4d/2d Correspondence: Instantons and $\mathcal{W}$-algebras

Thesis by

Jaewon Song

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2012

(Defended May 17, 2012)

Acknowledgements

The greatest gratitude goes to my advisor Hirosi Ooguri, who constantly supported and guided me. He was a role model in my graduate studies as a scientist. His deep knowledge and clear understanding always stimulated me and I wouldn't have been able to appreciate many aspects of physics without his guidance. I would like to thank John Schwarz for teaching me string theory and also for serving on my defense committee. His enthusiasm and keen insight attracted me to the realms of string theory. I thank Anton Kapustin for teaching me the essence and beauty of the quantum field theory and serving on my defense committee. I thank Sergei Gukov for sharing his insights on mathematical physics. His lecture was always fun and full of ideas, and I was very lucky to be able to share time with him. I thank Maria Spiropulu for being my committee member. It was really a delightful experience to work as her teaching assistant, and her enthusiasm and openness always impressed me.

I would like to thank my collaborators. Christoph Keller has been a very good friend during my graduate studies, and I was very fortunate to work on various projects with him. I was lucky to work with Lotte Hollands. Her knowledge and style greatly expanded the boundary of my knowledge that I didn't appreciate before. Chang-Soon Park was a very good friend, we talked a lot-in and out of physics. His solid understanding of physics and persistence always impressed me. I can't thank Yuji Tachikawa enough. He was very supportive on various stage of my research, and also it was very fortunate and enjoyable to work with him. I also thank Sangmin Lee for suggesting a project, and also being a mentor at various steps. I would like to express my gratitude to Jaehyung Choi. We have shared a tremendous amount of time discussing physics and life-from undergraduate years until now-even while living in different parts of the continent. I was happy to work with Yutaka Ookouchi who is always friendly and full of ideas. I also thank Noppadol Mekareeya for being in IPMU at the same time as I was so that we could start a project together.

I would like to thank Soo-Jong Rey, Piljin Yi, and Kimyeong Lee whose lectures attracted me to theoretical high-energy physics. I also thank them for their hospitality when I visited Korea. I also thank Abhijit Gadde, Dongmin Gang, Yu-tin Huang, Matt Johnson, Hee-Cheol Kim, Seok Kim, Eunkyung Koh, Sungjay Lee, Joseph Marsano, Yu Nakayama, Jaemo Park, Vasily Pestun, Sakura Schafer-Nameki, Jihye Seo, Piotr Sulkowski, and Meng-Chwan Tan for various discussions and advice.

I would also like to thank my fellow graduate students at the Caltech Theory Group: Miguel Bandres, Denis Bashkirov, Kimberly Boddy, Hee-Joong Chung, Paul Cook, Tudor Dimofte, Arthur Lipstein, Chan Youn Park, Heywood Tam, Chien-Yao Tseng, Ketan Vyas, Brian Willett, Itamar Yaakov, Masahito Yamazaki, and Jie Yang. I also thank Carol Silberstein for her help and assistance. The time I have spent in the Caltech Theory Group couldn't have been enjoyable without these people. Various discussions, informal seminars, and chat over lunch and dinner-all of those moments are valuable to me.

I would like to thank many of my friends who prayed for me and also who made my life rich and enjoyable.

Finally, I deeply thank my family. I could not have finished the PhD program without their love and support.

Abstract

In this thesis, we study the 4d/2d correspondence of Alday-Gaiotto-Tachikawa, which relates the class of 4-dimensional $\mathcal{N} = 2$ gauge theories (theories of class $\mathcal{S}$) to a 2-dimensional conformal field theory. The 4d gauge theories are obtained by compactifying 6-dimensional $\mathcal{N} = (2, 0)$ theory of type $\Gamma = A, D, E$ on a Riemann surface C. On the 2-dimensional side, we have Toda theory on the surface C with $\mathcal{W}$-algebra symmetry, which is an extension of the Virasoro symmetry. In particular, the instanton partition function of the 4d gauge theory is reproduced by a conformal/chiral block of Virasoro/$\mathcal{W}$-algebra. We develop techniques to compute the partition functions on 4d and 2d sides, for various gauge groups and matter fields.

We generalize the Alday-Gaiotto-Tachikawa 4d/2d correspondence to various cases. First, we study $\mathcal{N} = 2$ pure Yang-Mills theory with arbitrary gauge groups, including the exceptional groups. We explicitly construct the corresponding $\mathcal{W}$-algebra currents, and confirm the correspondence holds at 1-instanton level. Second, we study the conformal quiver theory with $Sp(1) - SO(4)$ gauge group. Finally, we study Sicilian gauge theories with trifundamental half-hypermultiplets. We also find that the conformal theories with $Sp(1)$ gauge group and $SU(2)$ gauge group have different instanton partition functions in terms of bare gauge couplings. We show this is an artifact of the renormalization scheme, by explicitly constructing a map between the bare couplings and studying its geometrical interpretations. This demonstrates the scheme independence of renormalization at the non-perturbative level. This thesis is mainly based on publications [1, 2, 3].

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Supersymmetry has been a great source for understanding the strong dynamics of quantum field theory. Ordinary field theory without supersymmetries, such as Yang-Mills theory, is known to be notoriously difficult to understand its low-energy behavior. It is believed to have a mass gap and to exhibit confinement. Yang-Mills theory is the basic building block of our current understanding of nature, but we do not have a proper analytic tool to understand its low energy or strong coupling phenomena. Supersymmetric gauge theories on the other hand, are much more tractable to analyze. Especially, the power of holomorphy [4] and duality [5] enables us to obtain exact results.

Generally, one can 'supersymmetrize' a given field theory with different numbers of supersymmetries. The number of supersymmetry usually refers to the number of conserved fermions. For example, in 4-dimensions, a single fermion has 4 real components, and $\mathcal{N}$ supersymmetry means having $4\mathcal{N}$ conserved fermionic charges or supercharges. In $d = 4$, one can get up to $\mathcal{N} = 4$ supersymmetry. The $\mathcal{N} = 1$ supersymmetry is the only phenomenologically viable extension for the Standard Model, since extended supersymmetries cannot have a chiral fermion, which is the crucial feature of the Standard Model. The $\mathcal{N} = 1$ extension of the Standard Model, has been a primary object of study beyond the Standard Model physics. It not only gives a natural cure to the instability of the Standard Model under the quantum loop corrections, but also contains a natural candidate for the dark matter.

On the other hand, $\mathcal{N} = 4$ supersymmetric theories have been a great source of theoretical studies. Most notably, $\mathcal{N} = 4$ Super-Yang-Mills theory [6] has vanishing beta function, and conformal. It is dual to the type IIB superstring theory on $AdS_5 \times S^5$, which provides the prototype of the gauge/gravity duality [7, 8, 9, 10]. Also, there is extensive evidence

that this theory in planar limit is quantum integrable [11]. Also the theory is believed to be S-dual to itself. It enjoys beautiful yet very non-trivial aspects of the conformal field theory. [1]

The $\mathcal{N} = 2$ supersymmetric gauge theories are in the middle. They are still not applicable to phenomenology since one cannot have a chiral theory.[2] Even though they are not immediately useful for building realistic models, $\mathcal{N} = 2$ theories are mathematically much more tractable than $\mathcal{N} = 1$ but sufficiently richer than $\mathcal{N} = 4$ theory. They can be asymptotically free, so that low-energy dynamics can be strongly coupled. Nevertheless, the power of holomorphy is greater than that of $\mathcal{N} = 1$ theories so that one can determine exact low-energy effective action including the Kahler potential. This is due to the seminal work by Seiberg and Witten [15, 16], which we will review in the later chapters. $\mathcal{N} = 2$ theory has also applications to mathematics of 4-manifolds. The topologically twisted theory on a curved 4-manifold computes topological invariants known as Donaldson invariants [17] or equivalently Seiberg-Witten invariants [18].

1.1 4d/2d correspondence: Instantons and W-algebras

The main topic we want to discuss in this thesis is a class of 4-dimensional $\mathcal{N} = 2$ supersymmetric gauge theories and its relation to 2-dimensional conformal field theory. It can be understood in terms of multiple M5-branes wrapped on a Riemann surface C. The world-volume theory of N M5-branes is described by $\mathcal{N} = (2, 0)$ theory of A_{N-1} type. Heuristically, the 4d/2d correspondence can be understood as a result of the path-integral version of the 'Fubuni theorem'. Let us put the $\mathcal{N} = (2, 0)$ theory of type $\Gamma = A, D, E$ on $M_4 \times C$. We choose the spacetime 4-manifold M_4 as a spin manifold in order to preserve supersymmetry. First we integrate along the C direction. The theory is partially topologically twisted along the C direction in order to preserve supersymmetry in 4-dimension. This 'integration' yields 4d $\mathcal{N} = 2$ gauge theory on M_4. Let's denote this theory as $T[\Gamma, C]$. If we do further integration, we get a number, namely the partition function $Z^{4d}_{T[\Gamma,C]}(M_4)$ of the gauge theory. On the other hand, we can integrate out the M_4 direction first to ob-

[1] The only known $\mathcal{N} = 4$ supersymmetric field theory is the maximally supersymmetric Yang-Mills theory with semi-simple gauge group G. It would be interesting to find a proof of the uniqueness or find another example of $\mathcal{N} = 4$ supersymmetric field theory. (J. Maldacena, "$\mathcal{N} = 4$ SYM, 35 Years After" conference)

[2] It is still possible to have a hidden sector with $\mathcal{N} = 2$ coupled to a $\mathcal{N} = 1$ visible sector. Such a model has been studied in, for example, [12, 13, 14].

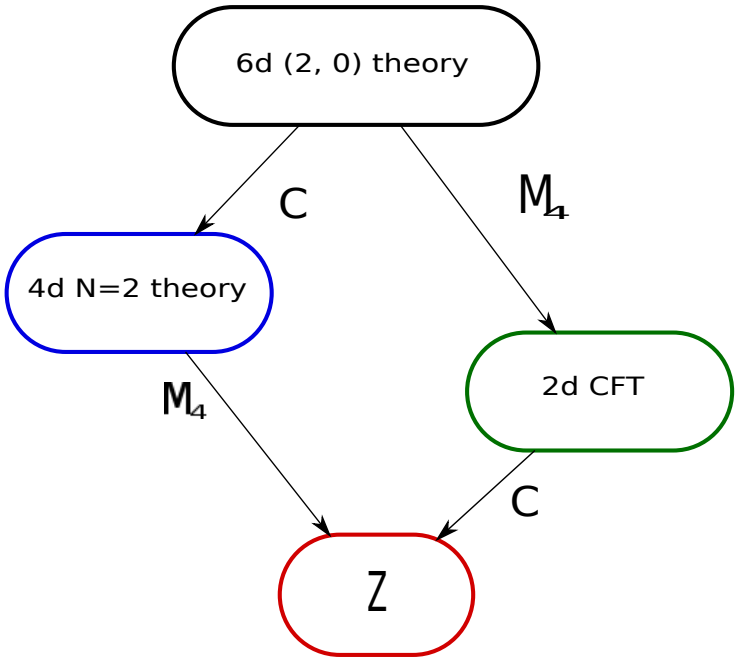


Figure 1.1: The 4d/2d correspondence as a Fubini theorem

tain a 2-dimensional theory on C. Now, the theory on C is determined by M_4. Let's call this 2-dimensional theory $\mathcal{T}[\Gamma, M_4]$. We can path-integrate this theory to obtain a partition function (or correlation function) $Z^{2d}_{\mathcal{T}[\Gamma,M_4]}(C)$. Since the order of integrations does not affect the end result, we get

$$\boxed{Z^{4d}_{T[\Gamma,C]}(M_4) = Z^{2d}_{\mathcal{T}[\Gamma,M_4]}(C).} \tag{1.1}$$

This is the 4d/2d correspondence; quantity in 4d gauge theory is given by another quantity in 2d theory. It also indicates that the 4d/2d correspondence depends on the choice of the spacetime manifold M_4.

The prototype of 4d/2d relation is the case when $C = T^2$. In this case, the 4d gauge theory is nothing but $\mathcal{N} = 4$ super Yang-Mills theory. The 4-dimensional gauge theory has S-duality, which means that the theory remains the same under the S-duality transformation, which sends the gauge coupling $\tau = \frac{\theta}{2\pi} + \frac{4\pi i}{g^2}$ to

$$\tau \to -\frac{1}{\tau}, \quad \tau \to \tau + 1, \tag{1.2}$$

and changes the gauge group G to its Langlands dual group ${}^{L}G$. The action on τ generates the well-known modular group $SL(2,\mathbb{Z})$. It generates the symmetry of the complex structure of a torus. So we can interpret the S-duality of the $\mathcal{N} = 4$ SYM as the equivalence of

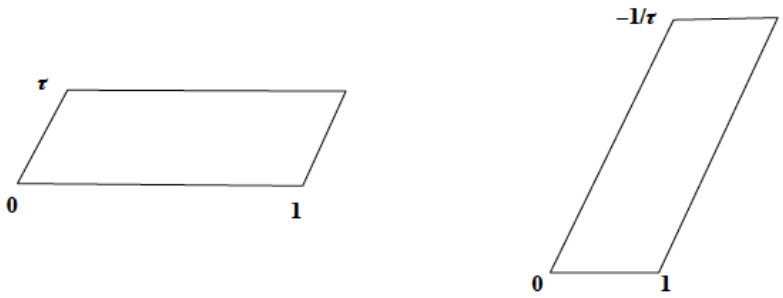


Figure 1.2: Modular transformation on a torus with complex structure τ

the geometric parameter τ under the modular group. The S-duality puts strong constraint on the 4-dimensional theory that the partition function has to be a modular form with certain weight. This is identified as a character of the affine Lie algebra $\hat{\mathfrak{g}}$ [19, 20, 21], and it is understood as free bosons living on the 2-dimensional surface C [22, 23, 24].

If we choose C to be generic and curved, the theory is no longer supersymmetric unless we do the partial topological twist along the surface C. Once this is done, we obtain $\mathcal{N} = 2$ supersymmetry in 4-dimensions. One can put number of punctures on the surface, change the topology to obtain various gauge theories [25]. One can generalize the notion of S-duality from this picture by looking at various degeneration limits of the same surface.

If we choose the 4-manifold M_4 as S^4, and the Riemann surface C as generic, the 2d theory turns out to be the Γ-type Toda theory. If $\Gamma = A_1$, then it is well-known Liouville theory. This is the case studied by Alday-Gaiotto-Tachikawa (AGT) [26]. The 4d side of the partition function was computed by Pestun [27], which has the following form

$$Z[S^4](\tau, m_i) = \int [da]\, |Z_{Nek}(\tau, \vec{a}, m_i; \epsilon_1 = 1/r, \epsilon_2 = 1/r)|^2 \tag{1.3}$$

where $\vec{a}$ is the Coulomb branch parameter and m_i are the masses of the hypermultiplets in the theory. Z_{Nek} is the Nekrasov partition function [28] on $\mathbb{R}^4$ in the Omega background [29, 23, 30], which we will discuss in the later chapters. The Nekrasov partition function is

decomposed into 3 pieces

$$Z_{Nek} = Z_{tree} Z_{1-loop} Z_{inst}. \tag{1.4}$$

where each piece corresponds to the partition function coming from the tree-level and 1-loop and the last one is the instanton contribution to the partition function. We call Z_{inst} the instanton partition function. The AGT correspondence not only holds at the level of the integral (1.3) but also at the level of integrands. The 2d side of the correspondence is given by a correlation function of the Liouville/Toda theory, which can be schematically written as

$$\langle V(h_{m_1}) \cdots V(h_{m_n}) \rangle_C = \int [da] C(h_{\vec{a}}, h_{m_i}) \left| \mathcal{F}(h_{\vec{a}}, h_{m_i}; b) \right|^2, \tag{1.5}$$

where $V(h_{m_i})$ is a $\mathcal{W}$-primary operator with conformal (or $\mathcal{W}$) weight determined by m_i and $\mathcal{F}$ is the n-point conformal block with external weights given by m_is and internal momentum given by $h_{\vec{a}}$. The parameter b determines the central charge of the CFT and it is mapped to the Omega deformation parameters ϵ_1, ϵ_2. The function $C(a, m_i)$ is the 3-point function of the CFT. The AGT correspondence states that the 1-loop part of the partition function is given by the 3-point function of CFT and the instanton partition function is given by the conformal block.

In this thesis, we will study the relation between instanton partition function and the conformal block:

$$\boxed{Z_{inst}(\vec{a}, m_i; \epsilon_1, \epsilon_2) = \mathcal{F}(h_a, h_m; b).} \tag{1.6}$$

Note that the conformal block doesn't depend on the dynamics of the underlying conformal field theory. In other words, it is completely determined by the representation of the Virasoro or $\mathcal{W}$-algebra. This is rather surprising from the 4d point of view, since the 4d gauge theory instanton partition function is given by an integral over the moduli space of instantons that depends on the geometry of the gauge theory. We will build language and techniques to study this relation in the following chapters.

1.2 Outline

In Chapter 2, we review $\mathcal{N} = 2$ supersymmetric gauge theories and the Seiberg-Witten solution. We will review the construction of $\mathcal{N} = 2$ theories obtained from M5-branes on a Riemann surface called the theories of class $\mathcal{S}$ or generalized quiver theories. We use the geometric language in terms of Hitchin systems, following [31, 25].

In Chapter 3, we study Seiberg-Witten theory microscopically. The method is developed by Nekrasov, using the idea of equivariant localization. Using the technique, we evaluate the integral over the moduli space of instantons with various matter content. Especially, we derive the contour integral formula with half-hypermultiplets such as $Sp - SO$ bifundamental [1, 2]. One of the surprising results we obtain is that the instanton partition function for conformal $Sp(1)$ gauge theory and $SU(2)$ gauge theory differs. We will show that this is an artifact of the renormalization scheme and show that physically they are indeed the same [1].

In Chapter 4, we focus on the 2d side of the correspondence. We will review $\mathcal{W}$-algebras and methods to compute the chiral block or correlation function of the given conformal field theory. Especially, we explicitly construct the $\mathcal{W}$-algebras corresponding to various Lie groups using free-field realizations, and study their representations [3].

Chapter 5 is the main part of the thesis. We will present the correspondence for arbitrary gauge groups including the exceptional groups [3]. In particular, we will show that gauge theories with non-simply-laced gauge groups are mapped into simply-laced $\mathcal{W}$-algebras, but in the twisted representations [1]. We will study the correspondence for non-linear quiver theories, called Sicilian or generalized quiver gauge theories [2]. We focus on A_1 theories, and verify that the correspondence between instanton partition function and Virasoro conformal block indeed works.

In Chapter 6, we make a remark on omitted topics including the other choice of the spacetime 4-manifold M_4 in (1.1), and conclude with summary and future directions.

Chapter 2

$\mathcal{N} = 2$ Supersymmetric Gauge Theories in 4-dimensions

2.1 $\mathcal{N} = 2$ Supersymmetry in general

There are two different types of $\mathcal{N} = 2$ SUSY multiplets, namely vector multiplet and hypermultiplet. An $\mathcal{N} = 2$ vector multiplet consists of a vector field A_μ, and two Weyl fermions $\psi_\alpha, \lambda_\alpha$, and a complex scalar field ϕ. In terms of $\mathcal{N} = 1$ multiplets, it consists of a $\mathcal{N} = 1$ vector $\mathcal{W}_\alpha$ containing A_μ, λ and a chiral multiplet Φ containing ϕ, ψ.

$$\begin{array}{ccc} & A_\mu & \\ \lambda & & \psi \\ & \phi & \end{array} \tag{2.1}$$

All of the fields are in the adjoint representation of the gauge group.

A hypermultiplet consists of 2 Weyl fermions $\psi_q, \psi_{\tilde{q}}$ and 2 complex scalars $q, \tilde{q}$. In $\mathcal{N} = 1$ language, it has 2 chiral multiplets Q containing q, ψ_q, and $\tilde{Q}$ containing $\tilde{q}, \psi_{\tilde{q}}$.

$$\begin{array}{ccc} & \psi_q & \\ q & & \tilde{q}^\dagger \\ & \psi^\dagger_{\tilde{q}} & \end{array} \tag{2.2}$$

Every $\mathcal{N} = 2$ theories has $SU(2)_R$ global symmetry which rotates the set of supercharges. We have arranged (2.1) and (2.2) in such a way that the $SU(2)_R$ rotates the horizontal components. That is (λ, ψ) and $(q, \tilde{q}^\dagger)$ form doublets and all the other component fields are singlets under $SU(2)_R$. In addition, one has $U(1)_R$ symmetry, which is anoma-

lous for asymptotically free theory. Under the R-symmetries, the components in a vector multiplet have charges $R(A_\mu) = 0, R(\lambda) = 1, R(\psi) = 1, R(\phi) = 2$, and the components in a hypermultiplet have charge $R(q) = 0, R(\psi_q) = 1, R(\tilde{q}) = 0, R(\tilde{\psi}_q) = 1$. In terms of superfields, we can write as

$$\mathcal{W}_\alpha \to \mathcal{W}_\alpha(e^{-i\alpha}\theta), \tag{2.3}$$

$$\Phi \to e^{2i\alpha}\Phi(e^{-i\alpha}\theta), \tag{2.4}$$

$$Q \to Q(e^{-i\alpha}\theta), \tag{2.5}$$

$$\tilde{Q} \to \tilde{Q}(e^{-i\alpha}\theta). \tag{2.6}$$

The Lagrangian of $\mathcal{N} = 2$ theory in terms of $\mathcal{N} = 1$ superfields when we have both vector and hypermultiplets, is written as

$$\begin{aligned}\mathcal{L} &= \frac{1}{4\pi}\mathrm{Im}\left[\int d^4\theta \left(\mathrm{Tr}\Phi^\dagger e^{-2V}\Phi + Q^\dagger e^{-2V}Q + \tilde{Q}e^{2V}\tilde{Q}^\dagger\right) + \int d^2\theta \mathrm{Tr}\frac{1}{2}\tau\mathcal{W}_\alpha\mathcal{W}^\alpha\right] \\ &\quad + \int d^2\theta\left(\sqrt{2}\tilde{Q}\Phi Q + m\tilde{Q}Q\right) + (c.c),\end{aligned} \tag{2.7}$$

where the complexified gauge coupling is given by

$$\tau = \frac{\theta}{2\pi} + \frac{4\pi i}{g^2}, \tag{2.8}$$

and m is a mass of the hypermultiplet.

When there is no hypermultiplet, the potential of $\mathcal{N} = 2$ Yang-Mills theory is given by

$$V = \frac{1}{g^2}\mathrm{Tr}[\phi, \phi^\dagger]^2. \tag{2.9}$$

The flat direction is simply given by $[\phi, \phi^\dagger] = 0$. Generically, the vacuum moduli space is given by non-zero vacuum expectation value of the scalar field ϕ in the vector multiplet, and gauge group is broken down to $U(1)^r$ where r is the rank of gauge group. For the unitary gauge group $SU(N)$, one can parametrize the flat direction by Casimir operators $\mathrm{tr}\phi^2, \mathrm{tr}\phi^3, \cdots \mathrm{tr}\phi^n$ up to a gauge transformation.

In general, the classical moduli space of $\mathcal{N} = 2$ theories has two different branches, namely Coulomb branch and Higgs branch. In the Higgs branch, the vev of adjoint scalar

field is zero $\langle\phi\rangle = 0$ but the 'squarks' q gets a vev, and the gauge symmetry is completely broken. The Higgs branch of the moduli space is a hyper-Kahler manifold and does not receive any quantum corrections.

The Coulomb branch is given by $\langle\phi\rangle \neq 0$, and $\langle q\rangle = 0$, and in this case the low-energy theory is given by $U(1)^r$ gauge theory. By the virtue of $\mathcal{N} = 2$ supersymmetry, the Low-energy effective action can be written in terms of the single holomorphic function $\mathcal{F}(a)$ called the prepotential

$$\mathcal{L}_{eff} = \frac{1}{4\pi}\mathrm{Im}\left[\int d^4\theta \frac{\partial\mathcal{F}(A)}{\partial A}\bar{A} + \int d^2\theta \frac{1}{2}\frac{\partial^2\mathcal{F}(A)}{\partial A^2}\mathcal{W}_\alpha\mathcal{W}^\alpha\right], \tag{2.10}$$

where A is the chiral superfield in the $\mathcal{N} = 2$ vector multiplet. Classically, the prepotential is given by

$$\mathcal{F}_{clas}(a) = \frac{1}{2}\tau a^2. \tag{2.11}$$

Of course it receives a quantum correction but $\mathcal{N} = 2$ supersymmetry restricts the form of the prepotential to be [32]

$$\mathcal{F}(\vec{a}) = \frac{i}{2\pi}a^2\log\frac{a^2}{\Lambda^2} + \sum_{k=1}^{\infty}\mathcal{F}_k\left(\frac{\Lambda}{a}\right)^{b_0 k}a^2, \tag{2.12}$$

where b_0 is the coefficient of the beta-function. For example, it is given by $2N_c - N_f$ for the $SU(N_c)$ gauge theory with N_f fundamental hypermultiplets. The first term of the above equation comes from 1-loop correction, which is exact in perturbation theory. The second term comes from non-perturbative effects, namely instanton corrections. We will discuss a systematic way to obtain instanton corrections to the prepotential from now.

2.2 Seiberg-Witten solution

In the seminal paper by Seiberg and Witten [15, 16], they solved the problem of obtaining the prepotential $\mathcal{F}(a)$ in terms of a family of auxiliary Riemann surface Σ_{SW}. If such curve is given, the prepotential can be obtained by evaluating certain period integrals. Suppose

we have a basis of 1-cycles $\{\alpha^1, \beta_1, \cdots, \alpha^r, \beta_r\}$ on Σ_{SW} so that

$$\alpha^i \cap \beta_j = \delta^i_j, \;\; \alpha^i \cap \alpha_j = \beta^i \cap \beta_j = 0. \tag{2.13}$$

Then, the instanton part of the prepotential

$$\mathcal{F}_{inst}(\vec{a}) = a^i_D a_i \tag{2.14}$$

is determined by the period integrals

$$a_i = \oint_{\alpha^i} \lambda_{SW}, \quad a^i_D = \oint_{\beta_i} \lambda_{SW}, \tag{2.15}$$

where λ_{SW} is certain meromorphic 1-form on the Seiberg-Witten curve Σ_{SW}. The Seiberg-Witten curve is a genus r curve parametrized by the Coulomb branch parameters. In the case of $SU(N)$ theory, the Coulomb branch is parametrized by $u_n = \mathrm{tr}\phi^n$ with $n = 2, 3, \cdots, N$.

Originally, the Seiberg-Witten curve was obtained by an educated guess based on the electric-magnetic duality and the singularity structure of the moduli space. For example, the $SU(2)$ gauge theory with no hypermultiplet has the curve of the form

$$y^2 = x(x-1)(x-u) \tag{2.16}$$

and the Seiberg-Witten form is given by

$$\lambda_{SW} = \frac{dx}{y}. \tag{2.17}$$

Of course, this is not the unique form satisfying the desired properties. We will use slightly different form in the next section based on M-theoretic construction of the Seiberg-Witten theory.

2.3 $\mathcal{N} = 2$ gauge theory from M5-branes on a Riemann surface

In this section we start by reviewing the construction of Gaiotto curves (or UV-curves) for $\mathcal{N} = 2$ gauge theories with a classical gauge group. We explain how the $\mathcal{N} = 2$ geometry is encoded in a ramified Hitchin system whose base is the Gaiotto curve. It can be understood as a six-dimensional theory compactified on the UV-curve. We will in particular write down the solutions for pure $\mathcal{N} = 2$ Yang-Mills theory with arbitrary gauge group.

2.3.1 UV-curves and Hitchin systems

Let us start this section with reviewing the constructions of Gaiotto curves for $SU(N)$ and $Sp(N)/SO(N)$ gauge groups and their embedding in a ramified Hitchin system. We then explain how to compare them in specific cases.

2.3.1.1 Unitary gauge group

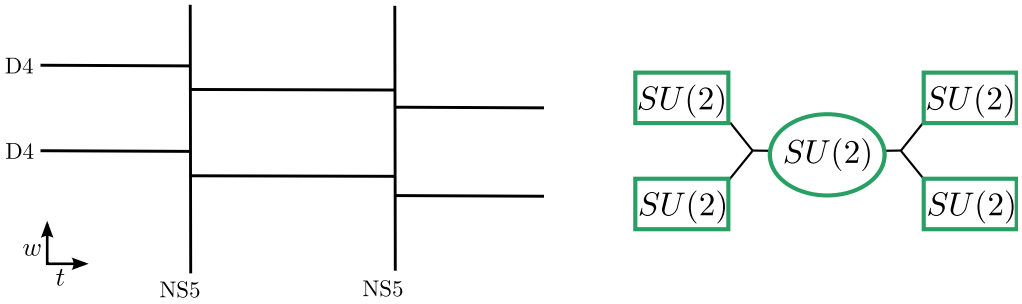


Figure 2.1: Illustrated on the left is an example of a D4/NS5 brane construction realizing the $SU(2)$ quiver gauge theory illustrated on the right. The Coulomb and mass parameters of the $SU(2)$ gauge theory parametrize the separation of the D4-branes, while the separation of the NS5-branes determines the microscopic coupling τ_{UV}. The Seiberg-Witten curve for this $SU(2)$ gauge theory is a torus with complex structure parameter τ_{IR}.

A special unitary quiver gauge theory can be realized in type IIA string theory using a D4/NS5 brane embedding [33]. See Figure 2.1 for the brane embedding of the $SU(2)$ gauge theory coupled to four hypermultiplets. From such a brane embedding one can read off the Seiberg-Witten curve Σ of the quiver gauge theory. It is, roughly speaking, a fattening of the D4/NS5 graph, as it is, for instance, illustrated in Figure 2.1.

The D4/NS5 brane embedding can be lifted to an M5-brane embedding in M-theory. The resulting ten-dimensional M-theory background is

$$\mathbb{R}^4 \times T^*\tilde{C} \times \mathbb{R}^2 \times S^1, \tag{2.18}$$

where we introduced a possibly punctured Riemann surface $\tilde{C}$ and its cotangent bundle $T^*\tilde{C}$. We insert a stack of N M5-branes that wraps the six-dimensional manifold $\mathbb{R}^4 \times \tilde{C}$. The positions of these M5-branes in the cotangent bundle determine the Seiberg-Witten curve Σ as a subspace of $T^*\tilde{C}$. In this perspective the Seiberg-Witten curve is given [25]

$$0 = \det(w - \boldsymbol{\phi}_U) = w^N + w^{N-2}\phi_2 + w^{N-3}\phi_3 + \ldots + \phi_N \tag{2.19}$$

as a branched degree N covering over the so-called Gaiotto curve $\tilde{C}$. The holomorphic differential w parametrizes the fiber direction of the cotangent bundle $T^*\tilde{C}$, whereas $\boldsymbol{\phi}_U$ is an $SU(N)$-valued differential on the curve $\tilde{C}$ of degree 1.

The degree d differentials $\phi_d = \mathrm{Tr}(\boldsymbol{\phi}_U^d)$ encode the classical vev's of the $SU(N)$ Coulomb branch operators of dimension d. They are allowed to have poles at the punctures of the Gaiotto curve. The coefficients at these poles encode the bare mass parameters of the gauge theory. To take care of these boundary conditions in the M-theory set-up, we need to insert additional M5-branes at the punctures of the Gaiotto curve. These M5-branes should intersect the Gaiotto curve transversally, and thus locally wrap the fiber of the cotangent bundle at the puncture [31].

Equation (2.19) determines the Seiberg-Witten curve Σ as an N-fold branched covering over the Gaiotto curve $\tilde{C}$. The Seiberg-Witten differential is simply

$$\lambda = w. \tag{2.20}$$

The $\mathcal{N} = 2$ geometry for unitary gauge groups is thus encoded in a ramified A_N Hitchin system on the punctured Gaiotto curve $\tilde{C}$, whose spectral curve is the Seiberg-Witten curve Σ and whose canonical 1-form is equal to the Seiberg-Witten differential [34].[1]

The topology of the Gaiotto curve is fully determined by the corresponding quiver

[1] A detailed discussion of boundary conditions for this Hitchin system can be found in [31] and references therein.

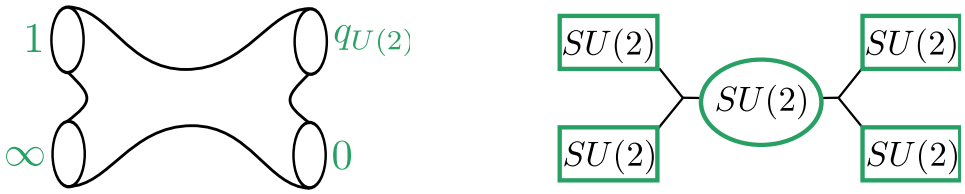


Figure 2.2: The left Figure illustrates the Gaiotto curve $\tilde{C}$ of the conformal $SU(2)$ quiver gauge theory that is illustrated on the right. The Gaiotto curve is a four-punctured sphere with complex structure parameter $q_{U(2)}$. The differential ϕ_2 has second order poles at the four punctures. The $SU(2)$ flavor symmetries are encoded in the coefficients of the differential ϕ_2 at these poles.

diagram. A gauge group translates into a tube of the Gaiotto curve, whereas a flavor group turns into a puncture. This is illustrated in Figure 2.2 for the $SU(2)$ gauge theory coupled to four flavors. The poles of the differentials ϕ_d determine the branch points of the fibration (2.19). Their coefficients encode the flavor symmetry of the quiver gauge theory. For gauge group $SU(N)$ the degree of the poles is integer. As we will see shortly this is not true for $Sp(N)$ and $SO(N)$ gauge theories.

2.3.1.2 Symplectic/orthogonal gauge group

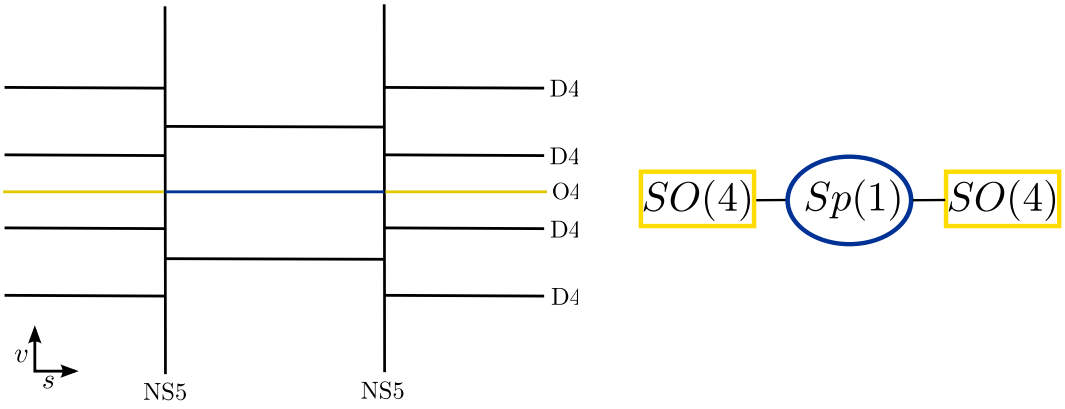


Figure 2.3: Illustrated on the left is an example of a D4/NS5 brane construction with O4$^{\pm}$ orientifold branes realizing the $Sp(1)$ quiver gauge theory illustrated on the right. The O4^{-} branes (in yellow) ensure that both flavor symmetry groups are $SO(4)$, whereas the O4^{+} brane (in blue) ensures that the gauge symmetry group is $Sp(1)$. The brane embedding of the conformal $SO(4)$ gauge theory is found by swapping the inner and the outer D4 and O4 branes.

For symplectic or orthogonal gauge theories a similar description exists. Engineering

these gauge theories in type IIA requires orientifold O4-branes in addition to the D4 and NS5-branes [35, 36, 37]. The orientifold branes are parallel to the D4 branes. They act on the string background as a combination of a worldsheet parity Ω and a spacetime reflection in the five dimensions transverse to it. The space-time reflection introduces a mirror brane for each D4-brane, whereas the worldsheet parity breaks the space-time gauge group. More precisely, there are two kinds of O4-branes, distinguished by the sign of $\Omega^2 = \pm 1$. The O4$^-$ brane breaks the $SU(N)$ gauge symmetry to $SO(N)$, whereas the O4$^+$ brane breaks it to $Sp(N/2)$. The brane construction that engineers the conformal $Sp(1)$ gauge theory is schematically shown in Figure 2.3.[2] Notice that there are two hidden D4-branes on top of the O4$^+$ brane, so that the number of D4-branes is equal at each point over the base.

From these brane setups we can extract the Seiberg-Witten curve for the $Sp(N-1)$ and $SO(2N)$ gauge theories coupled to matter. To find the Gaiotto curve we rewrite the Seiberg-Witten curve in the form [38]

$$0 = v^{2N} + \varphi_2 v^{2N-2} + \varphi_4 v^{2N-4} + \ldots + \varphi_{2N}, \tag{2.21}$$

where the differentials φ_k encode the Coulomb parameters and the bare masses. Equation (2.21) defines the Seiberg-Witten curve as a branched covering over the Sp/SO Gaiotto curve. More precisely, the Seiberg-Witten curve is embedded in the cotangent bundle T^*C of the Gaiotto curve C with holomorphic differential v. The Seiberg-Witten differential is simply

$$\lambda = v, \tag{2.22}$$

the canonical 1-form in the cotangent bundle T^*C.

Whereas for the $Sp(N-1)$ gauge theory there is an extra condition saying that the zeroes at $v = 0$ of the right-hand-side should be double zeroes, the $SO(2N)$ gauge theory requires these zeroes to be simple zeroes. These conditions come up somewhat ad-hoc in the type IIA description, but can be explained from first principles in an M-theory perspective [39]. The orientifold brane construction lifts in M-theory to a stack of M5-branes in a $\mathbb{Z}_2$-orbifold background. The orbifold acts on the five dimensions transverse to the

[2] The Sp and SO brane constructions illustrated here can be naturally extended to linear Sp/SO quivers. We will come back to this in section 5.4.

M5-branes, and in particular maps $v \mapsto -v$.

For the pure $Sp(N-1)$-theory the differential φ_{2N} vanishes, so that a factor v^2 in equation (2.21) drops out. The resulting Seiberg-Witten curve can be written in the form

$$0 = \det(v - \boldsymbol{\varphi}_{Sp}), \tag{2.23}$$

where $\boldsymbol{\varphi}_{Sp}$ is a $Sp(N-1)$-valued differential. The non-vanishing differentials φ_{2k} can thus be obtained from the Casimirs of the Lie algebra $\mathfrak{sp}(N-1)$. If we include massive matter to the $Sp(N-1)$ gauge theory, however, or consider an $SO(2N)$ gauge theory, equation (2.21) can be reformulated as

$$0 = \det(v - \boldsymbol{\varphi}_{SO}) \tag{2.24}$$

where the differential $\boldsymbol{\varphi}_{SO}$ is $SO(2N)$-valued. This equation is clearly characterized by the Casimirs of the Lie algebra $\mathfrak{so}(2N)$. More precisely, we recognize the D_N-invariants $\mathrm{Tr}(\Phi^{2k})$ and $\mathrm{Pfaff}(\Phi)$ in the differentials φ_{2k} and $\varphi_{\tilde{N}} = \sqrt{\varphi_{2N}}$, respectively. In general, the $\mathcal{N} = 2$ geometry for symplectic and orthogonal gauge groups is thus encoded in a ramified D_N Hitchin system based on the Sp/SO Gaiotto curve C, whose spectral curve is the Seiberg-Witten curve (2.21).

The Lie algebra D_N has a $\mathbb{Z}_2$ automorphism under which the invariants with exponent $2, \ldots, 2N-2$ are even and the invariant of degree N is odd. On the level of the differentials ϕ_k this translates into possible half-integer poles for the invariant $\varphi_{\tilde{N}}$. Going around such a pole the differential $\varphi_{\tilde{N}}$ has a $\mathbb{Z}_2$ monodromy. We will see explicitly in the examples. We call the puncture corresponding to such a pole a half-puncture. The half-punctures introduce $\mathbb{Z}_2$ twist-lines on the Gaiotto curve [38]. This is illustrated for the $Sp(1)$ and $SO(4)$ Gaiotto curve in Figure 2.4 and Figure 2.5.

Lastly, let us make a few remarks on the worldvolume theory on a stack of M5-branes. In the low energy limit this theory is thought to be described by a six-dimensional conformal $(2,0)$ theory of type ADE. For the M-theory background (2.18) it is of type A, whereas for the $\mathbb{Z}_2$-orbifolded M-theory background it is of type D. The $(2,0)$ theory has a "Coulomb branch" parametrized by the vev's of a subset of chiral operators whose conformal weights are given by the exponents d of the Lie algebra $\mathfrak{g}$. These operators parametrize

Figure 2.4: The left Figure illustrates the Gaiotto curve C of the conformal $Sp(1)$ quiver gauge theory that is illustrated on the right. The $Sp(1)$ Gaiotto curve differs from the $SU(2)$ Gaiotto curve by the $\mathbb{Z}_2$ twist-line that runs parallel to the tube. We will discuss the precise relation between the $Sp(1)$ and the $SU(2)$ Gaiotto curve in section 3.5.2.

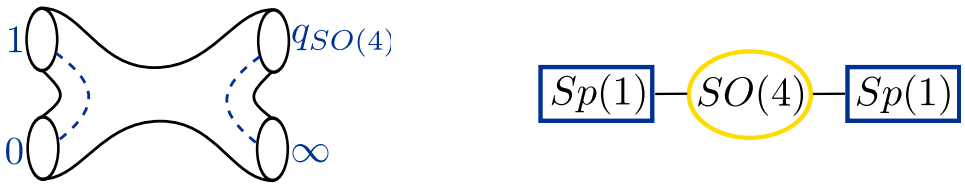


Figure 2.5: The left Figure illustrates the Gaiotto curve C of the conformal $SO(4)$ quiver gauge theory that is illustrated on the right. The $SO(4)$ Gaiotto curve differs from the $Sp(1)$ Gaiotto curve by a different configuration of $\mathbb{Z}_2$ twist-lines. In particular, the twist lines don't run through the tube.

the configurations of M5-branes in the M-theory background. In the Hitchin system they appear as the degree d differentials. Boundary conditions at the punctures of the Gaiotto curve are expected to lift to defect operators in the M5-brane worldvolume theory. We refer to [31] for a more detailed description.

G	$h^\vee$	w_i	G	$h^\vee$	w_i	$\Gamma^{(r)}$
A_{n-1}	n	$2,3,\ldots,n$	B_n	$2n-1$	$2,4,\ldots,2n$	$A^{(2)}_{2n-1}$
D_n	$2n-2$	$2,4,\ldots,2n-2;n$	C_n	$n+1$	$2,4,\ldots,2n$	$D^{(2)}_{n+1}$
E_6	12	$2,5,6,8,9,12$	F_4	9	$2,6,8,12$	$E^{(2)}_6$
E_7	18	$2,6,8,10,12,14,18$	G_2	4	$2,6$	$D^{(3)}_4$
E_8	30	$2,8,12,14,18,20,24,30$				

Table 2.1: The dual Coxeter number $h^\vee$ and the dimensions w_i of the Casimir invariants for all simple Lie groups G. Recall $A_{n-1} = \mathrm{SU}(n)$, $B_n = \mathrm{SO}(2n+1)$, $C_n = \mathrm{Sp}(n)$, $D_n = \mathrm{SO}(2n)$. For non-simply-laced G, the Langlands dual $(G^{(1)})^\vee = \Gamma^{(r)}$ of the affine G algebra is also shown.

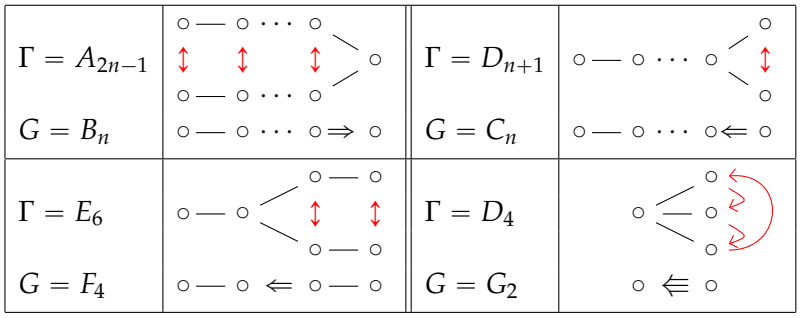


Figure 2.6: The relation between a non-simply-laced Lie algebra G, its associated simply-laced algebra Γ, and the outer automorphism used to fold Γ to obtain G.

2.3.2 Martinec-Warner solution of $\mathcal{N} = 2$ pure theory:

The Seiberg-Witten curve of pure $\mathcal{N} = 2$ gauge theory for arbitrary G was constructed in [40] as the spectral curve of a Toda lattice. (See also [41, 42, 43, 44].) From a modern perspective [45, 25, 31], their construction reads as follows. Let us first consider the case when G is simply-laced. Take 6d $\mathcal{N} = (2,0)$ theory of type G, and compactify this theory on $C = \mathbb{CP}^1$ parametrized by z, with two codimension 2 defects at $z = 0$ and $z = \infty$. The 6d theory has world-volume fields $\phi^{(w_i)}(z)$ on C, transforming as degree w_i multi-differentials, where w_i is the degree of the Casimir invariants of G given in Table 2.1. We then set

$$\phi^{(w_i)}(z) = u^{(w_i)}\,(dz/z)^{w_i}\,, \qquad (w_i \neq h^\vee); \tag{2.25}$$

$$\phi^{(h^\vee)}(z) = \left(\Lambda^{h^\vee} z + u^{(h^\vee)} + \frac{\Lambda^{h^\vee}}{z}\right)(dz/z)^{h^\vee} \qquad (w_i = h^\vee). \tag{2.26}$$

Here $u^{(w_i)}$ is the vev of the dimension w_i Coulomb branch operator, and Λ is the holomorphic dynamical scale of the gauge theory. From this data one can then construct the Seiberg-Witten curve Σ [46], or equivalently the fibration of the ALE space of type G [44, 47]. In the following we label the degrees w_i so that $h^\vee = w_n$.

When G is non simply-laced, we take a pair $(\Gamma, \mathbb{Z}_r)$ such that the twisted affine Lie algebra $\Gamma^{(r)}$ is the Langlands dual to $G^{(1)}$, the untwisted affine algebra of G. In other words, the Dynkin diagram of G is obtained by folding the Dynkin diagram of Γ as in Fig. 2.6.[3] For example, when $G = G_2$, $\Gamma = \mathrm{SO}(8)$ and $r = 3$. We then put 6d theory of type

[3] Γ is called the associated simply-laced algebra of G [48]. G is also known as the orbit Lie algebra of the pair $(\Gamma, \mathbb{Z}_r)$, see [49]. Note that G is *not* the $\mathbb{Z}_r$-invariant part of Γ in general, as explained in Appendix A.2.

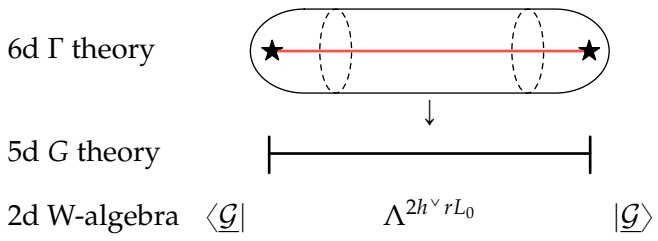


Figure 2.7: Top: the Seiberg-Witten solution of pure $\mathcal{N} = 2$ super Yang-Mills theory with gauge group G in terms of 6d $\mathcal{N} = (2,0)$ theory of type Γ on $C = \mathbb{CP}^1$ with the $\mathbb{Z}_r$ twist line from $z = 0$ to $z = \infty$. Middle: the S^1 reduction to the 5d maximally supersymmetric Yang-Mills theory with gauge group G on a segment, with a suitable half-BPS boundary condition on both ends. Bottom: In the 2d description, the coherent state $\langle \mathcal{G}|$ is produced by the BPS boundary condition. It is then propagated along the horizontal direction and annihilated by $|\mathcal{G}\rangle$.

Γ on $C = \mathbb{CP}^1$, with a twist by $\mathbb{Z}_r$ around two singularities at $z = 0, \infty$. The fields of the 6d theory are divided into two sets, $\phi^{(\hat{w}_i)}(z)$ which are invariant under $\mathbb{Z}_r$ action, and $\phi^{(\tilde{w}_i)}(z)$ which transform nontrivially under $\mathbb{Z}_r$. We then take

$$\phi^{(\hat{w}_i)}(z) = u^{(\hat{w}_i)} (dz/z)^{\hat{w}_i}, \tag{2.27}$$

$$\phi^{(\tilde{w}_i)}(z) = 0, \qquad (\tilde{w}_i \neq h^\vee), \tag{2.28}$$

$$\phi^{(\tilde{w}_i)}(z) = \left(\Lambda^{\tilde{w}_i} z^{1/r} + \frac{\Lambda^{\tilde{w}_i}}{z^{1/r}}\right) (dz/z)^{\tilde{w}_i} \qquad (\tilde{w}_i = h^\vee). \tag{2.29}$$

Here, Λ is the dynamical scale and $u^{(\hat{w}_i)}$ is the vev of the degree-$\hat{w}_i$ Coulomb branch operator; note that the degrees of Casimirs of G are exactly the degrees of Casimirs of Γ invariant under $\mathbb{Z}_r$. Note also that the dual Coxeter number of G are exactly the highest degree of Casimirs of Γ not invariant under $\mathbb{Z}_r$.

The construction is summarized in Fig. 5.2: the 6d $\mathcal{N} = (2,0)$ theory of type Γ on a circle with $\mathbb{Z}_r$ twist gives maximally supersymmetric 5d Yang-Mills theory of gauge group G. To obtain pure $\mathcal{N} = 2$ Yang-Mills theory, we need to put the 5d Yang-Mills theory on a segment with an appropriate half-BPS boundary condition on both ends. The boundary condition then becomes the prescribed singularity of the worldvolume fields $\phi^{(w_i)}(z)$. When G is classical, the 5d Yang-Mills theory can be realized on coincident D4-branes, possibly on top of an O4-plane. Then the BPS boundary condition comes from ending them on an NS5-brane [33, 35, 36, 37].

2.4 S-duality

The Gaiotto construction of $\mathcal{N} = 2$ superconformal field theories gives a nice geometric interpretation of the S-duality of $\mathcal{N} = 2$ gauge theories. One of the prime examples of the theories exhibit S-duality is $SU(2)$ gauge theory with 4 fundamental hypermultiplets. It has vanishing beta function, and believed to be exactly conformal. It can be obtained from 2 M5-branes wrapped on the 4-punctured sphere.

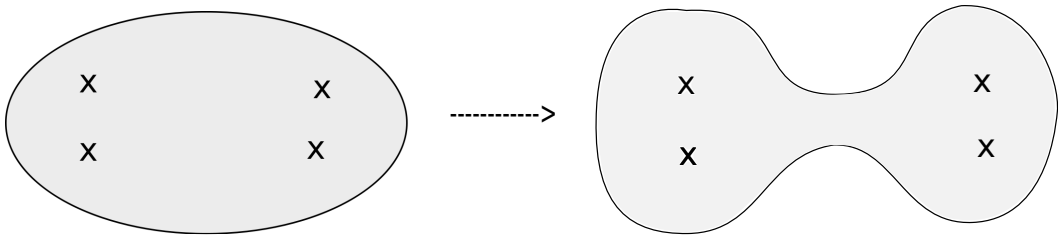


Figure 2.8: A weakly coupled description of the corresponding theory can be read off from going to a degenerate limit. Once this is done, a hhin tube corresponds to a gauge group, and punctures correspond to hypermultiplets. This gives $SU(2)$ gauge theory with $N_f = 4$ hypermultiplets

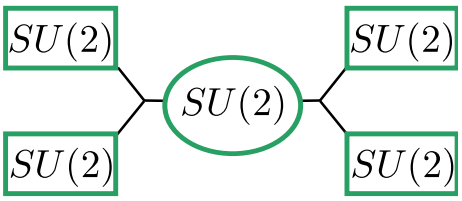


Figure 2.9: The quiver diagram corresponding to the 4-punctured sphere UV-curve. It can be easily read from considering the degenerate limit of the curve.

Note that any punctured Riemann surface can be constructed by gluing pair of pants. The basic building block is the three-punctured sphere. It has a trifundamental hypermul-

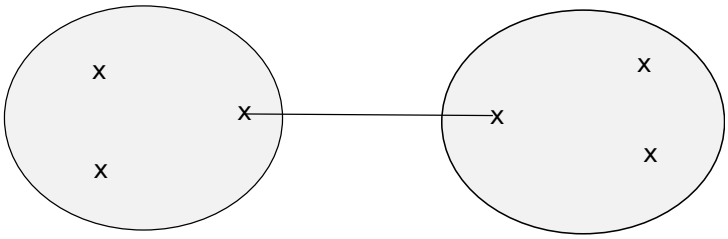


Figure 2.10: Obtaining a new curve from gluing pair of pants

tiplet which transform as $(2,2,2)$ under the flavor symmetry $SU(2)_1 \times SU(2)_2 \times SU(2)_3$. Connecting two punctures corresponds to gauging the flavor symmetry. S-dualities can be

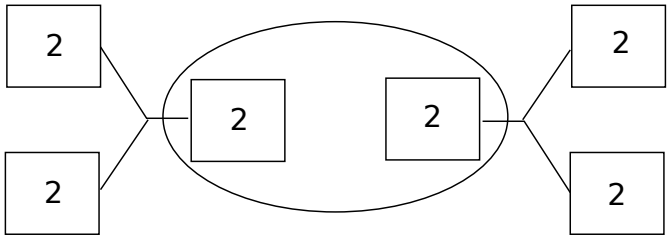


Figure 2.11: Gauging the flavor symmetry group can be thought of as gluing the pair of pants

understood as different choices of the degeneration limit coming from the single Riemann surface $\mathcal{C}$.

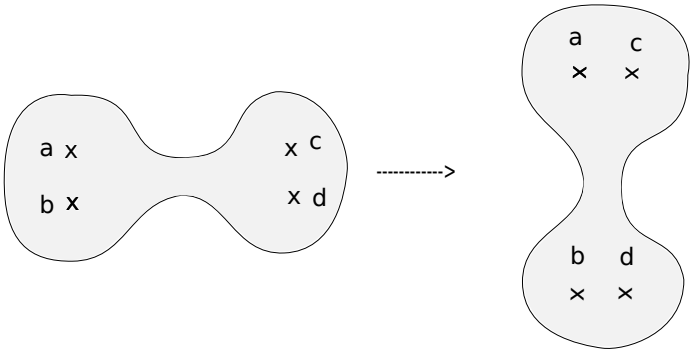


Figure 2.12: S-duality on the $SU(2)$ theory with 4 hypermultiplets. It can be understood as taking different degenerate limit. Since the both curves on the left and right are the same but rotated, the effective 4-dimensional theory has to be the same.

Since the S-duality can be concisely understood as an operation on the Riemann surface, it is natural to expect that S-duality is reflected in the 4d/2d correspondence as well. In the 2d side, the partition function of 4d gauge theory is given by a correlation function on 2d theory. Indeed, the S-duality of the theory on the 4-punctured sphere is reflected as the channel duality of the correlation function. The theory corresponding to the one-punctured torus is $\mathcal{N} = 2^*$ theory. When the mass of adjoint hypermultiplet is set to zero, it is nothing but $\mathcal{N} = 4$ theory. The S-duality can be thought as a modular invariance of the 2d CFT. In practice, it is very hard to evaluate the correlation function of 2d theory exactly. Therefore the direct check of S-duality from 4d/2d correspondence is still difficult, but we

can argue the duality on a more general ground.

Chapter 3

ABCDEFG of Instanton counting

In this chapter, we study methods of instanton counting to solve 4-dimensional gauge theories with $\mathcal{N} = 2$ supersymmetry for various gauge groups and matter fields. Especially in this chapter

- we find a uniform expression of 1-instanton partition function for arbitrary gauge groups in terms of root lattice [3],
- we find renormalization scheme dependence of $\mathcal{N} = 2$ SCFT and interpret it geometrically in terms of a map between two Gaiotto curves [1],
- we derive the contour integral formula for the half-hypermultiplets [1, 2].

3.1 Nekrasov's solution to $\mathcal{N} = 2$ gauge theory

Nekrasov partition function was introduced in [28], as a culmination of a long series of works, e.g.,[50, 51, 52] on the instanton calculation of the non-perturbative effects in $\mathcal{N} = 2$ gauge theory.

At low energies the four-dimensional $\mathcal{N} = 2$ gauge theory is governed by the prepotential $\mathcal{F}_0$, which determines the metric on the Coulomb branch of the gauge theory. Classically, the metric on the Coulomb branch is flat and the prepotential

$$\mathcal{F}_0^{\text{clas}} = 2\pi i\, \tau_{UV}\, \vec{a} \cdot \vec{a}, \tag{3.1}$$

is proportional to the microscopic coupling constant τ_{UV}. At the quantum level the prepotential receives both one-loop and non-perturbative instanton corrections, which give

corrections to the metric on the Coulomb moduli space. The instanton corrections to the prepotential can be computed as equivariant integrals over the instanton moduli space [28]. Let us briefly sketch how this comes about. We will discuss in much more detail in section 3.2.

Instantons on $\mathbb{R}^4$ are solutions of the self-dual instanton equation

$$F_A^+ = 0. \tag{3.2}$$

The instanton moduli space $\mathcal{M}^G$ parametrizes these solutions up to gauge transformations that leave the fiber at infinity fixed. The components $\mathcal{M}_k^G$ of the instanton moduli space are labeled by the topological instanton number $k = 1/8\pi^2 \int F_A \wedge F_A$. The instanton corrections to the prepotential for the pure $\mathcal{N} = 2$ gauge theory are captured by the instanton partition function

$$Z^{\text{inst}} = \sum_k q^k \oint\limits_{\mathcal{M}_k^G} 1, \tag{3.3}$$

where $\oint 1$ formally computes the volume of the moduli space. The parameter q can be considered as a formal parameter which counts the number of instantons. Physically, it is identified with a power $q = \Lambda^{b_0}$ of the dynamically generated scale Λ, when the gauge theory is asymptotically free. The power b_0 is determined by the one-loop β-function. It is identified with an exponent $q = \exp(2\pi i \tau_{\text{UV}})$ of the microscopic coupling τ_{UV} when the beta-function of the gauge theory vanishes.

If we introduce hypermultiplets to the pure $\mathcal{N} = 2$ gauge theory, the instanton correction to the prepotential are instead determined by solutions of the monopole equations

$$\begin{aligned} F^+_{A,\mu\nu} + \frac{i}{2}\, \overline{q}_\alpha \Gamma_{\mu\nu}{}^\alpha{}_\beta q^\beta &= 0, \\ \sum_\mu \Gamma^\mu_{\dot\alpha\alpha} D_{A,\mu} q^\alpha &= 0. \end{aligned} \tag{3.4}$$

In these equations Γ^μ are the Clifford matrices and $\sum_\mu \Gamma^\mu D_{A,\mu}$ is the Dirac operator in the instanton background for the gauge field A. Although there are no positive chirality solutions to the Dirac equation, the vector space of negative chirality solutions is k-dimensional. Because this vector space depends on the gauge background A, it is useful

to view it as a k-dimensional vector bundle over the instanton moduli space $\mathcal{M}_k^G$. We will call this vector bundle $\mathcal{V}$. More precisely, since the solutions to the Dirac equations are naturally twisted by the half-canonical line bundle $\mathcal{L}$ over $\mathbb{R}^4$ we will denote it by $\mathcal{V} \otimes \mathcal{L}$.

Instanton corrections to the $\mathcal{N} = 2$ gauge theory, with N_f hypermultiplets in the fundamental representation of the gauge group, are computed by the instanton partition function

$$Z^{\text{inst}} = \sum_k q^k \oint_{\mathcal{M}_k^G} e(\mathcal{V} \otimes \mathcal{L} \otimes M), \tag{3.5}$$

which is the integral of the Euler class of the vector bundle $\mathcal{V} \otimes \mathcal{L}$ of solutions to the Dirac equation over the moduli space $\mathcal{M}_k^G$. The flavor vector space $M = \mathbb{C}^{N_f}$ encodes the number of hypermultiplets in the gauge theory.

A difficulty in the evaluation of the instanton partition functions (3.3) and (3.5) is that the instanton moduli space $\mathcal{M}_k^G$ both suffers from an UV and an IR non-compactness. Instantons can become arbitrary small, as well as move away to infinity in $\mathbb{R}^4$. The IR non-compactness can be solved by introducing the Ω-background, which refers to the action of the torus

$$\mathbf{T}^2_{\epsilon_1,\epsilon_2} = U(1)_{\epsilon_1} \times U(1)_{\epsilon_2} \tag{3.6}$$

on $\mathbb{R}^4 = \mathbb{C} \oplus \mathbb{C}$ by a rotation $(z_1, z_2) \mapsto (e^{i\epsilon_1} z_1, e^{i\epsilon_2} z_2)$ around the origin with parameters $\epsilon_1, \epsilon_2 \in \mathbb{C}$. If we localize the instanton partition function equivariantly with respect to the $\mathbf{T}^2_{\epsilon_1,\epsilon_2}$-action, only instantons at the fixed origin will contribute, so that we can ignore the instantons that run off to infinity. The UV non-compactness can be cured for gauge group $U(N)$ by turning on an FI parameter. For Sp and SO gauge groups it is shown in [53] how to evaluate the instanton integrals, while implicitly curing the UV non-compactness of the instanton moduli space. Note that this effectively means that we have introduced a renormalization scheme.

Apart from the torus $\mathbf{T}^2_{\epsilon_1,\epsilon_2}$ there are a few other groups that act on the instanton moduli space $\mathcal{M}_k^G$. Their actions can be understood best from the famous ADHM construction of the instanton moduli space [54]. This construction gives the moduli space as the quotient of the solutions of the ADHM equations by the so-called dual group G_k^D with Cartan torus $\mathbf{T}^k_{\phi_i}$

whose weights we will call ϕ_i. There is a also natural action of the Cartan torus $\mathbf{T}^N_{\vec{a}}$ of the framing group G on the ADHM solution space, whose weights are given by the Coulomb branch parameters $\vec{a}$. Last, if the theory contains hypermultiplets, there is furthermore an action of the Cartan $\mathbf{T}^{N_f}_{\vec{m}}$ of the flavor symmetry group acting on M, whose weights correspond to the masses $\vec{m}$ of the hypers.

In total, we want to compute the partition function equivariantly with respect to the torus

$$\mathbf{T} = \mathbf{T}^2_{\epsilon_1,\epsilon_2} \times \mathbf{T}^N_{\vec{a}} \times \mathbf{T}^k_{\phi_i} \times \mathbf{T}^{N_f}_{\vec{m}}, \tag{3.7}$$

which comes down to computing the equivariant character of the action of those four tori. This results in a rational function $\mathbf{z}^k(\phi_i, \vec{a}, \vec{m}, \epsilon_1, \epsilon_2)$ of the weights. From the construction of the Dirac bundle it is clear that $\mathbf{z}^k$ factorizes if there are multiple hypers. Finally, we need to take into account the ADHM quotient. This we do by integrating over the dual group G^D_k. In total the instanton partition function is given by the integral

$$Z^{\text{inst}}_k = \int \prod_i d\phi_i \, \mathbf{z}^k_{\text{gauge}}(\phi_i, \vec{a}, \epsilon_1, \epsilon_2) \, \mathbf{z}^k_{\text{matter}}(\phi_i, \vec{a}, \vec{m}, \epsilon_1, \epsilon_2), \tag{3.8}$$

where all the $\mathcal{N} = 2$ multiplets in the gauge theory give a separate contribution. The instanton partition function of in principle any $\mathcal{N} = 2$ gauge theory with a Lagrangian prescription can be computed in this way. We will derive the explicit expressions in section 3.2

The integrand of (3.8) will have poles on the real axis. To cure this we will introduce small positive imaginary parts for the equivariance parameters. At least for asymptotically free theories we can then convert (3.8) into a contour integral, so that the problem reduces to enumerating poles and evaluating their residues. For $U(N)$ theory the poles are labeled by N Young diagrams with in total k boxes [28, 55, 56]: one way of phrasing this is that the $U(N)$ instanton splits into N non-commutative $U(1)$ instantons.

