Dissertation

Defects in Supersymmetric Gauge Theory and Integrable Lattice Models

Toshihiro Ota

Department of Physics, Osaka University, Toyonaka, Osaka 560-0043, Japan

Abstract

abstract

Acknowledgements

I would like to acknowledge kind support and help from many people, without those this thesis would not have been completed. Firstly, I would like to thank my Ph.D. advisor Koji Hashimoto who has given me great opportunity to start my study of string theory at Osaka university. In particular, in the first three years of my Ph.D. program I could learn a lot of things through discussion and collaboration with him. I owe deep gratitude to my collaborators Kazunobu Maruyoshi and Junya Yagi. This thesis is largely based on the joint work with them, and I have been enjoying discussion and chat with them every time. They brought me to the fascinating subject at the intersection of supersymmetric gauge theory, string theory, and integrable system. I am also grateful to Yasuhiro Akutsu, Norihiro Iizuka, Tetsuya Onogi and Satoshi Yamaguchi for careful reading of this thesis and giving me useful comments.

I would like to say how much of a pleasure it has been being a member of iTHEMS at RIKEN. In particular, I really appreciate Tetsuo Hatsuda, who is the director of iTHEMS, and Masato Taki, who was a member of iTHEMS and is also my collaborator on machine learning. Their advice, encouragement, and arrangement were always helping me a lot during my stay at iTHEMS. Over the past two years, I have been really benefited from stimulating discussions with colleagues in various fields of science at iTHEMS. Such an experience has made my Ph.D. course quite fruitful and way more enjoyable. I am especially indebted to mathematicians at iTHEMS; Hokuto Konno, Genki Ouchi, Kenta Sato and Masaki Taniguchi for communicating with and telling me lots of mathematical ideas. Inspiring discussions and interesting conversation with them were a great help of this thesis.

I would like to thank Tomoko Iwanami, Chikako Ota and Hitomi Wada, the secretaries at iTHEMS, and Mika Umetani and Satomi Suzuki, the secretaries of the particle physics theory group at Osaka university for arranging a lot of procedures so that I could do my research at RIKEN. They were always kind and I was helped many times. During my Ph.D. study, I have been supported in part by RIKEN Junior Research Associate program and by JSPS fellowships for Young Scientists. Finally, I wish to thank my parents for their understanding, love and constant support.

Contents

1 Introduction and Summary

2 Integrability from extra dimensions

Throughout this paper, we discuss correspondences between a certain class of supersymmetric gauge theories and integrable lattice models:

$$\mathcal{I}_{\mathsf{T}_{4d}[\mathsf{L}_{2d}]} = Z_{\mathsf{L}_{2d}[\mathsf{T}_{4d}]}, \tag{2.1}$$

where the left-hand side is the supersymmetric index of a four-dimensional gauge theory T_{4d} and the right-hand side denotes the statistical partition function of the corresponding integrable lattice model L_{2d} specified by the four-dimensional theory T_{4d}. As it turns out, the correspondence emerges from TQFT with extra dimensions [1]. The aim of this section is to give the step-by-step explanation of the above correspondence from an elementary level. Besides, to explain these, we will start by clarifying the precise meaning of the terms TQFT, lattice model, and integrability.

2.1 Preliminaries

In section 2.2, we will introduce lattice model as discrete quantum field theory (QFT). So let us start from the general settings of QFT. We first briefly review one-dimensional QFT, which is nothing but quantum mechanics, and then extend the discussion to general quantum field theory in $(d+1)$-dimensions. Though we do not have a complete definition of general QFT, a special class of QFT which is called *topological* QFT is axiomatized by Atiyah [2].

2.1.1 What is quantum field theory?

Let us begin with one-dimensional QFT, as known as quantum mechanics (QM). Suppose we have a one-dimensional manifold[1], which may be thought of as "time" $M^1 =$ interval or

[1]Throughout the paper, we assume all the manifold is smooth and oriented, unless otherwise stated.

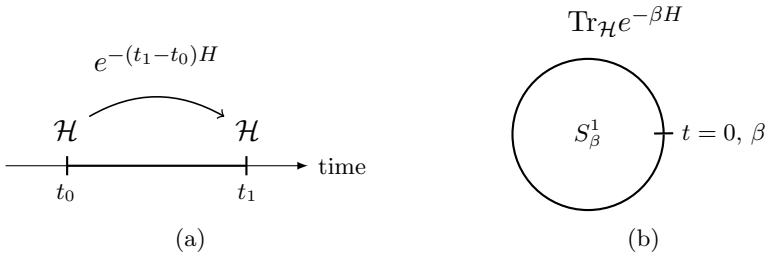


Figure 1: Two most crucial properties of quantum mechanics.

S^1 (or $\mathbb{R}$), and data to specify a QM:

- $\mathcal{H}$: a vector space (Hilbert space/state space),
- $\mathcal{O}$: a set of self-adjoint operators (observables, acting on $\mathcal{H}$),
- H : an operator called *Hamiltonian*,

where $H \in \mathcal{O}$, and the state space $\mathcal{H}$ is a finite or an infinite dimensional $\mathbb{C}$-vector space. We sometimes need to allow a set of operators $\mathcal{O}$ include not self-adjoint operators, but for simplicity we do not care about that at this moment. In undergraduate, we have learned that one should consider the Schrödinger equation and find a good basis in $\mathcal{H}$ which diagonalizes the Hamiltonian.

As is well known, the procedure of solving the Schrödinger equation, at least formally, leads to an expression of time evolution of states by a linear map[2] e^{-tH} in quantum mechanics, and the partition function in many-body statistical mechanics. In turn, we have two most crucial properties of QM. Given an interval $M^1 = [t_0, t_1]$, we have state vectors at the endpoints of the interval and a time evolution between the states:

$$\underset{t_0 \qquad t_1}{\longmapsto} \quad \rightsquigarrow \quad e^{-(t_1-t_0)H} : \mathcal{H} \longrightarrow \mathcal{H} \ , \tag{2.2}$$

and given a circle $M^1 = S^1_\beta$, we get a number called *partition function*:

$$S^1_\beta \quad \rightsquigarrow \quad \mathrm{Tr}_{\mathcal{H}} e^{-\beta H} \ . \tag{2.3}$$

In addition, from physical facts it is quite reasonable to assume that the time evolution and the partition function are compatible with cutting and gluing intervals and S^1s. This simply means that a time evolution from time t_0 to t_1 followed by another time evolution from t_1 to t_2 is equal to a single time evolution from t_0 to t_2, etc. Summarizing, the properties of quantum mechanics, or one-dimensional QFT, are rephrased by the language of given one-dimensional manifold M^1 (see figure 1):

[2] We here do not get into the argument of Euclideanization.

- Given an interval, QM produces state vectors at the endpoints and a linear map between them.
- Given an S^1, QM produces a number.
- These two properties are compatible with cutting and gluing intervals and S^1s.

Thus, in the abstract we conclude that quantum mechanics is a gadget satisfying the above properties for each given one-dimensional manifold M^1.

From these observations of quantum mechanics, we wish to extend the discussion to general $(d+1)$-dimensional QFT. The starting point of defining a $(d+1)$-dimensional QFT is the choice of a $(d+1)$-dimensional manifold M^{d+1}, which has d spatial directions and one "time" direction. For most QFTs the manifold M^{d+1} is viewed as a Riemannian manifold with a smooth metric on it. As already noticed, we will mostly consider a positive definite Riemannian metric, QFT on which is usually referred to as an Euclidean QFT, and hence precisely there is no notion of "time" in such a theory, at least globally. The manifold M^{d+1} may or may not have boundaries. In case it does have boundaries some additional information is needed at the boundaries to define the QFT. In addition to a Riemannian metric, depending on the situation one wants to consider, one often needs some more structures on the manifold M^{d+1}, e.g. smooth structure, conformal structure, spin structure, etc.

To obtain the data to specify a QFT, we now would like to extend the observations seen in QM. Suppose we have a $(d+1)$-dimensional manifold M^{d+1}, then we wish to "define" a QFT by a gadget Z, which should produce a vector when M^{d+1} has a boundary

$$Z\big(M^{d+1}\big) \in \mathcal{H}_{bdy}, \tag{2.4}$$

and should produce a number when M^{d+1} has no boundary

$$Z\big(M^{d+1}\big) \in \mathbb{C}. \tag{2.5}$$

In the case of M^{d+1} with a boundary, the vector space associated on the boundary is called the space of states or just Hilbert space in the physics literature. On the one hand, if M^{d+1} does not have boundary, the number $Z\big(M^{d+1}\big)$ is called the partition function. When the boundary of M^{d+1} consists of several disconnected components, namely the boundary is given by a disjoint union of simply connected d-dimensional closed manifolds $\big\{N_i^d\big\}$: $\partial M^{d+1} = \sqcup_i N_i^d$, Z should define a linear map among the vector spaces defined on the boundaries. In particular, in the case $M^{d+1} = N^d \times I$, where I is an interval of length T, M^{d+1} has two boundaries N^d and Z now gives rise to a linear map $Z\big(N^d \times I\big) =: U(T)$,

$$U(T) : \mathcal{H}_N \longrightarrow \mathcal{H}_N. \tag{2.6}$$

From physical facts, Z also should be compatible with cutting and gluing of $(d+1)$-dimensional manifolds, and thus we learn that the linear map given above satisfies $U(T_1)U(T_2) = U(T_1+T_2)$. This really defines an operator H as the generator of U,

$$U(T) = \exp(-TH). \tag{2.7}$$

H is called the Hamiltonian of the system, acting on the space of states $\mathcal{H}_N$. If one considers a manifold with Lorentzian metric, $U(T)$ is represented as

$$U(T) = \exp(-iTH), \tag{2.8}$$

and then the interval I is regarded as "physical time."

2.1.2 Construction of Z

So far we have discussed only in an abstract way what quantum field theory is, or in other words what the gadget Z should satisfy. Let us now see how we specify Z to define a QFT. Broadly speaking, there are three kinds of constructions of Z. They are not totally independent and have many aspects of applicabilities. In the rest of this section, M denotes a $(d+1)$-dimensional manifold with or without boundary, and N denotes a d-dimensional closed manifold without boundary.

Construct from an axiom

The first construction is in a sense the simplest one; we write down appropriate properties that a QFT must satisfy, and axiomatize them. Construction of Z is to give a mathematical formulation which satisfies the axioms as the data to specify the theory. This is actually the only way that one can define a QFT by mathematically rigorous procedures. Normally, such a theory is first studied by physicists as an ideal or a toy model from physical motivations, and then refined as rigorous mathematics by mathematicians.[3] There are several kinds of such theories. We now introduce typical three examples.

The first one is *free theories* in any dimensions, which are toy models of field theories and play important roles as a probe of more complicated QFTs. In any spacetime dimensions, for the Riemannian manifolds with or without spin structure, we can rigorously define the free field theories. One of them is the free scalar field theory. Besides a $(d+1)$-dimensional closed manifold M, pick a G-bundle $P \to M$ with a connection A and a G-vector space V. Then we construct the associated vector bundle $P \times_G V$, whose covariant derivative is denoted by D_A. We now have the Laplacian Δ_A given by the covariant derivative D_A, and then the free scalar field theory is defined by the partition function

$$Z_{\text{scalar}}(M; A) := 1/\det \Delta_A. \tag{2.9}$$

If the determinant of the Laplacian involves a divergence, it must be properly regularized.

Another free theory is the free fermion theory. In this case M also needs a spin structure and the spin representations S, S'. Then we construct the Dirac operator $\mathcal{D}_A$ on the spin bundles, and the partition function of the free fermion theory is given by

$$Z_{\text{fermion}}(M; A) := \det \mathcal{D}_A, \tag{2.10}$$

[3]The fact that a QFT can be treated in a mathematically rigorous way implies that the theory may have an enormous amount of symmetry, and the difficulties of the QFT are completely controlled by them. Even better, these theories can often be exactly solved in an appropriate sense.

again the determinant is taken appropriately.[4]

Next, in two dimensions, we have another axiomatic quantum field theory, that is *conformal field theory* (CFT). To discuss two-dimensional CFT, we need a conformal (complex) structure on M, i.e. M becomes a Riemann surface. CFT in two dimensions was formulated by Belavin, Polyakov, and Zamolodchikov [4] as a model of physical systems at critical points. They established the renowned Virasoro algebra as an infinite dimensional symmetry of the system and fully investigated the minimal model. The holomorphic part of the Virasoro algebra is captured by vertex operator algebras, and since then there are many mathematically rigorous discussions on them. Their kinematic behaviors on Riemann surfaces are governed by the conformal blocks, and its geometric meaning has been studied in [5]. In turn, the study of irrational CFTs has led to the AGT correspondence [6], which has brought to us large amount of developments in both physics and mathematics. We will give a general statement of AGT correspondence and apply it to our work in section 4.

The final example is *topological quantum field theories* (TQFTs). TQFT is one of the main focuses in this paper. These theories originate from Witten's proposals of topological field theories [7–9]. Inspired by Witten's works, Atiyah and Segal axiomatized the topological QFT,[5] and happily this also has brought to us a numerous amount of applications both to physics and to pure mathematics. For example, $(1+1)$-dimensional topological QFTs are known to be functorially equivalent to the category of commutative ¡¡¡¡¡¡¡ HEAD Frobenius algebras [11]. (2+1)-dimensional Chern-Simons theory for a ======= Frobenius algebras. $(2+1)$-dimensional Chern-Simons theory for a ¿¿¿¿¿¿¿ dissertation compact Lie group G is rigorously constructed by Kohno [12] using the monodromy representation of Knizhnik-Zamolodchikov equation [13], and Turaev-Viro, Reshetikhin-Turaev using quantum groups [14, 15]. $(2+1)$-dimensional Chern-Simons theory has many applications to knot or link invariants and 3-manifold invariants. Moreover, both 2d and 3d TQFTs have applications to the mirror symmetry, or topological string theory (see e.g. [16]), those have led to fruitful interactions between physics and mathematics. We will take more time for TQFT and introduce the Atiyah's axioms in some detail in section 2.1.3.

Path integrate a Boltzmann weight

Before going to Atiyah's TQFT axioms, let us see two more constructions of Z. These are no longer mathematically rigorous, but rather familiar constructions for physicists. The second construction is to define a partition function by *path integral*. The general prescription is given as follows: One first introduces an action functional over the classical field configurations, which deduces the classical equation of motion by the variational principle. Then one

[4] Fermion theories may have anomalies. If it is the case, the axiom needs to be somehow modified. In particular the partition function for a closed manifold is given by the eta invariant [3].

[5] Precisely speaking, what Segal axiomatized is the definition of conformal field theories [10]. Segal's definition of CFT is categorical and quite similar to the definition of Atiyah's TQFT. For the geometric definition of CFT, see [5].

exponentiates the action and integrate it over the space of fields.

The partition function for the free fields given above can also be defined by the path integral expression. For example, for the free complex scalar field theory consider a section ϕ of the vector bundle $P \times_G V$ over M, where V is a representation of a compact Lie group G. Then define an action functional

$$S : \Gamma(P \times_G V) \longrightarrow \mathbb{R}, \tag{2.11}$$

such that

$$S(\phi) = \int_M \frac{1}{2} D_A\phi \wedge *D_A\phi, \tag{2.12}$$

where $*$ is the Hodge star on M. D_A is again the covariant derivative given by the connection A. Using this action functional, physicists "define" its partition function by

$$Z_{\text{scalar}}(M; A) := \int_{\Gamma(P\times_G V)} \mathcal{D}\phi\, e^{-S(\phi)}. \tag{2.13}$$

The integrand e^{-S} of path integral is generically called the *Boltzmann weight*. For the free field theories, the path integral can be expressed as an infinite product of the Gaussian integral. Introducing an appropriate regularization, one can compute the exact partition function and it leads to the same result as mentioned above.

As another example, let us consider pure Yang-Mills theory. We introduce the kinetic term of the connection A, and define the action functional

$$S(A) = \int_M \frac{1}{4g^2} F_A \wedge *F_A, \tag{2.14}$$

where F_A is the curvature of the G-connection A. Define the partition function of Yang-Mills theory by

$$Z_{\text{YM}}(M) = \int_{\mathcal{A}_M/\mathcal{G}} \mathcal{D}A\, e^{-S(A)}, \tag{2.15}$$

where the integral is taken over the space of connections on M modulo gauge transformations. Unlike the free scalar theory, making this path integral mathematically precise is extremely difficult. Although it is still ill-defined, physicists have been working on this expression to understand many properties of quantum gauge theories. Experimentally, we discretize the manifold M to a $(d+1)$-dimensional lattice and put it on a supercomputer. At least, numerical calculations show the above construction may be a mathematically meaningful and reproduce many experimental results to reasonable accuracy.

Deduce from string/M-theory

The final construction of Z is to use string or M-theory. This is also our main tool to construct a QFT to realize the correspondence (2.1). Nonetheless, string or M-theory is less rigorous than path-integral expression of the construction of Z, and thus it is more hopeless

to give a mathematically precise meaning to the construction. We here would like to show just two examples which "define" a class of quantum field theories through string/M-theory.

The first example is the AdS/CFT correspondence.

The second example is the so-called six-dimensional $\mathcal{N} = (2,0)$ superconformal field theories (SCFT) [17]. We start from a 10-dimensional string theory called the type IIB string theory, which roughly speaking assigns the partition function $Z_{\text{IIB}}(\tilde{M})$ to a 10-dimensional manifold $\tilde{M}$. Pick a finite subgroup Γ_G of SU(2) of type $G = A_n$, D_n or $E_{6,7,8}$. We define a six-dimensional QFT Q_G by its partition function for a six-dimensional manifold M,

$$Z_{Q_G}(M) = Z_{\text{IIB}}\big(M \times \mathbb{C}^2/\Gamma_G\big). \tag{2.16}$$

They are examples of six-dimensional $\mathcal{N} = (2,0)$ superconformal QFTs. These theories are known not to have a Lagrangian description, and hence their partition function cannot be obtained from the path integral formulation. They have another description via M-theory, as the low-energy dynamics of M5-branes. The construction from M5-branes leads to the AGT correspondence and the notion of class-$\mathcal{S}$ theories, which we will explain in later sections.

2.1.3 Atiyah's topological QFT

As mentioned some times, in this paper TQFT in extra dimensions plays a crucial role to realize the correspondence between supersymmetric gauge theories and integrable lattice models. So now let us pause here and introduce the Atiyah's axioms of TQFT [2]. We first list the axioms, and right then give their physical meanings. Readers will notice that most of these axioms are physically quite natural and actually mathematical rephrasing of the properties that Z should satisfy, which we have already seen. In addition, we would like to define lattice model as a discrete version of QFT in the next subsection. Reviewing the definition of TQFT here will be a good help of the argument of lattice model.

Axiom (Atiyah's $(d+1)$-dimensional TQFT)

$(d+1)$-dimensional TQFT is defined by Z consisting of the following two data and five assignments:

- For each oriented d-dimensional closed manifold N, Z assigns a finite dimensional $\mathbb{C}$-vector space $\mathcal{H}_N$:
$$Z(N) = \mathcal{H}_N. \tag{2.17}$$

This corresponds to (a half of) canonical quantization, or geometric quantization, known in physics literature. Why we say it is "a half of" will be explained in a moment. In physics, especially for a field theory which has a Lagrangian description, we consider the phase space of fields, take a constant-time surface, and then perform canonical quantization by imposing canonical commutation relations on fields and their conjugate momenta. The constant-time surface is a codimension-1 hypersurface in $(d+1)$-dimensional manifold M, which in this

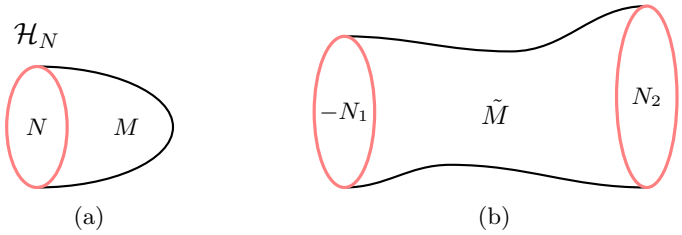


Figure 2: $(d+1)$-dimensional manifolds M and $\tilde{M}$. They have a single boundary and two disconnected boundaries. For each case, Z gives a vector and a linear map, respectively.

case is nothing but the d-dimensional closed manifold N. So, the manifold N is viewed as a collection of d spatial directions. One can think of the assigned vector space $\mathcal{H}_N$ as "the space of functionals on the classical fields on N," usually called the space of states or just Hilbert space of the system. The only major difference here is that for topological theory the Hilbert space is of finite dimension.[6]

- For each oriented $(d+1)$-dimensional manifold M with a boundary $\partial M = N$, Z assigns a vector

$$Z(M) = Z\left(\begin{array}{c}\text{N M}\end{array}\right) \in \mathcal{H}_N. \tag{2.18}$$

This expresses the path integral quantization on M with boundary. Recall that if one has a spacetime manifold with boundary, one needs to impose a boundary condition on fields on the boundary, and it leads to the state vector associated to the boundary condition by path integral expression; let φ be a fixed field configuration on N,

$$Z(M;\varphi) = \int_{X|_N=\varphi} \mathcal{D}X\, e^{-S(X)} \in \mathcal{H}_N. \tag{2.19}$$

In other words, the quantum state at a generic time t is given by the path integral of the Boltzmann weight over all the classical fields with time $< t$.

These data are subject to the following assignments:

1. (involutory) Let $-N$ denote a manifold N with the opposite orientation, then

$$\mathcal{H}_{-N} = \mathcal{H}_N^*, \tag{2.20}$$

   where $\mathcal{H}_N^*$ is the dual to $\mathcal{H}_N$.

2. (multiplicative) If N is a disjoint union of two d-dimensional closed manifolds N_1 and N_2, then the vector space associated on it is factorized to a tensor product:

$$\mathcal{H}_{N_1 \sqcup N_2} = \mathcal{H}_{N_1} \otimes \mathcal{H}_{N_2}. \tag{2.21}$$

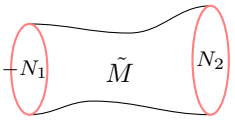
[6] One can remove the finiteness condition from the axiom. In TQFT we can in fact define a non-degenerate bilinear form from the other axioms, and from that as a corollary we can deduce the space of states is finite-dimensional.

This is really natural in the physical point of view. This condition means that if we have two physical systems defined on spatially disjoint union $N_1 \sqcup N_2$, then the space of states is represented as the tensor product of each state space. Also, from the conditions so far one finds that if a $(d+1)$-dimensional manifold M' has two boundaries such that

$$\partial M' = -N_1 \sqcup N_2, \tag{2.22}$$

then Z defines a linear map

$$Z\big(M'\big) = Z\left(\quad \right) \in \mathrm{Hom}_{\mathbb{C}}\left(\mathcal{H}_{N_1}, \mathcal{H}_{N_2}\right), \tag{2.23}$$



namely, for a cobordism between N_1 and N_2, we have a linear map between $\mathcal{H}_{N_1}$ and $\mathcal{H}_{N_2}$. This defines a time evolution, or transition amplitude, between the states in $\mathcal{H}_{N_1}$ and $\mathcal{H}_{N_2}$. When the field theory has a Hamiltonian, $Z\big(M'\big)$ may correspond to the time evolution operator

$$Z\big(M'\big) := e^{-tH} : \mathcal{H}_{N_1} \longrightarrow \mathcal{H}_{N_2}. \tag{2.24}$$

Since canonical quantization is a procedure to make an assignment of Hilbert spaces and the time evolution on them, as we saw in QM, this association of linear maps gives the other half of canonical quantization. One can also explicitly express $Z\big(M'\big)$ in the path integral expression by the Feynman kernel (also as known as propagator or Green's function). Suppose we have an initial state $\Psi_0 \in \mathcal{H}_{N_1}$, then the state $\Psi_t \in \mathcal{H}_{N_2}$ at time t is expressed by

$$\begin{aligned} \Psi_t(\varphi_t) &= \left(e^{-tH}\Psi_0\right)(\varphi_t) \\ &= \int K(\varphi_t, \varphi_0)\Psi_0(\varphi_0)\mathcal{D}\varphi_0, \end{aligned} \tag{2.25}$$

where

$$K(\varphi_t, \varphi_0) = \int_{X|_{N_1}=\varphi_0}^{X|_{N_2}=\varphi_t} \mathcal{D}X\, e^{-S(X)}. \tag{2.26}$$

3. For two cobordisms

$$\partial M_1 = -N_1 \sqcup N_2, \quad \partial M_2 = -N_2 \sqcup N_3, \tag{2.27}$$

it follows that

$$Z\big(M_1 \cup_{N_2} M_2\big) = Z(M_2)Z(M_1), \tag{2.28}$$

where the cobordisms M_1 and M_2 are glued along N_2 by a certain diffeomorphism on N_2. This asserts that the linear maps are transitive when we compose cobordisms. This is nothing but the physical requirement that the time evolution is compatible with the cutting and gluing the manifolds.

4. Given $N = \emptyset$ as an empty d-dimensional manifold, then the associated vector space is one-dimensional:

$$Z(\emptyset) = \mathbb{C}. \tag{2.29}$$

This is a non-triviality condition. From this condition, for each $(d+1)$-dimensional manifold M without boundary, $\partial M = \emptyset$, Z assigns a number:

$$Z(M) \in \mathbb{C}. \tag{2.30}$$

This number associated to a closed $(d+1)$-dimensional manifold is known as partition function. Not only is it an important quantity physically, but it also has mathematical applications, such as giving a topological invariant of the closed manifold M.

5. Let I be an interval, then for each cylinder $M = N \times I$, the linear map on $\mathcal{H}_N$ is trivial:

$$Z(N \times I) = \mathrm{id}_{\mathcal{H}_N}. \tag{2.31}$$

Since in a topological theory manifolds diffeomorphic to each other should be considered as the identical, each cobordism class defines a linear map. This means that the addition of a cylinder is a trivial operation; the associated linear map is the identity. This condition is another major difference from ordinary QFTs. For example, for $(2+1)$-dimensional Chern-Simons theory one can explicitly see its Hamiltonian is identically 0 and hence the time evolution of Chern-Simons theory is trivial.

This is all the assignments for Atiyah's TQFT. Although we gave the path integral expressions at the explanation of the physical meanings for some conditions, the axioms themselves are really the basis for the rigorous mathematical definition of Z. One can actually take an equivalent definition of TQFT in slightly different axioms. For more details, see the original paper by Atiyah [2].

At first sight, this might look a little bit complicated and too abstract. This definition, however, is really natural and works quite well in a sense that it gives a kind of "homomorphism" between geometry and algebra:

$$Z : \text{'geometry'} \longrightarrow \text{'algebra'}.$$

In fact, all in all the axioms of Atiyah's TQFT define a functor from the category of $(d+1)$-dimensional cobordisms to the category of finite-dimensional $\mathbb{C}$-vector spaces:[7]

$$Z : \mathrm{Bord}_{d+1} \longrightarrow \mathrm{Vect}_{\mathbb{C}}. \tag{2.32}$$

TQFT luckily has a mathematical precise definition. For a general QFT, i.e. not TQFT, in many cases there is no precise definition as we saw the examples of the construction of Z. However, a QFT is generically expected to be characterized by some functor from the geometric structure of spacetime manifolds to the algebraic description of physical states and observables.

[7] What is more, the category of cobordisms and the category of vector spaces are endowed with the product structures, that is, taking disjoint union and taking tensor product, which are really symmetric operations. As such, TQFTs are keeping such structures and then called symmetric monoidal functor.

2.2 Lattice model as discrete QFT

Now let us move on to the discussion on lattice model. First of all, let us begin with the question of "what is lattice model?" Probably there is no unique definition of lattice model, though. We here would like to characterize lattice model by a discrete version of QFT. What we have learned in the previous subsection is that a $(d+1)$-dimensional QFT may be defined by an appropriate functor Z_{Q}^{d+1}, which produces a number for each $(d+1)$-dimensional closed manifold M:

$$Z_{\mathrm{Q}}^{d+1}(M) \in \mathbb{C}, \tag{2.33}$$

which is called the partition function of the model, and Z_{Q}^{d+1} satisfies additional some reasonable conditions for each QFT of interest.

We introduce lattice model in a same manner. For a $(d+1)$-dimensional closed lattice L, a $(d+1)$-dimensional lattice model is defined by Z_{L}^{d+1}, which produces a number

$$Z_{\mathrm{L}}^{d+1}(L) \in \mathbb{C}, \tag{2.34}$$

which is again called the partition function, and Z_{L}^{d+1} satisfies additional some reasonable conditions. The most typical example of lattice model in particle physics is lattice gauge theory or lattice QCD.[8] For such theories, it is very clear that the model is defined by a discrete version of QFTs; One compactifies the spacetime manifold $\mathbb{R}^4$ to four-torus T^4, and then discretize the theory and put it on a lattice on the torus.

In either case, therefore, "to define a model" means to specify Z. Throughout the paper, we will consider only $d = 0$ or 1 case of lattice model, and we refer each case as 1d or 2d lattice model. To prepare for the discussion of integrability, we take the prominent example of statistical lattice model called the Ising model. We first briefly review the generality of the Ising model and compare with field theory, and then introduce transfer matrix which is essentially a time evolution operator in a discrete quantum system.

2.2.1 Prominent example – the Ising model

The Ising model is a very good introduction to integrable lattice model. The 1d and 2d Ising model is known to be exactly solvable, whose meaning we will clarify in a moment, and it is said to be integrable. To see the generality of the Ising model, we first define 2d lattice and spin configurations on it.

Define a 2d periodic lattice on two-torus T^2 by

$$\begin{aligned} L &:= \mathbb{Z}_m \times \mathbb{Z}_n \\ &= \{1, \dots, m\} \times \{1, \dots, n\}\,. \end{aligned} \tag{2.35}$$

[8] For sure, for any field theories when one computes a physical quantity on a computer, one needs to discretize the theory and put it on a lattice. In this sense, numerical analysis of field theory is always regarded as a lattice model.

A spin configuration on the lattice L is given by a map

$$\mathbf{s} : L \longrightarrow \{+1, -1\}\,, \tag{2.36}$$

where usually $+1$ is called "up spin" and -1 "down spin." The map defines a configuration of up and down spins at each site of the lattice L. So it can also be thought of as an assignment of $+1$ or -1 on all the sites of the lattice L. We often denote the spin at each site by its image $s_I := \mathbf{s}(I)$, $I \in L$. Let $S(L)$ be the set of all the spin configurations. Since spin configurations are defined on the $m \times n$ periodic lattice, the number of elements in $S(L)$ is 2^{mn}. In other words, the number of all the allowed configurations of spins on the lattice L is 2^{mn}.

For each spin configuration $\mathbf{s} \in S(L)$, define the energy functional of the Ising model by

$$\begin{aligned} E_{\text{Ising}}(\mathbf{s}) &= -J \sum_{\langle I, I' \rangle} s_I s_{I'} \\ &:= -J \left(\sum_{i,j=1}^{m,n} s_{ij} s_{i,j+1} + \sum_{i,j} s_{ij} s_{i+1,j} \right), \end{aligned} \tag{2.37}$$

where the $s_{ij} = s_{i+m,j} = s_{i,j+n}$, and $J \in \mathbb{R}_{>0}$ is a constant parameter. This is one of the simplest spin systems, in which only the nearest neighbor spins have interactions. From this energy functional, the partition function of the Ising model is defined by

$$Z\big(L, E_{\text{Ising}}; \beta\big) := \sum_{\mathbf{s} \in S(L)} e^{-\beta E_{\text{Ising}}(\mathbf{s})}, \tag{2.38}$$

where $\beta \in \mathbb{R}_{\geq 0}$ is an inverse temperature. The summand is generically called the Boltzmann weight as well as in field theory. In this expression, one recognizes that the right-hand side is a discrete version of path integral expression in field theory. The energy functional is corresponding to the action functional of a field theory, β^{-1} is the Planck's constant, and the sum over all the spin configurations is the path integral over all the field configurations. Given a periodic lattice L, the partition function returns a number, which is a discrete version of the partition function of QFT which returns a number as well if a closed manifold M is given. For later use, define another quantity called the free energy,

$$f(\beta) := -\frac{1}{\beta} \frac{1}{mn} \log Z(\beta), \tag{2.39}$$

where $Z(\beta)$ is the partition function defined above.

