9-2012

AdS/CFT in String Theory and M-Theory

Daniel R. Gulotta '03
Illinois Mathematics and Science Academy, dgulotta@alum.mit.edu

AdS/CFT in String Theory and M-Theory

Daniel R. Gulotta

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

Recommended for Acceptance
by the Department of
Physics
Adviser: Christopher P. Herzog

September 2012

Abstract

The AdS/CFT correspondence is a powerful tool that can help shed light on the relationship between geometry and field theory.

The first part of this thesis will focus on the construction of theories dual to Type IIB string theory on $AdS_5 \times Y^5$, where Y^5 is a toric Sasaki-Einstein manifold. This thesis will introduce a consistency condition called "proper ordering" and demonstrate that it is equivalent to several other previously known consistency conditions. It will then give an efficient algorithm that produces a consistent field theory for any toric Sasaki-Einstein Y^5.

The second part of this thesis will examine the large-N limit of the Kapustin-Willett-Yaakov matrix model. This model computes the S^3 partition function for a CFT dual to M-theory on $AdS_4 \times Y^7$. One of the main results will be a formula that relates the distribution of eigenvalues in the matrix model to the distribution of holomorphic operators on the cone over Y^7. A variety of examples are given to support this formula.

Acknowledgements

I would like to thank my advisor, Chris Herzog. He was always available to provide helpful advice and discuss ideas, and without him this thesis would not be possible. I would also like to thank Steve Gubser and Igor Klebanov for all that they have taught me over the past five years, as well as Amihay Hanany for welcoming me into his research group when I was an undergraduate at MIT.

I would also like to thank my other collaborators JP Ang, Matthias Kaminski, Tatsuma Nishioka, Silviu Pufu, and Fabio Rocha. I learned a great deal from them, and I am inspired by their enthusiasm and dedication.

I would like to thank my friends, especially Yaim Cooper, Randy Li, Shrenik Shah, Stephanie Wu, and the GC coffee house gamers for making the past five years more enjoyable.

Finally, I would like to thank my parents for their tremendous support and guidance over the years.

Contents

Contents

Chapter 1

Introduction

The figures in this chapter originally appeared in refs. [1, 2], written in collaboration with Christopher Herzog and Silviu Pufu.

1.1 AdS/CFT in string theory and M-theory

The AdS/CFT correspondence [3–5] is a conjectured duality that can help us to understand the connections between geometry and field theory. It is an example of the holographic principle, which proposes that in a quantum theory of gravity, all of the information about a particular region is encoded in the boundary of that region. The original motivation for the holographic principle is the observation that the entropy of a black hole is proportional to its surface area rather than its volume. AdS/CFT uses string theory to give specific examples of a gravity theory in an $n+1$-dimensional negatively curved spacetime, and a field theory on the n-dimensional boundary of the spacetime that is believed to be equivalent to the gravity theory. In addition to string theory, AdS/CFT has found some applications in the physics of superconductors and heavy ion collisions.

The AdS/CFT correspondence is so named because the original examples of the duality were between n-dimensional conformal field theories (CFT) and $n+1$-dimensional gravity theories on anti de Sitter space (AdS). A conformal field theory is a field theory that is invariant under the conformal group, i. e. the group of symmetries of n-dimensional spacetime that preserve angles but not necessarily lengths. Anti de sitter space is a homogeneous,

isotropic negatively curved spacetime. The group of isometries of $(n+1)$-dimensional AdS space (AdS_{n+1}) is the same as the group of conformal transformations of n-dimensional Minkowski space.

The most commonly studied objects in string theory are strings and D-branes. Excitations of a closed loop of string give rise to gravity, among other fields. Strings ending on a D-brane give rise to gauge fields on the D-brane. Consider Type IIB string theory on 10-dimensional Minkowski space. Suppose that we add a stack of N D3-branes. The field theory on the branes will be an $SU(N)$ gauge theory with $\mathcal{N} = 4$ supersymmetry. This theory is conformal. The D3-branes have mass and charge, so they will warp the surrounding spacetime. For N large, the geometry of spacetime approaches $AdS_5 \times S^5$, where S^5 is the five-sphere. Additionally, when N is large, quantum effects in the $AdS_5 \times S_5$ theory become small, allowing us to treat the theory classically. This is the original example of AdS/CFT.

Chapter 2 of this thesis will explore some related constructions. Rather than starting with 10-dimensional Minkowski space, we will start with $\mathbb{R}^{3,1} \times CY^3$, where CY^3 is a Calabi-Yau manifold with a singularity. When we place N D3-branes at the singularity, the geometry becomes $AdS_5 \times Y^5$, where Y^5 is a Sasaki-Einstein manifold. (That Y^5 is Sasaki-Einstein essentially just means that $AdS_5 \times Y^5$ preserves some supersymmetry.)

Chapters 3-4 will focus on a similar construction in M-theory. We start with M-theory on $\mathbb{R}^{2,1} \times CY^4$, and we can place N M2-branes at a singularity of CY^4. For N large, the geometry becomes $AdS_4 \times Y^7$, where Y^7 is Sasaki-Einstein.

Most of the CFTs that we study will be quiver gauge theories. A quiver gauge theory is a 3– or 4–dimensional supersymmetric field theory consisting of gauge fields and matter fields in fundamental, antifundamental, or bifundamental representations of the gauge groups. Many configurations in string theory and M-theory have low energy limits that are described by quiver gauge theories. In this thesis, we will only consider quivers with at least four supercharges; this ensures the existence of a $U(1)_R$ symmetry. We will also avoid fundamental and antifundamental representations in four dimensions; these would correspond to adding D7-branes to the Type IIB constructions. A Chern-Simons coupling for the gauge fields is possible in three dimensions, but not in four dimensions. Figure 1.1 gives an example of a quiver gauge theory.