For the $Sp(N)$ or $SO(N)$ theory it is not that simple to enumerate the poles of the contour integrals. Furthermore, not only the fixed points of the gauge multiplet are more complicated, but (in contrast to the $U(N)$ theory) also matter multiplets contribute additional poles. As an example, in appendix B we devise a technique to enumerate all the poles for

an $Sp(N)$ gauge multiplet. Each pole can still be expressed as a generalized diagram with signs, but the prescription is much more involved than in the $U(N)$ case.

The instanton partition function Z^{inst} in the Ω-background obviously depends on the equivariant parameters ϵ_1 and ϵ_2. In fact, it is rather easy to see that the series expansion of $\log(Z^{\text{inst}})$ starts out with a term proportional to $\frac{1}{\epsilon_1\epsilon_2}$, which is the regularized volume of the Ω-background. Even better, Z^{inst} has a series expansion[1]

$$Z^{\text{inst}} = \exp \mathcal{F}^{\text{inst}} = \exp\left(\sum_{g=0}^{\infty} \hbar^{2g-2} \mathcal{F}_g^{\text{inst}}(\beta)\right), \tag{3.9}$$

in terms of the parameter $\hbar^2 = -\epsilon_1\epsilon_2$ and $\beta = -\frac{\epsilon_1}{\epsilon_2}$. We call the exponent of the instanton partition function the instanton free energy $\mathcal{F}^{\text{inst}}$. As our notation suggests, we recover the non-perturbative instanton contribution to the prepotential $\mathcal{F}_0$ from the leading contribution of the exponent when $\hbar \to 0$. This has been showed in [23, 55, 60]. Let us emphasize that the prepotential $\mathcal{F}_0$ does not depend on the parameter β. The higher genus free energies $\mathcal{F}_{g\geqslant 1}(\beta)$ compute gravitational couplings to the $\mathcal{N} = 2$ gauge theory, and play an important role in, for example, (refined) topological string theory.

To recover the full prepotential, we need to add classical and 1-loop contributions to the instanton partition function. We call the complete partition function

$$Z^{\text{Nek}} = Z^{\text{clas}} Z^{1-\text{loop}} Z^{\text{inst}} \tag{3.10}$$

the *Nekrasov partition function*.

3.2 Instanton counting

In this section we discuss a method to derive the instanton counting formulae for general gauge groups and matters in various representations. [2]

[1] To find this expansion in merely even powers of $\hbar$ it is crucial to study the twisted kernel of the Dirac operator, in contrast to the kernel of the Dolbeault operator. This twist is ubiquitous in the theory of integrable systems. Mathematically, it has been emphasized in this setting in [57]. Physically, it corresponds to a mass shift $m \to m + \frac{\epsilon_1+\epsilon_2}{2}$. This mass shift was studied in several related contexts, see [58, 59]. An exception to the above expansion is the $U(N)$ theory which has a non-vanishing contribution $\frac{1}{\hbar}\mathcal{F}_{1/2}^{\text{inst}}(\beta)$.

[2] Additional explanations about the ADHM moduli space can for instance be found in [54, 61, 62], about instanton counting in the physics literature [63, 64, 28, 23, 30], and in the mathematics literature in [56, 65, 55, 66, 60, 67, 57], and about Sp/SO instanton counting in specific in [53, 68, 69, 70, 71].

3.2.1 ADHM construction

Let E be a rank N complex vector bundle on $\mathbb{R}^4$ with a connection A and a framing at infinity. The framing is an isomorphism of the fiber at infinity with $\mathbb{C}^N$. The ADHM construction studies the moduli space $\mathcal{M}_k$ of connections A on the bundle E that satisfy the self-dual instanton equation $F^+(A) = 0$, up to gauge transformations that are trivial at infinity. It turns out that this moduli space can be realized as a hyperkähler quotient of linear data.

$U(N)$ gauge group

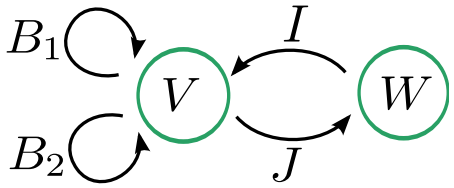


Figure 3.1: Quiver representation of the $U(N)$ ADHM quiver. The vector spaces V and W are k and N-dimensional, respectively, with a natural action of the dual group $U(k)$ and the framing group $U(N)$. The maps B_1, B_2, I and J are linear.

For the gauge group $G = U(N)$ the linear data consists of four linear maps

$$(B_1, B_2, I, J) \in \mathbf{X} = \mathrm{Hom}(V, V) \oplus \mathrm{Hom}(V, V) \oplus \mathrm{Hom}(W, V) \oplus \mathrm{Hom}(V, W), \tag{3.11}$$

where V and W are two complex vector spaces of dimension k and N, respectively. This linear data is summarized in an ADHM quiver diagram in Figure 3.1. The vector space W is isomorphic to the fiber of E (which in our case is of rank N). It is best thought of as the fiber at infinity, as there is a natural action of the framing group $U(N)$ on it, which physically can be thought of as the large gauge transformations at infinity. The tensor product of the vector space V with the half canonical bundle $K_{\mathbb{C}^2}^{1/2}$ on $\mathbb{C}^2 \cong \mathbb{R}^4$, on the other hand, can be identified with the space of normalizable solutions to the Dirac equation in the background of the instanton gauge field A. Since the instanton number $k = 1/8\pi^2 \int F_A \wedge F_A$ is given by the second Chern class of E, it follows from index theorems that this space has dimension k, as we advertised above. In particular it carries in a natural way the action of the dual group $U(k)$. More algebraically, the vector space V itself is isomorphic to the

cohomology group $H^1(E)$.

The framing group $U(N)$ and the dual group $U(k)$ thus act naturally on the linear ADHM data. Setting the three real moment maps

$$\mu_{\mathbb{R}} = [B_1, B_1^\dagger] + [B_2, B_2^\dagger] + II^\dagger - J^\dagger J \tag{3.12}$$

$$\mu_{\mathbb{C}} = [B_1, B_2] + IJ, \tag{3.13}$$

to zero gives the so-called *ADHM equations*. The ADHM construction identifies the instanton moduli space $\mathcal{M}_k^{U(N)}$ with the hyperkähler quotient of the solutions $\mathbf{X}$ to those equations by the dual group,

$$\mathcal{M}_k^{U(k)} = \mathbf{X}//U(k) \equiv \mu^{-1}(0)/U(k). \tag{3.14}$$

From a physical perspective the ADHM construction can be most natural understood using D-branes. We can engineer the moduli space of k instantons in the four-dimensional $U(N)$ theory by putting k D$(p-4)$-branes on top of N Dp-branes. The D$(p-4)$-branes appear as zero-dimensional instantons on the transverse four-dimensional manifold. The maps (B_1, B_2, I, J) can be understood as the zero-modes of D$(p-4)$−D$(p-4)$), Dp−D$(p-4)$ and D$(p-4)$−Dp open strings, respectively, and the ADHM equations are the D-term conditions. The ADHM quotient can thus be identified with the moduli space of the Higgs branch of the $U(k)$ gauge theory on the D$(p-4)$-branes.

Since the above quotient is highly singular due to small instantons, we change it by giving non-zero value to the Fayet-Illiopolous term ζ. This is equivalent to turning on NS 2-form field on the Dp-branes, and the resulting desingularized quotient can be interpreted as a moduli space of non-commutative instantons [72].

To perform our computations another, equivalent way of representing the ADHM construction will be useful. Let us introduce the spinor bundles $\mathcal{S}^\pm$ of positive and negative chirality on $\mathbb{R}^4$, and for brevity denote the half canonical bundle by $\mathcal{L} = K_{\mathbb{C}^2}^{1/2}$. Consider

the sequence

$$V \otimes L^{-1} \xrightarrow{\sigma} \begin{matrix} V \otimes S^- \\ \oplus \\ W \end{matrix} \xrightarrow{\tau} V \otimes L\,, \tag{3.15}$$

where S^- and L are the fibers of the bundles $\mathcal{S}^-$ and $\mathcal{L}$, respectively. Although these can all be trivialized, they are non-trivial equivariantly. We thus need to keep track of them for later. The mappings σ and τ are defined by

$$\sigma = \begin{pmatrix} z_1 - B_1 \\ z_2 - B_2 \\ J \end{pmatrix}, \qquad \tau = \begin{pmatrix} -z_2 + B_2, & z_1 - B_1, & I \end{pmatrix}, \tag{3.16}$$

where (z_1, z_2) are coordinates on $\mathbb{C}^2$. From the ADHM equations it follows that $\tau \circ \sigma = 0$, so that the sequence (3.15) is a chain complex. Since σ is injective and τ surjective, it is a so-called monad.

Notice that the vector space $V \otimes L^{-1}$ at the first position of the sequence (3.15) fixes the vector spaces at the remainder of the sequence. The fields B_1 and B_2 are coordinates on $\mathbb{C}^2$ and thus map $V \otimes L^{-1} \to V \otimes S^-$. The fields I and J are the two scalar components of an $\mathcal{N} = 2$ hypermultiplet, that properly speaking transform as sections of the line bundle $\mathcal{L}$.

To recover the vector bundle E, we vary the cohomology space $(\operatorname{Ker} \text{ø})/(\operatorname{Im} \text{œ})$ over $\mathbb{C}^2$, which gives indeed a vector bundle whose fiber at infinity is equal to W. One can also show that the curvature of this bundle is self-dual and that it has instanton number k. Even better, every solution of the self-dual instanton equations can be found in this way.

We are now ready to construct the main tool in our computation. This is the *universal bundle* $\mathcal{E}$ over the instanton moduli space $\mathcal{M}_k^{U(N)} \times \mathbb{R}^4$. The universal bundle is obtained by varying the ADHM-parameters of the maps in the complex (3.15). It has the property that

$$\mathcal{E}_{A,z} = E_z, \tag{3.17}$$

i.e., its fiber over an element $A \in \mathcal{M}_k^{U(N)}$ is the total space of the bundle E with connec-

tion A. Remember that the bundle E has fiber W at infinity in $\mathbb{R}^4$ and that the vector space V of solutions to the Dirac equations is related to its first cohomology $H^1(E)$. The vector spaces V and W can be extended to bundles $\mathcal{V}$ and $\mathcal{W}$ over the instanton moduli space. We can then easily compute the Chern character of the universal bundle $\mathcal{E}$ from its defining complex (3.15) as

$$\mathrm{Ch}(\mathcal{E}) = \mathrm{Ch}(\mathcal{W}) + \mathrm{Ch}(\mathcal{V})\left(\mathrm{Ch}(\mathcal{S}^-) - \mathrm{Ch}(\mathcal{L}) - \mathrm{Ch}(\mathcal{L}^{-1})\right). \tag{3.18}$$

SO/Sp **gauge groups**

The construction for $SO(N)$ and $Sp(N)$ gauge groups is very similar. We define $Sp(N)$ to be the special unitary transformations on $\mathbb{C}^{2N}$ that preserve its symplectic structure Φ_s, and $SO(N)$ the special unitary transformations on $\mathbb{C}^N$ that preserve its real structure Φ_r.

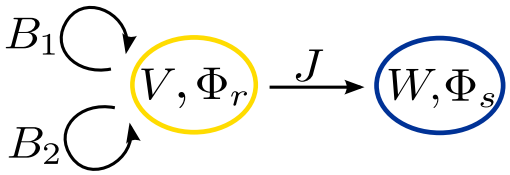


Figure 3.2: Quiver representation of the $Sp(N)$ ADHM quiver. The vector spaces V and W are k and $2N$-dimensional, respectively. V has a real structure Φ_r and a natural action of the dual group $SO(k)$, whereas W has a symplectic structure Φ_s and a natural action of the framing group $Sp(N)$. The maps B_1, B_2 and J are linear.

For $Sp(N)$ the linear data that is needed to define the ADHM complex consists of

$$(B_1, B_2, J) \in \mathbb{Y} = \mathrm{Hom}(V,V) \oplus \mathrm{Hom}(V,V) \oplus \mathrm{Hom}(V,W), \tag{3.19}$$

where V and W are a complex k and $2N$-dimensional vector space, resp., together with a real structure Φ_r on V and a symplectic structure Φ_s on W. This is illustrated as a quiver diagram in Figure 3.2. The dual group is given by $O(k)$, so that the moduli space of $Sp(N)$ instantons is given by

$$\mathcal{M}_k^{Sp(N)} = \{(B_1, B_2, J) \,|\, \Phi_r B_1, \Phi_r B_2 \in S^2 V^*,\ \Phi_r[B_1, B_2] - J^* \Phi_s J = 0\}/O(k). \tag{3.20}$$

For $SO(N)$ we just need to replace V and W by a complex $2k$ and N-dimensional vector space, resp., as well as change the role of symplectic structure and the real structure. This

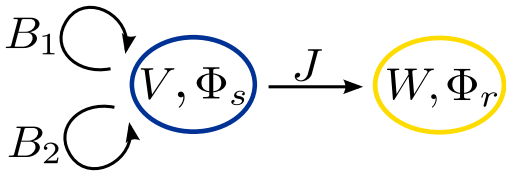


Figure 3.3: Quiver representation of the $SO(N)$ ADHM quiver. The vector spaces V and W are $2k$ and N-dimensional, respectively. V has a symplectic structure Φ_s and a natural action of the dual group $Sp(k)$, whereas W has a real structure Φ_r and a natural action of the framing group $SO(N)$. The maps B_1, B_2 and J are linear.

is illustrated as a quiver diagram in Figure 3.3. The dual group is given by $Sp(k)$, so that the moduli space of $SO(N)$ instantons is given by

$$\mathcal{M}_k^{SO(N)} = \{(B_1, B_2, J) \,|\, \Phi_s B_1, \Phi_s B \in \wedge^2 V^*,\ \Phi_s[B_1, B_2] - J^* \Phi_r J = 0\}/Sp(k). \tag{3.21}$$

A subtle issue for the above moduli spaces is that there is no appropriate Gieseker desingularization which resolves the singularity due to the zero-sized instantons (as in the case of $U(N)$). One way to understand this is by considering the string theory embedding. The above ADHM constructions can be obtained by considering Dp-D$(p-4)$ system and also adding an $O^{\pm}p$ plane on the top of the Dp branes. In the case of $U(N)$, the non-commutativity parameter that we introduce is coming from the NS 2-form field on the Dp-brane. But, the orientifold makes it impossible to turn on the background NS 2-form field. So we cannot resolve the singularity in the same way. An alternative way to resolve the singularity was studied by [73], but a physical understanding of this procedure is still lacking. Nevertheless, we will see that the equivariant volume of the moduli space can be obtained without explicitly resolving the singularity [53]. This formula is verified mathematically using Kirwan's formula of the equivariant volume of the symplectic quotient [69].

We can then again represent any $Sp(N)$ instanton solution E as the cohomology bundle of the sequence

$$V \otimes L^{-1} \xrightarrow{\sigma} \begin{matrix} V \otimes S^- \\ \oplus \\ W \end{matrix} \xrightarrow{\sigma^* \beta^*} V^* \otimes L, \tag{3.22}$$

where the mappings σ and β are defined by

$$\sigma = \begin{pmatrix} z_1 - B_1 \\ z_2 - B_2 \\ J \end{pmatrix}, \qquad \beta = \begin{pmatrix} 0 & \Phi_r & 0 \\ -\Phi_r & 0 & 0 \\ 0 & 0 & \Phi_s \end{pmatrix}. \tag{3.23}$$

The ADHM equations ensure that $\sigma^*\beta^*\sigma = 0$, and the sequence (3.22) is another monad. Analogously to the $U(N)$ example, when varying the ADHM-parameters in the complex (3.22) we find the universal bundle $\mathcal{E}_{Sp(N)}$. V is the k-dimensional solution space of the $Sp(N)$ Dirac operator on $\mathbb{C}^2$, which carries a real structure, and W is the fiber of E at infinity, and hence carries a symplectic structure.

Similarly, any $SO(N)$ instanton solution E can be represented as the cohomology bundle of the complex (3.22) as well, once we exchange Φ_r with Φ_s in the definition of the map β. The resulting complex is a short exact sequence, since according to the ADHM equations $\sigma^*\beta^*\sigma = 0$. By varying the ADHM-parameters we find the $SO(N)$ universal bundle $\mathcal{E}_{SO(N)}$. Note that V is the $2k$-dimensional solution space of the $SO(N)$ Dirac operator on $\mathbb{C}^2$, which carries a symplectic structure, whereas W is the fiber of E at infinity, and hence carries a real structure.

3.2.2 Ω-background and equivariant integration

Let us start with a supersymmetric $\mathcal{N} = 2$ gauge theory without matter. This theory can be topologically twisted (using the so-called Donaldson twist), so that its BPS equation is the instanton equation $F_A^+ = 0$. The instanton partition function is given by an integral over the moduli space $\mathcal{M}_k^{\text{inst}}$,

$$Z^{\text{inst}} = \sum_k q^k \oint\!\!\!\!\!\!\int_{\mathcal{M}_k^{\text{inst}}} 1\,, \tag{3.24}$$

where $\oint\!\!\!\!\!\!\int 1$ indicates the formal volume of the hyperkähler quotient.

If we add matter multiplets to the gauge theory, say a single $\mathcal{N} = 2$ hypermultiplet, the

BPS equations turn into the monopole equations

$$F^+_{A,\mu\nu} + \frac{i}{2}\,\overline{q}_\alpha \Gamma_{\mu\nu}{}^\alpha{}_\beta q^\beta = 0, \qquad \sum_\mu \Gamma^\mu_{\dot\alpha\alpha} D_{A,\mu} q^\alpha = 0. \tag{3.25}$$

Here, Γ^μ are the Clifford matrices and $\sum_\mu \Gamma^\mu D_{A,\mu}$ is the Dirac operator in the instanton background for the connection A. Furthermore, q^α is the lowest component of the twisted hypermultiplet. The representation of the connection A in the connection D_A is determined by the representation of the gauge group that the hypermultiplet is in. As is argued in section 3.4 of [68], it is possible to deform the action in a Q-exact way such that the first equation gets an extra factor

$$F^+_{A,\mu\nu} + \frac{i}{2t}\,\overline{q}_\alpha \Gamma_{\mu\nu}{}^\alpha{}_\beta q^\beta = 0,,$$

for an arbitrary value of t. Taking the limit $t \to \infty$ reduces this BPS equation to the selfdual instanton equation. Given an instanton solution A, the remainder of the action is forced to localize onto solutions of the Dirac equation in the background of A. The kernel of the Dirac operator thus forms a fiber over the instanton moduli space $\mathcal{M}^{\text{inst}}$.

Let us start out by adding a single hypermultiplet in the fundamental representation of the gauge group. Remember that the corresponding kernel was already encoded in the vector bundle $\mathcal{V} \otimes \mathcal{L}$ in the original ADHM construction. N_f hypers are simply described by the tensor product of N_f vector bundles $\mathcal{V}$. Each individual factor then carries the usual action of the dual group. Moreover, there is now also a natural action of the flavor symmetry group. By general arguments (see, for example, [74]) this partition function then localizes to the integral[3]

$$Z^{\text{inst}} = \sum_k q^k \oint_{\mathcal{M}^{\text{inst}}_k} e(\mathcal{V} \otimes \mathcal{L} \otimes M) \tag{3.26}$$

of the Euler class of the vector bundle $\mathcal{V} \otimes \mathcal{L}$ of solutions to the Dirac equation over the

[3]Here (and elsewhere in the paper) $\mathcal{L}$ is really just the fiber of the half-canonical bundle $\mathcal{L}$ at the origin of $\mathbb{R}^4$. More formally, we take the cup product of the bundle $\mathcal{V} \otimes \mathcal{L} \otimes M$ over $\mathcal{M}^{\text{inst}}_k \times \mathbb{R}^4$ with the push-forward i_* of the fundamental class $[\mathcal{M}^{\text{inst}}_k]$, where i embeds the moduli space $\mathcal{M}^{\text{inst}}_k$ in the product $\mathcal{M}^{\text{inst}}_k \times \mathbb{R}^4$ at the origin of $\mathbb{R}^4$.

moduli space $\mathcal{M}^{\text{inst}}$. The vector space $M = \mathbb{C}^{N_f}$ encodes the number of flavors.

Let us now discuss how to compute this partition function. First, however, note that (3.26) diverges. We will implicitly take care of the UV divergence in the next steps by computing a holomorphic character [53], but we also need to deal with IR divergences: the instanton moduli space has flat directions where the instantons move off to infinity. One way to regularize the instanton partition function is to introduce the so-called Ω-background, where we use equivariant integration with respect to the torus

$$\mathbf{T}^2_{\epsilon_1,\epsilon_2} = U(1)_{\epsilon_1} \times U(1)_{\epsilon_2}, \tag{3.27}$$

that acts on $\mathbb{R}^4 = \mathbb{C} \oplus \mathbb{C}$ by a rotation $(z_1, z_2) \to (e^{i\epsilon_1} z_1, e^{i\epsilon_2} z_2)$ around the origin. This forces the instantons to be localized at the origin of $\mathbb{R}^4$. The resulting Ω-background is denoted by $\mathbb{R}^4_{\epsilon_1,\epsilon_2}$. The partition function in the Ω-background is defined by equivariantly integrating with respect to the $\mathbf{T}^2_{\epsilon_1,\epsilon_2}$–action.

We have already introduced the other components of the equivariance group: The torus $\mathbf{T}^N_a$ of the gauge group G acting on the fiber W with weights a_l, the torus $\mathbf{T}^k_{\phi_i}$ of the dual group acting on V with weights ϕ_i, and lastly the torus $\mathbf{T}^{N_f}_m$ of the flavor symmetry group acting on the flavor vector space M with weights m_j. In total, we perform the equivariant integration with respect to the torus

$$\mathbf{T} = \mathbf{T}^2_{\epsilon_1,\epsilon_2} \times \mathbf{T}^N_a \times \mathbf{T}^k_\phi \times \mathbf{T}^{N_f}_m \,. \tag{3.28}$$

The instanton partition function in the Ω-background is thus defined as the equivariant integral

$$Z(a, m, \epsilon_1, \epsilon_2) = \sum_k q^k \oint_{\mathcal{M}_k} e_{\mathbf{T}}(\mathcal{V} \otimes \mathcal{L} \otimes M)\,, \tag{3.29}$$

where $e_{\mathbf{T}}$ is the equivariant Euler class with respect to the torus $\mathbf{T}$. We will evaluate (3.29) in two steps, using the fact that $\mathcal{M}_k$ is given by the solutions $\mu^{-1}(0)$ to the ADHM equations quotiented by the dual group G^D_k. We will thus first perform the equivariant integral over $\mu^{-1}(0)$, and then take care of the quotient by integrating out G^D_k, which gives a multiple integral over ϕ_i.

To perform the first part of the above integral, we apply the famous equivariant localization theorem, which tells us that the integral only depends on the fixed points of the equivariant group and its weights at those points. This then leads to a rational function in all the weights.

More precisely, suppose that the action of the element $t \in \mathfrak{t}$ on the integration space $\mathcal{M}$ (which is represented by a vector field V_t) has a discrete number of fixed points f. Then the equivariant localization theorem says that

$$\int_{\mathcal{M}} \alpha = \sum_f \frac{\iota^* \alpha}{\prod_k w_k[t](f)}, \tag{3.30}$$

where ι embeds the fixed point locus in $\mathcal{M}$ and where $w_k[t](f)$ are the weights of the action of the vector field V_t on the tangent space to the fixed point $f \in \mathcal{M}$. If we apply the localization theorem to the integral (3.29), the denominator of the resulting expression contains a product of weights of the torus action on the tangent bundle to the instanton moduli space. Its numerator is given by another product of weight of the torus action on the bundle $\mathcal{V} \otimes \mathcal{L} \otimes M$ of Dirac zero modes.

Let us start with computing the weights in the numerator, and for convenience restrict the matter content to a single hypermultiplet in the fundamental representation of the gauge group. Since the bundle in the numerator is the kernel of the Dirac operator, we can equally well obtain these weights from the equivariant index $\mathrm{Ind}_{\mathbf{T}} = \sum_k n_k e^{i w_k}$ of the Dirac operator. For the purpose of (3.30), the sum over weights can be translated into a product by the formula

$$\sum_k n_k e^{i w_k} \to \prod_i (w_k)^{n_k}. \tag{3.31}$$

To compute the equivariant index $\mathrm{Ind}_{\mathbf{T}}$ of the Dirac operator coupled to the instanton background, we make use of the equivariant version of Atiyah-Singer index theorem. It is given by

$$\mathrm{Ind}_{\mathbf{T}} = \int_{\mathbb{C}^2} \mathrm{Ch}_{\mathbf{T}}(\mathcal{E} \otimes \mathcal{L}) \, \mathrm{Td}_{\mathbf{T}}(\mathbb{C}^2) = \frac{\mathrm{Ch}_{\mathbf{T}}(\mathcal{E} \otimes \mathcal{L})|_{z_1 = z_2 = 0}}{(e^{i\epsilon_1} - 1)(e^{i\epsilon_2} - 1)}, \tag{3.32}$$

where $\mathcal{E}$ is the universal bundle over the instanton moduli space $\mathcal{M}_k$ that we constructed

in the previous section. Remember that the fiber of $\mathcal{E}$ over an element A in the instanton moduli space is given by the total space of the instanton bundle E with connection A. The second equality is obtained by applying the equivariant localization theorem and using the equivariant Todd class of $\mathbb{C}^2$ equals

$$\mathrm{Td}_{\mathbf{T}}(\mathbb{C}^2) = \frac{\epsilon_1\epsilon_2}{(e^{i\epsilon_1}-1)(e^{i\epsilon_2}-1)}, \tag{3.33}$$

where the weights of the action of $\mathbf{T}_{\epsilon_1,\epsilon_2}$ on $\mathbb{C}^2$ are ϵ_1 and ϵ_2.

The purpose of all of this was to reduce everything to the equivariant Chern character of the universal bundle $\mathcal{E}$, for which have found the simple expression (3.18) in terms of the Chern characters of the vector bundles $\mathcal{W}$, $\mathcal{V}$, $\mathcal{L}$ and $\mathcal{S}$. We can easily obtain the weights of the torus $\mathbf{T}$ on these bundles, so that we can compute the contribution of a fundamental hypermultiplet. We will write down explicit expressions in a moment, but let us first explain how to obtain the weights for other representations.

If we instead wish to extract the weights for an anti-fundamental hyper we just need replace the equivariant character for universal bundle $\mathcal{E}$ by its complex conjugate $\mathcal{E}^*$. Other representations that are tensor products of fundamentals and anti-fundamentals (or symmetric or antisymmetric combinations thereof) can be obtained similarly. For instance, the adjoint representation for a classical gauge group can be expressed as some product of the fundamental and the anti-fundamental representation. This product is the tensor product for $U(N)$, the anti-symmetric product for $SO(N)$ and the symmetric product for $Sp(N)$. Note that in those cases we also obtain the representations of the dual groups. We thus obtain the weights for an adjoint hypermultiplet by computing the character of the appropriate product of the universal bundle and its complex conjugate. The weights for the gauge multiplet are the same as for an adjoint hypermultiplet, but end up in the denominator of the contour integral instead of the numerator. This is consistent with the localization formula (3.29), as the tangent space to the instanton moduli space can be expressed as the same product of the universal bundle and its dual.

Once we obtain the index, we can extract the equivariant weights from it by using the rule (3.31). Finally, we need to integrate out the dual group G_k^D. This leads to a multiple integral over $d\phi_i$ along the real axis. We will absorb factors appearing from this integration such as the Vandermonde determinant of the Haar measure and the volume of the dual

group into the contribution of the gauge multiplet $\mathbf{z}^k_{\text{gauge}}$. The resulting integrand actually has poles on the real axis, which we cure by giving small imaginary parts to the equivariance parameters. We will describe this in more detail once we turn to the actual evaluation of such integrals. Since the integrand obtained is a rational function in the parameters of G^D_k, we can convert the integral into a contour integral around the poles of the integrand. In total the equivariant integral (3.29) over the moduli space thus reduces to

$$Z_k(a, m, \epsilon_1, \epsilon_2) = \oint \prod_i d\phi_i \prod_R \mathbf{z}^k_R(\phi_i, a, m, \epsilon_1, \epsilon_2), \tag{3.34}$$

where $\mathbf{z}^k_R$ are the integrands that represent the matter content of the gauge theory and where the ϕ_i's parametrize the dual group. We will discuss which poles (3.34) is integrated around shortly.

Equivariant index for $U(N)$ theories

Let us see how this works out explicitly for gauge group $U(N)$. We first have a look at the weights of the torus $\mathbf{T}^2_{\epsilon_1,\epsilon_2}$ at the fibers of the half-canonical bundle $\mathcal{L}$ and the spinor bundles $\mathcal{S}^\pm$ at the origin of $\mathbb{R}^4$. Remember that the torus $\mathbf{T}^2_{\epsilon_1,\epsilon_2}$ acts on the coordinates z_1 and z_2 with weights ϵ_1 and ϵ_2, respectively. It thus acts on local sections s of the half-canonical bundle as

$$s \in \mathcal{L}: \qquad s \mapsto e^{i\epsilon_+} s\,,$$

with $\epsilon_\pm = \frac{\epsilon_1 \pm \epsilon_2}{2}$. Local sections of the four-dimensional spinor bundles $\mathcal{S}^\pm$ can be written in terms of those of the two-dimensional spinor bundles on $\mathbb{R}^2$. Since the weights of the torus $\mathbf{T}^1_{\epsilon_j}$ on the local sections of the two spinor bundles on $\mathbb{R}^2$ are $\pm\frac{\epsilon_j}{2}$, the torus $\mathbf{T}^2_{\epsilon_1,\epsilon_2}$ acts on local sections $\psi_\pm$ of the four-dimensional spinor bundle as

$$\psi_\pm \in \mathcal{S}^\pm: \qquad \psi_\pm \mapsto \text{diag}(e^{i\epsilon_\pm}, e^{-i\epsilon_\pm})\, \psi_\pm\,.$$

Let us continue with the weights of the equivariant torus $\mathbf{T}^N_a \times \mathbf{T}^k_\phi$. The equivariant

torus then acts on the linear ADHM data as

$$\begin{aligned} v \in V: \qquad & v \mapsto \text{diag}\,(e^{i\phi_1}, \cdots, e^{i\phi_k})\, v\,, \\ w \in W: \qquad & w \mapsto \text{diag}\,(e^{ia_1}, \cdots, e^{ia_N})\, w\,. \end{aligned}$$

Combining all weights and using the formula (3.18), we find that the equivariant Chern character of the universal bundle $\mathcal{E}_{U(N)}$ is given by

$$\text{Ch}_{\mathbf{T}}(\mathcal{E}_{U(N)})|_{z_1=z_2=0} \;=\; \sum_{l=1}^{n} e^{ia_l} - (e^{i\epsilon_1}-1)(e^{i\epsilon_2}-1)\sum_{i=1}^{k} e^{i\phi_i - i\epsilon_+}. \tag{3.35}$$

Using the index formula (3.32) we have now computed the contribution for a fundamental massless hypermultiplet. As we explained before we can easily generalize this to other representations. In particular, we can give the hypermultiplet a mass by introducing a weight m for the flavor torus $\mathbf{T}^1_m$. This will act on the linear ADHM data as

$$\begin{aligned} v \in V: \qquad & v \mapsto \text{diag}\,(e^{im}, \cdots, e^{im})\, v\,, \\ w \in W: \qquad & w \mapsto \text{diag}\,(e^{im}, \cdots, e^{im})\, w\,. \end{aligned}$$

For gauge group $U(N)$ the poles of the resulting contour integral (3.34) can be labeled by a set of colored Young diagrams $\mathbf{Y} = (Y_1, Y_2, \cdots Y_N)$ [28, 55, 56]. Therefore, the partition function can be written as

$$Z(a, m, \epsilon_1, \epsilon_2) = \sum_{\mathbf{Y}} q^{|\mathbf{Y}|} \prod_{R} \mathbf{z}_{R,|\mathbf{Y}|}(\mathbf{Y}; a, m, \epsilon_1, \epsilon_2)\,. \tag{3.36}$$

When the gauge group is a product of M factors, the instanton partition function can be written as a sum over M colored Young diagrams $\mathbf{Y}$. For SO/Sp gauge groups, we will see that the structure of the contour integral is similar. However, the poles are no longer labeled by a simple set of colored Young diagrams.

Equivariant index for SO/Sp gauge theories

For $Sp(N)$ the weights of the equivariant torus action on the vector spaces V and W are given by

$$\begin{aligned} v \in V: \qquad v &\mapsto \text{diag}\left(e^{i\phi_1}, \cdots, e^{i\phi_n}, (1), e^{-i\phi_1}, \cdots, e^{-i\phi_n}\right) v \\ w \in W: \qquad w &\mapsto \text{diag}\left(e^{ia_1}, \cdots, e^{ia_N}, e^{-ia_1}, \cdots, e^{-ia_N}\right) w\,, \end{aligned}$$

where $k = 2n + \chi$, with $n = [k/2]$ and $\chi \equiv k \pmod 2$. The (1) is inserted when $\chi = 1$ and omitted when $\chi = 0$. The equivariant character of the universal bundle is thus given by

$$\begin{aligned} \text{Ch}_{\mathbf{T}}(\mathcal{E}_{Sp})|_{z_1=z_2=0} \;=\; & \sum_{l=1}^{N} \left(e^{ia_l} + e^{-ia_l}\right) \\ & - (e^{i\epsilon_1} - 1)(e^{i\epsilon_2} - 1) \left(\sum_{i=1}^{n} \left(e^{i\phi_i - i\epsilon_+} + e^{-i\phi_i - i\epsilon_+}\right) + \chi\, e^{-i\epsilon_+} \right). \end{aligned} \tag{3.37}$$

For $SO(N)$ the weights are given by

$$\begin{aligned} v \in V: \qquad v &\mapsto \text{diag}\left(e^{i\phi_1}, \cdots, e^{i\phi_k}, e^{-i\phi_1}, \cdots, e^{-i\phi_k}\right) v \\ w \in W: \qquad w &\mapsto \text{diag}\left(e^{ia_1}, \cdots, e^{ia_n}, (1), e^{-ia_1}, \cdots, e^{-ia_n}\right) w\,, \end{aligned}$$

where $N = 2n + \chi$, such that $n = [N/2]$ and $\chi \equiv N \pmod 2$. Again, (1) is inserted when $\chi = 1$ and omitted when $\chi = 0$. The equivariant character of the universal bundle is therefore equal to

$$\begin{aligned} \text{Ch}_{\mathbf{T}}(\mathcal{E}_{SO})|_{z_1=z_2=0} \;=\; & \sum_{l=1}^{n} \left(e^{ia_l} + e^{-ia_l}\right) + \chi \\ & - (e^{i\epsilon_1} - 1)(e^{i\epsilon_2} - 1) \sum_{i=1}^{k} \left(e^{i\phi_i - i\epsilon_+} + e^{-i\phi_i - i\epsilon_+}\right). \end{aligned} \tag{3.38}$$

Building from these expressions we can obtain the instanton partition functions of quiver gauge theories containing matter fields in various representations.

3.2.3 Contour integrals for various matter fields with different gauge groups

Let us collect various contour integrands for the instanton counting.

Fundamental of $U(N)$

The equivariant index for a fundamental $U(N)$ hypermultiplet of mass m is given by

$$\begin{aligned}\mathrm{Ind}_{\mathbf{T}} &= \int_{\mathbb{C}^2} \mathrm{Ch}_{\mathbf{T}}(\mathcal{E}_U \otimes \mathcal{L} \otimes M)\mathrm{Td}_{\mathbf{T}}(\mathbb{C}^2) \\ &= \frac{\sum_{l=1}^{N} e^{i(a_l+\epsilon_+ +m)}}{(e^{i\epsilon_1}-1)(e^{i\epsilon_2}-1)} - \sum_{i=1}^{k} e^{i(\phi_i+m)}. \end{aligned} \tag{3.39}$$

The first term corresponds to the 1-loop factor. We read the matter contribution to the contour integral as

$$\mathbf{z}_k^N = \prod_{i=1}^{k}(\phi_i + m). \tag{3.40}$$

The anti-fundamental matter can be obtained similarly:

$$\begin{aligned}\mathrm{Ind}_{\mathbf{T}} &= \int_{\mathbb{C}^2} \mathrm{Ch}_{\mathbf{T}}(\mathcal{E}_U^* \otimes \mathcal{L} \otimes M)\mathrm{Td}_{\mathbf{T}}(\mathbb{C}^2) \\ &= \frac{\sum_{l=1}^{N} e^{-i(a_l-\epsilon_+ -m)}}{(e^{i\epsilon_1}-1)(e^{i\epsilon_2}-1)} - \sum_{i=1}^{k} e^{-i(\phi_i+\epsilon_+ -m)} \end{aligned} \tag{3.41}$$

so that

$$\mathbf{z}_k^{\tilde{N}} = \prod_{i=1}^{k}(\phi_i - m). \tag{3.42}$$

Note that $\mathbf{z}_k^N(\phi_i, -m, \epsilon_1, \epsilon_2) = \mathbf{z}_k^{\tilde{N}}(\phi, m, \epsilon_1, \epsilon_2)$.

Adjoint of $U(N)$

The equivariant index is

$$\begin{aligned}\mathrm{Ind}_{\mathbf{T}} &= \int_{\mathbb{C}^2} \mathrm{Ch}_{\mathbf{T}}(\mathcal{E} \otimes \mathcal{E}^* \otimes \mathcal{L} \otimes M)\mathrm{Td}_{\mathbf{T}}(\mathbb{C}^2) \\ &= \frac{\sum_{l,m=1}^{N} e^{i(a_l-a_m+\epsilon_+ +m)}}{(e^{i\epsilon_1}-1)(e^{i\epsilon_2}-1)} - \sum_{i,l=1}^{k,N} e^{i(\phi_i-a_l+m)} - \sum_{i,l=1}^{k,N} e^{-i(\phi_i-a_l-m)} \\ &+ (e^{i\epsilon_1}-1)(e^{i\epsilon_2}-1)\sum_{i,j=1}^{k} e^{i(\phi_i-\phi_j-\epsilon_+ +m)}, \end{aligned} \tag{3.43}$$

and the integrand is given by

$$\mathbf{z}_k^{adj} = \prod_{i,l=1}^{k,N}(\phi_i - a_l + m)\prod_{i,l=1}^{k,N}(\phi_i - a_l - m)\prod_{i,j=1}^{k}\frac{(\phi_{ij}+m-\epsilon_-)(\phi_{ij}+m+\epsilon_-)}{(\phi_{ij}+m+\epsilon_+)(\phi_{ij}+m-\epsilon_+)}, \tag{3.44}$$

where $\phi_{ij} = \phi_i - \phi_j$ and $\epsilon_\pm = \frac{\epsilon_1 \pm \epsilon_2}{2}$.

$U(N)$ gauge multiplet

In general, the equivariant index of a gauge multiplet is given by the character of the tangent space of the moduli space [30],

$$\mathrm{Ind}_{\mathbf{T}}(T_p\mathcal{M}) = -\mathrm{Ind}_{\mathbf{T}}(\mathcal{E}\otimes\mathcal{E}^*). \tag{3.45}$$

Therefore,

$$\begin{aligned}
\mathrm{Ind}_{\mathbf{T}} &= -\int_{\mathbb{C}^2}\mathrm{Ch}_{\mathbf{T}}(\mathcal{E}\otimes\mathcal{E}^*)\mathrm{Td}_{\mathbf{T}}(\mathbb{C}^2) \\
&= -\frac{\sum_{l,m=1}^{N,N} e^{i(a_l-a_m)}}{(e^{i\epsilon_1}-1)(e^{i\epsilon_2}-1)} + \sum_{i,m=1}^{k,N} e^{i(\phi_i-\epsilon_+-a_m)} + \sum_{j,l=1}^{k,N} e^{-i(\phi_j+\epsilon_+-a_l)} \\
&- (e^{i\epsilon_1}-1)(e^{i\epsilon_2}-1)\sum_{i,j=1}^{k,k} e^{i(\phi_i-\phi_j-\epsilon)}.
\end{aligned} \tag{3.46}$$

The resulting contour integral is given by

$$\begin{aligned}
Z_k &= \frac{1}{k!}\left(\frac{\epsilon}{\epsilon_1\epsilon_2}\right)^k \oint\left(\prod_{i=1}^{k}\frac{d\phi_i}{2\pi i}\right) \\
&\times \frac{1}{\prod_{i,m=1}^{k,N}(\phi_i - a_m)\prod_{j,l=1}^{k,N}(\phi_j - a_l + \epsilon_1 + \epsilon_2)}\prod_{1\leqslant i<j\leqslant k}\frac{\phi_{ij}^2(\phi_{ij}^2-\epsilon^2)}{(\phi_{ij}^2-\epsilon_1^2)(\phi_{ij}^2-\epsilon_2^2)}.
\end{aligned} \tag{3.47}$$

where $\phi_{ij} = \phi_i - \phi_j$. The poles of the contour integral can be classified in terms of N-colored Young diagrams with k boxes. In the case of $U(N)$ linear quiver theory, there is no additional pole besides coming from the gauge multiplet. Therefore, the instanton partition function for the $U(N)$ theories can be simply written in terms of sum over colored Young diagrams $\vec{Y} = (Y_1, \cdots, Y_N)$.

Bifundamental $(N_1, \bar{N}_2)$ **of** $U(N_1) \times U(N_2)$

The index is

$$\begin{aligned}
\mathrm{Ind}_{\mathbf{T}} &= \int_{\mathbb{C}^2} \mathrm{Ch}_{\mathbf{T}}(\mathcal{E}_1 \otimes \mathcal{E}_2^* \otimes \mathcal{L} \otimes M)\mathrm{Td}_{\mathbf{T}}(\mathbb{C}^2) \\
&= \frac{\sum_{l,m=1}^{N_1,N_2} e^{i(a_l - b_m + m + \epsilon_+)}}{(e^{i\epsilon_1}-1)(e^{i\epsilon_2}-1)} - \sum_{i,l=1}^{k_1,N_2} e^{i(\phi_{1,i} - b_m + m)} - \sum_{i,l=1}^{k_2,N_1} e^{-i(\phi_{2,j} - a_l - m)} \\
&+ (e^{i\epsilon_1}-1)(e^{i\epsilon_2}-1) \sum_{i,j=1}^{k_1,k_2} e^{i(\phi_{1,i} - \phi_{2,j} - \epsilon_+ + m)}.
\end{aligned} \tag{3.48}$$

The corresponding integral is

$$\begin{aligned}
\mathbf{z}_{k_1,k_2}^{(N_1,\bar{N}_2)} &= \prod_{i,l=1}^{k_1,N_2} (\phi_{1,i} - b_l + m) \prod_{i,l=1}^{k_2,N_1} (\phi_{2,i} - a_l - m) \\
&\times \prod_{i,j=1}^{k_1,k_2} \frac{(\phi_{1,i} - \phi_{2,j} + m - \epsilon_-)(\phi_{1,i} - \phi_{2,j} + m + \epsilon_-)}{(\phi_{1,i} - \phi_{2,j} + m + \epsilon_+)(\phi_{1,i} - \phi_{2,j} + m - \epsilon_+)}.
\end{aligned} \tag{3.49}$$

We can obtain N_2 fundamental matter contributions (3.40) with masses $m - b_l$ by setting $k_1 = k, k_2 = 0$. Also, by setting $\phi_{1,i} = \phi_{2,i} = \phi_i$ and $a_i = b_i$, we obtain adjoint matter (3.43).

Fundamental of $Sp(N)$

The equivariant index for a fund. $Sp(N)$ hypermultiplet of mass m is given by

$$\begin{aligned}
\mathrm{Ind}_{\mathbf{T}} = &\int_{\mathbb{C}^2} \mathrm{Ch}_{\mathbf{T}}(\mathcal{E}_{\mathrm{Sp}} \otimes \mathcal{L} \otimes M)\, \mathrm{Td}_{\mathbf{T}}(\mathbb{C}^2) \\
= &\frac{1}{(e^{i\epsilon_1}-1)(e^{i\epsilon_2}-1)} \sum_{l=1}^{N} \left(e^{ia_l + im + i\epsilon_+} + e^{-ia_l + im + i\epsilon_+} \right) \\
&- \sum_{i=1}^{n} \left(e^{i\phi_i + im} + e^{-i\phi_i + im} + \chi e^{im} \right),
\end{aligned} \tag{3.50}$$

where $k = 2n + \chi$ and $\epsilon_+ = \frac{\epsilon_1 + \epsilon_2}{2}$. Here, we tensored the universal bundle $\mathcal{E}_{\mathrm{Sp}}$ by the vector space $M \cong \mathbb{C}$ on whose elements the flavor symmetry $U(1)_m$ acts by $v \mapsto e^{im}v$. Since the first term in the index computes perturbative terms in the free energy, the contour

integrand for the instanton contribution to the free energy is

$$\mathbf{z}_k^N = m^\chi \prod_{i=1}^{n} (\phi_i + m)(\phi_i - m). \tag{3.51}$$

$Sp(N)$ gauge multiplet

The equivariant index for an $Sp(N)$ gauge multiplet is given by

$$\mathrm{Ind}_{\mathbf{T}} = \quad -\int_{\mathbb{C}^2} \mathrm{Ch}_{\mathbf{T}}(\mathrm{Sym}^2 \mathcal{E}_{Sp(N)})\, \mathrm{Td}_{\mathbf{T}}(\mathbb{C}^2) \tag{3.52}$$

The resulting contour integral is

$$\begin{aligned} Z_k &= \frac{(-1)^n}{2^n n!} \left(\frac{\epsilon}{\epsilon_1 \epsilon_2}\right)^n \left[\frac{-1}{2\epsilon_1\epsilon_2 P(\epsilon_+)}\right]^\chi \\ &\quad \times \oint \left(\prod_{i=1}^{n} \frac{d\phi_i}{2\pi i}\right) \frac{\Delta(0)\Delta(\epsilon)}{\Delta(\epsilon_1)\Delta(\epsilon_2)} \prod_{i=1}^{n} \frac{1}{(2\phi_i^2 - \epsilon_1^2)(2\phi_i^2 - \epsilon_2^2) P(\phi_i + \epsilon_+) P(\phi_i - \epsilon_+)} \end{aligned} \tag{3.53}$$

with $\epsilon = \epsilon_1 + \epsilon_2$ and

$$\begin{aligned} P(x) &= \prod_{l=1}^{N} (x^2 - a_l^2)\,, \\ \Delta(x) &= \left[\prod_{i=1}^{n} (\phi_i^2 - x^2)\right]^\chi \prod_{i<j} \left((\phi_i + \phi_j)^2 - x^2\right) \left((\phi_i - \phi_j)^2 - x^2\right)\,. \end{aligned}$$

We describe a way to enumerate the poles of the above contour integral in appendix B, using what we call generalized Young diagrams.[4]

Bifundamental of $Sp(N_1) \times Sp(N_2)$

The equivariant index of a bifund. $Sp(N_1) \times Sp(N_2)$ hyper with mass m is given by

$$\mathrm{Ind}_{\mathbf{T}} = \int_{\mathbb{C}^2} \mathrm{Ch}_{\mathbf{T}}(\mathcal{E}^1_{Sp} \otimes \mathcal{E}^2_{Sp} \otimes \mathcal{L} \otimes M)\, \mathrm{Td}_{\mathbf{T}}(\mathbb{C}^2). \tag{3.54}$$

where $M \cong \mathbb{C}$ is acted upon by the flavor symmetry group $U(1)_m$. Here we extended the universal bundles $\mathcal{E}^1_{Sp}$ and $\mathcal{E}^2_{Sp}$ over the product $\mathcal{M}_{Sp(N_1),k_1} \times \mathcal{M}_{Sp(N_2),k_2} \times \mathbb{R}^4$ by pulling

[4]The cases where $\epsilon_1 = -\epsilon_2$ were derived and discussed in [70, 71]. But their derivation can not be easily generalized to the general ϵ_1, ϵ_2.

them back using the respective projection maps $\pi_i : \mathcal{M}_{Sp(N_1),k_1} \times \mathcal{M}_{Sp(N_2),k_2} \to \mathcal{M}_{Sp(N_i),k_i}$. Define

$$
\begin{aligned}
P_1(x,a) &= \prod_l^{N_1}(x^2 - a_l^2), \\
P_2(x,b) &= \prod_m^{N_2}(x^2 - b_m^2), \\
\Delta(x) &= \prod_{i,j=1}^{n_1,n_2}((\phi_i + \tilde{\phi}_j)^2 - x^2)((\phi_i - \tilde{\phi}_j)^2 - x^2), \\
\Delta_1(x) &= \prod_i(\phi_i^2 - x^2), \\
\Delta_2(x) &= \prod_i(\tilde{\phi}_i^2 - x^2),
\end{aligned}
$$

where $k_i = 2n_i + \chi_i$. Then the instanton contour integrand is given by

$$
\begin{aligned}
\mathbf{z}_{k_1,k_2}^{N_1,N_2} \quad = \quad & \prod_i^{n_1} P_2(\phi_i + m)P_2(\phi_i + m + \epsilon) \prod_j^{n_2} P_1(\tilde{\phi}_j + m)P_1(\tilde{\phi}_j + m + \epsilon) \qquad (3.55) \\
& \times \prod_l^{N_1}(a_l^2 - m^2)^{\chi_2} \prod_k^{N_2}(b_k^2 - m^2)^{\chi_1} \\
& \times \left(\frac{\Delta(m - \epsilon_-)\Delta(m + \epsilon_-)}{\Delta(m - \epsilon_+)\Delta(m + \epsilon_+)}\right)\left(\frac{\Delta(m - \epsilon_-)\Delta(m + \epsilon_-)}{\Delta(m - \epsilon_+)\Delta(m + \epsilon_+)}\right)^{\chi_2} \\
& \times \left(\frac{\Delta(m - \epsilon_-)\Delta(m + \epsilon_-)}{\Delta(m - \epsilon_+)\Delta(m + \epsilon_+)}\right)^{\chi_1}\left(\frac{(m - \epsilon_-)(m + \epsilon_-)}{(m - \epsilon_+)(m + \epsilon_+)}\right)^{\chi_1\chi_2},
\end{aligned}
$$

where $\epsilon = \epsilon_1 + \epsilon_2$. Note that there are additional poles that involve the mass parameter m. The contour prescription is to assume $\epsilon_3 = -m - \epsilon_+$ and $\epsilon_4 = m - \epsilon_+$ to have a positive imaginary value. This is the same prescription as for the massive adjoint hypermultiplet in the $\mathcal{N} = 2^*$ theory [75].

Fundamental of $SO(N)$

The equivariant index of a fund. $SO(N)$ hypermultiplet of mass m is given by

$$\begin{aligned}\mathrm{Ind}_{\mathbf{T}} = & \int_{\mathbb{C}^2} \mathrm{Ch}_{\mathbf{T}}(\mathcal{E}_{\mathrm{SO}} \otimes \mathcal{L} \otimes M)\, \mathrm{Td}_{\mathbf{T}}(\mathbb{C}^2) \\ = & \frac{1}{(e^{i\epsilon_1}-1)(e^{i\epsilon_2}-1)} \left(\chi e^{i\epsilon_+} + \sum_{l=1}^{n} \left(e^{ia_l+im+i\epsilon_+} + e^{-ia_l+im+i\epsilon_+} \right) \right) \\ & - \sum_{i=1}^{k} \left(e^{i\psi_i+im} + e^{-i\psi_i+im} \right),\end{aligned} \tag{3.56}$$

where $N = 2n + \chi$. The corresponding instanton integrand is

$$\mathbf{z}_k^N = \prod_{i=1}^{k} (\psi_i + m)(\psi_i - m). \tag{3.57}$$

$SO(N)$ **gauge multiplet**

The equivariant index for an $SO(N)$ gauge multiplet is given by

$$\mathrm{Ind}_{\mathbf{T}} = -\int_{\mathbb{C}^2} \mathrm{Ch}_{\mathbf{T}}(\wedge^2 \mathcal{E}_{SO(N)})\, \mathrm{Td}_{\mathbf{T}}(\mathbb{C}^2) \tag{3.58}$$

The resulting contour integral is

$$Z_k = \frac{(-1)^{k(N+1)}}{2^k k!} \left(\frac{\epsilon}{\epsilon_1 \epsilon_2} \right)^k \oint \left(\prod_{i=1}^{k} \frac{d\psi_i}{2\pi i} \right) \frac{\Delta(0)\Delta(\epsilon)}{\Delta(\epsilon_1)\Delta(\epsilon_2)} \frac{\psi_i^2(\psi_i^2 - \epsilon_+^2)}{P(\psi_i + \epsilon_+)P(\psi_i - \epsilon_+)} \tag{3.59}$$

where $\epsilon = \epsilon_1 + \epsilon_2$ and

$$\begin{aligned} P(x) &= x^\chi \prod_{l=1}^{n} \left(x^2 - a_l^2\right) , \\ \Delta(x) &= \prod_{i<j} \left((\psi_i - \psi_j)^2 - x^2\right) \left((\psi_i + \psi_j)^2 - x^2\right) . \end{aligned}$$

When $\epsilon_1 = -\epsilon_2$ the pole structure is simplified and described by a set of N-colored Young diagrams like for the gauge group $U(N)$.[5]

[5]The cases where $\epsilon_1 = -\epsilon_2$ were derived and discussed in [70, 71]. But their derivation can not be easily generalized to the general ϵ_1, ϵ_2.

3.3 One-instanton contribution for arbitrary gauge group

In this section, we study the 1-instanton partition function for arbitrary gauge group. We follow the presentation of [53] in this section. Consider the partition function of the 5d supersymmetric field theory with the same matter content on the spacetime of the form $\mathbb{C}^2 \times \mathbb{R}$ parameterized by (z_1, z_2, x_5), with the identification

$$(z_1, z_2, x_5) \sim (z_1 e^{\epsilon_1 \beta}, z_2 e^{\epsilon_2 \beta}, x_5 + \beta), \tag{3.60}$$

together with an appropriate $\mathrm{SU}(2)_R$ symmetry rotation to preserve supersymmetry. The vev of the 4d scalar field can be included either via the vev of the 5d scalar field, or via the Wilson line of the gauge field around x^5. Here we use the latter. Then the partition function is

$$Z_{5d} = \mathrm{tr}(-1)^F \exp\left[i\beta H + \beta(\epsilon_1 J_1 + \epsilon_2 J_2 + a_i H^i)\right] \tag{3.61}$$

where H is the Hamiltonian, $J_{1,2}$ are the rotations of $z_{1,2}$-planes corrected with an appropriate amount of the $\mathrm{SU}(2)$ R-symmetry to commute with the supercharge, and H^i are the generators of the Cartan of the gauge group. Here the trace is taken in the field theory Hilbert space.[6]

Using the localization, the partition function can be written as the product of the one-loop contribution and the instanton contribution. The contribution Z_k from k-instanton configurations is

$$Z_{k,5d} = \mathrm{tr}_{\mathcal{H}_{k,\mathrm{BPS}}} \exp \beta(\epsilon_1 J_1 + \epsilon_2 J_2 + a_i H^i) \tag{3.62}$$

where $\mathcal{H}_{k,\mathrm{BPS}}$ is the BPS subspace of the Hilbert space of the supersymmetric quantum mechanics from the k-instanton configurations. For the pure gauge theory, this is just the supersymmetric sigma model whose target space is the k-instanton moduli space $\mathcal{M}_{G,k}$ of gauge group G. Then the BPS subspace $\mathcal{H}_{k,\mathrm{BPS}}$ is the space of holomorphic functions on $\mathcal{M}_{G,k}$. Therefore $Z_{k,5d}$ is just the character of the holomorphic functions on $\mathcal{M}_{G,k}$ under the action of the spacetime rotation $\mathrm{U}(1)^2$ and the gauge rotation G. This quantity is also known as the Hilbert series.

[6] Physically, it would be more natural to take $\epsilon_{1,2}$ and a_i to be purely imaginary, but supersymmetry guarantees that Z is holomorphic with respect to them. For convenience we regard $\epsilon_{1,2}$ as real and a_i as purely imaginary.

In the $\beta \to 0$ limit, $Z_{k,5d}$ is known to behave as

$$Z_{k,5d} \sim \beta^{-2kh^\vee} Z_k(\vec{a}; \epsilon_{1,2}) \tag{3.63}$$

and then the 4d instanton partition function is given by

$$Z_{\text{inst}} = \sum_k \Lambda^{2kh^\vee} Z_k(\vec{a}; \epsilon_{1,2}). \tag{3.64}$$

Then the instanton part of the prepotential is given by

$$F_{\text{inst}} = \lim_{\epsilon_1, \epsilon_2 \to 0} \epsilon_1 \epsilon_2 \log Z_{\text{inst}}. \tag{3.65}$$

3.3.1 One-instanton contribution

Let us calculate the one-instanton contribution for arbitrary G. When G is classical, we can use the ADHM description of the instanton moduli space to obtain the contribution [53, 70, 1]. Here, we use a more direct approach.

The one-instanton configuration of arbitrary gauge group G is obtained by embedding an SU(2) BPST instanton into G via a map $\mathrm{SU}(2) \hookrightarrow G$ associated to the long root of G [76, 77, 78, 79, 80]. Therefore the one-instanton moduli space has a decomposition

$$\mathcal{M}_{G,1} = \mathbb{C}^2 \times \tilde{\mathcal{M}}_{G,1} \tag{3.66}$$

where the factor $\mathbb{C}^2$ stands for the center of the instanton, and $\tilde{\mathcal{M}}_{G,1}$ stands for the size and the gauge direction of the instanton. $\mathcal{M}_{G,1}$ is a hyperkähler cone of real dimension $4(h^\vee - 1)$.

Holomorphic functions on $\mathbb{C}^2$ are just polynomials of the coordinates z_1 and z_2, and the character under the rotations $\mathrm{U}(1)^2 \subset \mathrm{SO}(4)$ is just $(1 - e^{\beta\epsilon_1})^{-1}(1 - e^{\beta\epsilon_2})^{-1}$. The space of holomorphic functions on $\tilde{\mathcal{M}}_{G,1}$ was known to mathematicians, e.g., [81, 82] using the fact that $\tilde{\mathcal{M}}_{G,1}$ is the orbit of the highest weight vector in $\mathfrak{g}_{\mathbb{C}}$ under $G_{\mathbb{C}}$. The same space was also studied from a physical point of view by [83]. The conclusion is that as the representation of $\mathrm{U}(1)^2 \times G$, the space of the holomorphic functions on $\mathcal{M}_{G,1}$ is decomposed as

$$\bigoplus_m (T_1 T_2)^{\otimes m} \otimes V(-m\vec{\alpha}_0) \tag{3.67}$$

	$\lvert W\rvert$	$\lvert\Delta_l\rvert$	G_0		$\lvert W\rvert$	$\lvert\Delta_l\rvert$	G_0
A_n	$(n+1)!$	n^2+n	A_{n-2}	B_n	$2^n n!$	$2n(n-1)$	$A_1 \times B_{n-2}$
D_n	$2^{n-1}n!$	$2n(n-1)$	$A_1 \times D_{n-2}$	C_n	$2^n n!$	$2n$	C_{n-2}
E_6	$72 \cdot 6!$	72	A_5	F_4	1152	24	C_3
E_7	$72 \cdot 8!$	126	D_6	G_2	12	6	A_1
E_8	$192 \cdot 10!$	240	E_7				

Table 3.1: Additional data of groups, required for the analysis in Appendix 3.3.2.

where T_i is the 1-dimensional representation of U(1) with character $e^{\beta\epsilon_i}$, $V(\vec{w})$ is the irreducible representation of G of the highest weight $\vec{w}$, and $-\vec{\alpha}_0$ is the highest root of the root system of G. The factor $(T_1T_2)^{\otimes m}$ arises from the fact that the radial direction is generated by $e^{(\epsilon_1+\epsilon_2)/2}$ under the same $\mathrm{U}(1)^2 \subset \mathrm{SO}(4)$.

Using the Weyl character formula, the character of this representation can be expressed as a summation over Weyl group elements, which can then be simplified as a summation over long roots, as we derive in 3.3.2. The end result is that the 4d 1-instanton contribution, including the contribution from the centre of mass, is given by

$$\boxed{Z_{k=1} = -\frac{1}{\epsilon_1\epsilon_2} \sum_{\vec{\gamma}\in\Delta_l} \frac{1}{(\epsilon_1+\epsilon_2+\vec{\gamma}\cdot\vec{a})(\vec{\gamma}\cdot\vec{a}) \prod_{\vec{\gamma}^\vee\cdot\vec{\alpha}=1,\ \vec{\alpha}\in\Delta}(\vec{\alpha}\cdot\vec{a})}\,,} \tag{3.68}$$

where Δ and Δ_l are the sets of the roots and the long roots, respectively. The one-instanton contribution to the prepotential via (3.65) reproduces the instanton calculation by Ito and Sasakura [84]. Explicit results for individual G will be discussed in Sec. 5.2.3.

3.3.2 Hilbert series of the one-instanton moduli space

In this section we derive expression (3.68) by computing the Hilbert series of the one-instanton moduli space. Let $V(\vec{w})$ be the highest weight representation of G of highest weight $\vec{w}$. As is conventional, we denote the highest root by $-\vec{\alpha}_0$, so that $V(-\vec{\alpha}_0)$ is the adjoint representation. As explained in Sec. 3.3.1, the holomorphic function on the centered 1-instanton moduli space $\tilde{\mathcal{M}}_{G,1}$ has the irreducible decomposition

$$V = \bigoplus_{m=0}^{\infty} V(-m\vec{\alpha}_0) \otimes T^{\otimes m}\,, \tag{3.69}$$

under the action of $\mathrm{U}(1) \times G$, where T is a one-dimensional representation of $\mathrm{U}(1)$. The character, or equivalently the Hilbert series is then

$$Z = \mathrm{tr}_V e^{\vec{\phi}} e^{\mu} \tag{3.70}$$

where $\vec{\phi}$ is an element of the Cartan subalgebra of G, and e^{μ} is the $\mathrm{U}(1)$ action. We will abbreviate $e^{\vec{\alpha}\cdot\vec{\phi}}$ as $e^{\vec{\alpha}}$. The Weyl character formula then gives

$$\begin{aligned} Z &= \sum_{m=0}^{\infty} e^{m\mu} \frac{\sum_{w\in W} s(w) e^{-wm\vec{\alpha}_0 + w\vec{\rho}}}{\prod_{\vec{\alpha}\in\Delta^+}(e^{\vec{\alpha}/2} - e^{-\vec{\alpha}/2})} = \frac{\sum_{w\in W} s(w) e^{w\vec{\rho}}/(1 - e^{\mu - w\vec{\alpha}_0})}{\prod_{\vec{\alpha}\in\Delta^+}(e^{\vec{\alpha}/2} - e^{-\vec{\alpha}/2})} \\ &= \sum_{\vec{\gamma}\in\Delta_l} \frac{1}{(1 - e^{\mu+\vec{\gamma}}) \prod_{\vec{\alpha}\in\Delta^+}(e^{\vec{\alpha}/2} - e^{-\vec{\alpha}/2})} \sum_{\substack{w\in W \\ -w\vec{\alpha}_0 = \vec{\gamma}}} s(w) e^{w\vec{\rho}}. \end{aligned} \tag{3.71}$$

Here, Δ is the set of roots, Δ_l is the set of long roots, Δ^+ is the set of positive roots, $\vec{\rho}$ is the Weyl vector, and $s(w)$ is the sign of an element w of the Weyl group W.

Now, consider a subgroup G_0 of G, whose Dynkin diagram is formed by the nodes of the Dynkin diagram of G which is *not* connected to the extended node of the affine Dynkin diagram, see Figure A.1 and Table 3.1. The Weyl group of G_0 fixes $-\vec{\alpha}_0$ by construction. Furthermore, $|W(G)| = |\Delta_l| \cdot |W(G_0)|$. Therefore, $W(G_0)$ is exactly the subgroup of $W(G)$ which fixes $-\vec{\alpha}_0$.