According to the general story of statistical mechanics, the probability that a configuration $\mathbf{s}$ with energy $E(\mathbf{s})$ is realized is given by the canonical ensemble[9]

$$p\,(\mathbf{s}) = \frac{1}{Z} e^{-\beta E(\mathbf{s})}, \tag{2.40}$$

[9]To the end of this subsection, we argue for a general energy functional, but the reader may assume the Ising model.

where Z is the partition function of a statistical system. In a spin system, a physical observable is in general given by a functional on the space of spin configurations:

$$\mathcal{O} : S(L) \longrightarrow \mathbb{R}, \tag{2.41}$$

and its expectation value is given by the canonical ensemble as

$$\langle \mathcal{O} \rangle_p := \sum_{\mathbf{s} \in S(L)} \mathcal{O}(\mathbf{s}) p(\mathbf{s}) = \frac{1}{Z} \sum_{\mathbf{s} \in S(L)} \mathcal{O}(\mathbf{s}) e^{-\beta E(\mathbf{s})}. \tag{2.42}$$

In particular, an energy functional is an example of physical observable,

$$\frac{1}{mn} \langle E \rangle_p = \frac{1}{mn} \frac{1}{Z} \sum_{\mathbf{s} \in S(L)} E(\mathbf{s}) e^{-\beta E(\mathbf{s})}, \tag{2.43}$$

which is obtained from the derivative of the free energy,

$$\frac{1}{mn} \langle E \rangle_p = \frac{\partial}{\partial \beta} \big(\beta f(\beta) \big). \tag{2.44}$$

Generally speaking, the free energy provides us all the information of the system. For example, the state most likely to happen is governed by variational principle of the free energy, and we can in principle compute physical quantities such as expectation value of energy, fluctuation, specific heat, and so on. To make a lattice model as a physically meaningful system, however, one needs to take the thermodynamic limit; $m, n \to \infty$. Therefore, in this sense, integrability of lattice model is characterized by the calculability of an exact free energy at the thermodynamic limit. This may be rephrased as well by the calculability of the partition function. Based on these argument, we shall see the integrability of the Ising model in the next subsection.

2.2.2 Transfer matrix and integrability

It is well known that 1d and 2d Ising model is exactly solvable in the sense that one can exactly compute its free energy in the thermodynamic limit. Both 1d and 2d cases are solved by so-called the method of transfer matrix. To see this, for simplicity in this section we consider 1d Ising model.

The setup is almost the same as in the 2d case. A 1d lattice on torus and a spin configuration is defined by

$$L_{1d} = \mathbb{Z}_n = \{1, \ldots, n\}, \tag{2.45}$$

$$\mathbf{s} : L_{1d} \longrightarrow \{\pm 1\}. \tag{2.46}$$

Define the energy functional of 1d Ising model by

$$E_{1d\,\mathrm{Ising}}(\mathbf{s}) = -J \sum_{i=1}^{n} s_i s_{i+1}, \quad s_{n+1} = s_1. \tag{2.47}$$

(a) (b)

Figure 3: (a) Classical Ising spin chain. Spin variables s_i take values in $\{\pm 1\}$. They are just c numbers. (b) Quantum spins are vectors in $\mathbb{C}^2$ associated at each site. The discrete time evolution of quantum spins is given by the transfer matrix.

Then the partition function of 1d Ising model is given as

$$\begin{aligned} Z\big(L_{1d}, E_{1\mathrm{d}}; \beta\big) &:= \sum_{\mathbf{s}} e^{-\beta E_{1d}(\mathbf{s})} \\ &= \sum_{s_1,\ldots,s_n=\pm 1} e^{Ks_1s_n} \cdots e^{Ks_3s_2} e^{Ks_2s_1}, \end{aligned} \tag{2.48}$$

where $K := \beta J$.

Now introduce the *transfer matrix*,

$$T := \left(e^{Kss'}\right)_{s,s'=\pm 1} = \begin{pmatrix} e^K & e^{-K} \\ e^{-K} & e^K \end{pmatrix}, \tag{2.49}$$

where the indices of row and column are labeled by $+1$ and -1, respectively. Using this matrix T, the partition function is rewritten as

$$Z\left(\beta\right) = \sum_{s_1,\ldots,s_n=\pm 1} T_{s_1s_n} \cdots T_{s_3s_2} T_{s_2s_1}. \tag{2.50}$$

By definition of matrix multiplication, the partition function is eventually given by a trace:

$$\begin{aligned} Z(\beta) &= \sum_{s_1=\pm 1} \left(T^n\right)_{s_1s_1} \\ &= \mathrm{Tr}_{\mathbb{C}^2}\left(T^n\right). \end{aligned} \tag{2.51}$$

This expression tells us that we have the quantum Hilbert space $\mathbb{C}^2$ at the boundary of each 1d segment, and the transfer matrix T sends a state to the adjacent site, which is the discrete "time evolution" of this system; $\log T \propto$ Hamiltonian as we saw in 1d QFT in section 2.1.1. This corresponds to the analogue of Hamiltonian/operator formalism in field theory. In operator formalism, spins are replaced by the Pauli matrices, and original classical up spin and down spin are replaced by the eigenvalues and eigenvectors of the Pauli matrix σ^z. Further, the non-commutativity is now manifest in a way that the spins and their time evolution is given by matrices, this is the consequence of quantization.

In the expression above of the partition function, let us consider the thermodynamic limit $n \to \infty$. The eigenvalues of T are easily obtained and let them be $\lambda_0 > \lambda_1$, then we have the free energy

$$-\beta f(\beta) = \lim_{n\to\infty} \frac{1}{n} \log Z(\beta)$$

$$= \lim_{n\to\infty} \frac{1}{n} \log \lambda_0 \left(1 + \left(\frac{\lambda_1}{\lambda_0}\right)^n\right)$$
$$= \log \lambda_0. \tag{2.52}$$

We conclude that the free energy in the thermodynamic limit is just given by the largest eigenvalue of the transfer matrix. Finally, integrability of lattice model is rephrased as

One can exatly find the eigenvalues of transfer matrix.
$\Rightarrow$ Can exactly compute the free enrgy.
$\Rightarrow$ The system is exactly solvable.

The transfer matrices of 1d and 2d Ising model are diagonalizable and one can exactly find their eigenvectors and eigenvalues using, for example, algebraic Bethe ansatz. In this sense, 1d and 2d Ising model are said to be integrable, or exactly solvable. Then, a natural question arises; when can we diagonalize the transfer matrix of a lattice model? This actually leads to the most fundamental answer to the question of "what is integrable model." A canonical answer is *Yang-Baxter equation*. When the Boltzmann weight with spectral parameters, which take values in a Riemann surface, satisfies the Yang-Baxter equation, the lattice model is integrable. We are going to give an explanation of this definition of integrability in next subsection in some detail, with its origin from TQFT with extra dimensions.

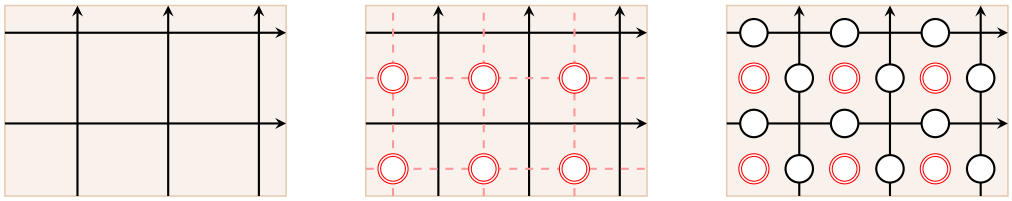
2.3 Integrability from TQFT in extra dimensions

So far, we have seen that a QFT is ideally given by some functor Z, and lattice model is also given by Z as a discrete version of QFT. Now, we would like to relate these two. The key concept is that "line operators in 2d TQFT form a lattice model." In particular, if the 2d TQFT has extra dimensions, a corresponding lattice model naturally turns out to be integrable.

The purpose of this section is to provide a general discussion to relate four-dimensional supersymmetric gauge theories and integrable lattice models. We first give a step-by-step explanation from two-dimensional TQFT to integrable lattice model, and then put the argument in the setup of brane tilings. Branes in string theory are powerful enough to yield the systematic method to relate a class of supersymmetric gauge theories with integrable lattice models, as we will see. To begin with, we present a general prescription of brane constructions at a formal level. Concrete setups and more details will be given in later sections. The discussion in this subsection is mainly based on [18].

2.3.1 Lattice models from TQFTs with line operators

Suppose that we have a two-dimensional TQFT T on a two-torus T^2 equipped with line operators $\mathcal{L}_i$, $i = 1, \ldots, l$. Wrap them around one-cycles C_i on the torus in such a way that they form an $m \times n$ lattice; see figure 4(a). We wish to consider the correlation function

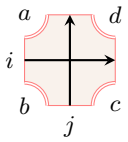


(a) (b) (c)

Figure 4: Construction of a lattice model from line operators. (a) A lattice of line operators on a torus. (b) The torus with holes obtained by gluing the pieces. (c) The corresponding spin model.

of this configuration of line operators. To compute it, our strategy is to break up the torus into pieces and perform in turn the path integral piecewise. Then the original correlation function is reconstructed by combining these pieces.[10]

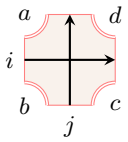
Let us consider the following piece of surface which contains an intersection of two line operators $\mathcal{L}_i$ and $\mathcal{L}_j$:



. (2.53)

If one thinks of the above picture as the world-sheet of two scattering open strings, one can regard the above picture as follows: The endpoints of the open strings sweep out the double-lined arcs, on which D-branes are there and the boundary conditions (Chan-Paton factors) are specified by labels a, b, c, d. The line operators are the world-lines of particles associated with the open strings.

The path integral for the piece (2.53) produces a linear map (recall an important property of quantum mechanics (2.2) and an assignment of TQFT)

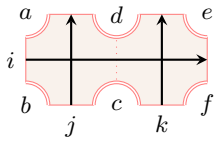


$$\rightsquigarrow \quad R_{ij}\begin{pmatrix} a & d \\ b & c \end{pmatrix} : V_{ab,i} \otimes V_{bc,j} \longrightarrow V_{ad,j} \otimes V_{dc,i}\,, \tag{2.54}$$

where $V_{ab,i}$ is the space of states of the open string propagating along $\mathcal{L}_i$ with the boundary conditions on the left and the right ends specified by a, b, respectively. We call this map the R-matrix (or R-operator) associated with this decorated surface. To reconstruct the original configuration of line operators, we glue these pieces together. Gluing them amounts to the

[10] The original Atiyah's TQFT is actually called "closed TQFT." In the following, we will compute the correlation function by embedding the closed TQFT with line operators into an open/closed TQFT with line operators, or so-called defect TQFT; see e.g. [19, 20].

composition of R-matrices. For example, gluing two pieces horizontally gives



$$\rightsquigarrow \quad R_{ik}\begin{pmatrix} d & e \\ c & f \end{pmatrix} \circ_{V_{dc,i}} R_{ij}\begin{pmatrix} a & d \\ b & c \end{pmatrix}. \tag{2.55}$$

The original configuration is thus obtained by gluing all the pieces. It, however, still has holes assigned boundary conditions as in figure 4(b), which we must fill by summing over the boundary conditions. To do this, let us recall that the path integral on a finite-length cylinder with boundary condition a imposed on one end defines a closed string state $|a\rangle$ as known as a boundary state. Similarly, the path integral on a disk with no insertion of operators defines a state $|1\rangle$ on the boundary. Assume that we have chosen the set of boundary conditions B to be sufficiently large so that the state $|1\rangle$ can be written as a superposition of boundary states:

$$|1\rangle = \sum_{a\in B} c_a\,|a\rangle\,. \tag{2.56}$$

Then, the sum over the boundary conditions gives the state $|1\rangle$ on the boundary of each hole, which is replaced with a disk:

$$\sum_{a\in B} c_a \Big(a \quad \Big) = \sum_{a\in B} c_a \Big(|a\rangle \quad \Big) = |1\rangle \quad = \quad . \tag{2.57}$$

Thus the holes are filled and the original configuration on the torus is reconstructed.

Now that we have understood how to reconstruct the correlation function of line operators in the field theoretic point of view, let us reinterpret this procedure as an operation in statistical lattice model. Rephrasing the procedures so far, it is clear that the above configuration defines a partition function of a statistical spin model. The model has spins located at two kinds of sites, ○ and ◎. They correspond to the open string states and the boundary states, see figure 4(c). A spin at ○ takes values in the chosen basis for the relevant open string states $V_{ab,i}$, while that at ◎ is valued in B. The Boltzmann weights for each spin configuration are given by the matrix elements of R-matrix and the coefficients of boundary states c_a. Thus, we conclude that the correlation function of line operators $\{\mathcal{L}_i\}$ on the torus coincides with the partition function of a statistical spin model defined on the periodic lattice formed by the line operators:

$$\left\langle \prod_{i=1}^{l} \mathcal{L}_i[C_i] \right\rangle_{\mathsf{T},T^2} = Z_{\mathsf{L}(\mathsf{T})}\big(\{\mathcal{L}_i[C_i]\}\big). \tag{2.58}$$

$\mathsf{L}(\mathsf{T})$ denotes the lattice model arising from the TQFT T with the line operators and $\{\mathcal{L}_i[C_i]\}$ is the lattice formed by line operators $\mathcal{L}_i$ wrapping around C_i. In other words, the partition function of lattice model (2.34) is in this case defined by the lattice $\{\mathcal{L}_i[C_i]\}$, and specified by the left-hand side. By this construction, we can view typical examples of lattice models.

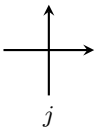
IRF model

If $\dim V_{ab,i} = 1$ for any a, b, i, in this case we can ignore the spins at $\bigcirc$. We just sum over the boundary conditions and such a spin system $\mathsf{L}(\mathsf{T})$ is called *interaction-round-a-face model*, or *IRF model* for short. The spins are placed on the faces of the lattice, and interaction takes place among four spins surrounding a vertex. The 2d Ising model actually can be formulated as an IRF model by taking vertex-face transformation.

Vertex model

If B consists of a single boundary condition, say a, we can simply write the open string state space by $V_i := V_{aa,i}$, and the R-matrix is represented only by a crossing of two lines:

$$R_{ij} = \quad , \tag{2.59}$$



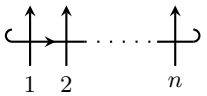
$$R_{ij} : V_i \otimes V_j \longrightarrow V_j \otimes V_i \ . \tag{2.60}$$

The space of states V_i can also be thought of as the Hilbert space of a particle propagating on the line operator $\mathcal{L}_i$. In this case we can ignore the spins at $\bigcirc\!\!\bigcirc$ since there is no summation over boundary conditions. This means that the lattice model $\mathsf{L}\,(\mathsf{T})$ is called a *vertex model*: Spins are living on the edges of the lattice and interact with each other at the vertices.

We can always recast our lattice model into a vertex model by setting $V_i := \bigoplus_{a,b\in B} V_{ab,i}$, at least formally, and declaring that all newly introduced R-matrix elements, which correspond to scattering processes with inconsistent Chan-Paton factors, vanish. We can also absorb the coefficients c_a into the R-matrix elements by appropriate rescaling. In what follows we implicitly perform this reformulation, and restrict ourselves to vertex model.

A remarkable aspect of this construction of lattice models is that it allows us to understand integrability from a higher-dimensional point of view. This is crucial observation by Costelllo [21]. In our lattice model, consider a horizontal line operator $\mathcal{L}_i$ intersecting vertical line operators $\mathcal{L}_j$, $j = 1, \ldots, n$. Concatenating the R-matrices in this row, we get the row-to-row transfer matrix:

$$T_i = \quad = \mathrm{Tr}_{V_i}\left(R_{i,n} \circ_{V_i} \cdots \circ_{V_i} R_{i,1}\right). \tag{2.61}$$



The hooks on the horizontal line are to remind us that the periodic boundary condition is imposed. This gives an endomorphism of $\bigotimes_{j=1}^{n} V_j$ which maps a state just below $\mathcal{L}_j$ to another state above it, namely the transfer matrix defines the discrete time evolution of this spin system and $\mathcal{H} := \bigotimes_{j=1}^{n} V_j$ is the total quantum Hilbert space. In terms of transfer matrices, the partition function is written as a trace,

$$Z_{\mathsf{L}(\mathsf{T})}\big(\{\mathcal{L}_i[C_i]\}\big) = \mathrm{Tr}_{\mathcal{H}}\left(T_{n+m} \cdots T_{n+1}\right). \tag{2.62}$$

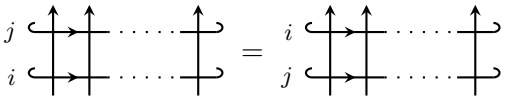
An important point is that the underlying field theory is a TQFT, and hence a state evolves trivially on a cylinder unless it hits something, line operators in the present case.

Now suppose that each line operator depends on a continuous parameter which is an element of some smooth manifold S. This parameter is called *spectral parameter* of the lattice model. Let u_i be the spectral parameter of $\mathcal{L}_i$, then the R-matrix and the transfer matrix are rewritten by

$$R_{ij} \longrightarrow R_{ij}(u_i, u_j), \tag{2.63}$$

$$T_i \longrightarrow T_i(u_i; u_1, \ldots, u_n). \tag{2.64}$$

To avoid clutter, we fix $u_1, \ldots, u_n$ and suppress them below. A vertex model is said to be integrable if the transfer matrices $T_i(u_i)$ are meromorphic functions of u_i and commute with each other:



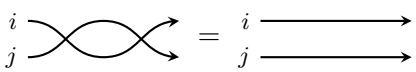
$$\iff [T_i(u_i), T_j(u_j)] = 0, \quad u_i \neq u_j. \tag{2.65}$$

When the transfer matrices with different spectral parameters commute, we can find a series of mutually commuting operators on the total Hilbert space $\bigotimes_{j=1}^{n} V_j$ by Laurent expansion of the transfer matrix. In particular, they commute with the transfer matrix itself, which means they produce an infinite number of conserved quantities. In fact, the commutativity of transfer matrix implies that one can find the exact eigenvalues and eigenvectors of the transfer matrix, which was the underlying reason of integrability of the Ising model.

The situation considered here, namely the addition of spectral parameters and commuting transfer matrix, is naturally realized if the TQFT has "extra dimensions." In this scenario, we really start with a higher-dimensional theory $\tilde{\mathsf{T}}$ formulated on the product space $S \times T^2$, which is topological on the torus T^2 but not on S. We wrap line operators $\mathcal{L}_i$ around $\{u_i\} \times C_i$, where u_i are points on S. If one can see only the torus T^2 and is unaware of the extra dimensions S, the theory seems to be the previous two-dimensional TQFT $\mathsf{T} \cong \tilde{\mathsf{T}}[S]$ that has parameters taking values in S. One finds that line operators $\mathcal{L}_i(u_i)$ wrapping around C_i carry continuous parameters u_i in the seemingly two-dimensional theory, and the correlation function for this configuration of line operators is given by the partition function of a lattice model $\mathsf{L}\big(\tilde{\mathsf{T}}[S]\big)$ defined on the lattice $\{\mathcal{L}_i(u_i)[C_i]\}$.

For a generic choice of $\{u_i\}$, the transfer matrices of the lattice model $\mathsf{L}\big(\tilde{\mathsf{T}}[S]\big)$ commute since the two horizontal line operators such as equation (2.65) can move freely and interchange their positions owing to the topological nature along T^2; no phase transition occurs when they pass each other as they do not meet in the full spacetime $S \times T^2$. Thus, integrability follows from the existence of extra dimensions, whose coordinates provide continuous spectral parameters.

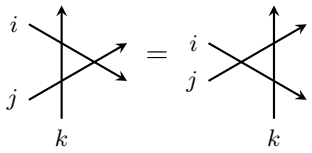
In fact, we can deduce integrability from another point of view. By the same logic, we have the unitarity relation



(2.66)

$$\iff R_{ji}(u_j,u_i)R_{ij}(u_i,u_j)=\mathrm{id}_{V_i\otimes V_j}, \tag{2.67}$$

and the Yang-Baxter equation



(2.68)

$$\iff R_{ij}(u_i,u_j)R_{ik}(u_i,u_k)R_{jk}(u_j,u_k)=R_{jk}(u_j,u_k)R_{ik}(u_i,u_k)R_{ij}(u_i,u_j)\ , \tag{2.69}$$

where the R-matrix R_{ij} acts on $V_i\otimes V_j$ as an intertwiner and trivial on V_k, etc. From these two relations we can reproduce the commutativity of the transfer matrix. We should emphasize that the Yang-Baxter equation is a local condition. When the Boltzmann weight locally satisfies the Yang-Baxter equation, it extends to the commutativity of the transfer matrix and hence to the integrability of the model. In this sense, the Yang-Baxter equation is the fundamental condition of integrability of a lattice model.

Before getting into the discussion of brane construction, we would like to generalize the above arguments to further higher-dimensional situations. First of all, replace the two-torus T^2 with a general two-dimensional surface Σ, along which line operators are wrapped. We now have $S\times\Sigma$, and similarly to the above a lattice model is defined on Σ by line operators with spectral parameters. Let us consider the case that we really have more extra dimensions and a higher-dimensional theory is formulated on $S\times M\times\Sigma$,[11] where M is some smooth manifold. In such a case, the line operators we had may descend from extended operators of dimension greater than one. Let T again be the new higher-dimensional theory, and suppose that it is topological on Σ and has extended operators $\mathcal{E}_i$ whose codimension is greater than $\dim S$. Place $\mathcal{E}_i$ on submanifolds of the form $\{u_i\}\times N_i\times C_i$. Since $\mathsf{T}[S\times M]$ is again regarded as a two-dimensional TQFT defined on Σ, the correlation function of the operators $\mathcal{E}_i$ in the full theory T still should coincide with the partition function of an integrable lattice model $\mathsf{L}\big(\mathsf{T}[S\times M]\big)$. The model is now defined on the lattice formed by the line operators $\mathcal{E}_i\big(u_i;N_i\big)[C_i]$ winding along one-cycles C_i on Σ, which are the image of $\mathcal{E}_i$ in the two-dimensional theory $\mathsf{T}[S\times M]$.

As well as we can view the higher-dimensional theory as a two-dimensional TQFT, we may also view it as a theory $\mathsf{T}[\Sigma]$, which is a QFT on $S\times M$ specified by the surface Σ. In this theory, the operators $\mathcal{E}_i$ are seen as extended operators $\mathcal{E}_i\big(C_i\big)$ supported on $\{u_i\}\times N_i$. Then, we have another relation

$$\left\langle\prod_{i=1}^{l}\mathcal{E}_i\big(C_i\big)\big[\,\{u_i\}\times N_i\big]\right\rangle_{\mathsf{T}[\Sigma],S\times M}=Z_{\mathsf{L}(\mathsf{T}[S\times M])}\big(\,\{\mathcal{E}_i\big(u_i;N_i\big)[C_i]\}\big). \tag{2.70}$$

Thus, we finally arrive at a correspondence between a QFT on $S\times M$ equipped with extended operators and an integrable lattice model on Σ.

[11] Manifolds M, N in this subsection are not necessarily the same as ones appeared in previous subsections.

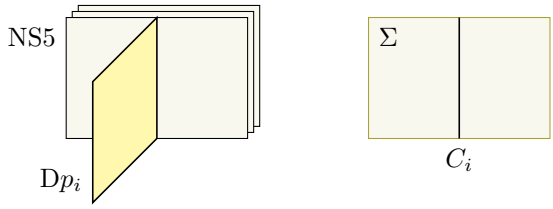


Figure 5: The Dp-brane Dp_i ending on the NS5-branes creates a defect $\mathcal{E}_{\mathrm{Dp_i}}$ along C_i.

2.3.2 Brane construction and correspondence

We have seen above that a lattice model is realized by a lattice of line operators in a two-dimensional TQFT, and it is integrable if the TQFT is embedded in higher dimensions and the line operators come from extended operators localized in some directions of the extra dimensions. Now we are ready to explain how to get such structures of correspondence (2.70) using branes in string theory. The brane construction here is still rather abstract. A little bit more concrete setups will be given in next section. The reader may skip this subsection for the first reading and jump to the next section.

Consider a type II string theory in a ten-dimensional spacetime

$$\mathbb{R}^4 \times T^*\Sigma \times \mathbb{R}^2, \tag{2.71}$$

where Σ is a two-dimensional surface embedded in $T^*\Sigma$ as a zero section. Introduce a stack of N NS5-branes supported on $\mathbb{R}^4 \times \Sigma \times \{0\}$ in this spacetime, and Dp-branes Dp_i on $\mathbb{R}^{p-1} \times \Sigma_i \times \{0\}$ ending on the NS5-branes, where $\mathbb{R}^{p-1}$ is a subspace of $\mathbb{R}^4$ (assuming $p \leq 5$) and Σ_i are surfaces in $T^*\Sigma$ such that $\Sigma_i \cap \Sigma = C_i$; see figure 5. Provided that Σ_i are suitably chosen, this brane system preserves four supercharges.

The low-energy dynamics of the NS5-branes is governed by a six-dimensional theory $\mathsf{T}_{\mathrm{NS5}}$ on $\mathbb{R}^4 \times \Sigma$. The theory $\mathsf{T}_{\mathrm{NS5}}$ is depending on whether IIA or IIB theory we are considering:

$$\begin{aligned}
&\text{IIA} : \mathcal{N} = (2,0) \text{ superconformal QFT of type } A_{N-1}, \\
&\text{IIB} : \mathcal{N} = (1,1) \text{ super Yang-Mills theory with gauge group SU}(N).
\end{aligned}$$

The theory $\mathsf{T}_{\mathrm{NS5}}$ is formulated on $\mathbb{R}^4 \times \Sigma$, with topological twist along Σ which breaks half of sixteen supercharges. In this twisted theory, Dp_i create p-dimensional defects $\mathcal{E}_{\mathrm{D}p_i}$ on $\mathbb{R}^{p-1} \times C_i$, reducing the number of unbroken supercharges to four. From the point of view of a four-dimensional observer, this brane configuration gives half-BPS defects $\mathcal{E}_{\mathrm{D}p_i}(C_i)$ in an $\mathcal{N} = 2$ theory $\mathsf{T}_{\mathrm{NS5}}[\Sigma]$. The total system is invariant under a U(1) R-symmetry originating from the rotational symmetry on the $\mathbb{R}^2$ factor of the ten-dimensional spacetime.

Let us take a three-manifold M and $(p-2)$-submanifolds N_i of M, and modify the above construction so that the world-volumes of the NS5-branes and the Dp-branes become $S^1 \times M \times \Sigma$ and $S^1 \times N_i \times \Sigma_i$, respectively. At low energies, we get the same theory $\mathsf{T}_{\mathrm{NS5}}$ formulated on $S^1 \times M \times \Sigma$ and defects $\mathcal{E}_{\mathrm{D}p_i}$ located on $S^1 \times N_i \times C_i$. In general, this modification completely breaks supersymmetry. For certain choices of M and N_i, however,

there is a string background in which a fraction of supersymmetry is still preserved. In such a background, the path integral computes the twisted partition function, or *supersymmetric index* of $\mathsf{T}_{\mathrm{NS5}}$, defined by a trace with respect to the Hilbert space on $M \times \Sigma$ in the presence of defects $\mathcal{E}_{\mathrm{D}p_i}$ inserted on $N_i \times C_i$ (recall that for a closed manifold with thermal circle, QFT produces a number represented as a trace, such as (2.3)).

A salient feature of supersymmetric indices is that they are protected against continuous changes of various parameters of the theory. This means that the index of our theory is invariant under deformations of the geometric data of Σ and C_i, such as the metric on Σ and the shapes of C_i. In other words, the theory $\mathsf{T}_{\mathrm{NS5}}$ on $S^1 \times M \times \Sigma$ is topological on Σ, as far as the computation of the index is concerned.

To relate the present setup to the the previous situation (2.70), we apply T-duality along the circle S^1. It turns Dp_i into D$(p-1)$-branes D$(p-1)_i$, localized at points u_i along the dual circle $\tilde{S}^1$, while sending the NS5-branes to those in the other type II string theory. The new NS5-branes produce the dual six-dimensional theory $\tilde{\mathsf{T}}_{\mathrm{NS5}}$ on $\tilde{S}^1 \times M \times \Sigma$, and in this theory D$(p-1)_i$ create $(p-1)$-dimensional defects $\mathcal{E}_{\mathrm{D}(p-1)_i}$ on $\{u_i\} \times N_i \times C_i$. Thus we are in the situation studied before, and the correlation function of $\mathcal{E}_{\mathrm{D}(p-1)_i}$ for this configuration coincides with the partition function of an integrable lattice model:

$$\left\langle \prod_{i=1}^{l} \mathcal{E}_{\mathrm{D}p_i}(C_i)\big[S^1 \times N_i\big] \right\rangle_{\mathsf{T}_{\mathrm{NS5}}[\Sigma], S^1 \times M} = Z_{\mathsf{L}\left(\tilde{\mathsf{T}}_{\mathrm{NS5}}\left[\tilde{S}^1 \times M\right]\right)}\big(\{\mathcal{E}_{\mathrm{D}(p-1)_i}(u_i; N_i)[C_i]\}\big). \quad (2.72)$$

Here the left-hand side is expressed in the original frame; it implicitly depends on each spectral parameter u_i through the holonomy $\exp(2\pi i u_i)$ around S^1 of the gauge field for the flavor symmetry $\mathrm{U}(1)_i$ supported on Dp_i. The holonomy appears in the index as a refinement parameter, or called *fugacity*, associated with $\mathrm{U}(1)_i$.

2.3.3 Defects as transfer matrices

Finally, we apply the construction developed so far to the main theme of this paper: integrable lattice models and defects as transfer matrices. Let us consider the brane construction for $p = 5$ case. To conform with the standard convention, take S-duality first and we still have D5- and NS5-branes

$$\begin{aligned} N\,\mathrm{D5} &\quad S^1 \times M \times \Sigma, \\ \mathrm{NS5}_i &\quad S^1 \times N_i \times \Sigma_i, \end{aligned}$$

where $\mathrm{NS5}_i$ create defects $\mathcal{E}_{\mathrm{NS5}_i}$ on $S^1 \times N_i \times C_i$.

One should notice that in this setup we necessarily have $N_i = M$ and thus the defects $\mathcal{E}_{\mathrm{NS5}_i}$ fill the whole $S^1 \times M$, which produces a four-dimensional theory

$$\mathsf{T}_{\mathrm{D5NS5}}[\Sigma]$$

with $\mathcal{N} = 1$ supersymmetry. Then we now have

$$\langle 1 \rangle_{\mathsf{T}_{\mathrm{D5NS5}}[\Sigma], S^1 \times M} = Z_{\mathsf{L}\left(\mathsf{T}_{\mathrm{D5NS5}}[S^1 \times M]\right)}\big(\{\mathcal{E}_i(S^1 \times M)[C_i]\}\big). \quad (2.73)$$

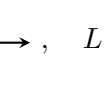
For example, when $M = S^3$, the left-hand side is given by the supersymmetric index for $\mathcal{N} = 1$ theory and the right-hand side corresponds to the partition function of Bazhanov-Sergeev integrable lattice model [22–25]. When $M = L(p, 1)$, lens space, the left-hand side is computed in [26], which defines a new integrable lattice model through this correspondence.

The brane tiling construction of integrable lattice models can be enriched by introducing additional defects. Besides the previously defined D5NS5-brane system, let us consider a D3-brane such as

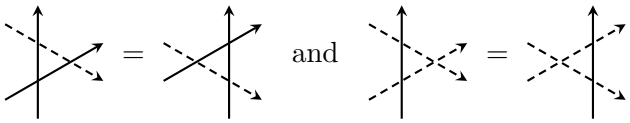
$$\text{D3} \quad S^1 \times N \times C \times \mathbb{R}_+, \tag{2.74}$$

$$\text{D3}' \quad S^1 \times \{0\} \times C' \times \mathbb{R}^2, \tag{2.75}$$

where N is a curve in M and C, C' are one-cycles on Σ. These D3-branes create new extended defect operators elongated in $S^1 \times N$ and S^1, respectively. A single D3-brane insertion corresponds to a new oriented line C, C' in the integrable lattice model, which we represent by a dashed line. Now that we have two kinds of lines originating from five-branes and a D3-brane, we can define three kinds of R-matrices:

$R =$  , $L =$  , $\mathcal{R} =$  . (2.76)

The middle one is usually called *L-operator*. Correspondingly, we have four Yang-Baxter equations, involving zero to three dashed lines. Those that involving one or two dashed line,



(2.77)

are called *RLL relations*. The effect of the insertion of such an additional defect on the lattice model is seen in terms of the L-operator. The neighborhood of the dashed line looks like

 . (2.78)

This diagram shows that the defect acts on the lattice model by a transfer matrix constructed from L-operators. Thus, the insertion of a defect operator in four-dimensional theory represented by (2.74) or (2.75) is mapped into lattice model side as the action of a transfer matrix constructed from L-operators.