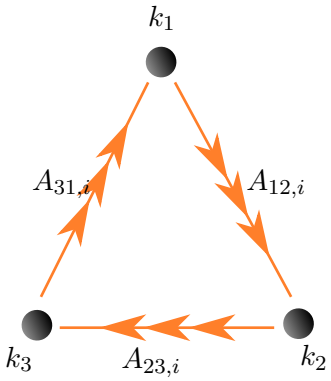


Figure 1.1: This quiver has three unitary gauge fields B_1, B_2, B_3 (depicted by black dots), and nine matter fields $A_{31,i}$, $A_{12,i}$, $A_{23,i}$ for $i \in \{1, 2, 3\}$ (depicted by arrows). The field $A_{mn,i}$ transforms in the fundamental of the gauge group B_n and the antifundamental of the gauge group B_m. Although it is not obvious from the diagram, there is a superpotential $\epsilon^{ijk} A_{31,i} A_{12,j} A_{23,k}$. In three dimensions, we can add Chern-Simons terms $\frac{k_i}{4\pi} B_i \wedge dB_i$ to the Lagrangian.

1.2 Dimer models and D3-branes

Chapter 2 will investigate the problem of constructing a quiver gauge theory dual to Type IIB string theory on $AdS_5 \times Y^5$ for a given Y^5. It appears to be prohibitively difficult to perform the construction for an arbitary Sasaki-Einstein manifold Y^5, so we will focus on the case where Y^5 is toric. A $2n-1$ dimensional Sasakian manifold is said to be toric if it possesses a $U(1)^n$ isometry group. This is the largest possible abelian isometry group for a Sasakian manifold. When Y^5 is toric, we can use a diagram called a dimer model [6–12] or brane tiling to describe the quiver gauge theory.

Not all dimer models describe Type IIB string theory on $AdS_5 \times Y^5$. If a dimer defines a field theory that does not describe any Y^5, then we call it inconsistent. There have been several previously proposed criteria for determining that a dimer is inconsistent. If a field theory is to describe some Y^5, then for $N = 1$, the Higgs branch of the moduli space should be isomorphic to the cone over Y^5. Once we have found the Y^5 that might be dual to the CFT, there are some tests that we can perform. AdS/CFT tells us that the number of gauge groups in the field theory should be equal to the number of even-dimensional cycles of Y^5 [13]. So we can reject any dimer that has the wrong number of gauge groups. Some other indications of inconsistency are the existence of operators with R-charge less than $\frac{2}{3}$ [10, 14, 15], cubic anomalies in the CFT not matching Chern-Simons coefficients in

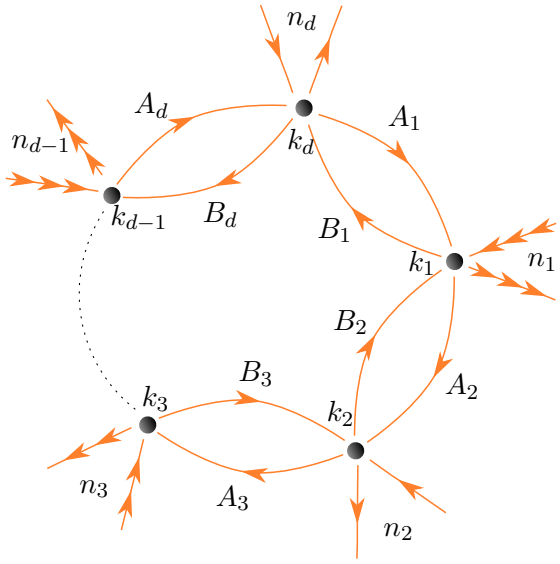


Figure 1.2: A necklace quiver gauge theory. Necklace quivers will be studied in chapters 3–4. The gauge sector consists of d $U(N)$ gauge groups with Chern-Simons coefficients k_a. The matter content consists of the bifundamental fields A_a and B_a, as well as n_a pairs of fundamentals and anti-fundamentals transforming under the ath gauge group.

AdS [5, 16], duplicate "corner perfect matchings" [10], and incorrect winding numbers of "zigzag paths" [10].

However, it was not clear whether there could be inconsistent dimers that these criteria fail to recognize. This thesis will propose a criterion called "proper ordering" and present evidence that a dimer is consistent iff it is properly ordered. It will show that proper ordering is equivalent to the criterion on the number of gauge groups, and is at least as strong as the criteria for R-charges, cubic anomalies, corner perfect matchings, and zigzag paths.

Given a toric Y^5, we would like to know how to construct a dimer model for Y^5. Previous attempts at an algorithm [10–12] suffered from being computationally inefficient, not being proven to produce a construction, and not being proven to produce only consistent constructions. This thesis will present an efficient algorithm that produces a properly ordered dimer for any toric Y^5.

1.3 Matrix models for M2-branes

Chapters 3–4 will use the matrix model of Kapustin, Willett, and Yaakov [17] to study various field theories dual to M-theory on $AdS_4 \times Y^7$. The matrix model uses the localization procedure of Pestun [18] to compute the partition function Z_{S^3} of the field theory on S^3. One reason for studying the matrix model is that it can help us understand nonperturbative dualities in three-dimensional field theories. In four dimensions, anomalies can be used to test nonperturbative dualities. In three dimensions, these anomalies do not exist. We would like to use the free energy $F = \ln |Z_{S^3}|$ as a replacement for the Weyl anomaly coefficient a. Along renormalization group flows, both a and F decrease monotonically [19–23]. In four dimensions, one can compute the exact infrared R-symmetry by maximizing a with respect to a set of trial R-charges [24]. In three dimensions, we can likewise find the infrared R-symmetry by maximizing F [25, 26].

Ref. [27] determined the large-N limit of the matrix model for certain quivers called "necklace quivers". These quivers have $\mathcal{N} = 3$ supersymmetry. Figure 1.2 shows an example of a necklace quiver. For such quivers, they find that the saddle point approximation successfully determines the free energy to leading order. This thesis will use the matrix model to test S-duality and Seiberg-like duality in necklace quivers. It will also analyze the large-N matrix model for several other theories, including some that have just $\mathcal{N} = 2$ supersymmetry.

