion in Knot Theory," Uspekhi Mat. Nauk 41 (1986), no. 1(247), 97–147, 240.

[103] G. Gotz, T. Quella and V. Schomerus, "Representation Theory of $\mathfrak{sl}(2|1)$," J. Algebra **312**, 829 (2007) [hep-th/0504234].

[104] K. A. Intriligator and N. Seiberg, "Mirror symmetry in three-dimensional gauge theories," Phys. Lett. B **387**, 513 (1996), [hep-th/9607207].

[105] A. Kapustin and M. J. Strassler, "On mirror symmetry in three-dimensional Abelian gauge theories," JHEP **9904**, 021 (1999), [hep-th/9902033].

[106] L. Nicolaescu, "Seiberg-Witten invariants of rational homology 3-spheres," Commun. Contemp. Math. 6 (2004), no. 6, 833–866.

[107] Yuhan Lim, "Seiberg-Witten invariants for 3-manifolds in the case b1=0 or 1," Pacific J. Math. 195 (2000), no. 1, 179–204.

[108] N. Seiberg and E. Witten, "Gauge dynamics and compactification to three-dimensions," in *Saclay 1996, The mathematical beauty of physics* 333-366 [hep-th/9607163].

[109] C.H. Taubes, "Casson's invariant and gauge theory," J. Differential Geom. 31 (1990), no. 2, 547–599.

[110] M.F. Atiyah, L. Jeffrey, "Topological Lagrangians and cohomology," J. Geom. Phys. 7 (1990), no. 1, 119–136.

[111] M. Blau and G. Thompson, "N=2 topological gauge theory, the Euler characteristic of moduli spaces, and the Casson invariant," Commun. Math. Phys. **152**, 41 (1993) [hep-th/9112012].

[112] G. W. Moore and E. Witten, "Integration over the u plane in Donaldson theory," Adv. Theor. Math. Phys. **1**, 298 (1997) [hep-th/9709193].

[113] E. Witten, "SL(2,Z) action on three-dimensional conformal field theories with Abelian symmetry," In *Shifman, M. (ed.) et al.: From fields to strings, vol. 2* 1173-1200 [hep-th/0307041].

[114] T. Creutzig and P. B. Ronne, "The GL(1|1)-symplectic fermion correspondence," Nucl. Phys. B **815**, 95 (2009) [arXiv:0812.2835 [hep-th]].

[115] V.G. Turaev, "Torsion invariants of Spinc-structures on 3-manifolds," Math. Res. Lett. 4 (1997), no. 5, 679–695.

[116] S. Axelrod, S. Della Pietra and E. Witten, "Geometric Quantization of Chern-Simons Gauge Theory," J. Diff. Geom. **33**, 787 (1991).

[117] J. Murakami, "On local relations to determine the multi-variable Alexander polynomial of colored links," Knots 90 (Osaka, 1990), 455–464, de Gruyter, Berlin, 1992.

[118] G. Torres, "On the Alexander polynomial," Ann. of Math. (2) 57, (1953).

[119] T. Kitao, K. Ohta and N. Ohta, "Three-dimensional gauge dynamics from brane configurations with (p,q) - five-brane," Nucl. Phys. B **539**, 79 (1999) [hep-th/9808111].

[120] O. Aharony, O. Bergman, D. L. Jafferis and J. Maldacena, "N=6 superconformal Chern-Simons-matter theories, M2-branes and their gravity duals," JHEP **0810**, 091 (2008) [arXiv:0806.1218 [hep-th]].

[121] O. Aharony, O. Bergman and D. L. Jafferis, "Fractional M2-branes," JHEP **0811**, 043 (2008) [arXiv:0807.4924 [hep-th]].

[122] O. Bergman, A. Hanany, A. Karch and B. Kol, "Branes and supersymmetry breaking in three-dimensional gauge theories," JHEP **9910**, 036 (1999) [hep-th/9908075].

[123] K. Ohta, "Supersymmetric index and s rule for type IIB branes," JHEP **9910**, 006 (1999) [hep-th/9908120].

[124] H. Yamagishi, "The Fermion Monopole System Reexamined," Phys. Rev. D **27**, 2383 (1983).

[125] B. Grossman, "Does A Dyon Leak?" Phys. Rev. Lett. **50**, 464 (1983).