What we will discuss in detail in the subsequent sections are the introduction of a single D3-brane (2.74) and (2.75) to the D5NS5-brane system, and investigate the correspondence between supersymmetric gauge theories and integrable lattice models with the additional defects. In the next two sections, we are exploring the followings:

1. For the case of single D3 (2.74), let $M = S^3$, $\Sigma = T^2$, and $N = S^1$. Then the D3 creates a surface defect on $S^1 \times S^1 \subset S^1 \times S^3$ and it acts on the supersymmetric index as a transfer matrix in the corresponding lattice model:

$$\langle 1 \rangle_{\mathsf{T}_{\text{D5NS5}}[\Sigma], S^1 \times S^3} = \mathcal{I}_{S^1 \times S^3}(p, q, t), \tag{2.79}$$

$$Z_{\mathsf{L}\left(\mathsf{T}_{\mathrm{D5NS5}}[S^1\times S^3]\right)}\left(\{\mathcal{E}_i(S^1\times S^3)[C_i]\}\right) = Z_{\mathrm{Bazhanov-Sergeev}}, \tag{2.80}$$

the surface defect index is represented as a difference operator acting on the original supersymmetric index [27, 28],

$$\left\langle S_{(r,s)}\right\rangle_{\mathsf{T}_{\mathrm{D5NS5}}[\Sigma]} = \mathfrak{S}_{(r,s)}\,\mathcal{I}_{S^1\times S^3}(p,q,t). \tag{2.81}$$

2. For the case of single D3′ (2.75), let $M=\mathbb{R}^3$ and $\Sigma = T^2$. D3′ creates a line defect wrapping the thermal circle S^1 and it acts on the quantum Hilbert space of a spin chain as a transfer matrix, since now we have non-compact three-manifold $\mathbb{R}^3$. This spin chain has a equivalent lattice model description, and it is really an integrable model of trigonometric type. The line defect in the four-dimensional theory is realized as a Wilson-'t Hooft line operator T, and its magnetic charge and electric charge $(\mathbf{m},\mathbf{e})$ is specified by the one-cycle C' winding around the torus T^2. As it turns out, the vacuum expectation values (vevs) of Wilson-'t Hooft lines realize the deformation quantization of the Hitchin moduli space, and are naturally quantized by the Weyl quantization:

$$\text{Weyl quantization of } \left\langle T_{(\mathbf{m},\mathbf{e})}\right\rangle = \text{trigonometric transfer matrix}. \tag{2.82}$$

In the following subsections, we study these correspondences. Sections 3 & 4 are almost independent.

3 Surface defects as transfer matrices

In this section, we mainly review the results in [29], in which the correspondence we introduced in section 2.3.3 was established. It turns out that a surface defect is identified with the transfer matrix constructed from Sklyanin's L-operator.

The structure of this section is as follows (see figure 6). We first apply to the brane tilings the general prescription to construct the correspondence between supersymmetric gauge theory and integrable lattice model developed in section 2.3. Brane tiling technique allows us to study quiver gauge theories in a systematic way, and taking the four-manifold to be $S^1\times S^3$, on which quiver gauge theories are defined, the path integral of the twisted partition function of the theory computes the supersymmetric index. For such a case, we will obtain a dictionary between supersymmetric gauge theory and integrable lattice model, which is called *Gauge/YBE correspondence*. The brane tiling construction can be mapped into the setups of the so-called class-$\mathcal{S}$ theories. We proceed to additional surface defects in the simplest case of the theories of class-$\mathcal{S}$. Then we find that the surface defects act on the supersymmetric index by difference operators (3.61) and they are identified with the transfer matrices of the corresponding integrable lattice model (3.62), which is the most important relation of this section:

$$\mathfrak{S}_{(0,1)} = \mathrm{Tr}\left(L^{\diamondsuit}(d,(c,b))\right), \tag{3.1}$$

with an appropriate parameter identification.

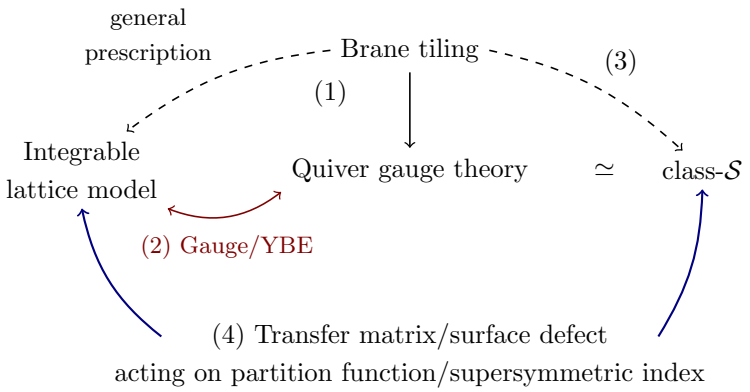


Figure 6: Structure of section 3. (1) & (2) Along the line of the general prescription developed in section 2.3, brane tiling technique realizes the Gauge/YBE correspondence. (3) Brane tiling construction of quiver gauge theories can be rephrased by the language of class-$\mathcal{S}$. (4) A surface defect acting on the index is identified with a transfer matrix acting on the partition function of a corresponding integrable lattice model.

3.1 Brane tilings and integrable lattice models

Let us begin with some generalities, before restricting ourselves to the case of $S^1 \times S^3$. In order to give the general argument of the correspondence, we make a brief review of the systematic brane construction of quiver gauge theories called *brane tilings*. For more details, the reader should be referred to the original papers and the excellent reviews [30–33]. Recall the 5-brane configuration given in section 2.3.3,

$$\begin{aligned} N\,\mathrm{D5} &\quad S^1 \times M \times \Sigma, \\ \mathrm{NS5}_i &\quad S^1 \times N_i \times \Sigma_i. \end{aligned}$$

For such a configuration, what one needs to notice is that when an NS5-brane meets N D5-branes, they combine to form a bound state. In the language of (p, q) 5-branes, this bound state is either an $(N, 1)$ or $(N, -1)$ 5-brane, depending on the relative positions of the branes; see figure 7. Therefore, the extended defects $\mathcal{E}_{\mathrm{NS5}_i}$ are domain walls in T_{D5} separate the spacetime into the regions with different values of the NS5-brane charge q. The curves C_i along which these domain walls are located are known as *zig-zag paths*. Across a zig-zag path the charge q jump by one.

Conversely, given a configuration of curves C_i on the two-dimensional surface Σ and a 5-brane charge assignment consistent with it, we can construct a 5-brane system whose zig-zag paths are C_i: we take NS5-branes approaching the D5-branes from transverse directions, and let them meet along C_i and form bound states over regions with $q \neq 0$. Such a 5-brane system is called a brane tiling on Σ [30, 31].

As we have explained in the last section, a brane tiling gives rise to a four-dimensional $\mathcal{N} = 1$ theory. A concrete description of this theory is known for the subset of brane tilings that involve only $(N, 0)$ 5-branes (i.e. N coincident D5-branes) and $(N, \pm 1)$ 5-branes. Given

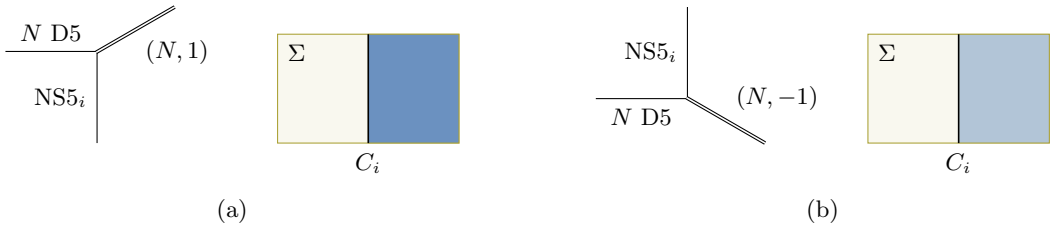


Figure 7: An NS5-brane combines with a stack of N D5-branes, forming (a) an $(N,1)$ 5-brane or (b) an $(N,-1)$ 5-brane. The 5-brane junction is a domain wall in T_{D5}. The shaded regions shown above support a nonzero NS5-brane charge $q = \pm 1$. In the context of brane tiling, the curves C_i are called zig-zag paths.

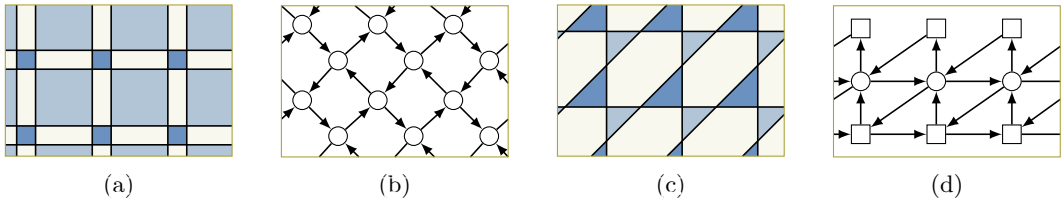
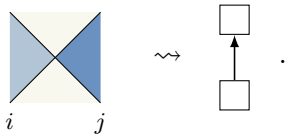


Figure 8: (a) A brane tiling on a torus. (b) The periodic quiver associated with (a). (c) A brane tiling on a finite-length cylinder. (d) the quiver for (c).

a brane tiling in this subset, we indicate $(N,1)$ and $(N,-1)$ 5-brane regions by dark and light shading, respectively, while leaving $(N,0)$ regions unshaded (see figure 7). After the shading, we get a checkerboard-like pattern on Σ where shaded faces adjoin unshaded ones and two shaded faces sharing a vertex are of different types, see figure 8(a) and 8(c).

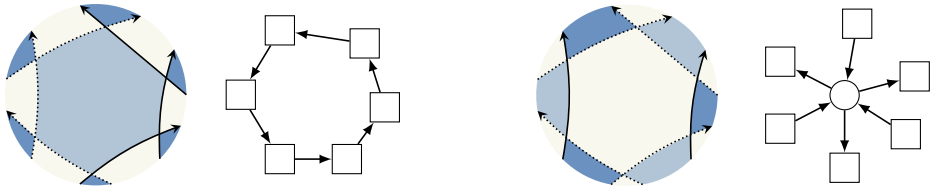
Each unshaded region supports N D5-branes, hence an $\mathrm{SU}(N)$ vector multiplet lives there. If the region contains some part of the boundary, the multiplet is frozen by boundary conditions and the associated symmetry is an $\mathrm{SU}(N)$ flavor symmetry; in quiver notation, we represent a dynamical vector multiplet by a gauge node ○ and a non-dynamical one by a flavor node □. From open strings stretched between two unshaded regions (namely ending on N D5-branes partitioned by NS5-brane) sharing a vertex, we get a chiral multiplet that transforms in the fundamental representation under one of the associated gauge or flavor groups and in the anti-fundamental representation under the other. We write it by an arrow between the two nodes:



(3.2)

The arrow points from the anti-fundamental side to the fundamental side. See figure 8 for examples of quivers obtained from brane tilings.

Moreover, for every set of zig-zag paths bounding a shaded region, we have a loop of arrows and world-sheet instantons generate a superpotential term given by the trace of the



(a) (b)

Figure 9: Zig-zag paths bounding (a) a shaded region and (b) an unshaded region. In either case, the R-charges of two of the arrows are different from those of the rest.

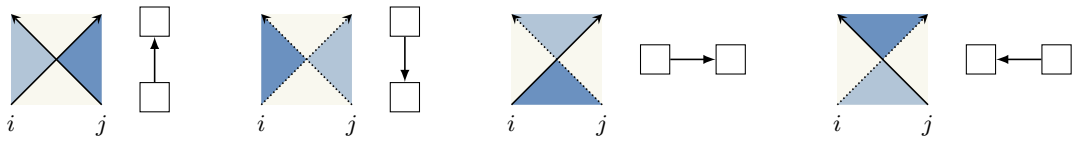
product of the bifundamental chiral multiplets in the loop. The coefficient of this term is positive or negative depending on whether the direction of the loop is clockwise or counter-clockwise. Thus, the four-dimensional theory realized by a brane tiling in the subset under consideration is an $\mathcal{N} = 1$ supersymmetric gauge theory described by a quiver with potential drawn on Σ.

Each $\mathrm{NS5}_i$ supports a U(1) flavor symmetry $\mathrm{U}(1)_i$. An arrow is charged under $\mathrm{U}(1)_i$ if it is crossed by C_i. The charge F_i of $\mathrm{U}(1)_i$ can be normalized in such a way that the arrow in (3.2) has $F_i = -1$ and $F_j = +1$. The diagonal combination of all $\mathrm{U}(1)_i$ acts on the theory trivially since every arrow is crossed by exactly two zig-zag paths from the opposite sides.

The theory also has an R-symmetry $\mathrm{U}(1)_R$. Its definition is not unique as the R-charge R can be shifted by a linear combination of U(1) flavor charges. However, the R-charge assignment is constrained by two conditions. The first is that $\mathrm{U}(1)_R$ must be unbroken by the superpotential and therefore the R-charges of the chiral multiplets contained in each superpotential term must add up to two. The second is that $\mathrm{U}(1)_R$ must be free of anomaly. This requires that for every gauge node, the sum of the R-charges of the arrows starting from or ending at that node must equal the number of the arrows minus two.

To fix the R-charge assignment, let us assume that we can orient the zig-zag paths and bound every shaded or unshaded region with zig-zag paths all heading upward, for some choice of the "vertical" direction in the neighborhood of that region. This is the case for the examples in figure 8. The zig-zag paths are thus oriented and fall into two groups; when a zig-zag path goes upward and we cross it from the left to the right, q increases by one. We distinguish the latter case from the former by drawing the zig-zag path with a dotted line. Then, we give an arrow $R = 0$ if it originates from a crossing of two zig-zag paths of the same type, and $R = 1$ otherwise. With this R-charge assignment the two conditions described above are satisfied; see figure 9.

Summarizing the rules for assignment of the charges, we can read off the quiver diagram from zig-zag paths as in figure 10.



(a) (b) (c) (d)

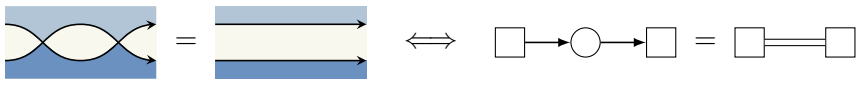
Figure 10: The rule for assigning a quiver to a brane tiling diagram. The arrows in (a) and (b) have $(R, F_i, F_j) = (0, -1, 1)$. Those in (c) and (d) have $(R, F_i, F_j) = (1, 1, -1)$.

3.1.1 Integrable lattice models from quiver gauge theories

From the supersymmetric index of the four-dimensional $\mathcal{N} = 1$ theory realized by a brane tiling, we obtain an integrable lattice model defined on the lattice $\{C_i\}$ consisting of the zig-zag paths. Each C_i carries a spectral parameter u_i. S-duality followed by T-duality on S^1 turns $\mathrm{NS5}_i$ into a D4-brane, and its coordinate on the dual circle $\tilde{S}^1$ is u_i. Instead, we can apply T-duality on S^1 and lift $\mathrm{NS5}_i$ to an M5-brane, then u_i becomes the coordinate on the M-theory circle. Either way, u_i is determined by the holonomy of the U(1) gauge field on $\mathrm{NS5}_i$ along S^1.

If the theory is described by a quiver diagram, translation between the gauge theory and the lattice model goes as follows [25, 26]. Nodes in the quiver diagram are interpreted as spin sites. For each flavor node, we can turn on a holonomy of the associated gauge field. The index depends on the conjugacy class of the holonomy, which is uniquely represented by a diagonal matrix $\mathrm{diag}(z_1, \ldots, z_N)$ up to permutations of the entries. The index is therefore a symmetric function of the U(1)-valued variables $(z_1, \ldots, z_N)$ obeying the constraint $z_1 \cdots z_N = 1$. These variables are fugacities for the $\mathrm{SU}(N)$ flavor symmetry and parametrize the value of the spin at this node, and thus the spins take values in the maximal torus $\mathrm{U}(1)^{N-1}$ of $\mathrm{SU}(N)$. For a gauge node, integration is performed over the fugacities since its gauge field is a path integral variable. This is the summation over the values of a spin placed on an internal face. Finally, arrows in the quiver diagram represent interactions between spins.

The unitarity relations are satisfied if the contributions to the index from arrows with $R = 0$ are properly normalized. For example, consider the relation[12]

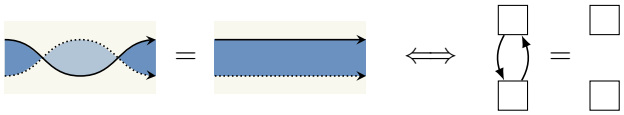


(3.3)

where the right-hand side is a "delta function" that equates two flavor nodes when one of them is gauged. The theory on the left-hand side is SQCD with N colors and N flavors. It exhibits confinement and has a vacuum in which the mesons take nonzero expectation values and the flavor symmetry $\mathrm{SU}(N) \times \mathrm{SU}(N)$ is broken to the diagonal subgroup [34].

[12] By an equality of two quiver diagrams, we mean that the supersymmetric indices of the theories described those quivers are equal. The precise meaning will be given for the $S^1 \times S^3$ case below.

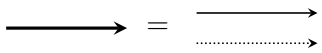
The index computed in this vacuum is given by the right-hand side, provided that we cancel the contributions from the surviving baryon and anti-baryon. Another unitarity relation



(3.4)

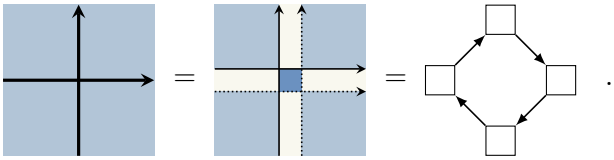
holds since the two arrows on the left-hand side form a loop and generates a mass term in the superpotential. We can send the mass to infinity so that these arrows decouple from the theory, and are left with the right-hand side.

The Yang-Baxter equation with three zig-zag paths is harder to understand, as it always involves (N, q) region with $|q| > 1$ and a quiver description is not available. The problem stems from the fact that our effects are domain walls across which q changes. To circumvent the difficulty, we take a pair of zig-zag paths of different types and think of it as a single line:



(3.5)

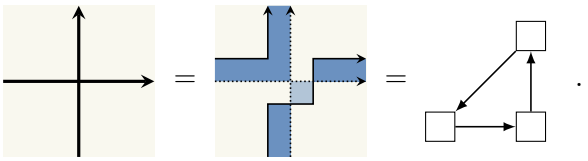
This line does not alter the value of q. Taking two copies of this line and placing them in an $(N, -1)$ background, we can make the R-matrix



(3.6)

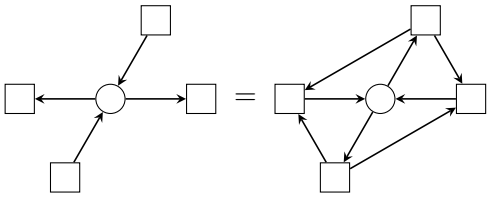
A lattice model constructed from this R-matrix is a vertex model whose quiver consists of diamonds of arrows; see figure 8(b). The vector space carried by a line is the space of symmetric functions of fugacities $(z_1, \ldots, z_N)$.

Alternatively, we can place these lines in an $(N, 0)$ background and force them to exchange their constituent zig-zag paths as they cross:



(3.7)

This R-matrix leads to an IRF model described by a quiver with triangles of arrows, as shown in figure 8(d). The corresponding Yang-Baxter equation, after cancellation of some factors with the help of the unitarity relation (3.3), reads



(3.8)

The two sides are related by Seiberg duality [35] for SQCD with N colors and $2N$ flavors, so their indices are indeed equal. The Yang-Baxter equation in the lattice model in this case

is hence also identified with the Seiberg duality in gauge theory side. The relation between the Yang-Baxter move and Seiberg duality was first pointed out in [36] and established in [25]. The Yang-Baxter equation for the R-matrix (3.6), which is our main focus in the next subsection though more complicated, also follows from this equality.

3.1.2 Supersymmetric index on $S^1 \times S^3$ and Gauge/YBE correspondence

Let us be more specific. Based on the observations so far, here we demonstrate that for the case of $M = S^3$, namely given the four-manifold $S^1 \times S^3$, the supersymmtric index indeed matches the partition function of an integrable lattice model. Furthermore, in this case the correspondence between supersymmetric gauge theory and integrable lattice model has a clear one-by-one dictionary, as we will see.

We now focus on the case $M = S^3$, in which geometry the supersymmetric index is well-studied from the works by [37–39]. Parametrize S^3 by two complex variables (ζ_p, ζ_q) satisfying $|\zeta_p|^2 + |\zeta_q|^2 = 1$, and denote the isometry groups acting on ζ_p and ζ_q by $\mathrm{U}(1)_p$ and $\mathrm{U}(1)_q$, respectively. We take $S^1 \times S^3$ to be a twisted product; we prepare a trivial S^3-fibration over an interval $[0, \beta]$ and identify the fibers at the ends of the base using an isometry $\left(e^{\mathrm{i}\theta_p}, e^{\mathrm{i}\theta_q}\right) \in \mathrm{U}(1)_p \times \mathrm{U}(1)_q$. On this spacetime, the partition function of the quiver gauge theory realized by a brane tiling gives the supersymmetric index refined by the isometries and the flavor symmetries.

The index is defined as a trace of refined Boltzmann weight over the space of states on S^3, which is computed exactly using state-operator correspondence in conformal case or by localization of the path integral. For $\mathcal{N} = 1$ quiver gauge theories on $S^1 \times S^3$, it is

$$\mathcal{I}(p, q, \{a_i\}) = \mathrm{Tr}_{\mathcal{H}_{S^3}} \left((-1)^F p^{j_1 + j_2 + R/2} q^{j_1 - j_2 + R/2} \prod_i a_i^{F_i} \right), \tag{3.9}$$

where the trace is taken over the space $\mathcal{H}_{S^3}$ of states on S^3. Here $(-1)^F$ is the fermion parity, and j_1, j_2 are generators of the maximal torus $\mathrm{U}(1)_1 \times \mathrm{U}(1)_2$ of the isometry group $\mathrm{Spin}(4) \simeq \mathrm{SU}(2)_1 \times \mathrm{SU}(2)_2$ of S^3, $\{a_i\}$ and p, q are complex parameters.

Building blocks of 4d $\mathcal{N} = 1$ supersymmetric quiver gauge theories are vector multiplets and bifundamental chiral multiplets. A vector multiplet is present at a gauge node. A bifundamental chiral multiplet has two flavor groups, say $\mathrm{SU}(N)_z$ and $\mathrm{SU}(N)_w$[13] . The full index is generically given by a combination of vector and chiral multiplets involved in the theory. These multiplets are expressed in terms of the elliptic gamma function

$$\Gamma(z; p, q) = \prod_{j,k=0}^{\infty} \frac{1 - p^{j+1} q^{k+1} z^{-1}}{1 - p^j q^k z} \tag{3.10}$$

with $p = e^{-\beta + \mathrm{i}\theta_p}$ and $q = e^{-\beta + \mathrm{i}\theta_q}$. To write down the formula, let $a_i = e^{2\pi \mathrm{i} u_i}$ be the fugacity

[13] We label gauge and flavor groups by their associated fugacities, say z. The same applies to the nodes appearing below, since the rank of them is fixed.

for the flavor group $\mathrm{U}(1)_i$ associated with the ith zig-zag path. Also, we introduce the Pochhammar symble $(z;q)_\infty = \prod_{k=0}^{\infty}\left(1-q^k z\right)$.

Given a theory $\mathcal{T}$ with flavor group $\mathrm{SU}(N)_w$ and another theory $\mathcal{T}'$ with flavor group $\mathrm{SU}(N)_{w'}$, we can couple them to obtain a new theory $\left(\mathcal{T}\times\mathcal{T}'\right)/\mathrm{SU}(N)_z$ by gauging the diagonal subgroup $\mathrm{SU}(N)_z$ of $\mathrm{SU}(N)_w \times \mathrm{SU}(N)_{w'}$. To construct a quiver gauge theory, we take a number of bifundamental chiral multiplets and couple them by gauging all or part of the flavor nodes.

The index $\mathcal{I}_\mathcal{T}$ of a 4d $\mathcal{N}=1$ theory $\mathcal{T}$ with flavor group $\mathrm{SU}(N)_z$ is a symmetric meromorphic function of the fugacities $z_1,\ldots,z_N$. This symmetric property reflects the gauge invariance of the index. At the level of the index, gauging of a flavor group is realized by introduction of the corresponding vector multiplet and integration over its fugacities. For example,

$$\mathcal{I}_{(\mathcal{T}\times\mathcal{T}')/\mathrm{SU}(N)_z} = \int_{\mathbb{T}^{N-1}} \prod_{I=1}^{N-1} \frac{dz_I}{2\pi i z_I} \mathcal{I}_V(z)\mathcal{I}_\mathcal{T}(z)\mathcal{I}_{\mathcal{T}'}(z), \tag{3.11}$$

with the integration performed over the unit circle $\mathbb{T}$ for each variable z_I. The index $\mathcal{I}_\mathcal{T}(z)$ of the vector multiplet is given by elliptic gamma functions:

$$\textcircled{z} \;= \mathcal{I}_V(z;p,q) = \frac{(p;p)_\infty^{N-1}(q;q)_\infty^{N-1}}{N!} \prod_{\substack{I,J=1\\ I\neq J}}^{N} \frac{1}{\Gamma(z_I/z_J;p,q)}. \tag{3.12}$$

Here and in the following, we are often using a quiver to mean the index of the corresponding theory. From now on we fix p, q and omit them from the notation unless needed.

The index of a bifundamental chiral multiplet with fugacity a is given by

$$\boxed{z} \xrightarrow{\;a\;} \boxed{w} \;= \mathcal{I}_B(z,w;a) = \prod_{I,J=1}^{N} \Gamma\left(a\frac{w_I}{z_J}\right). \tag{3.13}$$

This function satisfies

$$\mathcal{I}_B(z,w;a)\mathcal{I}_B\left(w,z;\frac{pq}{a}\right) = 1. \tag{3.14}$$

This identity corresponds the unitarity relation (3.4) and says that as far as the index is concerned, we can cancel a pair of arrows making a loop if their R-charges add up to 2 and flavor charges add up to 0:

$$\boxed{z} \underset{pq/a}{\overset{a}{\rightleftarrows}} \boxed{w} \;=\; \boxed{z} \qquad \boxed{w} \quad . \tag{3.15}$$

Physically, the reason is that we can turn on mass term for such a pair. The index is invariant under this deformation, and the bifundamental chiral multiplets decouple from the theory if we send the mass to infinity, leaving a trivial contribution to the index. We will make use of this identity frequently.

Another useful fact is that if we define the "delta function"

$$\boxed{z} \;=\!=\; \boxed{w} \tag{3.16}$$

by the relation

$$\mathcal{I}_{\mathcal{T}}(w) = \int_{\mathbb{T}^{N-1}} \prod_{I=1}^{N-1} \frac{dz_I}{2\pi i z_I} \mathcal{I}_V(z)\mathcal{I}_{\mathcal{T}}(z) \boxed{z} \!=\!=\!= \boxed{w} \,, \tag{3.17}$$

then we have

$$\begin{aligned} \boxed{z} \xrightarrow{a} \textcircled{x} \xrightarrow{a^{-1}} \boxed{w} &= \int_{\mathbb{T}^{N-1}} \prod_{I=1}^{N-1} \frac{dx_I}{2\pi i x_I} \mathcal{I}_V(x)\mathcal{I}_B(z,x;a)\mathcal{I}_B\left(x,w;a^{-1}\right) \\ &= \Gamma\left(a^{\pm N}\right) \boxed{z} \!=\!=\!= \boxed{w} \quad . \end{aligned} \tag{3.18}$$

This corresponds to the other unitarity relation (3.3) and implies a consequence of confinement and chiral symmetry breaking [34, 40]. At low energies the theory on the left-hand side is described by the mesons and the baryons. It has a vacuum in which the mesons take nonzero expectation values and the flavor symmetry $\mathrm{SU}(N)_w \times \mathrm{SU}(N)_z$ is broken to the diagonal subgroup. In this vacuum the fugacities w and z are identified, so we get the quiver on the second line. The $\Gamma(a^{\pm N}) := \Gamma(a^N)\Gamma(a^{-N})$ is the contribution from the baryons.

So, the unitarity relation (3.3) is satisfied if we normalize the contribution from each arrow with $R = 0$ by dividing it by the factor $\Gamma\big(\prod_i a_i^{NF_i}; p, q\big)$, which cancels the contribution from the corresponding baryon. The Yang-Baxter equation (3.8) is an integral identity obeyed by the elliptic gamma function [41–43].

We can readily write down the formula for the index of a general quiver gauge theory. For simplicity, suppose that the theory is described by a quiver that contains no flavor node. Then, the index is computed by

$$\prod_{\textcircled{z}} \int_{\mathbb{T}^{N-1}} \prod_{I=1}^{N-1} \frac{dz_I}{2\pi i z_I} \mathcal{I}_V(z) \prod_{\textcircled{x}\to\textcircled{y}} \mathcal{I}_B(x,y), \tag{3.19}$$

where the two products are taken over all nodes and all arrows, respectively. the index is a function of the parameters p, q and the flavor fugacities a_i, which are suppressed in the above expression. If the quiver contains flavor nodes, the index is also a function of their fugacities.

In the expression of the supersymmetric index (3.19), it is manifest that the index of a quiver gauge theory may be interpreted as the partition function of a statistical mechanical model with continuous spins. Indeed, this formula precisely computes the partition function of a spin model in which spins are placed at gauge nodes [1, 25]. The spin variables at the gauge nodes are the fugacities $z_1, \ldots, z_N$, and they interact among themselves as well as with spins at nearest-neighbor nodes, namely those connected by arrows. The Boltzmann weights for the self-interaction and the nearest-neighbor interaction are $\mathcal{I}_V$ and $\mathcal{I}_B$, respectively; see figure 11. What is more, the quantities in four-dimensional gauge theories really have one-to-one correspondence in the lattice model side. The dictionary is given by table 1 and such a correspondence is called *Gauge/YBE correspondence*.

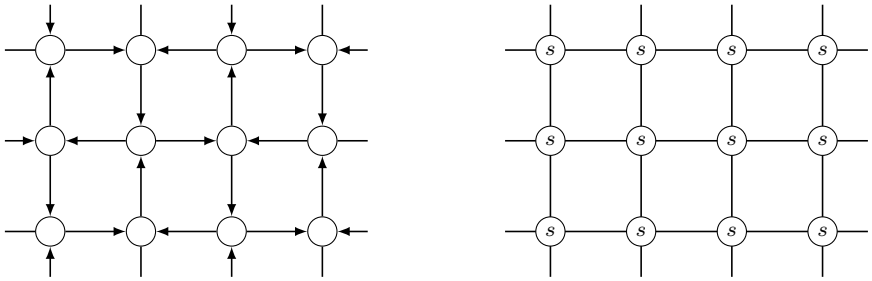


Figure 11: (a) A quiver diagram specifying a gauge theory. On each node and each edge, vector multiplet and bifundamental chiral multiplet contribute to the index. (b) A spin system corresponding to the quiver diagram (a), defined on the same lattice. Spin configurations and interaction specifies a lattice model. The nearest-neighbor interaction may be chiral.

Table 1: Dictionary for the Gauge/YBE correspondence

Integrable model	Quiver gauge theory
Spin lattice	Quiver diagram
Rapidity line	Zig-zag path
Spectral parameter	R-charge
Statistical partition function	Supersymmetric partition function (index)
Temperature-like parameters	Quantum parameters (such as p, q)
Spin variables	Gauge holonomies along non-contractible cycles
Number of spin components	Rank of a gauge group
Self-interaction	Vector multiplet
Nearest-neighbor interaction	Bifundamental matter multiplet
Star-star relation	Seiberg(-like) duality
R-matrix	
Composition of R-matrices	
Yang-Baxter equation	

Table 2: D4-NS5 brane configuration

	0	1	2	3	4	5	6	7	8	9
D4	×	×	×	×			×			
NS5	×	×	×	×	×	×				

3.2 Class-$\mathcal{S}$ theories in brane tilings

3.2.1 $\mathcal{N}=2$ quiver theories to the theories of class-$\mathcal{S}$

Building blocks of 4d $\mathcal{N}=1$ supersymmetric quiver gauge theories are vector multiplets and bifundamental chiral multiplets. A vector multiplet is present at a gauge node. A bifundamental chiral multiplet has two flavor groups, say $\mathrm{SU}(N)_z$ and $\mathrm{SU}(N)_w$.

Given a theory $\mathcal{T}$ with flavor group $\mathrm{SU}(N)_w$ and another theory $\mathcal{T}'$ with flavor group $\mathrm{SU}(N)_{w'}$, we can couple them to obtain a new theory $(\mathcal{T}\times\mathcal{T}')/\mathrm{SU}(N)_z$ by gauging the diagonal subgroup $\mathrm{SU}(N)_z$ of $\mathrm{SU}(N)_w\times\mathrm{SU}(N)_{w'}$. To construct a quiver gauge theory, we take a number of bifundamental chiral multiplets and couple them by gauging all or part of the flavor nodes.

Now we aim to check the proposal on surface defects and transfer matrices by comparing them with independent calculations. In this section we perform the simplest such check for surface defects in A_1 theories of class $\mathcal{S}$ [44, 45], which arise from compactification of the 6d $\mathcal{N}=(2,0)$ theory of type A_1 on punctured Riemann surfaces. The action of surface defects on the supersymmetric indices of class-$\mathcal{S}$ theories have been studied in [27,28,46,47]. Here we first review the computation for the surface defect labeled with the fundamental representation of SU(2) based on the method developed in [27], and show that the result agrees with the prediction from the transfer matrix (3.47).

Typical examples of class-$\mathcal{S}$ theories are $\mathcal{N}=2$ gauge theories characterized by linear and circular quivers with $\mathrm{SU}(N)$ nodes. They are actually also examples of brane tiling models discussed in the previous sections. As such, they allow us to translate key notions in class-$\mathcal{S}$ theories to the language of brane tilings, and vice versa. So let us first describe these theories as class-$\mathcal{S}$ theories as well as brane tiling models, and understand the relation between the two descriptions. Although we will mainly work with $N=2$, for now we keep N general.