This thesis will also propose a relationship between the eigenvalues of the large-N matrix model and the density of chiral operators in the field theory. In the theories that we will consider, the saddle point eigenvalues of the matrix model can be described by the functions $\rho(x)$ and $y_a(x)$. The function $\rho(x)$ gives the density of eigenvalues with real part $N^{1/2}x$; it is the same for each gauge group. The function $y_a(x)$ gives the imaginary parts of the eigenvalues with real part $N^{1/2}x$. Each gauge group has a different $y_a(x)$. In the theory described by figure 1.2, for example, the free energy functional is given by [27]

$$F[\rho, \delta y_a] = \pi N^{3/2} \int \rho(x) dx \left[\left(\sum_a n_a \right) |x| + 2x \sum_{a=1}^{d} q_a \delta y_a(x) + \rho(x) \sum_{a=1}^{d} f(y_{a-1}(x) - y_a(x)) \right], \tag{1.1}$$

where f is a periodic function of period one that satisfies $f(x) = \frac{1}{4} - x^2$ for $-\frac{1}{2} \leq x \leq \frac{1}{2}$. We can determine ρ, $y_a(x)$, and the free energy by extremizing this functional.

All Sasaki-Einstein Y^7 have a $U(1)_R$ symmetry, and those that can be described by quiver gauge theories have an additional magnetic $U(1)$. Let $\psi(r, m)$ be the number of gauge invariant chiral primary operators with R-charge $\leq r$ and magnetic charge $\leq m$. If X is a chiral operator that is charged only under the center of the gauge group, then let $\psi_X(r, m)$ count gauge invariant chiral operators modulo the relation $X = 0$. For large r, the functions ψ and ψ_X can be approximated by homogeneous, piecewise polynomial funtions. When we write ψ and ψ_X, we will sometimes refer to these approximations. The proposed relationship between matrix eigenvalues and operator counts is

$$\left.\frac{\partial^3 \psi}{\partial r^2 \partial m}\right|_{m=rx/\mu} = \frac{r}{\mu}\rho(x) \ , \tag{1.2}$$

$$\left.\frac{\partial^2 \psi_X}{\partial r \partial m}\right|_{m=rx/\mu} = \frac{r}{\mu}\rho(x)\left[R(X) + \sum y_a q_a\right] \ , \tag{1.3}$$

where the q_a are the gauge charges of X. These formulas are useful because the left-hand side can be computed directly from the geometry without needing to know anything about the field theory. In particular, for theories such as chiral theories where we are not yet sure how to analyze the matrix model, we can formulate a guess as to what the eigenvalue densities should be.

1.4 Issues and possible future directions

We can now construct a CFT dual to Type IIB string theory on $AdS_5 \times Y^5$ for toric Y^5, but many Y^5 of interest are not toric. It would be nice to know if there are efficient methods of constructing a field theory for more general Y^5.

A related problem is to construct a CFT dual to M-theory on $AdS_4 \times Y^7$ for arbitrary Y^7. Refs. [28–33] have made some progress on this problem. We will see some examples of these field theories in chapters 3–4, although we will not address the problem of constructing them. In general, this problem seems to be significantly more difficult than the problem of constructing a CFT dual to Type IIB string theory on $AdS_5 \times Y^5$.

One weakness of our methods for analyzing the matrix model is that they appear not to work for theories with chiral bifundamentals. It is still unknown whether the matrix model gives the expected $N^{3/2}$ scaling for the free energy in chiral theories. Our operator counting formulas offer a prediction of the saddle point eigenvalues, but in the case that we examined, we were not able to find a free energy functional that had the predicted eigenvalues as a saddle point. One promising result for chiral theories is the derivation of Seiberg duality invariance of the matrix model in [34].

Mariño and Putrov [35, 36] have had success in analyzing the matrix model using the methods of many-body quantum mechanics. They observed that the matrix model can be expressed as the partition function for a one-dimensional Fermi gas. Furthermore, for necklace quivers, the fermions do not interact with each other. They were able to find an estimate for the free energy with an error that decreases exponentially in N. In addition to the necklace quivers, they analyzed some non-chiral $\mathcal{N} = 2$ theories with long-range forces. These theories are quiver theories but are not believed to be dual to M-theory on AdS_4. For these theories, they were able to find the free energy to leading order. It is not obvious how to generalize their methods to other theories, however.

All of the quivers studied in this thesis have only unitary gauge groups, and these groups all have equal ranks. Ref. [37] examines the matrix model for $\mathcal{N} = 3$ theories with unitary groups of different ranks, and ref. [38] considers $\mathcal{N} = 3$ theories with orthogonal and symplectic gauge groups.

There are also many Y^7 that are described by something more exotic than a quiver gauge theory. We will analyze a few of these in chapters 3 and 4 for necklace quivers with $T(U(N))$ couplings. ($T(U(N))$ is a theory that can couple to two $U(N)$ gauge groups. It causes operators that are magnetically charged under one of the gauge groups to acquire electric charge in the other gauge group. Essentially, it causes the two gauge groups to become S-dual versions of each other.) However, there are many more possibilities that we haven't considered. As long as Y^7 has a $U(1)_R \times U(1)$ isometry and the field theory is non-chiral, we expect that the operator counting formulas (1.2) and (1.3) will allow us to determine the correct matrix model eigenvalues. For Y^7 without a $U(1)_R \times U(1)$, we do

not know of a Lagrangian description for the theory, and so it is not clear if it is possible to write down a matrix model.

We would like to understand better why the matrix model works from a geometric perspective. Extremizing a free energy functional such as (1.1) gives us equations of motion. We can use (1.2) and (1.3) to replace ρ and y_a with operator counts and thereby rewrite the equations of motion as purely geometric statements on Y^7, without any reference to the matrix model. For example, if we take the equation of motion that we obtain by extremizing (1.1) with respect to ρ and use (1.2) and (1.3) to replace the matrix model quantities with operator densities, we get

$$\sum_a \psi_{A_a}^{(1,1)} \psi_{B_a}^{(1,1)} = 2\psi^{(1,1)} \tag{1.4}$$

where A_a and B_a are bifundamental fields. It is not obvious why such a relation should be true. It would be interesting to know if there is a nice geometric explanation for this and similar relations. It would also be interesting to know if there is any deeper connection between the $\mathcal{N} = 3$ theories whose quivers are Dynkin diagrams and the corresponding Lie groups.