The difference $\vec{\rho}(G) - \vec{\rho}(G_0)$ of the Weyl vectors of G and G_0 is perpendicular to all $\vec{\alpha}_i$ of G_0, and therefore is proportional to $-\vec{\alpha}_0$. Therefore there is a constant c such that

$$\vec{\rho}(G) = -c\vec{\alpha}_0 + \vec{\rho}(G_0). \tag{3.72}$$

To fix c, take the inner product with respect to $-\vec{\alpha}_0$, using the expansion $-\vec{\alpha}_0^\vee = \sum n_i^\vee \alpha_i^\vee$ where $n_i^\vee$ are the comarks shown in Figure A.1. We find

$$2c = -\vec{\alpha}_0^\vee \cdot \vec{\rho}(G) = \sum n_i^\vee \vec{\alpha}_i^\vee \cdot \vec{\rho}(G) = \sum n_i^\vee = h^\vee - 1. \tag{3.73}$$

Here we used that the sum of the comarks is $h^\vee - 1$.

For a long root $\vec{\gamma}$, we pick a Weyl group element $w_{\vec{\gamma}}$ such that $-w_{\vec{\gamma}}\vec{\alpha}_0 = \vec{\gamma}$. The set of w

such that $-w\vec{\alpha}_0 = \vec{\gamma}$ is then simply $w_{\vec{\gamma}} W(G_0)$. Therefore,

$$\sum_{-w\vec{\alpha}_0=\vec{\gamma}} s(w)e^{w\rho} = \sum_{w\in W(G_0)} s(w_{\vec{\gamma}}w)e^{w_{\vec{\gamma}}w\rho} = e^{-(h^\vee-1)w_{\vec{\gamma}}\vec{\alpha}_0/2}s(w_{\vec{\gamma}}) \sum_{w\in W(G_0)} s(w)e^{w_{\vec{\gamma}}w\rho(G_0)}$$
$$= e^{(h^\vee-1)\vec{\gamma}/2}s(w_{\vec{\gamma}}) \prod_{\vec{\alpha}\in\Delta^+(G_0)} (e^{w_{\vec{\gamma}}\vec{\alpha}/2} - e^{-w_{\vec{\gamma}}\vec{\alpha}/2}) \,. \quad (3.74)$$

Plugging it in to (3.71), we have

$$\begin{aligned} Z &= \sum_{\vec{\gamma}\in\Delta_l} \frac{e^{(h^\vee-1)\vec{\gamma}/2}s(w_{\vec{\gamma}}) \prod_{\vec{\alpha}\in\Delta^+(G_0)}(e^{w_{\vec{\gamma}}\vec{\alpha}/2} - e^{-w_{\vec{\gamma}}\vec{\alpha}/2})}{(1-e^{\mu+\vec{\gamma}}) \prod_{\vec{\alpha}\in\Delta^+}(e^{\vec{\alpha}/2} - e^{-\vec{\alpha}/2})} \\ &= \sum_{\vec{\gamma}\in\Delta_l} \frac{e^{(h^\vee-1)\vec{\gamma}/2}s(w_{\vec{\gamma}}) \prod_{\vec{\alpha}\in\Delta^+(G_0)}(e^{w_{\vec{\gamma}}\vec{\alpha}/2} - e^{-w_{\vec{\gamma}}\vec{\alpha}/2})}{(1-e^{\mu+\vec{\gamma}})s(w_{\vec{\gamma}}) \prod_{\vec{\alpha}\in\Delta^+}(e^{w_{\vec{\gamma}}\vec{\alpha}/2} - e^{-w_{\vec{\gamma}}\vec{\alpha}/2})} \\ &= \sum_{\vec{\gamma}\in\Delta_l} \frac{e^{(h^\vee-1)\vec{\gamma}/2}}{(1-e^{\mu+\vec{\gamma}}) \prod_{\vec{\alpha}\in\Delta^+\backslash\Delta^+(G_0)}(e^{w_{\vec{\gamma}}\vec{\alpha}/2} - e^{-w_{\vec{\gamma}}\vec{\alpha}/2})} \,. \end{aligned} \quad (3.75)$$

Recall that the inner product $-\vec{\alpha}_0^\vee \cdot \vec{\alpha}$ is

- 2 if and only if $\vec{\alpha} = -\vec{\alpha}_0$,
- 1 if and only if $\vec{\alpha} \in \Delta^+\backslash\Delta^+(G_0)$ and $\vec{\alpha} \neq -\vec{\alpha}_0$,

so that (3.75) can be further written as

$$\begin{aligned} &= \sum_{\vec{\gamma}\in\Delta_l} \frac{e^{(h^\vee-1)\vec{\gamma}/2}}{(1-e^{\mu+\vec{\gamma}})(e^{-w_{\vec{\gamma}}\vec{\alpha}_0/2} - e^{w_{\vec{\gamma}}\vec{\alpha}_0/2}) \prod_{-\vec{\alpha}_0^\vee\cdot\vec{\alpha}=1}(e^{w_{\vec{\gamma}}\vec{\alpha}/2} - e^{-w_{\vec{\gamma}}\vec{\alpha}/2})} \\ &= \sum_{\vec{\gamma}\in\Delta_l} \frac{e^{(h^\vee-1)\vec{\gamma}/2}}{(1-e^{\mu+\vec{\gamma}})(e^{\vec{\gamma}/2} - e^{-\vec{\gamma}/2}) \prod_{\vec{\gamma}^\vee\cdot\vec{\alpha}=1}(e^{\vec{\alpha}/2} - e^{-\vec{\alpha}/2})} \,. \end{aligned} \quad (3.76)$$

Now the elements $w_{\vec{\gamma}}$ is gone.

Note that $|\{\vec{\alpha} \mid \vec{\gamma}^\vee \cdot \vec{\alpha} = 1\}|$ is $2h^\vee - 4$, so that (3.76) has $2h^\vee - 2$ terms in the denominator, which is the expected number for the centered instanton moduli space. To get the explicit expressions, we let

$$e^{\vec{\alpha}} \equiv e^{\vec{\alpha}\cdot\phi} = e^{\beta\vec{\alpha}\cdot\vec{a}}, \qquad e^{\mu} = e^{\beta(\epsilon_1+\epsilon_2)}. \quad (3.77)$$

In the papers on Hilbert series, e.g., [83], the variables $x_i = e^{a_i}$ were used instead. The 4d

version is obtained by taking the $\beta \to 0$ limit, giving

$$\beta^{2h^\vee-2} Z \to \sum_{\vec{\gamma}\in\Delta_l} \frac{-1}{(\epsilon_1+\epsilon_2+\vec{\gamma}\cdot\vec{a})(\vec{\gamma}\cdot\vec{a})\prod_{\vec{\gamma}^\vee\cdot\vec{\alpha}=1}(\vec{\alpha}\cdot\vec{a})}. \tag{3.78}$$

Together with the contribution of $\mathbb{C}^2$ we indeed get (3.68).

3.4 $\mathcal{N}=2$ SCFT and renormalization scheme

3.4.1 Infrared versus ultraviolet

In this section, we will compare the Nekrasov partition functions for gauge theories whose gauge group can be represented in two ways, possibly differing by a $U(1)$ factor. Think, for instance, of $Sp(1)$ versus $U(2)$ or $SO(4)$ versus $U(2)\times U(2)$. We are particularly keen on comparing $Sp(1)$ and $U(2)$ partition functions for conformally invariant theories. Naively, we would expect that the difference simply reproduces the "$U(1)$ factor". As we report in subsection 3.4.2, however, the Nekrasov partition functions of the $Sp(1)$ and $U(2)$ gauge theory coupled to four hypermultiplets are not at all related in such a simple way.

To find a resolution of this disagreement, we should keep in mind the difference between infrared and ultraviolet quantities. Whereas the Nekrasov partition function $Z^{\text{Nek}}(q)$ computes low-energy quantities, such as the prepotential $\mathcal{F}_0$, it is defined in terms of a series expansion in the exponentiated microscopic gauge coupling $q = \exp(2\pi i \tau_{UV})$. The gauge coupling τ_{UV}, however, is sensitive to the choice of the renormalization scheme and therefore cannot be assigned a physical (low-energy) meaning. It turns out that the renormalization schemes in the two instanton computations indeed differ. This means in particular that we should not identify the microscopic gauge couplings for the $U(2)$ and the $Sp(1)$ gauge theory. Instead, we should only expect to find agreement between the $Sp(1)$ and $U(2)$ Nekrasov partition functions when we express them in terms of physical low energy variables.

What are such low energy variables? Recall that in the low energy limit of the $\mathcal{N}=2$ gauge theory the Coulomb branch opens up, which is classically parametrized by the Casimirs of the gauge group. The prepotential $\mathcal{F}_0$ of the $\mathcal{N}=2$ gauge theory determines the corrections to the metric on the Coulomb branch, whose imaginary part in turn prescribes the period matrix τ_{IR} of the so-called Seiberg-Witten curve [15, 16]. As

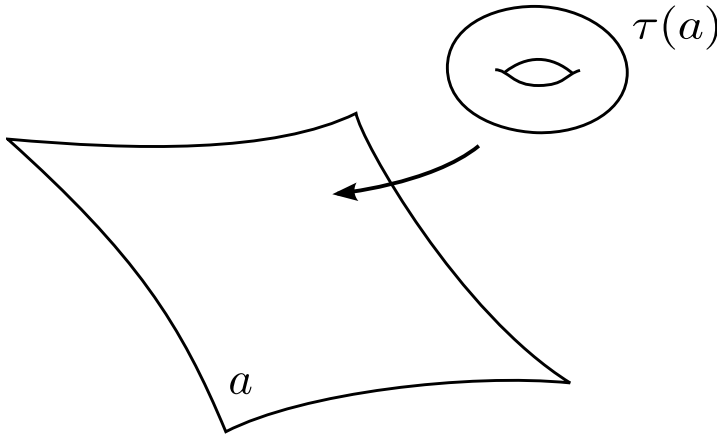


Figure 3.4: The period matrix $\tau_{\text{IR,ij}}$ of the Seiberg-Witten curve is equal to the second derivative $\partial_{a_i}\partial_{a_j}\mathcal{F}_0$ of the prepotential with respect to the Coulomb parameter a. The imaginary part of τ_{IR} determines the metric on the Coulomb branch.

the Seiberg-Witten curve changes along with the Coulomb moduli a, its Jacobian defines a torus-fibration over the Coulomb branch (see Figure 3.4). Whereas for asymptotically free gauge theories the Seiberg-Witten curve depends on the dynamically generated scale Λ, for conformally invariant gauge theories the Seiberg-Witten curve is dependent on the value of the microscopic gauge couplings τ_{UV}. Conformally invariant theories are characterized by a moduli space for the UV gauge couplings, from each element of which a Coulomb moduli space emanates in the low energy limit.

Since the Seiberg-Witten curve determines the masses of BPS particles in the low-energy limit of the $\mathcal{N} = 2$ gauge theory, its period matrix τ_{IR} is a physical quantity that should be independent of the chosen renormalization scheme. In contrast, the microscopic gauge couplings τ_{UV} are characteristics of the chosen renormalization scheme.

To be more concrete, we can use the prepotential $\mathcal{F}_0$ computed from the Nekrasov partition function to find the relation between τ_{IR} and τ_{UV} by

$$2\pi i\tau_{\text{IR}} = \frac{1}{2}\partial_a^2\mathcal{F}_0(\tau_{\text{UV}}, a) = \frac{1}{2}\partial_a^2(\mathcal{F}_{0,\text{pert}} + \mathcal{F}_{0,\text{inst}})(\tau_{\text{UV}}, a)\,. \tag{3.79}$$

Here $\mathcal{F}_{0,\text{pert}}$ contains the classical as well as 1-loop contribution to the prepotential, which, for instance, can be found in [68][7]. In particular, if two prepotentials that are computed using two different schemes differ by an a-dependent term, then the corresponding rela-

[7] Note that there is a typo in the expression for the gauge contribution in [68].

tions between τ_{IR} and τ_{UV} differ as well. If we invert the relation (3.79), and express both Nekrasov partition functions Z^{Nek} in terms of the period matrix τ_{IR}, we expect that they should agree up to a possible spurious factor that doesn't depend on the Coulomb parameters. This says that the two ways of instanton counting correspond to two distinct renormalization schemes.

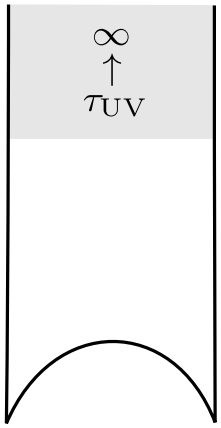


Figure 3.5: The marginal coupling τ_{UV} in the Nekrasov partition function defines a local coordinate on the moduli space of the $\mathcal{N} = 2$ conformal gauge theory near a weak-coupling point where $\tau_{\text{UV}} \to \infty$.

In fact, it is not quite obvious that the full Nekrasov partition functions, in contrast to just the prepotential, should agree when expressed in the period matrix τ_{IR}. It would have been possible that the relation between the prepotential τ_{IR} and the microscopic couplings τ_{UV} gets quantum corrections in terms of the deformation parameters ϵ_1 and ϵ_2, in such a way that only when expressed in terms of a quantum period matrix $\tau_{\text{IR}}(\epsilon_1, \epsilon_2)$ the Nekrasov partition functions do agree. In subsection 3.4.2 we will find however that this is not the case. The Nekrasov partition function agree when expressed in terms of the classical period matrix τ_{IR}. One possible argument for this is that the higher genus free energies $\mathcal{F}_{g\geqslant 1}$ are uniquely determined given the prepotential $\mathcal{F}_0$. In other words, that when the prepotentials (and thus Seiberg-Witten curves plus differentials) for two gauge theories agree we also expect the higher genus free energies to match up. This is reasonable to expect from several points of view, i.e., the interpretation of the $\mathcal{F}_g$s as free energies in an integrable hierarchy [85, 86].

The Nekrasov partition function is computed as a series expansion in $q = \exp(2\pi i \tau_{UV})$. The microscopic coupling τ_{UV} thus corresponds to a choice of local coordinate on the mod-

uli space of microscopic gauge couplings near a weak-coupling point (see Figure 3.5). An inequivalent renormalization scheme corresponds to a different choice of coordinate in that neighborhood. In particular, given two different renormalization schemes, by combining their respective IR-UV relations, we can obtain the relation between the two different microscopic couplings, and thus find the explicit coordinate transformation on the moduli space. Explicitly, we identify the infra-red couplings of two related theories by

$$\tau_{\text{IR}} = \frac{1}{2}\partial_a^2 \mathcal{F}_0^A(\tau_{\text{UV}}^A, a) = \frac{1}{2}\partial_a^2 \mathcal{F}_0^B(\tau_{\text{UV}}^B, a). \tag{3.80}$$

By inverting the IR-UV relation for the gauge theory characterized by the microscopic coupling τ_{UV}^A, we find the relation between the microscopic couplings τ_{UV}^A and τ_{UV}^B of both gauge theories. Since this is a relation between quantities in the ultra-violet, we expect it to be independent of infra-red parameters such as the masses and Coulomb branch parameters. Indeed, in all examples that we study in subsection 3.4.2, we will find that the moduli-independent UV-UV relation that follows from equation (3.80) relates the Nekrasov partition functions up to a spurious factor that is independent of the Coulomb parameters.[8]

We will often consider gauge theories as being embedded in string theories. Different models of the same gauge theory give different embeddings in string theory, which means that the results will differ when expressed in terms of UV variables, even though the IR results agree.

Take as an example the string theory realization of a supersymmetric $\mathcal{N} = 2$ $SU(2)$ gauge theory. The unitary point of view leads to a construction of D4, NS5 and D6-branes in type IIA theory [33], whereas the symplectic point of view introduces an orientifold in this picture and mirror images for all D4-branes [35, 36, 37]. Clearly, these are different realizations of the $SU(2)$ gauge theory. Nevertheless, both descriptions should give the same result in the infra-red.

Indeed, the two aforementioned string theory embeddings, based on either a $U(2)$ or a $Sp(1)$ gauge group, determine a physically equivalent Seiberg-Witten curve. For instance,

[8] Yet another example of a renormalization scheme for the four-dimensional $Sp(1)$ gauge theory with four flavors is found by counting string instantons in a system of D3 and D7 branes in Type I' [87, 88, 89].

the brane embedding of the pure $Sp(1)$ gauge theory determines the curve [36]

$$s^2 - s\left(v^2(v^2+u) + 2\Lambda^4\right) + \Lambda^8 = 0, \tag{3.81}$$

in terms of the covering space variables $s \in \mathbb{C}^*$, $v \in \mathbb{C}$ and the gauge invariant coordinate $u = \mathrm{Tr}(\Phi^2)$ on the Coulomb branch. This is merely a double cover [90] of the more familiar parametrization of the $SU(2)$ Seiberg-Witten curve

$$\Lambda^2 t^2 - t\left(w^2 + u\right) + \Lambda^2 = 0, \tag{3.82}$$

with $t \in \mathbb{C}^*$ and $w \in \mathbb{C}$, which follows from the unitary brane construction [33].

In fact, the choice for an instanton renormalization scheme is closely related to the choice for a brane embedding, as the precise parametrizations of the Seiberg-Witten curves (3.81) and (3.82) can be recovered in a thermodynamic (classical) limit by a saddle-point approximation of the $Sp(1)$ and the $U(2)$ Nekrasov partition functions respectively [23, 53].

3.4.2 Examples

Let us illustrate the above theory by a selection of examples. We start with comparing $Sp(1)/SO(4)$ and $U(2)$ partition functions in gauge theories with a single gauge group, and extend this to partition functions for more general linear and cyclic quivers. In particular, we find the identification of $Sp(1)/SO(4)$ and $U(2)$ instanton partition functions expressed in low-energy moduli and the relation between the $Sp(1)/SO(4)$ and $U(2)$ microscopic gauge couplings.

3.4.2.1 $Sp(1)$ versus $U(2)$: the asymptotically free case

First of all, let us consider the $Sp(1)$ theory with a single gauge group coupled to N_f massive hypermultiplets, where N_f runs from 1 to 4. In the asymptotically free theories, with $N_f \leqslant 3$, we find that the $Sp(1)$ Nekrasov partition function equals the $U(2)$ Nekrasov partition function – with Coulomb parameters $(a, -a)$ – up to a factor that doesn't depend on the Coulomb parameter and only contributes to the low genus refined free energies $\mathcal{F}_{0,\frac{1}{2},1}$. In the following we will call a factor with these two properties a *spurious* factor.

More precisely, we compute that [9]

$$Z^{N_f=0}_{U(2)}(q) = Z^{N_f=0}_{Sp(1)}(q) \tag{3.83}$$

$$Z^{N_f=1}_{U(2)}(q) = Z^{N_f=1}_{Sp(1)}(q) \tag{3.84}$$

$$Z^{N_f=2}_{U(2)}(q) = Z^{N_f=2}_{Sp(1)}(q)[Z^{\tilde{N}_f=0}_{U(1)}(q)]^{1/2} \tag{3.85}$$

$$Z^{N_f=3}_{U(2)}(q) = Z^{N_f=3}_{Sp(1)}(q)[Z^{\tilde{N}_f=1}_{U(1)}(q)]^{1/2} \exp\left(-\frac{q^2}{32\epsilon_1\epsilon_2}\right), \tag{3.86}$$

up to degree six in the $q = \Lambda^{4-N_f}$ expansion, where $Z^{\tilde{N}_f}_{U(1)}$ is the instanton partition function of the $U(1)$ gauge theory coupled to $\tilde{N}_f$ hypermultiplets with masses m_1 up to $m_{\tilde{N}_f}$. Explicitly,

$$Z^{\tilde{N}_f=0}_{U(1)}(q) = \exp\left(-\frac{q}{\epsilon_1\epsilon_2}\right), \tag{3.87}$$

$$Z^{\tilde{N}_f=1}_{U(1)}(q) = \exp\left(-\frac{mq}{\epsilon_1\epsilon_2}\right), \tag{3.88}$$

with $q = \Lambda^{2-\tilde{N}_f}$, $m = \mu_1 + \mu_2 + \mu_3 + \epsilon_1 + \epsilon_2$ where μ_i being the masses of the fundamental hypermultiplets in equation (3.86). Note that the equalities (3.83)–(3.86) can equally well be written down for any combination of hypers in the fundamental and anti-fundamental representation of the $U(2)$ gauge group. The contribution of a fundamental hypermultiplet just differs from that of an anti-fundamental hypermultiplet by mapping $\mu \mapsto -\mu$.

Let us make two more remarks about the formulas (3.83)–(3.86). First, the form of the spurious factor in the equalities (3.83)–(3.86) is close to what is called the $U(1)$ factor in the AGT correspondence: they agree for $N_f = 2$ and differ slightly for the $N_f = 3$ theory. In particular, both factors don't depend on the Coulomb parameters and only contribute to the lowest genus contributions $\mathcal{F}_{0,\frac{1}{2},1}$ of the refined free energy. Second, since the $U(2)$ and $Sp(1)$ Nekrasov partition functions coincide up to moduli-independent terms, the relation between IR and UV couplings is the same. It follows that they will agree up to spurious factors even when written in terms of IR couplings.

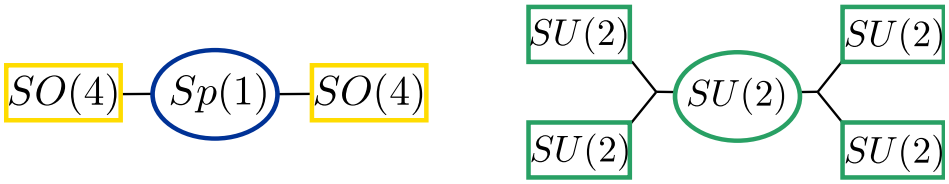


Figure 3.6: On the left: Quiver of the $Sp(1)$ gauge theory coupled to two fundamental and two anti-fundamental hypermultiplet. Since the (anti-)fundamental representation of $Sp(1)$ is pseudo-real, the flavor symmetry group of two hypermultiplets enhances to $SO(4)$. On the right: Quiver of the $SU(2)$ gauge theory coupled to two fundamental and two anti-fundamental hypermultiplets. The flavor symmetries of the hypermultiplets is enhanced to $SU(2)$.

3.4.2.2 $Sp(1)$ versus $U(2)$: the conformal case

Comparing the Nekrasov partition functions for conformal $Sp(1)$ and the $U(2)$ gauge theories, both coupled to four hypermultiplets, yields a substantially different result.[10,11] (The quivers of the respective gauge theories are illustrated in Figure 3.6 for later reference.) The $Sp(1)$ Nekrasov partition function does not agree with the $U(2)$ Nekrasov partition function up to a spurious factor, when expressed in the UV gauge couplings with the identification $q_{Sp(1)} = q_{U(2)}$. In particular, the prepotentials $\mathcal{F}_0$ differ, leading to a different relation between τ_{IR} and $q_{Sp(1)}$ than between τ_{IR} and $q_{U(2)}$.

When all the hypers are massless, or equivalently when sending the Coulomb parameter $a \to \infty$, we find that the map between UV gauge couplings and the period matrix τ_{IR} does not depend on a and is given by [12,13]

$$q^2_{Sp(1)} = 16\,\frac{\theta_2(q^2_{\text{IR}})^4}{\theta_3(q^2_{\text{IR}})^4} \tag{3.89}$$

$$q_{U(2)} = \frac{\theta_2(q_{\text{IR}})^4}{\theta_3(q_{\text{IR}})^4}, \tag{3.90}$$

where we define $q_{\text{IR}} = \exp(2\pi i \tau_{\text{IR}})$.[14] The $Sp(1)$ and $U(2)$ mappings just differ by doubling

[9] In this section we denote the Nekrasov partition function Z^{Nek} by Z. In the equations (3.83)-(3.86) there is an agreement for the instanton partition functions as well.

[10] The four flavor partition function is nevertheless perfectly consistent with the partition functions for fewer flavors. When we send the masses of the hypermultiplets to infinity, we do find the corresponding $N_f < 4$ partition functions.

[11] The expression for the ratio proposed in [91] does not hold beyond instanton number $k = 1$.

[12] All $Sp(1)$ results in this subsection have been checked up to order 6 in the $Sp(1)$ instanton parameter.

[13] Equation (3.90) was first found in [92].

[14] Here we use a convention different from appendix B of [26].

the value of the microscopic gauge coupling as well as the infra-red period matrix.

If we express the massless partition functions in terms of the low-energy variables by using (3.89) and (3.90), they agree up to a spurious factor (which is independent of the Coulomb parameter a and only contributes to the lower genus refined free energies $\mathcal{F}_{0,\frac{1}{2},1}$). In fact, it turns out that even if we reexpress the *massive* partition functions using the *massless* UV-IR mappings (3.89) and (3.90), we still find agreement up to a spurious factor.

On the other hand, even if we use the *massive* $Sp(1)$ and $U(2)$ IR-UV mappings which do depend on the Coulomb parameter a and the masses of the hypers, we find that the $Sp(1)$ and $U(2)$ renormalization scheme are related by the transformation

$$\boxed{q_{U(2)} = q_{Sp(1)} \left(1 + \frac{q_{Sp(1)}}{4}\right)^{-2}}, \tag{3.91}$$

which as expected is not dependent on the Coulomb branch moduli.

For completeness let us give the expression for the spurious factor once we express both full partition functions in terms of $q_{Sp(1)}$. For the unrefined case $\hbar = \epsilon_1 = -\epsilon_2$ we find

$$\frac{Z_{U(2)}(q_{U(2)}(q_{Sp(1)}))}{Z_{Sp(1)}(q_{Sp(1)})} = \left(1 + \frac{q_{Sp(1)}}{4}\right)^{M+N} \left(1 - \frac{q_{Sp(1)}}{4}\right)^{N-M}, \tag{3.92}$$

where $M = \frac{1}{\hbar^2}\sum_{i<j} \mu_i \mu_j$ and $N = -\frac{1}{2\hbar^2}\sum_i \mu_i^2 + \frac{1}{8}$. Here, we emphasize that this relation is between the full Nekrasov partition functions including the perturbative pieces and not just between the instanton parts. Notice that this spurious factor is quite close to, yet more complicated than the square-root of the unrefined $U(1)$ partition function

$$Z_{U(1)}^{N_f=2}(q) = (1-q)^{-\frac{m_1 m_2}{\hbar^2}},$$

of the $U(1)$ gauge theory coupled to two hypermultiplets with masses m_1 and m_2, the square of which entered the AGT correspondence as the "$U(1)$ factor". Similarly, we interpret the spurious factor (3.92) as a decoupled $U(1)$ factor.

3.4.2.3 $SO(4)$ versus $U(2) \times U(2)$ instantons

The instanton partition function for the pure $SO(4)$ gauge theory agrees with that of the pure $U(2) \times U(2)$ theory [15]

$$Z^{N_f=0}_{SO(4)}(q) = Z^{N_b=0}_{U(2)\times U(2)}(q)\,, \tag{3.93}$$

if we make the identifications

$$q_{U(2),1} = q_{U(2),2} = 16\, q_{SO(4)} \quad \text{and} \quad (b_1, b_2) = (a_1 + a_2, a_1 - a_2)\,. \tag{3.94}$$

Here, $b_{1,2}$ are the Coulomb parameters of the $SO(4)$ gauge theory and $a_{1.2}$ those of the $U(2) \times U(2)$ gauge theory. The second relation follows simply from the embedding of $\mathfrak{su}(2) \times \mathfrak{su}(2)$ in $\mathfrak{so}(4)$.

When we couple the $SO(4)$ theory to a single massive hypermultiplet, its instanton partition function matches with that of the $U(2) \times U(2)$ theory coupled to a massive bifundamental up to a spurious factor

$$Z^{N_f=1}_{SO(4)}(q) = Z^{N_b=1}_{U(2)\times U(2)}(q) \exp\left(-\frac{4q}{\hbar^2}\right), \tag{3.95}$$

for the unrefined case under the same identification (3.94).

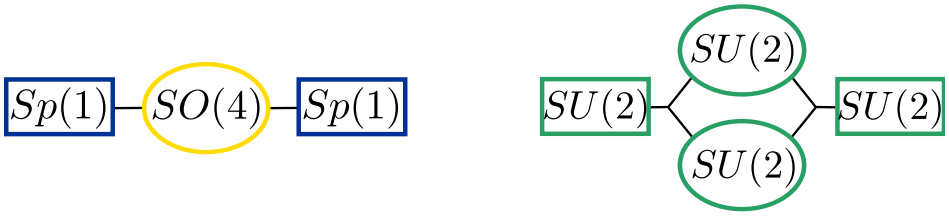


Figure 3.7: On the left: Quiver representation of the $SO(4)$ gauge theory coupled to one fundamental and one anti-fundamental hypermultiplet. Since the (anti-)fundamental representation of $SO(4)$ is real, the flavor symmetry group of each hypermultiplet enhances to $Sp(1)$. On the right: Quiver representation of the $SU(2) \times SU(2)$ gauge theory coupled to two bi-fundamental hypermultiplets. The flavor symmetry of the bifundamental field is enhanced to $SU(2)$.

Now, let us consider the conformal case. Naively comparing the Nekrasov partition

[15] We checked the $SO(4)$ results in this subsection up to order 2 for the refined $SO(4)$ partition functions and up to order 6 for the unrefined ones.

function of the conformal $SO(4)$ gauge theory coupled to two massive hypermultiplets with that of the $U(2) \times U(2)$ theory coupled by two massive bifundamentals shows a serious disagreement. (Their quivers are illustrated in Figure 3.7.) However, if we follow the same strategy as explained in the conformal $Sp(1)$ example, we see that it is once more simply a matter of different renormalization schemes. The $SO(4)$ and the $U(2)$ gauge theory are related by the change of marginal couplings

$$\boxed{q_{SO(4)} = \frac{\theta_2(q_{U(2)})^4}{\theta_3(q_{U(2)})^4},} \tag{3.96}$$

when we identify $q_{U(2)} = q_{U(2),1} = q_{U(2),2}$. Using this UV-UV relation and the relation between the Coulomb branch parameters (3.94), the $SO(4)$ and $U(2) \times U(2)$ Nekrasov partition functions agree up to a spurious factor

$$\frac{Z_{U(2)\times U(2)}(q_{U(2)}(q_{SO(4)}))}{Z_{SO(4)}(q_{SO(4)})} = 1 - \frac{4M}{\hbar^2}q + \frac{8M^2 + 2N\hbar^2}{\hbar^4}q^2 + \dots \tag{3.97}$$

that is similar to the $Sp(1) - U(2)$ spurious factor in equation (3.92), with now $M = m_1^2 - m_1 m_2 + m_2^2$ and $N = 3m_1^2 + m_1 m_2 + 3m_2^2$.

3.4.2.4 $Sp(1) \times Sp(1)$ versus $U(2) \times U(2)$ instantons

Next, we analyze $Sp(1) \times Sp(1)$ quiver gauge theories coupled to at most 4 massive hypermultiplets. The bifundamental multiplet, that couples the two $Sp(1)$ gauge groups, introduces new poles in the theory, similar to the adjoint multiplet in the $\mathcal{N} = 2^*$ gauge theory.

As expected, we find immediate agreement between the $Sp(1)$ and $U(2)$ instanton partition functions up to a spurious factor, when we couple fewer than two hypers to each multiplet.

If more than two hypers are coupled to one of the gauge groups, we need to express the partition function in terms of the physical period matrix τ_{IR}. Notice that the $Sp(1)$ (as well as $U(2)$) instanton partition function is a function of two UV gauge couplings, whereas the period matrix is a symmetric 3×3 matrix. It is nevertheless easy to find a bijective relation between the two diagonal entries of the period matrix and the two UV gauge couplings. The off-diagonal entry in the period matrix represents a mixing of the two gauge groups,

and can be expressed in terms of the diagonal entries.

Let us consider the conformal linear quiver with two $Sp(1)$ gauge groups as an example. We couple the two $Sp(1)$'s by a bifundamental and add two extra hypermultiplets to the first and to the second gauge group. The *moduli-independent* UV-IR relation for $Sp(1)$ has a series expansion [16]

$$q_{Sp(1),i} = q_{\mathrm{IR},ii} - \frac{1}{64} q_{\mathrm{IR},ii}^3 + \frac{1}{32}\, q_{\mathrm{IR},ii} q_{\mathrm{IR},jj}^2 + \mathcal{O}(q_{\mathrm{IR}}^4), \tag{3.98}$$

whereas the one for $U(2)$ has the form

$$\begin{aligned} q_{U(2),i} = q_{\mathrm{IR},ii} &- \frac{1}{2} q_{\mathrm{IR},ii}^2 + \frac{1}{2} q_{\mathrm{IR},ii} q_{\mathrm{IR},jj} + \frac{11}{64} q_{\mathrm{IR},ii}^3 \\ &- \frac{1}{2} q_{\mathrm{IR},ii}^2 q_{\mathrm{IR},jj} + \frac{3}{32} q_{\mathrm{IR},ii} q_{\mathrm{IR},jj}^2 + \mathcal{O}(q_{\mathrm{IR}}^4), \end{aligned} \tag{3.99}$$

for $i \in \{1,2\}$ and $i \neq j$.

As before we use the *moduli-independent* UV-IR mappings (3.98) and (3.99) to evaluate the *massive* partition function as a function of the physical IR moduli $\tau_{\mathrm{IR},11}$ and $\tau_{\mathrm{IR},22}$. Again this shows agreement of the $Sp(1)$ and $U(2)$ partition functions up to a spurious factor in the lower genus free energies.

Composing the *moduli-dependent* mappings between UV-couplings and the period matrix, we find that the two renormalization schemes are related by

$$\begin{aligned} q_{Sp(1),i} = q_{U(2),i} &+ \frac{1}{2} q_{U(2),i}^2 - \frac{1}{2} q_{U(2),i} q_{U(2),j} + \frac{5}{16} q_{U(2),i}^3 \\ &- \frac{1}{16} q_{U(2),i} q_{U(2),j}^2 + \mathcal{O}(q_{U(2)}^4), \end{aligned} \tag{3.100}$$

for $i \in \{1,2\}$ and $i \neq j$. Note that this mapping is independent of the Coulomb branch moduli and the mass parameters, as it should be. Notice as well that a mixing amongst the two gauge groups takes places, so that we cannot simply use the UV-UV mapping for a single gauge group twice. Substituting this relation into the $Sp(1)$ partition function indeed turns brings it into the form of the $U(2)$ partition function up to a spurious AGT-like factor.

The above procedure can be applied to any linear or cyclic quiver.[17] Most importantly,

[16] Here we have rescaled $q_{\mathrm{IR}} \to q_{\mathrm{IR}}/16$.

[17] For example, we also tested it for the $\mathcal{N} = 2^*$ $Sp(1)$ gauge theory.

the $Sp(1)$ and $U(2)$ partition function agree (up to a spurious factor) when expressed in IR coordinates, and, the mapping between $Sp(1)$ and $U(2)$ renormalization schemes is independent of the moduli in the gauge theory and mixes the gauge groups. Moreover, the spurious factors that we find are more complicated than the "$U(1)$-factors" that appear in the AGT correspondence.

3.5 UV-IR relation and $\mathcal{N} = 2$ geometry

In section 3.4 we found in which way Nekrasov partition functions for different models of the same underlying physical gauge theory are related, i.e., comparing the $U(2)$ versus the $Sp(1)$ method of instanton counting. The $\mathcal{N} = 2$ instanton counting defines a renormalization scheme for each model, such that, when expressed in terms of physical infra-red variables, the Nekrasov partition functions of two such models agree up to a spurious factor. When expressed in terms of the microscopic couplings, however, the Nekrasov partition functions are related by a non-trivial mapping. Our goal in this section is to explain this mapping geometrically.

In [25] the complex structure moduli of the Gaiotto curve are identified with the exactly marginal couplings of the conformal $\mathcal{N} = 2$ gauge theory. The Gaiotto curve therefore not only captures the infra-red data of the gauge theory, but also information about the chosen renormalization scheme. Different renormalization schemes are related by non-trivial mappings of exactly marginal couplings. Geometrically, it was argued that a choice of renormalization scheme corresponds to a choice of local coordinates on the complex structure moduli space of the Gaiotto curve.

A given physical gauge theory of course admits a countless number of renormalization schemes. On the other hand, our focus here is only to find a geometric interpretation of a small subset of such choices. We expect to find such an interpretation when different models can both be embedded by a brane construction in string theory. In such examples we can construct the corresponding Gaiotto curves and a mapping between them. This mapping geometrizes the mapping between the exactly marginal couplings.

Most of the examples we considered in the previous section are of this type. However, one of them is not. This is the comparison of the $SU(2) \times SU(2)$ gauge theory with the

$Sp(1) \times Sp(1)$ gauge theory.[18] Since we are not aware of a brane embedding of the latter theory, there is no immediate reason to expect an (obvious) geometric interpretation of the mapping between exactly marginal couplings in that example. In the other examples, for which brane constructions are well known, we do expect to find a geometric explanation for the UV-UV mappings.

As a last comment, let us emphasize that the validity of the mappings between exactly marginal couplings extends beyond the prepotential $\mathcal{F}_0$. As we found in the last section, the full Nekrasov partition functions Z^{Nek} are related by the mapping between UV couplings, which does not receive corrections in the deformation parameters ϵ_1 and ϵ_2. Geometrically, this implies that there are no quantum corrections to the mappings between the Gaiotto curves for different models of the same physical gauge theory.

3.5.1 SO/Sp versus U geometries

Suppose that we have two models for the same physical gauge theory, who both can be embedded as a ramified Hitchin system in M-theory. How are these models related geometrically?

First of all, we expect that the mapping between the exactly marginal couplings is reflected as a mapping between the complex structure parameters of the corresponding Gaiotto curves. Also, there should be an isomorphism between the spectral curves of the respective Hitchin systems, as these correspond to the Seiberg-Witten curves. Moreover, the boundary conditions of the Hitchin differentials at the punctures of both models should be related by the mapping that identifies the corresponding matter representations. In particular, this relates the eigenvalues of the Seiberg-Witten differential at the punctures of both models. In total we should thus find a bijective mapping between the complete ramified Hitchin system, including the Gaiotto curve as well as the Hitchin differentials.

As we will see in detail in the next subsection, it is easy to come up with such a mapping for $Sp(1)/SO(4)$ versus $SU(2)$ gauge theories [38]. We just interpret the $\mathbb{Z}_2$-twist lines on the $Sp(1)/SO(4)$ Gaiotto curve as branch-cuts, and its double cover as the $SU(2)$ Gaiotto curve. The latter curve is thus equipped with an involution that interchanges the two sheets of the cover. We can recover the Hitchin differentials on the $Sp(1)/SO(4)$ Gaiotto curve by splitting the Hitchin differential on its cover into even and odd parts under the

[18] The same holds for the comparison of the $\mathcal{N} = 2^*$ theory with gauge group $Sp(1)$ versus $U(2)$.

involution. Indeed, recall that the $SU(2)$ Seiberg-Witten curve is determined by a single differential ϕ_2 of degree 2, whereas the $Sp(1)/SO(4)$ Seiberg-Witten curve is defined by two degree 2 differentials φ_2 and $\varphi_{\tilde{2}}$, the first one being even under the $\mathbb{Z}_2$-automorphism and the second one odd.

Notice that this double construction doesn't work for any Sp/SO theory, as the differentials on the cover generically do not have a simple interpretation in terms of a set of differentials of a unitary theory. Two theories can only be related by a double covering if the Lie algebra underlying the Hitchin system of one of them splits into two copies of the Lie algebra underlying the Hitchin system of the other. Nonetheless, for any two models of the same gauge theory there should be a corresponding isomorphism of Hitchin systems.

Before going into the example-subsection, let us note that in the previous section we also encountered a gauge theory without an obvious brane embedding. This is the $Sp(1) \times Sp(1)$ gauge theory. It is not possible to realize this theory in the standard manner using NS5-branes, as the type of the orientifold has to differ on either side of the NS5-brane. Geometrically, this is reflected in the fact that there doesn't exist an involution on the $SU(2) \times SU(2)$ Gaiotto curve with the right properties. It would be interesting to find whether there exists a geometric interpretation of the mapping between the exactly marginal couplings of the $SU(2) \times SU(2)$ and the $Sp(1) \times Sp(1)$ gauge theory anyway.

3.5.2 Examples

Let us now return to the results of section 3.4. First of all, we can explain the appearance of the modular lambda function

$$\lambda = \frac{\theta_2^4}{\theta_3^4} : \mathbb{H} \to \mathbb{P}^1 \backslash \{0, 1, \infty\} \tag{3.101}$$

as the relation (3.90) between the infra-red coupling τ_{IR} and the exactly marginal coupling $q_{U(2)}$ in the conformal $SU(2)$ gauge theory. This modular function gives an explicit isomorphism between the quotient $\mathbb{H}/\Gamma(2)$ of the upper half plane $\mathbb{H}$ by the modular group $\Gamma(2)$ and $\mathbb{P}^1 \backslash \{0, 1, \infty\}$. Whereas the complex structure modulus τ_{IR} of the $SU(2)$ Seiberg-Witten curve takes values in $\mathbb{H}/\Gamma(2)$, the complex structure modulus $q_{U(2)}$ of the $SU(2)$ Gaiotto curve (which is the cross-ratio of the four punctures on the UV-curve) takes values in $\mathbb{P}^1 \backslash \{0, 1, \infty\}$. The modular lambda function thus determines the double cover map

between the Seiberg-Witten curve and the Gaiotto curve for the conformal $SU(2)$ gauge theory [26].

We continue with studying the $Sp(1)$ and $SO(4)$ geometry in detail. In appendix C we have summarized various existing descriptions of the $SU(2)$ geometry, and their relation to the Gaiotto geometry.

3.5.2.1 $Sp(1)$ versus $U(2)$ geometry

The $Sp(1)$ Seiberg-Witten curve can be derived from the orientifold brane construction that is illustrated in Figure 2.3. In Gaiotto form it reads

$$v^4 = \varphi_2(s)v^2 + \varphi_4(s) \tag{3.102}$$

with

$$\begin{aligned}\varphi_2(s) &= \frac{(\mu_1^2+\mu_2^2)s^2 + u(1+\tilde{q}_{Sp(1)})s + (\mu_3^2+\mu_4^2)\tilde{q}_{Sp(1)}}{(s-1)\left(s-\tilde{q}_{Sp(1)}\right)}\left(\frac{ds}{s}\right)^2\\ \varphi_4(s) &= -\frac{\mu_1^2\mu_2^2 s^2 + 2\prod_{i=1}^4 \mu_i \sqrt{\tilde{q}_{Sp(1)}}s + \mu_3^2\mu_4^2\tilde{q}_{Sp(1)}}{(s-1)\left(s-\tilde{q}_{Sp(1)}\right)}\left(\frac{ds}{s}\right)^4,\end{aligned}$$

where μ_i are the bare masses of the hypermultiplets, and u is the classical vev of the adjoint scalar Φ in the gauge multiplet. We also introduced a new parameter $\tilde{q}_{Sp(1)}$ that we will relate to the coupling $q_{Sp(1)}$ in a moment. The differential φ_2 corresponds to the D_2-invariant $\mathrm{Tr}(\Phi^2)$, whereas the square-root $\varphi_{\tilde{2}} = \sqrt{\varphi_4}$ corresponds to the D_2-invariant Pfaff(■). Note that the differential φ_4 vanishes if the masses are set to zero.

It follows that the $Sp(1)$ Seiberg-Witten curve is a branched fourfold cover over the UV-curve $\mathbb{P}^1$ with coordinate s. The UV-curve has four branch points at the positions

$$s \in \{0, \tilde{q}_{Sp(1)}, 1, \infty\}, \tag{3.103}$$

The complex structure of the $Sp(1)$ Gaiotto curve is parametrized by $\tilde{q}_{Sp(1)}$. Since the differential $\varphi_{\tilde{2}}$ has a pole of order half at the punctures at $s = 1$ and $s = \tilde{q}_{Sp(1)}$, these are half-punctures. There is a $\mathbb{Z}_2$-twist line running between the two half-punctures, as the differential $\varphi_{\tilde{2}}$ experiences a $\mathbb{Z}_2$ monodromy around them. In contrast, the punctures at $s = 0$ and $s = \infty$ are full punctures.

The Seiberg-Witten differential λ is a $SO(4,\mathbb{C})$-valued differential. It has nonzero residues at the poles $s = 0$ and $s = \infty$ only. This implies that the $SO(4)$ flavor symmetry is associated with these punctures. Indeed, the residues of the differential λ at $s = \infty$ are given by $\pm\mu_1$ and $\pm\mu_2$, whereas at $s = 0$ they are $\pm\mu_3$ and $\pm\mu_4$. So both at $s = 0$ and $s = \infty$ the residues parametrize the Cartan of $\mathfrak{su}(2) \times \mathfrak{su}(2) = \mathfrak{so}(4)$.

Summarizing, we have found that the $Sp(1)$ UV-curve is a four-punctured two-sphere with two half-punctures and two full punctures. The two $SO(4)$-flavor symmetry groups can be associated to the two full punctures. This is illustrated in the left picture in Figure 2.4.

Viewing the $\mathbb{Z}_2$-twist lines as branch-cuts, and the half-punctures as branch-points, it is natural to consider the double cover of the UV-curve. Let us use the $SL(2,\mathbb{C})$ freedom of the $Sp(1)$ theory to interchange the full-punctures with the half-punctures. Call the complex structure coordinate of the $Sp(1)$ UV-curve $\tilde{q}_{Sp(1)} = q^2$. The branched covering map is then simply given by

$$t^2 = s,$$

where s is the coordinate on the $Sp(1)$ UV-curve and t the coordinate on the double cover. The pre-images of the full punctures on the base are at $\pm 1, \pm q$ on the cover, and there is a $SU(2)$-flavor symmetry attached to them. The total flavor symmetry at both full-punctures adds up to $SO(4)$. This is illustrated in Figure 3.8.

Note that this double cover of the $Sp(1)$ UV-curve has exactly the same structure as the $SU(2)$ UV-curve. The only difference is that the punctures are at different positions, something which can be taken care of by a Möbius transformation. Since this leaves the complex structure of the UV-curve invariant, the gauge theory is invariant under such transformations. In particular the masses of the hypermultiplets, which are the residues of the Seiberg-Witten differential at its poles, remain the same. We can use the fact that the cross-ratios of the two configurations have to be equal to read off the relation between the $U(2)$ and $Sp(1)$ exactly marginal couplings,

$$\boxed{q_{U(2)} = 4q\,(1+q)^{-2}\,.} \tag{3.104}$$

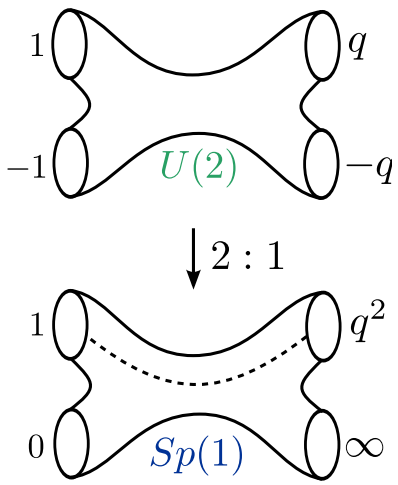


Figure 3.8: The UV-curve of the $SU(2)$ gauge theory coupled to 4 hypers is a double cover over the UV-curve of the $Sp(1)$ gauge theory with 4 hypers. We denote the complex structure parameter on the $Sp(1)$ Gaiotto curve by $\tilde{q}_{Sp(1)} = q^2$.

Explicitly, the Möbius transformation that relates the $SU(2)$ UV-curve and the double cover of the $Sp(1)$ UV-curve is given by the mapping

$$\gamma(z) = -\frac{z\,(1+q) - 2q}{z\,(1+q) - 2}\,, \tag{3.105}$$

that sends the four punctures at positions $\{0, 1, q_{U(2)}, \infty\}$ to four punctures at the positions $\{\pm q, \pm 1\}$.

We can now make contact between the geometry of the $SU(2)$ and $Sp(1)$ Gaiotto curves and the relation between their exactly marginal couplings. Indeed, we recover the UV-UV mapping (3.91) from the identification of cross-ratios in equation (3.104), when we identify

$$\boxed{q^2 = \tilde{q}_{Sp(1)} = \left(\frac{q_{Sp(1)}}{4}\right)^2.} \tag{3.106}$$

We should therefore choose the complex structure parameter $\tilde{q}_{Sp(1)}$ of the $Sp(1)$ Gaiotto-curve proportional to the square $q^2_{Sp(1)}$ of the $Sp(1)$ instanton parameter. The square is related to the $\mathbb{Z}_2$-twist line along the $Sp(1)$ Gaiotto curve. The proportionality constant is merely determined by requiring that $q_{U(2)} = q_{Sp(1)} + \ldots$, which is needed to make the classical contributions to the Nekrasov partition function agree.

3.5.2.2 $SO(4)$ versus $U(2) \times U(2)$ geometry

The $SO(4)$ Seiberg-Witten curve in Gaiotto form reads

$$v^4 = \varphi_2(s)v^2 + \varphi_4(s) \tag{3.107}$$

with

$$\varphi_2(s) = \frac{\mu_1^2 s^2 - (u_1+u_2)(1+\tilde{q}_{SO(4)})s + \mu_2^2\tilde{q}_{SO(4)}}{(s-1)\left(s-\tilde{q}_{SO(4)}\right)}\left(\frac{ds}{s}\right)^2$$
$$\varphi_4(s) = \frac{u_1 u_2 s}{(s-1)\left(s-\tilde{q}_{SO(4)}\right)}\left(\frac{ds}{s}\right)^4,$$

The Seiberg-Witten curve is a fourfold cover of the four-punctured sphere with complex structure parameter $\tilde{q}_{SO(4)}$. This time there are four half-punctures. The differential $\varphi_{\tilde{2}}$ not only has poles of order $1/2$ at $s = 1$ and $s = \tilde{q}_{SO(4)}$, but also poles of order $3/2$ at $s = 0$ and $s = \infty$. The residue of the Seiberg-Witten differential is only nonzero at $s = 0$ and $s = \infty$. Since its nonzero residues equal $\pm\mu_1$ at $s = \infty$ and $\pm\mu_2$ at $s = 0$, the $Sp(1)$ flavor symmetry is associated to these punctures.

Summarizing, the UV-curve corresponding to the conformal $SO(4)$ theory is a four-punctured sphere with four half-punctures, and $\mathbb{Z}_2$ twist-lines running between two pairs of half-punctures. This is illustrated on the left in Figure 2.5.

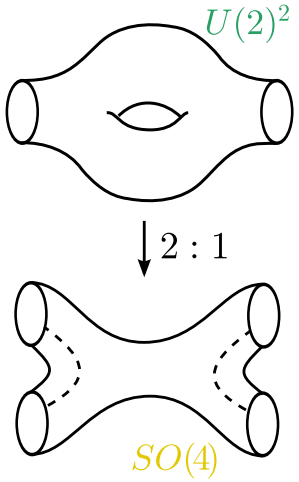


Figure 3.9: The UV-curve of the $SU(2) \times SU(2)$ gauge theory coupled by two bifundamentals is a double cover over the UV-curve of the $SO(4)$ gauge theory coupled to four hypers.

Let us again interpret the $\mathbb{Z}_2$ twist-lines as branch-cuts. As illustrated in Figure 3.9, this time the branched double cover of the UV-curve is a torus with two punctures. The punctures on the torus project to two of the four branch-points of the double covering, which means that there must be a $SU(2)$-flavor symmetry attached to them. This is precisely the the same structure as that of the Gaiotto curve of the $SU(2) \times SU(2)$ gauge theory coupled to two bifundamentals.

However, note that the $SU(2) \times SU(2)$ UV-curve has two complex structure parameters corresponding to the two marginal couplings $q_{U(2),1}$ and $q_{U(2),2}$, whereas the $SO(4)$ UV-curve has only a single complex structure parameter corresponding to the marginal coupling $q_{SO(4)}$ of the $SO(4)$ theory. The $\mathbb{Z}_2$-symmetry on the cover curve implies that we should identify $q_{U(2)} = q_{U(2),1} = q_{U(2),2}$ in order to relate the $SU(2) \times SU(2)$ Gaiotto curve to the $SO(4)$ Gaiotto curve.

In equation (3.96) we found that the relation between the marginal couplings of the $SO(4)$ and the $SU(2) \times SU(2)$-theory is given by the modular lambda mapping

$$\boxed{q_{SO(4)} = \lambda(2\tau_{U(2)})\,.} \tag{3.108}$$

The above covering relation between their Gaiotto curve explains the appearing of this lambda mapping geometrically, as it relates the complex structure parameter of the torus to the complex structure parameter of the four-punctured sphere, when we identify $\tilde{q}_{SO(4)} = q_{SO(4)}$.

Mainly for future convenience, let us make the covering map explicitly. Take a torus T^2 with half-periods ω_1 and ω_2 and consider the map $T^2 \to \mathbb{CP}^2$ given by

$$(\wp(z) : \wp'(z) : 1),$$

where $\wp(z)$ is the Weierstrass $\wp$-function. Since the Weierstrass $\wp$-function satisfies

$$\wp'(z)^2 = 4\wp(z)^3 - g_2\wp(z) - g_3,$$

it defines a branched double cover over the sphere $\mathbb{P}^1$. The branch points are determined by the zeroes and poles $\{0, \omega_1, \omega_2, \omega_3 = \omega_1 + \omega_2\}$ of the derivative $\wp'(z)$ of the Weierstrass

$\wp$-function. The double covering is thus given by the equation

$$t^2 = 4(s - e_1)(s - e_2)(s - e_3)$$

with $e_i = \wp(\omega_i)$. The points $\omega_3 = \tau$ and 2τ on the torus T^2 map onto the two branch points e_3 and ∞ on the sphere $\mathbb{P}^1$. Let us call $q_{SO(4)}$ the cross-ratio of the four branch points $0, e_1, e_2, e_3$. By the shift $s \mapsto s + \frac{(1+q_{SO(4)})}{3}$ we simply bring the curve into the form

$$t^2 = 4s(s - 1)(s - q_{SO(4)}).$$

The mapping between the $SU(2) \times SU(2)$ Gaiotto curve and the $SO(4)$ Gaiotto curve is thus a combination of the Weierstrass map $\wp$ and a simple Möbius transformation. This indeed determines (3.108) as the mapping between the respective complex structure parameters.

3.6 Instanton counting for half-hypermultiplets

The trifundamental fields that appear in Sicilian quiver gauge theories of type A_1 form half-hypermultiplet representations of the $\mathcal{N} = 2$ SUSY algebra, in contrast to the more common (full) hypermultiplets. In this section we review the basic properties of Sicilian quivers with trifundamental half-hypermultiplets and show that they preserve $\mathcal{N} = 2$ supersymmetry. Subsequently, we develop the tools for counting instantons in quiver gauge theories with half-hypermultiplets. We apply these tools in section 5.5.3 to compute the instanton partition functions of Sicilian quivers and compare them with CFT computations.

3.6.1 Half-hypermultiplets

There are two types of $\mathcal{N} = 2$ supersymmetry multiplets: the $\mathcal{N} = 2$ vector multiplet and the hypermultiplet. The former consists of a vector field A_μ, two Weyl fermions λ_α and ψ_α, and one complex scalar B. All of them transform in the adjoint representation of the gauge group. In $\mathcal{N} = 1$ language such an $\mathcal{N} = 2$ vector multiplet consists of one $\mathcal{N} = 1$ vector multiplet with component fields (A_μ, λ_α) and one chiral multiplet with components (B, ψ_α). A hypermultiplet requires a choice of representation R of the gauge group. In $\mathcal{N} = 1$ language it consists of two chiral multiplets Q and $\tilde{Q}$, the former transforming in

the representation R and the latter in its complex conjugate R^*. The chiral multiplet Q has component fields (ϕ, χ_α), and the anti-chiral multiplet $\tilde{Q}$ has components $(\tilde{\phi}, \tilde{\chi}_\alpha)$. We call both Q and $\tilde{Q}$ half-hypermultiplets.

The half-hypermultiplets Q and $\tilde{Q}$ form massless representations of the $\mathcal{N} = 2$ SUSY algebra. However, even though the helicities of the states in a half-hypermultiplet form a CPT complete distribution, the half-hypermultiplet does not transform as a real representation of the SUSY algebra. Indeed, notice that the massless $\mathcal{N} = 2$ SUSY algebra is equal to the Clifford algebra $\mathrm{Cl}_{4,0}$ with invariance group $SO(4)$. The four-dimensional representation of the Clifford algebra $\mathrm{Cl}_{4,0}$, under which the half-hypermultiplet transforms, is pseudo-real instead of (strictly) real. A generic half-hypermultiplet will therefore not be invariant under CPT.

It is possible though to circumvent this constraint. More precisely, an $\mathcal{N} = 2$ multiplet is CPT invariant if it transforms under a real representation of the product of the SUSY algebra, the gauge group and possible flavor groups. So, apart from the obvious possibility of combining a chiral and an anti-chiral multiplet into a full hypermultiplet, we can also consider a single half-hypermultiplet in a pseudo-real representation of the gauge group G.

Nevertheless, there is an additional requirement. Even when a single half-hyper-multiplet transforms in a real representation of $\mathrm{Cl}_{4,0} \times G$, such a theory may still be anomalous due to Witten's anomaly argument [93]. According to this argument, for example, a single half-hypermultiplet in the fundamental representation of $SU(2)$ is anomalous (its partition function vanishes) since it contains an odd number of chiral fermions. On the contrary, a single half-hypermultiplet in the fundamental representation of $SU(2)^3$ contains four chiral fermions in each $SU(2)$-representation. Therefore, quiver gauge theories with $SU(2)$ trifundamental half-hypermultiplets are free of Witten's $SU(2)$ anomaly as well as CPT invariant.

Other examples of consistent theories with half-hypermultiplets occur when we consider massless bifundamental couplings between SO and Sp gauge groups. A half-hypermultiplet transforming under the bifundamental of $SO \times Sp$ is in a pseudo-real representation of the gauge group $G = SO \times Sp$ and is free of the Witten anomaly as well.

Half-hypermultiplet as a constrained hypermultiplet

Since working with full hypermultiplets is often much more efficient than with half-hypermultiplets, in what follows we find an alternative method to deal with half-hypers. Instead of considering a half-hypermultiplet by itself, we start with a full hypermultiplet (consisting of two half-hypermultiplets) and impose a constraint on it which only leaves a half-hypermultiplet.

A full hypermultiplet can be thought of as a multiplet formed out of two $\mathcal{N} = 1$ chiral multiplets Q and $\tilde{Q}$. The chiral multiplet Q transforms in representation R and the anti-chiral multiplet $\tilde{Q}$ in its complex conjugate R^*. By the remarks above, for the theory of a single half-hyper to make sense, R needs to be a pseudoreal. Let σ_G be the anti-linear involution that maps the representation R to its complex conjugate R^*. Since the representation R is pseudo-real, it obeys $\sigma_G^2 = -1$. For example, in the case of the fundamental representation of $SU(2)$, the involution σ_G is given by the ϵ-tensor $i\sigma_2$. For the trifundamental representation of $SU(2)$ it is given by the product of three ϵ-tensors, one for each $SU(2)$ gauge group.

To impose our constraint, we need a map τ that relates Q to $\tilde{Q}$ and vice versa. Since $\tilde{Q}$ appears in the complex conjugate, τ needs to be anti-linear. Moreover, it needs to preserve the representation, which means that it must involve σ_G. Let us write a full hypermultiplet as Q^a whose two half-hypermultiplet components are given by $Q^1 = Q$ and $Q^2 = \tilde{Q}^*$. The involution τ is then defined by

$$Q^a \mapsto \tau(Q^a) = \sigma_G \otimes \sigma_I (Q^a)^* \,, \tag{3.109}$$

where

$$\sigma_I \begin{pmatrix} Q^1 \\ Q^2 \end{pmatrix} = \begin{pmatrix} -Q^2 \\ Q^1 \end{pmatrix} . \tag{3.110}$$

It is straightforward to check that indeed $\tau^2 = 1$. We can describe a half-hypermultiplet as a hypermultiplet that stays (anti-)invariant under τ, i.e., that is an eigenvector of τ of eigenvalue ± 1. More explicitly, such a hyper is given by $(Q, \pm\sigma_G Q^*)$.

This description of a half-hypermultiplet is, for instance, convenient to find the Lagrangian for a half-hyper Q starting from the Lagrangian of a full hyper. Recall that the

Lagrangian for the hypermultiplet $Q^a = (Q, \tilde{Q}^*)$ coupled to a vector multiplet (V, Φ) is given by

$$\mathcal{L}_{\text{fh}} = \int d^2\theta\, d^2\bar{\theta} \left(Q^\dagger e^V Q + \tilde{Q}^t e^{-V} \tilde{Q}^* \right) + 2\sqrt{2}\, \text{Re} \int d^2\theta \left(\tilde{Q}^t\, \Phi\, Q \right).$$

For pseudo-real representations we can apply the constraint $\tilde{Q} = \pm \sigma_G Q$ to recover the Lagrangian

$$\mathcal{L}_{\text{hh}} = \int d^2\theta\, d^2\bar{\theta} \left(Q^\dagger e^V Q \right) \pm \sqrt{2}\, \text{Re} \int d^2\theta \left(Q^t \sigma_G^t\, \Phi\, Q \right)$$

for a single half-hypermultiplet. Here we rescaled $Q \to \frac{1}{\sqrt{2}} Q$ to give the kinetic term in the Lagrangian a canonical coefficient. We also used

$$\tilde{Q}^t e^{-V} \tilde{Q}^* = Q^t \sigma_G^t e^{-V} \sigma_G Q^* = Q^t e^{V^t} Q^* = Q^\dagger e^V Q\,,$$

since $\sigma_G^{-1} T \sigma_G = -T^t$ for $T \in \mathfrak{g}$. Since we found the Lagrangian $\mathcal{L}_{\text{hh}}$ by starting out with the Lagrangian $\mathcal{L}_{\text{fh}}$ for a full hypermultiplet and then applying the constraint $\tilde{Q} = \pm \sigma_G Q$, it is automatically invariant under $\mathcal{N} = 2$ supersymmetry.

Let us spell this out in some more detail. Substituting the constraint $\tilde{Q} = \pm \sigma_G Q$ in the $\mathcal{N} = 2$ supersymmetry equations yields two identical copies of the supersymmetry variations for the components of Q, which depend on all of the $\mathcal{N} = 2$ supersymmetry parameters. The Lagrangian $\mathcal{L}_{\text{hh}}$ is obviously invariant under these variations. The $SU(2)_R$ symmetry now acts on the vector of complex scalars $(q, \pm \sigma_G q^*)^t$.

As an example, the Lagrangian for the trifundamental $SU(2)$ half-hypermultiplet reads in components

$$\begin{aligned} \mathcal{L}_{\text{trif}} = \int d^2\theta\, d^2\bar{\theta} \left(Q^*_{abc} e^{(V_1)^a{}_{a'}} Q^{a'bc} + Q^*_{abc} e^{(V_2)^b{}_{b'}} Q^{ab'c} + Q^*_{abc} e^{(V_3)^c{}_{c'}} Q^{abc'} \right) \\ \pm \sqrt{2}\, \text{Re} \int d^2\theta \left(\epsilon^{bb'} \epsilon^{cc'} Q_{abc}\, \Phi^{aa'}\, Q_{a'b'c'} + \epsilon^{aa'} \epsilon^{cc'} Q_{abc}\, \Phi^{bb'}\, Q_{a'b'c'} + \epsilon^{aa'} \epsilon^{bb'} Q_{abc}\, \Phi^{cc'}\, Q_{a'b'c'} \right). \end{aligned} \tag{3.111}$$

We can obtain the bifundamental hyper by demoting one of gauge groups to a flavor group. From this perspective it is clear that the bifundamental has an enhanced $SU(2)$ flavor symmetry, as we already knew from general principles. We discuss these and other aspects of the $SU(2)$ trifundamental in detail in appendix D.