Let us consider the standard type IIA brane configuration for an $\mathcal{N}=2$ linear quiver theory with $m+1$ nodes. It consists of N D4-branes spanning the 01236 directions, intersected by m -branes extending along the 012345 directions (see table 2). This brane configuration is lifted in M-theory to M5-branes, wrapped on a cylinder with m punctures created by intersecting M5-branes. Therefore, the $\mathcal{N}=2$ linear quiver theory is obtained by compactification of the 6d $\mathcal{N}=(2,0)$ theory of type A_{N-1} on a cylinder with m punctures, or a sphere with $m+2$ punctures. We distinguish the two punctures coming from the ends of the cylinder from the m punctures in between. They are referred to as maximal and minimal

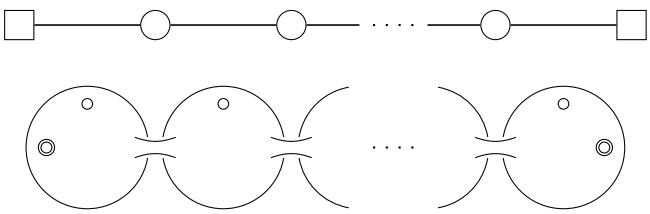


Figure 12: An $\mathcal{N} = 2$ linear quiver theory associated to a punctured sphere.

punctures, respectively. In the class-$\mathcal{S}$ language, the $\mathcal{N} = 2$ linear quiver theory is a theory associated to a sphere with 2 maximal and m minimal punctures, see figure.

R-symmetry of the theory is $\mathrm{SU}(2)_I \times \mathrm{U}(1)_r$, where $\mathrm{SU}(2)_I$ represents the rotation symmetry of the 789-space, and $\mathrm{U}(1)_r$ the rotation symmetry of the 45-plane. The $\mathrm{SU}(N)$ flavor node from each end of the quiver is associated to the maximal puncture on the corresponding side of the sphere. The ith gauge node is associated to the tube between the ith and $(i+1)$th minimal punctures. To the ith minimal puncture is associated a flavor symmetry $\mathrm{U}(1)_i$ which acts on the hypermultiplet charged under the $(i-1)$th and ith gauge nodes.

Following the philosophy of class-$\mathcal{S}$ theories, we decompose this theory into basic building blocks by decoupling gauge fields. Roughly, the gauge coupling of the ith gauge node is inversely proportional to the length between the ith and $(i+1)$th minimal punctures. To make the gauge couplings small, we take the minimal punctures far apart from one another. Then the surface looks like a string of m spheres, each containing a single minimal puncture, connected by long tubes. The smaller the gauge couplings get, the longer the tubes become, and eventually these spheres split up as the couplings go to zero. Each of the spheres represents a bifundamental hypermultiplet, which is a linear quiver with $m = 1$, so it has one minimal and two maximal punctures. The quiver thus breaks into a collection of three-punctured spheres, or trinions.

Conversely, a sphere with two maximal and m minimal punctures is obtained by gluing m trinions together, namely by replacing pairs of maximal punctures with tubes. In general, we can connect two Riemann surfaces with a tube at maximal punctures. From the point of view of gauge theory, gluing corresponds to gauging the diagonal combination of the $\mathrm{SU}(N)$ flavor symmetries associated to the maximal punctures involved. Using trinions with one minimal and two maximal punctures, we can obtain any linear quiver in this way, and for that manner also a circular quiver by further gluing the two ends of a linear quiver together. In this sense, these trinions are building blocks for linear and circular quivers. As these two kinds of quivers can be treated essentially in the same manner, we will focus on linear quivers in the followings.

To make contact with brane tilings, what we need to do is to find the counterpart of trinion, building block of quiver gauge theories, in brane tiling systems. To do this, we describe the $\mathcal{N} = 2$ linear quiver theory as an $\mathcal{N} = 1$ quiver gauge theory. In terms of $\mathcal{N} = 1$ supermultiplets, the $\mathcal{N} = 2$ vector multiplet for the ith gauge node decomposes into a vector

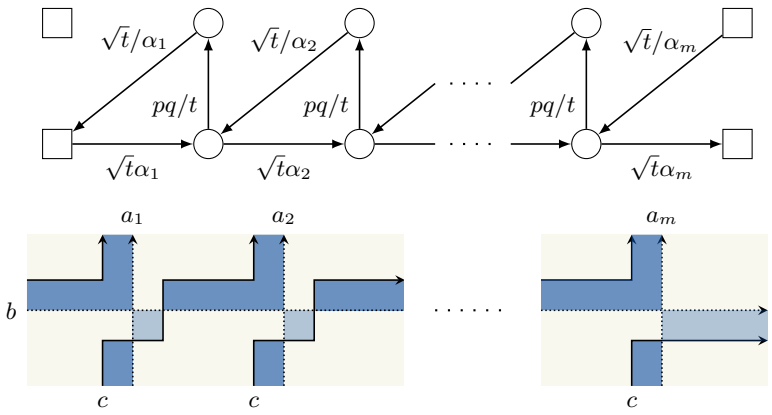


Figure 13: An $\mathcal{N}=2$ linear quiver as a brane tiling model. In the quiver, the two nodes in the same column are identified. In the brane tiling diagram, the vertical direction is periodic.

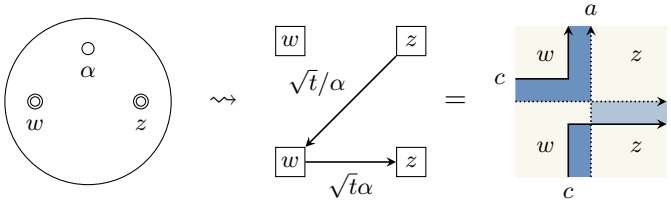
multiplet and a chiral multiplet Φ_i in the adjoint representation with $(r, I_3) = (-1, 0)$, while the ith hypermultiplet consists of two bifundamental chiral multiplets Q_i, $\tilde{Q}_i$ with $(r, I_3) = (0, 1/2)$. Here I_3 is a Cartan generator of $\mathrm{SU}(2)_I$. The pair $(Q_i, \tilde{Q}_i^\dagger)$ transforms in the doublet of $\mathrm{SU}(2)_I$ and have $\mathrm{U}(1)_i$ charge $F_i = -1$. From the point of view of $\mathcal{N}=1$ supersymmetry, the U(1) symmetry generated by the combination

$$\mathcal{F} = r + I_3 \tag{3.20}$$

is a flavor symmetry. We denote the fugacity for $\mathcal{F}$ by t. For the standard definition of the $\mathcal{N}=2$ index, r and $\mathcal{F}$ enter the trace through the combination $(pq)^{-r}t^{\mathcal{F}}$. Then, the fugacities of Q_i, $\tilde{Q}_i$ and Φ_i are $\sqrt{t}/\alpha_i$, $\sqrt{t}\alpha_i$ and pq/t, respectively.

It is helpful for us to prepare two copies for each node of the quiver and impose identification between them. We draw the arrows in such a way that Φ_i connects the two copies of the ith node and makes a triangle with Q_i and $\tilde{Q}_i$, as in figure 13. Drawn in this form, it is clear that the $\mathcal{N}=2$ linear quiver is a special case of the triangle quiver described in section 3.1, except that the vertical arrow is missing between the flavor nodes at the right end. The corresponding brane tiling diagram is therefore essentially the same, as shown in figure. Note that the cubic superpotentials, generated around the triangles by world-sheet instantons, are precisely what we need for the theory to have $\mathcal{N}=2$ supersymmetry.

As we can split the $(m+1)$-punctured sphere into a collection of m trinioins, we can also break the brane tiling diagram into basic pieces. Each piece represents a single trinioin and is made of three zig-zag paths:



(3.21)

Gluing two trinions corresponds to concatenating two such diagrams side by side. In the course of this operation, we must interchange the positions of the zig-zag paths labeled b and c near the glued side of one of the diagrams. This results in an additional vertical arrow in the combined quiver, which is the adjoint chiral multiplet in the $\mathcal{N}=2$ vector multiplet used in the gauging.

Let us find the relationship between the convention we use for brane tilings and that used above. The R-charge R in the brane tiling model is given in terms of the charges of the $\mathcal{N}=2$ theory by

$$R = R_0 + \frac{1}{2}\sum_i F_i, \quad R_0 = -r + I_3. \tag{3.22}$$

The flavor charges associated to the zig-zag paths can be written as

$$F_{a_i} = -F_i, \quad F_b = -\mathcal{F} + \frac{1}{2}\sum_i F_i, \quad F_c = \mathcal{F} + \frac{1}{2}\sum_i F_i. \tag{3.23}$$

Without loss of generality, we can set

$$a_i = \frac{1}{\alpha_i}. \tag{3.24}$$

Plugging these relations into the combination $(pq)^{R/2}\prod_i a_i^{F_{a_i}} b^{F_b} c^{F_c}$ that enters the indices of the bifundamental chiral multiplets, we deduce

$$b = \frac{1}{\sqrt{t}}, \quad c = \sqrt{\frac{t}{pq}}. \tag{3.25}$$

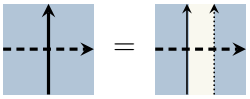
Before proceeding, we should mention a peculiarity in the A_1 case. When $N=2$, the U(1) flavor symmetry of a bifundamental hypermultiplet is enhanced to SU(2) due to the fact that the fundamental representation of SU(2) is pseudoreal. For this reason there is no distinction between minimal and maxmal punctures, and each trinion can be regarded as a half-hypermultiplet in the trifundamental representation of $\mathrm{SU}(2)^3$. This is reflected in the index of a trinion,

$$\mathcal{I}_B(w,z;\sqrt{t}a)\mathcal{I}_B(z,w;\sqrt{t}/a) = \Gamma(\sqrt{t}\alpha^{\pm 1}z^{\pm 1}w^{\pm 1}), \tag{3.26}$$

which is manifestly symmetric under permutation of a, z and w.

3.3 Surface defects as transfer matrices

There are two circles in S^3 around which we can place half-BPS surface defects without breaking the isometries, namely $\{\zeta_p = 0\}$ and $\{\zeta_q = 0\}$. Accordingly, dashed lines come in two types, related by an interchange of p and q. If a dashed line is in an n-dimensional representation, the L-operator in an $(N,-1)$ background

$L =$  (3.27)

may be represented as an $n \times n$ matrix, whose entries are difference operators acting on the fugacities for the flavor node associated with the $(N, 0)$ region below the dashed line. It satisfies the RLL relation with the R-matrix (3.6). This R-matrix defines the Bazhanov-Sergeev model of type SU(N) [22, 23].

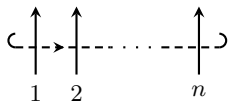
3.3.1 Surface defects and L-operators

Now we introduce a half-BPS surface defect, and put it on $S^1 \times S^1 \subset S^1 \times S^3$. The second S^1 factor may be taken to be either $\{\zeta_1 = 0\}$ or $\{\zeta_2 = 0\}$ in the parametrization $|\zeta_1|^2 + |\zeta_2|^2 = 1$ of S^3. These are the circles in S^3 that are left invariant under the action of the isometry group $\mathrm{U}(1)_p \times \mathrm{U}(1)_q$. As explained in section 2.3.3, the index in the presence of the surface defect is again given by a correlation function of line operators in a 2d TQFT on Σ. The correlator now contains a new line operator created by the D3-brane ending on the D5-branes, which was denoted by dashed line:

 . (3.28)

This line operator is specified by a representation of SU(N) [28, 48, 49]. In fact, it is labeled with a pair of representations (R_1, R_2) since in general we can take superposition of two surface defects, each wrapped around either circle in S^3.

In any case, the correlation function equals the partition function of a lattice model whose lattice is made of two kinds of lines, zig-zag paths coming from NS5-branes and the dashed line coming from the D3-branes. An extra dimension emerges as the M-theory circle if the brane system is embedded in the M-theory via T-duality along the first S^1 factor. Under this embedding, the D3-branes are mapped to M2-branes supported at points on the M-theory circle. Thus, the inclusion of the dashed line does not spoil the integrability of the lattice model. Moreover, by deforming the zig-zag paths near the dashed line, we can always make the neighborhood of the dashed line look like



(3.29)

in some (N, q) 5-brane background. Each crossing of a solid line and the dashed one gives us an R-operator, which we call L-operator. Thus we conclude that the surface defect is represented in the lattice model by the insertion of a transfer matrix constructed from L-operators.

3.3.2 Fundamental representation of SU(2)

Let us consider the simplest interesting setup where we have just two D5-branes and a single D3-brane, and identify the concrete form of the transfer matrix (3.29) in this case. For $N = 2$, (N, q) and $(N, q + 2)$ 5-branes are related by an $\mathrm{SL}(2; \mathbb{Z})$ transformation of type IIB string theory. Therefore, we can go to a duality frame in which the transfer matrix only involves either $(N, 0)$ or $(N, -1)$ regions or $(N, 0)$ or $(N, 1)$ regions. The two cases are on an

equal footing, and in fact related in a simple way. We first consider the transfer matrix in the $(N,-1)$ background.

We denote the L-operator in this case by $L^{\diamondsuit}$ since it is the operator that arises when a dashed line is inserted in a brane tiling model described by the diamond quiver constructed from the R-operator (3.6). In the situation under consideration, the gauge group of a brane tiling model is a product of SU(2) groups, and the surface defect is labeled by $(R_1,R_2)=(\emptyset,\square)$; the D3-brane wraps the circle $\{\zeta_2=0\}$ in S^3. Let $V^{\diamondsuit}$ be the space of meromorphic functions $f(z)$ such that $f(z)=f(1/z)$, and $W=\mathbb{C}^2$. Then we can represent $L^{\diamondsuit}: W\otimes V^{\diamondsuit}\to V^{\diamondsuit}\otimes W$ as 2×2 matrix whose entries are operators acting on functions in $V^{\diamondsuit}$. The R-matrix $\mathcal{R}^{\diamondsuit}_{ij}: W_i\otimes W_j\to W_j\otimes W_i$ is a 4×4 matrix. These operators, together with the R-operator $R^{\diamondsuit}$, satisfy the Yang-Baxter equations and RLL relations.

On the other hand, in the context of integrable lattice models Sklyanin constructed an L-operator $L^{\mathrm{S}}: W\otimes V^{\diamondsuit}\to V^{\diamondsuit}\otimes W$ that solves the RLL relation [50]

$$\mathcal{R}^{\mathrm{B}}_{12}(u_1,u_2)L^{\mathrm{S}}_1(u_1,(\nu,l))L^{\mathrm{S}}_2(u_2,(\nu,l))=L^{\mathrm{S}}_2(u_2,(\nu,l))L^{\mathrm{S}}_1(u_1,(\nu,l))\mathcal{R}^{\mathrm{B}}_{12}(u_1,u_2), \tag{3.30}$$

with Baxter's R-matrix $\mathcal{R}^{\mathrm{B}}_{ij}: W_i\otimes W_j\to W_j\otimes W_i$ for the eight-vertex model [51, 52]. Here u_i is a complex spectral parameter for W_i, and (ν,l) is a pair of complex spectral parameters for $V^{\diamondsuit}$. Baxter's R-matrix solves the Yang-Baxter equation

$$\mathcal{R}^{\mathrm{B}}_{12}(u_1,u_2)\mathcal{R}^{\mathrm{B}}_{13}(u_1,u_3)\mathcal{R}^{\mathrm{B}}_{23}(u_2,u_3)=\mathcal{R}^{\mathrm{B}}_{23}(u_2,u_3)\mathcal{R}^{\mathrm{B}}_{13}(u_1,u_3)\mathcal{R}^{\mathrm{B}}_{12}(u_1,u_2), \tag{3.31}$$

and the solution is given by

$$\mathcal{R}^{\mathrm{B}}_{ij}(u_i,u_j)=P\sum_{a=0}^{3}w_a(u_i-u_j)\sigma_a\otimes\sigma_a,\quad w_a(u)=\frac{\theta_{a+1}(u+\eta)}{\theta_{a+1}(\eta)}, \tag{3.32}$$

where $P: W_i\otimes W_j\to W_j\otimes W_i$ is the permutation operator, σ_a are the Pauli matrices and 2×2 unit matrix, $\theta_{a+1}(u)=\theta_{a+1}(u|\tau)$ are the Jacobi theta functions, and τ, η are complex parameters of the eight-vertex model. Sklyanin's L-operator is defined by

$$L^{\mathrm{S}}\left(u,(\nu,l)\right)=P\sum_{a=0}^{3}w_a(u+\eta)\sigma_a\otimes\mathbf{S}^{(l)}_a. \tag{3.33}$$

The operators $\mathbf{S}^{(l)}_a$ act on meromorphic functions $f(\zeta)$ as difference operators:

$$\left(\mathbf{S}^{(l)}_a f\right)(\zeta)=\mathrm{i}^{\delta_{a,0}}\frac{\theta_{a+1}(\eta)}{\theta_{a+1}(2\zeta)}\left(\theta_{a+1}(2\zeta-2\eta l)f(\zeta+\eta)-\theta_{a+1}(-2\zeta-2\eta l)f(\zeta-\eta)\right). \tag{3.34}$$

They generate the so-called Sklyanin algebra [53].

In [54], Derkachov and Spiridonov constructed an R-operator $R^{\mathrm{DS}}_{ij}: V^{\diamondsuit}_i\otimes V^{\diamondsuit}_j\to V^{\diamondsuit}_j\otimes V^{\diamondsuit}_i$ that satisfies the RLL relation

$$\begin{aligned}&R^{\mathrm{DS}}_{12}((\nu_1,l_1),\tilde{f}(\nu_2,l_2))L^{\mathrm{DS}}_2(u,(\nu_2,l_2))L^{\mathrm{DS}}_1(u,(\nu_1,l_1))\\&\qquad=L^{\mathrm{DS}}_1(u,(\nu_1,l_1))L^{\mathrm{DS}}_2(u,(\nu_2,l_2))R^{\mathrm{DS}}_{12}((\nu_1,l_1),(\nu_2,l_2)),\end{aligned} \tag{3.35}$$

The L-operator $L^{\mathrm{DS}} : W \otimes V^{\diamondsuit} \to V^{\diamondsuit} \otimes W$ is essentially Sklyanin's L-operator, differing only by an automorphism of the Sklyanin algebra:

$$L^{\mathrm{DS}}(u,(\nu,l)) = \varphi\sigma_3 L^{\mathrm{S}}\varphi^{-1}, \quad (\varphi f)(\zeta) := \exp(\pi \mathrm{i}\zeta^2/\eta) f(\zeta). \tag{3.36}$$

This L-operator also satisfies the RLL relation with Baxter's R-matrix $\mathcal{R}^{\mathrm{B}}$. At this point, what is important is that they show their R-operator R^{DS} is precisely the R-operator for the diamond quiver in the brane tiling model $R^{\diamondsuit}$, that is,

$$R^{\mathrm{DS}}_{ij}((\nu_i,l_i),(\nu_j,l_j)) = R^{\diamondsuit}_{ij}((a_i,b_i),(a_j,b_j)), \tag{3.37}$$

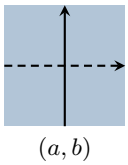
with the variables ζ and z are related by $z = \exp(2\pi\mathrm{i}\zeta)$ and the parameters matched as

$$a_i b_i = \exp(-2\pi\mathrm{i}\nu_i), \quad \frac{a_i}{b_i} = \exp(2\pi\mathrm{i}\eta(2l_i+1)), \tag{3.38}$$

$$(p,q) = (\exp(2\pi\mathrm{i}\tau), \exp(4\pi\mathrm{i}\eta)). \tag{3.39}$$

Based on these observations, we propose that the L-operator for the diamond quiver

$$L^{\diamondsuit}(c,(a,b)) = \quad c \tag{3.40}$$



is the L-operator of Derkachov and Spiridonov:

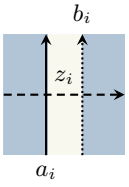
$$L^{\mathrm{DS}}(u,(\nu,l)) = L^{\diamondsuit}(c,(a,b)). \tag{3.41}$$

Requiring $L^{\diamondsuit}(c,(a,b)) = L^{\diamondsuit}(1,(a/c,b/c))$ fixes the relation between the two spectral parameters for the dashed line to be

$$c = \exp(\pi\mathrm{i}u). \tag{3.42}$$

For the computation of the transfer matrix, we exploit the fact that $L^{\diamondsuit}$ really consists of three parts separated by zig-zag paths:

$$L^{\diamondsuit}_i(c,(a_i,b_i)) = \quad c \tag{3.43}$$



Reflecting this structure, $L^{\diamondsuit}$ can be expressed in the following factorized form:

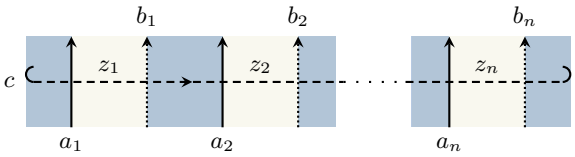
$$L^{\diamondsuit}_i(c,(a_i,b_i)) = B\left(z_i;\frac{b_i}{c}\right)\cdot\varphi(z_i)\frac{1}{\theta_1(z_i^2)}\begin{pmatrix}\Delta_i^{1/2} & 0\\ 0 & \Delta_i^{-1/2}\end{pmatrix}\varphi^{-1}(z_i)\cdot A\left(z_i;\frac{a_i}{c}\right). \tag{3.44}$$

In this expression, $\Delta_i^{\pm 1/2}$ are difference operators acting on functions of z_i as $\left(\Delta_i^{\pm 1/2} f\right)(z_i) = f(q^{\pm 1/2} z_i)$ and

$$A(z;a) = \begin{pmatrix}\bar\theta_4(a/z) & \bar\theta_3(a/z)\\ \bar\theta_4(az) & \bar\theta_3(az)\end{pmatrix}, \quad B(z;b) = \begin{pmatrix}\bar\theta_3(bz) & -\bar\theta_3(b/z)\\ \bar\theta_4(bz) & -\bar\theta_4(b/z)\end{pmatrix}, \tag{3.45}$$

where $\bar{\theta}_a(z) = \theta_a(z;\sqrt{p})$ and we used the multiplicative notation for the theta functions. Roughly speaking, one can think of the three matrices in the expression (3.44) as corresponding to the left, middle and right parts of the above diagram.

The transfer matrix (3.29) is obtained by concatenating n copies of the pieces (3.43) along a loop:



(3.46)

Thus, multiplying n copies of the L-operators (3.44) and using formulas in the appendix, we obtain the following formula for the transfer matrix:

$$\mathrm{Tr}_W \left(L_n^{\diamondsuit}(c,(a_n,b_n)) \circ_W \cdots \circ_W L_1^{\diamondsuit}(c,(a_1,b_1)) \right) = \sum_{s_1=\pm 1} \cdots \sum_{s_n=\pm 1} \prod_{i=1}^{n} \ell\left(z_{i-1}^{s_{i-1}}, z_i^{s_i}; \frac{b_{i-1}}{c}, \frac{a_i}{c} \right) \prod_{j=1}^{n} \Delta_j^{s_j/2}, \quad (3.47)$$

where

$$\ell(w,z;b,a) = \frac{1}{\theta(z^2)} \theta\left(\sqrt{\frac{p}{q}} ba \frac{w}{z} \right) \theta\left(\sqrt{\frac{p}{q}} \frac{a}{b} \frac{1}{wz} \right). \quad (3.48)$$

In this formula we have dropped off an overall constant independent of the spectral parameters. At any rate, the overall normalization of the L-operators is irrelevant and cannot be determined by the RLL relations. The RLL relations actually admit more degrees of freedom than just the overall normalization. For example, we can multiply $L^{\diamondsuit}(c,(a,b))$ by a function $f(c,(a,b))$ of its spectral parameters, and the new one still solves the RLL relations. In section 3.3.4 we will check the proposal by comparing it with independent computations from gauge theory.

So far we have considered the surface defect labeled $(R_1,R_2) = (\emptyset,\square)$. Of course, we may also consider the case with $(R_1,R_2) = (\square,\emptyset)$ in the same manner, by letting surface defects wrap around the other S^1 inside S^3. Hence, there are two sets of L-operators related by the symmetry exchanging p and q. The underlying algebraic structure is the product of two copies of the Sklyanin algebra, known as the *elliptic modular double* [55].

3.3.3 Surface defects in A_1 theories of class $\mathcal{S}$

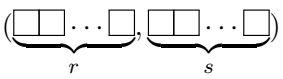
In [27], it was explained how to construct a surface defect labeled with a pair of integers (r,s), and how to determine its action on the supersymmetric index. Although the method applies to general $\mathcal{N}=2$ theories with SU(N) flavor symmetry, here we review it in the language of class-$\mathcal{S}$ theories.

Suppose we have a class-$\mathcal{S}$ theory $\mathcal{T}_{\mathrm{IR}}$ associated to a Riemann surface that contains a maximal puncture, whose flavor group we call $\mathrm{SU}(N)_z$. To this surface we introduce an extra minimal puncture. Concretely, we can do this as follows. First, we rename the flavor group $\mathrm{SU}(N)_z$ to $\mathrm{SU}(N)_{w'}$. Then, we take trinion representing a hypermultiplet $(Q,\tilde{Q})$ with

flavor symmetry $\mathrm{SU}(N)_{w''} \times \mathrm{SU}(N)_z \times \mathrm{U}(1)_\alpha$, and glue it to $\mathcal{T}_{\rm IR}$ by gauging the diagonal subgroup $\mathrm{SU}(N)_w$ of $\mathrm{SU}(N)_{w'} \times \mathrm{SU}(N)_{w''}$. The resulting theory $\mathcal{T}_{\rm UV}$ has one more flavor symmetry, $\mathrm{U}(1)_\alpha$, than $\mathcal{T}_{\rm IR}$. Correspondingly, the surface associated to $\mathcal{T}_{\rm UV}$ has one more minimal puncture than the original surface.

The theory $\mathcal{T}_{\rm UV}$ is related to $\mathcal{T}_{\rm IR}$ via the RG flow induced by a diagonal constant vev given to the quark Q, or equivalently, to the baryon $B = \det Q$. The vev higgses the gauge group $\mathrm{SU}(N)_w$ and breaks $\mathrm{SU}(N)_w \times \mathrm{SU}(N)_z$ down to the diagonal subgroup. Moreover, it turns the cubic superpotential $\tilde{Q}\Phi Q$ into a quadratic one that makes $\tilde{Q}$ and Φ massive, where Φ is the adjoint chiral multiplet introduced in the gluing. Up to Nambu-Goldstone multiplets that survive the higgsing, in the infrared the multiplets we added are gone and we recover $\mathcal{T}_{\rm IR}$, with $\mathrm{SU}(N)_w$ replaced with $\mathrm{SU}(N)_z$. In effect, the minimal puncture introduced by gluing the trinion is "closed." The R-charge I_3 is broken by the vev, but the combination $I_3 + F_\alpha/2$ is preserved and identified with a Cartan generator of the infrared SU(2) R-symmetry.

To create a surface defect in $\mathcal{T}_{\rm IR}$, we instead give the baryon a position-dependent vev $\langle B \rangle = \zeta_1^r \zeta_2^s$. Here, as vefore, ζ_1 and ζ_2 are complex coordinates of the two orthogonal planes rotated by $j_p = j_1 + j_2$ and $j_q = j_1 - j_2$, respectively. Away from the origin, the effect of the position-dependent vev is the same as that of the constant vev, so we get $\mathcal{T}_{\rm IR}$ in the infrared. If $r \neq 0$, however, the infrared theory is modified on the plane $\{\zeta_1 = 0\}$ since the vev vanishes there. By the same token, the theory is modified on the plane $\{\zeta_2 = 0\}$ if $s \neq 0$. Hence, in general we obtain $\mathcal{T}_{\rm IR}$ with the insertion of a surface defect labeled with the pair of integers (r, s), supported on the planes $\{\zeta_1 = 0\}$ and $\{\zeta_2 = 0\}$. This surface defect is to be identified with the surface defect labeled with the pair



of symmetric representation of $\mathrm{SU}(N)$ discussed in the previous section [28].

The index of $\mathcal{T}_{\rm UV}$ has a pole in the α-plane at $\alpha = \sqrt{t} p^{r/N} q^{s/N}$, and the residue there gives the index of $\mathcal{T}_{\rm IR}$ in the presence of the surface defect of type (r, s). The reason is the following. The position-dependent vev $\langle B \rangle = \zeta_1^r \zeta_2^s$ breaks $\mathrm{U}(1)_p$, $\mathrm{U}(1)_q$, and $\mathrm{SU}(2)_I$. At this value of α, however, the only combinations of charges that enter the trace defining the index are those that are preserved by the vev. Thus, we can still define the index in this background. As explained above, $\mathcal{T}_{\rm UV}$ flows to $\mathcal{T}_{\rm IR}$ plus Nambu-Goldstone multiplets in the infrared. The latter contains massless degrees of freedom, and they contribute to the index by a diverging factor, in fact a simple pole in the α-plane. Therefore, the residue at this pole gives the index of $\mathcal{T}_{\rm IR}$, together with some factor associated with the Nambu-Goldstone multiplets.

We wish to compute this residue and determine the action of the surface defect on the index in the simplest non-trivial case, namely when $N = 2$ and $(r, s) = (0, 1)$. But first, let us look at the trivial case $(r, s) = (0, 0)$ to gain intuition of the computation.

In the construction of a surface defect described above, $\tilde{Q}$ and Φ actually play no role. The essential point is that the vev given to the baryon built from Q replaces $\mathrm{SU}(N)_w$ with

$\mathrm{SU}(N)_z$ in the infrared. So we couple $\mathcal{T}_{\mathrm{IR}}$ just to Q for the moment. The index of the combined theory is given by

$$\int_{\mathbb{T}} \frac{dw}{2\pi \mathrm{i} w} \mathcal{I}_{\mathrm{V}}(w) \mathcal{I}_{\mathrm{B}}(z, w; \rho) \mathcal{I}_{\mathcal{T}_{\mathrm{IR}}}(w) = \kappa \int_{\mathbb{T}} \frac{dw}{2\pi \mathrm{i} w} \frac{\Gamma(\rho z^{\pm 1} w^{\pm 1})}{\Gamma(w^{\pm 2})} \mathcal{I}_{\mathcal{T}_{\mathrm{IR}}}(w), \tag{3.49}$$

where $\rho = \sqrt{t}/a$ is the fugacity of Q and $\kappa = (p;p)_\infty (q;q)_\infty/2$. In this integral, $|\rho| < 1$ is assumed, but we can analytically continue ρ to a complex parameter and study its pole structure. At $\rho = 1$, a constant vev be turned on for B without conflicting with the definition of the index. The integral should have a pole at this point in the p-plane, and we want to calculate the residue there.

The integrand has two pairs of poles in the w-plane at

$$w = \rho z,\ \rho^{-1} z; \qquad w = \rho z^{-1},\ \rho^{-1} z^{-1}. \tag{3.50}$$

As $\rho \to$, the first pair of poles collide and pinch the integration contour, and the integral diverges. Likewise, the second pair also collide in this limit. The pole of the integral in the ρ-plane arises from the contributions from these poles in w. Using formula (eq:appendix), we find that the contribution from the pole at $w = \rho z$ is

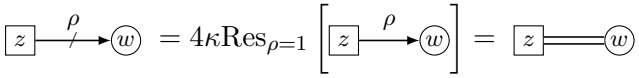
$$\frac{1}{2} \frac{\Gamma(\rho^2 z^2) \Gamma(z^{-2})}{\Gamma(\rho^2 z^2) \Gamma(\rho^{-2} z^{-2})} \Gamma(\rho^2) \mathcal{I}_{\mathcal{T}_{\mathrm{IR}}}(\rho z). \tag{3.51}$$

The last factor $\Gamma(\rho^2)$ indeed has a pole at $\rho = 1$, with residue $1/4\kappa$. The pole at $w = \rho z^{-1}$ makes an equal contribution, and we get

$$\mathrm{Res}_{\rho=1} \left[\int_{\mathbb{T}} \frac{dw}{2\pi \mathrm{i} w} \mathcal{I}_{\mathrm{V}}(w) \mathcal{I}_{\mathrm{B}}(z, w; \rho) \mathcal{I}_{\mathcal{T}_{\mathrm{IR}}}(w) \right] = \frac{1}{2} \frac{1}{2\kappa} \mathcal{I}_{\mathcal{T}_{\mathrm{IR}}}(z). \tag{3.52}$$

As expected, the residue reproduces the index of $\mathcal{T}_{\mathrm{IR}}$, multiplied by some factors. The factor of $1/2$ comes from the fact that B has fugacity ρ^2, and disappears if we add the equal contribution from the pole at $\rho = -1$. The factor $1/2\kappa$ is the contribution from a decoupled free chiral multiplet contained in a Nambu-Goldstone multiplet. It is the inverse of the index of a free vector multiplet since higgsing of a U(1) gauge theory with a single chiral multiplet leads to a trivial theory whose index is one.