Chapter 2

Properly ordered dimers, R-charges, and an efficient inverse algorithm

This chapter is an edited version of ref. [39].

2.1 Introduction

Given a Sasaki-Einstein manifold X_5, it is generally difficult to construct a field theory that describes Type IIB string theory on $AdS_5 \times X_5$. If the cone over X_5 is toric, then the problem becomes more manageable. In the toric case. dimer models [6–12] are a convenient way of encoding the field content and superpotential of the CFT. One can try to compute the geometry from the dimer or vice versa. There are algorithms for solving the former problem by taking the determinant of the Kasteleyn matrix [6–12] and by counting the windings of zigzag paths [10–12]. The latter problem can be solved by the "Fast Inverse Algorithm" [10–12], although the algorithm is computationally infeasible for all but very simple toric varieties due to the large amount of trial and error required. We resolve this problem by eliminating the need for trial and error. Our algorithm uses some ideas from the Fast Inverse Algorithm and the method of partial resolution of the toric singularity [40–43].

One difficulty in constructing dimers is that not every dimer describes a consistent field theory. One way of determining that a dimer is not consistent is by counting faces. Each face represents a gauge group, and a consistent theory should have as many gauge groups as there are cycles for Type IIB D-branes to wrap in the AdS theory. Previously there was not a simple, easy to check criterion for determining that a dimer is consistent. We propose that any dimer that has the correct number of faces and that has no nodes of valence one is consistent. We will present several pieces of evidence to support our proposal.

If the dimer is consistent, then the cubic anomalies of the CFT should be equal to the Chern-Simons coefficients of the AdS dual [5, 16]. We show that equality holds in dimers that meet our two criteria.

In a four-dimensional SCFT the unitarity bound says that each gauge invariant scalar operator should have dimension at least one [14], and the R-charge of a chiral primary operator is two-thirds of its dimension [15]. However, when we try to compute the R-charge of a gauge invariant chiral primary operator in an inconsistent dimer theory, the answer is sometimes less than two-thirds. We will show that in dimers that meet our two criteria, the R-charges of chiral primary operators are always at least two-thirds if the number of colors is sufficiently large.

We also show that dimers that meet our two criteria have the properties that corner perfect matchings are unique, and that the zigzag path windings agree with the (p, q)-legs of the toric diagram.

While studying R-charges, we prove that $\frac{27N^2K}{8\pi^2} < a \leq \frac{N^2K}{2}$ for toric theories, where a is the cubic 't Hooft anomaly $\frac{3}{32}(3\,\mathrm{Tr}\,R^3 - \mathrm{Tr}\,R)$, N is the number of colors of each gauge group, and K is the area of the toric diagram (which is half the number of gauge groups).

2.2 Definitions

A *dimer model* [6–12] consists of a graph whose vertices are colored black or white, and every edge connects a white vertex to a black vertex, i. e. the graph is bipartite. We will use dimer models embedded on the torus T^2 to describe toric quiver gauge theories.

A *perfect matching* of the dimer is a set of edges of the dimer such that each vertex is an endpoint of exactly one of the edges. The difference of two perfect matchings is the set of edges that belong to exactly one of the matchings.

The *Kasteleyn matrix* is a weighted adjacency matrix of the dimer. There is one row for each white vertex and one column for each black vertex. Let γ_w and γ_z be a pair of curves whose winding numbers generate the homology group $H^1(T^2)$. The weight of an edge is $cw^a z^b$ where c is an arbitrary nonzero complex number[1], w and z are variables, a is the number of times γ_w crosses the edge with the white edge endpoint on its left minus the number of times γ_w crosses the edge with the white endpoint on its right and b is defined similarly with γ_w replaced by γ_z. The determinant of this matrix tells us the geometry of the field configuration.

The *Newton polygon* of a multivariate polynomial is the convex hull of the set of exponents of monomials appearing in the polynomial. The Newton polygon of the determinant is known as the *toric diagram*. If we choose a different basis for computing the Kasteleyn matrix, then the toric diagram changes by an affine transformation.

A (p, q)-*leg* of a toric diagram is a line segment drawn perpendicular to and proportional in length to a segment joining consecutive boundary lattice points of the diagram.

A *zigzag path* is a path of the dimer on which edges alternate between being clockwise adjacent around a vertex and being counterclockwise adjacent around a vertex. A zigzag path is uniquely determined by a choice of an edge and whether to turn clockwise or counterclockwise to find the next edge. Therefore each edge belongs to two zigzag paths. (These paths could turn out to be the same, although it will turn out that we want to work with models in which they are always different.)

In [10] it is conjectured that in a consistent field theory, the toric diagram can also be computed by looking at the windings of the zigzag paths: they are in one-to-one correspondence with the (p, q)-legs. The conjecture was proved using mirror symmetry in [11].

[1]The original definition of the Kasteleyn matrix imposes constraints on c for the purpose of counting perfect matchings [7–10]. However, these constraints are not necessary for determining the Newton polygon. We follow the convention of [11], which points out that it is useful for the purposes of mirror symmetry to allow arbitrary nonzero coefficients.

The *unsigned crossing number* of a pair of closed paths on the torus is the number of times they intersect. The *signed crossing number* of a pair of oriented closed paths on the torus is the number of times they intersect with a positive orientation (the tangent vector to the second path is counterclockwise from the tangent to the first at the point of intersection) minus the number of times they intersect with a negative orientation. It is a basic fact from homology theory that the signed crossing number of a path with winding (a, b) and a path with winding (c, d) is $(a, b) \wedge (c, d) = ad - bc$.