3.6.2 Instanton counting for half-hypermultiplets

We now turn to instanton counting for half-hypermultiplets. Also for this purpose it is convenient to use the description of a half-hypermultiplet as a constrained full hypermultiplet. Instanton counting for any $\mathcal{N} = 2$ gauge theory with full hypermultiplets is developed in [28, 53] and spelled out in more detail in, for example, [68] which we reviewed in this chapter. It is performed by topologically twisting the $\mathcal{N} = 2$ gauge theory. The resulting instanton partition function is given by the integral

$$Z^{\text{inst}} = \sum_k q_{\text{UV}}^k \oint\limits_{\mathcal{M}_k^G} e(\mathcal{V}) \tag{3.112}$$

over the ADHM moduli space of instantons $\mathcal{M}_k^G$ for the gauge group G and instanton number k, where the Euler class $e(\mathcal{V})$ encodes the matter content of the gauge theory. More precisely, the vector bundle $\mathcal{V}$ is equal to the space of solutions to the Dirac equation for the chosen matter representation in the self-dual instanton background.

Let us emphasize that the $SU(2)$ R-symmetry is essential for performing the topological twist. We identify the new Lorentz group of the twisted $\mathcal{N} = 2$ theory as

$$L' = SU(2)_L \times \text{diag}(SU(2)_R \times SU(2)_I) \,,$$

where we denoted the R-symmetry group by $SU(2)_I$ to avoid confusion. After twisting the two complex scalars of a full hypermultiplet combine into a Weyl spinor

$$\Psi = (\psi_1, \psi_2) = (q, \tilde{q}^*) \,,$$

i.e., the R-symmetry index turns into a spinor index. The matter part of the theory localizes to solutions of the Dirac equation [19]

$$(i\sigma^\mu \partial_\mu + \sigma^\mu A_\mu)\Psi = 0 \tag{3.113}$$

in the self-dual instanton background determined by the gauge field A, with $\sigma^\mu = (\mathbf{1}, i\sigma^i)$

[19] Although we write down an explicit form of the Dirac equation for a spinor transforming in the fundamental representation of a single gauge group, equation (3.113), as well as the following equations, should be read abstractly and can easily be adapted to hold in a generic setting.

(note that we are in Euclidean signature). These solutions form a vector bundle over the moduli space of self-dual instantons, localizing the path-integral to the integral over the moduli space of instantons (3.112). Actually computing the instanton partition function (3.112) can then be reduced to evaluating the equivariant index of the Dirac operator with respect to a torus action on the ADHM moduli space.

For half-hypermultiplets the twisting works similar, since we have established the R-symmetry invariance of the half-hypermultiplet Lagrangian. Let us start with a twisted full hypermultiplet. Since the R-symmetry indices of the scalars in the full hypermultiplet turn into spinor indices, we can again define the map

$$\Psi \mapsto \tau(\Psi) = \sigma_G \otimes \sigma_I \Psi^* \,. \tag{3.114}$$

As before, the matrices σ_G and σ_I act on the gauge and spinor indices, respectively. In particular,

$$\sigma_I \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} = \begin{pmatrix} -\psi_2 \\ \psi_1 \end{pmatrix}$$

The path integral of the half-hypermultiplet theory localizes onto solutions of the Dirac equation (3.113) that are invariant under τ.

The involution τ indeed maps solutions of the Dirac equation (3.113) to solutions, as can be seen from

$$(i\sigma^\mu \partial_\mu + \sigma^\mu A_\mu)\tau(\Psi) = -\sigma_G \otimes \sigma_I((i\sigma^\mu \partial_\mu + \sigma^\mu A_\mu)\Psi)^* = 0 \,, \tag{3.115}$$

where we have used $\sigma_I^{-1}\sigma^\mu\sigma_I = (\sigma^\mu)^*$ and $\sigma_G^{-1}A_\mu\sigma_G = -A_\mu^*$. We can thus find a basis of the space of solutions to the Dirac equation on which τ acts with eigenvalue ± 1. The relevant solutions for the single half-hypermultiplet are given by those basis elements which all have eigenvalue $+1$ (or all eigenvalue -1) under τ, and form a half-dimensional vector bundle over the moduli space $\mathcal{M}_G^k$ of self-dual instantons.

As an intermezzo, remember that the space of fermionic solutions to the Dirac equation in a pseudo-real representation always admits a real structure. It is not hard to see that the anti-linear involution τ in fact defines this real structure. So let us consider a basis of

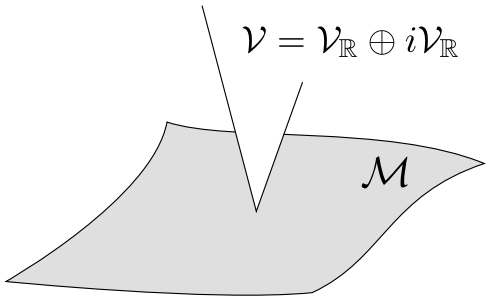


Figure 3.10: The solutions to the Dirac equation in a given representation of the gauge group form a vector bundle $\mathcal{V}$ over the ADHM moduli space $\mathcal{M}$. A pseudo-real representation induces a real structure τ on the vector bundle $\mathcal{V}$ that splits it into two copies $\mathcal{V} = \mathcal{V}_{\mathbb{R}} \oplus i\mathcal{V}_{\mathbb{R}}$. The relevant solutions for a half-hypermultiplet are either parametrized by $\mathcal{V}_{\mathbb{R}}$ or $i\mathcal{V}_{\mathbb{R}}$.

solutions on which τ acts with eigenvalues ± 1. Whereas for a theory with a hyper all solutions with eigenvalue $+1$ or -1 need to be taken into account, the theory with a half-hyper enforces a restriction to the solutions with eigenvalues either all $+1$ or all -1.

Let us name V the total vector space of solutions to the Dirac equation in a given instanton background. Then the real structure τ induces a splitting

$$V = V_R \oplus iV_R \, . \tag{3.116}$$

The vector space V_R (called the real form of τ) consists of solutions with eigenvalue $+1$, whereas iV_R consists of solutions with eigenvalue -1. Indeed, since τ is anti-linear, multiplying a solution $\Psi \in V_R$ by i yields a solution with eigenvalue -1. The real structure τ reduces the group of basis transformations acting on V from $U(d)$ to $SO(d)$, where d is the dimension of V. The action of $SO(d)$ leaves V_R invariant.

The two half-hypermultiplets that make up a hypermultiplet are defined by the two constraints $\tilde{Q} = \pm\sigma_G Q$. One half-hypermultiplet singles out the subspace $V_R \subset V$, and the other the subspace $iV_R \subset V$. So multiplying the solutions of the Dirac equation by i brings us from one half-hypermultiplet to the other.

Instanton partition functions for half-hypermultiplets

Let us summarize the above. Consider an $\mathcal{N} = 2$ gauge theory coupled to a full hypermultiplet in a pseudo-real representation of the gauge group. Its instanton partition function is given by equation (3.112). This is in fact equal to

$$Z^{\text{inst}}_{\text{fh}} = \sum_k q^k_{\text{UV}} \oint\limits_{\mathcal{M}^G_k} e(\mathcal{V}_R \oplus i\mathcal{V}_R), \tag{3.117}$$

since the pseudo-real representation defines a real structure on the complex vector bundle $\mathcal{V}$ of solutions to the Dirac equation. The bundle $\mathcal{V}_R$ is an oriented real bundle, whose Euler class is defined as the Pfaffian (this is only non-trivial when the rank of the bundle is even). The Euler class of its complexification $\mathcal{V} = \mathcal{V}_R \oplus i\mathcal{V}_R$ can then be expressed as the square of the Euler class of $\mathcal{V}_R$,

$$e(\mathcal{V}_R \oplus i\mathcal{V}_R) = e(\mathcal{V}_R)^2. \tag{3.118}$$

This equality continues to hold for the equivariant Euler classes $e_{\mathbf{T}}(\mathcal{V})$ and $e_{\mathbf{T}}(\mathcal{V}_R)$, with respect to the torus action $\mathbf{T} = T_{a_k} \times T_{\phi_i} \times U(1)_{\epsilon_1,\epsilon_2}$ on the ADHM moduli space, where T_{a_k} is the torus of the gauge group, T_{ϕ_i} the torus of the dual group, and the action of $U(1)_{\epsilon_1,\epsilon_2}$ on $\mathbb{R}^4$ defines the Omega-background.

In other words, when the rank of $\mathcal{V}_R$ is even (i.e., when the complex Dirac index of the pseudo-real representation is even), the instanton partition function for a half-hyper theory localizes as

$$Z^{\text{inst}}_{\text{hh}} = \sum_k q^k_{\text{UV}} \oint\limits_{\mathcal{M}^G_k} e(\mathcal{V}_R). \tag{3.119}$$

Since the involution τ commutes with the torus $\mathbf{T}$, we can compute the contribution of a half-hypermultiplet equivariantly by just taking the square-root of the product of weights for the full hypermultiplet theory.

In this manner we can compute the instanton partition function for the $Sp(1)$ trifundamental half-hypermultiplet and the $Sp(1) - SO(4)$ bifundamental half-hypermultiplet. Notice that in both examples the Dirac index is even for any instanton number k. In the

following, we apply this scheme to evaluate instanton partition functions corresponding to $Sp - SO$ quiver gauge theories.

3.6.3 Instanton contribution for the $Sp \times SO$ bifundamental

We continue with finding a contour integrand prescription for the contribution of the $Sp \times SO$ bifundamental matter multiplet. In general, the instanton partition function for a linear quiver gauge theory with gauge group $G = G_1 \times \cdots \times G_P$ can be formulated schematically as

$$Z_k^{\text{inst}} = \int \prod d\phi_i \; \mathbf{z}^k_{\text{gauge}}(\phi_{p,i}, \vec{a}_p) \; \mathbf{z}^k_{\text{bifund}}(\phi_{p,i}, \phi_{q,j}, \vec{a}_p, \vec{a}_q, \mu) \; \mathbf{z}^k_{\text{fund}}(\phi_{p,i}, \vec{a}_p), \tag{3.120}$$

where the three $\mathbf{z}$'s in the integrand refer to the contribution of gauge multiplets, bifundamental and fundamental matter fields, respectively. The only missing ingredient needed to compute the instanton partition function for a linear D_2-quiver is the contribution of the $Sp \times SO$ bifundamental half-hypermultiplet.

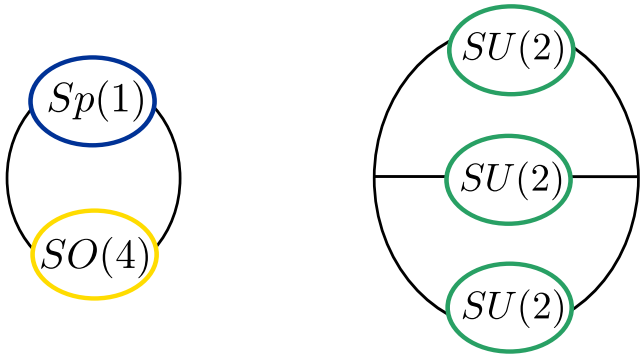


Figure 3.11: On the left: the quiver diagram for a cyclic $Sp(1)/SO(4)$ quiver gauge theory coupled to two $Sp(1) \times SO(4)$ bifundamentals. The bifundamentals do not have a flavor symmetry group. On the right: the corresponding generalized $SU(2)$ quiver including two $SU(2)^3$-trifundamentals.

Let us therefore consider the simplest gauge theory with *two* $Sp(1) \times SO(4)$ bifundamental half-hypermultiplets. This theory is illustrated on the left in Figure 3.11. As is illustrated on the bottom of Figure 3.12, the corresponding UV-curve is a two-punctured torus with a $\mathbb{Z}_2$-twist line running between the punctures. The corresponding A_1 theory is a generalized quiver theory with genus 2.

Let us define the vector bundles $\mathcal{V}_{Sp}$ and $\mathcal{V}_{SO}$ of solutions to the Dirac equation in the

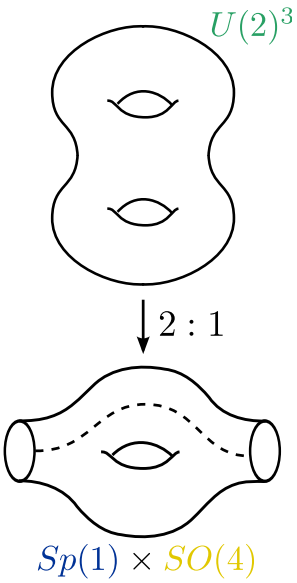


Figure 3.12: The cover and base Gaiotto-curve corresponding to the $Sp(1) \times SO(4)$ and $SU(2)$ quiver gauge theories illustrated in Figure 3.11.

fundamental representation of the Sp and SO gauge group, respectively. Since the double copy of a bifundamental half-hypermultiplet is a full bifundamental hypermultiplet, its contribution to the instanton partition function is given by the usual integral

$$\oint_{\mathcal{M}_{Sp} \times \mathcal{M}_{SO}} e(\mathcal{V}_{Sp} \otimes \mathcal{V}_{SO} \otimes \mathcal{L} \otimes M). \tag{3.121}$$

The integrand is the Euler class of the tensor product $\mathcal{V}_{Sp} \otimes \mathcal{V}_{SO} \otimes \mathcal{L} \otimes M$ over the product $\mathcal{M}_{Sp} \times \mathcal{M}_{SO}$ of the instanton moduli spaces.[20] In this expression, $M \cong \mathbb{C}$ is the flavor vector space and $\mathcal{L}$ the half-canonical line bundle over $\mathbb{R}^4$.

Following the same strategy as we did for a usual hypermultiplet, we obtain a contour integrand

$$\mathbf{z}^{Sp,SO}_{k_1,k_2,db}(\phi, \psi, \vec{a}, \vec{b}, m)$$

corresponding to the double copy of the bifundamental half-hypermultiplet. The parameters $\vec{a}$ and $\vec{b}$ are the Coulomb parameters of the Sp and the SO gauge theory, respec-

[20] Originally, $\mathcal{V}_{Sp}$ is a vector bundle over the instanton moduli space $\mathcal{M}_{Sp}$ and $\mathcal{V}_{SO}$ a vector bundle over the instanton moduli space $\mathcal{M}_{SO}$. However, we define both bundles as vector bundles over the product $\mathcal{M}_{Sp} \times \mathcal{M}_{SO}$, by pulling them back using the projection maps $\pi_{Sp/SO} : \mathcal{M}_{Sp} \times \mathcal{M}_{SO} \to \mathcal{M}_{Sp/SO}$.

tively, whereas m is the mass parameter for the double copy. Since the half-hypermultiplets should be massless we substitute $m = 0$. For precisely this value of the mass, the resulting expression indeed turns out to be a complete square

$$\mathbf{z}^{Sp,SO}_{k_1,k_2,\mathrm{db}}(\phi,\psi,\vec{a},\vec{b},m=0,\epsilon_1,\epsilon_2) = \left(\mathbf{z}^{Sp,SO}_{k_1,k_2,\mathrm{hb}}(\phi,\psi,\vec{a},\vec{b})\right)^2. \tag{3.122}$$

It is thus natural to identify the square-root $\mathbf{z}_{\mathrm{hb}}$ of the double bifundamental with the contribution coming from a $Sp \times SO$-bifundamental half-hypermultiplet. We derive the corresponding contour integrands in the following.

Double $Sp - SO$ half-bifundamental

Let us consider the equivariant index

$$\mathrm{Ind}_{\mathbf{T}} = \int_{\mathbb{C}^2} \mathrm{Ch}_{\mathbf{T}}(\mathcal{E}_{Sp}\otimes\mathcal{E}_{SO}\otimes\mathcal{L}\otimes M)\mathrm{Td}_{\mathbf{T}}(\mathbb{C}^2) = \frac{\mathrm{Ch}_{\mathbf{T}}(\mathcal{E}_{Sp}\otimes\mathcal{E}_{SO}\otimes\mathcal{L}\otimes M)}{(e^{i\epsilon_1}-1)(e^{i\epsilon_2}-1)}. \tag{3.123}$$

Suppose that we add this contribution to the instanton partition function for a quiver with an $Sp(N_1)$ and an $SO(N_2)$ node, which quiver gauge theory do we describe? The contour integrand corresponding to the above equivariant index equals

$$\begin{aligned}\mathbf{z}_{k_1,k_2} \;=\; & \prod_{l=1}^{n_2}\Delta_1(m\pm b_l)\prod_{k=1}^{N_1}\Delta_2(m\pm a_k) \\ & \times\left(\frac{\Delta(m-\epsilon_-)\Delta(m+\epsilon_-)}{\Delta(m-\epsilon_+)\Delta(m+\epsilon_+)}\right)\left(\frac{\Delta_2(m-\epsilon_-)\Delta_2(m+\epsilon_-)}{\Delta_2(m-\epsilon_+)\Delta_2(m+\epsilon_+)}\right)^{\chi_\phi} \\ & \times\Delta_1(m)^{\chi_b}P_2(m)^{\chi_\phi}(m)^{\chi_b\chi_\phi},\end{aligned} \tag{3.124}$$

where $\pm$ is again an abbreviation for a product over both terms. Here $k_1 = 2n_1 + \chi_\phi$ and $N_2 = 2n_2 + \chi_b$, whereas

$$\begin{aligned}
\Delta_1(x) &= \prod_{i=1}^{n_1}(\phi_i^2 - x^2) \\
\Delta_2(x) &= \prod_{j=1}^{k_2}(\psi_j^2 - x^2) \\
\Delta(x) &= \prod_{i,j=1}^{n_1,k_2}\left((\phi_i + \psi_j)^2 - x^2\right)\left((\phi_i - \psi_j)^2 - x^2\right) \\
P_1(x,a) &= \prod_{k=1}^{N_1}(a_k^2 - x^2) \\
P_2(x,b) &= \prod_{l=1}^{n_2}(b_l^2 - x^2).
\end{aligned}$$

What information can we extract from the terms in equation (3.124)? Notice that if we decouple the $SO(N_2)$ gauge group, the contour integrand (3.124) reduces to that for $2n_2$ fundamental $Sp(N_1)$ hypers with masses $m \pm b_l$. If we instead decouple the $Sp(N_1)$ gauge group, the contour integrand reduces that for $2N_1$ fundamental $SO(N_2)$ hypers with masses $m \pm a_k$. In other words, the equivariant index (3.123) contains twice the degrees of freedom of a half-bifundamental coupling between the $Sp(N_1)$ and the $SO(N_2)$ gauge group.

Furthermore, adding this contour integral to the contribution for a pure $Sp(N-1)$ and a pure $SO(2N)$ theory, yields a total contour integral with as many terms in the numerator as in the denominator. The corresponding quiver gauge theory is therefore conformal. This implies that equation (3.123) describes two copies of the $Sp(N_1) - SO(N_2)$ half-bifundamental.

If we specify to the $Sp(1) - SO(4)$ interaction we have $\chi_b = 0$, $N_1 = 1$ and $n_2 = 2$. Coupling this to a $Sp(1)$ and $SO(4)$ gauge group gives a quiver with two $Sp(1) - SO(4)$

half-bifundamentals. Explicitly, the instanton integrand is given by

$$\begin{aligned}\mathbf{z}^{Sp(1),SO(4)}_{k_1,k_2,\text{db}} &= \prod_{i=1}^{n_1}\prod_{l=1}^{2}\left(\phi_i^2-(m+b_l)^2\right)\left(\phi_i^2-(m-b_l)^2\right) \\ &\times\prod_{j=1}^{k_2}\left(\psi_j^2-(m+a)^2\right)\left(\psi_j^2-(m-a)^2\right) \\ &\times\left(\frac{\Delta(m-\epsilon_-)\Delta(m+\epsilon_-)}{\Delta(m-\epsilon_+)\Delta(m+\epsilon_+)}\right) \\ &\times\left(\prod_{l=1}^{2}\left(b_l^2-m^2\right)\frac{\Delta_2(m-\epsilon_-)\Delta_2(m+\epsilon_-)}{\Delta_2(m-\epsilon_+)\Delta_2(m+\epsilon_+)}\right)^{\chi_\phi}.\end{aligned} \tag{3.125}$$

Notice that there are additional poles that involve mass parameter m just like in the case of the $Sp(N_1)\times Sp(N_2)$ bifundamental.

$Sp(1)-SO(4)$ half-bifundamental

The $Sp-SO$ double bifundamental contribution (3.124) turns into a complete square when we choose the mass to be $m=0$. We therefore identify the square-root of this double half-bifundamental contribution for $m=0$ with the contour integral contribution of the half-bifundamental hypermultiplet:

$$\mathbf{z}^{Sp(N_1),SO(N_2)}_{k_1,k_2,\text{db}}(\phi,\psi,a,b,m=0,\epsilon_1,\epsilon_2)=\left(\mathbf{z}^{Sp(N_1),SO(N_2)}_{k_1,k_2,\text{hb}}(\phi,\psi,a,b)\right)^2. \tag{3.126}$$

For the $Sp(1)-SO(4)$ gauge theory the half-bifundamental contour integrand is explicitly given by

$$\mathbf{z}^{Sp(1),SO(4)}_{k_1,k_2,hb} = \prod_{i=1}^{n_1}\left(\phi_i^2-b_1^2\right)\left(\phi_i^2-b_2^2\right)\prod_{j=1}^{k_2}\left(a^2-\psi_j^2\right)\frac{\Delta(\epsilon_-)}{\Delta(\epsilon_+)}\left(b_1b_2\frac{\Delta_2(\epsilon_-)}{\Delta_2(\epsilon_+)}\right)^{\chi_\phi}, \tag{3.127}$$

where $k_1=2n_1+\chi$. There're many different choices of $\pm$ signs for each of the parenthesis in the expression, but we can fix the signs by studying the decoupling limit of one of the gauge groups and compare them with single gauge group computation.

Chapter 4

ADE of $\mathcal{W}$-algebras

In this chapter we discuss the 2d side of the correspondence, namely 2-dimensional conformal field theory with $\mathcal{W}$-algebra symmetry. Conformal symmetry is a powerful symmetry that includes usual Poincare spacetime symmetry and also scale symmetry. Critical phenomena in statistical mechanics are described by conformal field theory. Conformal field theory is absolutely crucial in the study of string theory since the world sheet theory of strings are given as a 2-dimensional conformal field theory. W-algebra is an extension of the Virasoro algebra, which is the underlying symmetry of the 2-dimensional conformal field theory. In the following, we review conformal and W-symmetries in 2 dimensions, especially focusing on the computation of the chiral blocks. For the readers who are interested in learning the subject in depth, we refer to the review paper [94].

4.1 Conformal symmetry in 2 dimensions and Virasoro algebra

In this section, we will review some of the basic properties of 2-dimensional conformal field theory. We refer to [95, 96] for the details. Conformal symmetries are generated by the following action

$$x \rightarrow x' = x + a, \tag{4.1}$$

$$x \rightarrow x' = \Lambda x \quad \Lambda \in SO(1, d-1) \text{ or } SO(d), \tag{4.2}$$

$$x \rightarrow x' = \lambda x \quad \lambda \in \mathbb{R} - \{0\}, \tag{4.3}$$

$$x \rightarrow x' = \frac{x + bx^2}{1 + 2b \cdot x + b^2 x^2}, \tag{4.4}$$

where the first two generate the Poincare group and the latter two are scale transformations and special conformal transformation, respectively. In d-dimensional Euclidean spacetime, the conformal group is $SO(1, d+1)$ and in Lorentzian spacetime, it is $SO(2, d)$. In 2-dimensions, the conformal group becomes infinite dimensional, since any holomorphic function defines a conformal map in 2-dimensions. If we write the coordinates as complex variables $z = x_1 + ix_2, \bar{z} = x_1 - ix_2$, for a given analytic function $f(z)$ on $\mathbb{C}$,

$$z \to f(z)\ ,\ \ \bar{z} \to \bar{f}(\bar{z}). \tag{4.5}$$

defines a conformal transformation so that the two-dimensional metric transforms as

$$ds^2 = dzd\bar{z} = \left|\frac{df}{dz}\right|^2 dzd\bar{z}. \tag{4.6}$$

The basis of such transformation are of the form $z \to z + \epsilon_n(z)$ where $\epsilon_n(z) = -z^{n+1}$ and also its anti-holomorphic pieces with $n \in \mathbb{Z}$. The infinitesimal generators of such transformations are of the form $l_n = -z^{n+1}\partial_z$ and $\bar{l}_n = -\bar{z}^{n+1}\partial_{\bar{z}}$. It is easy to check that the generators satisfy

$$[l_m, l_n] = (m-n)l_{m+n},\ \ [\bar{l}_m, \bar{l}_n] = (m-n)\bar{l}_{m+n},\ \ [l_m, \bar{l}_n] = 0. \tag{4.7}$$

We will see that the above relation is subject to quantum corrections by a central piece. A subalgebra spanned by $\{l_{-1}, l_0, l_1\}$ and $\{\bar{l}_{-1}, \bar{l}_0, \bar{l}_1\}$ generates a group of global conformal transformation $SL(2, \mathbb{C})$.

Suppose there is a field $\phi(z, \bar{z})$ in a conformal field theory which transform under the conformal transformations $z \to f(z), \bar{z} \to \bar{f})(\bar{z})$ as

$$\phi(z, \bar{z}) \to \left(\frac{df}{dz}\right)^h \left(\frac{d\bar{f}}{d\bar{z}}\right)^{\bar{h}} \phi(z, \bar{z}). \tag{4.8}$$

Then it is called a primary field of weight $(h, \bar{h})$. When it transforms in a same way under the global conformal transformation, it is called a quasi-primary field.

Any field theory has a conserved stress-energy tensor $T_{\mu\nu}(x)$ satisfying

$$\nabla^\mu T_{\mu\nu}(x) = 0. \tag{4.9}$$

When the theory is scale invariant[1], from the conservation of the dilatation current $j_\mu = T_{\mu\nu}x^\nu$, we get

$$T^\mu_\mu(x) = 0. \tag{4.10}$$

In the complex coordinates, $z, \bar{z}$, one can show that the only non-vanishing components of the stress energy tensor are $T(z) \equiv T_{zz}$ and $\bar{T}(\bar{z}) \equiv T_{\bar{z}\bar{z}}$ and they are functions of z and $\bar{z}$, respectively. The holomorphic component describes the left-moving modes and the anti-holomorphic component describes the right-moving modes. Since they are identical, we will focus on the holomorphic piece from now.

The operator product expansion (OPE) of the stress-energy tensor with other fields tells you how the fields transform under the conformal group. The primary fields of conformal dimension $(h, \bar{h})$ has the property that OPE with stress-energy tensor is given as

$$T(z)\phi_h(w) = \frac{h\phi_h(w)}{(z-w)^2} + \frac{\partial\phi_h(w)}{z-w} + \dots, \tag{4.11}$$

where we suppressed the right-moving coordinates here. We will frequently suppress right-moving coordinates.

The OPE of the stress-energy tensor with itself is given by

$$T(z)T(w) = \frac{c/2}{(z-w)^4} + \frac{2T(w)}{(z-w)^2} + \frac{\partial T(w)}{z-w} + \dots, \tag{4.12}$$

which shows that $T(z)$ is a quasi-primary field of weight 2. The constant c is called the central charge of a CFT. When evaluating OPE, one should keep in mind that we have to keep track of the ordering. Whenever we use operators, we will use the radial ordering for the time ordering, that is time flows in the $|z|$ direction.

It is useful to expand the stress-energy tensor in terms of Laurent modes L_n. We write the mode expansion as

$$T(z) = \sum_{n\in\mathbb{Z}} L_n z^{-n-2}, \quad L_n = \oint_{\mathcal{C}_0} \frac{dz}{2\pi i} z^{n+1} T(z) \tag{4.13}$$

[1]In 2-dimensions, it has been proved that scale invariance implies conformal invariance. [97] In higher dimensions, it is strongly believed that scale invariance implies the conformal invariance, but there is no proof or counter-example.

where $\mathcal{C}_0$ is the contour around the origin in counterclockwise direction. Now, one can work out the commutation relation among the modes using the OPE (4.12) to obtain

$$[L_m, L_n] = (m-n)L_{m+n} + \frac{c}{12}m(m^2-1)\delta_{m+n,0}. \tag{4.14}$$

This is the Virasoro algebra. We see that when the central charge c is zero, it reduces to the classical algebra of conformal transformations (4.7).

4.1.1 Representations of Virasoro Algebra

A physical state of conformal field theory has to be in the representation of the Virasoro algebra. One can construct a highest weight module starting from the highest weight state satisfying

$$L_0|h,c\rangle = h|h,c\rangle, \quad L_n|h,c\rangle = 0. \tag{4.15}$$

There is a one-to-one correspondence between the highest weight state $|h,c\rangle$ and a primary operator $\phi(z)$ of conformal weight h

$$|h,c\rangle = \lim_{z\to 0}\phi(z)|0\rangle, \tag{4.16}$$

where the vacuum state $|0\rangle$ is defined to be the state with $L_m|0\rangle = 0$ for $m \geqslant -1$. Hermicity of the stress-energy tensor $T(z)$ requires that $L_m^\dagger = L_{-m}$. The highest weight module consists of the descendant states

$$L_{-k_1}L_{-k_2}\cdots L_{-k_m}|h,c\rangle, \quad k_i > 0, \tag{4.17}$$

which is called the Verma module $M(h,c)$. The Verma module can be decomposed in terms of L_0-eigenspace

$$M(h,c) = \bigoplus_{N\geqslant 0} M(h,c)_{(N)}, \tag{4.18}$$

where

$$M(h,c)_{(N)} = \{v \in M(h,c) | L_0 v = (h+N)v\}. \tag{4.19}$$

This means that the states in $M(h,c)_{(N)}$ are of the form

$$L_{-k_1} \cdots L_{-k_m} |h,c\rangle, \quad \sum_{i=1}^{m} k_i = N, \;\; k_1 \geqslant k_2 \geqslant \cdots \geqslant k_m > 0. \tag{4.20}$$

The dimension of the space $M(h,c)_{(N)}$ is given by the Euler function $p(N)$ which counts the number of partitions of N. The number of partition function is given in arms of the generating function

$$\frac{1}{\prod_{n=1}^{\infty}(1-q^n)} = \sum_{N=0}^{\infty} p(N) q^N. \tag{4.21}$$

The positivity of the norm of physical states puts certain constraint. Let's consider a $p(N) \times p(N)$ matrix $\mathcal{M}(h,c)_N$ of inner products of the vectors in $M(h,c)_{(N)}$

$$\langle h,c | L_{i_m} \cdots L_{i_1} L_{-j_1} \cdots L_{-j_n} | h,c \rangle, \tag{4.22}$$

where $i_1 + \cdots i_m = j_1 + \cdots + j_n = N$. This matrix is called the Kac matrix and its determinant is given by

$$\det \mathcal{M}(h,c)_N = \prod_{rs \leqslant N} (h - h_{r,s}(m))^{p(N-rs)} \tag{4.23}$$

where the central charge is given by a parameter m satisfying

$$c = 1 - \frac{6}{m(m+1)} \;\text{ or }\; m = -\frac{1}{2} \pm \frac{1}{2}\sqrt{\frac{25-c}{1-c}}, \tag{4.24}$$

and $h_{r,s}(m)$ is given by

$$h_{r,s}(m) = \frac{[(m+1)r - ms]^2 - 1}{4m(m+1)}, \quad r,s \in \mathbb{Z}_{\geqslant 0}. \tag{4.25}$$

In general, the Verma module $M(h,c)$ is not irreducible. There may be a set of null vectors in the Verma module, which are orthogonal to all the states in $M(h,c)$. The physical

states should correspond to the quotients of the Verma module by the set of null vectors. When the highest weight of a Verma module is $h = h_{r,s}$, then there is a null vector at level rs. An operator with conformal weight $h = h_{r,s}$ is called a degenerate operator.

The positivity of the norm (or unitarity) restricts the value of central charge and highest weight to be

$$c \geqslant 1, \quad h \geqslant 0 \tag{4.26}$$

or

$$c = 1 - \frac{6}{m(m+1)}, \ \ h = h_{r,s}(m), \quad m = 2, 3, 4, \cdots, \ \ 1 \leqslant r \leqslant m-1, \ \ 1 \leqslant s \leqslant m. \tag{4.27}$$

Note that in the latter case, only finite number of highest weights allowed.

4.1.2 Free Field Theory

Here, we discuss the conformal field theory with single free boson to illustrate the free field construction of the simplest $\mathcal{W}$-algebra, namely the Virasoro algebra. One can write a mode expansion of the free boson by

$$i\partial\varphi(z) = \sum_{n\in\mathbb{Z}} \alpha_n z^{-n-1}, \tag{4.28}$$

with OPE given by

$$\partial\varphi(z)\partial\varphi(w) \sim \frac{-1}{(z-w)^2}. \tag{4.29}$$

One can derive the commutation relation of modes from above OPE to get

$$[\alpha_m, \alpha_n] = m\delta_{m+n,0}. \tag{4.30}$$

This defines the $U(1)$ affine Lie algebra. The stress energy tensor of the theory is given by

$$T(z) = -\frac{1}{2}(\partial\varphi)^2, \tag{4.31}$$

in terms of the free field. In terms of modes, we get in particular

$$L_0 = \frac{1}{2}\alpha_0^2 + \sum_{n\geqslant 1} \alpha_{-n}\alpha_n. \tag{4.32}$$

We also see that the free boson theory has central charge $c = 1$. One can easily write down a Verma module of the theory. Let's define the highest weight vector $|a\rangle$ as

$$\alpha_0|a\rangle = a|a\rangle, \;\; \alpha_n|a\rangle = 0, \;\; n > 0, \tag{4.33}$$

and the highest weight module is spanned by the states of the form

$$\alpha_{-k_1}\cdots\alpha_{-k_n}|a\rangle, \;\; k_1 \geqslant k_2 \geqslant \cdots \geqslant k_n > 0. \tag{4.34}$$

It is easy to see from (4.32) that the highest weight vector $|a\rangle$ has conformal weight $\frac{1}{2}a^2$. The corresponding vertex operator is given by $V_a(z) = e^{ia\varphi(z)}$.

So far, the central charge was fixed to be $c = 1$. We can modify the theory slightly by adding a background charge Q at $z = \infty$. This is due to Feigin-Fuchs [98]. It does not change the OPE of the fields, but changes the stress-energy tensor to

$$T(z) = -\frac{1}{2}(\partial\varphi)^2 + iQ\partial^2\varphi. \tag{4.35}$$

Now, the central charge becomes $c = 1 - 12Q^2$ so that it screens the central charge of the original system. It also changes the conformal weight of the vertex operators $V_a(z)$ by

$$h(a) = a\left(\frac{a}{2} - Q\right). \tag{4.36}$$

We will use analogous construction to realize the $\mathcal{W}$-algebras of various types. One advantage of free boson construction in the 4d/2d correspondence is that the relation between the Coulomb branch parameters of 4d gauge theory is directly given by the zero modes of the free bosons. It gives much simpler and direct relation between the two without using indirect map of parameters between the Coulomb branch parameters and the weights of the $\mathcal{W}$-primary states.

4.1.3 Conformal Block

Another powerful consequence of the conformal symmetry is that n-point correlation function of a CFT is fully determined by the set of 3-point functions of primary fields and the n-point conformal block [99, 100]. Conformal blocks are fully determined by the representation of Virasoro algebra. For example, 4-point function can be written as

$$\langle V_1(\infty)V_2(1)V_3(q)V_4(0)\rangle = \sum_i C_{12h_i}C_{h_i34}\mathcal{F}_{1,2,3,4;i}(q)\bar{\mathcal{F}}_{1,2,3,4;i}(\bar{q}), \tag{4.37}$$

where C_{abc} is the 3-point function of the vertex operators a, b, c located at $z = 0, 1, \infty$, and we are summing over the primary fields labeled by i with weights h_i. The conformal block $\mathcal{F}$ is given by

$$\mathcal{F} = \frac{\sum_{\vec{I},\vec{J}}\langle V_1(\infty)V_2(1)L_{-\vec{I}}V_h(0)\rangle K^{-1}_{IJ}\langle L_{-\vec{J}}V_h(\infty)V_3(q)V_4(0)\rangle}{\langle V_1(\infty)V_2(1)V_h(0)\rangle\langle V_h(\infty)V_3(1)V_4(0)\rangle} \tag{4.38}$$

where $L_{-\vec{I}} = L_{-i_1}L_{-i_2}\cdots L_{-i_m}$ for $\vec{I} = (i_1, i_2, \cdots, i_m)$ and $K_{\vec{I},\vec{J}}$ is the Kac matrix. Note that this definition differs slightly from the usual definition, as we have not divided out a factor $q^{h-h_3-h_4}$. Our definition will be slightly more convenient to work with when we compare with the instanton partition function.

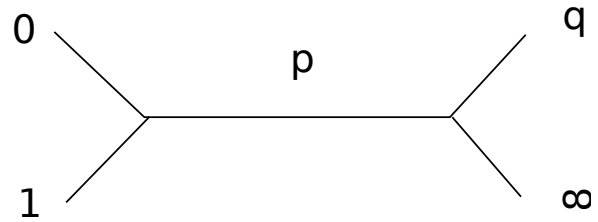


Figure 4.1: The 4-point conformal block.

One can obtain arbitrary n-point conformal block from sewing the elementary building blocks. The building blocks are 3-point functions of the form

$$R_M(h_1, h_2, h_3) = \frac{\langle V_1V_2L_{-\vec{m}}V_3\rangle}{\langle V_1V_2V_3\rangle}, \tag{4.39}$$

$$S_{M,N}(h_1, h_2, h_3) = \frac{\langle L_{-\vec{m}}V_1V_2L_{-\vec{n}}V_3\rangle}{\langle V_1V_2V_3\rangle}. \tag{4.40}$$

One can glue two vertices by putting the propagator which is given by the inverse of Kac

matrix K^{-1}. For example, the 4-point function given as above can be written as

$$R_M(h_1, h_2, h)K^{-1}(h)R_M(h, h_3, h_4). \tag{4.41}$$

One can also have a building block with all three vertices being descendant states. We will discuss this in section 5.5.1 where we relate it to the trifundamental half-hypermultiplet appears in Sicilian gauge theories. We will also see that it is possible to generalize the concept of conformal block to $\mathcal{W}$-algebras in the next section.

4.2 $\mathcal{W}$-algebras

A $\mathcal{W}$-algebra is an extension of the Virasoro algebra, with additional conserved currents $W^{(s_i)}(z)$ which are the quasi-primaries of conformal dimensions s_i. It includes $W^{(2)}(z) = T(z)$ and the OPE of $\mathcal{W}$-currents has be closed, which means that every terms should be written in terms of $W^{(s_i)}(z)$ and their derivatives. One can make a formal definition of the $\mathcal{W}$-algebras using the notion of meromorphic conformal field theory [94].

The simplest example of the $\mathcal{W}$-algebra called $\mathcal{W}_3$ is first written by Zamolodchikov [101]. It consists of two currents $T(z)$ and $W(z)$ which are quasi-primaries of dimension 2 and 3. The OPE of dimension 3 currents are

$$\begin{aligned} W(z)W(w) \quad \sim \quad & \frac{c/3}{(z-w)^6} + \frac{2T(w)}{(z-w)^4} + \frac{\partial T(w)}{(z-w)^3} \\ & + \frac{1}{(z-w)^2}\left(2\beta\Lambda(w) + \frac{3}{10}\partial^2 T(w)\right) \\ & + \frac{1}{(z-w)^2}\left(\beta\Lambda(w) + \frac{1}{15}\partial^2 T(w)\right), \end{aligned} \tag{4.42}$$

where

$$\Lambda(w) = (TT)(w) - \frac{3}{10}\partial^2 T(w), \quad \beta = \frac{16}{22+5c}. \tag{4.43}$$

The form of OPE is fixed by requiring the crossing symmetry of the correlation function. In terms of Laurent modes

$$T(z) = \sum_{n\in\mathbb{Z}} \frac{L_n}{z^{n+2}}, \quad W(z) = \sum_{n\in\mathbb{Z}} \frac{W_n}{z^{n+3}}, \tag{4.44}$$

the commutation relations are given by

$$[L_m, W_n] = (2m - n)W_{m+n} \tag{4.45}$$

and

$$\begin{aligned}[W_m, W_n] &= \frac{c}{360} m(m^2-1)(m^2-4)\delta_{m+n,0} \\ &\quad + (m-n)\left[\frac{1}{15}(m+n+3)(m+n+2) - \frac{1}{6}(m+2)(n+2)\right] L_{m+n} \\ &\quad + \beta(m-n)\Lambda_{m+n}\end{aligned} \tag{4.46}$$

where

$$\Lambda_m = \sum_{n\in\mathbb{Z}} (L_{m-n}L_n) - \frac{3}{10}(m+3)(m+2)L_m. \tag{4.47}$$

One can use these commutation relations and construct the Verma module corresponding to the primary states with respect to both $T(z)$ and $W(z)$. One can start with a state with

$$L_0|h,w\rangle = h|h,w\rangle, \quad W_0|h,w\rangle = w|h,w\rangle, \quad L_n|h,w\rangle = W_n|h,w\rangle = 0, \ \ (n>0) \tag{4.48}$$

and then form a W-descendeant states by applying W_{-k} or $L_{-k'}$ successively

$$W_{-k_1}\cdots L_{-k'_1}\cdots|h,w\rangle, \ \ k_i, k'_j \in \mathbb{N}. \tag{4.49}$$

These states form the highest weight module of the $\mathcal{W}_3$-algebra.

Even though it is possible to directly construct of $\mathcal{W}$-algebra by starting with currents with higher conformal dimension and try to make the algebra close, it is very hard to construct the $\mathcal{W}$-algebra with many high-dimensional currents. The direct construction has been done only up to the case with 3 generators [102, 103], but not beyond.

There is no complete classification all $\mathcal{W}$ algebras, but many examples are known and have been studied. One particular family of examples are the the so-called Casimir algebras, which are based on simply laced Lie algebras. Its generators are constructed from the $\mathfrak{g}$-invariant contractions of the current field $J(z)$ of the affine Lie algebra $\mathfrak{g}$. The series of $\mathcal{W}_N$-algebras, for instance, is related to the A_N Lie algebras. In [91] $\mathcal{W}_N$-blocks have

been related to the instanton partition functions corresponding to $U(N)$ gauge theories. It is natural to expect that also the other Casimir algebras appear as dual descriptions of instanton counting.

There is a more systematic method of obtaining the $\mathcal{W}$-algberas called the quantum Drinfeld-Sokolov (DS) reduction [104, 105, 106]. It starts from an affine Lie algebra $\hat{\mathfrak{g}}$ and impose some constraints on the generators by the BRST procedure. The reduced algebra $\mathcal{W}(\mathfrak{g})$ is given by the BRST cohomology.

4.2.1 Chiral Blocks and Twisted Representations

Since the spectrum of the CFT decomposes into representations of the $\mathcal{W}$ algebra, we can use generalized Ward identities to relate correlation functions of ($\mathcal{W}$-)descendant fields to correlation functions of ($\mathcal{W}$-)primary fields. In the case of the Virasoro algebra, we can always reduce them to functions of primary fields only. For general $\mathcal{W}$-algebras this is only possible if one restricts to primary fields on which the $\mathcal{W}$-fields satisfy additional null relations.

We can make use of this property by computing chiral blocks. For a given configuration of a punctured Riemann surface, we define the chiral block by picking a representation ϕ for every tube, inserting the projector on the representation $P_{\mathcal{H}_\phi}$ at that point in the correlator, and dividing by the product of all three point functions of the primary fields. By the above remarks the result is then independent of the three-point functions of the theory, i.e., it only depends on the kinematics of the theory.

In the simplest configuration, the sphere with four punctures, the chiral block is thus given by

$$\mathcal{F} = \frac{\langle V_1(\infty)V_2(1)P_{\mathcal{H}_\phi}V_3(q)V_4(0)\rangle}{\langle V_1(\infty)V_2(1)|\phi\rangle\langle\phi|V_3(1)V_4(0)\rangle}\,. \tag{4.50}$$

On the gauge theory side it corresponds to the partition function including the perturbative contribution. Also, let us take the convention in what follows that whenever we write a correlator, we assume that it is divided by the appropriate primary three point functions. The projector is usually written as

$$P_{\mathcal{H}_\phi} = \sum_{I,J} |\phi_I\rangle\langle\phi_J|(K^{-1})_{I,J} \tag{4.51}$$

where $I = (i_1, i_2, \ldots)$ denotes the $\mathcal{W}$ descendants, such that $\phi_I = W_{-i_1} W_{-i_2} \cdots \phi$ is a $\mathcal{W}$ descendant. K is the inner product matrix and the sum runs over all $\mathcal{W}$ descendants of ϕ.

A representation ϕ of $\mathcal{W}$ is called untwisted if it is local with respect to $\mathcal{W}$, so that one can freely move $\mathcal{W}$-fields around it. The $\mathcal{W}$-fields then have integer mode representations around that representation.

More generally, the $\mathcal{W}$-fields can pick up phases when circling around ϕ, so that the correlation function has a branch cut extending from ϕ. Such ϕ are called twisted representations. Because of the phase α picked up by the $\mathcal{W}$-fields, their modes are no longer integer, but given by $r \in \mathbb{Z} + \alpha$.

A particular case of twisted representations can appear when the $\mathcal{W}$ has an outer automorphism such as a $\mathbb{Z}_N$-symmetry. Let us say that by circling around a twisted representation the algebra $\mathcal{W}$ gets mapped to an image under $\mathbb{Z}_N$, such as

$$W_k \mapsto W_{k+1} , \qquad k = 1, \ldots N . \tag{4.52}$$

By choosing linear combinations $W^{(k)}$ of the modes W_k that are eigenvectors under the automorphisms, the $W^{(k)}$ indeed pick up phases $2\pi ik/N$. For $N = 2$, the case that we are interested in below, the $\mathcal{W}$-algebra thus decomposes into generators W^+, W^- of integer and half-integer modes, respectively.

Let us finally note that in the case of Liouville theory the $\mathcal{W}$-algebra is simply the Virasoro algebra. Example of conformal field theories with bigger $\mathcal{W}$-algebras are Toda theories.

4.3 Free field realization of $\mathcal{W}$-algebras

We will construct our $\mathcal{W}$-algebras of type $\Gamma = A_n, D_n$ and E_n from free fields using the quantum version of the Drinfeld-Sokolov reduction. For our purpose this boils down to the following steps: We introduce the free bosons $\vec{\varphi}$ living in the weight space of the semisimple Lie group Γ of rank n. We normalize the OPE of the free bosons so that $J^k = i\partial\varphi^i$ satisfies

$$J^k(z) J^l(w) \sim \frac{\delta^{kl}}{(z-w)^2} . \tag{4.53}$$

We assume Γ is simply-laced, and normalize the roots to have squared length 2. The W-algebra is then given by the centralizer of the screening charges $Q_i^{\pm}$ defined as follows: For each simple root $\vec{\alpha}_i$ there are charges

$$Q_i^{\pm} = \oint \tilde{s}_i^{\pm} dz \,, \qquad \tilde{s}_i^{\pm} = \exp(b^{\pm 1}\vec{\alpha}_i \cdot \vec{\varphi}) \,. \tag{4.54}$$

Then we find the operators constructed from the bosons $\vec{\varphi}$ which commute with all the screening charges. It is known that there are n independent generators with weights w_i, tabulated in Table. 2.1. The weight 2 operator is the energy-momentum tensor given by

$$T(z) = W^{(2)}(z) = -\frac{1}{2}(\partial\vec{\varphi} \cdot \partial\vec{\varphi})(z) + Q\vec{\rho} \cdot \partial^2\vec{\varphi}(z) \,, \tag{4.55}$$

where $\vec{\rho}$ is the Weyl vector and $Q = b + 1/b$.

Our computations were all done in Mathematica using the package `OPEdefs.m` developed by K. Thielemans [107]. In the following, all composite operators are understood to be OPE-normal ordered.

4.3.1 Simply laced W-algebras

4.3.1.1 A_n

For A_n we make use of the quantum Miura transform [108, 109]. Let $\vec{e}_i, i = 1, \ldots n+1$ be the weights of the fundamental $n+1$ dimensional representation of A_n, as in Appendix A. We then construct a set of generators $U^{(k)}(z)$ from

$$R^{(n+1)}(z) = -\sum_{k=0}^{n+1} U^{(k)}(z)(Q\partial)^{n+1-k} = (Q\partial - \vec{e}_1 \cdot \partial\vec{\varphi}(z)) \cdots (Q\partial - \vec{e}_{n+1} \cdot \partial\vec{\varphi}(z)) \,. \tag{4.56}$$

One can show that the singular part of the OPE of $R^{(n+1)}$ with $\tilde{s}_i^{\pm}$ is a total derivative, which means that the $U^{(k)}(z)$ are indeed in the centralizer of the screening charges. Since $U^{(1)}(z)$ vanishes, the remaining set of generators has the correct dimensions w_i. One can also show that they are independent, from which we conclude that we have a full set of generators. Note however that these generators are certainly not unique, as we can always add suitable products and derivatives of lower order generators.

4.3.1.2 D_n

For D_n we introduce $\vec{e}_i, i = 1, \dots n$ such that $\pm\vec{e}_i$ form the weights of the fundamental $2n$ dimensional representation. We repeat the construction of $R^{(n)}(z)$ as

$$R^{(n)}(z) = -\sum_{k=0}^{n} V^{(k)}(z)(Q\partial)^{n-k} = (Q\partial - \vec{e}_1 \cdot i\partial\vec{\varphi}(z)) \cdots (Q\partial - \vec{e}_n \cdot i\partial\vec{\varphi}(z)) \,. \tag{4.57}$$

In this case however it turns out that only $V^{(n)}(z)$ commutes with the screening charges. To obtain the rest of the generators, we can take the OPE of $V^{(n)}$ with itself,

$$V^{(n)}(z)V^{(n)}(w) = \frac{a_n}{(z-w)^{2n}} + \sum_{k=1}^{n-1} \frac{a_{n-k}}{(z-w)^{2(n-k)}} \left(U^{(2k)}(z) + U^{(2k)}(w)\right) \,, \tag{4.58}$$

where we choose the normalization $a_k = \prod_{j=1}^{k-1}(1 - 2j(2j+1)Q^2)$. Once again one can show that $V^{(n)}$ and the $U^{(2k)}$ are independent, which gives us a set of n generators of the correct weights [110, 111].

4.3.1.3 E_6

For W-algebra of type E_6, the concise Miura transforms such as (4.56) for type A_n and (4.57) for type D_n are not known. Thus, one is forced to construct the commutants of the screening operators (4.54) directly. Note that an operator $O(z)$ commute with the screening charge Q_i if and only if it has the form

$$O(z) = \sum_a X_a(T_j)(z) Y_a(\partial\varphi^1, \dots, \partial\varphi^{i-1}, \partial\varphi^{i+1}, \dots, \partial\varphi^n)(z) \tag{4.59}$$

where

- X_a and Y_a stand for normal-ordered polynomials, possibly with derivatives, constructed from their respective arguments,
- T_i is the energy momentum tensor for the boson $\varphi_i = \vec{\alpha}_i \cdot \vec{\varphi}$ along the root $\vec{\alpha}_i$, i.e. $T_i(z) = -\partial\varphi_i\partial\varphi_i(z) + Q\partial^2\varphi_i/2$ with central charge $1 + 6Q^2$,
- and $\varphi^i = \vec{w}^i \cdot \vec{\varphi}$ where $\vec{w}^i$ are the fundamental weights, so that Y_a are constructed from bosons perpendicular to the root $\vec{\alpha}_i$.

Therefore, $O(z)$ is in the W-algebra of type E_n if and only if $O(z)$ has the decomposition (4.59) for each simple root $\vec{\alpha}_j$. Details of the construction of the $W(E_6)$ algebra are presented in Appendix G.

4.3.2 Twisted sectors of the simply-laced W-algebras

We need to consider the sectors of the simply-laced algebras twisted by their outer automorphisms to compare with the instanton partition function for non-simply-laced gauge groups.[2]

The $\mathbb{Z}_r$ outer automorphism acts on the simple roots as shown in Fig. 2.6. This induces an action on the free bosons $\vec{\varphi}$. Since this is also a symmetry of the W-algebra, and we can consider a $\mathbb{Z}_r$-twisted state. In practice we pick new linear combinations of bosons $\tilde{\varphi}_j$ which are eigenstates of the $\mathbb{Z}_r$ action:

$$\partial\tilde{\varphi}_j \mapsto e^{2\pi i k_j/r}\partial\tilde{\varphi}_j\,, \qquad k_j = 0, \dots r-1\,. \tag{4.60}$$

Their modes are therefore in $\mathbb{Z} + k_j/r$. The set of generators $W^{(w_i)}$ therefore decompose into generators $W^{(\hat{w}_i)}$ with integer modes, and generators $W^{(\check{w}_i)}$ with non-integer modes. The former correspond to the invariants of the non-simply-laced gauge group that we want to construct, and the states of the lowest level in the twisted Verma module is generated by $W^{(\check{w}_i)}_{-1/r}$.

The actions on the W-generators are given explicitly as follows:

- The $\mathbb{Z}_2$ action on $W(A_n)$ maps $\tilde{U}^{(k)} \to (-1)^k\tilde{U}^{(k)}$ where $\tilde{U}^{(k)}$ is a suitable redefinition of $U^{(k)}$ defined in (4.56). For example, in the case of $W(A_5)$,

$$\begin{aligned}
&\tilde{U}^{(2)} = U^{(2)}, && \tilde{U}^{(3)} = U^{(3)} - 2Q\partial\tilde{U}^{(2)},\\
&\tilde{U}^{(4)} = U^{(4)} - \frac{3}{2}Q\partial\tilde{U}^{(3)}, && \tilde{U}^{(5)} = U^{(5)} - Q\partial\tilde{U}^{(4)} + Q^3\partial^3\tilde{U^{(2)}},\\
&\tilde{U}^{(6)} = U^{(6)} - \frac{1}{2}Q\partial\tilde{U}^{(5)} + \frac{1}{4}Q^3\partial^3\tilde{U}^{(3)}. &&
\end{aligned} \tag{4.61}$$

  The redefined currents are determined by requiring W-generators to have definite Z_2 eigenvalues.

[2]The W-algebras for non-simply-laced groups B_n, C_n, G_2, F_4 can also be determined via Drinfeld-Sokolov reduction, see e.g., [112], but that is not what we use.

- The $\mathbb{Z}_2$ action on $W(D_n)$ maps $V^{(n)} \to -V^{(n)}$, $U^{(2k)} \to U^{(2k)}$ where $V^{(n)}$ and $U^{(2k)}$ are defined in (4.57), (4.58).

- The $\mathbb{Z}_2$ action on $W(E_6)$ maps $W^{(n)} \to (-1)^n W^{(n)}$, where $W^{(n)}$ is defined in Appendix G.

- The $\mathbb{Z}_3$ action on $W(D_4)$ is given is induced from the $\mathbb{Z}_3$ action on the four free bosons (4.60). Explicitly, we first define $W^{(2,4,6)}$ and $\tilde{W}^{(4)}$ by

$$\begin{aligned} W^{(2)} &= U^{(2)}, \qquad & W^{(4)} &= U^{(4)} + \frac{1}{4}\partial^2 W^{(2)} - \frac{1}{2}U^{(4)}W^{(2)} + \frac{9Q^2(1-4Q^2)}{2-12Q^2}\partial^2 W^{(2)}, \\ \tilde{W}^{(4)} &= \sqrt{3}V^{(4)}, \qquad & W^{(6)} &= U^{(6)} + \frac{1}{2}\partial^2 W^{(4)} - \frac{1}{3}W^{(4)}W^{(2)} + \frac{5}{6}Q^2\partial^2 W^{(4)} \,. \end{aligned} \tag{4.62}$$

  where $V^{(4)}$ and $U^{(2,4,6)}$ are defined in (4.57), (4.58). Then $W^{(2,6)}$ is invariant and $W^{(4)} + i\tilde{W}^{(4)} \to e^{2\pi i/3}(W^{(4)} + i\tilde{W}^{(4)})$ under the $\mathbb{Z}_3$ action.

4.3.3 Basic properties of the Verma module

Before continuing, we recall two basic features of the untwisted and twisted Verma module. The first is their Weyl invariance. The zero modes of the W-generators are given in terms of the zero modes $\vec{J}_0$ of the free bosons $\vec{\varphi}$. If we define $\vec{a}$ by

$$\vec{a} = \vec{J}_0 - Q\vec{\rho} \tag{4.63}$$

then the zero modes of W-generators are Weyl-invariant polynomials of $\vec{a}$. For twisted sectors, the twisted bosons do not have zero modes. Correspondingly, $\vec{a}$ and $\vec{J}_0$ are invariant under the twist; $\vec{\rho}$ is automatically invariant.

The second is the Kac determinant at the lowest level, which we detail in Appendix F. Here we just quote the result, which is given by

$$(\text{Kac determinant at level } -1/r) \propto \prod_{\vec{\gamma}\in\Delta_l} (\vec{\gamma}\cdot\vec{a} + Q) \tag{4.64}$$

where

$$\Delta_l = \begin{cases} \Delta, & (r=1) \\ \{\vec{\alpha} + o(\vec{\alpha})\}, & (r=2) \\ \{\vec{\alpha} + o(\vec{\alpha}) + o^2(\vec{\alpha})\}. & (r=3) \end{cases} \tag{4.65}$$

Here $\vec{\alpha}$ runs over roots with $\vec{\alpha} \neq o(\vec{\alpha})$. As explained in Appendix A.2, Δ_l can be identified with the set of long roots of G, which is the S-dual of the $\mathbb{Z}_r$ invariant subgroup of Γ.

Chapter 5

$ABCDEFG$ of Instantons and ADE of $\mathcal{W}$-algebras

5.1 Introduction

It has been almost twenty years since it was realized that the quantum dynamics of 4d $\mathcal{N} = 2$ gauge theory is encoded in the classical geometry of a two-dimensional Riemann surface Σ, called the Seiberg-Witten curve [15, 16]. The curve Σ was originally introduced as an auxiliary construct, but it was later recognized [45, 33, 25, 31] that Σ is a branched covering of another two-dimensional surface C called Gaiotto curve or UV-curve on which a 6d $\mathcal{N} = (2, 0)$ theory is compactified to obtain the 4d $\mathcal{N} = 2$ theory.

From this point of view, it is not surprising that the 2d quantum dynamics on Σ or C have some bearing on the 4d gauge dynamics. Indeed the role of 2d free bosons in this setup has long been recognized: see e.g. [113, 114, 115] for the identification of the gravitational factor on the u-plane [116, 117] as the one-loop determinant of a free boson on Σ, and [22, 23, 24] for a more general analysis.

Compactifying the six-dimensional superconformal $(2, 0)$ theory of type ADE on either a two-dimensional Riemann surface or a four-manifold, suggests that there should be a correspondence between the following two systems. The first system is a four-dimensional superconformal $\mathcal{N} = 2$ gauge theory with ADE gauge group and whose UV-curve is equal to the Riemann surface. The second system is a two-dimensional field theory that lives on the Riemann surface and should be characterized by an ADE type.

Remember that the tubes of the UV-curve are associated to the ADE gauge group of the $\mathcal{N} = 2$ gauge theory, and the punctures on the UV-curve to matter. Decomposing the

UV-curve into pairs of pants suggests that the marginal gauge couplings τ_{UV} should be identified with the sewing parameters $q = \exp(2\pi i \tau_{\text{UV}})$ of the curve. The symmetry of the two-dimensional theory should be related to the ADE gauge group. Furthermore, two-dimensional operators that are inserted at punctures of the UV-curve should encode the flavor symmetries of the corresponding matter multiplets.

A particular instance of such a 4d–2d connection was discovered in [26]. It was found that instanton partition functions in the Ω-background $\mathbb{R}^4_{\epsilon_1,\epsilon_2}$ for linear and cyclic $U(2)$ quiver gauge theories are closely related to Virasoro conformal blocks of the pair of pants decomposition of the corresponding UV-curves. In this so-called AGT correspondence the central charge of the Virasoro algebra is determined by the value of the two deformation parameters ϵ_1 and ϵ_2 as

$$c = 1 + \frac{6(\epsilon_1 + \epsilon_2)^2}{\epsilon_1 \epsilon_2}. \tag{5.1}$$

The conformal weights of the vertex operators at the punctures of the UV-curve are specified by the masses of the hypermultiplets in the quiver theory, and the conformal weights of the fields in the internal channels are related in the same way to the Coulomb branch parameters.

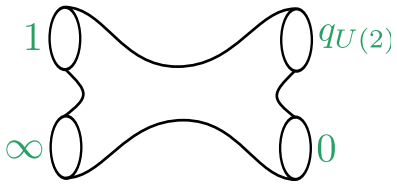


Figure 5.1: The AGT correspondence relates the instanton partition function of the $U(2)$ gauge theory coupled to four hypermultiplets to a Virasoro conformal blocks on the four-punctured sphere with vertex operator insertions at the four punctures.

More precisely, the instanton partition function can be written as the product of the Virasoro conformal block times a factor that resembles a $U(1)$ partition function. The interpretation for this is that the single Casimir of degree 2 of the subgroup $SU(2) \subset U(2)$ corresponds to the energy-stress tensor of the CFT, and the overall $U(1)$ factorizes. This picture was made more explicit in [118].

For general gauge groups with more Casimirs we expect the CFT to have a bigger symmetry group. Since the symmetries of a CFT are captured by its so-called $\mathcal{W}$ or chiral

algebra, we expect that the Nekrasov partition functions should be identified with $\mathcal{W}$-blocks instead of Virasoro blocks. For $SU(N)$, this picture was proposed and checked in [91]. The relation between Hitchin systems and $\mathcal{W}$-algebras was also discussed in [119].

In this chapter, we will study the correspondence for every gauge groups for the pure Yang-Mills case, and for the $Sp - SO$ linear quivers, and finally to the $SU(2)$ non-linear or Sicilian quiver gauge theories.

5.2 Pure YM case

The essence of the 4d/2d correspondence is that the correlator of the 2d Toda theory equals the Nekrasov partition function of the 4d theory with Ω deformation parameters $\epsilon_{1,2}$, and that the vevs of its W-currents $W^{(w_i)}(z)$ become the world-volume fields $\phi^{(w_i)}$ determining the Seiberg-Witten curve:

$$\lim_{\epsilon_{1,2}\to 0} \langle W^{(w_i)}(z)\rangle (dz)^{w_i} \to \phi^{(w_i)}(z)\ . \tag{5.2}$$

In particular, the instanton contribution to Nekrasov's partition function should equal the conformal block of the W-algebra. With the help of (5.2) we translate the conditions on the singularities of $\phi^{(w_i)}(z)$ at $z = 0, \infty$ into conditions on the state $|\mathcal{G}\rangle$, which we call the Gaiotto-Whittaker state, inserted at $z = 0, \infty$. This state turns out to be a certain coherent state in the Verma module of the untwisted sector of the $W(G)$-algebra when G is simply-laced, and in the $\mathbb{Z}_r$-twisted sector of the $W(\Gamma)$-algebra when G is non-simply-laced. We come back to the details in Sec. 5.2.1. The most important relation is

$$W_{1/r}^{(h^\vee)}|\mathcal{G}\rangle = \Lambda^{h^\vee}|\mathcal{G}\rangle. \tag{5.3}$$

Under a suitable identification of parameters we should then have the equality

$$Z_{\text{inst}}(\vec{a};\epsilon_{1,2}) = \langle \mathcal{G}|\mathcal{G}\rangle = \langle \underline{\mathcal{G}}|\Lambda^{2h^\vee r L_0}|\underline{\mathcal{G}}\rangle\ . \tag{5.4}$$

On the right hand side, $|\underline{\mathcal{G}}\rangle$ is the coherent state (5.3) defined by setting $\Lambda = 1$. The relation can be understood as in Fig. 5.2: the boundary condition creates the state $\langle \underline{\mathcal{G}}|$, which is then propagated by the distance $\propto \log \Lambda$, then is annihilated by $|\underline{\mathcal{G}}\rangle$. Indeed, $\log \Lambda$ is the

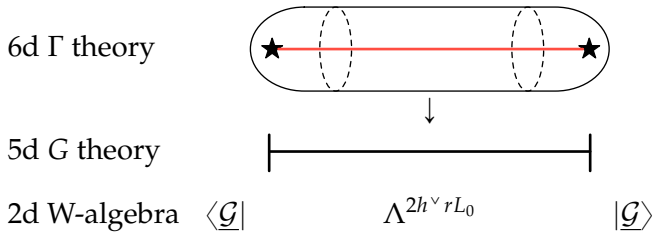


Figure 5.2: Top: the Seiberg-Witten solution of pure $\mathcal{N} = 2$ super Yang-Mills theory with gauge group G in terms of 6d $\mathcal{N} = (2,0)$ theory of type Γ on $C = \mathbb{CP}^1$ with the $\mathbb{Z}_r$ twist line from $z = 0$ to $z = \infty$. Middle: the S^1 reduction to the 5d maximally supersymmetric Yang-Mills theory with gauge group G on a segment, with a suitable half-BPS boundary condition on both ends. Bottom: In the 2d description, the coherent state $\langle \mathcal{G}|$ is produced by the BPS boundary condition. It is then propagated along the horizontal direction and annihilated by $|\mathcal{G}\rangle$.