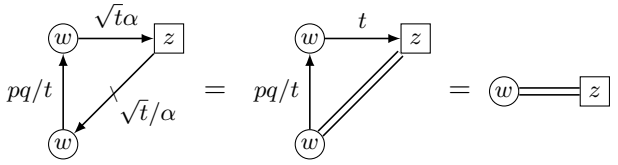
In order to express this result in a concise form, we introduce the notation of "striking out an arrow" in a quiver diagram to indicate that a constant vev is given to the baryonic operator built from the bifundamental chiral multiplet represented by that arrow, and the notation, what we just found is the identity



, (3.53)

where the right-hand side is the delta function defined by the relation (3.17). This identity holds when the index of any theory with SU(2) flavor symmetry (or more generally, any meromorphic function $f(w)$ such that $f(w) = f(1/w)$) is coupled to the right node.

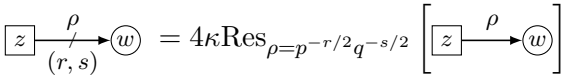
With the help of this identity, we can readily show that when a constant vev is turned on for B, the index of $\mathcal{T}_{\rm UV}$ reduces to that of $\mathcal{T}_{\rm IR}$. All we have to do is to look at the part of $\mathcal{T}_{\rm UV}$ describing the coupling to the trinion, and compute the relevant residue:



(3.54)

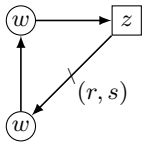
In the first equality we use the identity (3.53) and set $\rho = 1$, and in the second we canceled the pair of arrows making a loop. Thus, the vev transforms the trinion into the original flavor node of $\mathcal{T}_{\rm IR}$.

We can compute the index of $\mathcal{T}_{\rm IR}$ in the presence of a surface defect in a similar manner. To indicate that the position-dependent vev $\langle B \rangle = \zeta_1^r \zeta_2^s$ is turned on, we put the label (r, s) on the struck-out arrow:



(3.55)

Then the action of the surface defect of type (r, s) on the index is encoded in the diagram



(3.56)

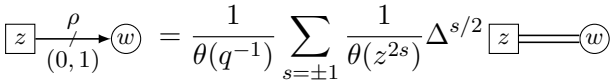
Let us calculate the residue (3.56) for $(r, s) = (0, 1)$. At $\rho = q^{-1/2}$, the index $\mathcal{I}_B(z, w; \rho) = \Gamma(\rho z^{\pm 1} w^{\pm 1})$ of Q has four sets of colliding poles in the w-plane. Two of them are

$$w = \rho q z,\ \rho^{-1} z; \qquad w = \rho z^{-1},\ \rho^{-1} q^{-1} z^{-1}, \tag{3.57}$$

while the other two are

$$w = \rho q z^{-1},\ \rho^{-1} z^{-1}; \qquad w = \rho z,\ \rho^{-1} q^{-1} z. \tag{3.58}$$

The contributions to the residue come from these poles. A small calculation shows that the first two sets of poles contribute in the same way: they set $w = q^{1/2} z$ and give a factor of $1/\theta(q^{-1})\theta(z^2)$ in total. Similarly, the contributions from the last two set $w = q^{-1/2} z$ and give a factor of $1/\theta(q^{-1})\theta(z^{-2})$. Altogether, we find that the result can be expressed as



(3.59)

We remind the reader that $\theta(z) = \theta(z; p)$ and $\Delta^{\pm 1/2}$ act on functions of z as $\left(\Delta^{\pm 1/2} f\right)(z) = f(q^{\pm 1/2} z)$.

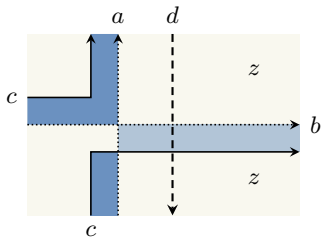
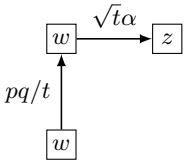


Figure 14: The brane tiling representation of a surface defect acting on a maximal puncture.

Unlike the case of the constant vev, this identity does not cause a complete cancelation of the indices of $\tilde{Q}$ and Φ. Rather, for $\rho = q^{-1/2}$ and $w = q^{\pm 1/2} z$, we have



$$= \theta\left(\frac{t}{q} z^{\mp 2}\right) \theta(t). \tag{3.60}$$

Therefore, the effect of introducing the surface defect of type $(0,1)$ on the index is realized by the difference operator

$$\mathfrak{S}_{(0,1)} = \frac{\theta(t)}{\theta(q^{-1})} \sum_{s=\pm 1} \frac{1}{\theta(z^{2s})} \theta\left(\frac{t}{q} z^{-2}\right) \Delta^{s/2}. \tag{3.61}$$

The prefactor $\theta(t)/\theta(q^{-1})$ is equal to the index of a free chiral field in two dimensions, and represents the center-of-mass degree of freedom of the surface defect.

The difference operator $\mathfrak{S}_{(0,1)}$ acts on the fugacity for the maximal puncture on which the surface defect was constructed. This fact has a natural interpretation. To construct the surface defect, we first introduced an extra minimal puncture, and then took the residue of a pole in the fugacity of the associated flavor symmetry. The latter step can be thought of as transforming the minimal puncture to another kind of puncture which represents the surface defect. By construction, this puncture is located in the neighborhood of a maximal puncture contained in a trinion. We can take the surface defect puncture and collide it to the maximal puncture. The collision produces a new puncture, and defines the action of the surface defect on the maximal puncture.

3.3.4 Comparison with the transfer matrix

Let us finally compare the result with the proposal. For clarity of presentation, take a minimal puncture in $\mathcal{T}_{\rm IR}$ and move it close to the maximal puncture on which the surface defect acts. Then the neighborhood of these punctures looks like a trinion glued to another maximal puncture, and is represented by zig-zag paths as in figure 14.

According to the proposal, the surface defect creates a dashed line with some spectral

parameter d, also drawn in the picture. It acts on the lattice model as the transfer matrix

$$\mathrm{Tr}\left(L^{\diamondsuit}(d,(c,b))\right) = \sum_{s=\pm 1} \frac{1}{\theta(z^{2s})}\theta\left(\sqrt{\frac{p}{q}}\frac{bc}{d^2}\right)\theta\left(\sqrt{\frac{p}{q}}\frac{c}{b}z^{-2s}\right)\Delta^{s/2}. \tag{3.62}$$

From the relation (3.25), we see that if we set

$$d = \frac{1}{\sqrt{qt}}, \tag{3.63}$$

the transfer matrix indeed reproduces the difference operator (3.61), up to an overall factor which cannot be fixed by the Yang-Baxter equations.

As noted in [27], the above transfer matrix is essentially the Hamiltonian of elliptic Ruijsenaars-Schneider model [56, 57] of type A_1. This fact follows from a general result obtained in [58].

Here we have considered only the surface defect of type $(0,1)$, but the general story is similar. The surface defect of type (r,s) acts on the index by a difference operator $\mathfrak{S}_{(r,s)}$. This operator is expected to coincide with the transfer matrix for an appropriate L-operator. If so, by the RLL relation, the operators $\mathfrak{S}_{(r,s)}$ for all (r,s) should commute with each other. This is indeed true [27]. From the class-$\mathcal{S}$ point of view, the mutual commutativity is guaranteed by the fact that the index is independent of the positions of punctures representing surface defects. Therefore, the order in which they act on a maximal puncture is irrelevant. Note that this argument also exploits the existence of an extra dimension, which is the M-theory circle that emerges as the type IIA brane configuration is lifted to M-theory.

For the same reason, a surface defect puncture can be placed between any two punctures, whether minimal or maximal, and still yield the same result. From the point of view of the type IIA system, this property appears to be quite non-trivial and is known as the "hopping invariance" of the index [28]. From the lattice model viewpoint, this is guaranteed by the other kind of RLL relation.

For the fundamental representation of SU(2), the above L-operator is essentially identified with Sklyanin's L-operator, which satisfies the RLL relation with Baxter's R-matrix for the eight-vertex model and generates the so-called Sklyanin algebra. For the fundamental representation of SU(N) with general N, we get the L-operator for Belavin's elliptic R-matrix [59]. If instead placed in an $(N,0)$ background, the L-operator gives a representation of Felder's elliptic quantum group for $\mathfrak{sl}_N$ [60–62]. For the details, see [63].

4 Line defects as transfer matrices

4.1 Integrable lattice models of elliptic and trigonometric type

In this section we discuss the integrable system side of the correspondence. After reviewing L-operators, transfer matrices and their relation to quantum integrable systems, we introduce an L-operator for the elliptic dynamical R-matrix. Then we define fundamental trigonometric

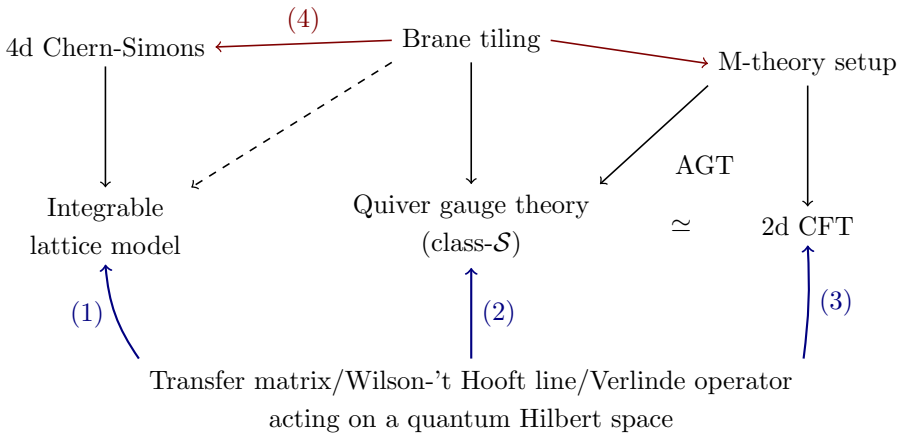


Figure 15: Structure of section 4. (1)

L-operators as certain limits of the elliptic L-operator. These fundamental L-operators are building blocks of transfer matrices that correspond to Wilson–'t Hooft lines in $\mathcal{N} = 2$ supersymmetric circular quiver theories.

4.1.1 L-operator and quantum integrable system

Let $\mathfrak{h}$ be a finite-dimensional commutative complex Lie algebra and V a finite-dimensional diagonalizable $\mathfrak{h}$-module. Choosing a basis $\{v_i\}$ of V that is homogeneous with respect to weight decomposition, we denote the weight of v_i by h_i and the (i,j)th entry of a matrix $M \in \mathrm{End}(V)$ by M^i_j. We write $\mathcal{M}_{\mathfrak{h}^*}$ for the field of meromorphic functions on the dual space $\mathfrak{h}^*$ of $\mathfrak{h}$.

Let $R\colon \mathbb{C} \times \mathfrak{h}^* \to \mathrm{End}(V \otimes V)$ be an $\mathrm{End}(V \otimes V)$-valued meromorphic function on $\mathbb{C} \times \mathfrak{h}^*$ that is invertible at a generic point $(z, a) \in \mathbb{C} \times \mathfrak{h}^*$. The coordinate z is called the *spectral parameter* and a is called the *dynamical parameter*.

In the discussions that follow, fundamental roles will be played by L-operators. By an *L-operator* for R, we mean a map $L\colon \mathbb{C} \to \mathrm{End}(V \otimes \mathcal{M}_{\mathfrak{h}^*} \otimes \mathcal{M}_{\mathfrak{h}^*})$, which we think of as a matrix whose entries are linear operators on meromorphic functions on $\mathfrak{h}^* \times \mathfrak{h}^*$. It must satisfy two conditions.

First, its matrix elements act on $f \in \mathcal{M}_{\mathfrak{h}^*} \otimes \mathcal{M}_{\mathfrak{h}^*}$ as

$$L(z)^j_i f(a^1, a^2) = L(z; a^1, a^2)^j_i \Delta^1_i \Delta^2_j f(a^1, a^2)\,, \tag{4.1}$$

where $L(z; a^1, a^2)^j_i$ is a meromorphic function on $\mathbb{C} \times \mathfrak{h}^* \times \mathfrak{h}^*$ and Δ^1_i, Δ^2_j are difference operators such that

$$\Delta^1_i f(a^1, a^2) = f(a^1 - \epsilon h_i, a^2)\,, \qquad \Delta^2_j f(a^1, a^2) = f(a^1, a^2 - \epsilon h_j)\,. \tag{4.2}$$

Here ϵ is a fixed complex parameter.

Second, the L-operator satisfies the *RLL relation*

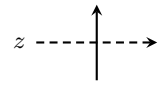
$$\sum_{k,l} R(z-z',a^2)^{mn}_{kl} L(z;a^1,a^2)^k_i L(z';a^1-\epsilon h_i,a^2-\epsilon h_k)^l_j$$
$$= \sum_{k,l} L(z';a^1,a^2)^n_l L(z;a^1-\epsilon h_l,a^2-\epsilon h_n)^m_k R(z-z',a^1)^{kl}_{ij} . \quad (4.3)$$

Equivalently, the operator relation

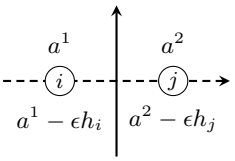
$$\sum_{k,l} R(z-z',a^2)^{mn}_{kl} L(z)^k_i L(z')^l_j = \sum_{k,l} R(z-z',a^1)^{kl}_{ij} L(z')^n_l L(z)^m_k \quad (4.4)$$

holds on any meromorphic function $f(a^1,a^2)$.

It is helpful, and will turn out to be physically meaningful, to represent the L-operator graphically as two crossing oriented line segments:

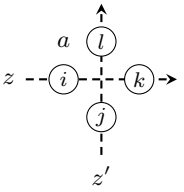
$L(z) =$  . (4.5)

The dashed line extending in the horizontal direction has a spectral parameter. The graphical representation of a matrix element of the L-operator is

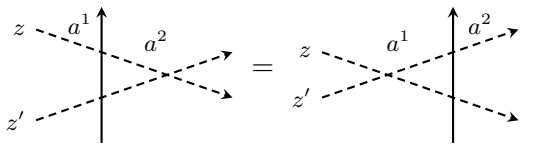
$L(z;a^1,a^2)^j_i =$  . (4.6)

Each edge of a dashed line carries a state in V, and the state may change when the line crosses another line. To each region separated by lines, a dynamical parameter is assigned. The values of dynamical parameters on the two sides of a dashed line carrying state v_i differ by ϵh_i.

We also represent the operator R as two crossing dashed lines:

$R(z-z',a)^{kl}_{ij} =$  . (4.7)

Then, the RLL relation (4.3) simply means an equality between two configurations involving two dashed and one solid lines:

 . (4.8)

The states carried by the internal dashed edges are summed over.

By comparing the values of the dynamical parameter assigned to the lower right regions of the two sides, we see that for R to satisfy the RLL relation, it must commute with

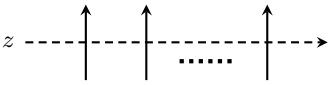
$h \otimes 1 + 1 \otimes h$ for all $h \in \mathfrak{h}$; in other words, $R(z,a)_{ij}^{kl} = 0$ unless $h_i + h_j = h_k + h_l$. This is a consistency condition for the rule that determines how dynamical parameters change across dashed lines.

Associated with an L-operator, there is an integrable quantum mechanical system consisting of particles moving in the space $\mathfrak{h}^*$. The Hilbert space of each particle is $\mathcal{M}_{\mathfrak{h}^*}$. (This is quantum mechanics in which real variables are analytically continued to complex ones.) The Hilbert space of the system is $\mathcal{M}_{\mathfrak{h}^*}^{\otimes n}$ if n is the number of particles.

To construct this system, define the *monodromy matrix* $M\colon \mathbb{C} \to \mathrm{End}(V \otimes \mathcal{M}_{\mathfrak{h}^*}^{\otimes n+1})$ by the product of n copies of the L-operator: its matrix elements are given by

$$M(z)_{i^1}^{i^{n+1}} = \sum_{i^2,\dots,i^n} \prod_{r=1}^{n} L(z; a^r, a^{r+1})_{i^r}^{i^{r+1}} \prod_{s=1}^{n+1} \Delta_{i^s}^s \,, \tag{4.9}$$

acting on any meromorphic function $f(a^1, \dots, a^{n+1})$. (The superscript on Δ_i specifies the variable on which the difference operator acts.) This is a dashed line crossing n double lines:

$M(z) =$  . (4.10)

Identifying $a^{n+1} = a^1$ and taking the trace, one obtains the *transfer matrix* $T\colon \mathbb{C} \to \mathrm{End}(\mathcal{M}_{\mathfrak{h}^*}^{\otimes n})$:

$$T(z) = \sum_{i^1,\dots,i^n} \prod_{r=1}^{n} L(z; a^r, a^{r+1})_{i^r}^{i^{r+1}} \prod_{s=1}^{n} \Delta_{i^s}^s \,, \qquad i^{n+1} = i^1 \,. \tag{4.11}$$

Graphically, $T(z)$ is represented by the same picture as above but with the horizontal direction made periodic.

By construction, T is an $\mathrm{End}(\mathcal{M}_{\mathfrak{h}^*}^{\otimes n})$-valued meromorphic function. As such, each coefficient T_m in the Laurent expansion $T(z) = \sum_{m\in\mathbb{Z}} T_m z^m$ is an operator acting on the Hilbert space $\mathcal{M}_{\mathfrak{h}^*}^{\otimes n}$. Then, one may pick a particular linear combination of these coefficients and declare that it is the Hamiltonian of the quantum mechanical system. The Hamiltonian thus obtained is a difference operator, which is typical of relativistic systems.

Alternatively, one may think of this system as a one-dimensional periodic quantum spin chain. This spin chain is constructed from n solid lines extending in the longitudinal direction of a cylinder, as shown in figure 16(a). The dynamical parameter a^r resides in the region sandwiched by the rth and the $(r+1)$th solid lines. One regards the n dynamical parameters $a^1, \dots, a^n$ as continuous spin variables; see figure 16(b). Thinking of the longitudinal direction as the time direction, the Hilbert space of the spin chain is again $\mathcal{M}_{\mathfrak{h}^*}^{\otimes n}$. An action of $T(z)$ on the Hilbert space is induced by an insertion of a dashed line with spectral parameter z in the circumferential direction of the cylinder, as in figure 16(c).

The integrability of the system is a consequence of the RLL relation. By repeated use of the RLL relation, one deduces that the monodromy matrix satisfies a similar relation:

$$\sum_{k,l} R(z-z', a^{n+1})_{kl}^{mn} M(z)_i^k M(z')_j^l = \sum_{k,l} R(z-z', a^1)_{ij}^{kl} M(z')_l^n M(z)_k^m \,. \tag{4.12}$$

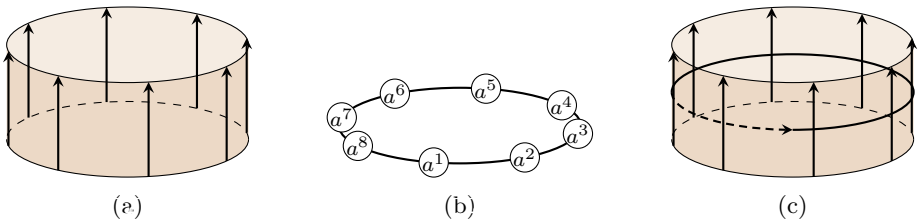


Figure 16: (a) Double lines in the longitudinal direction of a cylinder. (b) The corresponding quantum spin chain with continuous spin variables. (c) A dashed line winding around the cylinder acts on the spin chain by the transfer matrix.

Multiplying both sides by $R^{-1}(z-z',a^1)^{ij}_{mn}$, setting $a^{n+1}=a^1$ and summing over i, j, m, n, one finds

$$T(z)T(z')=T(z')T(z)\,. \tag{4.13}$$

In other words, transfer matrices at different values of the spectral parameter commute. It follows that the Laurent coefficients $\{T_m\}$ mutually commute and, in particular, commute with the Hamiltonian. Hence, the system has a series of commuting conserved charges.

There is a slight generalization of the above construction of commuting transfer matrices. Suppose that $g\in \mathrm{End}(V)$ satisfies

$$(g\otimes g)R(z,a)=R(z,a)(g\otimes g) \tag{4.14}$$

and a subspace W of V is invariant under R, R^{-1}, L and g. (For instance, the invariance of W under R means that $R(z,a)(W\otimes V)\subset W\otimes V$ and $R(z,a)(V\otimes W)\subset V\otimes W$ for all z, a.) Then, the trace can be twisted by g and restricted to W:

$$T_{g,W}=\mathrm{Tr}_W(gM)\,. \tag{4.15}$$

If W_1, W_2 are such invariant subspaces, then

$$[T_{g,W_1}(z),T_{g,W_2}(z')]=0\,. \tag{4.16}$$

Thus, we get different kinds of transfer matrices labeled by invariant subspaces, and they commute with each other. A typical situation in which this construction applies is when $\mathfrak{h}$ is a Cartan subalgebra of a complex Lie algebra $\mathfrak{g}_{\mathbb{C}}$, V is a direct sum of irreducible representations of $\mathfrak{g}_{\mathbb{C}}$, and g is an element of $\mathfrak{g}_{\mathbb{C}}$.

Algebraically, L-operators give representations of *dynamical quantum groups* [60,61,64]. As an algebra, the dynamical quantum group corresponding to R is generated by the meromorphic functions on $\mathbb{C}\times\mathfrak{h}^*\times\mathfrak{h}^*$, together with additional generators $l(z)^i_j$, $l^{-1}(z)^i_j$. The generators $l(z)^i_j$ are to be understood as the matrix elements of an abstract L-operator and satisfy the same relations as above; $l^{-1}(z)^i_j$ are the elements of the inverse matrix. This algebra has further structures (coproduct and counit) which make it an $\mathfrak{h}$-bialgebroid.

4.1.2 Elliptic L-operators

An important example of an L-operator is one for the elliptic dynamical R-matrix [65–67], which is a representation of the elliptic quantum group for $\mathfrak{sl}_N$. In this example, $\mathfrak{h}$ is the Cartan subalgebra of $\mathfrak{sl}_N$ and $V = \mathbb{C}^N$ is the vector representation of $\mathfrak{sl}_N$.

The Lie algebra $\mathfrak{sl}_N$ consists of the traceless complex $N \times N$ matrices and $\mathfrak{h}$ is the subalgebra of diagonal elements. We denote by $E_{ij} \in \mathfrak{gl}_N$ the matrix that has 1 in the (i,j)th entry and 0 elsewhere, and by E^*_{ij} the element of $\mathfrak{gl}^*_N = \mathrm{Hom}(\mathfrak{gl}_N, \mathbb{C})$ such that $\langle E_{ij}, E^*_{kl}\rangle = \delta_{ik}\delta_{jl}$. (The bilinear map $\langle -, -\rangle \colon \mathfrak{gl}_N \times \mathfrak{gl}^*_N \to \mathbb{C}$ is the natural pairing.) The elements of $\mathfrak{h}$ are matrices of the form $\sum_{i=1}^N b_i E_{ii}$, with $\sum_{i=1}^N b_i = 0$. Since $\mathfrak{h}$ is isomorphic to the quotient of the subspace of $\mathfrak{gl}_N$ consisting of the diagonal matrices by the subspace spanned by the identity matrix $I = \sum_{i=1}^N E_{ii}$, the dual space $\mathfrak{h}^*$ is isomorphic to the subspace of $\mathfrak{gl}^*_N$ consisting of elements of the form $\sum_{i=1}^N a_i E^*_{ii}$ such that $\langle I, \sum_{i=1}^N a_i E^*_{ii}\rangle = \sum_{i=1}^N a_i = 0$. Thus, $\mathfrak{h}^*$ may also be identified with the space of traceless diagonal matrices.

The natural action of $\mathfrak{sl}_N$ on $\mathbb{C}^N$ defines the vector representation of $\mathfrak{sl}_N$. In terms of the standard basis $\{e_1, \dots, e_N\}$ of $\mathbb{C}^N$, we have $\sum_{j=1}^N a_j E_{jj} e_i = a_i e_i$. The weight of e_i is therefore

$$h_i = E^*_{ii} - \frac{1}{N}\sum_{j=1}^N E^*_{jj}\,. \tag{4.17}$$

For $a \in \mathfrak{h}^*$, we write $a_i = \langle E_{ii}, a\rangle$. Then, $\sum_{i=1}^N a_i = 0$ and $a = \sum_{i=1}^N a_i E^*_{ii} = \sum_{i=1}^N a_i h_i$.

Fix a point τ in the upper half plane, $\mathrm{Im}\,\tau > 0$, and let

$$\theta_1(z) = -\sum_{j \in \mathbb{Z}+\frac{1}{2}} e^{\pi \mathrm{i} j^2 \tau + 2\pi \mathrm{i} j(z+\frac{1}{2})} \tag{4.18}$$

be Jacobi's first theta function. The *elliptic dynamical R-matrix* R^{ell} is defined by [60,61,64]

$$R^{\mathrm{ell}}(z,a) = \sum_{i=1}^N E_{ii} \otimes E_{ii} + \sum_{i \neq j} \alpha(z, a_{ij}) E_{ii} \otimes E_{jj} + \sum_{i \neq j} \beta(z, a_{ij}) E_{ji} \otimes E_{ij}\,, \tag{4.19}$$

where $a_{ij} = a_i - a_j$ and

$$\alpha(z,a) = \frac{\theta_1(a+\epsilon)\theta_1(-z)}{\theta_1(a)\theta_1(\epsilon - z)}\,, \qquad \beta(z,a) = \frac{\theta_1(a-z)\theta_1(\epsilon)}{\theta_1(a)\theta_1(\epsilon - z)}\,. \tag{4.20}$$

The *elliptic L-operator* L^{ell}, which satisfies the RLL relation with R^{ell}, has the matrix elements given by [58]

$$L^{\mathrm{ell}}_{w,y}(z; a^1, a^2)^j_i = \frac{\theta_1(z - w + a^2_j - a^1_i)}{\theta_1(z-w)} \prod_{k(\neq i)} \frac{\theta_1(a^1_k - a^2_j - y)}{\theta_1(a^1_{ki})}\,. \tag{4.21}$$

The complex numbers w, y may be thought of as spectral parameters for the corresponding solid line. The presence of the two parameters is a consequence of the fact that $R^{\mathrm{ell}}(z,a)$ is invariant under shift of a by a multiple of the identity matrix I and in the RLL relation (4.3)

the spectral parameters z, z' enter the R-matrix only through the difference $z - z'$; note also that the L-operator can be multiplied by any function of the spectral parameter.

The elliptic dynamical R-matrix and the elliptic L-operator have many more properties than just that they satisfy the RLL relation. Most importantly, the R-matrix is a solution of the dynamical Yang–Baxter equation and encodes the Boltzmann weights for a two-dimensional integrable lattice model. This model is equivalent to the eight-vertex model (or more precisely, the Belavin model [59] which is an $\mathfrak{sl}_N$ generalization of the eight-vertex model [51,52]) in the sense that the transfer matrices of the two models are related by a similarity transformation. The elliptic L-operator, on the other hand, satisfies the RLL relation with another R-matrix which describes an integrable lattice model called the Bazhanov–Sergeev model [22, 23], whose spins variables take values in $\mathfrak{h}^*$. We will not discuss these aspects in this paper. The interested reader is referred to [63] for more details.

4.1.3 Trigonometric L-operators

The L-operators that appear in the correspondence with Wilson–'t Hooft lines are obtained from the elliptic L-operator L^{ell} via the trigonometric limit $\tau \to \mathrm{i}\infty$. For comparison with gauge theory results, we actually need to express these L-operators in somewhat different forms.

First, we describe L-operators in a quantum mechanical language. Let us explain this description in the case in which $\mathfrak{h}$ is the Cartan subalgebra of $\mathfrak{sl}_N$. Recall that $\mathfrak{sl}_N$ has simple coroots

$$\alpha_i^\vee = E_{ii} - E_{i+1,i+1}\,, \qquad i = 1, \ldots, N-1\,, \tag{4.22}$$

and the fundamental weights

$$\omega_i = (\alpha_i^\vee)^* = \sum_{j=1}^{i} h_j\,. \tag{4.23}$$

Consider quantum mechanics of a particle living in $\mathfrak{h}^* \times \mathfrak{h}^*$, with Planck constant

$$\hbar = -\frac{\epsilon}{2\pi}\,. \tag{4.24}$$

If $(a^1, a^2) \in \mathfrak{h}^* \times \mathfrak{h}^*$ is the position of the particle, we write $a^r = \sum_{i=1}^{N-1} q_i^r \omega_i$, $r = 1, 2$. Similarly, we write the momenta $(b^1, b^2) \in \mathfrak{h} \times \mathfrak{h}$ of the particle as $b^r = \sum_{i=1}^{N-1} p_i^r \alpha_i^\vee$. The corresponding position and momentum operators $\hat{q}_i^r$, $\hat{p}_i^s$ satisfy the canonical commutation relations:

$$[\hat{q}_i^r, \hat{p}_j^s] = \mathrm{i}\hbar\delta^{rs}\delta_{ij}\,, \qquad i, j = 1, \ldots, N-1\,. \tag{4.25}$$

(As before, we are treating q_i^r, p_i^r as analytically continued variables.)

To rewrite the commutation relations in a form that is invariant under the action of the Weyl group, we make a change of basis

$$a^r = \sum_{i=1}^{N} a_i^r E_{ii}^*\,, \qquad b^r = \sum_{i=1}^{N} b_i^r E_{ii}\,. \tag{4.26}$$

Then, the corresponding observables $\hat{a}_i^r$, $\hat{b}_i^r$ obey the traceless condition, $\sum_{i=1}^N \hat{a}_i^r = \sum_{i=1}^N \hat{b}_i^r = 0$, and satisfy the commutation relations

$$[\hat{a}_i^r, \hat{b}_j^s] = \mathrm{i}\hbar\delta^{rs}\Big(\delta_{ij} - \frac{1}{N}\Big)\,, \qquad i,\, j = 1,\, \dots,\, N\,. \tag{4.27}$$

Using these observables we can identify the matrix elements of an L-operator L with an operator in the Hilbert space of this quantum mechanical system:

$$L(z)_i^j = L(z; \hat{a}^1, \hat{a}^2)_i^j e^{2\pi\mathrm{i}(\hat{b}_i^1 + \hat{b}_j^2)}\,. \tag{4.28}$$

In quantum mechanics, there is an invertible map from functions on the classical phase space to operators in the Hilbert space, known as the Weyl transform: if q and p are canonically conjugate variables, it maps

$$f(q,p) \mapsto \hat{f}(\hat{q},\hat{p}) = \int_{\mathbb{R}^4} \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}p\,\mathrm{d}q\, f(q,p) e^{\mathrm{i}(x(\hat{q}-q)+y(\hat{p}-p))}\,. \tag{4.29}$$

The inverse map is the *Wigner transform*, which we denote by $\langle - \rangle$:

$$f(\hat{q},\hat{p}) \mapsto \langle f(\hat{q},\hat{p})\rangle = \int_{\mathbb{R}} \mathrm{d}x\, e^{\mathrm{i}px/\hbar}\Big\langle q + \frac{1}{2}x\Big|\hat{f}(\hat{q},\hat{p})\Big|q - \frac{1}{2}x\Big\rangle\,. \tag{4.30}$$

In the situation at hand, if we rewrite the expression (4.28) as

$$L(z)_i^j = e^{\pi\mathrm{i}(\hat{b}_i^1 + \hat{b}_j^2)} \widetilde{L}(z; \hat{a}^1, \hat{a}^2)_i^j e^{\pi\mathrm{i}(\hat{b}_i^1 + \hat{b}_j^2)}\,, \tag{4.31}$$

then we have

$$\langle L(z)_i^j \rangle = e^{2\pi\mathrm{i}(b_i^1 + b_j^2)} \widetilde{L}(z; a^1, a^2)_i^j\,. \tag{4.32}$$

Next, we apply a similarity transformation to the elliptic L-operator. Assume $\mathrm{Im}\,\epsilon > 0$ and let

$$\Gamma(z,\tau,\epsilon) = \prod_{m,n=0}^{\infty} \frac{1 - e^{2\pi\mathrm{i}((m+1)\tau + (n+1)\epsilon - z)}}{1 - e^{2\pi\mathrm{i}(m\tau + n\epsilon + z)}} \tag{4.33}$$

be the elliptic gamma function. Then, $\overline{\Gamma}(z) = e^{\pi\mathrm{i}z^2/2\epsilon}\Gamma(z,\tau,\epsilon)$ has the property that $\overline{\Gamma}(z+\epsilon,\tau,\epsilon) = g(\tau,\epsilon)\theta_1(z)\overline{\Gamma}(z,\tau,\epsilon)$ for some function $g(\tau,\epsilon)$. We define the conjugated L-operator $\mathcal{L}^{\mathrm{ell}}_{w,m}(z)$ by

$$\mathcal{L}^{\mathrm{ell}}_{w,m}(z)_i^j = \Phi_{m-\frac{1}{2}\epsilon} L^{\mathrm{ell}}_{w,m-\frac{1}{2}\epsilon}(z)_i^j \Phi^{-1}_{m-\frac{1}{2}\epsilon}\,, \tag{4.34}$$

where

$$\Phi_y = \prod_{k,l=1}^{N} \overline{\Gamma}(\hat{a}_k^1 - \hat{a}_l^2 - y)^{\frac{1}{2}} \prod_{k\neq l} \overline{\Gamma}(\hat{a}_{kl}^1)^{-\frac{1}{2}}\,. \tag{4.35}$$

It has the Wigner transform

$$\begin{aligned}\langle \mathcal{L}^{\mathrm{ell}}_{w,m}(z)_i^j \rangle = e^{2\pi\mathrm{i}(b_i^1 + b_j^2)} &\frac{\theta_1(z - w + a_j^2 - a_i^1)}{\theta_1(z-w)} \\ &\times \left(\frac{\prod_{k(\neq i)} \theta_1(a_k^1 - a_j^2 - m) \prod_{l(\neq j)} \theta_1(a_i^1 - a_l^2 - m)}{\prod_{k(\neq i)} \theta_1(a_{ki}^1 - \frac{1}{2}\epsilon)\theta_1(a_{ik}^1 - \frac{1}{2}\epsilon)}\right)^{\frac{1}{2}}\,.\end{aligned} \tag{4.36}$$

With these preparations, let us finally take the trigonometric limit to define the trigonometric L-operator:

$$\mathcal{L}_{w,m} = \lim_{\tau\to \mathrm{i}\infty} \mathcal{L}^{\mathrm{ell}}_{w,m}\,. \tag{4.37}$$

The trigonometric L-operator satisfies the RLL relation with the trigonometric limit R^{trig}. of the elliptic R-matrix R^{ell}. Concretely, $\mathcal{L}_{w,m}$ and R^{trig} are obtained from $\mathcal{L}^{\mathrm{ell}}_{w,m}$ and R^{ell} by the replacement $\theta_1(z) \to \sin(\pi z)$.