We will work with the zigzag path diagrams of [10] (referred to there as rhombus loop diagrams). We obtain a zigzag path diagram from a dimer as follows. For each edge of the dimer we draw a vertex of the zigzag path diagram at a point on that edge. To avoid confusion between the vertices of this diagram and the vertices of the dimer we will call the latter nodes. We connect two vertices of the zigzag path diagram if the dimer edges they represent are consecutive along a zigzag path. (This is equivalent to them being consecutive around a node and also to them being consecutive around a face.) We orient the edges of the zigzag path diagram as follows. If the endpoints lie on dimer edges that meet at a white (resp. black) node, then the edge should go counterclockwise (resp. clockwise) as seen from that node. With this definition, each node of the dimer becomes a face of the zigzag path diagram, with all edges oriented counterclockwise for a white node, or clockwise for a black node. The other faces of the zigzag path diagram correspond to faces of the dimer, and the orientations of their edges alternate. Figure 2.17 shows an example of a dimer and its corresponding zigzag path diagram.

Conversely, we can obtain a dimer from a zigzag path diagram provided that the orientations of the intersections alternate along each path. Around each vertex of such a zigzag path diagram, there is one face with all counterclockwise oriented edges, one face with all clockwise oriented edges, and two faces whose edge orientations alternate. Draw a white node at each counterclockwise oriented face and a black node at each clockwise oriented face, and connect nodes whose faces share a corner.

2.3 Consistency of dimer field theories

2.3.1 Criteria for consistency and inconsistency

One difficulty in dealing with dimer models is that not all of them produce valid field theories. While there are a number of ways of determining that a dimer produces an invalid field theory there has not yet been a simple criterion for showing that a dimer theory is valid.

One way of proving that a dimer produces an invalid field theory is by counting the number of faces of the dimer, i. e. the number of gauge groups. If the dimer theory is consistent, then the number of gauge groups should equal the number of 0, 2, and 4-cycles in the Calabi-Yau around which D3, D5, and D7-branes, respectively, can wrap [13]. The Euler characteristic of the Calabi-Yau is the number of even dimensional cycles minus the number of odd dimensional cycles. There are no odd dimensional cycles, so the number of gauge groups should be equal to the Euler characteristic. The Euler characteristic of a toric variety equals twice the area of the toric diagram [44].

We propose that a dimer will produce a valid field theory if the dimer has no nodes of valence one and it has a number of faces equal to twice the area of the lattice polygon whose (p, q)-legs are the winding numbers of the zigzag paths. (Recall that this polygon is the same as the Newton polygon of the determinant of the Kasteleyn matrix for consistent theories.) In this section, we will show that dimers satisfying our two criteria also have the properties that their cubic anomalies agree with the Chern-Simons coefficients of the AdS dual, the R-charges of gauge invariant chiral primary operators are greater than or equal to two-thirds, the windings of the zigzag paths are in one-to-one correspondence with the (p, q)-legs of the toric diagram, and the corner perfect matchings are unique.

It will be convenient to introduce a property that we call "proper ordering", which will turn out to be equivalent to the property of having the correct number of faces and no valence one nodes. We call a node of the dimer *properly ordered* if the order of the zigzag paths around that node is the same as the circular order of the directions of their windings. (We do not allow two zigzag paths with the same winding to intersect, nor do we allow

zigzag paths of winding zero, since these scenarios make the ordering ambiguous.) We call a dimer properly ordered if each of its nodes is properly ordered.

Theorem 2.3.1. *A connected dimer is properly ordered iff it has no valence one nodes and it has a number of faces equal to twice the area of the convex polygon whose (p, q)-legs are the winding numbers of the zigzag paths of the dimer.*

Proof. A properly ordered dimer cannot have a valence one node, since such a node would be the endpoint of an edge that is an intersection of a zigzag path with itself. Therefore it suffices to prove that a dimer with no valence one nodes is properly ordered iff it has a number of faces equal to twice the area of the convex polygon whose (p, q)-legs are the winding numbers of the zigzag paths of the dimer.

Define the "winding excess" of a node v of the dimer as follows. Let $\mathbf{w}_0, \mathbf{w}_1, ..., \mathbf{w}_{n-1}$ be the winding numbers of the zigzag paths passing through v (in the order that the paths appear around v). Start at $\mathbf{w}_0$ and turn counterclockwise to $\mathbf{w}_1$, then counterclockwise to $\mathbf{w}_2$, etc., and finally counterclockwise back to $\mathbf{w}_0$. Then the winding excess is defined as the number of revolutions that we have made minus one. (In the special case where $\mathbf{w}_i$ and $\mathbf{w}_{i+1}$ are equal or one of them is zero, we count one-half of a revolution.) A node is properly ordered iff it has winding excess zero and none of the $\mathbf{w}_i$ are zero and no two consecutive windings are equal. A node with a $\mathbf{w}_i = 0$ or $\mathbf{w}_i = \mathbf{w}_{i+1}$ can have winding excess zero only if it has exactly two edges (and hence two zigzag paths passing through it). There must be some other node that is an endpoint of one of the edges where the two zigzag paths intersect, and that has more than two edges (since the graph is connected). This node cannot have winding excess zero. So all nodes are properly ordered iff all nodes have winding excess zero. A node has negative winding excess iff it has just one edge, and we have assumed that the dimer has no such nodes. Therefore the dimer is properly ordered iff the sum of all of the winding excesses is zero.