UV coupling, and is proportional to the length of the fifth direction.

This relation was first considered for SU(2) in [120], and its calculation was later streamlined in [121]. The check for SU(3) was performed in [122] using the known explicit commutation relation of the W_3-algebra.

In this section we will check (5.4) at the one-instanton level uniformly for all G. We will use the free-field realization of the W-algebra, without explicitly writing down the complicated commutation relation of the modes of its modes. Before proceeding, it is to be noted that relation (5.4) when we have a full surface operator has already been rigorously proved for all groups to all order in [66, 60].

5.2.1 Identification of the coherent state

Under the correspondence of Nekrasov's partition functions and conformal blocks of W-algebras, the key relation is that the vev of the W-currents should become the fields $\phi^{(w_i)}(z)$ in the limit $\epsilon_{1,2} \ll a$:

$$\lim_{\epsilon_{1,2}\to 0} \langle W^{(w_i)}(z)\rangle (dz)^{w_i} \to \phi^{(w_i)}(z). \tag{5.5}$$

The fields $\phi^{(w_i)}(z)$ have two singularities $z = 0, \infty$, which means that there is a state $\langle \mathcal{G}|$ at $z = \infty$ and a state $|\mathcal{G}\rangle$ at $z = 0$. The behavior of $\phi^{(w_i)}(z)$ at $z = \infty$ has the same form as the behavior at $z = 0$ by the map $w = 1/z$. Therefore the state $\langle \mathcal{G}|$ is a conjugate of the state $|\mathcal{G}\rangle$.

When G is simply-laced, the conditions (2.25) and (2.26) imply that $|\mathcal{G}\rangle$ is in the Verma

module of $W(G)$-algebra generated from the highest weight state $|\vec{w}\rangle$ with

$$W_0^{(w_i)}|\vec{w}\rangle = w^{(w_i)}|\vec{w}\rangle \tag{5.6}$$

where the eigenvalues $w^{(w_i)}$ should equal the vev $u^{(w_i)}$ up to some quantization error involving $\epsilon_{1,2}$. The condition (2.26) then tells us that

$$W_\ell^{(w_i)}|\mathcal{G}\rangle = 0 \quad \text{for } \ell > 0 \text{ unless} \quad W_1^{(h^\vee)}|\mathcal{G}\rangle = \Lambda^{h^\vee}|\mathcal{G}\rangle. \tag{5.7}$$

When G is non-simply-laced, the conditions (2.27) through (2.29) imply that $|\mathcal{G}\rangle$ is in the Verma module generated by the $\mathbb{Z}_r$-twisted vacuum $|\vec{w}\rangle$ of $W(\Gamma)$-algebra determined by

$$W_0^{(\hat{w}_i)}|\vec{w}\rangle = w^{(\hat{w}_i)}|\vec{w}\rangle \tag{5.8}$$

where the eigenvalues $w^{(\hat{w}_i)}$ should equal the vev $u^{(\hat{w}_i)}$ up to the quantization error involving $\epsilon_{1,2}$. The condition (2.29) then says

$$W_\ell^{(w_i)}|\mathcal{G}\rangle = 0 \quad \text{for } \ell > 0 \text{ unless} \quad W_{1/r}^{(h^\vee)}|\mathcal{G}\rangle = \Lambda^{h^\vee}|\mathcal{G}\rangle. \tag{5.9}$$

Then we should have the relation

$$Z(\vec{a}, \epsilon_{1,2}) = \langle \mathcal{G}|\mathcal{G}\rangle. \tag{5.10}$$

Both sides are power series; the k-instanton contribution on the left hand side corresponds to the level-k/r contribution on the right hand side.

Note that the norm of $|\mathcal{G}\rangle$ does not change if we change the definition of $W^{(h^\vee)}$ by adding products and derivatives of lower degree generators, since $W_{-\ell}^{(h^\vee)}$ only changes by negative modes of lower generators, which annihilate $|\mathcal{G}\rangle$ anyway.

5.2.2 Coherent state at the lowest level

Here, we discuss the procedure to compute the norm of the Gaiotto-Whittaker vector $\langle \mathcal{G}|\mathcal{G}\rangle$ using the free-bosons representation of the W-algebras. The method is the same as in [121, 122]. We stick to the untwisted representations of the W-algebras for the moment.

Let us expand $|\mathcal{G}\rangle$ in terms of levels of descendants

$$|\mathcal{G}\rangle = |\vec{w}\rangle + \Lambda^{h^\vee}|\mathcal{G}_1\rangle + (\Lambda^{h^\vee})^2|\mathcal{G}_2\rangle + \cdots \tag{5.11}$$

so that $|\mathcal{G}_\ell\rangle$ has conformal weight ℓ. The condition (5.9) is now

$$W_\ell^{(w_i)}|\mathcal{G}_\ell\rangle = 0 \quad \text{for } n > 0 \text{ except for} \qquad W_1^{(h^\vee)}|\mathcal{G}_\ell\rangle = |\mathcal{G}_{\ell-1}\rangle \tag{5.12}$$

with $|\mathcal{G}_0\rangle = |\vec{w}\rangle$. In what follows, we will compute $\langle\mathcal{G}_1|\mathcal{G}_1\rangle$ and compare against the 1-instanton computations. Expressed in terms of descendants at level 1 $|\mathcal{G}_1\rangle$ is

$$|\mathcal{G}_1\rangle = \sum_i A_i W_{-1}^{(w_i)}|\vec{w}\rangle. \tag{5.13}$$

Now, use (5.12) to get

$$0 = \langle\vec{w}|W_1^{(w_i)}|\mathcal{G}_1\rangle = \sum_j A_j\langle\vec{w}|W_1^{(w_i)}W_{-1}^{(w_j)}|\vec{w}\rangle = \sum_j K_{ij}^{(1)}A_j \text{ for } w_i \neq h^\vee, \tag{5.14}$$

$$1 = \langle\vec{w}|\vec{w}\rangle = \langle\vec{w}|W_1^{(h^\vee)}|\mathcal{G}_1\rangle = \sum_j A_j\langle\vec{w}|W_1^{(h^\vee)}W_{-1}^{(w_j)}|\vec{w}\rangle = \sum_j K_{nj}^{(1)}A_j \tag{5.15}$$

where $K^{(\ell)}$ is the Kac-Shapovalov matrix at level ℓ. We can solve for A so that $A_i = (K^{(1)})_{in}^{-1}$, and the norm is given by

$$\langle\mathcal{G}_1|\mathcal{G}_1\rangle = \sum_{i,j} A_i K_{ij}^{(1)} A_j = (K^{(1)})_{nn}^{-1}. \tag{5.16}$$

It can be easily generalized to arbitrary level to get $\langle\mathcal{G}_\ell|\mathcal{G}_\ell\rangle = (K^{(n)})_{n^\ell,n^\ell}^{-1}$ where the index n^ℓ means we pick the element corresponds to $(W_{-1}^{(w_n)})^\ell|\vec{w}\rangle$. In order to get the norm of the Gaiotto states for the $\mathbb{Z}_r$ twisted sector, we simply look for the descendants of level-$1/r$, and take the corresponding element of the Kac-Shapovalov matrix.

To evaluate the Kac-Shapovalov matrix write the W generators in terms of free bosons and expand as

$$J^k(z) = i\partial\varphi^k(z) = \sum_{m\in\mathbb{Z}} J_m^k z^{-m-1} \tag{5.17}$$

with the usual commutation relation

$$[J_m^k, J_n^l] = m\delta^{k,l}\delta_{m+n,0}. \tag{5.18}$$

Then the W-algebra vacuum $|\vec{w}\rangle$ is represented by the free-boson vacuum $|\vec{a}\rangle$ where $\vec{a} = \vec{J}_0 - Q\vec{\rho}$ is the shifted zero mode of the free bosons. Note that the bra $\langle\vec{w}|$ then corresponds to $\langle -\vec{a}|$ due to the background charge $Q\vec{\rho}$ and the shift.

Let us first discuss the simply-laced case. As we are dealing with normal ordered products, the descendant states at level one can be expressed as

$$W_{-1}^{(w_i)}|\vec{a}\rangle = \sum_j M_{ij}(\vec{a})J_{-1}^j|\vec{a}\rangle, \qquad \langle -\vec{a}|W_1^{(w_i)} = \sum_j \langle -\vec{a}|J_1^i M_{ij}(-\vec{a}). \tag{5.19}$$

The coefficient $M_{ij}(\vec{a})$ is a polynomial in $\vec{a}$ and Q. The Kac-Shapovalov matrix is given by

$$K_{ij}(\vec{a}) = \sum_k M_{ik}(-\vec{a})M_{kj}(\vec{a}). \tag{5.20}$$

The twisted case is slightly more involved. Again we know that the lowest descendant states can be written as

$$W_{-1/r}^{(\tilde{w}_i)}|\vec{a}\rangle = \sum_j \tilde{M}_{ij}(\vec{a})J_{-1/r}^j|\vec{a}\rangle, \tag{5.21}$$

where $\tilde{M}_{ij}(\vec{a})$ is again a polynomial in the zero modes of the untwisted bosons. To compute it, we need to find modes of the form $:(J^j)^m:_{-1/r}$, which can be found from the original prescription of OPE-normal ordering, that is by subtracting the singular part of the correlator of the m bosons. For instance, to obtain the constant C_{2m+1} of the state $:(J^j)^{2m+1}:_{-1/2}|\vec{a}\rangle = C_{2m+1}J_{-1/2}|\vec{a}\rangle$ we extract the regular part of the correlator

$$\lim_{z_i\to z_1}\langle -\vec{a}|J_{1/2}^j J^j(z_{2m+1})J^j(z_{2m})\cdots J^j(z_1)|\vec{a}\rangle\Big|_{reg} = \frac{C_{2m+1}}{2z_1^{2m+1/2}}\,, \tag{5.22}$$

and similarly for the zero modes of even powers of J.

5.2.3 Comparison

After all these preparations, now we can compare the norm of the Gaiotto-Whittaker vector and the one-instanton partition function. Obviously, the norm of the Gaiotto-Whittaker vector can have poles only at the zero of the Kac determinant, (4.64), i.e. when $\vec{\beta}\cdot\vec{a}_{\text{boson}} + Q = 0$ for a long root β. We also have the formula of the one-instanton expression in the gauge theory side, (3.68), which has apparent poles when $\vec{\beta}\cdot\vec{a}_{\text{gauge}} = \epsilon_1+\epsilon_2$ for a long root β, or when $\vec{\gamma}\cdot\vec{a}_{\text{gauge}} = 0$ for an arbitrary root γ. In order for them to have any chance of agreement, we need to identify

$$\vec{a}_{\text{boson}} = \frac{\vec{a}_{\text{gauge}}}{\sqrt{\epsilon_1\epsilon_2}}, \qquad Q = \frac{\epsilon_1+\epsilon_2}{\sqrt{\epsilon_1\epsilon_2}}. \tag{5.23}$$

Using the procedure outlined above, we have checked the agreement between the norm of the coherent state and instanton partition function at 1-instanton level

- for simply-laced algebras $A_{1,2,3,4,5,6}$, D_4, and E_6,
- and for non-simply-laced algebras $B_{2,3}$, C_n, F_4 and G_2.

In general, the agreement comes with a multiplicative ambiguity due to the normalization of W-currents. It can be easily absorbed into the redefinition of the expansion parameter Λ.[1] For higher rank algebras, such as A_5, A_6, F_4 and E_6, due to the computational complexity, we checked the agreement by plugging in several set of test numbers for the zero modes and Q parameter instead of leaving it as a symbolic expression. Let us now discuss the cases A_n, D_n, B_n, C_n, G_2 and F_4 in this order.

5.2.3.1 A_n

The W-algebra calculation leads to the following explicit form of the Gaiotto-Whittaker vector at level one:

$$|\mathcal{G}_1\rangle = \sum_i v_i(\vec{a}) J_{i,-1}|\vec{a}\rangle \tag{5.24}$$

where

$$v_i(\vec{a}) = \sum_j C_{ij}(\vec{a}) w_j(\vec{a}), \quad w_i(\vec{a}) = \frac{1}{\prod_{x<i}(a_x-a_i)\prod_{i<x}(Q-a_i+a_x)} \tag{5.25}$$

[1] When the underlying gauge theory is conformal, one may encounter much more intricate map between expansion parameters as we will see in the later sections.

where

$$C_{ij}(\vec{a}) = 0, \qquad (i < j) \tag{5.26}$$

$$C_{ij}(\vec{a}) = 1, \qquad (i = j) \tag{5.27}$$

$$C_{ij}(\vec{a}) = (-1)^{i-j} Q \frac{\prod_{j<k<i}(Q - a_j + a_k)}{\prod_{j<k\leqslant i}(a_j - a_k)}. \qquad (i > j) \tag{5.28}$$

Here $a_i = \vec{e}_i \cdot \vec{a}$. We checked the validity for small n; we believe it is true in general.

The corresponding bra is given by

$$\langle \mathcal{G}_1| = \sum_i \langle -\vec{a}| J_{i,1} v_i(-\vec{a}). \tag{5.29}$$

Now, it can be checked that

$$v_i(-\vec{a}) = \sum_j \tilde{w}_j(\vec{a})(C^{-1}(\vec{a}))_{ji}, \text{ with } \tilde{w}_i(\vec{a}) = \frac{1}{\prod_{x<i}(Q - a_x + a_i)\prod_{i<x}(a_i - a_x)}. \tag{5.30}$$

Therefore,

$$\langle \mathcal{G}_1|\mathcal{G}_1\rangle = \sum_i \prod_{i\neq j} \frac{1}{(a_i - a_j)(Q - a_i + a_j)} \tag{5.31}$$

which is indeed the one-instanton contribution of Nekrasov's partition function [75, 55, 122] calculated from the geometry of the one-instanton moduli space. For example,

$$Z_{\mathrm{SU}(2),1} = -\frac{2}{4a^2 - Q^2}\,, \tag{5.32}$$

$$Z_{\mathrm{SU}(3),1} = \frac{6\left(a_1^2 + a_1 a_2 + a_2^2 - Q^2\right)}{[(2a_1 + a_2)^2 - Q^2][(a_1 + 2a_2)^2 - Q^2][(a_1 - a_2)^2 - Q^2]}\,. \tag{5.33}$$

Two comments are in order. First, note that the uniform formula (3.68) instead gives the following form

$$Z_{k=1} = \sum_{i\neq j} \frac{1}{(a_i - a_j)(Q - a_j - a_i)} \prod_{k\neq i} \frac{1}{a_i - a_k}. \tag{5.34}$$

The agreement of (5.31) and (5.34) are not easy to see, but they are equal nonetheless. Second, recall that each summand in the formula (5.31) comes from the contribution of a fixed point in the resolution of the one-instanton moduli space. Then the relation (5.25) means that the free boson basis $J^i_{-1}|\vec{a}\rangle$ is an upper-triangular redefinition of the basis formed by

the fixed points, explicitly confirming the results of Maulik and Okounkov [123].

5.2.3.2 D_n

The $W(D_n)$ contains W-generators of dimension $2, 4, \cdots, 2n-2$ and additionally of dimension n. The Nekrasov partition function for $D_n = SO(2n)$ can be obtained easily by evaluating the contour integral expression in [53]. We have,

$$Z_{SO(2n),1} = -\sum_{i=1}^{n} \left[\frac{(a_i \pm Q)(2a_i \pm Q)}{\prod_{j\neq i}(a_i^2 - a_j^2)\left((a_i \pm Q)^2 - a_j^2\right)} \right], \tag{5.35}$$

where $\pm$ means we sum over both signs and $a_i = \vec{e}_i \cdot \vec{a}$. We have checked the agreement up to D_4.

5.2.3.3 B_n

To get the coherent state for B_n, we start from $W(A_{2n-1})$. We need to evaluate the n-dimensional Kac-Shapovalov matrix, since all the odd-dimensional W-currents are twisted by $\mathbb{Z}_2$ automorphism. The $\mathbb{Z}_2$ action maps φ_i to $-\varphi_{2n+1-i}$. Then the eigenstates are

$$\tilde{\varphi}_i^+ = \varphi_i - \varphi_{2n+1-i} \text{ and } \tilde{\varphi}_i^- = \varphi_i + \varphi_{2n+1-i}\,. \tag{5.36}$$

Then, the twisted W-currents can be written as

$$\tilde{U}^{(m)}_{-1/2}|\vec{a}\rangle = \sum_i B_{m,i}(\vec{a}) J^{-,i}_{-1/2}|\vec{a}\rangle, \tag{5.37}$$

where $B_{m,i}$ is a function of zero modes of the untwisted free bosons. Now, we evaluate the Kac-Shapovalov matrix to obtain the norm of the coherent state. For example, when $n = 3$,

$$K_{B_3} = \begin{pmatrix} \langle -\vec{a}|\tilde{U}^{(3)}_{1/2}\tilde{U}^{(3)}_{-1/2}|\vec{a}\rangle & \langle -\vec{a}|\tilde{U}^{(3)}_{1/2}\tilde{U}^{(5)}_{-1/2}|\vec{a}\rangle \\ \langle -\vec{a}|\tilde{U}^{(5)}_{1/2}\tilde{U}^{(3)}_{-1/2}|\vec{a}\rangle & \langle -\vec{a}|\tilde{U}^{(5)}_{1/2}\tilde{U}^{(5)}_{-1/2}|\vec{a}\rangle \end{pmatrix} \tag{5.38}$$

and take the inverse of K_{B_3} and read off the $(2,2)$ component of it.

On the instanton side, we have,

$$Z_{SO(2n+1),1} = \sum_{i=1}^{n} \left[\frac{(2a_i \pm Q)}{a_i \prod_{j\neq i}(a_i^2 - a_j^2)\left((a_i \pm Q)^2 - a_j^2\right)} \right], \tag{5.39}$$

where $\pm$ means that we sum over both signs and $a_i = \vec{\epsilon}_i \cdot \vec{a}/\sqrt{2}$. We find that they agree with the norm of the corresponding coherent states up to numeric constants.

5.2.3.4 C_n

The Gaiotto-Whittaker vector corresponds to C_n can be obtained from the $W(D_{n+1})$-algebra. We can just follow the same procedure as B_n case, but in this case we can do much easily. There is only one $W(D_n)$-generator that is not invariant under $\mathbb{Z}_2$ which is $V^{(n)}$. In terms of free bosons, it only shifts the sign of one of the bosons. Therefore, the Kac-Shapovalov matrix is just a number, we can simply use (4.64). We get the norm of the Gaiotto-Whittaker vector to be

$$\langle \mathcal{G}_1 | \mathcal{G}_1 \rangle \propto \frac{1}{\prod_{i=1}^{n}(Q^2 - 4a_i^2)}. \tag{5.40}$$

It is known that the moduli space (neglecting the center of mass contribution) of one $\mathrm{Sp}(n)$ instanton is $\mathbb{C}^{2n}/\mathbb{Z}_2$, and the corresponding Hilbert series is given by (3.32) of [83], whose $\beta \to 0$ limit can be easily taken. Or, equivalently, note that $\mathrm{U}(1)^2 \times \mathrm{Sp}(n)$ acts on $\mathbb{C}^{2n}$ with the eigenvalues $(\epsilon_1 + \epsilon_2)/2 \pm a_i$. Therefore the integral is just

$$Z_{\mathrm{Sp}(n),1} = \frac{1}{2} \frac{1}{\prod_{i=1}^{n}(Q^2/4 - a_i^2)} \tag{5.41}$$

where $a_i = \vec{e}_i \cdot \vec{a}$ and the factor $1/2$ comes from the orbifolding. This expression can, of course, also be obtained from Formula (3.78). We see that they agree completely up to a multiplicative constant.

5.2.3.5 G_2

There is only one W-generator that has a $-1/3$ mode: $W^{(4,\frac{2}{3})} = W^{(4)} + i\tilde{W}^{(4)}$. Therefore, the Kac-Shapovalov matrix is one-dimensional, which is given by

$$\langle \vec{a}|W^{(4,\frac{2}{3})}_{1/3}W^{(4,\frac{2}{3})}_{-1/3}|\vec{a}\rangle \propto \frac{1}{4}\left(Q^2-6a_2^2\right)\left[4Q^4-12Q^2\left(3a_1^2+a_2^2\right)+9\left(3a_1^2-a_2^2\right)^2\right]. \tag{5.42}$$

We can use the formula (4.64) for this example as well. The norm of Gaiotto-Whittaker state is given as the inverse of the the above expression.

The Hilbert series was given in (5.48) and (5.49) of [83]. Taking the limit $\beta \to 0$, we obtain

$$Z_{G_2,1} = \frac{72}{\left(Q^2-6a_2^2\right)\left[4Q^4-12Q^2\left(3a_1^2+a_2^2\right)+9\left(3a_1^2-a_2^2\right)^2\right]} \tag{5.43}$$

in our variables.[2] Here, $a_i = \vec{\epsilon}_i \cdot \vec{a}$. This expression can also be obtained from Formula (3.78). See the section A.2 for G_2 for explicit expression for the roots and their basis. We see that the Hilbert series result completely agrees with the norm of Gaiotto-Whittaker state up to a multiplicative constant.

5.2.3.6 F_4

The Kac-Shapovalov matrix we need to compute is

$$K_{F_4} = \begin{pmatrix} \langle -\vec{a}|W^{(5)}_{1/2}W^{(5)}_{-1/2}|\vec{a}\rangle & \langle -\vec{a}|W^{(5)}_{1/2}W^{(9)}_{-1/2}|\vec{a}\rangle \\ \langle -\vec{a}|W^{(9)}_{1/2}W^{(5)}_{-1/2}|\vec{a}\rangle & \langle -\vec{a}|W^{(9)}_{1/2}W^{(9)}_{-1/2}|\vec{a}\rangle \end{pmatrix}. \tag{5.44}$$

Honest computation of this matrix is too time consuming for a desktop computer of 2011, due to the complication in the evaluation of the normal ordering of twisted bosons in the expressions involving $W^{(9)}$ as in (5.22). Thankfully, the Kac determinant is known in closed form in (4.64). Therefore we find

$$\langle \mathcal{G}_1|\mathcal{G}_1\rangle = \langle -\vec{a}|W^{(5)}_{1/2}W^{(5)}_{-1/2}|\vec{a}\rangle/(\text{Kac determinant}), \tag{5.45}$$

which is fairly straightforward to compute. It was checked that it agrees with the instanton

[2] Here $\vec{a}_{\text{ours}} = \sqrt{3}\vec{a}_{\text{theirs}}$

expression (3.68). The explicit results for E_6 and F_4 are too lengthy to put here.

5.3 $SO - Sp$ quiver

5.3.1 The $SO(4)$ and $Sp(1)$ AGT correspondence

Remember that the $\mathcal{N} = 2$ geometry is characterized by a ramified Hitchin system on the UV-curve. For conformal $SO(2N)$ and $Sp(N-1)$ gauge theories the Hitchin system is described in terms of the differentials ϕ_{2k} (for $k = 1, \ldots, N-1$) and $\tilde{\phi}_N$ that can be constructed out of the D_N-invariants $\mathrm{Tr}\left(\Phi^{2k}\right)$ and $\mathrm{Pfaff}(\Phi)$, respectively. In the six-dimensional $(2,0)$ theory these differentials appear as a set of chiral operators whose conformal weights are equal to the exponents of the Lie algebra. When we reduce the six-dimensional theory over a four-manifold we expect these operators to turn into the Casimir operators $\mathcal{W}^{(2k)}$ and $\tilde{\mathcal{W}}^{(N)}$ of the $\mathcal{W}(D_N)$-algebra. The $\mathbb{Z}_2$-automorphism of the D_N-algebra translates to an additional $\mathbb{Z}_2$-symmetry on the level of the CFT, and we thus expect a relation to a twisted $\mathcal{W}(D_N)$-algebra. In other words, we expect that the Lie algebra underlying the Hitchin system is precisely reflected in the Casimir operators of the corresponding $\mathcal{W}$-algebra on the UV-curve.

Let us now put all the pieces together and formulate the AGT correspondence for $SO(4)$ and $Sp(1)$. Since the definition of the UV-curve for both theories involves the $SO(4)$-invariants ϕ_2 and $\tilde{\phi}_2$ of degree two, we expect this CFT to have an underlying $\mathcal{W}(D_2)$-algebra. We will denote this algebra by $\mathcal{W}(2,2)$, as it contains two Casimir operators of weight two. In fact those operators correspond to two copies T^A and T^B of the Virasoro algebra.

Similar to the correspondence between $U(2)$ instanton partition function and Virasoro conformal blocks, the new correspondence is between $SO(4)/Sp(1)$ instanton partition functions and twisted $\mathcal{W}(2,2)$-algebra blocks. The configuration of the block is given in the following way: At the full punctures of the $SO(4)/Sp(1)$ UV-curve insert untwisted vertex operators, whose weights correspond to the masses of the $Sp(1)$ fundamental hypers. At the half punctures, insert twisted vertex operators. Whenever a half-puncture lifts to a regular point on the cover, we should insert the vacuum of the twist sector σ, which we will describe later on. For any half-puncture that lifts to a puncture on the cover we may insert a general twisted field, whose single weight corresponds to the mass of the $SO(4)$

N=2 gauge theory	CFT
$SO(4)/Sp(1)$ quiver	$SO(4)/Sp(1)$ UV-curve
$Sp(1)$ fund. hyper (μ_1, μ_2)	untwisted $\mathcal{W}(2,2)$ representation (h_{μ_1}, h_{μ_2})
$SO(4)$ fund. hyper μ	twisted $\mathcal{W}(2,2)$ representation h_μ
$Sp(1) - SO(4)$ bifund. hyper	twist vacuum σ
$Sp(1)$ Coulomb par. a	weight of twisted int. channel h_a
$SO(4)$ Coulomb pars. (a_1, a_2)	weights of untwisted int. channel (h_{a_1}, h_{a_2})

Table 5.1: The AGT correspondence for $SO(4)/Sp(1)$.

fundamental hyper.

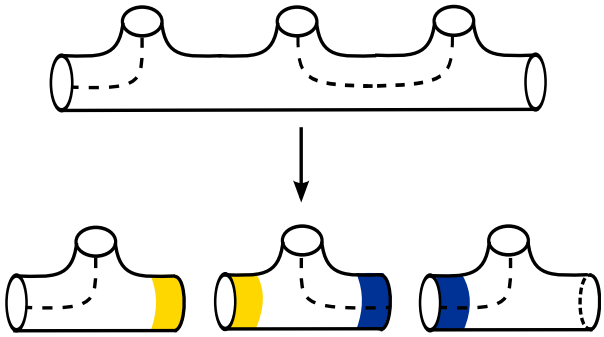


Figure 5.3: Twisted $\mathcal{W}(2,2)$-algebra blocks can be computed by decomposing the Riemann surface into pairs of pants. Internal tubes that have a $\mathbb{Z}_2$ twist line (blue) correspond to an $Sp(1)$ gauge group and carry twisted representations of $\mathcal{W}(2,2)$. Internal tubes without a twist line (yellow) correspond to $SO(4)$ and carry two copies of the Virasoro algebra.

When decomposing the UV-curve into pair of pants, we cut tubes with or without twist lines (see Figure 5.3). A tube with a twist line corresponds to a $Sp(1)$ gauge group, and the weight of the twisted primary in the channel corresponds to its single Coulomb branch parameter a. A tube without a twist line corresponds to a $SO(4)$ gauge group, and the two weights of the untwisted primary in the channel correspond to the two Coulomb branch parameters a_1 and a_2. All of this is summarized in table 5.1.

The detailed identification of parameters can be found in the examples we will work out. These examples will show that $Sp(1)/SO(4)$ instanton partition functions agree with the twisted $\mathcal{W}(2,2)$ up to a spurious factor that is independent of the Coulomb and mass parameters of the gauge theory. Note in particular that, unlike in the original $U(2)$ case, no $U(1)$ prefactor appears, which is exactly what one would expect.

We can also use this correspondence to explain the relation between $Sp(1)/SO(4)$ and $U(2)$ theories. More precisely, given an $Sp(1)/SO(4)$ UV-curve, we first map the corresponding chiral block to its double cover. This is in fact a well-known method to compute twisted correlators. The resulting configuration can then be mapped to a $U(2)$ configuration by a suitable conformal coordinate transformation. We will argue below that such a coordinate transformation only introduces a spurious factor. It thus follows that the partition function of the $U(2)$ configuration agrees up to a spurious factor with the partition function of the $Sp(1)/SO(4)$ configuration once expressed in terms of the same coupling constants. To put it another way, the difference between $U(2)$ and $Sp(1)/SO(4)$ partition functions is indeed only a reparametrization of the moduli space caused by choosing a different renormalization scheme.

5.3.2 Correlators for the $\mathcal{W}(2,2)$ algebra and the cover trick

As mentioned above, the algebra $\mathcal{W}(2,2)$ contains two Casimir operators of weight two. These operators can be identified with two Virasoro tensors $T^A(z)$ and $T^B(z)$. The $\mathcal{W}$-algebra thus decomposes into two copies of the Virasoro algebra. This reflects the decomposition of the Lie algebra $\mathfrak{so}(4) \cong \mathfrak{su}(2)_A \times \mathfrak{su}(2)_B$. Geometrically, the fact that we find two copies of the Virasoro algebra follows simply from the double covering that relates the $U(2)$ and the $Sp(1)/SO(4)$ UV-curves. A single copy of the Virasoro algebra associated to the cover $U(2)$ UV-curve descends to two copies of the Virasoro algebra on the base $Sp(1)/SO(4)$ UV-curve.

As illustrated on the left in Figure 5.4, this is in particular the case for an internal tubular neighborhood of the $SO(4)$ Gaiotto curve. The single copy of the Virasoro algebra on its inverse image descends to two copies of the Virasoro algebra on the tubular neighborhood itself. The $SO(4)/Sp(1)$ UV-curves additionally contain $\mathbb{Z}_2$ twist-lines. When crossing such a twist-line the two copies of the Virasoro algebra get interchanged. We thus propose an underlying twisted $\mathcal{W}(2,2)$-algebra. The actual energy-stress tensor $T^+(z) = T^A(z) + T^B(z)$ is of course invariant, whereas $T^-(z) = T^A(z) - T^B(z)$ picks up a minus sign.

To compute the corresponding chiral blocks, we decompose the Gaiotto curves into pair of pants and sum over all $\mathcal{W}$-descendants of a given channel. The only difference with Virasoro correlators is that there are now two types of tubes to cut, those with $\mathbb{Z}_2$

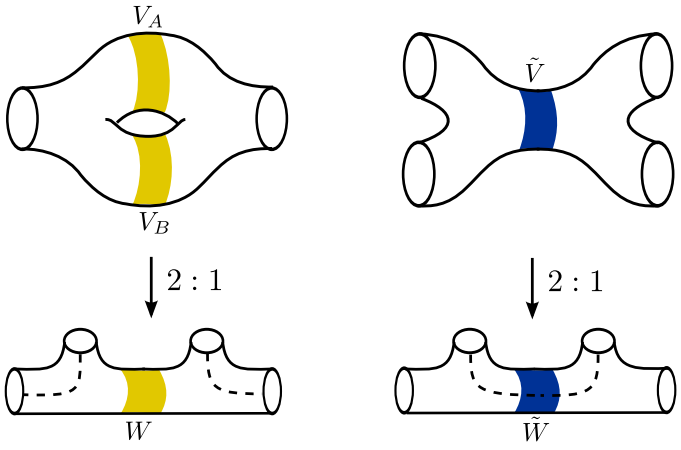


Figure 5.4: On the left (right): Illustration of the branched double covering of the $SU(2)$ UV-curve over the $SO(4)$ UV-curve ($Sp(1)$ UV-curve). The yellow (blue) tubular neighborhoods W and $\tilde{W}$ on the base curves are part of internal tubes without (with) a $\mathbb{Z}_2$ twist line. On the cover the yellow (blue) patches illustrate their respective inverse images. The $\mathcal{W}$-algebra modes associated to both base tubes lift to a single copy of the Virasoro algebra on their inverse images.

twist-lines and those without. This is illustrated in Figure 5.3.

When cutting open a tube without a $\mathbb{Z}_2$ twist-line, the intermediate fields are given by L^A and L^B-descendants of an untwisted representation, characterized by the conformal weights (D_A, D_B) under L_0^A and L_0^B. We can therefore associate the Hilbert space

$$\mathcal{H}_{SO(4)} = \{L^A_{-m_1} \cdots |\phi_A\rangle \otimes L^B_{-n_1} \cdots |\phi_B\rangle \ : m_i \in \mathbb{N}, n_i \in \mathbb{N}\}, \tag{5.46}$$

where $|\phi_{A/B}\rangle$ has weight $D_{A/B}$, to a tube without a $\mathbb{Z}_2$ twist-line. On the other hand, if we cut a tube with a $\mathbb{Z}_2$ twist-line, the intermediate fields are in a twisted representation of the $\mathcal{W}$-algebra. It is then most convenient to describe them in terms of descendants of L^+ and L^-,

$$\mathcal{H}_{\widetilde{Sp}(1)} = \{L^+_{-m_1} \cdots L^-_{-r_1} \cdots |\phi_C\rangle \ : m_i \in \mathbb{N}, r_i \in \frac{1}{2} + \mathbb{N}\} \, . \tag{5.47}$$

Note that since L^- has no zero mode, the representation ϕ_C is characterized by just a single weight.

To actually compute three point functions with twist fields, we can use the well-known cover trick, which is nicely explained in, e.g., [124, 125] : We find a function that maps

the punctured Riemann surface with branch cuts to a cover surface which does not have any branch cuts. Since the theory is conformal, we know how the correlation functions transform under this map. On the cover we can then evaluate a correlation function with no twisted fields and no branch cuts in the usual way. In order for this to work, we need to find a cover map that has branch points where the twist fields are inserted. The precise map from the base to the cover thus depends on the positions of the branch cuts. On the cover there is then only a single copy of the Virasoro algebra.

To illustrate all of this, let us take the following simple model as a map from the cover to the base:

$$\tilde{z} \mapsto z = \tilde{z}^2 \,. \tag{5.48}$$

This particular map has branch cuts at 0 and ∞, and is thus suitable to deal with correlation functions that have twist fields at those two points. We can relate the stress-energy tensor on the cover $T(\tilde{z})$ to the two copies on the base in the following manner. The stress-energy tensor on the cover transforms to

$$T(\tilde{z}) = \left(\frac{d\tilde{z}}{dz}\right)^{-2} \left[T(z) - \frac{c}{12}\{\tilde{z}; z\}\right] \,, \tag{5.49}$$

where the Schwarzian derivative given by

$$\{\tilde{z}; z\} = \frac{\tilde{z}'''}{\tilde{z}} - \frac{3}{2}\left(\frac{\tilde{z}''}{\tilde{z}'}\right)^2 \tag{5.50}$$

appears because T is not a primary field. Around the branch point 0 on the base we can then define two $\mathcal{W}$-fields T^+ and T^- by picking out the even and odd modes of T,

$$L_n^+ = 2\oint \frac{dz}{z^{n-1}} T(z) = \frac{1}{2}L_{2n} + \frac{3c}{48}\delta_{n,0} \qquad (\text{for } n \in \mathbb{Z}), \tag{5.51}$$

$$L_r^- = 2\oint \frac{dz}{z^{r-1}} T(z) = \frac{1}{2}L_{2r} \qquad (\text{for } r \in \frac{1}{2} + \mathbb{Z}). \tag{5.52}$$

The L^+ then form a Virasoro algebra with central charge $2c$, and T^- is a primary field of weight 2. As discussed above, the twisted field ϕ at the point 0 is a twisted representation of T^- and T^+ which has only one weight, namely the eigenvalue of L_0^+. This means that on the cover point there sits a field ϕ which is an untwisted representation of the Virasoro

algebra of the corresponding weight, and its L^+ and L^- descendants are given by even and odd L descendants.

There is one special twist field σ which has the property that its lift to the cover gives the vacuum. It has the lowest possible conformal weight for a twist field and serves in some sense as the vacuum of this particular twist sector.

Around any other puncture on the base that is not a branch point, we simply obtain two independent copies L^A and L^B of the Virasoro algebra, coming from the two pre-images of the punctures on the cover. As long as we stay away from branch points, the Virasoro tensor $T(\tilde{z})$ on the cover is given by $T^A(\tilde{z})$ on the first and by $T^B(\tilde{z})$ on the second sheet of the cover. Since T^A and T^B commute on the base, a field $\phi^{A,B}$ on the base factorizes into representations of T^A and T^B, $\phi^{A,B} = \phi^A \otimes \phi^B$ with conformal weights (D_A, D_B) under both copies. On the cover this leads to two untwisted fields ϕ^A and ϕ^B sitting at the two images of the cover map, both of which are again untwisted representations of the Virasoro algebra.

Let us now to turn to some more technical points. The map from the base to the cover in general introduces corrections to the three point functions. In particular since one or more of those fields are descendants, they will exhibit more complicated transformation properties than we are used to from primary fields. Let us therefore briefly discuss how conformal blocks behave under coordinate transformations.

When dealing with descendants fields, it will be useful to use the notation $\phi(z) = V(\phi, z)$, which we shorten to $V_i(z)$ if $\phi_i(z)$ is a primary field. The transformation of a general descendant field ϕ under a general coordinate transformation $z \mapsto f(z)$ is given by [126]

$$D_f V(\phi, z) D_f^{-1} = V\left(f'(z)^{L_0} \prod_{n=1}^{\infty} e^{T_n(z) L_n} \phi, f(z)\right), \tag{5.53}$$

where the operator D_f is given by

$$D_f = e^{f(0) L_{-1}} f'(0)^{L_0} \prod_{n=1}^{\infty} e^{T(0)_n L_n} . \tag{5.54}$$

Here we take all products to go from left to right. The functions $T_n(z)$ are defined recur-

sively. The first two are given by

$$T_1(z) = \frac{f''(z)}{2f'(z)}\,, \qquad T_2(z) = \frac{1}{3!}\left(\frac{f'''(z)}{f'(z)} - \frac{3}{2}\left(\frac{f''(z)}{f'(z)}\right)^2\right)\,. \tag{5.55}$$

First note that if ϕ is a primary field, (5.117) reduces to the standard expression $\phi \mapsto (f'(z))^h\phi$. For general descendants however, the result will be a linear combination of correlators of lower descendant fields. Also note that $T_2(z)$ is in fact a multiple of the Schwarzian derivative. It is actually true that all higher $T_n(z)$ are sums of products of derivatives of the Schwarzian derivative. Since the Schwarzian derivative of a Möbius transformation vanishes, those transformations lead to much simpler expressions.

In some cases however we can avoid having to transform descendant fields. Assume that we want to compute the chiral block of a configuration for which we know the base to cover map f. When we go to the cover, we can use the fact that D_f is a function of the Virasoro modes L_n only. This means that it does not mix different representations, so that, more formally,

$$D_f^{-1} P_{\mathcal{H}_\phi} D_f = P_{\mathcal{H}_\phi}\,. \tag{5.56}$$

From this it follows that conformal block has the same transformation properties as the underlying correlation function, as can be seen, e.g., in the simplest case

$$\begin{aligned}&\langle V_1(z_1)V_2(z_2)P_{\mathcal{H}_\phi}V_3(z_3)V_4(z_4)\rangle \\ &\quad = \prod_{i=1}^{4}(f(z_i))^{h_i}\langle V_1(f(z_1))V_2(f(z_2))P_{\mathcal{H}_\phi}V_3(f(z_3))V_4(f(z_4))\rangle\,.\end{aligned} \tag{5.57}$$

Note that what we have said here is strictly speaking true only for global coordinate transformations f, i.e., for Möbius transformations

$$z \mapsto \gamma(z) = \frac{az+b}{cz+d}\,, \qquad a,b,c,d \in \mathbb{C}\,. \tag{5.58}$$

Other transformations, in particular also cover maps, must be treated with more caution, as they can introduce new singularities. On a technical level this means that at some points f is no longer locally invertible and D_f does no longer annihilate the vacuum.

From these remarks it follows that conformal blocks exhibit the same behavior as correlation functions under coordinate transformations. This does not mean, however, that their behavior under channel crossing is the same. In particular, the full partition function must be crossing symmetric, whereas individual conformal blocks will transform into each others in a very complicated manner. More precisely, if we expand the analytic continuation of the full partition function around 0 or ∞, then the resulting power series has essentially the same form as the original expansion. This is simply a consequence of covariance under Möbius transformations and the fact that we can change the order of operators in the correlation function, as they are mutually local. In contrast, even though the conformal block still transforms nicely under coordinate changes, the projector in it is not local, so that we cannot change the order of the fields at will, which means that expansions around different points will look different.

Coming back to the computation of the conformal block, if we do not know the full cover map, then we need to decompose the conformal block into three point functions with twist fields. We then evaluate these three point functions by mapping them to their appropriate covers. Note that in that case the cover maps are different for the individual three point functions, and no longer defined for the entire configuration. This means that the above arguments no longer apply, and that we must take into account the transformation properties of the descendant fields.

Let us make one more remark concerning prefactors in the AGT correspondence. From (5.57) we see that any coordinate transformation on the UV-curve leads to a product of prefactors of the form $(f')^h$. From the way h is related to the gauge theory masses, it follows that this factor does not depend on the Coulomb branch parameters, and that it only contributes to $\mathcal{F}_0$ and $\mathcal{F}_1$. Nevertheless the structure of the exponent of the $U(1)$ prefactor found in [26] is different, so that it cannot be transformed away in this way, which is in line with what was expected on physical grounds.

5.3.3 Examples

We proceed to verify the correspondence in detail in a few examples, the $Sp(1)$ gauge theory coupled to four hypermultiplets and the $SO(4)$ gauge theory coupled to two hypermultiplets.

5.3.3.1 $Sp(1)$ versus $U(2)$ correlators

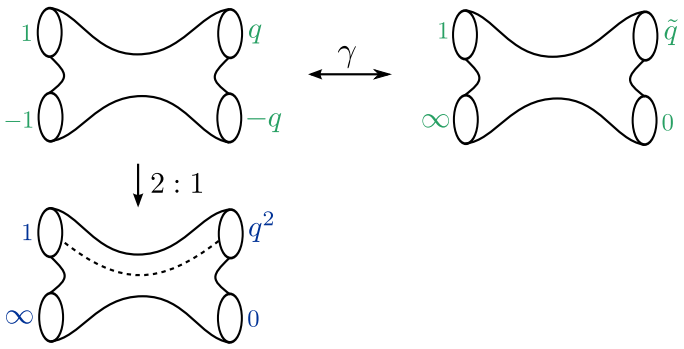


Figure 5.5: On the left, the UV-curve of the $Sp(1)$ gauge theory coupled to 4 hypers and its double cover. The Möbius transformation γ relates the double cover to the $SU(2)$ UV-curve.

Recall that the UV-curve for the $Sp(1)$ gauge theory coupled to four massive hypers is given by a four-punctured sphere, as illustrated on the bottom left of Figure 5.5. The two half-punctures at 1 and q^2 are connected by a branch cut. As we have found in section 3.5, the cross-ratio q^2 of the four punctures can be expressed in terms of the $Sp(1)$ instanton coupling $q_{Sp(1)}$ as

$$q^2 = \left(\frac{q_{Sp(1)}}{4}\right)^2. \tag{5.59}$$

The chiral block we need to evaluate is obtained by cutting the tube with the twist line, so that

$$\mathcal{F}_{Sp(1)}(q) = \langle V_1^{A,B}(\infty)\sigma(1)P_{\mathcal{H}_\phi}\sigma(q^2)V_2^{A,B}(0)\rangle \tag{5.60}$$

where the vertex operators $V_{1,2}^{A,B}$ factorize into representations of T^A and T^B of weight (h_1, h_2) and (h_3, h_4),

$$V_1^{A,B} = V_1^A V_1^B$$
$$V_2^{A,B} = V_2^A V_2^B.$$

Here we identified the half-integral mode $L^-_{-\frac{1}{2}}$ of the twisted $\mathcal{W}(2,2)$ algebra with the one-instanton modulus $q_{Sp(1)}$ of the $Sp(1)$ theory.

To evaluate the correlator with twist fields, we want to go to the double cover. The base has half-punctures at 1 and q^2, so that we map it to the double cover by

$$z \mapsto \tilde{z} = \pm\sqrt{\frac{z-q^2}{z-1}} \,. \tag{5.61}$$

This maps has indeed branch points at 1 and q^2, and it maps the fields at 0 and ∞ to $\pm q$ and ± 1. The block (5.60) on the base thus becomes the block on the cover (up to some constant prefactor)

$$\mathcal{F}_{Sp(1)}(q) = \left(1-q^2\right)^{\sum_i h_i} \langle V_1^B(-1) V_1^A(1) P_{\mathcal{H}_\phi} V_2^A(q) V_2^B(q) \rangle_C \,. \tag{5.62}$$

Note that since we know the full base-cover map, we were able to make use of (5.57) without worrying about descendant fields. To evaluate (5.62), we write it as a sum over three point functions in the usual manner. Let us therefore define three point coefficients on the cover by

$$\tilde{\mathbf{C}}^{h_1,h_2;h_3}_{I_1,I_2,I_3} = \langle V^1_{I_1}(1) V^2_{I_2}(-1) V^3_{I_3}(0) \rangle$$

where I_i gives the Virasoro descendants acting on V^i, which we will usually denote by Young diagrams. The conformal block can then be evaluated as

$$\begin{aligned}\mathcal{F}_{Sp(1)}(q) = \left(1-q^2\right)^{\sum_i h_i} \sum_{I_a, J_a} \tilde{\mathbf{C}}^{h_1,h_2;h_a}_{\bullet,\bullet,I_a} \tilde{\mathbf{C}}^{h_3,h_4;h_a}_{\bullet,\bullet,J_a} (\langle V^a_{I_a} | V^a_{J_a} \rangle)^{-1} q^{h_a+|I_a|} \\ = q^{h_a} \left(1 + \frac{2(h_1-h_2)(h_3-h_4)}{h_a} q + \dots\right) \,.\end{aligned} \tag{5.63}$$

Using the identification of parameters

$$\begin{aligned} h_i &= \frac{1}{\epsilon_1\epsilon_2}\left(\frac{Q^2}{4} - m_i^2\right), \\ h_a &= \frac{1}{\epsilon_1\epsilon_2}\left(\frac{Q^2}{4} - a^2\right), \end{aligned}$$

where $Q = \epsilon_1 + \epsilon_2$ and the momenta m_i are related to the mass parameters μ_i as

$$\begin{aligned} m_1 &= \tfrac{\mu_1+\mu_2}{2} & m_3 &= \tfrac{\mu_3+\mu_4}{2} \\ m_2 &= \tfrac{\mu_1-\mu_2}{2} & m_4 &= \tfrac{\mu_3-\mu_4}{2}, \end{aligned}$$

we find that equation (5.62) is indeed equal to the $Sp(1)$ instanton partition function up to a spurious factor independent of a and μ_i,[3]

$$Z^{Sp(1)}(q_{Sp(1)}) = \left(1 - \left(\frac{q_{Sp(1)}}{4}\right)^2\right)^{-\frac{1}{16}(c+1)} \mathcal{F}_{Sp(1)}(q_{Sp(1)})\,. \tag{5.64}$$

Note in particular that this spurious factor is independent of the masses of the hypermultiplets, in contrast to the spurious factor in the AGT correspondence for unitary gauge groups. This is indeed as expected, as the latter should come from the decoupled $U(1)$ in the $U(2)$, whereas there is no such $U(1)$ in the $Sp(1)$ setup. This fact will also be important for extending the $Sp(1)$-correspondence to linear $Sp(1) - SO(4)$ quivers.

Relation to the $U(2)$ correlator

We already knew that the full $Sp(1)$ and $U(2)$ Nekrasov partition function are related by the change of parameters (5.59). To understand this better from the conformal field theory perspective, let us study the relation between the $Sp(1)$ conformal block (5.60) and the $U(2)$ conformal block.

We have used the fact that the conformal block on the base (5.60) can be related to the block on the cover (5.62). The block on the cover is obviously very closely related to the original $U(2)$ configuration depicted on the right of Figure 3.8. We can map one to the other using the Möbius transformation γ given by equation (3.105). This is of course only possible provided that we make their cross ratios agree by identifying

$$q_{U(2)} = \frac{q_{Sp(1)}}{\left(1 + \frac{q_{Sp(1)}}{4}\right)^2}\,,$$

which is exactly the relation found on the gauge theory side.

From equation (5.57) we also know that this transformation only introduces an overall

[3] We checked this result up to order 6 in the instanton parameter.

prefactor which does not depend on the weight of the intermediate channel, so that it is an a-independent prefactor. If we were interested in the relation between instanton partition functions without the perturbative part, we would need to divide both (5.62) and the $U(2)$ block by $q^{h_a-h_1-h_4}$. The a-dependent part of the ratio of the two is then

$$\left(\frac{q_{U(2)}}{q_{Sp(1)}}\right)^{h_a}.$$

On the gauge theory side, this factor originates from the difference in the instanton part of F_0. By construction, $F_0^{\text{inst}} - \tilde{F}_0^{\text{inst}} = -a^2(\log q_{U(2)} - \log q_{Sp(1)})$, which agrees with the above factor (for $Q = 0$).

5.3.3.2 $SO(4)$ versus $U(2) \times U(2)$ correlators

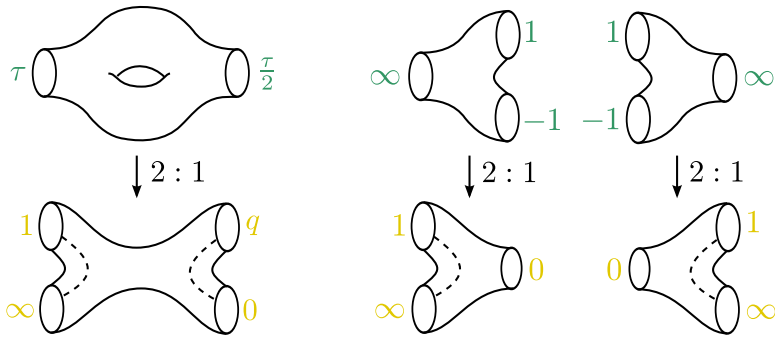


Figure 5.6: The UV-curve for the $SO(4)$ coupled to two hypermultiplets and its double cover. The left picture illustrates the global mapping between the $SO(4)$ UV-curve and its double cover, whereas the right picture illustrates the local mappings that we use to compute the twisted $\mathcal{W}(2,2)$ conformal block on the $SO(4)$ UV-curve.

Let us now turn to the $SO(4)$ case. The UV-curve for the $SO(4)$ theory with two hypers in the fundamental is given in the lower left of Figure 5.6. The chiral block we need to evaluate is

$$\mathcal{F}_{SO(4)}(q_{SO(4)}) = \langle \hat{V}_1(\infty)\sigma(1)P_{\mathcal{H}_\phi}\sigma(q_{SO(4)})\hat{V}_2(0)\rangle\,. \tag{5.65}$$

Note that we identified the integral modes $L_{-1}^{A/B}$ with the one-instanton parameter $q_{SO(4)}$.

Similarly to the $Sp(1)$ example that we discussed previously, there is an elegant way of obtaining the chiral block that makes use of the fact that we know the double cover map

of the full configuration (5.65). This double covering was described in section 3.5.2.2. In particular it maps the punctures

$$(\infty, 1, q_{SO(4)}, 0) \mapsto (1/2, 0, \tau/2, (1+\tau)/2) \, .$$

The configuration on the cover is a torus with two punctures at 0 and $q_{U(2)}$. The conformal block for this configuration has been computed in [26] and agrees with the $U(2) \times U(2)$ instanton partition function. Up to spurious prefactors introduced by the mapping to the cover, (5.65) is thus given by the conformal block of the two punctured torus expressed in terms of $q_{SO(4)}$.

However, in more general examples (i.e.the ones that we encounter in section 5.4) it will be much harder to find the global mapping between the SO/Sp UV-curve and its double cover. We thus need to develop a method that doesn't require this global information, and computes the twisted $\mathcal{W}(2,2)$ block from a simple decomposition of the UV-curve into pairs of pants. Let us exemplify this for the $SO(4)$ UV-curve.

Evaluating the $SO(4)$-block (5.65) is more complicated than the $Sp(1)$-block we considered previously. Since it has two branch cuts, there is no longer a simple square root map that maps the block (5.65) to its double cover, the torus. What we will do instead is to first decompose the block into three point functions, and then map those three point functions individually to their covers, as depicted on the right side of Figure 5.6.

Let us also define twisted three point coefficients on the base as

$$\mathbf{C}^{h_A,h_B;h_1}_{I_A,I_B,\bullet} := \langle \hat{V}_1(\infty)\sigma(1)V^{A,B}_{I_A,I_B}(0)\rangle, \tag{5.66}$$

so that

$$\mathcal{F}_{SO(4)}(q_{SO(4)}) = \sum_{I_A,J_A,I_B,J_B} \mathbf{C}^{h_A,h_B;h_1}_{I_A,I_B,\bullet}\mathbf{C}^{h_A,h_B;h_2}_{J_A,J_B,\bullet}(\langle V^{2,3}_{I_A,I_B}|V^{2,3}_{J_A,J_B}\rangle)^{-1} q^{h_A+|I_A|+h_B+|I_B|}_{SO(4)} \, .$$

Note that we have used that $\sigma(1)$ is a primary field, so that we can exchange the fields at 0 and ∞ at will. Our task is now to evaluate (5.66). This we do by mapping to it the double cover. Since 1 and ∞ are branch points, we use the map

$$z \mapsto \tilde{z} = \pm(1-z)^{-1/2} \, , \tag{5.67}$$

which maps (5.66) to three point functions on the cover of the form

$$\langle V^A_{I_A}(1) V^B_{I_B}(-1) \hat{V}^1(0) \rangle \,. \tag{5.68}$$

To find the precise relation between (5.66) and (5.68) however we need to take into account the transformation properties of all the fields under the map from the transformation (5.67). This is no issue for σ and $\hat{V}^1$, since those fields are always primary fields, so that any overall prefactors will always be cancelled once we divide by the primary three point function. In what follows, we will always omit these factors. It is however an issue for V^A and V^B, since those fields are descendants. Using (5.117) we can thus express the field on the base by the field on the cover as

$$\begin{aligned} &V^{A,B}(0) = \\ &= V\left(\left(\frac{1}{2}\right)^{L_0} e^{3L_1/4} e^{L_2/16} \cdots \phi^A_{I_A}, 1\right) V\left(\left(-\frac{1}{2}\right)^{L_0} e^{3L_1/4} e^{L_2/16} \cdots \phi^B_{I_B}, -1\right) \end{aligned} \tag{5.69}$$

where we have only included terms that are relevant up to second level descendants.

Let us show how to compute the first order term of the chiral block. To fix the normalization, we use that the primary three point function transforms as

$$\mathbf{C}^{h_A,h_B;h_1}_{\bullet,\bullet,\bullet} = \left(\frac{1}{2}\right)^{h_A} \left(-\frac{1}{2}\right)^{h_B} \tilde{\mathbf{C}}^{h_A,h_B;h_1}_{\bullet,\bullet,\bullet} \,.$$

The normalized coefficients for the first level descendants can then be computed to be

$$\begin{aligned} \mathbf{C}^{h_A,h_B;h_1}_{\square,\bullet,\bullet} &= \frac{1}{2}\tilde{\mathbf{C}}^{h_A,h_B;h_1}_{\square,\bullet,\bullet} + \frac{3h_A}{2} \\ \mathbf{C}^{h_A,h_B;h_1}_{\bullet,\square,\bullet} &= -\frac{1}{2}\tilde{\mathbf{C}}^{h_A,h_B;h_1}_{\bullet,\square,\bullet} + \frac{3h_B}{2} \end{aligned}$$

Using

$$\tilde{\mathbf{C}}^{h_A,h_B;h_1}_{\square,\bullet,\bullet} = \frac{1}{2}(-h_1 - 3h_A + h_B) \,, \qquad \tilde{\mathbf{C}}^{h_A,h_B;h_1}_{\bullet,\square,\bullet} = \frac{1}{2}(h_1 - h_A + 3h_B) \,,$$

we obtain

$$\mathcal{F}_{SO(4)} = q_{SO(4)}^{h_A+h_B}\Big(1+\Big(\frac{(3h_A+h_B-h_1)(3h_A+h_B-h_2)}{2h_A}+ \frac{(3h_B+h_A-h_1)(3h_B+h_A-h_2)}{2h_B}\Big)\frac{q_{SO(4)}}{16}+\dots\Big)\,. \quad (5.70)$$

Using the identification of parameters

$$h_i = \frac{1}{\epsilon_1\epsilon_2}\left(\frac{Q^2}{4}-\mu_i^2\right),$$
$$h_{A/B} = \frac{1}{\epsilon_1\epsilon_2}\left(\frac{Q^2}{4}-\beta_{A/B}^2\right),$$

where $\beta_{A/B} = \frac{b_1 \pm b_2}{2}$, we have indeed checked up to order 2 that (5.65) agrees with the $SO(4)$ partition function up to a spurious prefactor given by

$$Z_{\text{sp}} = (1-q)^{\frac{3}{8}Q^2}\,. \quad (5.71)$$

5.4 Linear Sp/SO quivers

In this section we discuss the generalization of the correspondence for single Sp and SO gauge groups to linear quiver gauge theories involving both Sp and SO gauge groups. This process will involve new elements from the instanton counting perspective, which we introduce in this section.

The SO/Sp correspondence that we studied in the previous sections can be naturally extended to linear quiver gauge theories with alternating Sp and SO gauge groups. The reason for requiring the SO and Sp gauge groups to alternate is that only such gauge theories can be engineered using an orientifold D4/NS5-brane set-up. These configurations are natural from the gauge theory perspective as well. Remember that the flavor symmetry for an $Sp(N-1)$-fundamental hyper enhances to $SO(2N)$, while the flavor symmetry for $SO(N)$-fundamental hyper enhances to $Sp(N-1)$. So, for general N, only linear quivers with alternating gauge groups $Sp(N-1)$ and $SO(2N)$ correctly reproduce the flavor symmetry of the bifundamental fields. An example of a linear Sp/SO quiver is illustrated in Figure 5.7.

Special about linear Sp/SO quivers is that the bifundamental fields are not full hyper-

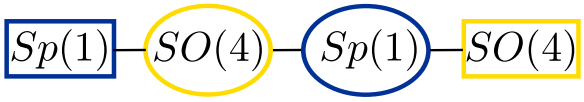


Figure 5.7: Example of a linear $Sp(1)/SO(4)$ quiver gauge theory with a single $Sp(1)$ and $SO(4)$ gauge group, one $SO(4)$-fundamental hyper, two $Sp(1)$-fundamental hypers and one $SO(4) \times Sp(1)$-bifundamental hyper (consisting of eight half-hypermultiplets).

multiplets, but half-hypermultiplets. Let us discuss this briefly. Usually, a hypermultiplet of representation R of a gauge group G consists of two $\mathcal{N} = 1$ chiral multiplets: one chiral multiplet in the representation R and the other in the complex conjugate representation $\bar{R}$ of G. When the representation R is pseudoreal, however, a single chiral superfield already forms an $\mathcal{N} = 2$ hypermultiplet. This is called a half-hypermultiplet in R. The half-hypermultiplets must be massless, as it is not possible to construct a gauge invariant mass-term in the Lagrangian for a half-hypermultiplet.

Even though a half-hypermultiplet is CPT invariant, it is not always possible to add them to an $\mathcal{N} = 2$ gauge theory due to the Witten anomaly [93]. Because an $Sp \times SO$ bifundamental multiplet contains an even number of half-hypermultiplet components, we can circumvent the anomaly. Indeed, the $Sp(N) \times SO(M)$ bifundamental is the tensor product of $2N$ half-hypermultiplets corresponding to the (anti-)fundamental $Sp(N)$ flavor symmetry, and M half-hypermultiplets corresponding to the fundamental $SO(M)$ flavor symmetry. In total this gives $2NM$ half-hypermultiplets.

Our goal in this section is to write down Nekrasov contour integrands for linear SO/Sp quivers and verify the correspondence with chiral blocks of the $\mathcal{W}$-algebra. Before getting there, let us first discuss some of the geometry of linear SO/Sp quivers.

5.4.1 UV-curves for linear Sp/SO quivers

As illustrated in Figure 5.8, the orientifold D4/NS5 brane constructions for Sp and SO gauge theories can be naturally extended to any linear quiver theory with alternating $Sp(N-1)$ and $SO(2N)$ gauge groups by introducing an extra NS5-brane for every bifundamental field. For this construction to work it is necessary that the gauge groups alternate as crossing an NS5-brane exchanges one type of orientifold brane with the other. From this string theory embedding we can simply read off the Seiberg-Witten curve.

The Seiberg-Witten curve corresponding to a linear quiver with $Sp(N-1)$ as well as

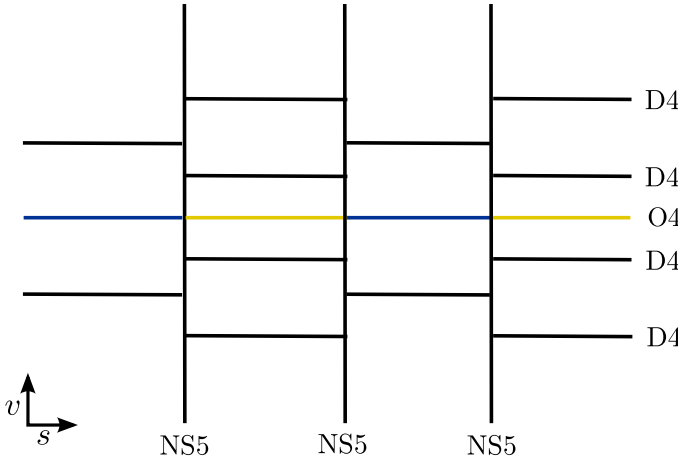


Figure 5.8: Orientifold D4/NS5-brane embedding of the linear Sp/SO quiver theory of Figure 5.7.

$SO(2N)$ gauge groups can be written in the Hitchin-form [38]

$$0 = \det(v - \boldsymbol{\varphi}_{Sp/SO}) = v^{2N} + \varphi_2 v^{2N-2} + \varphi_4 v^{2N-4} + \ldots + \varphi_{2N}. \tag{5.72}$$

As before, this equation determines the Seiberg-Witten curve as a degree $2N$ covering over the UV-curve. The Hitchin differentials φ_{2k} (for $1 \leqslant k \leqslant N-1$) are of degree $2k$ and encode the vev's of the Coulomb branch operators $\mathrm{Tr}(\Phi^{2k})$ of the adjoint scalar Φ for all gauge groups in the linear quiver. On the other hand, the degree N differential $\varphi_{\tilde{N}} = \sqrt{\varphi_{2N}}$ encodes the vev's of the operators Pfaff(■) for the SO gauge groups in the quiver only. All differentials are also functions of the exactly marginal coupling constants τ_{UV} and the bare mass parameters, in such a way that the residue of the matrix-valued differential $\boldsymbol{\varphi}_{Sp/SO}$ at each puncture encodes the flavor symmetry of the corresponding matter multiplet.

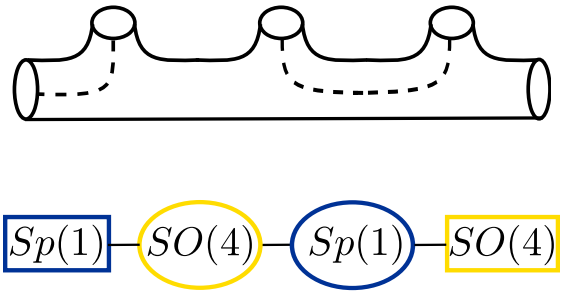


Figure 5.9: The Riemann surface on top is the UV-curve corresponding to the linear $Sp(1)/SO(4)$ quiver at the bottom.