Once we are in the trigonometric setup, the quasi-periodicity in $z \to z+\tau$ is lost and we can further take the limits $w \to \pm \mathrm{i}\infty$. This allows us to introduce more fundamental L-operators:

$$\mathcal{L}_{\pm,m} = \lim_{w\to \pm\mathrm{i}\infty} \mathcal{L}_{w,m}\,. \tag{4.38}$$

These L-operators do not depend on the spectral parameters z, w, and their matrix elements have the Wigner transforms

$$\langle (\mathcal{L}_{\pm,m})^j_i \rangle = e^{2\pi\mathrm{i}(b^1_i + b^2_j)} e^{\pm\pi\mathrm{i}(a^2_j - a^1_i)} \ell_m(a^1,a^2)^j_i\,, \tag{4.39}$$

with

$$\ell_m(a^1,a^2)^j_i = \left(\frac{\prod_{k(\neq i)} \sin\pi(a^1_k - a^2_j - m) \prod_{l(\neq j)} \sin\pi(a^1_i - a^2_l - m)}{\prod_{k(\neq i)} \sin\pi(a^1_{ki} - \frac{1}{2}\epsilon) \sin\pi(a^1_{ik} - \frac{1}{2}\epsilon)} \right)^{\frac{1}{2}}. \tag{4.40}$$

The L-operator for arbitrary parameters z, w can be realized as a linear combination of $\mathcal{L}_{\pm,m}$:

$$\mathcal{L}_{w,m}(z) = \frac{e^{\pi\mathrm{i}(z-w)}\mathcal{L}_{+,m} - e^{-\pi\mathrm{i}(z-w)}\mathcal{L}_{-,m}}{\sin\pi(z-w)}\,. \tag{4.41}$$

The monodromy matrix $\mathcal{M}_{\sigma,m}$ constructed from $\mathcal{L}_{\pm,m}$ is labeled by an n-tuple of signs $\sigma = (\sigma^1,\dots,\sigma^n) \in \{\pm\}^n$ and an n-tuple of complex numbers $m = (m^1,\dots,m^n)$:

$$\langle (\mathcal{M}_{\sigma,m})^{i^{n+1}}_{i^1} \rangle = \sum_{i^2,\dots,i^n} \prod_{s=1}^{n+1} e^{2\pi\mathrm{i} b^s_{i^s}} \prod_{r=1}^{n} e^{\sigma^r \pi\mathrm{i}(a^{r+1}_{i^{r+1}} - a^r_{i^r})} \ell_{m^r}(a^r, a^{r+1})^{i^{r+1}}_{i^r}\,. \tag{4.42}$$

The corresponding transfer matrix $\mathcal{T}_{\sigma,m}$ has the Wigner transform

$$\langle \mathcal{T}_{\sigma,m} \rangle = \sum_{i^1,\dots,i^n} \prod_{r=1}^{n} e^{2\pi\mathrm{i} b^r_{i^r}} e^{\sigma^r \pi\mathrm{i}(a^{r+1}_{i^{r+1}} - a^r_{i^r})} \ell_{m^r}(a^r, a^{r+1})^{i^{r+1}}_{i^r}\,, \tag{4.43}$$

with $a^{n+1} = a^1$, $i^{n+1} = i^1$. Our claim is that these quantities equal the vevs of Wilson–'t Hooft lines in $\mathcal{N}=2$ supersymmetric gauge theories.

4.2 Wilson-'t Hooft lines as transfer matrices

In the previous section we defined the fundamental trigonometric L-operators (4.38) and calculated transfer matrices constructed from them. As explained in section 1, these transfer matrices are expected to have interpretations as Wilson–'t Hooft lines in $\mathcal{N}=2$ supersymmetric gauge theories described by a circular quiver. In this section we verify this expectation by computing the vevs of the corresponding Wilson–'t Hooft lines.

4.2.1 Wilson–'t Hooft lines in $S^1 \times_\epsilon \mathbb{R}^2 \times \mathbb{R}$

Consider a four-dimensional gauge theory whose gauge group is a compact Lie group G with Lie algebra $\mathfrak{g}$. Choosing a maximal torus $T \subset G$ with Lie algebra $\mathfrak{t}$, we let $\Lambda_{\mathrm{r}}(\mathfrak{g}) \subset \mathfrak{t}^*$ and $\Lambda_{\mathrm{cr}}(\mathfrak{g}) \subset \mathfrak{t}$ be the root lattice and the coroot lattice of $\mathfrak{g}$, respectively. Their duals are the coweight lattice $\Lambda_{\mathrm{cw}}(\mathfrak{g}) = \Lambda_{\mathrm{r}}(\mathfrak{g})^\vee \subset \mathfrak{t}$ and the weight lattice $\Lambda_{\mathrm{w}}(\mathfrak{g}) = \Lambda_{\mathrm{cr}}(\mathfrak{g})^\vee \subset \mathfrak{t}^*$.

An 't Hooft line is the worldline of a very heavy monopole, that is, a nondynamical magnetically charged particle. In the presence of an 't Hooft line, the gauge field of the theory has a singularity at the location of the monopole: in terms of the polar angle θ and the azimuthal angle ϕ of the spherical coordinates centered at the monopole, the gauge field behaves as

$$A = \frac{\mathbf{m}}{2}(1 - \cos\theta)\mathrm{d}\phi + \cdots , \tag{4.44}$$

where $\cdots$ represents less singular terms. (For simplicity we are setting the gauge theory theta angles to zero.) The coefficient $\mathbf{m}$ is the magnetic charge of the monopole. Different singular gauge field configurations of the above form describe the same monopole if their magnetic charges are related by gauge transformation. It follows that $\mathbf{m}$ can be chosen from $\mathfrak{t}$, and the choice is meaningful only up to the action of the Weyl group $W(G)$ of G.

The above expression of A is valid in a trivialization over a coordinate patch that contains the point $\theta = 0$ of a two-sphere surrounding the monopole. At $\theta = \pi$, there is a "Dirac string" which supports an unphysical magnetic flux. For the Dirac string to be invisible, we must have

$$\langle \mathbf{m}, w \rangle \in \mathbb{Z} \tag{4.45}$$

for every weight $w \in \mathfrak{t}^*$ of the representation of every field in the theory. This is simply the condition that the holonomy of A around the point $\theta = \pi$ is trivial in the bundles of which the fields are sections. The theory always contains fields in the adjoint representation, so $\mathbf{m}$ belongs to the coweight lattice:[14]

$$\mathbf{m} \in \Lambda_{\mathrm{cw}}(\mathfrak{g})/W(G)\,. \tag{4.46}$$

Equivalently, $\mathbf{m}$ is specified by an irreducible representation of the Langlands dual ${}^L\mathfrak{g}$ of $\mathfrak{g}$. In general, $\mathbf{m}$ lies in a sublattice of $\Lambda_{\mathrm{cw}}(\mathfrak{g})/W(G)$ determined by the matter content.

We can also consider heavy particles that carry both magnetic and electric charges. The worldline of such a dyon is called a Wilson–'t Hooft line. In the path integral formalism, a Wilson–'t Hooft line is realized by an insertion of a Wilson line

$$\mathrm{Tr}_R\, P \exp\left(\mathrm{i}\int_L A\right) \tag{4.47}$$

and a singular boundary condition on the support L of the line as specified by the magnetic charge. The prescribed singularity (4.44) breaks the gauge symmetry to the stabilizer $G_{\mathbf{m}}$

[14] Further, $\mathbf{m}$ belongs to the cocharacter lattice $\{v \in \mathfrak{t} \mid \exp(2\pi\mathrm{i}v) = \mathrm{id}_G\}$, which is a sublattice of $\Lambda_{\mathrm{cw}}(\mathfrak{g})$. If we take G to be the adjoint group, the cocharacter lattice coincides with $\Lambda_{\mathrm{cw}}(\mathfrak{g})$.

of $\mathbf{m}$, so R is an irreducible representation of $G_{\mathbf{m}}$. (More precisely, R is an irreducible representation of the stabilizer of $\mathbf{m}$ in the universal cover $\widetilde{G}$ of G [68].)

The data specifying such a pair $(\mathbf{m}, R)$ is actually the same as a pair $(\mathbf{m}, \mathbf{e})$ of coweight $\mathbf{m}$ and weight $\mathbf{e}$ modulo the Weyl group action:

$$(\mathbf{m}, \mathbf{e}) \in \big(\Lambda_{\mathrm{cw}}(\mathfrak{g}) \times \Lambda_{\mathrm{w}}(\mathfrak{g})\big)/W(G)\,. \tag{4.48}$$

As emphasized in [68], this data has more information than a pair of irreducible representations of $\mathfrak{g}$ and ${}^L\mathfrak{g}$.

In [69], the vevs of Wilson–'t Hooft lines in $\mathcal{N} = 2$ supersymmetric gauge theories on $S^1 \times_\epsilon \mathbb{R}^2 \times \mathbb{R}$ in the Coulomb phase were computed via localization of the path integral. The geometry $S^1 \times_\epsilon \mathbb{R}^2$ is a twisted product of S^1 and $\mathbb{R}^2$, constructed from $[0, 2\pi r] \times \mathbb{R}^2$ by the identification $(2\pi r, z) \sim (0, e^{2\pi \mathrm{i}\epsilon} z)$, where z is the complex coordinate of $\mathbb{R}^2 \cong \mathbb{C}$. These Wilson–'t Hooft lines wind around S^1, and are located at the origin of $\mathbb{R}^2$ and a point in $\mathbb{R}$. In order to preserve half of the eight supercharges, they require the complex scalar field ϕ in the vector multiplet to also have a singular behavior and replace the gauge field in the Wilson line (4.47) with $A + \mathrm{i}\,\mathrm{Re}\,\phi$. The vevs depend holomorphically on parameters

$$a \in \mathfrak{t}_{\mathbb{C}}\,, \qquad b \in \mathfrak{t}^*_{\mathbb{C}}\,, \tag{4.49}$$

which are set by the values of the gauge field and the vector multiplet scalar at spatial infinity. Essentially, a is given by the holonomy around S^1 at infinity of the gauge field, while b is that of the dual gauge field. We refer the reader to [69] for the precise definitions.

The vev of a Wilson line W_R in representation R is simply given by the classical value of the holonomy:

$$\langle W_R \rangle = \mathrm{Tr}_R\, e^{2\pi \mathrm{i} a}\,. \tag{4.50}$$

The vevs of 't Hooft lines are much more involved. For an 't Hooft line $T_{\mathbf{m}}$ with magnetic charge $\mathbf{m}$, the vev takes the form

$$\langle T_{\mathbf{m}} \rangle = \sum_{\substack{v \in \Lambda_{\mathrm{cr}}(\mathfrak{g}) + \mathbf{m} \\ \|v\| \le \|\mathbf{m}\|}} e^{2\pi \mathrm{i}\langle v, b\rangle} Z_{\text{1-loop}}(a, m, \epsilon; v) Z_{\mathrm{mono}}(a, m, \epsilon; \mathbf{m}, v)\,, \tag{4.51}$$

where m collectively denotes complex mass parameters. The summation over the coweights v in the shifted coroot lattice $\Lambda_{\mathrm{cr}} + \mathbf{m}$ accounts for the so-called "monopole bubbling," a phenomenon in which smooth monopoles are absorbed by the 't Hooft line and screen the magnetic charge. The norm $\|v\|$ with respect to a Killing form is bounded by $\|\mathbf{m}\|$, so this is a finite sum. The first two factors in the summand are the classical action and the one-loop determinant in the screened monopole background, respectively. The last factor is the nonperturbative contributions coming from degrees of freedom trapped on the 't Hooft line due to monopole bubbling.

Suppose that the theory under consideration consists of a vector multiplet and N_F hypermultiplets in representations R_f with mass parameters m_f, $f = 1, \ldots, N_F$. The one-loop

determinant $Z_{\text{1-loop}}$ is then the product of the contributions from the vector multiplet and the hypermultiplets:

$$Z_{\text{1-loop}}(a,m,\epsilon;v) = Z^{\text{vm}}_{\text{1-loop}}(a,\epsilon;v)\prod_{f=1}^{N_F} Z^{\text{hm},R_f}_{\text{1-loop}}(a,m_f,\epsilon;v)\,. \tag{4.52}$$

The two functions are given by

$$Z^{\text{vm}}_{\text{1-loop}}(a,\epsilon;v) = \prod_{\alpha\in\Phi(\mathfrak{g})}\prod_{k=0}^{|\langle v,\alpha\rangle|-1}\sin^{-\frac{1}{2}}\Big(\pi\langle a,\alpha\rangle + \pi\Big(\frac{1}{2}|\langle v,\alpha\rangle| - k\Big)\epsilon\Big)\,, \tag{4.53}$$

$$Z^{\text{hm},R}_{\text{1-loop}}(a,m,\epsilon;v) = \prod_{w\in P(R)}\prod_{k=0}^{|\langle v,w\rangle|-1}\sin^{\frac{1}{2}}\Big(\pi\langle a,w\rangle - \pi m + \pi\Big(\frac{1}{2}|\langle v,w\rangle| - \frac{1}{2} - k\Big)\epsilon\Big)\,. \tag{4.54}$$

Here, $\Phi(\mathfrak{g})$ is the set of roots of $\mathfrak{g}$ and $P(R)$ is the set of weights of R.

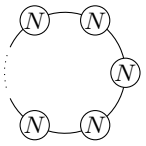
The factor Z_{mono} is subtle. The original computation in [69] did not give an answer that completely matches predictions from the AGT correspondence. The subtleties have been addressed in subsequent works [70–73] but not resolved in full generality.

Fortunately, for Wilson–'t Hooft lines that are of interest to us, the screened magnetic charges are in the same $W(G)$-orbit as $\mathbf{m}$. The corresponding contributions are therefore obtained by the $W(G)$-action from the perturbative term, for which $v = \mathbf{m}$ and $Z_{\text{mono}} = 1$.

To our knowledge, a formula for the vevs of dyonic Wilson–'t Hooft lines generalizing the expressions (4.50) and (4.51) has not been derived. Nevertheless, for the same reason as mentioned, we can calculate the vev of a relevant Wilson–'t Hooft line by first writing down its perturbative contribution, which is simply the product of the perturbative vevs of the corresponding purely electric and purely magnetic lines, and then summing over the contributions from the nonperturbative sectors related by the $W(G)$-action.

4.2.2 Transfer matrices from circular quiver theories

The Wilson–'t Hooft line that corresponds to the transfer matrix (4.43) is one in an $\mathcal{N}=2$ supersymmetric gauge theory that is described by a circular quiver with n nodes:



. (4.55)

Each node represents a vector multiplet for an $\mathrm{SU}(N)$ gauge group,[15] and each edge a hypermultiplet that transforms in the bifundamental representation under the gauge groups of the nodes it connects.

Let us first consider the case in which the quiver consists of a single node and a single edge. In this case, the gauge group $G = \mathrm{SU}(N)$ and the only hypermultiplet is in the adjoint representation. This theory is known as $\mathcal{N} = 2^*$ theory.

[15] More precisely, the gauge group is a product of $\mathrm{PSU}(N)$.

The roots of $\mathfrak{g} = \mathfrak{su}_N$ are $\alpha_{ij} = E^*_{ii} - E^*_{jj} = h_i - h_j$, $i \neq j$. The positive roots are α_{ij}, $i < j$, and the simple roots are $\alpha_i = \alpha_{i,i+1}$, $i = 1, , \ldots, N-1$. The fundamental coweights are $\omega_i^\vee = (\alpha_i^\vee)^* = \sum_{j=1}^i h_j^\vee$, with

$$h_i^\vee = E_{ii} - \frac{1}{N}\sum_{j=1}^N E_{jj}\,. \tag{4.56}$$

The various lattices are

$$\Lambda_{\rm r} = \bigoplus_{i=1}^{N-1} \mathbb{Z}\alpha_i\,, \qquad \Lambda_{\rm cr} = \bigoplus_{i=1}^{N-1} \mathbb{Z}\alpha_i^\vee\,, \qquad \Lambda_{\rm w} = \bigoplus_{i=1}^{N-1} \mathbb{Z}\omega_i\,, \qquad \Lambda_{\rm cw} = \bigoplus_{i=1}^{N-1} \mathbb{Z}\omega_i^\vee\,. \tag{4.57}$$

We recall that $\alpha_i^\vee$ are the simple coroots and $\omega_i = (\alpha_i^\vee)^*$ are the fundamental weights.

For $\mathcal{N} = 2^*$ theory with $G = \mathrm{SU}(N)$, minimal magnetic charges are $\mathbf{m} = \omega_1^\vee = h_1^\vee$ and $\mathbf{m} = \omega_{N-1}^\vee = -h_N^\vee$. These magnetic charges are the highest weights of the fundamental representation and the antifundamental representation of ${}^L\mathfrak{su}_N \cong \mathfrak{su}_N$, respectively.

Let us consider the 't Hooft line with $\mathbf{m} = h_1^\vee$. The vev of this 't Hooft line is expressed as a sum over the screened magnetic charges $v = h_1^\vee, h_2^\vee, \ldots, h_N^\vee$. The term for $v = h_1^\vee$ is the perturbative contribution and given by

$$e^{2\pi i b_1} \prod_{j=2}^N \sin^{-\frac{1}{2}}\Big(\pi a_{1j} + \frac{1}{2}\pi\epsilon\Big) \sin^{-\frac{1}{2}}\Big(\pi a_{j1} + \frac{1}{2}\pi\epsilon\Big) \sin^{\frac{1}{2}}(\pi a_{1j} - \pi m) \sin^{\frac{1}{2}}(\pi a_{j1} - \pi m)\,, \tag{4.58}$$

where $a_i = \langle a, h_i\rangle$, $b_i = \langle h_i^\vee, b\rangle$, $a_{ij} = a_i - a_j$ and m is the mass of the adjoint hypermultiplet. The other terms are related to this perturbative term by the Weyl group action which permutes $(h_1^\vee, \ldots, h_N^\vee)$, so we find

$$\langle T_{h_1^\vee}\rangle = \sum_{i=1}^N e^{2\pi i b_i} \prod_{j(\neq i)} \left(\frac{\sin\pi(a_{ij} - m)\sin\pi(a_{ji} - m)}{\sin\pi(a_{ij} - \frac{1}{2}\epsilon)\sin\pi(a_{ji} - \frac{1}{2}\epsilon)}\right)^{\frac{1}{2}}. \tag{4.59}$$

The vev of $T_{-h_N^\vee}$ is obtained from $\langle T_{h_1^\vee}\rangle$ by the replacement $b_i \to -b_i$.

Now, let us turn to a circular quiver with n nodes. For this theory, we have $G = \mathrm{SU}(N)^n$ and $\Lambda_{\rm cw}(\mathfrak{g}) = \Lambda_{\rm cw}(\mathfrak{su}_N)^{\oplus n}$. We consider the 't Hooft line with

$$\mathbf{m} = h_1^\vee \oplus \cdots \oplus h_1^\vee\,, \tag{4.60}$$

charged equally under the $\mathrm{SU}(N)$ factors of G. This time, the summation is over all coweights of the form $v = h_{i^1}^\vee \oplus \cdots \oplus h_{i^n}^\vee$. The perturbative term, for which $i^1 = \cdots = i^n = 1$, is given by

$$\prod_{r=1}^n e^{2\pi i b_1^r} \prod_{j=2}^N \sin^{-\frac{1}{2}}\Big(\pi a_{1j}^r + \frac{1}{2}\pi\epsilon\Big) \sin^{-\frac{1}{2}}\Big(\pi a_{j1}^r + \frac{1}{2}\pi\epsilon\Big) \times \sin^{\frac{1}{2}}\big(\pi(a_j^r - a_1^{r+1}) - \pi m^r\big) \sin^{\frac{1}{2}}\big(\pi(a_1^r - a_j^{r+1}) - \pi m^r\big)\,. \tag{4.61}$$

The superscript r refers to the rth SU(N) factor of G, with $a^{n+1} = a^1$. Collecting the contributions from the other coweights, we get

$$\langle T_{h_1^\vee \oplus \cdots \oplus h_1^\vee} \rangle = \sum_{i^1,\ldots,i^n} \prod_{r=1}^{n} e^{2\pi \mathrm{i} b_{i^r}^r} \ell_{m^r}(a^r, a^{r+1})_{i^r}^{i^{r+1}}, \tag{4.62}$$

where we have used the functions (4.40).

Comparing this expression with the Wigner transform (4.43) of the trigonometric transfer matrix $\mathcal{T}_{\sigma,m}$, we see

$$\langle T_{h_1^\vee \oplus \cdots \oplus h_1^\vee} \rangle = \langle \mathcal{T}_{(+,\ldots,+),m} \rangle = \langle \mathcal{T}_{(-,\ldots,-),m} \rangle \tag{4.63}$$

under the obvious identification of parameters.

In order to reproduce $\langle \mathcal{T}_{\sigma,m} \rangle$ for a general choice of the signs σ, we add to the 't Hooft line the electric charge

$$\mathbf{e} = \sum_{r=1}^{n} \sigma^r \frac{1}{2}(h_1^{r+1} - h_1^r) = \sum_{r=1}^{n} (\sigma^r 1 - \sigma^{r+1} 1) \frac{1}{2} h_1^{r+1}. \tag{4.64}$$

This electric charge is in a sense a minimal one that is compatible with the Dirac–Schwinger–Zwanziger quantization condition for locality: the charges $(\mathbf{m}, \mathbf{e})$ and $(\mathbf{m}', \mathbf{e}')$ of two dyons must satisfy $\langle \mathbf{m}, \mathbf{e}' \rangle - \langle \mathbf{m}', \mathbf{e} \rangle \in \mathbb{Z}$. In section 4.3, we will see the geometric meaning of this "minimality" in connection with the AGT correspondence.

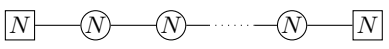
The magnetic charge (4.60) breaks the gauge group to $\mathrm{S}(\mathrm{U}(1) \times \mathrm{U}(N-1))^n$, and we are turning on a Wilson line that is charged under the U(1) factors with charges proportional to $(\sigma^r 1 - \sigma^{r+1} 1)/2$. The Wilson line multiplies the perturbative term (4.61) by the phase factor

$$\prod_{r=1}^{n} e^{\sigma^r \pi \mathrm{i} \langle a, h_1^{r+1} - h_1^r \rangle} = \prod_{r=1}^{n} e^{\sigma^r \pi \mathrm{i} (a_1^{r+1} - a_1^r)}. \tag{4.65}$$

Hence, the term with $v = h_{i^1}^\vee \oplus \cdots \oplus h_{i^n}^\vee$ gets the phase factor $e^{\sigma^r \pi \mathrm{i} (a_{i^{r+1}}^{r+1} - a_{i^r}^r)}$, and the vev of this Wilson–'t Hooft line matches the Wigner transform of $\mathcal{T}_{\sigma,m}$.

4.2.3 Monodromy matrices from linear quiver theories

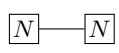
We have considered the Wilson–'t Hooft lines in the circular quiver theory and showed that their vevs match the Wigner transforms of the trigonometric transfer matrices. What correspond to the monodromy matrices then? In view of the fact that summing over the weights of the representation $V = \mathbb{C}^N$ in the integrable model amounts to summing over the different screened magnetic charges, natural candidates are Wilson–'t Hooft lines in a theory described by a linear quiver with $n+1$ nodes:



(4.66)

The leftmost and the rightmost nodes represent SU(N) flavor groups, which are not gauged.

In particular, we expect that the fundamental trigonometric L-operators (4.38) arise from the vevs of Wilson–'t Hooft lines of the theory of a bifundamental hypermultiplet:



(4.67)

Let us see if this is the case.

We introduce nondynamical vector multiplets for the SU(N) flavor groups, and consider the Wilson–'t Hooft lines with magnetic charge

$$\mathbf{m} = h_i^\vee \oplus h_j^\vee \tag{4.68}$$

and electric charges

$$\mathbf{e} = \mp\frac{1}{2}h_i \oplus \pm\frac{1}{2}h_j\,. \tag{4.69}$$

Note that the electric charges are fractional. The vevs of these Wilson–'t Hooft lines are

$$e^{2\pi\mathrm{i}(b_i^1+b_j^2)}e^{\pm\pi\mathrm{i}(a_j^2-a_i^1)}\prod_{k(\neq i)}\prod_{l(\neq j)}\left(\sin\pi(a_k^1-a_j^2-m)\sin\pi(a_i^1-a_l^2-m)\right)^{\frac{1}{2}}. \tag{4.70}$$

The vevs do not quite match the Wigner transforms (4.39) of $(\mathcal{L}_{\pm,m})_i^j$. They differ by the factor in the denominator of the function (4.40).

This factor is the one-loop determinant associated with the first node; it would have been present had the SU(N) flavor group been gauged and the vector multiplet been dynamical. From the gauge theory point of view, it is natural to think of this factor as a weight accompanying the summation over the screened magnetic charges. On the integrable system side, we could as well omit the denominator in question from the definitions of the L-operators and adopt the convention that the same weight is included when operators are multiplied within V. The L-operators would still satisfy the RLL relation.

To get the monodromy matrix (4.42), we take n L-operators and multiply them inside V. The gauge theory counterpart of this operation is to connect n copies of the two-node quiver (4.67), in the presence of appropriate Wilson–'t Hooft lines of the type considered above, by identifying and gauging flavor nodes. This produces the $n+1$ node linear quiver (4.66) and the Wilson–'t Hooft lines with magnetic charge

$$\mathbf{m} = h_{i^1}^\vee \oplus h_1^\vee \oplus h_1^\vee \oplus \cdots \oplus h_1^\vee \oplus h_{i^{n+1}}^\vee \tag{4.71}$$

and electric charge

$$\mathbf{e} = \sigma^1\frac{1}{2}(h_1^2 - h_{i^1}^1) + \sum_{r=2}^{n-1}\sigma^r\frac{1}{2}(h_1^{r+1}-h_1^r) + \sigma^n\frac{1}{2}(h_{i^{n+1}}^{n+1} - h_1^n)\,. \tag{4.72}$$

The vev of this Wilson–'t Hooft line reproduces the Wigner transform (4.42), except that a factor corresponding to the one-loop determinant for the vector multiplet for the first node is missing.

4.2.4 Other representations

The magnetic charge (4.60) of the above Wilson–'t Hooft lines is the highest weight of the representation $(\mathbb{C}^N)^{\oplus n}$ of the Langlands dual ${}^L\mathfrak{g}_{\mathbb{C}} \cong \mathfrak{sl}_N^{\oplus n}$ of $\mathfrak{g}_{\mathbb{C}}$. The corresponding transfer matrix (4.43) is represented graphically as n solid lines intersected by a single dashed loop carrying the representation $V = \mathbb{C}^N$, as shown in figure 16(c). The n regions sandwiched between solid lines correspond to the n copies of $\mathfrak{sl}_N$.

Both sides of the correspondence have a generalization in which the vector representation $\mathbb{C}^N$ is replaced by another representation R of $\mathfrak{sl}_N$. On the gauge theory side, we can change the magnetic charge of the Wilson–'t Hooft lines to the highest weight λ_R of $R^{\oplus n}$ while keeping the electric charges intact. On the integrable system side, the counterpart of this operation is the fusion procedure, which allows one to construct a dashed line in an arbitrary finite-dimensional representation of $\mathfrak{sl}_N$ from a collection of dashed lines in the vector representation, with the spectral parameters suitably adjusted.

We naturally expect that the vev of the Wilson–'t Hooft line with magnetic charge $\mathbf{m} = \lambda_R^{\oplus n}$ is equal to the Wigner transform of a transfer matrix constructed from L-operators in representation R, obtained by fusion from the L-operators (4.38) in the vector representation.

For $n = 1$ and $R = \wedge^k\mathbb{C}^N$, this equality can be verified from known results. In this case, the transfer matrix is the trigonometric limit of Ruijsenaars' difference operator [57]

$$\sum_{\substack{I\subset\{1,\dots,N\}\\|I|=k}} \Delta_I^{\frac12} \prod_{\substack{i\in I\\ j\notin I}} \sqrt{\frac{\theta_1(a_{ji}-m)\theta_1(a_{ij}-m)}{\theta_1(a_{ji}-\frac12\epsilon)\theta_1(a_{ij}-\frac12\epsilon)}}\,\Delta_I^{\frac12}\,, \qquad \Delta_I = \prod_{i\in I}\Delta_i\,, \tag{4.73}$$

and is related to the Macdonald operator by a similarity transformation [58]. On the other hand, the exterior power $\wedge^k\mathbb{C}^N$ being a minuscule representation (that is, all weights are related by the action of the Weyl group), the vev of the 't Hooft line with $\mathbf{m} = \omega_k^\vee = h_1^\vee + \cdots + h_k^\vee$ in $\mathcal{N} = 2^*$ theory can be computed from the perturbative term:

$$\langle T_{\omega_k^\vee}\rangle = \sum_{\substack{I\subset\{1,\dots,N\}\\|I|=k}} \prod_{\substack{i\in I\\ j\notin I}} e^{2\pi\mathrm{i}b_i}\left(\frac{\sin\pi(a_{ij}-m)\sin\pi(a_{ji}-m)}{\sin\pi(a_{ij}-\frac12\epsilon)\sin\pi(a_{ji}-\frac12\epsilon)}\right)^{\frac12}. \tag{4.74}$$

(For $\mathcal{N} = 2^*$ theory the choice of the signs $\sigma = \pm$ is irrelevant.) The vev matches the Wigner transform of the trigonometric Ruijsenaars operator.

In the case of symmetric powers, we find a discrepancy between this proposal and the formula (4.51). The vev of the 't Hooft line with $\mathbf{m} = kh_1^\vee$ in $\mathcal{N} = 2^*$ theory, as computed by that formula, can be expressed as the Moyal product of k copies of the vev for the vector representation [69, 74]:

$$\langle T_{kh_1^\vee}\rangle = \langle T_{h_1^\vee}\rangle \star \cdots \star \langle T_{h_1^\vee}\rangle\,. \tag{4.75}$$

The Moyal product $\star$ has the property that $\langle f\rangle\star\langle g\rangle = \langle fg\rangle$ with respect to the Wigner transform, so this equation means that we have

$$\langle T_{kh_1^\vee}\rangle = \langle \mathcal{T}_m^k\rangle\,. \tag{4.76}$$

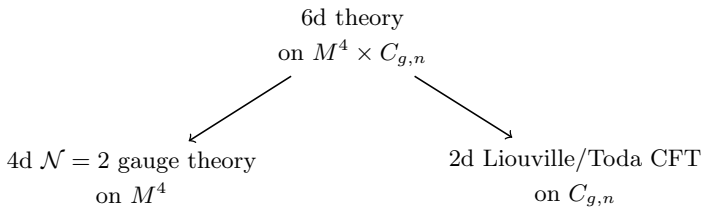
However, $\mathcal{T}_m^k$ is the transfer matrix in the tensor product representation $(\mathbb{C}^N)^{\otimes k}$, not the kth symmetric power of $\mathbb{C}^N$. The discrepancy might be ascribed to subtle monopole contributions to the vev of the 't Hooft line.

4.3 Transfer matrices from Verlinde operators

Here we briefly review the role of line and surface operators in the AGT correspondence. Let us first give the general statements of the correspondence and the notion of class $\mathcal{S}$ theories, which already appeared in section 3. For more details, see excellent reviews [75–77] and the references therein.