If we choose a node and draw all of the wedges between the consecutive winding numbers, then the winding excess is the number of wedges containing any given ray minus one. (We can think of the special case of consecutive winding numbers being the same as the average of a full wedge and an empty wedge, and the case of a zero winding number as the average

of wedges of all angles.) Now consider the sum of the winding excess over all vertices. A pair of oppositely oriented intersections between two zigzag paths forms two full wedges and therefore contributes two to the sum. The sum of the contributions from unpaired intersections can be computed as follows. Label the winding numbers $\mathbf{w}_0, \mathbf{w}_1, ..., \mathbf{w}_{n-1}$, ordered by counterclockwise angle from some ray $\mathcal{R}$. (A zigzag path with zero winding number has no unpaired intersections, so it is not included.) Then for $i < j$ the unpaired wedges formed by $\mathbf{w}_i$ and $\mathbf{w}_j$ will contain $\mathcal{R}$ iff $\mathbf{w}_i \wedge \mathbf{w}_j < 0$. There are $2|\mathbf{w}_i \wedge \mathbf{w}_j|$ unpaired wedges ($|\mathbf{w}_i \wedge \mathbf{w}_j|$ unpaired crossings of the zigzag paths, and each appears in two vertices). So the number of unpaired wedges formed by $\mathbf{w}_i$ and $\mathbf{w}_j$ containing $\mathcal{R}$ equals $\max(-2\mathbf{w}_i \wedge \mathbf{w}_j, 0) = |\mathbf{w}_i \wedge \mathbf{w}_j| - \mathbf{w}_i \wedge \mathbf{w}_j$. Since $\sum_{i<j} |\mathbf{w}_i \wedge \mathbf{w}_j|$ is the number of unpaired edges, it follows that the number of wedges containing $\mathcal{R}$ is the number of paired edges plus the number of unpaired edges minus $\sum_{i<j} \mathbf{w}_i \wedge \mathbf{w}_j$, or $E - \sum_{i<j} \mathbf{w}_i \wedge \mathbf{w}_j$, where E is the total number of edges of the dimer. The sum of the winding excesses is $E - V - \sum_{i<j} \mathbf{w}_i \wedge \mathbf{w}_j = F - \sum_{i<j} \mathbf{w}_i \wedge \mathbf{w}_j$, where V and F are the number of nodes and faces of the dimer, respectively. We have $\sum_{i<j} \mathbf{w}_i \wedge \mathbf{w}_j = \sum_i \mathbf{w}_i \wedge \sum_{j>i} \mathbf{w}_j$. If we lay the winding vectors tip-to-tail, then $\mathbf{w}_i \wedge \sum_{j>i} \mathbf{w}_j$ is twice the area of the triangle formed by the tail of $\mathbf{w}_0$ and the tip and tail of $\mathbf{w}_i$. Hence $\sum_i \mathbf{w}_i \wedge \sum_{j>i} \mathbf{w}_j$ is twice the area of the convex polygon formed by all the winding vectors. If we rotate the polygon 90 degrees then we get a polygon whose (p, q)-legs are the winding numbers. So the sum of the winding deficiencies of the nodes is zero iff F equals twice the area of the lattice polygon whose (p, q)-legs are the zigzag path winding numbers. □

2.3.2 Some perfect matchings of properly ordered dimers

We will construct some perfect matchings that will turn out to correspond to the corners of the toric diagram. Our construction of the perfect matchings is similar to Theorem 7.2 of [12]. Let $\mathcal{R}$ be any ray whose direction does not coincide with that of the winding number of any zigzag path. For any node v, consider the zigzag paths passing through v whose winding numbers make the smallest clockwise and smallest counterclockwise angles with $\mathcal{R}$. (These paths are unique because the proper ordering condition requires that all paths through v have different winding numbers.) By proper ordering, these two zigzag paths

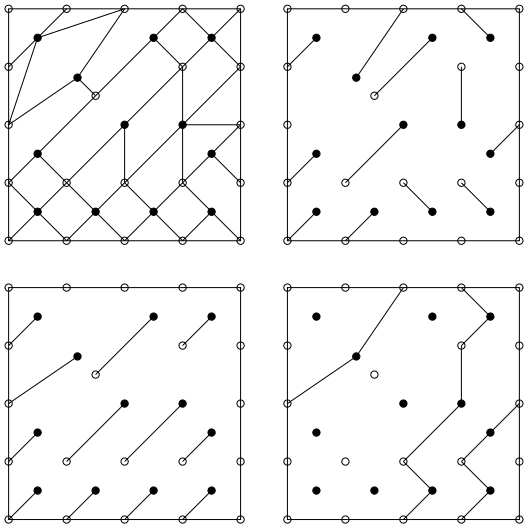


Figure 2.1: A dimer, two of its corner perfect matchings, and their difference, which is a zigzag path.

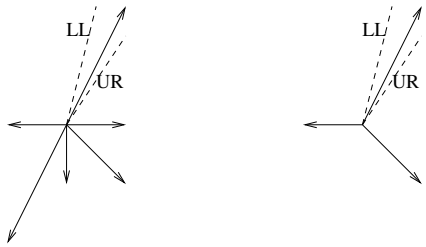


Figure 2.2: Left: The windings of the zigzag paths of the dimer in figure 2.3.2. The dotted lines labeled UR and LL are rays that yield perfect matchings shown in the upper right and lower left quadrants of figure 2.3.2, respectively. Right: The windings of the paths passing through the bottom right black node. For any node and any edge ending at that node, the proper ordering criterion implies that the two zigzag paths to which the edge belongs have adjacent winding directions. Therefore in the right diagram, there is a natural correspondence between edges passing through the node and wedges formed by consecutive arrows. When constructing a perfect matching $M(\mathcal{R})$, we choose the wedge containing $\mathcal{R}$. In the left diagram, there is a one-to-one correspondence between wedges and corners of the toric diagram.

must be consecutive around v. Therefore they share an edge that has v as an endpoint. Call this edge $e(v)$. Let v' be the other endpoint of $e(v)$. The same two zigzag paths must be consecutive about v' since they form the edge e. Since v' is properly ordered it must then be the case that those two paths make the smallest clockwise and counterclockwise angles with $\mathcal{R}$ among all paths passing through v'. Hence $e(v) = e(v')$. So the pairing of v with v' is a perfect matching. We will call this matching $M(\mathcal{R})$. Figure 2.3.2 depicts the relationship between rays and perfect matchings.

The following characterization of the boundary perfect matchings containing a given edge follows immediately from our definition and will be useful later.