It follows from equation (5.72) that the UV-curve for a linear $Sp(N-1)/SO(2N)$ quiver theory is a genus zero Riemann surface with punctures. For the $Sp(1)/SO(4)$ theory these punctures can be of two types. Either the differential $\varphi_{\tilde{2}} = \sqrt{\varphi_4}$ experiences a $\mathbb{Z}_2$-monodromy when going around the puncture, or it does not. As before, we call these punctures half-punctures and full punctures, respectively. The puncture representing a bifundamental matter field is a half-puncture. One way to understand this, is to compare the gauge theory quiver to the corresponding UV-curve. As is illustrated in Figure 5.9 each $Sp(1)$ gauge group corresponds to a tube with a $\mathbb{Z}_2$-twist line on the UV-curve, whereas each $SO(4)$ gauge group corresponds to a tube without a $\mathbb{Z}_2$-twist line. This implies that at each puncture corresponding to a bifundamental field a $\mathbb{Z}_2$-twist line has to end. Notice that the differential $\boldsymbol{\varphi}_{Sp/SO}$ should have a vanishing residue at this half-puncture, since the bifundamental is forced to have zero mass.

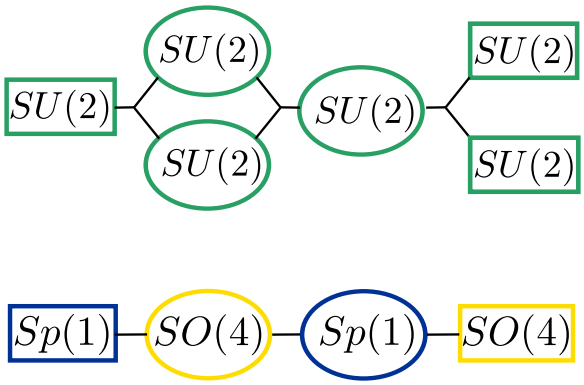


Figure 5.10: The generalized $SU(2)$ quiver theory, depicted at the top, has isomorphic gauge and flavor symmetries to the linear $SO(4)/Sp(1)$ quiver gauge theory, depicted on the bottom. This picture in particular relates the $SO(4) \times Sp(1)$ bifundamental to the $SU(2)^3$ trifundamental.

A remarkable feature of *linear* $Sp(1)/SO(4)$ quivers is that they are closely related to *generalized* $SU(2)$ quiver theories [38]. We have already seen that an $Sp(1)$-fundamental can be equivalently represented by an $SU(2)$-fundamental, and an $SO(4)$-fundamental by an $SU(2)^2$-bifundamental, as their representations are isomorphic. Even more interestingly, by the same argument an $Sp(1) \times SO(4)$-bifundamental is closely related to a matter multiplet with flavor symmetry group $SU(2)^3$. The elementary field with this property is known as the $SU(2)^3$-trifundamental, and it consists of eight free half-hypermultiplets in the fundamental representation of the three $SU(2)$ gauge groups [25]. Since the $Sp(1) \times SO(4)$

bifundamental contains eight half-hypermultiplets as well, we expect it to be equivalent to the $SU(2)^3$-trifundamental. One example of the relation between linear $Sp(1)/SO(4)$ quivers and generalized $SU(2)$ quiver theories is illustrated in Figure 5.10.

Similar to our discussion in section 3.5, we can interpret the $\mathbb{Z}_2$-twist lines on the Sp/SO UV-curve as branch cuts. As is illustrated in Figure 5.11, we find that the UV-curve for the generalized A_1 quiver theory is a double cover of the UV-curve for the corresponding linear D_2 quiver theory. This is consistent with the fact that the bifundamental fields cannot carry a mass. Indeed, each half-puncture corresponding to a $Sp(1) \times SO(4)$ bifundamental field lifts to a regular point on the $SU(2)$ UV-curve. Its flavor symmetry had thus better be trivial.

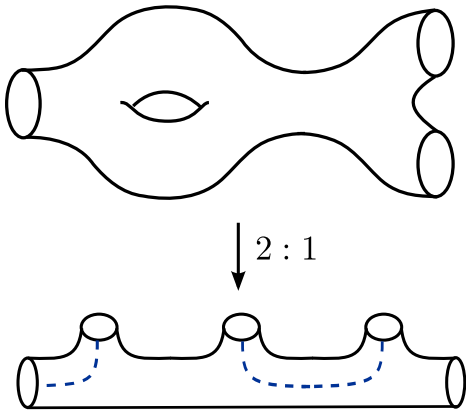


Figure 5.11: The top picture represents the UV-curve of a generalized $SU(2)$ quiver theory. It is a branched double cover over the UV-curve of the linear $SO(4)/Sp(1)$ quiver theory illustrated in the bottom.

Let us stress that, according to the arguments of section 3.6.1, the instanton partition function of a linear quiver theory that contains the $Sp(1) \times SO(4)$ bifundamental will be related to the instanton partition function containing the $SU(2)^3$-trifundamental by a non-trivial mapping of marginal gauge couplings. This mapping has a geometric interpretation, according to section 3.5, as it will relate the complex moduli of the corresponding UV-curves. Studying the instanton partition function of linear D_2 quiver theories thus sheds light on the understanding of non-linear A_1 quiver theories. We will discuss it in section 5.5

5.4.2 $Sp \times SO$ bifundamental

In analogy to [118], we can view the bifundamental half-hypermultiplet as a mapping

$$\Phi_{a,b_1,b_2} : \hat{\mathcal{H}}_{\widetilde{Sp}(1)} \to \hat{\mathcal{H}}_{SO(4)} \tag{5.73}$$

between two vector spaces $\hat{\mathcal{H}}_{\widetilde{Sp}(1)}$ and $\hat{\mathcal{H}}_{SO(4)}$, whose bases are parametrized by the poles of the respective gauge multiplet integrands. It is natural to expect that these vector spaces are related to the $\mathcal{W}$-representation spaces $\mathcal{H}_{\widetilde{Sp}(1)}$ and $\mathcal{H}_{SO(4)}$ that we encountered in the previous section. Note that for Sp and SO gauge groups we expect the spaces to be related without any additional $U(1)$ factors. For the original $U(2)$ AGT correspondence this was recently made precise [127]. The structure of the $Sp(1)$ poles is more complicated, however, and it would be interesting to find the exact mapping between the two spaces.

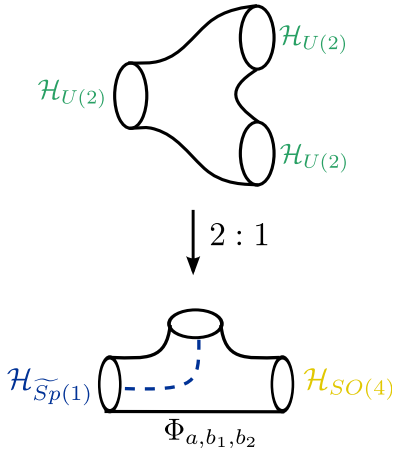


Figure 5.12: The instanton contribution for the $U(2)^3$ trifundamental can be represented as a linear map $Z^{\text{inst}} : \mathcal{H}_{U(2)} \to \mathcal{H}_{U(2)} \times \mathcal{H}_{U(2)}$. Similarly, and correspondingly, the $Sp(1) \times SO(4)$ bifundamental field defines a linear map $Z^{\text{inst}} : \hat{\mathcal{H}}_{\widetilde{Sp}(1)} \to \hat{\mathcal{H}}_{SO(4)}$.

5.4.3 Test of the $Sp(1) \times SO(4)$ AGT correspondence

Since we know the UV-curve corresponding to alternating SO/Sp quiver theories, we can extend the correspondence between SO/Sp gauge groups and $\mathcal{W}$ blocks. We are now ready to check this correspondence. At the same time this will also serve as an additional check that our expression for the half-hypermultiplet is correct. Consider thus the SO/Sp theory illustrated in Figure 5.7, with a single bifundamental half-hypermultiplet, two fundamen-

tal $Sp(1)$-hypermultiplets and one fundamental $SO(4)$-hypermultiplet.

5.4.3.1 $Sp(1) \times SO(4)$ instantons

Let us first compute the instanton partition function of this linear quiver theory. The first non-trivial term comes from $(k_1, k_2) = (1, 1)$. In the unrefined case $\epsilon_1 = -\epsilon_2 = \hbar$, it is given by

$$Z^{\text{inst}}_{1,1} = -\frac{m_1 m_2 b_1 b_2 \left(\left(m_3^2 - b_1^2\right)\left(-a^2 + b_1^2\right)\left(-\hbar^2 + b_1^2\right) + \left(a^2 - b_2^2\right)\left(-m_3^2 + b_2^2\right)\left(-\hbar^2 + b_2^2\right)\right)}{8a^2\hbar^4\left(b_1^2 - b_2^2\right)^2}$$

where m_1, m_2 are the masses of the $Sp(1)$-fundamentals, m_3 is the mass of the $SO(4)$-fundamental and a, b_i are the Coulomb branch parameters of $Sp(1)$ and $SO(4)$, respectively.

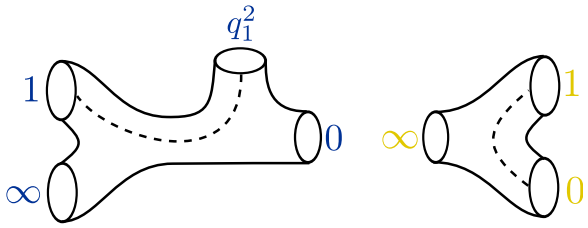


Figure 5.13: Decomposition of the $Sp(1) \times SO(4)$ UV-curve that we used for computing the corresponding $\mathcal{W}(2,2)$-block.

5.4.3.2 $Sp(1) \times SO(4)$ correlators

The correlator that corresponds to the (mirror of the) quiver illustrated in Figure 5.7 is the following:

$$\langle V_1(\infty)\sigma(1)\sigma(q_1^2)\sigma(q_1^2 q_2)\hat{V}_2(0)\rangle \tag{5.74}$$

For notational simplicity we have introduced the variables $q_1 = q_{Sp(1)}/4$ and $q_2 = q_{SO(4)}$. A single term in the chiral block is then expanded as (see Figure 5.13)

$$\begin{aligned}\langle V_1(\infty)\sigma(1)\sigma(q_1)V^{A,B}(0)\rangle\langle V^{A,B}(\infty)\sigma(q_1^2 q_2)\hat{V}_2(0)\rangle \\ = \langle V_1(\infty)\sigma(1)\sigma(q_1^2)V^{A,B}(0)\rangle(q_1^2 q_2)^{h_A+n_A+h_B+n_B}\langle V^{A,B}(\infty)\sigma(1)\hat{V}_2(0)\rangle\end{aligned} \tag{5.75}$$

To compute the rightmost correlator, we can proceed as in subsection 5.3.3.2 and map it to a cover correlator of the form (5.68). Due to the corrections we will get something of the form

$$\langle V^{A,B}(\infty)\sigma(1)\hat{V}_2(0)\rangle = \left(\frac{1}{2}\right)^{n_A+n_B} \tilde{\mathbf{C}}^{h_A,h_B;h_1}_{I_A,I_B,\bullet} + \text{lower descendant corrections}\,. \tag{5.76}$$

The correlator is thus the same as the one we computed in the section on $SO(4)$.

The four point correlator on the left we treat as in the $Sp(1)$ case, i.e., we apply the cover map (5.61). The only difference is then that $V^{A,B}$ is a descendant field, and thus, as in the $SO(4)$ computation, picks up corrections from the map:

$$\begin{aligned} V^{A,B}(0) = V\left(\left(\frac{-1+q_1^2}{2q_1}\right)^{L_0} \exp\left[\left(\frac{3}{4}+\frac{1}{4q_1^2}\right)L_1\right]\exp\left[\frac{(-1+q_1^2)^2}{16q_1^4}L_2\right]\cdots\phi^A, q_1\right) \\ \times V\left(\left(-\frac{-1+q_1^2}{2q_1}\right)^{L_0} \exp\left[\left(\frac{3}{4}+\frac{1}{4q_1^2}\right)L_1\right]\exp\left[\frac{(-1+q_1^2)^2}{16q_1^4}L_2\right]\cdots\phi^B, -q_1\right) \end{aligned} \tag{5.77}$$

This expression is however different in two ways from the corresponding $SO(4)$ expression (5.69). First, the vertex operators are at the positions $\pm q_1$. We decompose the four punctured correlator on the cover in usual way, and move them to the standard positions ± 1 using the map $z \mapsto zq_1^{-1}$. This simply leads to an additional prefactor $q_1^{-L_0}$ in equation (5.77). To pull out the standard prefactor $q_1^{h_{A,B}}$, it is useful to commute this prefactor all the way to the left, which we can do by using the identity

$$x^{L_0}L_n = \frac{L_n}{x^n}x^{L_0}\,. \tag{5.78}$$

This leads to the expression

$$\begin{aligned} &\left(\frac{-1+q_1^2}{2q_1^2}\right)^{h_A+n_A+h_B+n_B} \times \\ &\langle\phi|V\left(\exp\left[\frac{1+3q_1^2}{2(-1+q_1^2)}L_1\right]e^{L_2/4}\cdots\phi^A, 1\right)V\left(\exp\left[-\frac{1+3q_1^2}{2(-1+q_1^2)}L_1\right]e^{L_2/4}\cdots\phi^B, -1\right)\rangle\,. \end{aligned} \tag{5.79}$$

Note that the $q_1^{2(h_A+n_A+h_B+n_B)}$ in the denominator of the prefactor exactly cancels the corresponding factor of q_1 in equation (5.75). The numerator of the prefactor on the other hand is the same prefactor that we already found in the $Sp(1)$ computation.

Let us now actually compute the first few terms of the chiral block. The terms where of order zero in either $q_{Sp(1)}$ or $q_{SO(4)}$ are are simply the same as in the $SO(4)$ and $Sp(1)$ computation. We therefore consider the simplest new term, $q_{Sp(1)}q_{SO(4)}$. This means in particular that we can neglect all terms of order q_1^2 in the expression (5.79), so that the vertex operators no longer depend on q_1. Since we would like to rewrite this expression in terms of the three point coefficients $\tilde{\mathbf{C}}$ defined above, we need move the field ϕ from ∞ to 0 by applying the map $z \mapsto z^{-1}$. Note that we pick up some additional corrections due to the fact that the fields at ± 1 are descendant fields. In total (5.79) thus becomes

$$(2^{-n_A-n_B})\langle V\left(e^{3L_1/2}e^{L_2/4}\cdots\phi^A,1\right)V\left(e^{-3L_1/2}e^{L_2/4}\cdots\phi^B,-1\right)|\phi\rangle\,. \tag{5.80}$$

Not surprisingly, this is the same expression that we had found in the $SO(4)$ case. From this, the term of order $q_{Sp(1)}q_{SO(4)}$ is

$$\begin{aligned}\frac{\tilde{\mathbf{C}}^{h_1,h_2;h_a}_{\bullet,\bullet,\square}}{2h_a}\frac{1}{4}\Big(&\frac{(\tilde{\mathbf{C}}^{h_A,h_B;h_a}_{\square,\bullet,\square}+3h_A\tilde{\mathbf{C}}^{h_A,h_B;h_a}_{\bullet,\bullet,\square})(\tilde{\mathbf{C}}^{h_A,h_B;h_3}_{\square,\bullet,\bullet}+3h_A)}{2h_A}\\&+\frac{(\tilde{\mathbf{C}}^{h_A,h_B;h_a}_{\bullet,\square,\square}-3h_B\tilde{\mathbf{C}}^{h_A,h_B;h_a}_{\bullet,\bullet,\square})(\tilde{\mathbf{C}}^{h_A,h_B;h_3}_{\bullet,\square,\bullet}-3h_B)}{2h_B}\Big).\end{aligned} \tag{5.81}$$

Similar computations lead to higher order terms. We have checked that the instanton partition function and the chiral block agrees up to order $(k_1,k_2)=(1,2)$ up to a moduli independent spurious factor, using the same identifications of the previous examples.

5.5 Sicilian quivers

5.5.1 CFT building blocks for Sicilian quivers

Let us now discuss the building blocks that are needed for the AGT correspondence. In the correspondence for conformal $SU(2)$ quiver gauge theories hypermultiplets are given by punctures on the UV-curve. Gluing the neighborhoods of two punctures to create a tube gauges the flavor symmetry group of the two hypermultiplets into an $SU(2)$ gauge group. The masses of the two hypermultiplets have to be opposite to perform the gluing, since they correspond to the residue of the Seiberg-Witten 1-form at the puncture. The masses then turn into the Coulomb parameters $\pm a$ of the $SU(2)$ gauge group after the gluing.

On the CFT side, hypermultiplets correspond to insertions of primary fields ϕ^i whose

conformal weights are related to the masses of the hypermultiplets. A gauge group corresponds to inserting a complete set of descendants of a given primary field. We recall that an arbitrary Virasoro descendant ϕ_I at level N is given by a partition I of N by $\phi_I = \prod_j L_{-I_j}\phi$. For ease of notation we will also just write I for N. The projector on a particular representation that we insert can thus be written as

$$P_{\mathcal{H}_\phi} = \sum_{I,J} K^{-1}_{IJ} |\phi_J\rangle\langle\phi_I| \,, \tag{5.82}$$

where K^{-1} is the inverse of the Kac matrix $(K)_{IJ} = \langle\phi_I|\phi_J\rangle$. The modulus of the tube corresponds to the coupling of the gauge group. From this it is clear that if we decouple the gauge group by sending $q \to 0$ we recover the original expression for the ungauged theory, since the contributions of the descendants vanish and only the primary field survives.

The complete instanton partition function can thus be obtained from a pair of pants decomposition of the UV-curve. Its building blocks are given by three-point functions containing one or more descendant fields, and the total expression is obtained by summing over all descendant fields in the channels. This sum corresponds to the sum over the fixed points in the instanton counting. For linear and cyclic quivers, the only building blocks needed are hypermultiplets in the fundamental and hypermultiplets in the bifundamental. The corresponding CFT expressions are three-point functions with one or two descendant fields.

For Sicilian quivers such as in figure 5.14, however, we also need hypers in the fundamental of three different gauge groups. The corresponding CFT building block should then be described by the three-point function with three descendant fields inserted,

$$\langle V(\phi^1_{I_1}, z_1) V(\phi^2_{I_2}, z_2) V(\phi^3_{I_3}, z_3)\rangle \,. \tag{5.83}$$

Here we have used the notation $V(\phi, z)$ for the the vertex operator corresponding to the field ϕ inserted at z. The weights of the fields ϕ^i are related to the Coulomb branch parameters $a_{1,2,3}$ of the three $SU(2)$ gauge groups involved. Choosing the insertion points z_i is quite subtle and affects the outcome, as we will now discuss.

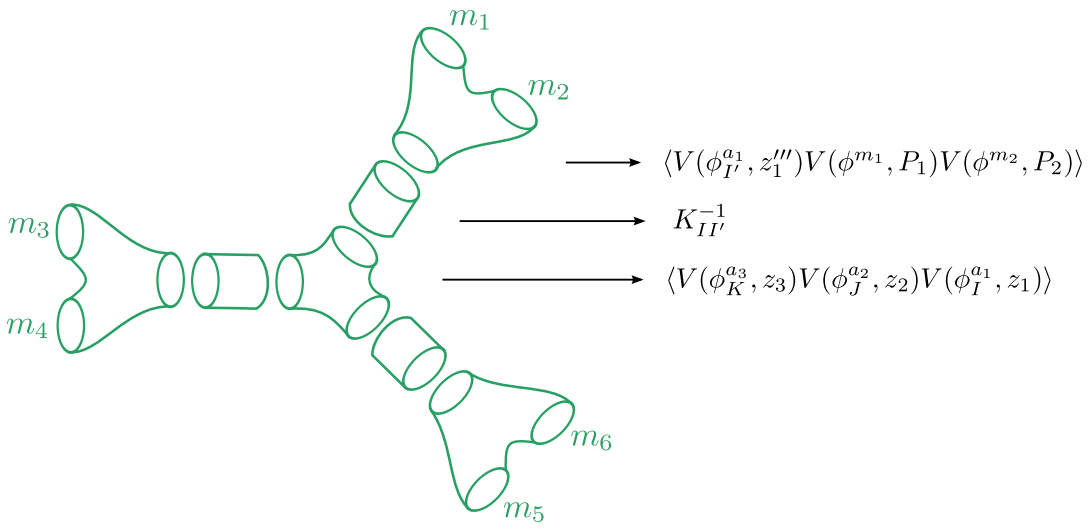


Figure 5.14: Decomposition of the sphere with six punctures into three-punctured spheres and tubes, and the corresponding conformal blocks

5.5.1.1 Three-point functions

Let us start with a reminder about three-point functions and some of their properties. For three primary fields the three-point function is fixed up to a constant C_{123}. The coordinate dependence itself is fixed covariance under by Möbius transformations, i.e., the global conformal symmetry.

Local conformal symmetry allows us to compute three-point functions of arbitrary descendants of those primary fields as well. In principle, this is straightforward: the only thing needed is the OPE of the stress energy tensor $T(z)$ with the primary fields with itself. We can then use

$$\langle V(L_{-n}\phi, z)\ldots\rangle = \oint_z dw\,(w-z)^{-n+1}\langle T(w)\,V(\phi, z)\cdots\rangle \tag{5.84}$$

to reduce the three-point function to contour integrals of the correlator of three primary fields and several energy stress tensors. This correlator is a meromorphic function on a Riemann surface and thus determined by its poles. We can thus consecutively eliminate the $T(z)$ by summing their OPEs with the other $T(w)$ and the primary fields, until we are left with just the three-point function of the primary fields. We can then evaluate the contour integral.

Though conceptually simple, in practice this procedure is quite cumbersome. Since most of the time we are interested in very specific values of z_i only, it can be more efficient to phrase the computation in terms of operators on the Hilbert space of a Virasoro representation. The operator-state correspondence tells us that

$$\lim_{z\to 0}\phi(z)|0\rangle = |\phi\rangle\,. \tag{5.85}$$

The corresponding bra state is given by the operator at infinity. More precisely, it is obtained from the ket state using the Möbius transformation $z \mapsto 1/z$:

$$\lim_{z\to 0}\langle 0|V(z^{-2L_0}e^{-\frac{1}{z}L_1}\phi, 1/z) = \langle\phi|\,. \tag{5.86}$$

The three-point function with primary fields at 0,1,∞ can then be computed as [4]

$$\langle\phi^1|V(\phi^2,1)|\phi^3\rangle = \langle\phi^1|\phi^2_{h_3-h_1}|\phi^3\rangle = C_{123}\,. \tag{5.87}$$

We can compute such three-point functions with descendant bra and ket states by commuting through all Virasoro operators using

$$[L_n,\phi_m] = (n(h-1)-m)\phi_{m+n} \tag{5.88}$$

for primary fields ϕ^2. If ϕ^2 is a descendant, then we first need to express it in terms of Virasoro operators and modes of the primary field, which we do by using the following expression for the $-N_1$ mode of a $V(L_{-N_2}\phi,z)$ [128]

$$\begin{aligned} V_{-N_1}(L_{-N_2}\phi) = &\sum_{l\geqslant 0}\binom{N_2-2-l}{l}L_{-N_2-l}V_{-N_1+N_2+l}(\phi) \\ &+(-1)^{N_2}\sum_{l\geqslant 0}\binom{N_2-2-l}{l}V_{-N_1-l+1}(\phi)L_{l-1}\,. \end{aligned} \tag{5.89}$$

Note that even though the sums are infinite, they reduce to finite sums when acting on any particular state.

Consider a theory with several identical gauge groups. One would expect that the par-

[4] Strictly speaking we can only do this for $h_3 - h_1 \in \mathbb{Z}$. From general arguments we know however that the coefficients of the conformal block are given by rational functions in h_i and c. The expressions we obtain thus continue to be valid for arbitrary values of h.

tition function should be symmetric under suitable permutations of the gauge group. On the CFT side this means that the three-point function should be symmetric under permutations of the insertion points.

For instance, if a theory contains a hypermultiplet in the fundamental of two $SU(2)$ groups, then the three-point function must be symmetric under exchanging the two. This is indeed the case, as follows from

$$\langle \phi^1_{I_1} | V(\phi^2, 1) | \phi^3_{I_3} \rangle = \langle \phi^3_{I_3} | V(\phi^2, 1) | \phi^1_{I_1} \rangle \,. \tag{5.90}$$

To see that (5.90) indeed holds we can use the Möbius transformation $z \mapsto 1/z$. A general field transforms under a Möbius transformation γ as

$$V(\phi, z) \mapsto V\left((\gamma(z)')^{L_0} e^{\frac{\gamma''}{2\gamma'} L_1} \phi, \gamma(z) \right) \,. \tag{5.91}$$

From this we see that as long as ϕ^2 is a primary field, it does not pick up any correction terms from this transformation.

On the other hand, if we consider the case of a hyper in the fundamental of three gauge groups, we need to insert three descendants, and $V(\phi^2, 1)$ will no longer transform in such a simple way. The usual vertex is then no longer symmetric under permutations, as the Möbius transformations that exchange punctures introduce corrections. This means that the standard CFT vertex must correspond to a regularization scheme of the gauge theory which treats the gauge groups differently.

More generally, if we use any Möbius transformation to change the insertion points of a three-point function with descendants, then due to (5.91) we will pick up corrections. This means that the detailed expression for the three-point function greatly depends on the choice of insertion points z_i in (5.83). It turns out that these issues are less severe for asymptotically free theories. Let us therefore turn to those cases.

5.5.1.2 Partition function for the trifundamental coupling

Conformal blocks a priori correspond to conformal gauge theories, as the flavor symmetries always work out in such a way that there are four fundamental hypers per $SU(2)$ gauge group. We can however obtain asymptotically free theories by sending the mass of

hypers to infinity and so decoupling them. More precisely, to decouple a hyper of mass m in the fundamental of a gauge group of coupling q we take

$$q \to \Lambda/m\,, \qquad m \to \infty\,. \tag{5.92}$$

Here Λ is the scale of the newly asymptotically free theory. In this way we can obtain any asymptotically free partition function from a conformal block.

Let us use the procedure outlined above to compute the partition function of a half-hypermultiplet in the trifundamental of $SU(2)$. We start out with conformal theory which corresponds to a sphere with six punctures (see figure 5.14), but decompose it in a symmetric (i.e., non-linear) way:

$$\begin{aligned} Z = \sum_{I_1,I_2,I_3} \sum_{J_1,J_2,J_3} &\langle \phi^{m_1} | V(\phi^{m_2},1) | \phi^{a_1}_{I_1} \rangle \langle \phi^{m_3} | V(\phi^{m_4},1) | \phi^{a_2}_{I_2} \rangle \langle \phi^{m_5} | V(\phi^{m_6},1) | \phi^{a_3}_{I_3} \rangle \\ &\times K^{-1}_{I_1 J_1} K^{-1}_{I_2 J_2} K^{-1}_{I_3 J_3} \langle \phi^{a_1}_{J_1} | V(\phi^{a_2}_{J_2},1) | \phi^{a_3}_{J_3} \rangle q_1^{I_1} q_2^{I_2} q_3^{I_3}\,. \end{aligned} \tag{5.93}$$

Note that we have chosen more or less by fiat that the trifundamental vertex, i.e., the three-punctured sphere in the center of the decomposition, is given by the sphere with punctures at $0, 1, \infty$. In view of the remarks in the previous section the result is certainly not symmetric under permutation of the gauge groups. To obtain the asymptotically free theory, that is the result for a single half-hyper in the trifundamental, we apply (5.92). It turns out that the resulting expression *is* symmetric under permutations up to spurious terms (which we explain in a moment). It is moreover independent on the choice of punctures of the three-punctured sphere in the center of the decomposition, up to a simple rescaling of the couplings q.

This rather surprising result can be better understood when computing asymptotically free theories using Gaiotto states [120]. Such a state $|h, \Lambda\rangle$ is an eigenstate of the Virasoro mode L_1 with eigenvalue Λ,

$$L_1 |h,\Lambda\rangle = \Lambda |h,\Lambda\rangle \qquad L_n |h,\Lambda\rangle = 0 \quad n \geqslant 2\,. \tag{5.94}$$

More concretely such as state can be written as a power series in Λ

$$|h, \Lambda\rangle = \sum_{n=0}^{\infty} \Lambda^n |v_n\rangle \,, \tag{5.95}$$

where $|v_0\rangle = |h\rangle$ and $|v_n\rangle$ is a specific linear combination of Virasoro descendants of $|h\rangle$ at level n. These states can then be used to compute instanton partition functions for asymptotically free $SU(2)$ theories. The norm of such a state, for instance, gives the instanton partition function of pure $SU(2)$ gauge theory. Both states in this norm originate from decoupling a pair of hypermultiplets in the conformal $SU(2)$ gauge theory. The conditions (5.94) come from the poles of the quadratic differential $\phi_2(z)$ on the UV-curve. (See [121] for a proof that this is equivalent to the infinite mass limit.)

It is natural to use the same strategy also for multiple gauge groups. The $SU(2)$ trifundamental can be obtained by decoupling three pairs of hypers in the conformal $SU(2)$ gauge theory corresponding to the six-punctured sphere. We thus compute the three-point function

$$Z_{\text{CFT}} = \langle h_1, \Lambda_1 | V(|h_2, \Lambda_2\rangle, 1) | h_3, \Lambda_3 \rangle \,. \tag{5.96}$$

This gives indeed the same expression as the one we obtained above. Now we can also explain why (5.96) is invariant under permutation of the three gauge groups (up to some trivial factors). As usual we use a Möbius transformation γ to exchange the three insertion points. From (5.91) and (5.94) it follows that the Gaiotto state $|h, \Lambda\rangle$ transforms to

$$e^{\frac{\gamma''}{2\gamma'}\Lambda} (\gamma')^h |h, \gamma' \Lambda\rangle \,, \tag{5.97}$$

so that after a redefinition of Λ the two three-point functions only differ by a spurious prefactor. Since this holds for any Möbius transformation, the result is essentially independent of the insertion points.

We propose that (5.96) is equal to the instanton partition function of a half-hyper in the trifundamental representation of $SU(2)$ (up to a spurious factor [5]). Even though we did not compute this partition function directly, we can perform several consistency checks on

[5] In the following we define a spurious factor as a factor that does not depend on the Coulomb branch parameters and only contributes to the first terms of the genus expansion of the free energy.

(5.96). First note that it has indeed a proper $\mathcal{F}_g$ expansion, i.e., that it can be written

$$Z = \exp\left(\sum_{g\geqslant 0} \hbar^{2g-2}\mathcal{F}_g\right), \tag{5.98}$$

with no higher negative powers of $\hbar$ appearing. Second, (5.96) reduces correctly to the $SU(2)$ bifundamental when we decouple one of the gauge groups. Finally, when setting $\Lambda_2 = \Lambda_3$ it agrees with the partition function of a $Sp(1) - SO(4)$ gauge theory with a single hyper in the bifundamental (details of this check can be found in section 5.5.3).

5.5.2 Towards a 4d/2d correspondence for Sicilian quivers

The simplest way to define a conformal $\mathcal{N} = 2$ Sicilian $SU(2)$ quiver gauge theory is through its M-theory construction. Wrap two M5 branes on a Riemann surface with punctures $\mathcal{C}$. The quiver theory corresponding to a particular duality frame is obtained from a decomposition of $\mathcal{C}$ into pairs of pants. The punctures of $\mathcal{C}$ correspond to hypermultiplets, and the tubes connecting the different pants correspond to $SU(2)$ gauge groups whose microscopic coupling constants are given by the complex structure moduli of the tubes.

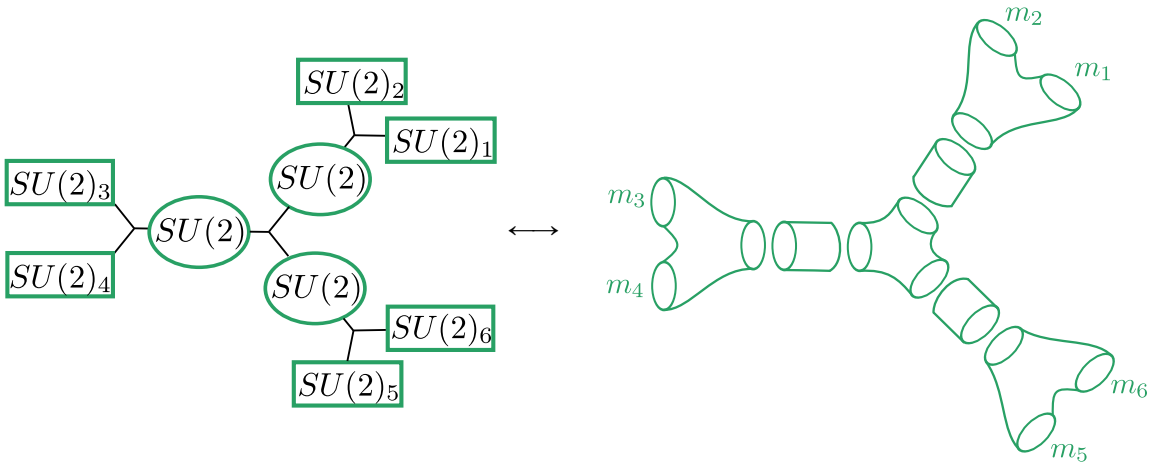


Figure 5.15: Illustration of the correspondence between instanton partition functions of Sicilian $SU(2)$ quiver gauge theories and Virasoro conformal blocks on the corresponding UV-curve for the six-punctured sphere. Each $SU(2)$ gauge group in the quiver is mapped to a tube in the UV-curve, whereas $SU(2)$ matter is represented by three-punctured spheres.

The building blocks are thus spheres with three punctures or tubes. There are three

different configurations. The sphere with two punctures and one tube corresponds to two hypermultiplets in the fundamental. The sphere with one puncture and two tubes corresponds to a hyper in the bifundamental of the two $SU(2)$. Finally, as a new element, there is the the sphere with three tubes. It corresponds to a half-hyper in the trifundamental. Since the half-hyper is massless, it is natural not to have a puncture for it in this building block. See figure 5.15 for an example.

Quivers with asymptotically free gauge groups can always be obtained from conformal theories by sending the mass of one of the hypers to infinity.

We can then compute the conformal block for this quiver in the following way. First, at every puncture insert a primary field whose conformal weight is given by the mass of the hyper in the usual way. Second, for every tube insert a projector

$$P_{\mathcal{H}_\phi} = \sum_{I,J} K_{IJ}^{-1} |\phi_J\rangle\langle\phi_I|$$

onto the channel that corresponds to the Coulomb branch parameter of the $SU(2)$ gauge group. The bra and ket state of that projector are inserted in the respective building blocks. The problem thus reduces to computing various three point functions

$$\langle V(\phi^1_{I_1}, z_1) V(\phi^2_{I_2}, z_2) V(\phi^3_{I_3}, z_3)\rangle \,,$$

of primary or descendant fields. As pointed out above, the subtlety lies in the choice of insertion the points z_i. For linear and cyclic quivers, all the building blocks have only one or two descendant fields inserted. Using the usual coordinates on the sphere or torus, the descendant fields are always inserted at 0 or ∞, that is as bra and ket states, and there is always a primary field inserted at 1. Using this prescription the conformal block agrees with the $SU(2)$ instanton partition function.

For trifundamental hypers the situation is more subtle. We can insert three descendant fields at the points $0, 1, \infty$, but in general the result will not agree with the instanton computation, since we are using a different parametrization of the moduli space. Once expressed in IR variables, the results will agree. To put it another way, there will be a map between the moduli space coordinates and the microscopic gauge coupling that will make them agree. Or more geometrically, the CFT correlators define a unique object on the

Seiberg-Witten curve, that is independent on the chosen parametrization of the complex structure moduli space of the UV-curve.

The situation is much simpler for asymptotically free gauge groups. In this case the conformal block will agree with the instanton partition function immediately, and will be essentially independent of the choice of insertion points.

Comparison with Nekrasov partition function: one-loop factor

As we consider theories with $\mathcal{N} = 2$ supersymmetry, the full Nekrasov partition function has tree-level, one-loop and instanton contributions,

$$Z_{\text{Nek}} = Z_{\text{clas}}(\tau_{\text{UV}})\, Z_{\text{1-loop}}\, Z_{\text{inst}}(\tau_{\text{UV}})\,. \tag{5.99}$$

The 4d/2d correspondence relates the purely representation-dependent piece of the Liouville correlator on the Gaiotto curve, that is the the conformal block, to the instanton partition function of the corresponding gauge theory in the Omega-background. Adding the classical contributions to the instanton partition function is crucial for finding good properties under coordinate changes on the complex structure moduli space of the Gaiotto curve (we spell this out explicitly in section 5.5.3). The one-loop factor can be identified with the three point function of Liouville theory. More properly, the full conformal block on the Gaiotto curve should be identified with the Nekrasov partition function on S^4 [26].

Let us check that this agreement continues to hold for Sicilian quivers. The one-loop factor can be found as a four-dimensional boson-fermion determinant in the Omega-background. Equivalently, it may be obtained from the equivariant index of the Dirac operator in the instanton background (see appendix E). The resulting contribution for the (full) $SU(2)$ trifundamental hypermultiplet is

$$\begin{aligned} Z^{\text{2trif}}_{\text{1-loop}} &= \prod_{n,m=1}^{\infty} \prod_{i,j,k=1}^{2} \left(a_i + b_j + c_k + \frac{Q}{2} + n\epsilon_1 + m\epsilon_2 \right)^{-1} \\ &\propto \prod_{i,j,k=1}^{2} \Gamma_2 \left(a_i + b_j + c_k + \frac{Q}{2} \middle| \epsilon_1, \epsilon_2 \right), \end{aligned} \tag{5.100}$$

where we take the Coulomb branch parameters $a_i = \pm a$, $b_j = \pm b$ and $c_k = \pm c$ of the three $SU(2)$ gauge groups and $Q = \epsilon_1 + \epsilon_2$. The Barnes' double gamma function $\Gamma_2(x|\epsilon_1, \epsilon_2)$ reg-

ularizes the infinite product. The one-loop partition function for the $SU(2)$ trifundamental half-hypermultiplet is given by a square-root of the above expression.

Agreement with the three-point function of Liouville theory follows by the same argument as for linear quivers [26]. Namely, the numerator of the DOZZ formula for the Liouville three-point function contains the product

$$\prod_{i,j,k=1}^{2} \Gamma_2\left(a_i + b_j + c_k + Q/2\right), \tag{5.101}$$

which equals the double trifundamental contribution in equation (5.100). Remember that the product (5.101) corresponds to the one-loop contribution of the Nekrasov partition function on S^4, which splits into a chiral and anti-chiral contribution on $\mathbb{R}^4$. Indeed, it is equal to the absolute value squared of the one-loop contribution for the $SU(2)$ trifundamental half-hyper, which, for example, can be written as

$$Z^{\text{trif}}_{\text{1-loop}} = \Gamma_2(a+b+c+Q/2)\Gamma_2(a+b-c+Q/2)\Gamma_2(a-b+c+Q/2)\Gamma_2(-a+b+c+Q/2).$$

5.5.3 Examples

In this section we test our proposal for extending the 4d/2d AGT correspondence to Sicilian quivers in the two examples illustrated in figure 5.16 and figure 5.17.

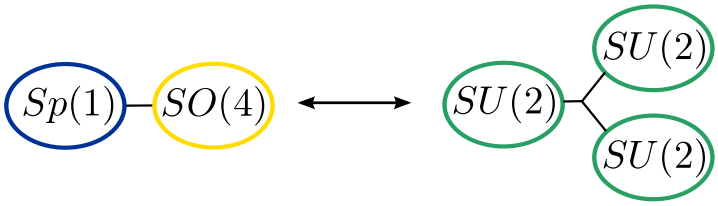


Figure 5.16: From a gauge theory perspective the $Sp(1) - SO(4)$ bifundamental, which is illustrated on the left, is equivalent to the $SU(2)$ trifundamental, which is illustrated on the right, once we identify two of the $SU(2)$ gauge couplings.

The quiver on the left in figure 5.16 consists of a single $Sp(1)$ gauge group and a single $SO(4)$ gauge group coupled by a bifundamental $Sp(1) - SO(4)$ half-hypermultiplet. It is equivalent to an $SU(2)$ Sicilian quiver gauge theory consisting of three $SU(2)$ gauge groups coupled by an $SU(2)$ trifundamental half-hypermultiplet, illustrated on the right in figure 5.16. The gauge couplings of both quivers are asymptotically free, so that the

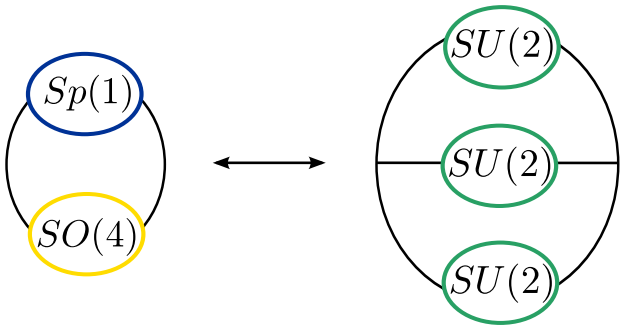


Figure 5.17: From a gauge theory perspective the cyclic $Sp(1) - SO(4)$ quiver, which is illustrated on the left, is equivalent to the genus 2 $SU(2)$ quiver, which is illustrated on the right, once we identify two of the $SU(2)$ gauge couplings.

instanton partition function should agree directly with the CFT block (5.96) without any subtleties involving a choice of coordinates. We check that this is indeed the case up to order 3.

The quiver in figure 5.17 is a conformal $Sp(1) - SO(4)$ gauge theory with two bifundamental $Sp(1) - SO(4)$ half-hypermultiplets, which is equivalent to a conformal $SU(2)$ Sicilian quiver gauge theory with three $SU(2)$ gauge groups coupled by two trifundamental $SU(2)$ half-hypermultiplets. Since the gauge theory is conformal, the results will depend on the choice of complex structure on the Gaiotto curve, and on the instanton counting scheme. Our proposal tells us which CFT configuration to choose to give direct agreement with the $Sp(1) - SO(4)$ instanton partition function, and we check that this indeed works up to order 3.

The conformal $SU(2)$ gauge theory can alternatively be described in terms of a massless full $SU(2)$ trifundamental hyper. So we can find its instanton partition function as well using the more conventional $U(2)$ instanton counting scheme.[6] We check that if we use the $U(2)$ trifundamental instanton counting scheme or choose different coordinates in the conformal block the results do agree in the IR. This confirms the general philosophy outlined above.

We have the instanton counting formulae for the $Sp(1) - SO(4)$ quiver gauge theories in chapter 3. They are given by a multiple contour integral of a meromorphic integrand.

[6]Notice that when we turn on the mass of this hypermultiplet, the theory does not have a string embedding anymore. This implies that we cannot find a Gaiotto curve. The Seiberg-Witten curve does exist, nevertheless, and can, for example, be found through a semi-classical approximation of the instanton partition function. See [129] for a related discussion.

This integrand consists of building blocks, each piece coming from a component of the quiver gauge theory. Especially, we find the contribution for the $Sp(1) - SO(4)$ bifundamental half-hypermultiplet as outlined in section 3.6.1. We also make a proposal the integrand for the full $SU(2)$ trifundamental hypermultiplet in appendix E. To actually evaluate these contour integrands, that is to find which of the poles contribute and to compute their residues, is an elaborate process, which we will describe later on.

5.5.3.1 The $SU(2)$ trifundamental as a $Sp(1) - SO(4)$ bifundamental

In this section we compute the instanton partition function of the $Sp(1) - SO(4)$ quiver gauge theory with a single bifundamental half-hypermultiplet. The quiver diagram is given in figure 5.16 and the corresponding Gaiotto curve is illustrated in figure 5.18.

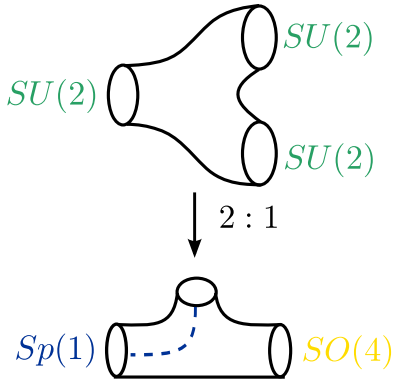


Figure 5.18: The UV-curve for the $SU(2)$ trifundamental is a double covering of the Gaiotto curve for the $Sp(1) - SO(4)$ bifundamental. (The corresponding quiver diagrams are illustrated in figure 5.16.)

Computing the instanton partition function

The instanton partition function of this theory is given by

$$Z_{\text{inst}}(q_1, q_2) = \sum_{k_1,k_2} q_1^{k_1} q_2^{k_2} Z_{k_1,k_2} \tag{5.102}$$

with

$$Z_{k_1,k_2} = \oint \prod_{i=1}^{n_1} d\phi_i \prod_{j=1}^{k_2} d\psi_j \, z_{\text{vec},k_1}^{Sp(1)}(\phi) \, z_{\text{vec},k_2}^{SO(4)}(\psi) \, z_{\text{bifund},k_1,k_2}^{Sp(1)-SO(4)}(\phi, \psi). \tag{5.103}$$

Here q_1 and q_2 correspond to the exponentiated gauge couplings of the $Sp(1)$ and $SO(4)$ gauge group, respectively, and $k_1 = 2n_1 + \chi_1$. As mentioned above, the main problem is to find the correct prescription for the contour integral, and to evaluate the residues of the poles in question. In the case of ordinary $SU(N)$ quiver gauge theories, the poles of the integrand only come from the vector multiplet contribution, and can be labeled by colored Young diagrams. In the case at hand, however, the $Sp(1) - SO(4)$ bifundamental does introduce additional poles, so that evaluating the contour integral becomes much more complicated. More precisely, besides the poles coming from the $Sp(1)$ and the $SO(4)$ vector multiplet, there are also poles

$$\psi_j = \pm\epsilon_+ \text{ (when } k_2 \text{ is odd)} \tag{5.104}$$

$$\phi_i = \pm\psi_j \pm \epsilon_+ \tag{5.105}$$

from the $Sp(1) - SO(4)$ bifundamental. Note that these poles intertwine $Sp(1)$ poles and $SO(4)$ poles.

A priori the integrals in (5.103) are over the real axis. We need to make a choice in moving the poles away from the real axis and the closing the contour. The usual prescription is to move $\epsilon_{1,2} \mapsto \epsilon_{1,2} + i0$ and then close the contour in the upper-half plane. We use this convention to deal with the vector multiplet poles. For the bifundamental poles, however, we need to choose the opposite prescription $\epsilon_+ \mapsto \epsilon_+ - i0$. This recipe originates from the description of the poles for the massive full bifundamental hypermultiplet. Similarly to the pole prescription in the $\mathcal{N} = 4$ ADHM construction, we introduce two additional equivariant parameters $\epsilon_3 = -\mu - \epsilon_+$ and $\epsilon_4 = \mu - \epsilon_+$, which we assume to have positive imaginary parts. [7] To find the pole prescription for the bifundamental half-hypermultiplet, we just set the mass μ to zero (which identifies $\epsilon_3 = \epsilon_4$). We furthermore encounter poles of the form $n\epsilon_1 - m\epsilon_3$ with $n \in \frac{1}{2}\mathbb{N}$, $m \in \mathbb{N}$, which we also need to include. Our prescription is to take $\mathrm{Im}(\epsilon_\alpha - \epsilon_\beta) \gg 0$ if $\alpha > \beta$ as in the reference [64].

With this recipe we are set to evaluate the integral as the sum of pole residues. For each integration variable ϕ_i or ψ_j we have a precise prescription, so that we can proceed integral by integral. In practice it is useful to replace $\epsilon_+ \mapsto -\epsilon_3$ in the equations (5.104) and (5.105) to avoid any source of confusion. After identifying the additional poles coming

[7] See, for example, [75] for a detailed discussion of the $\mathcal{N} = 4$ ADHM construction.

from the bifundamental, we substitute back $\epsilon_3 \mapsto -\epsilon_+$ to evaluate the integral. Note that the unrefined partition function can only be obtained by setting $\epsilon_1 + \epsilon_2 = 0$ after we have performed the integration.

For the quiver gauge theory with a single $Sp(1) - SO(4)$ bifundamental half-hyper (see figure 5.16) the additional bifundamental poles start to contribute at instanton number $k = (k_1, k_2) = (2, 1)$. There are 12 new poles at this order. To get agreement with the conformal block (5.96) it is essential to include these extra poles. Interestingly, in the unrefined limit their contribution happens to vanish, so that instanton counting becomes much simpler.

Comparison with the three-point function (5.96)

We identify the parameters of the conformal field theory and gauge theory to be

$$\begin{aligned} \Lambda_1 &= -\frac{q_1}{\epsilon_1\epsilon_2},\ \Lambda_2 = -\frac{q_2}{16\epsilon_1\epsilon_2},\ \Lambda_3 = \frac{q_2}{16\epsilon_1\epsilon_2} \\ h_1 &= \frac{1}{\epsilon_1\epsilon_2}\left(\frac{Q^2}{4} - a^2\right), \\ h_2 &= \frac{1}{\epsilon_1\epsilon_2}\left(\frac{Q^2}{4} - \left(\frac{b_1+b_2}{2}\right)^2\right), \\ h_3 &= \frac{1}{\epsilon_1\epsilon_2}\left(\frac{Q^2}{4} - \left(\frac{b_1-b_2}{2}\right)^2\right), \\ c &= 1 + \frac{6(\epsilon_1+\epsilon_2)^2}{\epsilon_1\epsilon_2}, \end{aligned} \tag{5.106}$$

where c is the central charge and h_i are the conformal weights of the vertex operators. Comparing this with the three-point function (5.96) we find

$$\boxed{Z_{\text{inst}}(q_1, q_2) = Z_{\text{CFT}}(q_1, q_2) Z^{(1)}_{\text{spur}}\,,} \tag{5.107}$$

up to $k = (k_1, k_2) = (2, 2)$, with the spurious factor

$$Z^{(1)}_{\text{spur}} = \exp\left(\frac{q_1}{2\epsilon_1\epsilon_2}\right)\exp\left(\frac{q_2}{8\epsilon_1\epsilon_2}\right).$$

5.5.3.2 Genus two quiver through $Sp(1) - SO(4)$ instanton counting

The Sicilian quiver theory for genus two Gaiotto curve is an $SU(2)^3$ theory with two trifundamental half-hypers. We can also view it as a $Sp(1) - SO(4)$ theory with two bifun-

damental half-hypers (see figure 5.17). As illustrated in figure 5.19, the corresponding Gaiotto curve is a torus with two punctures. The genus two curve is the double cover of the genus one curve with two branch points.

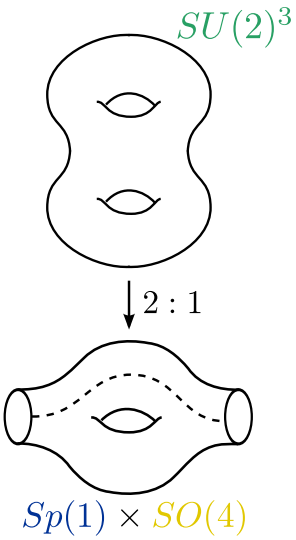


Figure 5.19: The UV-curve for the $SU(2)$ genus 2 quiver is a double covering of the Gaiotto curve for the $Sp(1) - SO(4)$ cyclic quiver. (The quiver diagrams are illustrated in figure 5.17.)

The partition functions will depend on the choice of coordinates of the Riemann surface, because the gauge theory is conformal. Using our proposal for the $Sp(1) - SO(4)$ AGT correspondence, we can find a choice of coordinates on the complex moduli space of the genus two Gaiotto curve that matches the instanton partition function with the conformal block directly, without a UV-UV map.

Computing the instanton partition function

Since a single full hypermultiplet can be obtained by combining two half-hypermultiplets, there are two different ways to compute the instanton partition function: either as a single full bifundamental hyper or as two bifundamental half-hypers. Using the first method we start from the massive full hyper. The term at order (k_1, k_2) is given by

$$Z^{(1)}_{k_1,k_2} = \oint \prod_{i=1}^{n_1} d\phi_i \prod_{j=1}^{k_2} d\psi_j \, z^{Sp(1)}_{\text{vec},k_1}(\phi, a) \, z^{SO(4)}_{\text{vec},k_2}(\psi, b) \, z^{Sp(1),SO(4)}_{\text{2bif},k_1,k_2}(\phi, \psi, \mu). \tag{5.108}$$

where the explicit form of the integrand is given in appendix E.

We can evaluate the contour integral using the two equivariant parameters, $\epsilon_3 = -\mu - \epsilon_+$ and $\epsilon_4 = \mu - \epsilon_+$. The massive full bifundamental introduces the additional poles

$$\begin{aligned} \psi_j &= \pm\epsilon_3, \pm\epsilon_4 \ \text{(when } k_2 \text{ is odd)} && (5.109)\\ \phi_i &= \pm\psi_j \pm \epsilon_3 && (5.110)\\ \phi_i &= \pm\psi_j \pm \epsilon_4. && (5.111) \end{aligned}$$

Compared to the single bifundamental half-hyper in the previous example the additional parameter ϵ_4 introduces extra poles. For example, there are in total 28 new poles with non-vanishing residues at $k = (1,2)$. Again the contribution from these new poles happens to vanish in the unrefined limit, whereas it is crucial to include them in the refined setup.

Using the second method we start with the contour integral corresponding to two massless bifundamental half-hypers

$$Z^{(2)}_{k_1,k_2} = \oint \prod_{i=1}^{n_1} d\phi_i \prod_{j=1}^{k_2} d\psi_j \, z^{Sp(1)}_{\text{vec},k_1}(\phi, a)\, z^{SO(4)}_{\text{vec},k_2}(\psi, b) \left(z^{Sp(1),SO(4)}_{\text{bif},k_1,k_2}(\phi,\psi) \right)^2, \tag{5.112}$$

where the explicit form of each of the building blocks is given in appendix E. The poles of this integral are simpler to enumerate, since there is just one new equivariant parameter $\epsilon_3 = -\epsilon_+$. There are 10 new poles with non-vanishing residues at $k = (1,2)$. As they should, both computations indeed give the same result once we set the mass $\mu = 0$.

CFT computation

Let us now compute the conformal block for the genus two surface. The most straightforward guess is to imitate (5.93) by taking

$$Z = \sum_{I_1,I_2,I_3} \sum_{J_1,J_2,J_3} K^{-1}_{I_1J_1} K^{-1}_{I_2J_2} K^{-1}_{I_3J_3} \langle \phi^{a_1}_{I_1} | V(\phi^{a_2}_{I_2}, 1) | \phi^{a_3}_{I_3} \rangle \langle \phi^{a_1}_{J_1} | V(\phi^{a_2}_{J_2}, 1) | \phi^{a_3}_{J_3} \rangle q_1^{I_1} q_2^{I_2} q_3^{I_3}. \tag{5.113}$$

By the general remarks above, this expression corresponds to a particular parametrization of the moduli space of the genus two surface. Presumably there should be a corresponding regularization scheme on the gauge theory side. In particular, the conformal block (5.113) should agree with any instanton computation in the IR.

We want to do a bit better than that however: we want to find an expression which

agrees with our instanton computation of the genus two surface as a cyclic $Sp(1) - SO(4)$ quiver in the UV. The AGT correspondence for Sp/SO quivers was worked out in the previous section. Let us briefly summarize the relevant facts.

The Gaiotto curve of the cyclic Sp/SO theory is a torus with a $\mathbb{Z}_2$ branch cut running between two branch points. The double cover of this curve is the genus two curve where two of the moduli are equal, see figure 5.19. The $\mathcal{W}$-algebra of the theory is a double copy of the Virasoro algebra, where the $\mathbb{Z}_2$-twist exchanges the two copies. The conformal block of this configuration on the torus with total modulus $q_1^2 q_2$ is given by

$$\mathrm{Tr}\left[\sigma(1) P_{a_1} \sigma(q_1^2) P_{a_2,a_3} (q_1^2 q_2)^{L_0}\right] . \tag{5.114}$$

Here σ is the $\mathbb{Z}_2$-twist vacuum, and we take the branch cut to go from $\sigma(1)$ through P_{a_1} to $\sigma(q_1^2)$. P_{a_1} is the projector onto the twisted representation coming from the primary field ϕ^1. As the primary field ϕ^1 transforms in a twisted representation, it is indeed characterized by a single parameter a_1. On the other hand, P_{a_2,a_3} is the projector onto the untwisted representation characterized by two parameters a_2 and a_3.

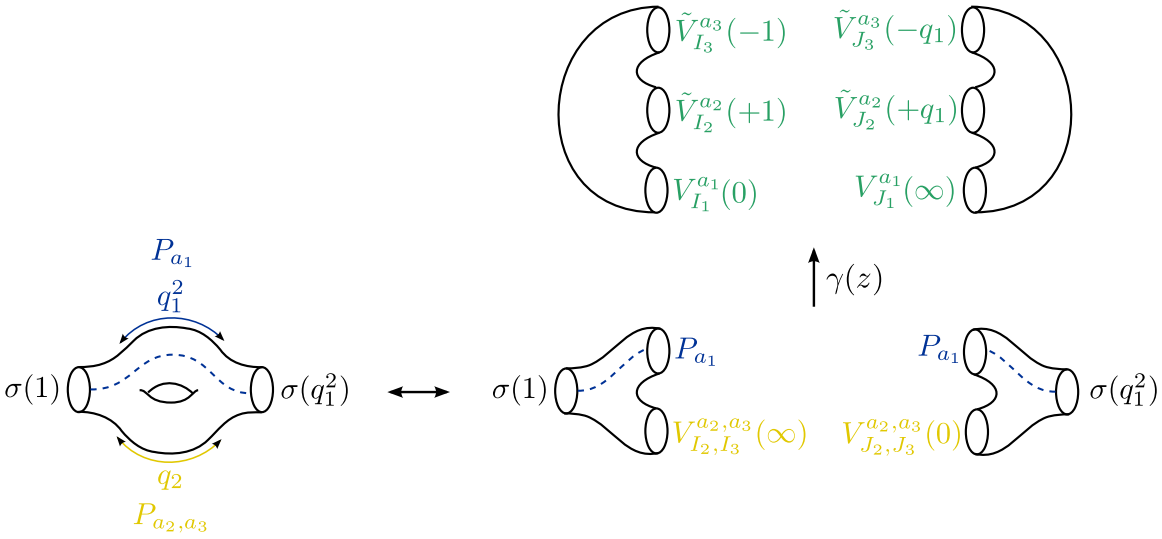


Figure 5.20: The map $\gamma(z)$ relates the $\mathcal{W}$-block on the twice punctured torus (which computes the double $Sp(1) - SO(4)$ instanton partition function) to a Virasoro block on its double cover, a genus two curve. More precisely, we first cut open the torus along the $SO(4)$ tube and insert a complete basis of states $V^{a_2,a_3}_{I_2,I_3}$ in the untwisted representation labeled by a_2 and a_3. Then we map this onto a genus two surface using the map γ, and insert a complete basis of states $V^{a_1}_{I_1}$ in the Virasoro representation labeled by a_1.

$$\langle V^{a_2,a_3}_{I_2,I_3}(\infty)\sigma(1)P_{a_1}\sigma(q_1^2)V^{a_2,a_3}_{J_2,J_3}(0)\rangle$$

$$\downarrow \quad z\mapsto\gamma(z)$$

$$\langle \tilde{V}^{a_2}_{I_2}(1)\tilde{V}^{a_3}_{I_3}(-1)|V^{a_1}_{I_1}\rangle\langle V^{a_1}_{J_1}|\tilde{V}^{a_2}_{J_2}(q_1)\tilde{V}^{a_3}_{J_3}(-q_1)\rangle$$

$$\swarrow \qquad\qquad\qquad \searrow z\mapsto q_1/z$$

$$\langle \tilde{V}^{a_2}_{I_2}(1)\tilde{V}^{a_3}_{I_3}(-1)|V^{a_1}_{I_1}\rangle \qquad\qquad \langle \tilde{V}^{a_2}_{J_2}(1)\tilde{V}^{a_3}_{J_3}(-1)|V^{a_1}_{J_1}\rangle$$

Table 5.2: Sequence of maps used for computing the cyclic $Sp(1)-SO(4)$ quiver

.

Computing correlators of twisted representations can be done by going to the cover of the surface. Here, the $\mathcal{W}$-algebra on the cover is a single copy of the Virasoro algebra, and the problem reduces to the computation of the standard conformal block on the genus two surface. Although conceptually straightforward, this procedure leads to some technical subtleties. The main problem is that we apply the cover map to compute correlators with Virasoro descendants, which leads to correction terms. Let us therefore spell out precisely the cover map and the ensuing correlation functions.

To map the four point function (5.114) to the cover, we use the map γ

$$z\mapsto\gamma(z)=\pm\sqrt{\frac{z-q_1^2}{z-1}}\,, \tag{5.115}$$

where the sign determines the branch of the cover. This map indeed has branch points at $z=1$ and $z=q_1^2$, and it maps the operators at 0 and ∞ to $\pm q_1$ and ± 1. This is illustrated in figure 5.20. In a second step, we want to reduce everything to standard building blocks, that is, three-point functions on the sphere with operators inserted at $1,-1$, and 0.

The total sequence of maps is shown in table 5.2. [8] Introducing the notation

$$\mathbf{C}^{a_2,a_3;a_1}_{I_2,I_3,I_1}(q_1)=\langle \tilde{V}^{a_2}_{J_2}(1)\tilde{V}^{a_3}_{J_3}(-1)|V^{a_1}_{J_1}\rangle q_1^{2(h_{a_2}+h_{a_3}+I_2+I_3)},$$

[8] The $\tilde{V}(\pm q_1)$s in the second line of table 5.2 are slightly different from the $\tilde{V}(\pm 1)$s, which are defined in equation (5.118).

the conformal block is given by

$$Z_{\rm CFT} = \sum_{I_i,J_i} K^{-1}_{I_1J_1} K^{-1}_{I_2J_2} K^{-1}_{I_3J_3} \mathbf{C}^{a_2,a_3;a_1}_{I_2,I_3,I_1}(q_1) \mathbf{C}^{a_2,a_3;a_1}_{J_2,J_3,J_1}(q_1) q_1^{I_1} q_2^{I_2+I_3}. \tag{5.116}$$

Note that the three-point functions $\mathbf{C}(q_1)$ depends on q_1 in a non-trivial way. They are obtained by acting with q_1-dependent coordinate transformations, so that the vertex operators $\tilde{V}$ depend on q_1. Since (5.115) is not a Möbius transformation, we need the following generalization [126] of (5.91):

$$V(\phi, z) \mapsto V\left(\left[\prod_{n=1}^{\infty} \exp\left(\frac{T_n(z)}{f'(z)^n} L_n\right)\right] f'(z)^{L_0}\phi, f(z)\right), \tag{5.117}$$

which holds for z away from singular points. Here we take all products to go from left to right. The functions $T_n(z)$ are defined recursively. The first two are given by

$$T_1(z) = \frac{f''(z)}{2f'(z)}, \qquad T_2(z) = \frac{1}{3!}\left(\frac{f'''(z)}{f'(z)} - \frac{3}{2}\left(\frac{f''(z)}{f'(z)}\right)^2\right).$$

For the transformations in (5.116) the new vertex operators are given by

$$\tilde{V}^a_I(\pm 1) = \left(\pm\frac{1-q_1^2}{2q_1^2}\right)^{h_a+I} V\left(\exp\left[\pm\frac{3+q_1^2}{2(1-q_1^2)} L_1\right] \exp\left[\frac{1}{4}L_2\right] \cdots \phi^a_I, \pm 1\right), \tag{5.118}$$

where the dots signify exponential factors involving higher L_n.

When we compare the genus two conformal block (5.116) with the instanton counting result, we find that they indeed agree up to order $k = (1, 2)$:

$$\boxed{Z_{\rm inst}(q_1, q_2) = Z_{\rm CFT}(q_1, q_2) Z^{(2)}_{\rm spur},} \tag{5.119}$$

where the spurious factor

$$Z^{(2)}_{\rm spur} = 1 - \frac{3\left(\epsilon_1+\epsilon_2\right)^2}{8\left(\epsilon_1\epsilon_2\right)} q_2 + \frac{3\left(\epsilon_1+\epsilon_2\right)^2\left(3\epsilon_1^2 - 2\epsilon_1\epsilon_2 + 3\epsilon_2^2\right)}{128\epsilon_1^2\epsilon_2^2} q_2^2 + \ldots$$

does not depend on physical parameters.