The AGT correspondence was found heuristically by Alday, Gaiotto and Tachikawa [6], based on the observations made in [44]. It was derived basically by comparing the instanton partition function [78] of the $\mathcal{N} = 2$ SU(2) gauge theory with the Liouville conformal block. Soon after their work, many checks and generalizations have been made, and it has been known that the statement applies to the correspondence between a large class of four-dimensional $\mathcal{N} = 2$ supersymmetric gauge theories on four-manifold M^4, which is now called the theories of class $\mathcal{S}$, labeled by punctured Riemann surfaces $C_{g,n}$ and two-dimensional non-supersymmetric QFTs defined on the Riemann surfaces.

The origin of the AGT correspondence is the six-dimensional $(2, 0)$ SCFT. Let the four-manifold M^4 be $\mathbb{R}^4_{\epsilon_1,\epsilon_2}$ (or $S^4_{\epsilon_1,\epsilon_2}$). Then this higher dimensional theory explains the correspondence such that the two theories are originally a single theory in six dimensions:



(4.77)

where ϵ_1 and ϵ_2 are called Ω-deformation parameters in the literature, which turn out to realize the deformation quantization of the moduli spaces. For $M^4 = \mathbb{R}^4_{\epsilon_1,\epsilon_2}$, the Ω-deformation causes two-dimensional rotations on the planes $\mathbb{R}^2_{\epsilon_1}$ and $\mathbb{R}^2_{\epsilon_2}$:

$$\mathbb{R}^4_{\epsilon_1,\epsilon_2} := \mathbb{R}^2_{\epsilon_1} \times \mathbb{R}^2_{\epsilon_2}. \tag{4.78}$$

Hence, the Ω-deformation breaks the whole rotation symmetry SO(4) of $\mathbb{R}^4$ to its subgroup $\mathrm{SO}(2)_{\epsilon_1} \times \mathrm{SO}(2)_{\epsilon_2}$. Ω deformation leads to the Gaussian regularization of $\mathbb{R}^4$ and $\mathbb{R}^2$, and their infinite volumes are regularized such that

$$\mathrm{Vol}\left(\mathbb{R}^4_{\epsilon_1,\epsilon_2}\right) = \int_{\mathbb{R}^4_{\epsilon_1,\epsilon_2}} 1 = \frac{1}{\epsilon_1\epsilon_2}, \qquad \mathrm{Vol}\left(\mathbb{R}^2_{\epsilon}\right) = \int_{\mathbb{R}^2_{\epsilon}} 1 = \frac{1}{\epsilon} \tag{4.79}$$

Such a twist allows us to calculate quantities that are invariant under the $\mathrm{U}(1) \simeq \mathrm{SO}(2)$ action on each plane, using the Duistermaat-Heckman fixed point formula [79] (or more

generally Atiyah-Bott localization formula [80]). Nekrasov applied these techniques to compute instanton partition function, which is an integral over the instanton moduli space, and obtained the exact results [78, 81] which plays a crucial role in the AGT correspondence.

4.3.1 Review of AGT correspondence

We have computed the vevs of a class of Wilson–'t Hooft lines in $\mathcal{N}=2$ supersymmetric gauge theories described by a circular quiver, and found that they match the Wigner transforms of transfer matrices constructed from the fundamental trigonometric L-operators. In this section, we show that these transfer matrices can also be identified with Verlinde operators in Toda theory on a punctured torus. The result is in keeping with the relation proposed in [69] based on the AGT correspondence [6] between Toda theory and $\mathcal{N}=2$ supersymmetric field theories.

The AGT correspondence originates from six-dimensional $\mathcal{N}=(2,0)$ supersymmetric field theory, of type A_{N-1} in our case, placed on $S^4_{\mathsf{b}} \times C_{g,n}$. Here, S^4_{b} is an ellipsoid, defined as a submanifold of $\mathbb{R}^5$ by the equation

$$(x^1)^2 + \mathsf{b}^{-2}\big((x^2)^2 + (x^3)^2\big) + \mathsf{b}^2\big((x^4)^2 + (x^5)^2\big) = r^2\,, \tag{4.80}$$

and $C_{g,n}$ is a Riemann surface of genus g with n punctures. With partial topological twisting along $C_{g,n}$, this system preserves eight of the sixteen supercharges of $\mathcal{N}=(2,0)$ supersymmetry in six dimensions.

In the limit in which $C_{g,n}$ shrinks to a point, the six-dimensional theory reduces to a four-dimensional $\mathcal{N}=2$ supersymmetric field theory on S^4_{b}, whose gauge and matter contents are determined by the choice of a pants decomposition of $C_{g,n}$ and boundary conditions at the punctures [44, 45]. The theories discussed in section 4.2 can all be obtained in this way. If one instead integrates out the modes along S^4_{b}, one is left with A_{N-1} Toda theory on $C_{g,n}$ with central charge

$$c = 1 + 6q^2\,, \qquad q = \mathsf{b} + \mathsf{b}^{-1}\,, \tag{4.81}$$

with vertex operators V_{β^r}, $r=1,\,\ldots,\,n$, inserted at the punctures. According to the AGT correspondence, the partition function of the theory on S^4_{b} equals the correlation function of Toda theory on $C_{g,n}$:

$$\langle 1\rangle_{S^4_{\mathsf{b}}} = \Big\langle \prod_r V_{\beta^r} \Big\rangle_{C_{g,n}}\,. \tag{4.82}$$

Let the vertex operators at the punctures be primary fields V_{β^r}, $r=1,\,\ldots,\,n$, labeled by momenta $\beta^r \in \mathfrak{h}^*$ valued in the dual of the Cartan subalgebra $\mathfrak{h}$ of $\mathfrak{sl}_N$. Given a pants decomposition of $C_{g,n}$, the Toda correlation function takes the form

$$\Big\langle \prod_r V_{\beta^r} \Big\rangle_{C_{g,n}} = \int [\mathrm{d}\alpha]\, \mathcal{C}(\alpha;\beta)\overline{\mathcal{F}(\alpha;\beta)}\mathcal{F}(\alpha;\beta)\,, \tag{4.83}$$

where $[\mathrm{d}\alpha]$ is a measure of integration over the set $\alpha = \{\alpha^1,\ldots,\alpha^{3g-3+n}\}$ of momenta assigned to the internal edges of the pants decomposition, $\beta = \{\beta^1,\ldots,\beta^n\}$ is the set of

momenta assigned to the external edges, $\mathcal{C}(\alpha;\beta)$ is the product of relevant three-point functions, and $\mathcal{F}(\alpha;\beta)$ is the corresponding conformal block which is a meromorphic function of α and β.

On the gauge theory side, $\mathcal{C}(\alpha;\beta)$ is interpreted as the product of the classical and the one-loop contributions to the partition function on S^4_{b}, whereas $\mathcal{F}(\alpha;\beta)$ and $\overline{\mathcal{F}(\alpha;\beta)}$ represent the nonperturbative contributions from instantons localized at the two poles at $x^2 = x^3 = x^4 = x^5 = 0$. The internal momenta α are related to the zero modes a of scalar fields in the vector multiplets by

$$\alpha = Q + \mathsf{i}\mathsf{a}\,, \tag{4.84}$$

and the external momenta β are identified with mass parameters for matter multiplets.

4.3.2 Verlinde operators and Wilson–'t Hooft lines

To incorporate Wilson–'t Hooft lines in the gauge theory, one introduces Verlinde loop operators in the Toda theory. We will explain the construction of relevant Verlinde operators in concrete examples. For the moment, it suffices to say that they are specified by a momentum of the form $\mu = -\mathsf{b}\lambda$ and a one-cycle γ in $C_{g,n}$, where λ is the highest weight of a representation of $\mathfrak{sl}_N$.[16] In the presence of a Verlinde operator $\Phi_\mu(\gamma)$, the Toda correlation function is modified to

$$\left\langle \Phi_\mu(\gamma) \prod_r V_{\beta^r} \right\rangle_{C_{g,n}} = \int [\mathrm{d}\alpha]\, \mathcal{C}(\alpha;\beta)\overline{\mathcal{F}(\alpha;\beta)}\big(\Phi_\mu(\gamma)\cdot\mathcal{F}(\alpha;\beta)\big)\,. \tag{4.85}$$

The AGT correspondence asserts [82, 83] that this is equal to the vev of a Wilson–'t Hooft line $T_{\mu,\gamma}$ winding around a circle S^1_{b} where $x^4 = x^5 = 0$ (at $x^1 = 0$, say):

$$\langle T_{\mu,\gamma}\rangle_{S^4_{\mathsf{b}}} = \left\langle \Phi_\mu(\gamma) \prod_r V_{\beta^r} \right\rangle_{C_{g,n}}. \tag{4.86}$$

It turns out that $\Phi_\mu(\gamma)$ acts on conformal blocks as a difference operator shifting the internal momenta α, just as Wilson–'t Hooft lines in $\mathcal{N} = 2$ supersymmetric gauge theories on $S^1 \times_\epsilon \mathbb{R}^2 \times \mathbb{R}$ shift Coulomb branch parameters. Indeed, it was argued in [69] that if one defines the modified Verlinde operator

$$\mathcal{L}_\mu(\gamma) = \mathcal{C}(\alpha;\beta)^{\frac{1}{2}}\Phi_\mu(\gamma)\mathcal{C}(\alpha;\beta)^{-\frac{1}{2}}\,, \tag{4.87}$$

then its Wigner transform is equal to the vev of the Wilson–'t Hooft line in the theory on $S^1 \times_\epsilon \mathbb{R}^2 \times \mathbb{R}$, up to an appropriate identification of parameters:

$$\langle T_{\mu,\gamma}\rangle_{S^1\times_\epsilon\mathbb{R}^2\times\mathbb{R}} = \langle \mathcal{L}_\mu(\gamma)\rangle\,. \tag{4.88}$$

Therefore, we expect that for suitable choices of $C_{g,n}$, β, μ and γ, the modified Verlinde operator $\mathcal{L}_\mu(\gamma)$ coincides with a transfer matrix constructed from the trigonometric L-operator.

[16] More generally, the momentum takes the form $\mu = -\mathsf{b}\lambda_1 - \mathsf{b}^{-1}\lambda_2$, where λ_1, λ_2 are the highest weights of a pair of representations of $\mathfrak{sl}_N$. The corresponding Wilson–'t Hooft line is a superposition of lines wrapping S^1_{b} and another circle $S^1_{\mathsf{b}^{-1}}$ where $x^2 = x^3 = 0$.

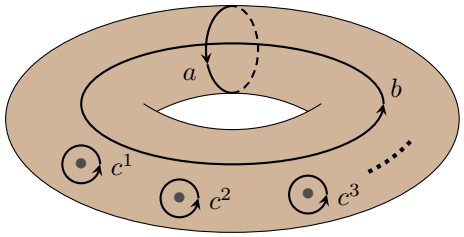


Figure 17: One-cycles on a punctured torus. The cycle c^r goes around the rth puncture.

4.3.3 Verlinde operators on a punctured torus

To reproduce the transfer matrix (4.43), we consider Toda theory on an n-punctured torus $C_{1,n}$ and insert vertex operators V_{β^r} with

$$\beta^r = -N\Big(\frac{q}{2} + \mathsf{i}\mathsf{m}^r\Big)h_N\,. \tag{4.89}$$

The corresponding four-dimensional theory on S^4_{b} is the one described by an n-node circular quiver, which we studied in section 4.2.2. The parameter m^r is the mass of the bifundamental hypermultiplet between the rth and $(r+1)$th nodes.

To this setup we introduce the Verlinde operator $\Phi_\mu(\gamma)$ with

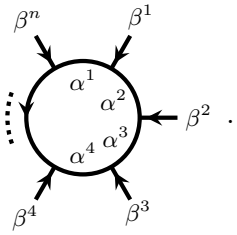
$$\mu = -\mathsf{b}\omega_1 = -\mathsf{b}h_1 \tag{4.90}$$

and γ being a cycle γ_σ specified by an n-tuple of signs $\sigma \in \{\pm\}^n$. If b and c^r are the cycles shown in figure 17, then

$$\gamma_\sigma = b + \sum_r \frac{1-\sigma^r 1}{2} c^r\,. \tag{4.91}$$

In other words, the curve γ_σ passes "above" or "below" the rth puncture depending on whether $\sigma^r = +$ or $-$. In the gauge theory, this operator corresponds to the Wilson–'t Hooft line with magnetic charge (4.60) and electric charge (4.64).

Let us explain the construction of this Verlinde operator step by step, following the treatment in [84]. To this end, it is convenient to represent the conformal block graphically as

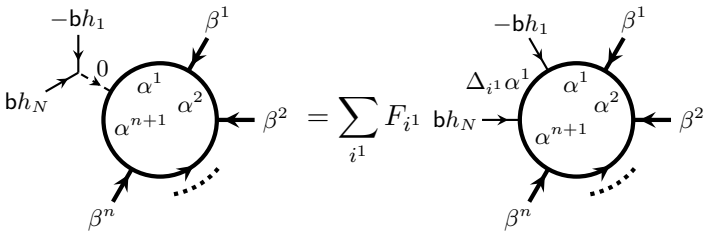


(4.92)

The internal momenta are α^r, $r = 1, \ldots, n+1$, with $\alpha^{n+1} = \alpha^1$.

The first step is to insert the identity operator between β^n and β^1, and resolve it into

the chiral vertex operators $V_{-\mathsf{b}h_1}$ and $V_{\mathsf{b}h_N}$ by fusion. This step gives the equality



(4.93)

The difference operator Δ_i acts on internal momenta by

$$\Delta_i \alpha = \alpha - \mathsf{b}h_i \,. \tag{4.94}$$

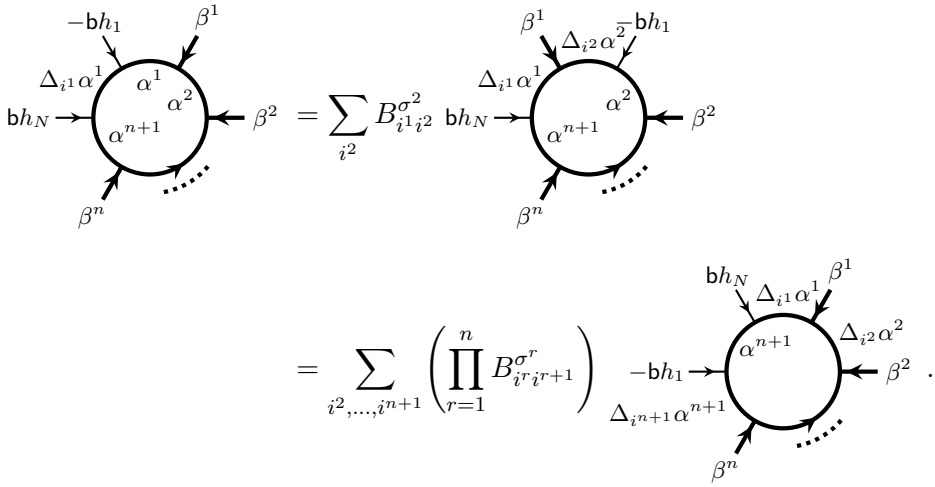
The function F_{i^1} is given by

$$F_{i^1} = \frac{\Gamma(N\mathsf{b}q)}{\Gamma(\mathsf{b}q)} \prod_{j^1(\neq i^1)} \frac{\Gamma(\mathsf{iba}^1_{j^1 i^1})}{\Gamma(\mathsf{b}q + \mathsf{iba}^1_{j^1 i^1})} \,, \tag{4.95}$$

with

$$Q = q\rho \,, \qquad \rho = \sum_{i=1}^{N-1} \omega_i \,. \tag{4.96}$$

Next, we transport $V_{-\mathsf{b}h_1}$ along γ_σ. Graphically, we move the external edge labeled $-\mathsf{b}h_1$ clockwise. Every time the line passes another external edge we get a braiding factor:

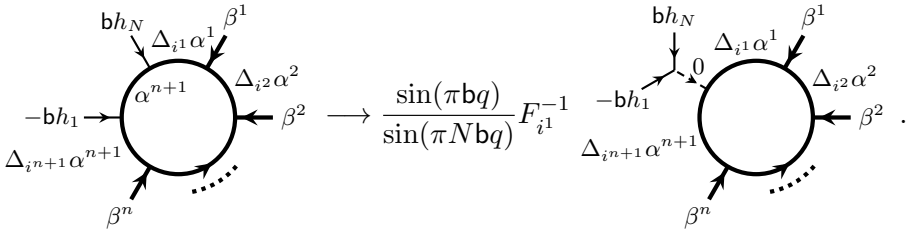


(4.97)

The function $B^{\sigma^r}_{i^r i^{r+1}}$ depends on the sign σ^r, which specifies the direction of the braiding moves:

$$B^{\sigma^r}_{i^r i^{r+1}} = e^{-\sigma^r \pi \mathsf{b}(\mathsf{a}^r_{i^r} - \mathsf{a}^{r+1}_{i^{r+1}})} \prod_{j^r(\neq i^r)} \frac{\Gamma(\mathsf{b}(q + \mathsf{ia}^r_{j^r i^r}))}{\Gamma(\mathsf{b}(\frac{1}{2}q + \mathsf{ia}^r_{j^r} - \mathsf{ia}^{r+1}_{i^{r+1}} - \mathsf{im}^r))}$$
$$\times \prod_{j^{r+1}(\neq i^{r+1})} \frac{\Gamma(\mathsf{iba}^{r+1}_{j^{r+1} i^{r+1}})}{\Gamma(\mathsf{b}(\frac{1}{2}q + \mathsf{ia}^{r+1}_{j^{r+1}} - \mathsf{ia}^r_{i^r} + \mathsf{im}^r))} \,. \tag{4.98}$$

Finally, we fuse $V_{-\mathsf{b}h_1}$ and $V_{\mathsf{b}h_N}$ and project the result to the channel in which the intermediate state is the identity operator:



(4.99)

Note that the right-hand side vanishes unless $i^{n+1} = i^1$ since $\alpha^{n+1} = \alpha^1$.

Thus, dropping the overall factor $\sin(\pi\mathsf{b}q)/\sin(\pi N\mathsf{b}q)$, we find that the Verlinde operator is the difference operator

$$\Phi_{-\mathsf{b}h_1}(\gamma_\sigma) = \sum_{i^1,\dots,i^n} \left(\prod_r B^{\sigma^r}_{i^r i^{r+1}}\right) \Delta_{\{i^1,\dots,i^n\}}\,, \tag{4.100}$$

where $i^{n+1} = i^1$ and

$$\Delta_{\{i^1,\dots,i^n\}} = \prod_r \Delta^r_{i^r}\,. \tag{4.101}$$

Before we compare the Verlinde operator with the transfer matrix, we must perform a change of basis and find the modified Verlinde operator (4.87). For the correlation function at hand, the product of three-point function factors is

$$\mathcal{C}(\alpha;\beta) = \prod_r \frac{\prod_{i<j}\Upsilon(\mathsf{i}a^r_{ij})\Upsilon(-\mathsf{i}a^{r+1}_{ij})}{\prod_{i,j}\Upsilon(\frac{1}{2}q + \mathsf{i}m^r - \mathsf{i}a^r_i + \mathsf{i}a^{r+1}_j)}\,. \tag{4.102}$$

The precise definition of the function Υ is not important for us; we just need to know that it satisfies the identity

$$\frac{\Upsilon(x+\mathsf{b})}{\Upsilon(x)} = \frac{\Gamma(\mathsf{b}x)}{\Gamma(1-\mathsf{b}x)}\mathsf{b}^{1-2\mathsf{b}x}\,, \tag{4.103}$$

where Γ is the gamma function.

Let us calculate $\mathcal{C}(\alpha;\beta)\Delta_{\{i^1,\dots,i^n\}}\mathcal{C}(\alpha;\beta)^{-1}$. The only nontrivial contributions come from the Υ-factors in which either of i or j (but not both) in a^r_{ij} is equal to i^r:

$$\begin{aligned}
&\mathcal{C}(\alpha;\beta)\Delta_{\{i^1,\dots,i^n\}}\mathcal{C}(\alpha;\beta)^{-1}\\
&= \prod_r \prod_{(i^r<)j} \frac{\Upsilon(\mathsf{i}a^r_{i^r j})}{\Upsilon(\mathsf{i}a^r_{i^r j} - \mathsf{b})} \prod_{i(<i^r)} \frac{\Upsilon(\mathsf{i}a^r_{ii^r})}{\Upsilon(\mathsf{i}a^r_{ii^r} + \mathsf{b})}\\
&\quad\times \prod_{(i^{r+1}<)j} \frac{\Upsilon(-\mathsf{i}a^{r+1}_{i^{r+1}j})}{\Upsilon(-\mathsf{i}a^{r+1}_{i^{r+1}j} + \mathsf{b})} \prod_{i(<i^{r+1})} \frac{\Upsilon(-\mathsf{i}a^{r+1}_{ii^{r+1}})}{\Upsilon(-\mathsf{i}a^{r+1}_{ii^{r+1}} - \mathsf{b})}\\
&\quad\times \prod_{i(\neq i^r)} \frac{\Upsilon(\frac{1}{2}q + \mathsf{i}m^r - \mathsf{i}a^r_i + \mathsf{i}a^{r+1}_{i^{r+1}} - \mathsf{b})}{\Upsilon(\frac{1}{2}q + \mathsf{i}m^r - \mathsf{i}a^r_i + \mathsf{i}a^{r+1}_{i^{r+1}})} \prod_{j(\neq i^{r+1})} \frac{\Upsilon(\frac{1}{2}q + \mathsf{i}m^r - \mathsf{i}a^r_{i^r} + \mathsf{i}a^{r+1}_j + \mathsf{b})}{\Upsilon(\frac{1}{2}q + \mathsf{i}m^r - \mathsf{i}a^r_{i^r} + \mathsf{i}a^{r+1}_j)}\,.
\end{aligned} \tag{4.104}$$

Combining the first two lines and using the aforementioned identity, we can rewrite this quantity as

$$\begin{aligned}&\mathcal{C}(\alpha;\beta)\Delta_{\{i^1,\dots,i^n\}}\mathcal{C}(\alpha;\beta)^{-1}\\&=\prod_r\prod_{j^r(\neq i^r)}\frac{\Gamma(1-\mathsf{b}(q+\mathsf{i}\mathsf{a}^r_{j^ri^r}))\Gamma(\mathsf{b}(\frac{1}{2}q+\mathsf{i}\mathsf{a}^r_{j^r}-\mathsf{i}\mathsf{a}^{r+1}_{i^{r+1}}-\mathsf{i}\mathsf{m}^r))}{\Gamma(\mathsf{b}(q+\mathsf{i}\mathsf{a}^r_{j^ri^r}))\Gamma(1-\mathsf{b}(\frac{1}{2}q+\mathsf{i}\mathsf{a}^r_{j^r}-\mathsf{i}\mathsf{a}^{r+1}_{i^{r+1}}-\mathsf{i}\mathsf{m}^r))}\\&\quad\times\prod_{j^{r+1}(\neq i^{r+1})}\frac{\Gamma(1-\mathsf{i}\mathsf{b}\mathsf{a}^{r+1}_{j^{r+1}i^{r+1}})\Gamma(\mathsf{b}(\frac{1}{2}q+\mathsf{i}\mathsf{a}^{r+1}_{j^{r+1}}-\mathsf{i}\mathsf{a}^r_{i^r}+\mathsf{i}\mathsf{m}^r))}{\Gamma(\mathsf{i}\mathsf{b}\mathsf{a}^{r+1}_{j^{r+1}i^{r+1}})\Gamma(1-\mathsf{b}(\frac{1}{2}q+\mathsf{i}\mathsf{a}^{r+1}_{j^{r+1}}-\mathsf{i}\mathsf{a}^r_{i^r}+\mathsf{i}\mathsf{m}^r))}\,.\end{aligned}\tag{4.105}$$

Plugging this expression into the formula for the modified Verlinde operator, we see that the various factors of gamma functions combine nicely into sine functions via Euler's reflection formula

$$\Gamma(x)\Gamma(1-x)=\frac{\pi}{\sin(\pi x)}\,.\tag{4.106}$$

The final result is

$$\begin{aligned}\mathcal{L}_{-\mathsf{b}h_1}(\gamma_\sigma)=\sum_{i^1,\dots,i^n}\Bigg(\prod_r e^{-\sigma^r\pi\mathsf{b}(\mathsf{a}^r_{i^r}-\mathsf{a}^{r+1}_{i^{r+1}})}\prod_{j^r(\neq i^r)}\left(\frac{\sin\pi\mathsf{b}(\frac{1}{2}q+\mathsf{i}\mathsf{a}^r_{j^r}-\mathsf{i}\mathsf{a}^{r+1}_{i^{r+1}}-\mathsf{i}\mathsf{m}^r)}{\sin\pi\mathsf{b}(q+\mathsf{i}\mathsf{a}^r_{j^ri^r})}\right)^{\frac{1}{2}}\\\times\prod_{j^{r+1}(\neq i^{r+1})}\left(\frac{\sin\pi\mathsf{b}(\frac{1}{2}q+\mathsf{i}\mathsf{a}^{r+1}_{j^{r+1}}-\mathsf{i}\mathsf{a}^r_{i^r}+\mathsf{i}\mathsf{m}^r)}{\sin\pi\mathsf{i}\mathsf{b}\mathsf{a}^{r+1}_{j^{r+1}i^{r+1}}}\right)^{\frac{1}{2}}\Bigg)\Delta_{\{i^1,\dots,i^n\}}\,.\end{aligned}\tag{4.107}$$

The above expression can be written in terms of the functions (4.40) as

$$\begin{aligned}&\mathcal{L}_{-\mathsf{b}h_1}(\gamma_\sigma)\\&\quad=\sum_{i^1,\dots,i^n}\Delta^{\frac{1}{2}}_{\{i^1,\dots,i^n\}}\Bigg(\prod_r\ell_{\mathsf{i}\mathsf{b}\mathsf{m}^r+\frac{1}{2}}(\mathsf{i}\mathsf{b}\mathsf{a}^r,\mathsf{i}\mathsf{b}\mathsf{a}^{r+1})^{i^{r+1}}_{i^r}e^{\sigma^r\pi\mathsf{i}(\mathsf{i}\mathsf{b}\mathsf{a}^r_{i^r}-\mathsf{i}\mathsf{b}\mathsf{a}^{r+1}_{i^{r+1}})}\Bigg)\Delta^{\frac{1}{2}}_{\{i^1,\dots,i^n\}}\,.\end{aligned}\tag{4.108}$$

Comparing this expression with the Wigner transform (4.43) of the trigonometric transfer matrix, we deduce that the modified Verlinde operator coincides with the transfer matrix,

$$\mathcal{L}_{-\mathsf{b}h_1}(\gamma_\sigma)=\mathcal{T}_{\sigma,m}\,,\tag{4.109}$$

under the identification

$$\epsilon=\mathsf{b}^2\,,\qquad a^r=\mathsf{i}\mathsf{b}\mathsf{a}^r\,,\qquad m^r=\mathsf{i}\mathsf{b}\mathsf{m}^r+\frac{1}{2}\,.\tag{4.110}$$

It has been proposed in [69] that precisely under this identification of parameters, a modified Verlinde operator in Toda theory corresponding to a Wilson–'t Hooft line in the AGT-dual theory on S^4_{b} reproduces the Weyl quantization of the same Wilson–'t Hooft line in the same theory, but placed in the spacetime $S^1\times_\epsilon\mathbb{R}^2\times\mathbb{R}$. Therefore, we again reach the conclusion that the vev of the Wilson–'t Hooft line with charge (4.60) and (4.64) are equal to the Wigner transform of the trigonometric transfer matrix (4.43).

4.4 Brane realization and string dualities

The AGT correspondence between Wilson–'t Hooft lines and Verlinde operators, which we exploited in section 4.3, can be realized in terms of branes in string theory. String dualities relate the brane configuration for the AGT correspondence to another configuration that realizes four-dimensional Chern–Simons theory, and in the latter setup the emergence of quantum integrability can be seen more transparently. Another chain of dualities relate these setups to the one studied in [29, 63], which provided the initial motivation for the present work. In this last section we discuss these brane constructions.

4.4.1 M-theory setup and brane realization

As explained in section 4.3, the field theoretic origin of the AGT correspondence is six-dimensional $\mathcal{N} = (2, 0)$ superconformal field theory, which in our context is of type A_{N-1} and compactified on an n-punctured torus $C_{1,n}$. This theory describes the low-energy dynamics of a stack of N M5-branes (modulo the center-of-mass degrees of freedom), intersected by n M5-branes.

Consider M-theory in the eleven-dimensional spacetime

$$M_{11} = \mathbb{R}_0 \times \mathbb{R}^2_{12} \times_\epsilon S^1_3 \times_{-\epsilon} \mathbb{R}^2_{45} \times S^1_6 \times \mathbb{R}_7 \times \mathbb{R}_8 \times \mathbb{R}_9 \times S^1_{10} \,. \tag{4.111}$$

(The subscripts indicate the directions in which the spaces extend.) We put N M5-branes $\mathrm{M5}_i$, $i = 1, \ldots, N$, on

$$M_{\mathrm{M5}_i} = \mathbb{R}_0 \times \mathbb{R}^2_{12} \times_\epsilon S^1_3 \times \{0\} \times S^1_6 \times \{0\} \times \{0\} \times \{0\} \times S^1_{10} \,. \tag{4.112}$$

They realize $\mathcal{N} = (2, 0)$ superconformal field theory on $\mathbb{R}_0 \times \mathbb{R}^2_{12} \times_\epsilon S^1_3 \times C_1$, with

$$C_1 = S^1_6 \times S^1_{10} \,. \tag{4.113}$$

Further, we introduce n M5-branes $\mathrm{M5}^r$, $r = 1, \ldots, n$, with worldvolumes

$$M_{\mathrm{M5}^r} = \mathbb{R}_0 \times \mathbb{R}^2_{12} \times_\epsilon S^1_3 \times \{0\} \times \{l^r\} \times \{0\} \times \mathbb{R}_8 \times \mathbb{R}_9 \times \{\theta^r\} \,. \tag{4.114}$$

These M5-branes create codimension-two defects in the six-dimensional theory, located at n points (l^r, θ^r) on C_1, making C_1 an n-punctured torus $C_{1,n}$.

The two sets of M5-branes share a four-dimensional part of the spacetime, $\mathbb{R}_0 \times \mathbb{R}^2_{12} \times_\epsilon S^1_3$, and on this four-dimensional spacetime we get an $\mathcal{N} = 2$ supersymmetric gauge theory with gauge group $G = \mathrm{SU}(N)^n$, described by the circular quiver with n nodes. (More precisely, the gauge group is $\mathrm{SU}(N)^n \times \mathrm{U}(1)$ but the U(1) factor is associated with the center-of-mass and decoupled from the rest of the theory.) In fact, reduction on S^1_{10} turns $\mathrm{M5}_i$ into D4-branes $\mathrm{D4}_i$ on

$$M_{\mathrm{D4}_i} = \mathbb{R}_0 \times \mathbb{R}^2_{12} \times_\epsilon S^1_3 \times \{0\} \times S^1_6 \times \{0\} \times \{0\} \times \{0\} \tag{4.115}$$

and $\mathrm{M5}^r$ into NS5-branes $\mathrm{NS5}^r$ on

$$M_{\mathrm{NS5}^r} = \mathbb{R}_0 \times \mathbb{R}^2_{12} \times_\epsilon S^1_3 \times \{0\} \times \{l^r\} \times \{0\} \times \mathbb{R}_8 \times \mathbb{R}_9 \,, \tag{4.116}$$

and the above brane configuration becomes the well-known D4–NS5 brane configuration for the circular quiver theory [85]. The difference $l^{r+1} - l^r$ in the x^6-coordinate between $\mathrm{NS5}^{r+1}$ and $\mathrm{NS5}^r$ is inversely proportional to the square of the gauge coupling for the rth gauge group, whereas the difference $\theta^{r+1} - \theta^r$ in the x^{10}-coordinate is the θ-angle for the rth gauge group.[17]

A Wilson–'t Hooft line in this four-dimensional theory is realized by an M2-brane on

$$M_{\mathrm{M2}} = \{t_0\} \times \{0\} \times S^1_3 \times \{0\} \times S^1_6 \times \{0\} \times \mathbb{R}^{\geq 0}_8 \times \{x_0\} \times \{\theta_0\}\,, \tag{4.117}$$

where $\mathbb{R}^{\geq 0}_8$ is the nonnegative part of $\mathbb{R}_8$. Upon reduction on S^1_{10}, this M2-brane becomes a D2-brane on

$$M_{\mathrm{D2}} = \{t_0\} \times \{0\} \times S^1_3 \times \{0\} \times S^1_6 \times \{0\} \times \mathbb{R}^{\geq 0}_8 \times \{x_0\} \tag{4.118}$$

and creates a Wilson–'t Hooft line of the type considered in section 4.2. It corresponds to a Verlinde operator in Toda theory on $C_{1,n}$, constructed from a vertex operator transported along the path $S^1_6 \times \{\theta_0\}$. We will explain in a moment how to get the other relevant Verlinde operators.