Lemma 2.3.2. *For any edge e of the dimer, let Z_r and Z_s be the zigzag paths such that e is a positively oriented intersection of Z_r with Z_s. (Equivalently, e is a negatively oriented intersection of Z_s with Z_r.) Let $\mathbf{w}_r$ and $\mathbf{w}_s$ be the windings of Z_r and Z_s, respectively. Let $\mathcal{R}$ be a ray. Then e is in $M(\mathcal{R})$ iff $\mathcal{R}$ is in the wedge that goes counterclockwise from $\mathbf{w}_r$ to $\mathbf{w}_s$.*

In particular each edge is in at least one corner perfect matching.

2.3.3 Zigzag paths and (p, q)-legs

As we mentioned in Section 2.2, it is known [10, 11] that dimers that produce a consistent field theory have the property that the (p, q)-legs of the toric diagram are in one-to-one correspondence with the winding numbers of the zigzag paths.

Theorem 2.3.3. *In a dimer with properly ordered nodes, the zigzag paths are in one-to-one correspondence with the (p, q)-legs of the toric diagram.*

Our proof of Theorem 2.3.3 resembles that of Theorem 9.3 of [12].

Lemma 2.3.4. *For any zigzag path Z in any dimer, the number of intersections of a perfect matching with Z is a degree one polynomial function of its coordinates.*

Proof. In computing the Kasteleyn matrix we can choose the path γ_z to follow Z, so that the number of times γ_z intersects a perfect matching M is just the number of edges that M and Z have in common. (See figure 2.3.3.) For this choice of γ_z, the point corresponding

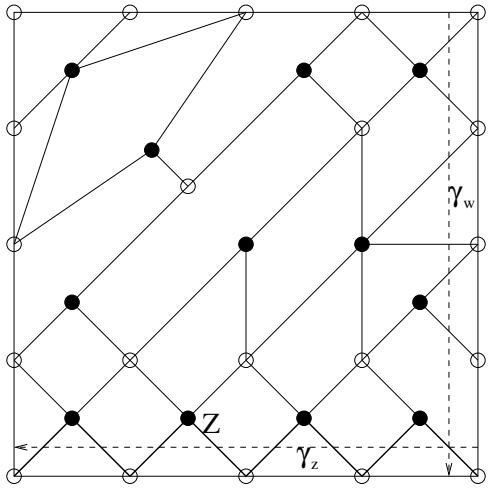


Figure 2.3: The path γ_z intersects each edge of the zigzag path Z and no other edges. We may choose any path γ_w that completes the basis.

to M has y-coordinate equal to $|M \cap Z|$. For a different choice of γ_z, the coordinates differ by an affine transformation. □

Lemma 2.3.5. *Let Z be a zigzag path of a properly ordered dimer, and let $\mathcal{R}_1$ and $\mathcal{R}_2$ be rays such that the winding direction of Z lies between them and all of the other winding directions do not. Then there exists a boundary line of the toric diagram passing through $M(\mathcal{R}_1)$ and $M(\mathcal{R}_2)$ such that all perfect matchings on this line intersect Z exactly $\frac{|Z|}{2}$ times, and all perfect matchings not on the line intersect Z fewer than $\frac{|Z|}{2}$ times.*

Proof. Since the winding number of Z is adjacent to $\mathcal{R}_1$, $M(\mathcal{R}_1)$ must choose one of the two Z-edges of each node that has them. Hence $|M(\mathcal{R}_1) \cap Z| = \frac{|Z|}{2}$ and similarly $|M(\mathcal{R}_2) \cap Z| = \frac{|Z|}{2}$. No perfect matching can contain more than half of the edges of the path. Therefore the toric diagram lies in the half plane that, in the coordinate system of Lemma 2.3.4, is given by the equation $y \leq \frac{|Z|}{2}$. $M(\mathcal{R}_1)$ and $M(\mathcal{R}_2)$ are both on the boundary. □

Proposition 2.3.6. *The matchings $M(\mathcal{R})$ lie on the corners of the toric diagram. The order of the corners around the boundary is the same as the order of the ray directions.*

Proof. The intersection of all half planes described in the proof of Lemma 2.3.5 is the convex hull of all of the $M(\mathcal{R})$'s. Conversely, each $M(\mathcal{R})$ is in the toric diagram. So the toric diagram must be the convex hull of the $M(\mathcal{R})$'s.

Each $M(\mathcal{R})$ must be at a corner of the toric diagram since it is contained in two different boundary lines (one for the first counterclockwise zigzag path direction from $\mathcal{R}$ and another for the first clockwise zigzag path direction). Furthermore, if $\mathcal{R}_1$ and $\mathcal{R}_2$ have only one winding direction between them, then they share a boundary line and hence $M(\mathcal{R}_1)$ and $M(\mathcal{R}_2)$ lie on consecutive corners. □

Proof of Theorem 2.3.3. Let $\mathbf{w}$ be the winding number of a zigzag path, and let n be the number of zigzag paths with that winding. Let $\mathcal{R}_1$ and $\mathcal{R}_2$ be rays such that $\mathbf{w}$ lies between them and all other winding directions do not. By Proposition 2.3.6, $M(\mathcal{R}_1)$ and $M(\mathcal{R}_2)$ lie on consecutive corners of the toric diagram. An edge belonging to one of the zigzag paths of winding $\mathbf{w}$ will be in either $M(\mathcal{R}_1)$ or $M(\mathcal{R}_2)$ but not both, while all other edges are in neither or both perfect matchings. Therefore the difference of the two perfect matchings is just the union of the zigzag paths with winding $\mathbf{w}$. Therefore the toric diagram points corresponding to $M(\mathcal{R}_1)$ and $M(\mathcal{R}_2)$ are separated by $-n\mathbf{w}^{\perp}$, where $-\mathbf{w}^{\perp}$ is the 90 degree clockwise rotation of $\mathbf{w}$. This proves the theorem. □

2.3.4 Unique corner perfect matchings

It is generally believed that dimers that have more than one perfect matching at a corner of the toric diagram are inconsistent [7,10,45]. We show that properly ordered dimers have unique corner perfect matchings.