5.5.3.3 Alternative prescriptions for the genus two quiver

In the previous example we chose the coordinates of the genus two Gaiotto curve in such a way that we obtained direct agreement between the conformal block and the $Sp(1) - SO(4)$ instanton counting. If we choose different coordinates, or if we use a different instanton counting scheme, then the result will be different. In the infrared, however, all versions should agree. Put differently, we should be able to find a map between the UV couplings that make two results agree. Let us show that this philosophy is correct in two examples.

Comparison using a different conformal block

First, we can use the "naive" conformal block (5.113) to compute the genus two correlator. [9] We indeed find agreement between the $Sp(1) - SO(4)$ instanton partition function and this genus two conformal block using the UV-UV map

$$\begin{aligned}
\tilde{q}_1 &= q_1 - \frac{1}{4}q_1q_2 - \frac{1}{2}q_1^2 - \frac{1}{16}q_1q_2^2 + \frac{1}{4}q_1^2q_2 + \frac{3}{16}q_1^3 + O(q^4), \\
\tilde{q}_2 &= \frac{1}{16}q_2 - \frac{1}{16}q_1q_2 + \frac{3}{128}q_2^2 - \frac{1}{128}q_1q_2^2 + \frac{3}{128}q_1^2q_2 + \frac{55}{4096}q_2^3 + O(q^4), \\
\tilde{q}_3 &= \frac{1}{16}q_2 + \frac{1}{16}q_1q_2 + \frac{3}{128}q_2^2 + \frac{1}{128}q_1q_2^2 + \frac{3}{128}q_1^2q_2 + \frac{55}{4096}q_2^3 + O(q^4).
\end{aligned} \tag{5.120}$$

To compare the instanton partition functions with different parametrization of couplings, it is important to include tree-level pieces.[10] Similarly we also need to include the tree-level piece of the conformal block. In the case at hand we have

$$Z_{\text{tree}}^{Sp-SO}(q) = q_1^{-\frac{a^2}{\epsilon_1\epsilon_2}}\left(\frac{q_2}{16}\right)^{-\frac{1}{2\epsilon_1\epsilon_2}(b_1^2+b_2^2)} \quad \text{and} \quad Z_{\text{tree}}^{\text{CFT}}(\tilde{q}) = \tilde{q}_1^{-\frac{a_1^2}{\epsilon_1\epsilon_2}}\,\tilde{q}_2^{-\frac{a_2^2}{\epsilon_1\epsilon_2}}\,\tilde{q}_3^{-\frac{a_3^2}{\epsilon_1\epsilon_2}}\,.$$

Note that these two factors are related by the identifications (5.106) and the mappings $a_2 \to \frac{1}{2}(b_1+b_2)$ and $a_3 \to \frac{1}{2}(b_1-b_2)$. Using the UV-UV mapping (5.120) and the above identification of Coulomb parameters, we find that the ratio of Z_{Nek} and Z_{CFT} is given by the spurious factor

$$Z_{\text{spur}}^{(3)} = 1 - \frac{(\epsilon_1+\epsilon_2)^2}{8\epsilon_1\epsilon_2}q_1 - \frac{(\epsilon_1+\epsilon_2)^2}{4\epsilon_1\epsilon_2}q_2 + \frac{(\epsilon_1+\epsilon_2)^2(\epsilon_1^2+3\epsilon_1\epsilon_2+\epsilon_2^2)}{32\epsilon_1^2\epsilon_2^2}q_1q_2 + \cdots\,.$$

[9] We additionally checked that genus two correlator is independent on the choice of internal punctures up to a UV-UV mapping.

[10] One-loop factors are not relevant since they do not involve the gauge couplings.

Comparison using a different method of instanton counting

As another test of this philosophy on the gauge theory side, we can compute the instanton using the $U(2)$ instanton counting scheme. This is possible since two $SU(2)$ trifundamental half-hypers combine into a single massless $SU(2)$ full trifundamental hyper. Apart from a non-trivial UV-UV mapping, we expect the $U(2)$ instanton and the CFT computation to differ by a non-trivial $U(1)$ factor. This $U(1)$ factor should not depend on any of the Coulomb parameters, after we enforce the tracelessness condition on all three gauge groups.

Since the trifundamental hyper is in the fundamental representation of the three gauge groups $SU(2)_A \otimes SU(2)_B \otimes SU(2)_C$, we find the contour integrand by considering the tensor product

$$\mathcal{E}_A \otimes \mathcal{E}_B \otimes \mathcal{E}_C \otimes \mathcal{L}$$

of the three $U(2)$ universal bundles $\mathcal{E}_{A/B/C}$ over the product of ADHM moduli spaces $\mathcal{M}_{U(2)_A} \times \mathcal{M}_{U(2)_B} \times \mathcal{M}_{U(2)_C} \times \mathbb{R}^4$ with the half-canonical bundle $\mathcal{L}$ over $\mathbb{R}^4$. Recall that the restriction of each universal bundle to a self-dual connection $\mathcal{A}$ is just the corresponding instanton bundle $\mathcal{E}|_{\mathcal{A}} = E$ over $\mathbb{R}^4$ The equivariant weights contributing to the Euler class $e_{\mathbf{T}}(\mathcal{V})$ of the bundle $\mathcal{V}$ of Dirac zero modes can then be found from the equivariant Dirac index

$$\mathrm{Ind}_{\mathbf{T}}(\not{D})^{g=2}_{U(2)^3} = \int_{\mathbb{C}^2} \mathrm{Ch}_{\mathbf{T}}(\mathcal{E}_A \otimes \mathcal{E}_B \otimes \mathcal{E}_C \otimes \mathcal{L})\, \mathrm{Td}_{\mathbf{T}}(\mathbb{C}^2). \tag{5.121}$$

The resulting contour integrand can be found in appendix E. For up to two instantons, it reproduces the partition functions for the bifundamental. For three non-zero instanton numbers the evaluation of the contour integral becomes tricky, because, unlike for $U(2)$ bifundamentals, many additional poles appear. It would be interesting to find an elegant prescription for the additional poles that yields agreement with the CFT.

We can still compare the instanton partition function up to second order. We find that

up to this order it agrees with the conformal block (5.113), when we use the map

$$\begin{aligned}
\tilde{q}_A &= q_A + 2q_A^2 q_B + 6q_A q_B q_C + \mathcal{O}(q^4) \\
\tilde{q}_B &= q_B + 2q_A q_B + 2q_B q_C + 8q_A q_B q_C + 3q_A^2 q_B - 2q_A q_B^2 + 2q_B^2 q_C + 3q_B q_C^2 + \mathcal{O}(q^4) \\
\tilde{q}_C &= q_C + 2q_B q_C + 6q_A q_B q_C + 3q_B^2 q_C + 2q_B q_C^2 + \mathcal{O}(q^4)
\end{aligned}$$

between gauge coupling constants $\tilde{q}_{A/B/C}$ and complex structure parameters $q_{A/B/C}$, and up to a (unrefined) $U(1)$ factor

$$Z_{\mathrm{U}(1)} = 1 + q_A q_B + q_B q_C + q_C q_A + \cdots \tag{5.122}$$

Again, we need to include the classical contributions here. The non-trivial UV-UV mapping is expected since the conformal three-point function only reduces directly to a bifundamental contribution when two of its punctures are set at the positions 0 and ∞, and the primary vertex operator is inserted at 1.

Chapter 6

Conclusion

In this thesis, we studied the relation between instanton moduli space and $\mathcal{W}$-algebras. We studied the theories of class $\mathcal{S}$, which arises from M5-branes wrapped on a UV-curve and showed that the partition functions of such theories are reproduced by the appropriate chiral blocks of the $\mathcal{W}$-algebra. We have done the direct microscopic calculation of the instanton partition functions using the equivariant localization techniques for many different examples. In particular, we carefully studied the effect of renormalization scheme at non-perturbative level and showed that it admits a nice geometric interpretation. We have also confirmed that 4d/2d correspondence holds for all the choice of $\Gamma = A, D, E$.

We haven't yet completed the story for all the possible theories that comes from $\Gamma = A, D, E$. On the 4d side, the gauge theory coming from E_n type $\mathcal{N} = (2, 0)$ theory has not been studied. Also, the instanton moduli spaces of the exceptional gauge groups are not known in general.[1] In the 2d side, the 3-point function of Toda theory is not fully known.

There are many other approaches to 4d/2d correspondence, for example, using the matrix model [133, 134], and geometric engineering using the refined topological strings [135]. There is also a correspondence between 5d version of the instanton partition function to the q-deformed version of the Virasoro algebra [136, 137]. There are proofs of certain special cases by showing that the same recursion relation holds on the both side of the correspondence [138, 139].

One of the biggest things we have omitted in this thesis is the study of the non-local BPS operators such as line operator [140, 141, 142] and surface operator [140, 143, 118, 144], and domain wall [145, 146]. Such non-local operators arise when we consider the

[1] Very recently, there has been some progress in this direction recently using string duality [130]. It has been checked by a direct computation of the partition function [131] using the recursive structure of the instanton partition function [55, 132].

	M_4	C	operator
M2	2	0	minimal surface operator
	1	1	line operator
	0	2	local operator, change 2d theory
M5	4	0	change 4d theory
	3	1	domain wall
	2	2	surface operator, change 2d theory

Table 6.1: Various types of defect on M5-branes. When M2 ends on M5, they share 2 dimensions so that M5 brane theory has 2-dimensional defect. When M5 and M5 intersects, we have 4-dimensional defect.

configurations of M2-branes ending on M5-branes, or intersecting M5-branes.

Especially, there has been a lot of work regarding the surface operator. The quantum Drinfeld-Sokolov reduction provides many construction of various types of $\mathcal{W}$-algebras, which does not belong to $\Gamma = A, D, E$ type. If we start with an affine Lie algebra of type $\hat{\mathfrak{g}}$, we not only can get the $\mathcal{W}(\mathfrak{g})$, but also other types. It is determined by embedding of $sl(2)$ to $\mathfrak{g}$, which has one-to-one correspondence with the Levi subgroup of Γ. It is exactly the subgroup which parametrizes the type of surface operator in a gauge theory with gauge group Γ and therefore there must be a corresponding 2d CFT interpretation. It has been studied in the papers [147, 148, 149, 150]. It would be interesting to extend these analyses of non-local BPS operators to non-unitary gauge groups.

As we have discussed in the introduction, one way to under the 4d/2d correspondence is through the M5-branes wrapped on $M_4 \times C$. We have been exclusively focused on the case where M_4 is either $\mathbb{R}^4$ or S^4. But what happens for other choices? Since we have

$$Z^{4d}_{T[\Gamma,C]}(M_4) = Z^{2d}_{\mathcal{T}[\Gamma,M_4]}(C), \tag{6.1}$$

we should find some appropriate relations for other choices as well. Not many examples are known, but there has been some study for the following cases: When $M_4 = S^1 \times S^3$, the 4d partition function computes the superconformal index and the corresponding 2d side is the topological field theory [151, 152, 153], where in certain limit reduces to the q-deformed Yang-Mills theory [154]. When $M_4 = \mathbb{R}^4/\mathbb{Z}_p$ or $S^4/\mathbb{Z}_p$ and $\Gamma = A_1$, it is conjectured to be dual to para-Liouville theory [155]. The special case $p = 2$ is (accidentally) the super-Liouville theory [156, 157].

Mathematically, the AGT conjecture can be formalized as follows: Let G be the compact Lie group, and $\mathcal{M}_k^G$ be some appropriate compactification of the G-instanton moduli space with second Chern number (instanton number) k. Let $\mathfrak{g}$ be the associated simply-laced Lie algebra of G and $\mathcal{W}(\mathfrak{g})$ be the corresponding $\mathcal{W}$-algebra. Let $\mathbf{T} = \mathbf{T}^2_{\epsilon_1,\epsilon_2} \times \mathbf{T}^r_{\vec{a}}$ where r is the rank of G. Then (pure YM version of) the AGT conjecture says that there is a natural $\mathcal{W}(\mathfrak{g})$ action on the direct sum of the (intersection) equivariant cohomology of $\mathcal{M}_k^G$:

$$\mathcal{W}(\mathfrak{g}) \circlearrowright \bigoplus_{k=1}^{\infty} H^*_{\mathbf{T}}(\mathcal{M}_k^G). \tag{6.2}$$

It is known to us that the above statement has been proved by Maulik and Okounkov [123][2]. Even though we have a mathematical proof that this version of AGT conjecture is true, we are still left with many questions both in physics and mathematics. Especially, it is unclear how to form the correspondence for general non-linear quiver gauge theories. In general, the quiver gauge theories of class $\mathcal{S}$ do not admit Lagrangian descriptions. In these cases, it is unclear how to define the corresponding chiral block of $\mathcal{W}$-algebras for non-Lagrangian SCFT. It would be interesting to formulate the conjecture appropriately which will shed light on both 4d superconformal field theories theories and 2d conformal field theory with $\mathcal{W}$-algebra symmetry.

More recently, there has been growing work on 3d/3d correspondence. We can replace M_4 by M_3 by reducing the dimension by one and also replace C by M_3' by increasing the dimension by one. When M_3 is given by (Ω-deformed) S^3, the correspondence has been studied by [146, 159, 160, 161, 162, 163]. The superconformal index on the theory given by compactifying M5-branes on a 3-manifold is studied in [164].

Perhaps one of the most interesting directions for the 4d/2d relation is to use it to understand the dynamics of the M5-brane theory. Even though the 4d/2d relation itself is independent of the details of the M5-branes, but it does show interesting aspects of M5-branes. What we have been studying can be rephrased as the study of partially topologically twisted version of the multiple M5-brane theory. In general, the topological version of a quantum field theory encodes the vacuum sector of the untwisted theory. It is not unreasonable to expect that the detailed study of the theory reveal some aspects of the mysterious M5-brane theory.

[2]There is also a recent proof [158] for the A-series

Appendix A

Roots of simple Lie algebras

Here we list the roots for all Lie algebras, emphasizing how to embed the root space of a non-simply-laced algebra G to that of a simply-laced algebra Γ. First let us present the simply-laced ones in detail. Note that the simple roots are named as in Figure A.1.

A.1 Simply-laced algebras

Roots of $A_n = \mathrm{SU}(n+1)$. We let $\{\vec{e}_i : 1 \leqslant i \leqslant n+1\}$ be an orthonormal basis. The positive roots are

$$\Delta^+ = \{\vec{e}_i - \vec{e}_j : 1 \leqslant i < j \leqslant n+1\} . \tag{A.1}$$

Note that the span of roots is only n-dimensional. The simple roots are

$$\vec{\alpha}_1 = \vec{e}_1 - \vec{e}_2, \quad \vec{\alpha}_2 = \vec{e}_2 - \vec{e}_3, \quad \ldots, \quad \vec{\alpha}_n = \vec{e}_n - \vec{e}_{n+1} . \tag{A.2}$$

Roots of $D_n = \mathrm{SO}(2n)$. The positive roots are

$$\Delta^+ = \{\vec{e}_i \pm \vec{e}_j : 1 \leqslant i < j \leqslant n\} . \tag{A.3}$$

The simple roots are

$$\vec{\alpha}_1 = \vec{e}_1 - \vec{e}_2, \quad \vec{\alpha}_2 = \vec{e}_2 - \vec{e}_3, \quad \ldots, \quad \vec{\alpha}_{n-1} = \vec{e}_{n-1} - \vec{e}_n, \, \vec{\alpha}_n = \vec{e}_{n-1} + \vec{e}_n. \tag{A.4}$$

Roots of E_6. The 36 positive roots are

$$\Delta^+ = \{\vec{e}_i + \vec{e}_j, \vec{e}_i - \vec{e}_j\}_{i<j\leqslant 5} \cup \left\{\frac{1}{2}(\pm\vec{e}_1 \pm \vec{e}_2 \pm \vec{e}_3 \pm \vec{e}_4 \pm \vec{e}_5 + \sqrt{3}\vec{e}_6)\right\}_{\text{\# minus signs even}} \tag{A.5}$$

The simple roots are

$$\begin{aligned} \vec{\alpha}_1 &= \frac{1}{2}\left(\vec{e}_1 - \vec{e}_2 - \vec{e}_3 - \vec{e}_4 - \vec{e}_5 + \sqrt{3}e_6\right) , \\ \vec{\alpha}_i &= \vec{e}_i - \vec{e}_{i-1}\ (i = 2,3,4,5), \quad \vec{\alpha}_6 = \vec{e}_1 + \vec{e}_2 . \end{aligned} \tag{A.6}$$

Roots of E_7. The 63 positive roots are

$$\Delta^+ = \{\vec{e}_i + \vec{e}_j, \vec{e}_i - \vec{e}_j\}_{i<j\leqslant 6} \cup \{\sqrt{2}\vec{e}_7\} \cup \left\{\frac{1}{2}(\pm\vec{e}_1 \pm \ldots \pm \vec{e}_6 + \sqrt{2}\vec{e}_7)\right\}_{\text{\# minus signs odd}} , \tag{A.7}$$

Roots of E_8. The 120 positive roots are

$$\Delta^+ = \{\vec{e}_i + \vec{e}_j\}_{i<j\leqslant 8} \cup \{\vec{e}_i - \vec{e}_j\}_{i<j\leqslant 8} \cup \left\{\frac{1}{2}(\pm\vec{e}_1 \pm \ldots \pm \vec{e}_7 + \vec{e}_8)\right\}_{\text{\# minus signs even}} . \tag{A.8}$$

A.2 Non-simply-laced algebras

Let Γ be a simply-laced algebra which is not A_{2n}, and let o be a symmetry of the Dynkin diagram. Note that o can be viewed as an outer-automorphism acting on the Lie algebra of type Γ. It can be checked that α and $o(\alpha)$ are perpendicular when $\alpha \neq o(\alpha)$. Let us take

$$\Delta_s = \{\vec{\alpha} \qquad : \vec{\alpha} \in \Delta,\ \vec{\alpha} = o(\vec{\alpha})\} , \tag{A.9}$$

$$\Delta_l = \{\vec{\alpha} + o(\vec{\alpha}) \qquad : \vec{\alpha} \in \Delta,\ \vec{\alpha} \neq o(\vec{\alpha})\} \ \text{ when } r = 2 , \tag{A.10}$$

$$\Delta_l = \{\vec{\alpha} + o(\vec{\alpha}) + o^2(\vec{\alpha}) : \vec{\alpha} \in \Delta,\ \vec{\alpha} \neq o(\vec{\alpha})\} \ \text{ when } r = 3 . \tag{A.11}$$

Then, $\Delta' = \Delta_s \sqcup \Delta_l$ is a non-simply-laced root system, with Δ_s and Δ_l short and long roots, respectively. In this normalization, the short root has length $\sqrt{2}$. Possible outer-automorphisms of simply-laced algebras that can be used are depicted in Figure 2.6. Every non-simply-laced algebra arises in this manner.

Note that one can also take the averages $\frac{1}{2}[\vec{\alpha} + o(\vec{\alpha})]$ and $\frac{1}{3}[\vec{\alpha} + o(\vec{\alpha}) + o^2(\vec{\alpha})]$ in (A.10)

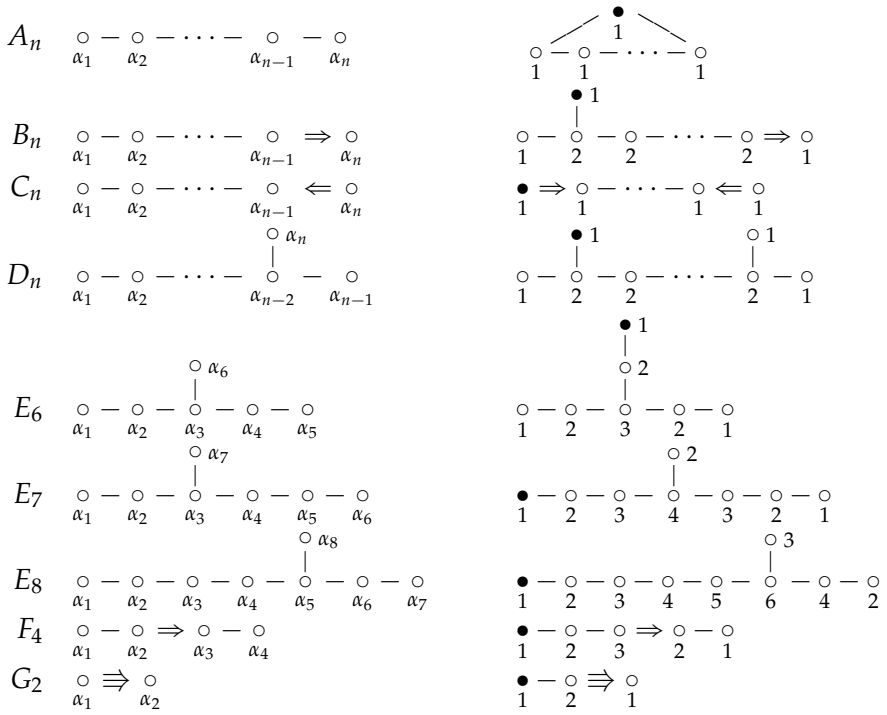


Figure A.1: Dynkin diagrams of simple Lie algebras, our labeling of the simple roots, and the comarks. The extended node is shown by a black blob.

and (A.11), respectively. A root system obtained this way corresponds to the subalgebra of Γ invariant under the $\mathbb{Z}_r$ action, and is Langlands dual to the one obtained via (A.9)–(A.11). It is natural in this latter convention for the long roots to have length $\sqrt{2}$. Even when the resulting root system is the same in two ways of folding, as in $G = G_2$ and $\Gamma = \mathrm{SO}(8)$, the embedding of the root space of G into the root space of Γ is different in the two cases. In this paper, we adhere to the former convention of folding. This convention is more natural in the context of singularity theory, see, e.g., [48, 165]. Let us now go over the root systems of the non-simply-laced algebras one by one.

Roots of $B_n = \mathrm{SO}(2n+1)$. The short positive roots and the long positive roots are

$$\Delta_s^+ = \{\vec{\epsilon}_i, : 1 \leqslant i \leqslant n\}, \qquad \Delta_l^+ = \{\vec{\epsilon}_i \pm \vec{\epsilon}_j : 1 \leqslant i < j \leqslant n\}\,. \tag{A.12}$$

where $\vec{\epsilon}_i \cdot \vec{\epsilon}_j = 2\delta_{ij}$ so that the lengths of the short roots are $\sqrt{2}$.

This root system comes from a $\mathbb{Z}_2$ outer-automorphism o of the root system of $A_{2n-1} = \mathrm{SU}(2n)$ with the action mapping the simple root $-\alpha_i$ to $-\alpha_{2n-i}$. The short and long positive

roots of B_n are then respectively

$$\Delta_s^+ = \{-\vec{e}_i + \vec{e}_{2n+1-i} : i = 1, \ldots, n\} \,, \tag{A.13}$$

$$\Delta_l^+ = \{(-\vec{e}_i + \vec{e}_{2n+1-i}) \pm (-\vec{e}_j + \vec{e}_{2n+1-j}) : 1 \leqslant i < j \leqslant n\}. \tag{A.14}$$

Setting

$$\vec{\epsilon}_i = -\vec{e}_i + \vec{e}_{2n+1-i}, \quad i = 1, \ldots, n \,, \tag{A.15}$$

the positive roots (A.13) and (A.14) become (A.12) as required.

Roots of $C_n = \mathrm{Sp}(n)$. The short positive roots and the long positive roots are

$$\Delta_s^+ = \{\vec{e}_i \pm \vec{e}_j : 1 \leqslant i < j \leqslant n\}, \qquad \Delta_l^+ = \{2\vec{e}_i : 1 \leqslant i \leqslant n\} \,, \tag{A.16}$$

where $\vec{e}_i \cdot \vec{e}_j = \delta_{ij}$ so that the lengths of the short roots are $\sqrt{2}$.

This root system is obtained by applying a $\mathbb{Z}_2$ outer-automorphism to the root system of $D_{n+1} = \mathrm{SO}(2n+2)$, which acts as follows:

$$o : -\alpha_{n-1} \mapsto -\alpha_n, \quad -\alpha_n \mapsto -\alpha_{n-1}, \quad -\alpha_i \mapsto -\alpha_i \text{ for } i \neq n, n-1 \,. \tag{A.17}$$

From (A.4), it is clear that o maps $\vec{e}_n$ to $-\vec{e}_n$ and leaves other e_i invariant. Following the procedure, we easily get (A.16).

Roots of F_4. The short positive roots are

$$\Delta_s^+ = \{\vec{\epsilon}_i\}_{i \leqslant 4} \cup \left\{ \frac{1}{2}(\vec{\epsilon}_1 \pm \vec{\epsilon}_2 \pm \vec{\epsilon}_3 \pm \vec{\epsilon}_4) \right\} \,, \tag{A.18}$$

and the long positive roots are

$$\Delta_l^+ = \{\vec{\epsilon}_i + \vec{\epsilon}_j\}_{i<j \leqslant 4} \cup \{\vec{\epsilon}_i - \vec{\epsilon}_j\}_{i<j \leqslant 4} \,, \tag{A.19}$$

where $\vec{\epsilon}_i \cdot \vec{\epsilon}_j = 2\delta_{ij}$ so that the length of the short roots is $\sqrt{2}$.

This root system can be obtained by applying a $\mathbb{Z}_2$ outer-automorphism to the root

system of E_6 which maps $(-\alpha_1, -\alpha_2, -\alpha_3, -\alpha_4, -\alpha_5, -\alpha_6)$ to $(-\alpha_5, -\alpha_4, -\alpha_3, -\alpha_2, -\alpha_1, -\alpha_6)$. Therefore the simple roots of F_4 are given by

$$\widehat{\vec{\alpha}}_1 = -\alpha_1 + -\alpha_5, \quad \widehat{\vec{\alpha}}_2 = -\alpha_2 + -\alpha_4, \quad \widehat{\vec{\alpha}}_3 = -\alpha_3, \quad \widehat{\vec{\alpha}}_4 = -\alpha_6. \tag{A.20}$$

We then introduce new basis $\epsilon_{1,2,3,4}$ via

$$\widehat{-\alpha_1} = \epsilon_2 - \epsilon_3\,, \quad \widehat{-\alpha_2} = \epsilon_3 - \epsilon_4\,, \quad \widehat{-\alpha_3} = \epsilon_4\,, \quad \widehat{-\alpha_4} = \frac{1}{2}(\epsilon_1 - \epsilon_2 - \epsilon_3 - \epsilon_4)\,. \tag{A.21}$$

These simple roots give rise to the positive roots listed in (A.18) and (A.19).

Roots of G_2. The short positive roots and long positive roots are

$$\Delta_s^+ = \left\{\sqrt{2}\vec{\epsilon}_1,\ \pm\frac{1}{\sqrt{2}}\vec{\epsilon}_1 + \sqrt{\frac{3}{2}}\vec{\epsilon}_2\right\}\,, \qquad \Delta_l^+ = \left\{\pm\frac{3}{\sqrt{2}}\vec{\epsilon}_1 + \sqrt{\frac{3}{2}}\vec{\epsilon}_2,\ \sqrt{6}\vec{\epsilon}_2\right\}, \tag{A.22}$$

where $\vec{\epsilon}_i \cdot \vec{\epsilon}_j = \delta_{ij}$ so that the length of the short roots is $\sqrt{2}$.

This root system is obtained by applying a $\mathbb{Z}_3$ outer-automorphism to the root system of D_4, which maps the set of simple roots $(-\alpha_1, -\alpha_2, -\alpha_3, -\alpha_4)$ to $(-\alpha_3, -\alpha_2, -\alpha_4, -\alpha_1)$. The simple roots of G_2 is then

$$\widehat{\alpha}_1 = -\alpha_1 + -\alpha_3 + -\alpha_4, \quad \widehat{\alpha}_2 = -\alpha_2. \tag{A.23}$$

We then let

$$\vec{\epsilon}_1 = \frac{1}{\sqrt{2}}(\vec{e}_1 - \vec{e}_3), \quad \vec{\epsilon}_2 = \frac{1}{\sqrt{6}}(\vec{e}_1 + 2\vec{e}_2 - \vec{e}_3), \tag{A.24}$$

resulting in (A.22).

Appendix B

Evaluating contour integrals

In this appendix we explain in more detail how to evaluate contour integrals for the $Sp(1)$ gauge group. (Evaluating $SO(4)$ contour integrals works similarly. It is somewhat simpler because there are no fractional instantons.) In the case of $U(N)$ gauge groups, [28, 55, 56] found closed expressions for the contribution of k instantons in terms of sums over Young diagrams.

Unfortunately the pole structure of $Sp(N)$ gauge groups is much more complicated. In the literature it has mostly only been evaluated up to three instantons, which only requires to perform one contour integral [68, 71].[1] It is possible to describe $Sp(N)$ instantons in terms of orientifolding the $U(N)$ setup. In general, there are poles involving Coulomb branch parameters, as well as poles that just involve the deformation parameters ϵ_1, ϵ_2. The former *regular* poles are similar to the $U(N)$ instantons, while the latter *fractional* are new. In the brane engineering picture, the former can be thought of as an instanton bound to D4-branes that are separated from the center at positions $\pm a_n$. The latter can be understood as an instanton stuck at the orientifold brane at the center. When we specialize to the case of $\epsilon_1 + \epsilon_2 = 0$, the analysis of the poles become simpler, and it reduces to the ordinary colored Young diagrams plus fractional instantons as in [70, 71]. However, this method does not work for general ϵ_1, ϵ_2.

For our analysis we will use the expressions obtained for the $Sp(1)$ gauge multiplet in [53, 68]. The k instanton contribution is given by the integral over the real axis of the variables ϕ_i of the integrand z_k. Let $n = \lfloor \frac{k}{2} \rfloor$, $\chi = k \mod 2$ such that $k = 2n + \chi$. It is useful

[1] [70] computed up to four instantons, but could only give an ad-hoc prescription for which poles to include.

to define $\epsilon = \epsilon_1 + \epsilon_2$ and $\epsilon_+ = \epsilon/2$. Define

$$\Delta(x) = \prod_{i<j\leqslant n} ((\phi_i + \phi_j)^2 - x^2)((\phi_i - \phi_j)^2 - x^2) \,, \tag{B.1}$$

$$P(x) = x^2 - a^2 \,. \tag{B.2}$$

Then z_k is given by

$$z_k(a, \phi, \epsilon_1, \epsilon_2) = \frac{(-1)^n}{2^{n+\chi} n!} \frac{\epsilon^n}{\epsilon_1^n \epsilon_2^n} \left[\frac{1}{\epsilon_1 \epsilon_2 (\epsilon_+^2 - a^2)} \prod_{i=1}^{n} \frac{\phi_i^2 (\phi_i^2 - \epsilon^2)}{(\phi_i^2 - \epsilon_1^2)(\phi_i^2 - \epsilon_2^2)} \right]^{\chi} \times \frac{\Delta(0)\Delta(\epsilon)}{\Delta(\epsilon_1)\Delta(\epsilon_2)} \prod_{i=1}^{n} \frac{1}{P(\phi_i - \epsilon_+) P(\phi_i + \epsilon_+)(4\phi_i^2 - \epsilon_1^2)(4\phi_i^2 - \epsilon_2^2)} \tag{B.3}$$

The contribution is obtained by integrating the ϕ_i along the real axis.

Hypermultiplets in the fundamental representation never contribute any new poles. As pointed out above, hypers in the adjoint representation of $Sp(N)$ do introduce new poles. This will not be covered here.

B.1 The ϵ prescription

To render this integral well-defined, we specify that $\epsilon_{1,2} \in \mathbb{R} + i0$. We can then close the integrals in the upper half plane and simply evaluate all residues. This prescription can be obtained by e.g.going to the five-dimensional theory, and requiring that the original integral converge.

At first sight one may be worried that we need additional information on the imaginary part of the ϵ if we want to evaluate the integral. More precisely, the following situation might arise: Let us take the residue of ϕ_1 around the pole $\phi_2 + a$ where a is some linear combination of $\epsilon_{1,2}$. If the original integrand had a pole $(\phi_3 + \phi_1 - b)^{-1}$, then the resulting expression seems to have a pole at $\phi_2 = -\phi_3 + b - a$, which would not longer have clearly

defined imaginary part. To see that this situation never occurs, note that

$$\begin{aligned}\oint_{-\phi_3+b-a} d\phi_2 \oint_{\phi_2+a} d\phi_1 \frac{1}{(\phi_1-\phi_2-a)(\phi_1+\phi_3-b)} F(\phi_1,\phi_2,\phi_3) \\ = F(-\phi_3+b,-\phi_3+b-a,\phi_3) \\ = -\oint_{-\phi_3+b-a} d\phi_2 \oint_{-\phi_3+b} d\phi_1 \frac{1}{(\phi_1-\phi_2-a)(\phi_1+\phi_3-b)} F(\phi_1,\phi_2,\phi_3)\,,\end{aligned} \tag{B.4}$$

i.e.when evaluating the poles in both ways (as we must) the contributions cancel. Note that this argument is also valid if ϕ_3 is a constant or zero. Also note that this does not imply that all residues vanish. The point is that if either a or b has negative imaginary part, then by our ϵ prescription we only evaluate one residue, which is therefore not cancelled. The upshot of this discussion is thus that whenever we evaluate the residues, we only need to include poles which have clearly defined positive imaginary part.

The k instanton contribution Z_k is thus a sum over positive poles $(\tilde{\phi}_i)_{i=1,\dots,n}$

$$Z_k = \sum_{(\tilde{\phi}_i)_{i=1,\dots,k}} \oint_{\phi_k=\tilde{\phi}_n} d\phi_n \dots \oint_{\phi_1=\tilde{\phi}_1} d\phi_1\, z_k(a,\phi,\epsilon_1,\epsilon_2)\,, \tag{B.5}$$

where $\tilde{\phi}_i$ is a linear combination of $b, \epsilon_1, \epsilon_2$ and possibly $\phi_j, j > i$ with positive imaginary part. Note that different poles can give the same contribution. In what follows we give an algorithm to obtain those poles and their combinatorial weight.

B.2 Chains

If k is even, then $\chi = 0$. The possible poles for ϕ_i are

$$\phi_i = \pm\epsilon_1/2\,, \qquad \phi_i = \pm\epsilon_2/2 \tag{B.6}$$

$$\phi_i = a \pm \epsilon_+ \qquad \phi_i = -a \pm \epsilon_+ \tag{B.7}$$

$$\phi_i = \phi_j \pm \epsilon_{1,2} \qquad \phi_i = -\phi_j \pm \epsilon_{1,2} \tag{B.8}$$

We will call poles as in (B.6) and (B.7) 'roots'. Due to poles of the form (B.8), the ϕ_i will take values in chains, just as in the $U(N)$ case. If a chain contains a root, we will call it an anchored chain. Note that unlike the $U(N)$ case there can also be chains that have no roots.

If k is odd, then there are the additional poles

$$\phi_i = \pm\epsilon_1 \,, \qquad \phi_i = \pm\epsilon_2 \tag{B.9}$$

Note that the numerator has a double zero for $\phi_i = \phi_j$ and $\phi_i = -\phi_j$.

For a more uniform treatment in the spirit of the $U(2)$ analysis, we define the set of roots b_l

$$b_l \in \{a + \epsilon_+, -a + \epsilon_+, \epsilon_1/2, \epsilon_2/2\} \tag{B.10}$$

for n even, and similarly for n odd. A general pole can consist of multiple chains that are independent of each other. It is thus possible to describe our algorithm by using the following toy model which only contains one root,

$$z_k = \frac{\Delta(0)\Delta(\epsilon)}{\Delta(\epsilon_1)\Delta(\epsilon_2)} \prod_{i=1}^{n} \frac{1}{\phi_i^2 - b^2} \,. \tag{B.11}$$

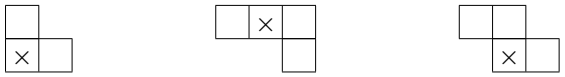
When going back to the full $Sp(1)$ integrand, one sums over all decompositions of k into chains with different roots. Note that not all roots can appear as anchors for a given pole. In particular, due to the numerator in (B.3) there cannot be two chains with roots $a + \epsilon_+$ and $-a + \epsilon_+$ in the same pole. Also note that one has to be somewhat careful when to exactly specialize the values b_l. To get the correct result, one has to take the residue of the expression with general b_l, and only afterward specialize to (B.10).

Anchored poles

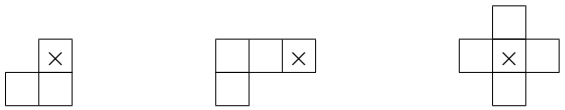
Anchored poles can be described in the following way: First, pick a 'generalized Young diagram' with n boxes. A generalized Young diagram is a set of connected boxes, one of them marked by $\times$, which we take to be the origin. The upper and lower edge of such a diagram must be monotonically decreasing, i.e.must slope from the upper left to the lower

right. To illustrate this, here are some examples:

allowed :



not allowed :



This is due to the zeros of $\Delta(0)$ and $\Delta(\epsilon)$ in the numerators, which can only be cancelled by double poles. To obtain the actual pole corresponding to a diagram, we consider signed diagrams, i.e.diagrams where each box comes with a sign. Let us denotes the signed diagrams corresponding to Y by $\tilde{Y}$, and let $\tilde{Y}_0$ be the diagram with all plus signs. The value of a variable ϕ_i for a box with sign $\pm$ is then

$$\phi_i = \pm(b + m\epsilon_1 + n\epsilon_2) \,, \tag{B.12}$$

where m and n are the horizontal and vertical positions of the box, respectively. The advantage of this description is that the even though a given signed diagram can arise from many different poles, the numerical values of the ϕ_i are determined by it. Moreover, because the integrand is invariant under $\phi_i \leftrightarrow -\phi_i$, the contribution of a signed diagram is the same of the unsigned diagram, up to an overall sign.

More precisely, each signed diagram contribues with a certain combinatorial weight, given by the (signed) number of the poles that contribute, so that the total contribution of Y is

$$I_Y = I_{\tilde{Y}_0} n_Y = I_{\tilde{Y}_0} \sum_{\tilde{Y}} n_{\tilde{Y}} \,. \tag{B.13}$$

It remains to compute the $n_{\tilde{Y}}$. To do this, write down all $n!$ numbered diagrams corresponding to $\tilde{Y}$, and check which ones give a contribution, i.e.are obtained during the evaluation of the contour integrals. The number i in each box indicates which ϕ_i takes this value. We then perform the contour integral consecutively, starting from ϕ_1. For a given ϕ_i, three things can happen: If it is at the origin, then it can take the value of the root b. This simply means that we the pole comes from the factor$\phi_i - b$. If it is not at the origin, we can connect it to one of its neighbors. If this neighbor has not been evaluated yet, then

the pole comes from the factor $(\phi_i \pm \phi_j - \epsilon_{1,2})$, $i < j$. If $j < i$, then the pole comes from the same factor, but we have already plugged in the value $\tilde{\phi}_j$ for ϕ_j. Finally, if it has already been connected to other boxes, we can also connect it to neighbors of those boxes.

All this is obviously subject to the constraint that the relative signs of the two boxes are correct, and that the imaginary part of the pole be positive. We can thus deduce some rules on evaluating numbered diagrams. In the following, an arrow over the boxes shows which way we can connect them.

- The highest number n must always appear in a box of positive sign in the upper right quadrant.
- We can only connect boxes in the following way: $\overrightarrow{\boxed{-}\boxed{-}}$, $\overleftarrow{\boxed{+}\boxed{+}}$, $\overleftrightarrow{\boxed{-}\boxed{+}}$.
- We can never connect the boxes $\boxed{+}\boxed{-}$.
- If there is a single rightmost box, its sign must be positive.
- If there is a single leftmost box in the negative quadrant, its sign must be negative.

The last four rules were stated for horizontally connected boxes. Of course equivalent rules also hold for vertically connected ones.

Cycles

A cycle is a chain that contains no root. Let us concentrate for the moment on its 'circular part' of length n. We start by integrating out ϕ_1, ϕ_2, and so on, and for ϕ_i we pick the pole

$$\phi_i = \sigma_i \phi_{i+1} + \delta_i \,, \qquad i = 1, \ldots, n-1 \,, \tag{B.14}$$

where periodicity $\phi_{n+1} = \phi_1$ is implied, and $\sigma_i = \pm 1$, $\delta_i = \epsilon_{1,2}$. For ϕ_n we then pick the pole whose numerical value is determined in such a way that $\phi_n = \sigma_n \phi_1 + \delta_n$. This value can be determined by noting that the variable ϕ_l then takes the value

$$\phi_l = \left(\prod_{i=1}^{l-1} \sigma_i \right) \phi_1 - \sum_{j=1}^{l-1} \left(\prod_{i=j}^{l-1} \sigma_i \right) \delta_j \tag{B.15}$$

so that the cycle only gives a contribution if $\prod_{i=1}^{n} \sigma_i = -1$, as otherwise there is either no solution, or there is a double pole which gives no contribution. The total pole is thus given

by

$$(\tilde{\phi}) = (\sigma_1\phi_2 + \delta_1, \cdots, \sigma_{n-1}\phi_n + \delta_{n-1}, -\frac{1}{2}\sum_{j=1}^{n-1}\left(\prod_{i=j}^{n-1}\sigma_i\right)\delta_j + \frac{1}{2}\delta_n) \,. \tag{B.16}$$

Again, this only contributes if the value of the last entry has a well-defined positive imaginary part.

B.3 Some examples

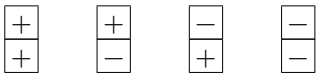
Let us now explain this more explicitly for the first low lying terms. For $k = 0$ and $k = 1$ there are no integrals. For $k = 2$ and $k = 3$ there is just one integral, so that one can simply sum over all poles. This has been treated in [68, 70, 71].

Four and five instantons: $n = 2$

Let us consider the case $n = 2$. There are four unsigned diagrams,

	$(\tilde{\phi}_1, \tilde{\phi}_2)$	n_Y
	$(b, b+\epsilon_2)$	3
	$(b, b+\epsilon_1)$	3
	$(b, b-\epsilon_2)$	-1
	$(b, b-\epsilon_1)$	-1

To arrive at the combinatorial weights, we first write down all signed versions of e.g.the first diagram:

 (B.17)

For each signed diagram we then write down all possible numbered diagrams and see if they are allowed. From the rules given above it is straightforward to see that only

$$\begin{array}{|c|}\hline + \\ \hline + \\ \hline \end{array} \qquad \begin{array}{|c|}\hline \mathbf{1} \\ \hline \mathbf{2} \\ \hline \end{array}\,,\ \begin{array}{|c|}\hline \mathbf{2} \\ \hline \mathbf{1} \\ \hline \end{array} \qquad (\phi_2+\epsilon_2, b), (b, b+\epsilon_2) \tag{B.18}$$

$$\begin{array}{|c|}\hline + \\ \hline - \\ \hline \end{array} \qquad \begin{array}{|c|}\hline \mathbf{2} \\ \hline \mathbf{1} \\ \hline \end{array} \qquad (-\phi_2+\epsilon_2, b+\epsilon_2) \tag{B.19}$$

are allowed. It is clear that the first diagram in the first line gives the same contribution as $(b, b+\epsilon_2)$. The diagram can be reduced to $(-b, b+\epsilon_2)$ by the same procedure as in (B.4). Since we pick up a minus sign in this process, the total combinatorial weight is $n_Y = 2+1 = 3$. Similarly, for the third diagram we obtain

$$\begin{array}{|c|}\hline + \\ \hline - \\ \hline \end{array} \qquad \begin{array}{|c|}\hline \mathbf{2} \\ \hline 1 \\ \hline \end{array} \qquad (-\phi_2+\epsilon_2, b)$$

This time we do not pick up a sign, so that $n_Y = -1$. Let us turn to the cyclic chains. If we choose $\sigma_1 = 1$, then $\phi_2 = \frac{1}{2}(\delta_1 - \delta_2)$, which we know does not contribute. The contributions thus come from $\sigma_1 = -1$, $\sigma_2 = 1$ and are given by

$$\begin{array}{cccc} (-\phi_2+\epsilon_1, \epsilon_1) & (-\phi_2+\epsilon_2, \epsilon_2) & (-\phi_2+\epsilon_1, \frac{1}{2}(\epsilon_1+\epsilon_2)) & (-\phi_2+\epsilon_2, \frac{1}{2}(\epsilon_1+\epsilon_2)) \\ \begin{array}{|c|c|}\hline - & + \\ \hline \end{array} & \begin{array}{|c|}\hline + \\ \hline - \\ \hline \end{array} & \begin{array}{|c|c|}\hline - & + \\ \hline \end{array} & \begin{array}{|c|}\hline + \\ \hline - \\ \hline \end{array} \end{array}$$

where we have represented the first two cycles by signed diagrams with root 0, and the second two by diagrams with root $\pm\frac{1}{2}(\epsilon_1 - \epsilon_2)$.

Six instantons: $n = 3$

Let us turn to $n = 3$ now. For completeness, we have listed all generalized Young diagrams, their values of the ϕ and the combinatorial weights in table B.1.

As an example, let us explain how to obtain the combinatorial weight for some of those cases. Take, for instance, the diagram $\begin{array}{|c|c|c|}\hline \times & & \\ \hline \end{array}$ and write down all signed diagrams. By the rules given above can immediately exclude all diagrams that have a minus sign in the rightmost box. The diagram $\tilde{Y}_0$ gives the same contribution as in the $U(N)$ case and has

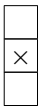
Diagram	(ϕ_1, ϕ_2, ϕ_3)	n_Y
	$(b, b+\epsilon_2, b+2\epsilon_2)$	15
	$(b, b+\epsilon_2, b+\epsilon_1)$	10
	$(b, b+\epsilon_1+2\epsilon_1)$	15
	$(b, b+\epsilon_2, b-\epsilon_2)$	-6
	$(b, b+\epsilon_1, b-\epsilon_1)$	-6
	$(b, b-\epsilon_2, b-2\epsilon_2)$	3
	$(b, b-\epsilon_1, b-2\epsilon_1)$	3
	$(b, b-\epsilon_2, b+\epsilon_1-\epsilon_2)$	-1
	$(b, b-\epsilon_1, b-\epsilon_1+\epsilon_2)$	-1
	$(b, b-\epsilon_1, b-\epsilon_2)$	2
	$(b, b+\epsilon_1, b+\epsilon_1-\epsilon_2)$	-5
	$(b, b+\epsilon_2, b+\epsilon_2-\epsilon_1)$	-5

Table B.1: Generalized Young diagrams, values of ϕ, and combinatorial weights

weight 6. The remaining three diagrams are

$\boxed{-}\boxed{+}\boxed{+}$	4	$\boxed{1}\boxed{2}\boxed{3}$	$(-\phi_2+\epsilon_1, b+\epsilon_1, b+2\epsilon_1)$
		$\boxed{2}\boxed{1}\boxed{3}$	$(-\phi_2+\epsilon_1, -\phi_3+2\epsilon_1, b+2\epsilon_1)$
		$\boxed{1}\boxed{3}\boxed{2}$	$(-\phi_3+\epsilon_1, \phi_3+\epsilon_1, b+\epsilon_1)$
		$\boxed{2}\boxed{3}\boxed{1}$	$(\phi_3+\epsilon_1, -\phi_3+\epsilon_1, b+\epsilon_1)$
$\boxed{-}\boxed{-}\boxed{+}$	2	$\boxed{1}\boxed{2}\boxed{3}$	$(\phi_2+\epsilon_1, -\phi_3+\epsilon_1, b+2\epsilon_1)$
		$\boxed{2}\boxed{1}\boxed{3}$	$(-\phi_3+\epsilon_1, -\phi_3+2\epsilon_1, b+2\epsilon_1)$
$\boxed{+}\boxed{-}\boxed{+}$	3	$\boxed{1}\boxed{2}\boxed{3}$	$(b, -\phi_3+\epsilon_1, b+2\epsilon_1)$
		$\boxed{2}\boxed{1}\boxed{3}$	$(-\phi_3+\epsilon_1, b, b+2\epsilon_1)$
		$\boxed{3}\boxed{1}\boxed{2}$	$(-\phi_2+\epsilon_1, \phi_3+2\epsilon_1, b)$

The top diagram is exactly as in the $U(N)$ case, so its combinatorial weight is 6. For the other diagrams, we have listed all numbered diagrams that contribute together with the precise pole they correspond to. Note that when converting the poles to the form of the table, it turns out that minus signs appear in such a fashion that all diagrams give positive contribution. The total combinatorial weight of $\boxed{\times}\boxed{}\boxed{}$ is thus 15.

Another example is $\boxed{}\boxed{}\boxed{\times}$. The rightmost box must have a positive sign, and the leftmost box is in a negative quadrant and must therefore have a negative sign. This leaves just two possibilities,

$\boxed{-}\boxed{+}\boxed{+}$	1	$\boxed{2}\boxed{1}\boxed{3}$	$(-\phi+\epsilon_1, -\phi_3+2\epsilon_1, b)$
$\boxed{-}\boxed{-}\boxed{+}$	2	$\boxed{1}\boxed{2}\boxed{3}$	$(\phi_2+\epsilon_1, -\phi_3+\epsilon_1, b)$
		$\boxed{2}\boxed{1}\boxed{3}$	$(-\phi_3+\epsilon_1, -\phi_3+2\epsilon_1, b)$

which give combinatorial weight 3.

Let us briefly describe the cycles now. For 3-cycles we get

$$(\tilde{\phi}_1, \tilde{\phi}_2, \tilde{\phi}_3) = (\sigma_1\phi_2+\delta_1, \sigma_2\phi_3+\delta_2, -\frac{1}{2}(\sigma_1\sigma_2\delta_1+\sigma_2\delta_2)+\frac{1}{2}\delta_3) \tag{B.20}$$

A priori, the allowed solutions are

$$\sigma_1 = -1, \sigma_2 = 1, \sigma_3 = 1, \delta_2 = \delta_3$$
$$\sigma_1 = 1, \sigma_2 = -1, \sigma_3 = 1, \delta_i = \epsilon_{1,2}$$
$$\sigma_1 = -1, \sigma_2 = -1, \sigma_3 = -1, \delta_1 = \delta_3$$

A closer analysis reveals however that all such solutions lead to zeros in the numerator due to the factor $\Delta(0)\Delta(\epsilon)$.

The only contribution thus comes from poles corresponding to 2-cycles with one attached arm. This means take the diagrams of the 2-cycles with root 0 and $\pm\frac{1}{2}(\epsilon_1 - \epsilon_2)$ and attach one box to it. Again, we want to compute the combinatorial weight of these configurations. Note however that in this case $\tilde{\phi}$ cannot be reduced to its numerical values.

For the extra root 0, note that the diagram $\square\,|\,\times\,|\,\square$ gives a vanishing contribution. We thus consider only $\boxed{\times}\,\square\,\square$. We have

$\boxed{-}\boxed{+}\boxed{+}$	4	$\boxed{1}\boxed{2}\boxed{3}$	$(-\phi_2 + \epsilon_1, \epsilon_1, 2\epsilon_1)$
		$\boxed{2}\boxed{1}\boxed{3}$	$(-\phi_2 + \epsilon_1, -\phi_3 + 2\epsilon_1, 2\epsilon_1)$
		$\boxed{2}\boxed{3}\boxed{1}$	$(\phi_3 + \epsilon_1, -\phi_3 + \epsilon_1, \epsilon_1)$
		$\boxed{1}\boxed{3}\boxed{2}$	$(-\phi_3 + \epsilon_1, \phi_3 + \epsilon_1, \epsilon_1)$
$\boxed{-}\boxed{-}\boxed{+}$	2	$\boxed{1}\boxed{2}\boxed{3}$	$(\phi_2 + \epsilon_1, -\phi_3 + \epsilon_1, 2\epsilon_1)$
		$\boxed{2}\boxed{1}\boxed{3}$	$(-\phi_3 + \epsilon_1, -\phi_3 + 2\epsilon_1, 2\epsilon_1)$

For $\begin{array}{|c|c|}\hline&\\\hline\times&\\\hline\end{array}$, the only signed diagram is $\begin{array}{|c|c|}\hline+&\\\hline-&+\\\hline\end{array}$, which has weight 4,

$$\begin{array}{|c|c|}\hline3&\\\hline1&2\\\hline\end{array} \quad \begin{array}{|c|c|}\hline3&\\\hline2&1\\\hline\end{array} \quad \begin{array}{|c|c|}\hline2&\\\hline1&3\\\hline\end{array} \quad \begin{array}{|c|c|}\hline1&\\\hline2&3\\\hline\end{array}$$

$$(-\phi_2 + \epsilon_1, \epsilon_1, \epsilon_2) \quad (-\phi_2 + \epsilon_1, -\phi_3 + \epsilon_2, \epsilon_2) \quad (-\phi_2 + \epsilon_2, \epsilon_2, \epsilon_1) \quad (-\phi_2 + \epsilon_2, -\phi_3 + \epsilon_1, \epsilon_1)$$

and for $\begin{array}{|c|c|}\hline \times & \\ \hline & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array}$ the only signed diagram $\begin{array}{|c|c|}\hline - & + \\ \hline & \multicolumn{1}{c}{-} \\ \cline{1-1}\end{array}$ has weight 2:

$$\begin{array}{|c|c|}\hline 1 & 3 \\ \hline & \multicolumn{1}{c}{2} \\ \cline{2-2}\end{array} \qquad\qquad \begin{array}{|c|c|}\hline 2 & 3 \\ \hline & \multicolumn{1}{c}{1} \\ \cline{2-2}\end{array}$$

$$(-\phi_3+\epsilon_1, -\phi_3+\epsilon_2, \epsilon_1) \qquad\qquad (-\phi_3+\epsilon_2, -\phi_3+\epsilon_1, \epsilon_1)$$

All other configurations can be obtained by exchanging $\epsilon_1 \leftrightarrow \epsilon_2$.

For the root $\pm\frac{1}{2}(\epsilon_1-\epsilon_2)$, note that no box can be attached to the box $\frac{\epsilon_1+\epsilon_2}{2}$ because of the zeros in the numerator. This leaves just three configurations

$\begin{array}{\|c\|c\|c\|}\hline - & - & + \\ \hline \end{array}$	2	$\begin{array}{\|c\|c\|c\|}\hline 1 & 2 & 3 \\ \hline \end{array}$	$(\phi_2+\epsilon_1, -\phi_3+\epsilon_1, \frac{\epsilon_1+\epsilon_2}{2})$
		$\begin{array}{\|c\|c\|c\|}\hline 2 & 1 & 3 \\ \hline \end{array}$	$(-\phi_3+\epsilon_1, -\phi_3+2\epsilon_1, \frac{\epsilon_1+\epsilon_2}{2})$
$\begin{array}{\|c\|c\|c\|}\hline - & + & + \\ \hline \end{array}$	1	$\begin{array}{\|c\|c\|c\|}\hline 2 & 1 & 3 \\ \hline \end{array}$	$(-\phi_2+\epsilon_1, -\phi_3+2\epsilon_1, \frac{\epsilon_1+\epsilon_2}{2})$
$\begin{array}{\|c\|c\|}\hline + & \\ \hline - & + \\ \hline \end{array}$	1	$\begin{array}{\|c\|c\|}\hline 1 & \\ \hline 2 & 3 \\ \hline \end{array}$	$(-\phi_2+\epsilon_2, -\phi_3+\epsilon_1, \frac{\epsilon_1+\epsilon_2}{2})$

and similarly for their mirror images under $\epsilon_1 \leftrightarrow \epsilon_2$.

To obtain the full contribution for the toy model (B.11) we also need to include all poles that consist of a 2-cycle and the root b. In total there are thus 112 poles.

Appendix C

$SU(2)$ Seiberg-Witten curves

Since the discovery of Seiberg-Witten theory a few different (yet physically equivalent) parametrizations for the $SU(2) = Sp(1)$ Seiberg-Witten curve have appeared in the literature. Let us summarize these different approaches here.

First of all, the Seiberg-Witten curve for the $SU(2)$ theory coupled to four hypermultiplets can be witten in the hyperelliptic form [166]

$$y^2 = P_{U(2)}(w)^2 - fQ, \tag{C.1}$$

where

$$P_{U(2)}(w) = w^2 - \tilde{u}, \quad f = \frac{4q_{U(2)}}{(1+q_{U(2)})^2}, \quad q_{U(2)} = \frac{\theta_2^4(\tau_{\rm IR})}{\theta_3^4(\tau_{\rm IR})}, \quad Q = \prod_{j=1}^{4}(w - \tilde{m}_j).$$

We should be careful that the mass parameters $\tilde{m}_j$ are *not* exactly the hypermultiplet masses $\tilde{\mu}_j$. Instead, they are related to the hypermultiplet masses $\tilde{\mu}_j$ as

$$\tilde{m}_j = -\tilde{\mu}_j + \frac{q_{U(2)}}{2(1+q_{U(2)})}\sum_k \tilde{\mu}_k. \tag{C.2}$$

Indeed, the meromorphic Seiberg-Witten differential

$$\lambda = \frac{-w + \frac{q_{U(2)}}{2(1+q_{U(2)})}\sum_k \tilde{\mu}_k}{2\pi i} d\log\left(\frac{P_{U(2)}(w) - y}{P_{U(2)}(w) + y}\right) \tag{C.3}$$

has residues $\pm\tilde{\mu}_j$ at the position $w = \tilde{\mu}_j$, so that the parameters $\tilde{\mu}_j$ are the hypermultiplet masses. These are also the parameters that appear in the Nekrasov formalism.

Another parametrization is found by D4/NS5 brane engineering in type IIA. The Seiberg-Witten curve (C.1) can be rewritten in the MQCD form [33]

$$(w-\tilde{m}_1)(w-\tilde{m}_2)t^2-(1+q_{U(2)})(w^2-u)t+q_{U(2)}(w-\tilde{m}_3)(w-\tilde{m}_4)=0 \tag{C.4}$$

by the coordinate transformation

$$t=-\frac{(1+q_{U(2)})(y-P(w))}{2(w-\tilde{m}_1)(w-\tilde{m}_2)}.$$

In this parametrization the meromorphic Seiberg-Witten 1-form can simply be taken to be

$$\lambda=w\frac{dt}{t}. \tag{C.5}$$

This differential differs from the one in equation (C.3) by an exact 1-form. It has first order poles at the positions $t\in\{0,q_{U(2)},1,\infty\}$. At $t=\infty$ and $t=0$ the residues are given by the mass parameters $\{\tilde{m}_1,\tilde{m}_2\}$ and $\{\tilde{m}_3,\tilde{m}_4\}$ respectively, whereas at $t=1$ and $t=q_{U(2)}$ there is only a single nonzero residue. The mass-parameters at $t=0,\infty$ parametrize the Cartan of the flavor symmetry group $SU(2)$, whereas the single residue at the other two punctures is an artifact of the chosen parametrization (that only sees a $U(1)\subset SU(2)$).

To restore the $SU(2)$ flavor symmetry at each of the four punctures, Gaiotto introduced the parametrization [25]

$$\tilde{w}^2=\left(\frac{(\tilde{m}_1+\tilde{m}_2)t^2+q_{U(2)}(\tilde{m}_3+\tilde{m}_4)}{4t(t-1)(t-q_{U(2)})}\right)^2+\frac{\tilde{m}_1\tilde{m}_2t^2+(1+q_{U(2)})\tilde{u}t+q_{U(2)}\tilde{m}_3\tilde{m}_4}{t^2(t-1)(t-q_{U(2)})}. \tag{C.6}$$

This is found from equation (C.4) by eliminating the linear term in w and mapping $w\mapsto\tilde{w}=tw$. By writing equation (C.6) in the form

$$\tilde{w}^2=\varphi_2(t), \tag{C.7}$$

it is clear that the Seiberg-Witten curve is a branched double cover over a two-sphere $\mathbb{P}^1$

with punctures at $t = 0, q_{U(2)}, 1, \infty$. The coefficients of φ_2 at the punctures are given by

$$\begin{aligned} t = 0: \qquad & \varphi_2 \sim \frac{(\tilde{\mu}_3 - \tilde{\mu}_4)^2}{4} \frac{dt^2}{t^2} \\ t = q_{U(2)}: \qquad & \varphi_2 \sim \frac{(\tilde{\mu}_3 + \tilde{\mu}_4)^2}{4} \frac{dt^2}{(t - q_{U(2)})^2} \\ t = 1: \qquad & \varphi_2 \sim \frac{(\tilde{\mu}_1 + \tilde{\mu}_2)^2}{4} \frac{dt^2}{(t-1)^2} \\ t = \infty: \qquad & \varphi_2 \sim \frac{(\tilde{\mu}_1 - \tilde{\mu}_2)^2}{4} \frac{dt^2}{t^2} \ . \end{aligned} \tag{C.8}$$

So if we keep

$$\lambda = \tilde{w} \frac{dt}{t} \tag{C.9}$$

as the Seiberg-Witten differential (which is allowed since it only differs from (C.5) by a shift of the flavor current by a multiple of the gauge current), we find that its residues are given by the square-roots of the coefficients of φ_2 in equation (C.8).

The Seiberg-Witten curve in the Gaiotto parametrization (C.7) is invariant under Möbius transformations, and therefore completely symmetric in all four punctures. This follows automatically as λ and φ_2 are respectively a 1-form and a 2-form on the two-sphere $\mathbb{P}^1$.

Furthermore, an $Sp(1)$ parametrization of the Seiberg-Witten curve is given by [90]

$$xy^2 = P_{Sp(1)}(x)^2 - g^2 \prod \prod (x - \mu_j^2), \tag{C.10}$$

where

$$P_{Sp(1)}(x) = x(x-u) + g \prod \prod \mu_j, \quad g^2 = \frac{4\tilde{q}_{Sp(1)}}{\left(1 + \tilde{q}_{Sp(1)}\right)^2},$$

Since $Sp(1) = SU(2)$, the parametrization (C.10) should be equivalent to the curve defined by (C.1). Indeed, comparing the discriminants of these two curves yields non-trivial relations between the Coulomb parameters and the masses, that become trivial in the weak-coupling limit [90].

In fact, there is a simple relation between the $Sp(1)$ parametrization (C.10) and the $SU(2)$ curve that Seiberg and Witten originally proposed [16]. By expanding equation (C.10)

and dividing out the constant term in x, we find an equation of the form $y^2 = x^3 + \ldots$. After some redefinitions this gives the original $SU(2)$ parametrization [90].

Let us finally mention that by the coordinate transformation $x = v^2$, $y = \tilde{y}/v$ and

$$\frac{2s}{\left(1 + \tilde{q}_{Sp(1)}\right)} = -\frac{\tilde{y} - P_{Sp(1)}(v^2)}{(v^2 - \mu_1^2)(v^2 - \mu_2^2)},$$

the $Sp(1)$ Seiberg-Witten curve (C.10) can be written in the Witten-form

$$(v^2 - \mu_1^2)(v^2 - \mu_2^2)s^2 - (1 + \tilde{q}_{Sp(1)})P_{Sp(1)}(v^2)s + \tilde{q}_{Sp(1)}(v^2 - \mu_3^2)(v^2 - \mu_4^2) = 0. \qquad \text{(C.11)}$$

This representation of the $Sp(1)$ SW curve is a double cover over the original $Sp(1)$ SW curve (C.10), because of the coordinate transformation $x = v^2$.

The above curve describes the embedding of the $Sp(1)$ gauge theory in string theory using a D4/NS5 brane construction including orientifold branes [36]. It should be viewed as being embedded in the covering space of the orientifold. For each D4-brane at position $v = v^*$ there is a mirror brane at position $v = -v^*$. The extra factor v^2 in the polynomial P can be identified with two extra D4-branes that are forced to sit at the orientifold at $v = 0$.

Appendix D

More about the trifundamental half-hypermultiplet

In this appendix we explicitly show the reduction of the Lagrangian for the $SU(2)$ trifundamental half-hypermultiplet to bifundamental and fundamental hypermultiplets when we Higgs one or more of the $SU(2)$ gauge groups. Before starting this argument, let us quickly remind ourselves about flavor symmetry enhancement for (pseudo-)real representations.