To understand the relation to quantum integrable systems, let us compactify $\mathbb{R}_9$ to a circle S^1_9 of radius R_9. By doing so, we are uplifting the four-dimensional gauge theory to a five-dimensional one, compactified on a circle. Indeed, by T-duality on S^1_9 we get D5-branes $\check{\mathrm{D}}5_i$, NS5-branes $\check{\mathrm{NS}}5^r$ and a D3-brane $\check{\mathrm{D}}3$ with worldvolumes

$$M_{\check{\mathrm{D}}5_i} = \mathbb{R}_0 \times \mathbb{R}^2_{12} \times_\epsilon S^1_3 \times \{0\} \times S^1_6 \times \{0\} \times \{0\} \times \check{S}^1_9\,, \tag{4.119}$$

$$M_{\check{\mathrm{NS}}5^r} = \mathbb{R}_0 \times \mathbb{R}^2_{12} \times_\epsilon S^1_3 \times \{0\} \times \{l^r\} \times \{0\} \times \mathbb{R}_8 \times \check{S}^1_9\,, \tag{4.120}$$

$$M_{\check{\mathrm{D}}3} = \{t_0\} \times \{0\} \times S^1_3 \times \{0\} \times S^1_6 \times \{0\} \times \mathbb{R}^{\geq 0}_8 \times \check{S}^1_9\,. \tag{4.121}$$

The D5- and NS5-branes intersect along $\mathbb{R}_0 \times \mathbb{R}^2_{12} \times_\epsilon S^1_3 \times \check{S}^1_9$, where a five-dimensional circular quiver theory arises. Recall that the radius $\check{R}_9$ of the dual circle $\check{S}^1_9$ is inversely proportional to the original radius, $\check{R}_9 = \alpha'/R_9$.

Going back to the M-theory setup, we reduce it on S^1_3 and apply T-duality on S^1_9. Then, $\mathrm{M5}_i$ become D5-branes $\widetilde{\mathrm{D5}}_i$ on

$$M_{\widetilde{\mathrm{D5}}_i} = \mathbb{R}_0 \times \mathbb{R}^2_{12} \times \{0\} \times S^1_6 \times \{0\} \times \{0\} \times \check{S}^1_9 \times S^1_{10}\,, \tag{4.122}$$

$\mathrm{M5}^r$ become D3-branes $\widetilde{\mathrm{D3}}^r$ on

$$M_{\widetilde{\mathrm{D3}}^r} = \mathbb{R}_0 \times \mathbb{R}^2_{12} \times \{0\} \times \{l^r\} \times \{0\} \times \mathbb{R}_8 \times \{\phi^r\} \times \{\theta^r\}\,, \tag{4.123}$$

and M2 becomes a fundamental string $\widetilde{\mathrm{F1}}$ on

$$M_{\widetilde{\mathrm{F1}}} = \{t_0\} \times \{0\} \times \{0\} \times S^1_6 \times \{0\} \times \mathbb{R}^{\geq 0}_8 \times \{\phi_0\} \times \{\theta_0\}\,. \tag{4.124}$$

[17] To realize nonzero values for the parameters a^r and b^r, we break each D4-brane $\mathrm{D4}_i$ into n segments $\mathrm{D4}^r_i$ suspended between neighboring NS5-branes and allow these segments to be located anywhere on $\mathbb{R}_8 \times \mathbb{R}_9$. Then, a^r_i is a complex linear combination of the x^9-coordinate of $\mathrm{D4}^r_i$ and the background holonomy of the U(1) gauge field on $\mathrm{D4}^r_i$ around S^1_3. The definition of b^r_i is similar, but involves both the x^8- and x^9-coordinates as well as a chemical potential for the magnetic charge at infinity which does not have a simple interpretation in this brane system.

4.4.2 String duality and 4d Chern-Simons theory

The N D5-branes $\widetilde{\text{D5}}_i$ support $\mathcal{N} = (1,1)$ super Yang–Mills theory with gauge group SU(N) on $\mathbb{R}_0 \times \mathbb{R}^2_{12} \times S^1_6 \times \check{S}^1_9 \times S^1_{10}$. Crucially, this theory is deformed in the present case due to the fact that the product of S^1_3 and $\mathbb{R}^2_{12}$ was twisted. This is a deformation of the type studied in [86], and in the sector in which the relevant supersymmetry is preserved, the deformed theory is actually equivalent to a four-dimensional variant of Chern–Simons theory, with Planck constant $\hbar \propto \epsilon$ [87]. Four-dimensional Chern–Simons theory, here placed on $\mathbb{R}_0 \times S^1_6 \times \check{S}^1_9 \times S^1_{10}$, depends topologically on the cylinder

$$\Sigma = \mathbb{R}_0 \times S^1_6 \tag{4.125}$$

and holomorphically on the torus

$$E = \check{S}^1_9 \times S^1_{10}\,. \tag{4.126}$$

The D3-branes $\widetilde{\text{D3}}^r$ create line defects extending in the longitudinal direction of Σ and located at the points

$$w^r = \phi^r + \mathrm{i}\theta^r \tag{4.127}$$

on E. The fundamental string $\widetilde{\text{F1}}$, on the other hand, creates a Wilson line in the vector representation that winds around the circumferential direction and is located at

$$z_0 = \phi_0 + \mathrm{i}\theta_0 \tag{4.128}$$

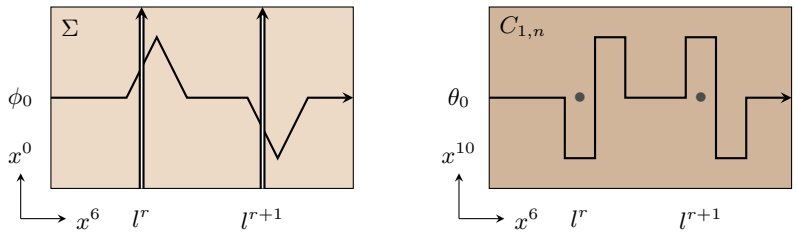
on E. Thus, on the cylinder Σ, we have the same situation as in figure 16(c), in which a quantum spin chain was described in terms of lines on a cylinder.

Indeed, a quantum integrable system emerges from such a configuration of line operators in four-dimensional Chern–Simons theory [88]. The Hilbert space of the integrable system is the space of states of the field theory on a time slice (where the x^0-coordinate is constant) intersected by line operators extending in the time direction. The integrability is a consequence of the topological–holomorphic nature of the theory: by the topological invariance on Σ, one can slide line operators winding around the cylinder continuously along the longitudinal direction; and if two such line operators are located at different points on E, one can move them past each other without encountering a phase transition, thereby establishing the commutativity of transfer matrices.

It was argued in [87], based on the earlier work [29, 63], that a crossing of line defects created by a D3-brane and a fundamental string produces the elliptic L-operator (4.21) with $z = z_0$, $w = w^r$ and $y = 0$[18] (up to shifts by constants). The parameter τ is the modulus of E:

$$\tau = \mathrm{i}\frac{R_{10}}{\check{R}_9}\,. \tag{4.129}$$

[18] More generally, D3^r can be split into two semi-infinite D3-branes D3$^r_+$ and D3$^r_-$, each ending on the stack of D5-branes at $x^8 = 0$. The parameter y is given by the separation of these two halves in E. In the five-dimensional circular quiver theory, the separation is proportional to the complex mass parameter m^r for the bifundamental hypermultiplet charged under the rth and $(r+1)$th gauge groups.



(a) (b)

Figure 18: (a) A path in Σ bending near solid lines. (b) The corresponding path in $C_{1,n}$ detours around the punctures.

Now, take the limit $\check{R}_9 \to 0$, in which $\check{S}^1_9$ shrinks to a point, S^1_9 decompactifies, and the five-dimensional circular quiver theory reduces to the four-dimensional one. This is the trigonometric limit $\tau \to \mathrm{i}\infty$, so we conclude that the transfer matrix constructed from the trigonometric L-operator arises from a Wilson–'t Hooft line in the four-dimensional circular quiver theory.

In the previous sections we studied the transfer matrix $\mathcal{T}_{\sigma,m}$ associated with the cycle γ_σ in $C_{1,n}$ specified by an n-tuples of signs σ. The Wilson–'t Hooft line considered above corresponds to a specific choice of σ. Those corresponding to the other choices can also be constructed in a similar manner, but the construction is a little more subtle. Let us explain how this construction works from the point of view of four-dimensional Chern–Simons theory.

For simplicity, let us set all $\theta^r = \theta_0$. (Since $\mathcal{T}_{\sigma,m}$ is independent of the spectral parameters z_0 and w^r, we do not lose anything by this specialization.) According to the analysis of [89], framing anomaly requires that if a Wilson line curves by an angle φ, its coordinate on S^1_{10} must be shifted by $-\epsilon N\varphi/2\pi$. We can make use of this property to get a Wilson line supported on the cycle γ_σ: fix a small value φ_0 and let the Wilson line bends by the angle $\sigma^r\varphi_0$ right before it crosses the rth solid line, as illustrated in figure 18.

The trigonometric limit $\check{R}_9 \to 0$ is equivalent to the limit $R_{10} \to \infty$. In this limit, $C_{1,n}$ is elongated by an infinite factor in the x^{10}-direction and the dashed line is located at $z = l^r - \sigma^r \mathrm{i}\infty$ when it crosses the rth solid line. This is precisely the limit that appears in the definitions of the fundamental L-operators (4.38), from which $\mathcal{T}_{\sigma,m}$ is constructed.

The D5–NS5–D3 brane system (4.119)–(4.121) is another interesting duality frame. It is actually possible to introduce an additional set of NS5-branes so that the 5-brane system realizes a four-dimensional $\mathcal{N} = 1$ supersymmetric gauge theory on $\mathbb{R}^2_{12} \times_\epsilon S^1_3 \times \check{S}^1_9$. The D3-brane creates a surface defect in this theory. As expected, it acts on the partition function of the theory as an elliptic transfer matrix [29, 63].

5 Outlook

A Special functions and some formulas

A.1 Theta functions

The theta function with characteristics is defined by

$$\theta\left[\begin{array}{c} a \\ b \end{array}\right](\zeta|\tau) = \sum_{n\in\mathbb{Z}} e^{\pi \mathrm{i}(n+a)^2\tau + 2\pi\mathrm{i}(n+a)(\zeta+b)}, \tag{A.1}$$

where ζ is a complex variable and τ is a complex parameter in the upper half-plane: $\operatorname{Im}\tau > 0$. The Jacobi's theta functions are defined by

$$\theta_1(\zeta|\tau) = -\theta\left[\begin{array}{c} 1/2 \\ 1/2 \end{array}\right](\zeta|\tau), \tag{A.2}$$

$$\theta_2(\zeta|\tau) = \theta_1(\zeta + 1/2|\tau), \tag{A.3}$$

$$\theta_3(\zeta|\tau) = e^{\pi\mathrm{i}(\zeta+\tau/4)}\theta_2(\zeta+\tau/2|\tau), \tag{A.4}$$

$$\theta_4(\zeta|\tau) = \theta_3(\zeta + 1/2|\tau). \tag{A.5}$$

The first of these, θ_1, is an odd function of ζ and satisfies

$$\theta_1(\zeta+1|\tau) = -\theta_1(\zeta|\tau), \tag{A.6}$$

$$\theta_1(\zeta+\tau|\tau) = -e^{\pi\mathrm{i}(2\zeta-\tau)}\theta_1(\zeta|\tau). \tag{A.7}$$

The other three are even functions. We have

$$2\theta_1(\zeta+\zeta')\theta_1(\zeta-\zeta') = \bar\theta_4(\zeta)\bar\theta_3(\zeta') - \bar\theta_4(\zeta')\bar\theta_3(\zeta), \tag{A.8}$$

$$2\theta_2(\zeta+\zeta')\theta_2(\zeta-\zeta') = \bar\theta_3(\zeta)\bar\theta_3(\zeta') - \bar\theta_4(\zeta')\bar\theta_4(\zeta), \tag{A.9}$$

$$2\theta_3(\zeta+\zeta')\theta_3(\zeta-\zeta') = \bar\theta_3(\zeta)\bar\theta_3(\zeta') + \bar\theta_4(\zeta')\bar\theta_4(\zeta), \tag{A.10}$$

$$2\theta_4(\zeta+\zeta')\theta_4(\zeta-\zeta') = \bar\theta_4(\zeta)\bar\theta_3(\zeta') + \bar\theta_4(\zeta')\bar\theta_3(\zeta). \tag{A.11}$$

with $\theta_a(\zeta) := \theta_a(\zeta|\tau)$ and $\bar\theta_a(\zeta) := \theta_a(\zeta|\tau/2)$.

It is useful to define another kind of theta function which we call the modified theta function:

$$\theta(z;p) = (z;p)_\infty (p/z;p)_\infty, \tag{A.12}$$

$$(z;p)_\infty := \prod_{k=0}^{\infty}(1-p^k z), \quad |p|<1. \tag{A.13}$$

It satisfies

$$\theta(z;p) = \theta(p/z;p). \tag{A.14}$$

Set $z = e^{2\pi\mathrm{i}\zeta}$, $p = e^{2\pi\mathrm{i}\tau}$ and introduce multiplicative notation

$$\theta_a(z;p) := \theta_a(\zeta|\tau). \tag{A.15}$$

Then the Jacobi's theta functions are rewritten in terms of the modified theta function as

$$\theta_1(z;p) = \mathrm{i}p^{1/8}(p;p)_\infty z^{-1/2}\theta(z;p), \tag{A.16}$$

$$\theta_2(z;p) = p^{1/8}(p;p)_\infty z^{-1/2}\theta(-z;p), \tag{A.17}$$

$$\theta_3(z;p) = (p;p)_\infty \theta(-\sqrt{p}z;p), \tag{A.18}$$

$$\theta_4(z;p) = (p;p)_\infty \theta(\sqrt{p}z;p). \tag{A.19}$$

A.2 Elliptic gamma function

The elliptic gamma function is closely related to the triple gamma function and depends on two complex parameters p and q:

$$\Gamma(z;p,q) = \prod_{j,k=0}^{\infty} \frac{1-p^{j+1}q^{k+1}z^{-1}}{1-p^jq^kz}; \quad |p|,\,|q| < 1. \tag{A.20}$$

It satisfies the identities

$$\Gamma(z;p,q)\Gamma(pq/z;p,q) = 1 \tag{A.21}$$

and

$$\Gamma(pz;p.q) = \theta(z;q)\Gamma(z;p,q), \tag{A.22}$$

$$\Gamma(qz;p.q) = \theta(z;p)\Gamma(z;p,q). \tag{A.23}$$

The function $\Gamma(z;p,q)$ has a pole at $z = p^{-j}q^{-k}$, where j, k are non-negative integers. The residue at this pole is given by

$$\mathrm{Res}_{z=p^{-j}q^{-k}}\left[\Gamma(z;p,q)\right] = \frac{(-1)^{jk+j+k}p^{(k+1)j(j+1)/2}q^{(j+1)k(k+1)/2}}{(p;p)_\infty(q;q)_\infty\theta(p,\ldots,p^j;q)\theta(q,\ldots,q^k;p)}, \tag{A.24}$$

where we have introduced the notation $\theta(z_1,\ldots,z_n;q) := \theta(z_1;q)\cdots\theta(z_n;q)$.

Let t_j, $j = 1,\ldots,6$ be six complex parameters such that $|t_j| < 1$ and $\prod_{j=1}^6 t_j = pq$. Then, we have the following identity proved in [90].

$$\frac{(p;p)_\infty(q;q)_\infty}{2}\int_{\mathbb{T}}\frac{dz}{2\pi\mathrm{i}z}\frac{\prod_{j=1}^6\Gamma(t_jz^{\pm1};p,q)}{\Gamma(z^{\pm2};p,q)} = \prod_{1\le j<k\le 6}\Gamma(t_jt_k;p,q). \tag{A.25}$$

Here $\mathbb{T}$ is the unit circle with counterclockwise orientation and

$$\Gamma(z^{\pm n};p,q) := \Gamma(z^n;p,q)\Gamma(z^{-n};p,q). \tag{A.26}$$

The left-hand side of the above formula is known as the elliptic beta integral.

A.3 The function Γ_{b} and upsilon function

The function Γ_{b} is related to the double gamma function, which is another relative of the ordinary gamma function. It can be defined by an integral representation

$$\log\Gamma_{\mathsf{b}}(x) = \int_0^\infty \frac{dt}{t}\left(\frac{e^{-xt}-e^{-Qt/2}}{(1-e^{-\mathsf{b}t})(1-e^{-t/\mathsf{b}})} - \frac{(Q-2x)^2}{8e^t} - \frac{Q-2x}{t}\right), \tag{A.27}$$

where x is a complex variable and b is a complex parameter such that $\operatorname{Re} x$, $\operatorname{Re}\mathsf{b} > 0$, and $Q = \mathsf{b} + \mathsf{b}^{-1}$.

It is self-dual:

$$\Gamma_{\mathsf{b}}(x) = \Gamma_{1/\mathsf{b}}(x), \tag{A.28}$$

and satisfies the identity

$$\frac{\Gamma_{\mathsf{b}}(x+\mathsf{b})}{\Gamma_{\mathsf{b}}(x)} = \sqrt{2\pi}\frac{\mathsf{b}^{\mathsf{b}x-1/2}}{\Gamma(\mathsf{b}x)}, \tag{A.29}$$

where the denominator in the right-hand side is the ordinary gamma function.

Through the function Γ_{b}, we define upsilon function as

$$\Upsilon_{\mathsf{b}}(x) = \frac{1}{\Gamma_{\mathsf{b}}(x)\Gamma_{\mathsf{b}}(Q-x)}. \tag{A.30}$$

The upsilon function also has an integral representation, which is convergent in the strip $0 < \operatorname{Re} x < Q$:

$$\log\Upsilon_{\mathsf{b}}(x) = \int_0^\infty \frac{dt}{t}\left[\left(\frac{Q}{2}-x\right)^2 e^{-t} - \frac{\sinh^2\left((Q/2-x)t/2\right)}{\sinh(\mathsf{b}t/2)\sinh(t/2\mathsf{b})}\right]. \tag{A.31}$$

The upsilon function is again self-dual:

$$\Upsilon_{\mathsf{b}}(x) = \Upsilon_{1/\mathsf{b}}(x), \tag{A.32}$$

and satisfies the identity

$$\frac{\Upsilon_{\mathsf{b}}(x+\mathsf{b})}{\Upsilon_{\mathsf{b}}(x)} = \frac{\Gamma(\mathsf{b}x)}{\Gamma(1-\mathsf{b}x)}\mathsf{b}^{1-2\mathsf{b}x}. \tag{A.33}$$

References

[1] J. Yagi, *Quiver gauge theories and integrable lattice models*, *JHEP* **10** (2015) 065, [1504.04055].

[2] M. Atiyah, *Topological quantum field theories*, *Inst. Hautes Etudes Sci. Publ. Math.* **68** (1989) 175–186.

[3] X.-z. Dai and D. S. Freed, *eta invariants and determinant lines*, *J. Math. Phys.* **35** (1994) 5155–5194, [hep-th/9405012].

[4] A. Belavin, A. M. Polyakov and A. Zamolodchikov, *Infinite Conformal Symmetry in Two-Dimensional Quantum Field Theory*, *Nucl. Phys. B* **241** (1984) 333–380.

[5] D. Friedan and S. H. Shenker, *The Analytic Geometry of Two-Dimensional Conformal Field Theory*, *Nucl. Phys. B* **281** (1987) 509–545.

[6] L. F. Alday, D. Gaiotto and Y. Tachikawa, *Liouville Correlation Functions from Four-dimensional Gauge Theories*, *Lett. Math. Phys.* **91** (2010) 167–197, [0906.3219].

[7] E. Witten, *Topological Sigma Models*, *Commun. Math. Phys.* **118** (1988) 411.

[8] E. Witten, *Topological Quantum Field Theory*, *Commun. Math. Phys.* **117** (1988) 353.

[9] E. Witten, *Quantum Field Theory and the Jones Polynomial*, *Commun. Math. Phys.* **121** (1989) 351–399.

[10] G. Segal, *The definition of conformal field theory*, in *Symposium on Topology, Geometry and Quantum Field Theory*, pp. 421–575, 6, 2002.

[11] J. Kock, *Frobenius algebras and 2D topological quantum field theories*, vol. 59 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 2004.

[12] T. Kohno, *Topological invariants for 3-manifolds using representations of mapping class groups. I*, *Topology* **31** (1992) 203–230.

[13] V. Knizhnik and A. Zamolodchikov, *Current Algebra and Wess-Zumino Model in Two-Dimensions*, *Nucl. Phys. B* **247** (1984) 83–103.

[14] V. Turaev and O. Viro, *State sum invariants of 3 manifolds and quantum 6j symbols*, *Topology* **31** (1992) 865–902.

[15] N. Reshetikhin and V. Turaev, *Invariants of three manifolds via link polynomials and quantum groups*, *Invent. Math.* **103** (1991) 547–597.

[16] K. Hori, S. Katz, A. Klemm, R. Pandharipande, R. Thomas, C. Vafa et al., *Mirror symmetry*, vol. 1 of *Clay mathematics monographs*. AMS, Providence, USA, 2003.

[17] E. Witten, *Some comments on string dynamics*, in *STRINGS 95: Future Perspectives in String Theory*, pp. 501–523, 7, 1995, hep-th/9507121.

[18] J. Yagi, *Branes and integrable lattice models*, *Mod. Phys. Lett. A* **32** (2016) 1730003, [1610.05584].

[19] G. W. Moore and G. Segal, *D-branes and K-theory in 2D topological field theory*, hep-th/0609042.

[20] N. Carqueville, *Lecture notes on 2-dimensional defect TQFT*, *Banach Center Publ.* **114** (2018) 49–84, [1607.05747].

[21] K. Costello, *Integrable lattice models from four-dimensional field theories*, *Proc. Symp. Pure Math.* **88** (2014) 3–24, [1308.0370].

[22] V. V. Bazhanov and S. M. Sergeev, *A Master solution of the quantum Yang-Baxter equation and classical discrete integrable equations*, *Adv. Theor. Math. Phys.* **16** (2012) 65–95, [1006.0651].

[23] V. V. Bazhanov and S. M. Sergeev, *Elliptic gamma-function and multi-spin solutions of the Yang-Baxter equation*, *Nucl. Phys. B* **856** (2012) 475–496, [1106.5874].

[24] V. Spiridonov, *Elliptic beta integrals and solvable models of statistical mechanics*, *Contemp. Math.* **563** (2012) 181–211, [1011.3798].

[25] M. Yamazaki, *Quivers, YBE and 3-manifolds*, *JHEP* **05** (2012) 147, [1203.5784].

[26] M. Yamazaki, *New Integrable Models from the Gauge/YBE Correspondence*, *J. Statist. Phys.* **154** (2014) 895, [1307.1128].

[27] D. Gaiotto, L. Rastelli and S. S. Razamat, *Bootstrapping the superconformal index with surface defects*, *JHEP* **01** (2013) 022, [1207.3577].

[28] A. Gadde and S. Gukov, *2d Index and Surface operators*, *JHEP* **03** (2014) 080, [1305.0266].

[29] K. Maruyoshi and J. Yagi, *Surface defects as transfer matrices*, *PTEP* **2016** (2016) 113B01, [1606.01041].

[30] A. Hanany and K. D. Kennaway, *Dimer models and toric diagrams*, hep-th/0503149.

[31] S. Franco, A. Hanany, K. D. Kennaway, D. Vegh and B. Wecht, *Brane dimers and quiver gauge theories*, *JHEP* **01** (2006) 096, [hep-th/0504110].

[32] K. D. Kennaway, *Brane Tilings*, *Int. J. Mod. Phys. A* **22** (2007) 2977–3038, [0706.1660].

[33] M. Yamazaki, *Brane Tilings and Their Applications*, *Fortsch. Phys.* **56** (2008) 555–686, [0803.4474].

[34] N. Seiberg, *Exact results on the space of vacua of four-dimensional SUSY gauge theories*, *Phys. Rev. D* **49** (1994) 6857–6863, [hep-th/9402044].

[35] N. Seiberg, *Electric - magnetic duality in supersymmetric nonAbelian gauge theories*, *Nucl. Phys. B* **435** (1995) 129–146, [hep-th/9411149].

[36] A. Hanany and D. Vegh, *Quivers, tilings, branes and rhombi*, *JHEP* **10** (2007) 029, [hep-th/0511063].

[37] C. Romelsberger, *Counting chiral primaries in $N = 1$, d=4 superconformal field theories*, *Nucl. Phys. B* **747** (2006) 329–353, [hep-th/0510060].

[38] J. Kinney, J. M. Maldacena, S. Minwalla and S. Raju, *An Index for 4 dimensional super conformal theories*, *Commun. Math. Phys.* **275** (2007) 209–254, [hep-th/0510251].

[39] G. Festuccia and N. Seiberg, *Rigid Supersymmetric Theories in Curved Superspace*, *JHEP* **06** (2011) 114, [1105.0689].

[40] V. Spiridonov and G. Vartanov, *Vanishing superconformal indices and the chiral symmetry breaking*, *JHEP* **06** (2014) 062, [1402.2312].

[41] V. P. Spiridonov, *Theta hypergeometric integrals*, *Algebra i Analiz* **15** (2003) 161–215.

[42] E. M. Rains, *Transformations of elliptic hypergeometric integrals*, *Ann. of Math. (2)* **171** (2010) 169–243.

[43] F. Dolan and H. Osborn, *Applications of the Superconformal Index for Protected Operators and q-Hypergeometric Identities to N=1 Dual Theories*, *Nucl. Phys. B* **818** (2009) 137–178, [0801.4947].

[44] D. Gaiotto, $N = 2$ *dualities*, *JHEP* **08** (2012) 034, [0904.2715].

[45] D. Gaiotto, G. W. Moore and A. Neitzke, *Wall-crossing, Hitchin Systems, and the WKB Approximation*, 0907.3987.

[46] L. F. Alday, M. Bullimore, M. Fluder and L. Hollands, *Surface defects, the superconformal index and q-deformed Yang-Mills*, *JHEP* **10** (2013) 018, [1303.4460].

[47] M. Bullimore, M. Fluder, L. Hollands and P. Richmond, *The superconformal index and an elliptic algebra of surface defects*, *JHEP* **10** (2014) 062, [1401.3379].

[48] S. Gukov and E. Witten, *Gauge Theory, Ramification, And The Geometric Langlands Program*, hep-th/0612073.

[49] S. Gukov and E. Witten, *Rigid Surface Operators*, *Adv. Theor. Math. Phys.* **14** (2010) 87–178, [0804.1561].

[50] E. K. Sklyanin, *Some algebraic structures connected with the Yang-Baxter equation*, *Funktsional. Anal. i Prilozhen.* **16** (1982) 27–34, 96.

[51] R. Baxter, *Eight-Vertex Model in Lattice Statistics*, *Phys. Rev. Lett.* **26** (1971) 832–833.

[52] R. J. Baxter, *Partition function of the eight vertex lattice model*, *Annals Phys.* **70** (1972) 193–228.

[53] E. K. Sklyanin, *Some algebraic structures connected with the Yang-Baxter equation. Representations of a quantum algebra*, *Funktsional. Anal. i Prilozhen.* **17** (1983) 34–48.

[54] S. Derkachov and V. Spiridonov, *Yang-Baxter equation, parameter permutations, and the elliptic beta integral*, *Russ. Math. Surveys* **68** (2013) 1027–1072, [1205.3520].

[55] V. P. Spiridonov, *The continuous biorthogonality of an elliptic hypergeometric function*, *Algebra i Analiz* **20** (2008) 155–185.

[56] S. N. M. Ruijsenaars and H. Schneider, *A new class of integrable systems and its relation to solitons*, *Ann. Physics* **170** (1986) 370–405.

[57] S. Ruijsenaars, *Complete Integrability of Relativistic Calogero-moser Systems and Elliptic Function Identities*, *Commun. Math. Phys.* **110** (1987) 191.

[58] K. Hasegawa, *Ruijsenaars' commuting difference operators as commuting transfer matrices*, *Comm. Math. Phys.* **187** (1997) 289–325.

[59] A. Belavin, *Dynamical Symmetry of Integrable Quantum Systems*, *Nucl. Phys. B* **180** (1981) 189–200.

[60] G. Felder, *Conformal field theory and integrable systems associated to elliptic curves*, hep-th/9407154.

[61] G. Felder, *Elliptic quantum groups*, in *11th International Conference on Mathematical Physics (ICMP-11) (Satellite colloquia: New Problems in the General Theory of Fields and Particles, Paris, France, 25-28 Jul 1994)*, pp. 211–218, 7, 1994, hep-th/9412207.

[62] G. Felder and A. Varchenko, *Elliptic quantum groups and Ruijsenaars models*, *J. Statist. Phys.* **89** (1997) 963–980.

[63] J. Yagi, *Surface defects and elliptic quantum groups*, *JHEP* **06** (2017) 013, [1701.05562].

[64] P. Etingof and A. Varchenko, *Solutions of the quantum dynamical Yang-Baxter equation and dynamical quantum groups*, *Comm. Math. Phys.* **196** (1998) 591–640.

[65] R. J. Baxter, *Eight vertex model in lattice statistics and one-dimensional anisotropic Heisenberg chain. 2. Equivalence to a generalized ice-type lattice model*, *Annals Phys.* **76** (1973) 25–47.

[66] M. Jimbo, T. Miwa and M. Okado, *Solvable lattice models whose states are dominant integral weights of* $A^{(1)}_{n-1}$, *Lett. Math. Phys.* **14** (1987) 123–131.

[67] M. Jimbo, T. Miwa and M. Okado, *Local state probabilities of solvable lattice models: an* $A^{(1)}_{n-1}$ *family*, *Nucl. Phys. B* **300** (1988) 74–108.

[68] A. Kapustin, *Wilson-'t Hooft operators in four-dimensional gauge theories and S-duality*, *Phys. Rev. D* **74** (2006) 025005, [hep-th/0501015].

[69] Y. Ito, T. Okuda and M. Taki, *Line operators on* $S^1 \times \mathbb{R}^3$ *and quantization of the Hitchin moduli space*, *JHEP* **04** (2012) 010, [1111.4221].

[70] T. D. Brennan, A. Dey and G. W. Moore, *On 't Hooft defects, monopole bubbling and supersymmetric quantum mechanics*, *JHEP* **09** (2018) 014, [1801.01986].

[71] T. D. Brennan, *Monopole Bubbling via String Theory*, *JHEP* **11** (2018) 126, [1806.00024].

[72] T. D. Brennan, A. Dey and G. W. Moore, *'t Hooft defects and wall crossing in SQM*, *JHEP* **10** (2019) 173, [1810.07191].

[73] B. Assel and A. Sciarappa, *On monopole bubbling contributions to 't Hooft loops*, *JHEP* **05** (2019) 180, [1903.00376].

[74] H. Hayashi, T. Okuda and Y. Yoshida, *Wall-crossing and operator ordering for 't Hooft operators in* $\mathcal{N} = 2$ *gauge theories*, *JHEP* **11** (2019) 116, [1905.11305].

[75] B. Le Floch, *A slow review of the AGT correspondence*, 2006.14025.

[76] T. Okuda, *Line operators in supersymmetric gauge theories and the 2d-4d relation*, pp. 195–222. 2016. 1412.7126. .

[77] S. Gukov, *Surface Operators*, pp. 223–259. 2016. 1412.7127. .

[78] N. A. Nekrasov, *Seiberg-Witten prepotential from instanton counting*, *Adv. Theor. Math. Phys.* **7** (2003) 831–864, [hep-th/0206161].

[79] J. Duistermaat and G. Heckman, *On the Variation in the cohomology of the symplectic form of the reduced phase space*, *Invent. Math.* **69** (1982) 259–268.

[80] M. Atiyah and R. Bott, *The Moment map and equivariant cohomology*, *Topology* **23** (1984) 1–28.

[81] N. Nekrasov and A. Okounkov, *Seiberg-Witten theory and random partitions*, *Prog. Math.* **244** (2006) 525–596, [hep-th/0306238].

[82] L. F. Alday, D. Gaiotto, S. Gukov, Y. Tachikawa and H. Verlinde, *Loop and surface operators in $\mathcal{N} = 2$ gauge theory and Liouville modular geometry*, *JHEP* **01** (2010) 113, [0909.0945].

[83] N. Drukker, J. Gomis, T. Okuda and J. Teschner, *Gauge Theory Loop Operators and Liouville Theory*, *JHEP* **02** (2010) 057, [0909.1105].

[84] J. Gomis and B. Le Floch, *'t Hooft Operators in Gauge Theory from Toda CFT*, *JHEP* **11** (2011) 114, [1008.4139].

[85] E. Witten, *Solutions of four-dimensional field theories via M theory*, *Nucl. Phys. B* **500** (1997) 3–42, [hep-th/9703166].

[86] J. Yagi, *Ω-deformation and quantization*, *JHEP* **08** (2014) 112, [1405.6714].

[87] K. Costello and J. Yagi, *Unification of integrability in supersymmetric gauge theories*, 1810.01970.

[88] K. Costello, *Supersymmetric gauge theory and the Yangian*, 1303.2632.

[89] K. Costello, E. Witten and M. Yamazaki, *Gauge Theory and Integrability, I*, 1709.09993.

[90] V. P. Spiridonov, *On the elliptic beta function*, *Uspekhi Mat. Nauk* **56** (2001) 181–182.