Flavor symmetry enhancement

Let us briefly explain a way to understand the enhancement of flavor symmetries for matter transforming in a (pseudo-)real representation. Although this is not of direct importance for this paper, it will be useful as background and in the following.

First of all, recall the familiar statement of flavor symmetry enhancement. The flavor symmetry group of N hypers in a real representation of the gauge group is enhanced from $U(1)^N$ to $Sp(N)$, whereas the flavor symmetry of N hypers in a pseudo-real representation is enhanced to $SO(2N)$. The cases that are important for us is the single hyper in the bifundamental of $SU(2)$ which has enhanced flavor symmetry $Sp(1) = SU(2)$, and two hypers in the fundamental which enhance to $SO(4) = SU(2) \times SU(2)$.

A pseudoreal representation is characterized by an antilinear map σ_G such that $\sigma_G^2 = -1$. This map corresponds to the complex conjugation, so that in case of the fundamental or adjoint representation $\sigma_G^{-1} T \sigma_G = T^* = -T^t$ for $T \in \mathfrak{g}$. For real representations the only difference is that $\sigma_G^2 = 1$. Note that σ_G is automatically unitary.

The basic idea behind the enhancement is that the N-dimensional 'flavor vector' Q_i is enlarged to a $2N$-dimensional vector $(Q_i, \sigma_G \tilde{Q}_i)$ (which still is in the representation R of the

gauge group). What needs to be shown is that the terms in the Lagrangian are invariants of $SO(2N)$ or $Sp(N)$.

For a single hypermultiplet in the fundamental representation of $SU(2)$, which is pseudo-real, the kinetic term in the Lagrangian $\mathcal{L}_{\text{fh}}$ can be rewritten as

$$\begin{aligned} Q^\dagger e^V Q + \tilde{Q}^t e^{-V} \tilde{Q}^* &= Q^\dagger e^V Q + \tilde{Q}^\dagger e^{-V^t} \tilde{Q} \\ &= Q^\dagger e^V Q + \tilde{Q}^\dagger \sigma_G^{-1} e^V \sigma_G \tilde{Q} \\ &= \begin{pmatrix} Q^\dagger, & \tilde{Q}^\dagger \sigma_G^\dagger \end{pmatrix} e^V \begin{pmatrix} Q \\ \sigma_G \tilde{Q} \end{pmatrix}. \end{aligned} \tag{D.1}$$

The Yukawa coupling in the Lagrangian $\mathcal{L}_{\text{fh}}$ is proportional to

$$\begin{aligned} 2\,\tilde{Q}^t \Phi Q &= Q^t \Phi^t \tilde{Q} + \tilde{Q}^t \Phi Q \\ &= Q^t \sigma_G \Phi \sigma_G \tilde{Q} + \tilde{Q}^t \Phi Q \\ &= \begin{pmatrix} Q^t, & \tilde{Q}^t \sigma_G^t \end{pmatrix} \sigma_G \Phi \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} Q \\ \sigma_G \tilde{Q} \end{pmatrix}. \end{aligned} \tag{D.2}$$

In both cases we used the fact $\sigma_G^{-1} T \sigma_G = -T^t$ for $T \in \mathfrak{g}$.

To see that (D.2) is an $SO(2)$ invariant, we can make a change of basis to $Q^\pm = Q \pm i\sigma_G \tilde{Q}$. The enhanced flavor group $SO(2)$ then acts in the fundamental on this new basis, and both (D.1) and (D.2) are the standard diagonal invariants. This argument generalizes in a straightforward way to an arbitrary number of hypers Q_i.

A similar argument works for real representations, such as for the bifundamental in $SU(2)_A \times SU(2)_B$. In this case the kinetic term can written exactly in the form (D.1) as well. The Yukawa term picks up a minus sign, due to the fact that now $\sigma_G^2 = 1$, so that the invariant is found by replacing

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \mapsto \begin{pmatrix} 0 & \mathbf{1}_N \\ -\mathbf{1}_N & 0 \end{pmatrix}$$

in (D.2). The Lagrangian is thus indeed invariant under $Sp(N)$. (Note that by $Sp(N)$ we actually mean $USp(N) = U(2N) \cap Sp(2N, \mathbb{C})$ here.)

Reduction of the $SU(2)$ trifundamental half-hypermultiplet

We first show that the Lagrangian for the $SU(2)$ trifundamental half-hyper reduces to that of a massive $SU(2)$ bifundamental hyper when one of the three $SU(2)$ gauge groups is Higgsed. Start with the Yukawa terms

$$W = \epsilon^{bb'}\epsilon^{cc'}Q_{abc}\Phi_1^{aa'}Q_{a'b'c'} + \epsilon^{aa'}\epsilon^{cc'}Q_{abc}\Phi_2^{bb'}Q_{a'b'c'} + \epsilon^{aa'}\epsilon^{bb'}Q_{abc}\Phi_3^{cc'}Q_{a'b'c'} \tag{D.3}$$

in the Lagrangian of the $SU(2)$ trifundamental half-hyper (as derived in section 3.6.1). The first gauge group is Higgsed by setting $(\Phi_1)^a_{\ a'} = m_1(\sigma_3)^a_{\ a'}$, so that the superpotential W reduces to

$$W = m_1(\sigma_3)^{a'}_{\ a}Q^{abc}Q_{a'bc} - \epsilon^{cc'}Q^a_{\ bc}\Phi_2^{bb'}Q_{ab'c'} - \epsilon^{bb'}Q^a_{\ bc}\Phi_3^{cc'}Q_{ab'c'}. \tag{D.4}$$

When identifying

$$\begin{aligned} Q_{bc} &\equiv Q_{1bc} = -Q^2_{\ bc} \\ \tilde{Q}_{bc} &\equiv Q_{2bc} = Q^1_{\ bc} \end{aligned} \tag{D.5}$$

in equation (D.4), we indeed recover the Yukawa terms

$$W = 2m_1\tilde{Q}^{bc}Q_{bc} - 2\epsilon^{cc'}\tilde{Q}_{bc}\Phi_2^{bb'}Q_{b'c'} - 2\epsilon^{bb'}\tilde{Q}_{bc}\Phi_3^{cc'}Q_{b'c'} \tag{D.6}$$

of the bifundamental hyper of mass m_1. Here we made use of the identity $\sigma_G^{-1}T\sigma_G = -T^t$.

Remark that the identifications in equation (D.5) reduce the $SU(2)$ R-symmetry for the trifund to that of the bifund, while identifying the gauge symmetry of the first $SU(2)$ gauge group with the enhanced flavor symmetry of the bifund. Also notice that the mass-term breaks the enhanced flavor symmetry of the bifund.

Let us continue by Higgsing the second gauge group by setting $(\Phi_2)^b_{\ b'} = m_2(\sigma_3)^b_{\ b'}$. The single bifundamental hyper turns into two fundamental hypers

$$\begin{aligned} Q_{(k)c} &\equiv Q_{kc} \\ \tilde{Q}_{(k)c} &\equiv \tilde{Q}_{kc}, \end{aligned} \tag{D.7}$$

where a subscript (.) refer to a flavor index. The Yukawa terms can be repackaged as

$$W = m_1 \, (\sigma_3)^g_{\ f} \, \delta^l_{\ k} \, Q^{(f)(k)c} Q_{(g)(l)c} + m_2 \, \delta^g_{\ f} \, (\sigma_3)^l_{\ k} \, Q^{(f)(k)c} Q_{(g)(l)c} \\ + \epsilon^{fg} \, \epsilon^{kl} Q_{(f)(k)c} \Phi_3^{cc'} Q_{(g)(l)c'}, \tag{D.8}$$

if we furthermore make the identifications $Q = Q_{(f=1)}$ and $\tilde{Q} = Q_{(f=2)}$. This superpotential describes two fundamental hypers whose flavor symmetry enhances to $SO(4)$ when the masses are turned off.

As a consistency check let us Higgs both bifundamental gauge groups by setting $(\Phi_2)^b_{\ b'} = m_2(\sigma_3)^b_{\ b'}$ and $(\Phi_3)^c_{\ c'} = m_3(\sigma_3)^c_{\ c'}$ in equation (D.6). This results in the Yukawa terms

$$W = m_1 \, \delta^l_{\ k} \, \delta^n_{\ m} \, (\sigma_3)^g_{\ f} \, Q^{(f)(k)(m)} Q_{(g)(l)(n)} + m_2 \, (\sigma_3)^l_{\ k} \, \delta^n_{\ m} \, \delta^g_{\ f} \, Q^{(f)(k)(m)} Q_{(g)(l)(n)} \\ + m_3 \, \delta^l_{\ k} \, (\sigma_3)^n_{\ m} \, \delta^g_{\ f} \, Q^{(f)(k)(m)} Q_{(g)(l)(n)}, \tag{D.9}$$

corresponding to the superpotential of eight half-hypers with a diagonal mass matrix with eigenvalues $\pm m_1 \pm m_2 \pm m_3$, as expected.

Appendix E

Contour integrands for Sicilian quivers

In this appendix we summarize the contour integrand formulae for the Sicilian quiver gauge theories.

$Sp(1) - SO(4)$ bifundamental full hypermultiplet

The contour integrand for the massive full $Sp(1) - SO(4)$ bifundamental hyper is

$$\begin{aligned} \mathbf{z}_{\text{2bif,k}_1,\text{k}_2}^{Sp(1)-SO(4)}(\phi,\psi,\mu,a,b_1,b_2) &= \left(\prod_{l=1}^{2}\Delta_1(\mu\pm b_l)\right)\Delta_2(\mu\pm a)P_2(\mu)^{\chi_\phi} \\ &\times\left(\frac{\Delta(\mu-\epsilon_-)\Delta(\mu+\epsilon_-)}{\Delta(\mu+\epsilon_+)\Delta(\mu-\epsilon_+)}\right)\left(\frac{\Delta_2(\mu-\epsilon_-)\Delta_2(\mu+\epsilon_-)}{\Delta_2(\mu+\epsilon_+)\Delta_2(\mu-\epsilon_+)}\right)^{\chi_\phi}, \end{aligned} \tag{E.1}$$

where the $Sp(1)$ instanton parameter $k_1 = 2n_1 + \chi_\phi$, the deformation parameters $\epsilon_\pm = \frac{\epsilon_1 \pm \epsilon_2}{2}$ and $\pm$ is an abbreviation for a product over both terms. Furthermore, μ is the physical mass parameter and

$$\begin{aligned} \Delta_1(x) &= \prod_{i=1}^{n_1}(\phi_i^2 - x^2) \\ \Delta_2(x) &= \prod_{j=1}^{k_2}(\psi_j^2 - x^2) \\ \Delta(x) &= \prod_{i,j=1}^{n_1,k_2}\left((\phi_i+\psi_j)^2 - x^2\right)\left((\phi_i-\psi_j)^2 - x^2\right) \\ P_2(x,b) &= \prod_{l=1}^{n_2}(b_l^2 - x^2). \end{aligned}$$

$Sp(1) - SO(4)$ single bifundamental half-hypermultiplet

The $Sp(1) - SO(4)$ double bifundamental integrand (E.1) becomes a complete square when the bifundamental mass vanishes. The integrand for the bifundamental half-hyper is thus simply

$$\mathbf{z}^{Sp(1)-SO(4)}_{\text{bif},\mathfrak{k}_1,\mathfrak{k}_2}(\phi,\psi,a,b_1,b_2) \;=\; \prod_{i=1}^{n_1}\left(\phi_i^2-b_1^2\right)\left(\phi_i^2-b_2^2\right)\prod_{j=1}^{k_2}\left(a^2-\psi_j^2\right)\frac{\Delta(\epsilon_-)}{\Delta(\epsilon_+)}\left(b_1b_2\frac{\Delta_2(\epsilon_-)}{\Delta_2(\epsilon_+)}\right)^{\chi_\phi}.$$

The factor $\Delta(\epsilon_+)$ in the denominator cannot be canceled by a contribution from the gauge multiplets, and therefore brings in additional poles.

$U(2)^3$ trifundamental full hypermultiplet

Starting from the equivariant index

$$\text{Ind}_{\mathbf{T}} = \int_{\mathbb{C}^2} \text{Ch}_{\mathbf{T}}(\mathcal{E}_{U(N_1)}\otimes\mathcal{E}_{U(N_2)}\otimes\mathcal{E}_{U(N_3)}\otimes\mathcal{L}\otimes M)\,\text{Td}_{\mathbf{T}}(\mathbb{C}^2)\,, \tag{E.2}$$

we obtain the contour integrand

$$\begin{aligned}
\mathbf{z}^{(N_1,N_2,N_3)}_{k_1,k_2,k_3} \;=\;& \prod_{i,m,n}^{k_1,N_2,N_3}(\phi_{1,i}+b_m+c_n+\mu)\prod_{j,l,n}^{k_2,N_1,N_3}(\phi_{2,j}+a_l+c_n+\mu)\prod_{k,l,m}^{k_3,N_1,N_2}(\phi_{3,k}+a_l+b_m+\mu)\\
&\times\prod_{i,j,n}^{k_1,k_2,N_3}\frac{(\phi_{1,i}+\phi_{2,j}+c_n-\epsilon_-+\mu)(\phi_{1,i}+\phi_{2,j}+c_n+\epsilon_-+\mu)}{(\phi_{1,i}+\phi_{2,j}+c_n+\mu+\epsilon_+)(\phi_{1,i}+\phi_{2,j}+c_n+\mu-\epsilon_+)}\\
&\times\prod_{i,k,m}^{k_2,k_3,N_1}\frac{(\phi_{1,i}+\phi_{3,k}+b_m-\epsilon_-+\mu)(\phi_{1,i}+\phi_{3,k}+b_m+\epsilon_-+\mu)}{(\phi_{1,i}+\phi_{3,k}+b_m+\mu+\epsilon_+)(\phi_{1,i}+\phi_{3,k}+b_m+\mu-\epsilon_+)}\\
&\times\prod_{j,k,l}^{k_2,k_3,N_1}\frac{(\phi_{2,j}+\phi_{3,k}+a_l-\epsilon_-+\mu)(\phi_{2,j}+\phi_{3,k}+a_l+\epsilon_-+\mu)}{(\phi_{2,j}+\phi_{3,k}+a_l+\mu+\epsilon_+)(\phi_{2,j}+\phi_{3,k}+a_l+\mu-\epsilon_+)}\\
&\times\prod_{i,j,k}^{k_1,k_2,k_3}\frac{(\phi^{123}_{ijk}+\epsilon+\mu)(\phi^{123}_{ijk}+\epsilon_1-\epsilon_2+\mu)(\phi^{123}_{ijk}-\epsilon_1+\epsilon_2+\mu)(\phi^{123}_{ijk}+\mu)^4(\phi^{123}_{ijk}+\mu-\epsilon)}{(\phi^{123}_{ijk}+\epsilon_1+\mu)^2(\phi^{123}_{ijk}+\epsilon_2+\mu)^2(\phi^{123}_{ijk}-\epsilon_1+\mu)^2(\phi^{123}_{ijk}-\epsilon_2+\mu)^2}
\end{aligned} \tag{E.3}$$

where $\phi^{123}_{ijk} = \phi_{1,i}+\phi_{2,j}+\phi_{3,k}$ and $\epsilon = \epsilon_1+\epsilon_2$ and $\epsilon_\pm = \frac{\epsilon_1\pm\epsilon_2}{2}$. The contour integrand for the massive full $SU(2)$ trifundamental hypermultiplet is found by setting $N_{1/2/3} = 2$. If we set the instanton parameter $k_3 = 0$, we recover the contour integrand for two copies of the bifundamental of mass $\pm c$.

The one-loop contribution to the Nekrasov partition function

$$Z^{\text{2trif}}_{\text{1-loop}} \quad = \quad \prod_{i,j=1}^{\infty} \prod_{l,m,n=1}^{2} (a_l + b_m + c_n + \mu + \epsilon_+ + i\epsilon_1 + j\epsilon_2)^{-1}$$

is obtained from the perturbative contribution to the equivariant index (E.2).

$Sp(1)^3$ trifundamental full hypermultiplet

The equivariant index of the Dirac operator for the $Sp(1)^3$ trifundamental full hyper is

$$\text{Ind}_{\mathbf{T}} \quad = \quad \int_{\mathbb{C}^2} \text{Ch}_{\mathbf{T}}(\mathcal{E}^1_{Sp(1)} \otimes \mathcal{E}^2_{Sp(1)} \otimes \mathcal{E}^3_{Sp(1)} \otimes \mathcal{L} \otimes M)\, \text{Td}_{\mathbf{T}}(\mathbb{C}^2). \tag{E.4}$$

Denote the instanton numbers to be (k_1, k_2, k_3) and write $k_\alpha = 2n_\alpha + \chi_\alpha$ for $\alpha = 1, 2, 3$. The corresponding contour integrand is given by

$$\begin{aligned} \mathbf{z}^{k_1,k_2,k_3}_{\text{2trif}}(\vec{\phi}, \vec{a}, m) \quad &= \quad H_{1;23}(m) H_{2;31}(m) H_{3;12}(m) \\ &\times \quad \frac{H_{12;3}(m-\epsilon_-) H_{12;3}(m+\epsilon_-) H_{23;1}(m-\epsilon_-)}{H_{12;3}(m-\epsilon_+) H_{12;3}(m-\epsilon_+) H_{23;1}(m+\epsilon_+)} \\ &\times \quad \frac{H_{23;1}(m+\epsilon_-)}{H_{23;1}(m-\epsilon_+)} \frac{H_{31;2}(m-\epsilon_-) H_{31;2}(m+\epsilon_-)}{H_{31;2}(m+\epsilon_+) H_{31;2}(m-\epsilon_-)} \\ &\times \quad \frac{H(m-\epsilon) H(m+\epsilon_1-\epsilon_2) H(m-\epsilon_1+\epsilon_2) H(m)^4 H(m+\epsilon)}{H(m-\epsilon_1)^2 H(m-\epsilon_2)^2 H(m+\epsilon_1)^2 H(m+\epsilon_2)^2} \end{aligned} \tag{E.5}$$

where

$$
\begin{aligned}
P_\alpha(x) &= (x^2 - a_\alpha)^2 \\
P_{\alpha\beta}(x) &= ((a_\alpha + a_\beta)^2 - x^2)((a_\alpha - a_\beta)^2 - x^2) \\
\Delta_\alpha(x) &= \prod_{i=1}^{n_\alpha} (\phi_{\alpha,i}^2 - x^2) \\
\Delta_{\alpha\beta}(x) &= \prod_{i,j=1}^{n_\alpha,n_\beta} \left((\phi_{\alpha,i} + \phi_{\beta,j})^2 - x^2\right) \left((\phi_{\alpha,i} - \phi_{\beta,j})^2 - x^2\right) \\
\Delta_{123}(x) &= \prod_{i,j,k} \left(x^2 - (\phi_{1,i} + \phi_{2,j} + \phi_{3,k})^2\right) \left(x^2 - (\phi_{1,i} + \phi_{2,j} - \phi_{3,k})^2\right) \\
&\quad \times \left(x^2 - (\phi_{1,i} - \phi_{2,j} - \phi_{3,k})^2\right) \left(x^2 - (\phi_{1,i} - \phi_{2,j} + \phi_{3,k})^2\right) \\
H_{\alpha;\beta\gamma}(x) &= \left(\prod_{i=1}^{n_\alpha} P_{\beta\gamma}(\phi_{\alpha,i} - x) P_{\beta\gamma}(\phi_{\alpha,i} + x)\right) \left(P_{\beta\gamma}(x)\right)^{\chi_\alpha} \\
H_{\alpha\beta;\gamma}(x) &= \Delta_{\alpha\beta}(a_\gamma + x)\Delta_{\alpha\beta}(a_\gamma - x)\Delta_\alpha(a_\gamma + x)^{\chi_\beta}\Delta_\alpha(a_\gamma - x)^{\chi_\beta} \\
&\quad \times \Delta_\beta(a_\gamma + x)^{\chi_\alpha}\Delta_\beta(a_\gamma - x)^{\chi_\alpha} \left(P_\gamma(x)\right)^{\chi_\alpha\chi_\beta} \\
H(x) &= \Delta_{123}(x)\Delta_{12}(x)^{\chi_3}\Delta_{13}(x)^{\chi_2}\Delta_{23}(x)^{\chi_1} \\
&\quad \times \Delta_1(x)^{\chi_2\chi_3}\Delta_2(x)^{\chi_3\chi_1}\Delta_3(x)^{\chi_1\chi_2} x^{\chi_1\chi_2\chi_3} \,.
\end{aligned}
$$

$Sp(1)$ trifundamental half-hypermultiplet

Setting the mass m to zero in (E.5), we find a complete square. The contour integrand for the massless $Sp(1)$ trifundamental half-hyper is its square-root

$$
\begin{aligned}
\mathbf{z}_{\text{trif}}^{k_1,k_2,k_3}(\vec{\phi},\vec{a}) &= P_{23}(\phi_1)P_{31}(\phi_2)P_{12}(\phi_3)(a_1^2 - a_2^2)^{\chi_3}(a_2^2 - a_3^2)^{\chi_1}(a_3^2 - a_1^2)^{\chi_2} \\
&\times \left(\frac{H_{23;1}(\epsilon_-)H_{31;2}(\epsilon_-)H_{12;3}(\epsilon_-)}{H_{23;1}(\epsilon_+)H_{31;2}(\epsilon_+)H_{12;3}(\epsilon_+)}\right) \left(\frac{H(0)^2 H(\epsilon_1+\epsilon_2)H(\epsilon_1-\epsilon_2)}{H(\epsilon_1)^2 H(\epsilon_2)^2}\right),
\end{aligned}
\tag{E.6}
$$

up to a sign ambiguity. This expression does not yield a physical result, however, since $H(0)$ contains a vanishing contribution when $\chi_{1/2/3} = 1$.

Appendix F

Kac determinant at the lowest level

Here we consider the Kac determinant at the lowest level. The result for the untwisted case is well known, see e.g. [167, 94]. The twisted case also follows straightforwardly by modifying the derivation for the untwisted case. For example, the Kac determinant for the Drinfeld-Sokolov reduction with respect to the minimal nilpotent was determined in [168]. Here we need to perform it with respect to the principal nilpotent.

First, consider the untwisted Verma module. Let the zero modes of the free bosons be $\vec{J}_0$. Pick a simple root $\vec{\alpha}_i$, and let $\varphi_i = \vec{\alpha}_i \cdot \vec{\varphi}$. As explained in Sec. 4.3.1.3, elements in the W-algebra can be expanded in terms of the energy-momentum tensor T_i for φ_i and the free bosons perpendicular to φ_i. At level one, $(T_i)_{-1} = (\vec{\alpha}_i \cdot \vec{J}_0) J_{i,-1}$, and similarly, the (-1)-mode of any operator constructed out of T_i comes with a factor of $\vec{\alpha}_i \cdot \vec{J}_0$. Thus, there is a null state in the W-algebra Verma module when $\vec{\alpha}_i \cdot \vec{J}_0 = 0$. Therefore the Kac determinant at level one is divisible by $\vec{\alpha}_i \cdot \vec{J}_0 = \vec{\alpha}_i \cdot \vec{a} + Q$. Due to the shifted Weyl invariance, the Kac determinant is divisible by $\vec{\alpha} \cdot \vec{a} + Q$ for all roots α. The Kac determinant has degree $2\sum(w_i - 1)$ in $\vec{a}$, which equals the number of roots. It follows that

$$\text{(Kac determinant at level 1)} \propto \prod_{\vec{\alpha} \in \Delta} (\vec{\alpha} \cdot \vec{a} + Q) \,. \tag{F.1}$$

Next, consider the states at level $-1/r$ of the $\mathbb{Z}_r$-twisted Verma module. Denote by o the $\mathbb{Z}_r$ action, under which the zero mode $\vec{J}_0$ is invariant. We regard it as in the Cartan of G, the S-dual of the $\mathbb{Z}_r$-invariant subalgebra of Γ. Recall that the long simple roots γ_i of G are given by $\gamma_i = \vec{\alpha}_i + \vec{\alpha}_{o(i)}$ when $r = 2$ and similarly for $r = 3$, (A.10), (A.11).

Take a simple root such that $\vec{\alpha}_i \neq \vec{\alpha}_{o(i)}$. Assuming $\Gamma \neq A_{2n}$, we have $\vec{\alpha}_i \cdot \vec{\alpha}_{o(i)} = 0$. Therefore the operators in the W-algebra can be written in terms of the energy-momentum

tensors T_i, $T_{o(i)}$ (and $T_{o^2(i)}$ if $r = 3$) for the free bosons φ_i, $\varphi_{o(i)}$ (and $\varphi_{o^2(i)}$) and free bosons perpendicular to φ_i, $\varphi_{o(i)}$ (and $\varphi_{o^2(i)}$). Now, the level $-(1/r)$ states arise from $T_i - T_{o(i)}$ when $r = 2$, and from $T_i + e^{2\pi i/3}T_{o(i)} + e^{4\pi i/3}T_{o^2(i)}$ when $r = 3$. Recall that $T_i \propto -(\partial\varphi_i\partial\varphi_i) + Q\partial^2\varphi_i/2$. Therefore the $(-1/r)$-mode always arises with the combination

$$(\vec{\alpha}_i \cdot \vec{J}_0 - Q(1 - 1/r))J_{i,-1/r}. \tag{F.2}$$

It follows the Kac determinant has a zero when $\vec{\alpha}_i \cdot \vec{J}_0 = Q(1 - 1/r)$ for a non-invariant simple root $\vec{\alpha}_i$, or in other words $\vec{\gamma}_i \cdot \vec{a} + Q = 0$ for a long root $\vec{\gamma}_i$ and the shifted zero mode $\vec{a}$. From Weyl invariance, the Kac determinant has a factor $\vec{\gamma} \cdot \vec{a} + Q$ for each long root γ. The Kac determinant has degree $2\sum_i(w_i - 1)$ where the sum is over the degrees w_i of $W(\Gamma)$-generators not invariant under $\mathbb{Z}_r$. This equals the number of long roots of G. We thus conclude that

$$(\text{Kac determinant at level } 1/r) \propto \prod_{\vec{\gamma} \in \Delta_l} (\vec{\gamma} \cdot \vec{a} + Q). \tag{F.3}$$

Appendix G

Construction of the W-algebra of type E_6

We provide some more details on the construction of the $W(E_6)$-algebra. We first construct the two lowest generators $W^{(2)}$ and $W^{(5)}$ utilizing $W(A_5)$ subalgebra. The higher generators $W^{(6,8,9,12)}$ we then obtain from suitable OPEs of $W^{(5)}$.

Let $J^{1,\dots,6} = i\partial\varphi^{1,\dots,6}$ be six orthonormal free bosons. Take the $A_2 \times A_2 \times A_1$ subalgebra of E_6, corresponding to the nodes of the Dynkin diagram except the central node. We can then choose the bosons in such a way that $J^{1,2}$, $J^{3,4}$ span the Cartan of the two A_2, and J_6 spans the Cartan of the A_1 respectively. Note that the $\mathbb{Z}_2$ outer automorphism exchanges the two A_2 subalgebras.

Consider now the A_5 subalgebra of E_6. It contains $A_2 \times A_2$, but not the A_1 described above. Construct the generators $U_{2,\dots,5}(z)$ of the $W(A_5)$ algebra via (4.56), and obtain the $\mathbb{Z}_2$ eigenstates (4.61). We introduce the boson

$$K(z) = (\sqrt{3}J^5 + J^6)/2\sqrt{2}\,, \tag{G.1}$$

which is perpendicular to the Cartan of A_5. The most general ansatz for $W^{(2)}$ and $W^{(5)}$ compatible with $\mathbb{Z}_2$ is then

$$W^{(2)} = U^{(2)} + c_1K^2 + c_2\partial K, \tag{G.2}$$

$$W^{(5)} = \tilde{U}^{(5)} + c_3\tilde{U}^{(3)}U^{(2)} + c_4\tilde{U}^{(3)}K^2 + c_5\partial^2\tilde{U}^{(3)} + c_6\partial\tilde{U}^{(3)}K + c_7\tilde{U}^{(3)}\partial K\,. \tag{G.3}$$

The constants $c_{1,\dots,7}$ can be determined by the method explained in section 4.3.1.3: we expand (G.3) in the J^i, and require that J^6 only appears in the form of its energy momentum

tensor

$$T(z) = -(J^6)^2/2 + (Q/\sqrt{2})\partial J^6 \,. \tag{G.4}$$

This fixes the constants to be

$$c_1 = 1,\ c_2 = -11Q,\ c_3 = 0,\ c_4 = 1,\ c_5 = \frac{3}{2}Q^2,\ c_6 = -3Q,\ c_7 = 2Q. \tag{G.5}$$

The higher generators can now be obtained by repeatedly taking the OPEs of $W^{(5)}$. Let $[O_1O_2]_p(w)$ be the coefficient of $(z-w)^{-p}$ of the OPE $O_1(z)O_2(w)$. We let

$$W^{(6)} = (1+12Q^2)^{-1}[W^{(5)}W^{(5)}]_4, \qquad W^{(8)} = [W^{(5)}W^{(5)}]_2, \tag{G.6}$$

$$W^{(9)} = [W^{(5)}W^{(6)}]_2, \qquad W^{(12)} = [W^{(6)}W^{(8)}]_2. \tag{G.7}$$

These OPEs a priori are not guaranteed to generate independent fields, but could contain just products of lower lying fields, possibly with derivatives. Independence can be confirmed as follows. Let $P^{(d)}$ be the terms in $W^{(d)}$ which do not contain derivatives, regarded as polynomials in J^i. It then suffices to check that they are algebraically independent, for which we can compute the Jacobian $\det(\partial P^{(d)}/\partial J^i)$ and check that it is non-zero. Note that W-generators constructed in this manner automatically have definite parity under $\mathbb{Z}_2$.

Unfortunately these OPE calculations, when expressed in terms of six free bosons, are too much for a laptop computer of 2011. It is thus necessary to use the $A_2 \times A_2 \times A_1$ subalgebra to organize the computation. The generator of $W(A_1)$ is $T(z)$ in (G.4). Let $U^{(2,3)}$ be the generators of the first $W(A_2)$, and $\tilde{U}^{(2,3)}$ be those of the second, as defined in (4.56). Then, we can rewrite $W^{(2)}$ and $W^{(5)}$ determined in (G.3) in terms of $U^{(2,3)}$, $\tilde{U}^{(2,3)}$, J_5, and T. Now the OPEs in (G.7) can be and were performed using the known OPEs of $W(A_2) \times W(A_2) \times W(A_1)$ in about three hours on a 2 GHz machine with 3 Gbytes of memory; the independence of the resulting generators can be and was checked. The denominator $(1 + 12Q^2)$ for $W^{(6)}$ in (G.7) is inserted because the OPE turns out to be divisible by this factor.

Bibliography

[1] L. Hollands, C. A. Keller, and J. Song, "From SO/Sp Instantons to W-Algebra Blocks," *JHEP* **03** (2011) 053, arXiv:1012.4468 [hep-th].

[2] L. Hollands, C. A. Keller, and J. Song, "Towards a 4d/2d correspondence for Sicilian quivers," *JHEP* **10** (2011) 100, arXiv:1107.0973 [hep-th].

[3] C. A. Keller, N. Mekareeya, J. Song, and Y. Tachikawa, "The ABCDEFG of Instantons and W-algebras," *JHEP* **03** (2012) 045, arXiv:1111.5624 [hep-th].

[4] N. Seiberg, "The Power of holomorphy: Exact results in 4-D SUSY field theories," arXiv:hep-th/9408013 [hep-th].

[5] N. Seiberg, "The Power of duality: Exact results in 4-D SUSY field theory," *Int.J.Mod.Phys.* **A16** (2001) 4365–4376, arXiv:hep-th/9506077 [hep-th].

[6] L. Brink, J. H. Schwarz, and J. Scherk, "Supersymmetric Yang-Mills Theories," *Nucl.Phys.* **B121** (1977) 77.

[7] J. M. Maldacena, "The Large N limit of superconformal field theories and supergravity," *Adv.Theor.Math.Phys.* **2** (1998) 231–252, arXiv:hep-th/9711200 [hep-th].

[8] S. Gubser, I. R. Klebanov, and A. M. Polyakov, "Gauge theory correlators from noncritical string theory," *Phys.Lett.* **B428** (1998) 105–114, arXiv:hep-th/9802109 [hep-th].

[9] E. Witten, "Anti-de Sitter space and holography," *Adv.Theor.Math.Phys.* **2** (1998) 253–291, arXiv:hep-th/9802150 [hep-th].

[10] O. Aharony, S. S. Gubser, J. M. Maldacena, H. Ooguri, and Y. Oz, "Large N field theories, string theory and gravity," *Phys.Rept.* **323** (2000) 183–386, arXiv:hep-th/9905111 [hep-th].

[11] N. Beisert, C. Ahn, L. F. Alday, Z. Bajnok, J. M. Drummond, *et al.*, "Review of AdS/CFT Integrability: An Overview," *Lett.Math.Phys.* **99** (2012) 3–32, arXiv:1012.3982 [hep-th].

[12] H. Ooguri, Y. Ookouchi, and C.-S. Park, "Metastable Vacua in Perturbed Seiberg-Witten Theories," *Adv.Theor.Math.Phys.* **12** (2008) 405–427, arXiv:0704.3613 [hep-th].

[13] J. Marsano, H. Ooguri, Y. Ookouchi, and C.-S. Park, "Metastable Vacua in Perturbed Seiberg-Witten Theories. Part 2. Fayet-Iliopoulos Terms and Kahler Normal Coordinates," *Nucl.Phys.* **B798** (2008) 17–35, arXiv:0712.3305 [hep-th].

[14] H. Ooguri, Y. Ookouchi, C.-S. Park, and J. Song, "Current Correlators for General Gauge Mediation," *Nucl. Phys.* **B808** (2009) 121–136, arXiv:0806.4733 [hep-th].

[15] N. Seiberg and E. Witten, "Monopole Condensation, and Confinement in $\mathcal{N} = 2$ Supersymmetric Yang-Mills Theory," *Nucl. Phys.* **B426** (1994) 19–52, arXiv:hep-th/9407087.

[16] N. Seiberg and E. Witten, "Monopoles, Duality and Chiral Symmetry Breaking in $\mathcal{N} = 2$ Supersymmetric QCD," *Nucl. Phys.* **B431** (1994) 484–550, arXiv:hep-th/9408099.

[17] E. Witten, "Topological Quantum Field Theory," *Commun.Math.Phys.* **117** (1988) 353.

[18] E. Witten, "Monopoles and four manifolds," *Math. Res. Lett.* **1** (1994) 769–796, arXiv:hep-th/9411102.

[19] C. Vafa and E. Witten, "A Strong coupling test of S duality," *Nucl. Phys.* **B431** (1994) 3–77, arXiv:hep-th/9408074.

[20] H. Nakajima, "Instantons on ALE spaces, quiver varieties, and Kac-Moody algebras," *Duke Math. J.* **76** no. 2, (1994) 365–416.

[21] H. Nakajima, "Instantons and affine Lie algebras," in *eprint arXiv:alg-geom/9510003*, p. 10003. Nov., 1995.

[22] M. Aganagic, A. Klemm, M. Mariño, and C. Vafa, "The Topological Vertex," *Commun. Math. Phys.* **254** (2005) 425–478, arXiv:hep-th/0305132.

[23] N. A. Nekrasov and A. Okounkov, "Seiberg-Witten theory and random partitions," in *The unity of mathematics*, vol. 244 of *Progr. Math.*, pp. 525–596. Birkhäuser Boston, Boston, MA, 2006. arXiv:hep-th/0306238.

[24] R. Dijkgraaf, L. Hollands, P. Sulkowski, and C. Vafa, "Supersymmetric Gauge Theories, Intersecting Branes and Free Fermions," *JHEP* **02** (2008) 106, arXiv:0709.4446 [hep-th].

[25] D. Gaiotto, "N=2 dualities," arXiv:0904.2715 [hep-th].

[26] L. F. Alday, D. Gaiotto, and Y. Tachikawa, "Liouville Correlation Functions from Four-Dimensional Gauge Theories," *Lett. Math. Phys.* **91** (2010) 167–197, arXiv:0906.3219 [hep-th].

[27] V. Pestun, "Localization of gauge theory on a four-sphere and supersymmetric Wilson loops," arXiv:0712.2824 [hep-th].

[28] N. A. Nekrasov, "Seiberg-Witten Prepotential From Instanton Counting," *Adv. Theor. Math. Phys.* **7** (2004) 831–864, arXiv:hep-th/0206161.

[29] G. W. Moore, N. Nekrasov, and S. Shatashvili, "Integrating over Higgs branches," *Commun.Math.Phys.* **209** (2000) 97–121, arXiv:hep-th/9712241 [hep-th].

[30] A. S. Losev, A. Marshakov, and N. A. Nekrasov, "Small instantons, little strings and free fermions," arXiv:hep-th/0302191.

[31] D. Gaiotto, G. W. Moore, and A. Neitzke, "Wall-Crossing, Hitchin Systems, and the WKB Approximation," arXiv:0907.3987 [hep-th].

[32] N. Seiberg, "Supersymmetry and Nonperturbative beta Functions," *Phys. Lett.* **B206** (1988) 75.

[33] E. Witten, "Solutions of Four-Dimensional Field Theories via M-theory," *Nucl. Phys.* **B500** (1997) 3–42, arXiv:hep-th/9703166.

[34] N. Hitchin, "Stable bundles and integrable systems," *Duke Math. J.* **54** no. 1, (1987) 91–114.

[35] N. J. Evans, C. V. Johnson, and A. D. Shapere, "Orientifolds, Branes, and Duality of 4D Gauge Theories," *Nucl. Phys.* **B505** (1997) 251–271, arXiv:hep-th/9703210.

[36] K. Landsteiner, E. Lopez, and D. A. Lowe, "$\mathcal{N} = 2$ Supersymmetric Gauge Theories, Branes and Orientifolds," *Nucl. Phys.* **B507** (1997) 197–226, arXiv:hep-th/9705199.

[37] A. Brandhuber, J. Sonnenschein, S. Theisen, and S. Yankielowicz, "M-theory and Seiberg-Witten Curves: Orthogonal and Symplectic Groups," *Nucl. Phys.* **B504** (1997) 175–188, arXiv:hep-th/9705232.

[38] Y. Tachikawa, "Six-dimensional D_N theory and four-dimensional SO-USp quivers," *JHEP* **07** (2009) 067, arXiv:0905.4074 [hep-th].

[39] K. Hori, "Consistency condition for fivebrane in M-theory on R**5/Z(2) orbifold," *Nucl. Phys.* **B539** (1999) 35–78, arXiv:hep-th/9805141.

[40] E. J. Martinec and N. P. Warner, "Integrable Systems and Supersymmetric Gauge Theory," *Nucl. Phys.* **B459** (1996) 97–112, arXiv:hep-th/9509161.

[41] A. Klemm, W. Lerche, S. Yankielowicz, and S. Theisen, "Simple singularities and N=2 supersymmetric Yang-Mills theory," *Phys.Lett.* **B344** (1995) 169–175, arXiv:hep-th/9411048 [hep-th].

[42] A. Gorsky, I. Krichever, A. Marshakov, A. Mironov, and A. Morozov, "Integrability and Seiberg-Witten Exact Solution," *Phys. Lett.* **B355** (1995) 466–474, arXiv:hep-th/9505035.

[43] H. Itoyama and A. Morozov, "Integrability and Seiberg-Witten Theory: Curves and Periods," *Nucl. Phys.* **B477** (1996) 855–877, arXiv:hep-th/9511126.

[44] W. Lerche and N. P. Warner, "Exceptional SW Geometry from ALE Fibrations," *Phys. Lett.* **B423** (1998) 79–86, arXiv:hep-th/9608183.

[45] A. Klemm, W. Lerche, P. Mayr, C. Vafa, and N. P. Warner, "Self-Dual Strings and $\mathcal{N} = 2$ Supersymmetric Field Theory," *Nucl. Phys.* **B477** (1996) 746–766, arXiv:hep-th/9604034.

[46] T. J. Hollowood, "Strong Coupling $\mathcal{N} = 2$ Gauge Theory with Arbitrary Gauge Group," *Adv. Theor. Math. Phys.* **2** (1998) 335–355, arXiv:hep-th/9710073.

[47] D.-E. Diaconescu, R. Donagi, and T. Pantev, "Intermediate Jacobians and ADE Hitchin Systems," arXiv:hep-th/0607159.

[48] P. Slodowy, *Simple Singularities and Simple Algebraic Groups*, vol. 815 of *Lecture Notes in Mathematics*. Springer, 1980.

[49] J. Fuchs, B. Schellekens, and C. Schweigert, "From Dynkin Diagram Symmetries to Fixed Point Structures," *Commun. Math. Phys.* **180** (1996) 39–98, arXiv:hep-th/9506135.

[50] D. Finnell and P. Pouliot, "Instanton Calculations Versus Exact Results in Four-Dimensional SUSY Gauge Theories," *Nucl. Phys.* **B453** (1995) 225–239, arXiv:hep-th/9503115.

[51] K. Ito and N. Sasakura, "One-Instanton Calculations in $\mathcal{N} = 2$ Supersymmetric $SU(N_C)$ Yang-Mills Theory," *Phys. Lett.* **B382** (1996) 95–103, arXiv:hep-th/9602073.

[52] N. Dorey, T. J. Hollowood, V. V. Khoze, and M. P. Mattis, "The Calculus of Many Instantons," *Phys. Rept.* **371** (2002) 231–459, arXiv:hep-th/0206063.

[53] N. Nekrasov and S. Shadchin, "ABCD of Instantons," *Commun. Math. Phys.* **252** (2004) 359–391, arXiv:hep-th/0404225.

[54] M. F. Atiyah, N. J. Hitchin, V. G. Drinfeld, and Y. I. Manin, "Construction of instantons," *Phys. Lett.* **A65** (1978) 185–187.

[55] H. Nakajima and K. Yoshioka, "Instanton counting on blowup. I. 4-dimensional pure gauge theory," *Invent. Math.* **162** no. 2, (2005) 313–355, arXiv:math/0306198.

[56] H. Nakajima, *Lectures on Hilbert schemes of points on surfaces*, vol. 18 of *University Lecture Series*. American Mathematical Society, Providence, RI, 1999.

[57] L. Gottsche, H. Nakajima, and K. Yoshioka, "Donaldson = Seiberg-Witten from Mochizuki's formula and instanton counting," *Publ. Res. Inst. Math. Sci. Kyoto* **47** (2011) 307–359, arXiv:1001.5024 [math.DG].

[58] T. Okuda and V. Pestun, "On the instantons and the hypermultiplet mass of N=2* super Yang-Mills on S^4," *JHEP* **03** (2012) 017, arXiv:1004.1222 [hep-th].

[59] D. Krefl and J. Walcher, "Shift versus Extension in Refined Partition Functions," arXiv:1010.2635 [hep-th].

[60] A. Braverman and P. Etingof, "Instanton Counting via Affine Lie Algebras. II: from Whittaker Vectors to the Seiberg-Witten Prepotential," in *Studies in Lie Theory: dedicated to A. Joseph on his 60th birthday*, J. Bernstein, V. Hinich, and A. Melnikov, eds. Birkhäuser, 2006. arXiv:math/0409441.

[61] E. Corrigan and P. Goddard, "Construction of Instanton and Monopole Solutions and Reciprocity," *Ann. Phys.* **154** (1984) 253.

[62] J. Bryan and M. Sanders, "Instantons on S^4 and $\overline{\mathbf{C}\mathrm{P}}^2$, rank stabilization, and Bott periodicity," *Topology* **39** no. 2, (2000) 331–352.

[63] A. Losev, N. Nekrasov, and S. L. Shatashvili, "Issues in topological gauge theory," *Nucl. Phys.* **B534** (1998) 549–611, arXiv:hep-th/9711108.

[64] G. W. Moore, N. Nekrasov, and S. Shatashvili, "D-particle bound states and generalized instantons," *Commun. Math. Phys.* **209** (2000) 77–95, arXiv:hep-th/9803265.

[65] H. Nakajima and K. Yoshioka, "Lectures on instanton counting," in *Algebraic structures and moduli spaces*, vol. 38 of *CRM Proc. Lecture Notes*, pp. 31–101. Amer. Math. Soc., Providence, RI, 2004. arXiv:math/0311058.

[66] A. Braverman, "Instanton Counting via Affine Lie Algebras I: Equivariant J-Functions of (Affine) Flag Manifolds and Whittaker Vectors," in *Workshop on algebraic structures and moduli spaces: CRM Workshop*, J. Hurturbise and E. Markman, eds. AMS, July, 2003. arXiv:math/0401409.

[67] A. Licata and A. Savage, "Vertex operators and the geometry of moduli spaces of framed torsion-free sheaves," *Selecta Math. (N.S.)* **16** no. 2, (2010) 201–240.

[68] S. Shadchin, "On certain aspects of string theory / gauge theory correspondence," arXiv:hep-th/0502180.

[69] J. Martens, "Equivariant volumes of non-compact quotients and instanton counting," *Comm. Math. Phys.* **281** no. 3, (2008) 827–857.

[70] M. Marino and N. Wyllard, "A note on instanton counting for N = 2 gauge theories with classical gauge groups," *JHEP* **05** (2004) 021, arXiv:hep-th/0404125.

[71] F. Fucito, J. F. Morales, and R. Poghossian, "Instantons on quivers and orientifolds," *JHEP* **10** (2004) 037, arXiv:hep-th/0408090.

[72] N. Nekrasov and A. S. Schwarz, "Instantons on noncommutative R**4 and (2,0) superconformal six dimensional theory," *Commun. Math. Phys.* **198** (1998) 689–703, arXiv:hep-th/9802068.

[73] F. C. Kirwan, "Partial desingularisations of quotients of nonsingular varieties and their Betti numbers," *Ann. of Math. (2)* **122** no. 1, (1985) 41–85.

[74] S. Cordes, G. W. Moore, and S. Ramgoolam, "Lectures on 2-d Yang-Mills theory, equivariant cohomology and topological field theories," *Nucl. Phys. Proc. Suppl.* **41** (1995) 184–244, arXiv:hep-th/9411210.

[75] U. Bruzzo, F. Fucito, J. F. Morales, and A. Tanzini, "Multi-instanton calculus and equivariant cohomology," *JHEP* **05** (2003) 054, arXiv:hep-th/0211108.

[76] C. W. Bernard, N. H. Christ, A. H. Guth, and E. J. Weinberg, "Pseudoparticle Parameters for Arbitrary Gauge Groups," *Phys. Rev.* **D16** (1977) 2967.

[77] A. I. Vainshtein, V. I. Zakharov, V. A. Novikov, and M. A. Shifman, "ABC of Instantons," *Sov. Phys. Usp.* **25** (1982) 195.

[78] P. B. Kronheimer, "Instantons and the geometry of the nilpotent variety," *J. Diff. Geom.* **32** (1990) 473.

[79] R. Brylinski, "Instantons and Kähler geometry of nilpotent orbits," in *Representation theories and algebraic geometry*, vol. 514 of *NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci.*, pp. 85–125. Kluwer, 1998. `math.SG/9811032`.

[80] P. Kobak and A. Swann, "The hyperkähler geometry associated to Wolf spaces," *Boll. Unione Mat. Ital. Serie 8, Sez. B Artic. Ric. Mat.* **4** (2001) 587, `math.DG/0001025`.

[81] E. B. Vinberg and V. L. Popov, "On a class of quasihomogeneous affine varieties," *Math. USSR-Izv.* **6** (1972) 743.

[82] quoted in Chap. III of D. Garfinkle, *A new construction of the Joseph ideal*. PhD thesis, MIT, 1982.

[83] S. Benvenuti, A. Hanany, and N. Mekareeya, "The Hilbert Series of the One Instanton Moduli Space," *JHEP* **06** (2010) 100, `arXiv:1005.3026 [hep-th]`.

[84] K. Ito and N. Sasakura, "Exact and Microscopic One-Instanton Calculations in $\mathcal{N} = 2$ Supersymmetric Yang-Mills Theories," *Nucl. Phys.* **B484** (1997) 141–166, `arXiv:hep-th/9608054`.

[85] M. Aganagic, R. Dijkgraaf, A. Klemm, M. Marino, and C. Vafa, "Topological strings and integrable hierarchies," *Commun. Math. Phys.* **261** (2006) 451–516, `arXiv:hep-th/0312085`.

[86] B. Eynard and N. Orantin, "Invariants of algebraic curves and topological expansion," *Commun. Number Theory Phys.* **1** no. 2, (2007) 347–452.

[87] M. Billo, L. Gallot, A. Lerda, and I. Pesando, "F-theoretic vs microscopic description of a conformal N=2 SYM theory," *JHEP* **11** (2010) 041, `arXiv:1008.5240 [hep-th]`.

[88] M. Billo, M. Frau, F. Fucito, A. Lerda, J. F. Morales, and R. Poghossian, "Stringy instanton corrections to N=2 gauge couplings," *JHEP* **05** (2010) 107, `arXiv:1002.4322 [hep-th]`.

[89] M. Billo, L. Ferro, M. Frau, L. Gallot, A. Lerda, and I. Pesando, "Exotic instanton counting and heterotic/type I-prime duality," *JHEP* **0907** (2009) 092, `arXiv:0905.4586 [hep-th]`.

[90] P. C. Argyres and A. D. Shapere, "The Vacuum Structure of N=2 SuperQCD with Classical Gauge Groups," *Nucl. Phys.* **B461** (1996) 437–459, arXiv:hep-th/9509175.

[91] N. Wyllard, "A_{N-1} Conformal Toda Field Theory Correlation Functions from Conformal $\mathcal{N} = 2$ $SU(N)$ Quiver Gauge Theories," *JHEP* **11** (2009) 002, arXiv:0907.2189 [hep-th].

[92] T. W. Grimm, A. Klemm, M. Marino, and M. Weiss, "Direct integration of the topological string," *JHEP* **08** (2007) 058, arXiv:hep-th/0702187.

[93] E. Witten, "An SU(2) anomaly," *Phys. Lett.* **B117** (1982) 324–328.

[94] P. Bouwknegt and K. Schoutens, "W symmetry in conformal field theory," *Phys. Rept.* **223** (1993) 183–276, arXiv:hep-th/9210010.

[95] P. H. Ginsparg, "Applied Conformal Field Theory," arXiv:hep-th/9108028 [hep-th].

[96] P. Di Francesco, P. Mathieu, and D. Senechal, *Conformal field theory*. New York, USA: Springer, 1997.

[97] J. Polchinski, "Scale and conformal invariance in quantum field theory," *Nucl. Phys.* **B303** (1988) 226.

[98] B. Feigin and D. Fuchs, "Verma Modules over the Virasoro Algebra," *PREPRINT - FEIGIN, B.L. (REC.OCT.88)* (1988) 16.

[99] A. A. Belavin, A. M. Polyakov, and A. B. Zamolodchikov, "Infinite conformal symmetry in two-dimensional quantum field theory," *Nucl. Phys.* **B241** (1984) 333–380.

[100] A. B. Zamolodchikov and A. B. Zamolodchikov, "Structure constants and conformal bootstrap in Liouville field theory," *Nucl. Phys.* **B477** (1996) 577–605, arXiv:hep-th/9506136.

[101] A. B. Zamolodchikov, "Infinite Additional Symmetries in Two-Dimensional Conformal Quantum Field Theory," *Theor. Math. Phys.* **65** (1985) 1205–1213.

[102] R. Blumenhagen, M. Flohr, A. Kliem, W. Nahm, A. Recknagel, and R. Varnhagen, "W algebras with two and three generators," *Nucl. Phys.* **B361** (1991) 255–289.

[103] H. G. Kausch and G. M. T. Watts, "A Study of W algebras using Jacobi identities," *Nucl. Phys.* **B354** (1991) 740–768.

[104] M. Bershadsky and H. Ooguri, "Hidden $SL(N)$ Symmetry in Conformal Field Theories," *Commun. Math. Phys.* **126** (1989) 49.

[105] B. Feigin and E. Frenkel, "Quantization of the Drinfeld-Sokolov Reduction," *Phys. Lett.* **B246** (1990) 75–81.

[106] V. G. Drinfeld and V. V. Sokolov, "Lie algebras and equations of Korteweg-de Vries type," *J. Sov. Math.* **30** (1984) 1975–2036.

[107] K. Thielemans, "A Mathematica Package for Computing Operator Product Expansions," *Int. J. Mod. Phys.* **C2** (1991) 787–798.

[108] V. A. Fateev and S. L. Lukyanov, "The Models of Two-Dimensional Conformal Quantum Field Theory with $\mathbb{Z}_N$ Symmetry," *Int. J. Mod. Phys.* **A3** (1988) 507.

[109] S. L. Lukyanov and V. A. Fateev, "Conformally invariant models of two-dimensional QFT with $\mathbb{Z}_N$ symmetry," *Sov. Phys. JETP* **67** (1988) 447.

[110] S. L. Lukyanov and V. A. Fateev, "Exactly solvable models of conformal quantum theory associated with the simple Lie algebra D_N," *Sov. J. Nucl. Phys.* **49** (1989) 925–932. [Yad.Fiz.49:1491-1504,1989].

[111] S. L. Lukyanov and V. A. Fateev, "Physics reviews: Additional symmetries and exactly soluble models in two-dimensional conformal field theory,". Chur, Switzerland: Harwood (1990) 117 p. (Soviet Scientific Reviews A, Physics: 15.2).

[112] K. Ito and S. Terashima, "Free Field Realization of WBC_N and WG_2 Algebras," *Phys. Lett.* **B354** (1995) 220–231, `arXiv:hep-th/9503165`.

[113] A. Klemm, M. Mariño, and S. Theisen, "Gravitational Corrections in Supersymmetric Gauge Theory and Matrix Models," *JHEP* **03** (2003) 051, `arXiv:hep-th/0211216`.

[114] R. Dijkgraaf, A. Sinkovics, and M. Temurhan, "Matrix Models and Gravitational Corrections," *Adv. Theor. Math. Phys.* **7** (2004) 1155–1176, arXiv:hep-th/0211241.

[115] H. Fuji and S. Mizoguchi, "Gravitational Corrections for Supersymmetric Gauge Theories with Flavors via Matrix Models," *Nucl. Phys.* **B698** (2004) 53–91, arXiv:hep-th/0405128.

[116] E. Witten, "On S-Duality in Abelian Gauge Theory," *Selecta Math.* **1** (1995) 383, arXiv:hep-th/9505186.

[117] G. W. Moore and E. Witten, "Integration over the u-plane in Donaldson Theory," *Adv. Theor. Math. Phys.* **1** (1998) 298–387, arXiv:hep-th/9709193.

[118] L. F. Alday and Y. Tachikawa, "Affine $SL(2)$ Conformal Blocks from 4D Gauge Theories," *Lett. Math. Phys.* **94** (2010) 87–114, arXiv:1005.4469 [hep-th].

[119] G. Bonelli and A. Tanzini, "Hitchin systems, N=2 gauge theories and W-gravity," *Phys. Lett.* **B691** (2010) 111–115, arXiv:0909.4031 [hep-th].

[120] D. Gaiotto, "Asymptotically Free $\mathcal{N} = 2$ Theories and Irregular Conformal Blocks," arXiv:0908.0307 [hep-th].

[121] A. Marshakov, A. Mironov, and A. Morozov, "On Non-Conformal Limit of the AGT Relations," *Phys. Lett.* **B682** (2009) 125–129, arXiv:0909.2052 [hep-th].

[122] M. Taki, "On AGT Conjecture for Pure Super Yang-Mills and W- algebra," *JHEP* **05** (2011) 038, arXiv:0912.4789 [hep-th].

[123] D. Maulik and A. Okounkov , in preparation.

[124] L. J. Dixon, D. Friedan, E. J. Martinec, and S. H. Shenker, "The Conformal Field Theory of Orbifolds," *Nucl. Phys.* **B282** (1987) 13–73.

[125] J. L. Cardy, O. A. Castro-Alvaredo, and B. Doyon, "Form factors of branch-point twist fields in quantum integrable models and entanglement entropy," *J. Stat. Phys.* **130** (2008) 129–168, arXiv:0706.3384 [hep-th].

[126] M. Gaberdiel, "A General transformation formula for conformal fields," *Phys. Lett.* **B325** (1994) 366–370, arXiv:hep-th/9401166.

[127] V. A. Alba, V. A. Fateev, A. V. Litvinov, and G. M. Tarnopolskiy, "On combinatorial expansion of the conformal blocks arising from AGT conjecture," *Lett. Math. Phys.* **98** (2011) 33–64, arXiv:1012.1312 [hep-th].

[128] M. R. Gaberdiel and A. Neitzke, "Rationality, quasirationality and finite W-algebras," *Commun. Math. Phys.* **238** (2003) 305–331, arXiv:hep-th/0009235.

[129] S. Cecotti and C. Vafa, "Classification of complete N=2 supersymmetric theories in 4 dimensions," arXiv:1103.5832 [hep-th].

[130] D. Gaiotto and S. S. Razamat, "Exceptional Indices," arXiv:1203.5517 [hep-th].

[131] C. A. Keller and J. Song, "Counting Exceptional Instantons," arXiv:1205.4722 [hep-th].

[132] H. Nakajima and K. Yoshioka, "Instanton counting on blowup. II. *K*-theoretic partition function," *Transform. Groups* **10** no. 3-4, (2005) 489–519, arXiv:math/0505553 [math-ag].

[133] R. Dijkgraaf and C. Vafa, "Toda Theories, Matrix Models, Topological Strings, and $\mathcal{N} = 2$ Gauge Systems," arXiv:0909.2453 [hep-th].

[134] A. Morozov, "Challenges of beta-deformation," arXiv:1201.4595 [hep-th].

[135] A. Iqbal, C. Kozcaz, and C. Vafa, "The refined topological vertex," *JHEP* **10** (2009) 069, arXiv:hep-th/0701156.

[136] H. Awata and Y. Yamada, "Five-dimensional AGT Conjecture and the Deformed Virasoro Algebra," *JHEP* **01** (2010) 125, arXiv:0910.4431 [hep-th].

[137] H. Awata and Y. Yamada, "Five-Dimensional AGT Relation and the Deformed β-Ensemble," *Prog. Theor. Phys.* **124** (2010) 227–262, arXiv:1004.5122 [hep-th].

[138] V. Fateev and A. Litvinov, "On AGT conjecture," *JHEP* **1002** (2010) 014, arXiv:0912.0504 [hep-th].

[139] L. Hadasz, Z. Jaskolski, and P. Suchanek, "Proving the AGT relation for $N_f = 0, 1, 2$ antifundamentals," *JHEP* **1006** (2010) 046, arXiv:1004.1841 [hep-th].

[140] L. F. Alday, D. Gaiotto, S. Gukov, Y. Tachikawa, and H. Verlinde, "Loop and surface operators in N=2 gauge theory and Liouville modular geometry," *JHEP* **01** (2010) 113, arXiv:0909.0945 [hep-th].

[141] N. Drukker, J. Gomis, T. Okuda, and J. Teschner, "Gauge Theory Loop Operators and Liouville Theory," *JHEP* **1002** (2010) 057, arXiv:0909.1105 [hep-th].

[142] J. Gomis, T. Okuda, and V. Pestun, "Exact Results for 't Hooft Loops in Gauge Theories on S^4," arXiv:1105.2568 [hep-th].

[143] C. Kozcaz, S. Pasquetti, and N. Wyllard, "A & B model approaches to surface operators and Toda theories," *JHEP* **08** (2010) 042, arXiv:1004.2025 [hep-th].

[144] T. Dimofte, S. Gukov, and L. Hollands, "Vortex Counting and Lagrangian 3-manifolds," *Lett. Math. Phys.* **98** (2011) 225–287, arXiv:1006.0977 [hep-th].

[145] N. Drukker, D. Gaiotto, and J. Gomis, "The Virtue of Defects in 4D Gauge Theories and 2D CFTs," *JHEP* **1106** (2011) 025, arXiv:1003.1112 [hep-th].

[146] K. Hosomichi, S. Lee, and J. Park, "AGT on the S-duality Wall," *JHEP* **1012** (2010) 079, arXiv:1009.0340 [hep-th].

[147] C. Kozcaz, S. Pasquetti, F. Passerini, and N. Wyllard, "Affine sl(N) conformal blocks from N=2 SU(N) gauge theories," *JHEP* **01** (2011) 045, arXiv:1008.1412 [hep-th].

[148] N. Wyllard, "W-algebras and surface operators in N=2 gauge theories," *J.Phys.A* **A44** (2011) 155401, arXiv:1011.0289 [hep-th].

[149] N. Wyllard, "Instanton partition functions in N=2 SU(N) gauge theories with a general surface operator, and their W-algebra duals," *JHEP* **02** (2011) 114, arXiv:1012.1355 [hep-th].

[150] H. Kanno and Y. Tachikawa, "Instanton counting with a surface operator and the chain-saw quiver," *JHEP* **1106** (2011) 119, arXiv:1105.0357 [hep-th].

[151] A. Gadde, E. Pomoni, L. Rastelli, and S. S. Razamat, "S-duality and 2d Topological QFT," *JHEP* **03** (2010) 032, arXiv:0910.2225 [hep-th].

[152] A. Gadde, L. Rastelli, S. S. Razamat, and W. Yan, "The Superconformal Index of the E_6 SCFT," *JHEP* **08** (2010) 107, arXiv:1003.4244 [hep-th].

[153] A. Gadde, L. Rastelli, S. S. Razamat, and W. Yan, "Gauge Theories and Macdonald Polynomials," arXiv:1110.3740 [hep-th].

[154] A. Gadde, L. Rastelli, S. S. Razamat, and W. Yan, "The 4d Superconformal Index from q-deformed 2d Yang- Mills," *Phys. Rev. Lett.* **106** (2011) 241602, arXiv:1104.3850 [hep-th].

[155] T. Nishioka and Y. Tachikawa, "Para-Liouville/Toda central charges from M5-branes," *Phys. Rev.* **D84** (2011) 046009, arXiv:1106.1172 [hep-th].

[156] V. Belavin and B. Feigin, "Super Liouville conformal blocks from N=2 SU(2) quiver gauge theories," *JHEP* **07** (2011) 079, arXiv:1105.5800 [hep-th].

[157] G. Bonelli, K. Maruyoshi, and A. Tanzini, "Gauge Theories on ALE Space and Super Liouville Correlation Functions," arXiv:1107.4609 [hep-th].

[158] O. Schiffmann and E. Vasserot, "Cherednik algebras, W algebras and the equivariant cohomology of the moduli space of instantons on A^2," arXiv:1202.2756 [math.QA].

[159] N. Hama, K. Hosomichi, and S. Lee, "SUSY Gauge Theories on Squashed Three-Spheres," *JHEP* **05** (2011) 014, arXiv:1102.4716 [hep-th].

[160] T. Dimofte and S. Gukov, "Chern-Simons Theory and S-duality," arXiv:1106.4550 [hep-th].

[161] T. Dimofte, D. Gaiotto, and S. Gukov, "Gauge Theories Labelled by Three-Manifolds," arXiv:1108.4389 [hep-th].

[162] S. Cecotti, C. Cordova, and C. Vafa, "Braids, Walls, and Mirrors," arXiv:1110.2115 [hep-th].

[163] Y. Terashima and M. Yamazaki, "SL(2,R) Chern-Simons, Liouville, and Gauge Theory on Duality Walls," *JHEP* **08** (2011) 135, arXiv:1103.5748 [hep-th].

[164] T. Dimofte, D. Gaiotto, and S. Gukov, "3-Manifolds and 3d Indices," arXiv:1112.5179 [hep-th].

[165] V. I. Arnol'd, "Critical points of functions on a manifold with boundary, the simple Lie groups B_k, C_k, F_4 and singularities of evolutes," *Uspekhi Mat. Nauk* **33** no. 5(203), (1978) 91.

[166] P. C. Argyres, M. R. Plesser, and A. D. Shapere, "The Coulomb phase of N=2 supersymmetric QCD," *Phys. Rev. Lett.* **75** (1995) 1699–1702, arXiv:hep-th/9505100.

[167] S. Mizoguchi, "Determinant Formula and Unitarity for the W_3 Algebra," *Phys. Lett.* **B222** (1989) 226.

[168] V. G. Kac and M. Wakimoto, "Quantum reduction in the twisted case," arXiv:math-ph/0404049.