Theorem 2.3.7. *If a dimer is properly ordered, then each corner of the toric diagram has just one perfect matching.*

Proof. Suppose there exists a perfect matching M' that shares a toric diagram point with $M(\mathcal{R})$ but is not equal to $M(\mathcal{R})$. Consider the set of zigzag paths that contain an edge that is in $M(\mathcal{R})$ or M' but not both. Let Z be one with minimal counterclockwise angle from $\mathcal{R}$. Let v be a node of the dimer through which Z passes. If v includes a zigzag path with winding between $\mathcal{R}$ and that of Z, then $M(\mathcal{R})$ and M' are the same at that vertex. If not, then $M(\mathcal{R})$ chooses one of the edges of Z at v. Recall that Lemma 2.3.4 says that the number of intersections with Z depends only on the toric diagram point. Therefore M' has the same number of edges in Z as $M(\mathcal{R})$. Since $M(\mathcal{R})$ chooses an edge of Z at every

node where $M(\mathcal{R})$ and M' differ, equality can hold only if M' chooses the other edge of Z at every such node. If we start at an edge of Z that is in $M(\mathcal{R})$ but not M' and alternately follow edges of the $M(\mathcal{R})$ and M', then we will traverse a cycle that lies entirely in Z. Since zigzag paths in properly ordered dimers do not intersect themselves, the cycle must be Z. Then both $M(\mathcal{R})$ and M' contain half the edges of Z. So the winding number of Z is either the closest or farthest from $\mathcal{R}$ in the counterclockwise direction. If Z were the farthest, then $M(\mathcal{R})$ and M' would have to be the same because every edge of the dimer would be in at least one zigzag path whose winding is closer to $\mathcal{R}$ in the counterclockwise direction than Z's. So Z must be the closest in the counterclockwise direction.

Now let Z' be a zigzag path with minimal clockwise angle from $\mathcal{R}$ on which $M(\mathcal{R})$ and M' differ. By the same reasoning as above, we find the winding direction of Z' is the closest to $\mathcal{R}$ in the clockwise direction and that $M(\mathcal{R})$ and M' have no edges of Z' in common. Since Z and Z' represent consecutive sides of the toric diagram, the crossing number of Z and Z' must be nonzero. A node can have two edges belonging to both Z and Z' only if they have opposite orientations, i. e. they contribute zero to the signed crossing number. Therefore there must be a node with only one Z-Z' intersection. M' must include this edge because it includes an edge of Z and Z' at every node that has one, but it cannot include this edge because it does not share any edges of Z with $M(\mathcal{R})$. Therefore our assumption that there existed a matching M' differing from $M(\mathcal{R})$ but sharing the same toric diagram point must be false. □

Once we know that the corner matchings are unique, we can also classify all of the boundary perfect matchings.

Corollary 2.3.8. *Consider a point A on the boundary of the toric diagram such that the nearest corner B in the counterclockwise direction is p segments away and the nearest corner C in the clockwise direction is q segments away. Then each perfect matching at A may be obtained from the perfect matching associated to B by flipping p zigzag paths and from the perfect matching associated to C by flipping q zigzag paths. The number of perfect matchings at A is $\binom{p+q}{q}$.*

Proof. For any boundary perfect matching M there exists a winding $\mathbf{w}$ such that M contains half the edges of each zigzag path of winding $\mathbf{w}$. For any zigzag path Z of winding $\mathbf{w}$, we can delete the half of the Z-edges that are in M and add the other half. This operation moves the perfect matching one segment along the boundary of the toric diagram. There can be at most p zigzag paths for which the operation moves the toric diagram point counterclockwise and at most q zigzag paths for which the operation moves the point clockwise. But there are a total of $p+q$ zigzag paths of winding $\mathbf{w}$, so there must be exactly p of the former and q of the latter. Consequently we see that M can be obtained from a corner perfect matching by flipping p zigzag paths (or from a different corner perfect matching by flipping q zigzag paths). The number of ways of choosing the paths to flip is $\binom{p+q}{p}$. □

2.3.5 R-charges and cubic anomalies

The R-charges of the fields may be determined by *a-maximization* [24]. First, we impose the constraint that the R-charge of each superpotential term should be two. We also impose the constraint that the beta function of each gauge group should be zero. These conditions can be expressed as

$$\sum_{e \in v} R(e) = 2 \tag{2.1}$$

$$\sum_{e \in f} [1 - R(e)] = 2. \tag{2.2}$$

Among all $U(1)$ symmetries satisfying these constraints, the R-symmetry is the one that locally maximizes the cubic 't Hooft anomaly

$$a = \frac{9N^2}{32}\left[F + \sum_e (R(e)-1)^3\right]. \tag{2.3}$$

Butti and Zaffaroni [45] have proposed some techniques for simplifying the computation of the R-charge. For any perfect matching M we can define a function δ_M that takes the value 2 on all edges in the perfect matching and zero on all other edges. Any such δ_M automatically satisfies (2.1). Butti and Zaffaroni noted that in some cases the perfect matchings on the boundary of the toric diagram yield functions that also satisfy (2.2),

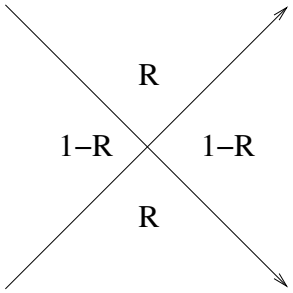


Figure 2.4: The contribution of the vertex to the equations (2.4) for the four surrounding faces.

and these functions span the set of solutions to (2.1) and (2.2). We will show that their observation is true for properly ordered dimers.

Theorem 2.3.9. *In a dimer with properly oriented nodes, the solutions to (2.1) and (2.2) are precisely the linear combinations of δ_M, for boundary perfect matchings M.*

We first determine the dimension of the solution space of (2.1) and (2.2), so that we will be able to show that there are not any more solutions beyond the boundary δ_M.

Lemma 2.3.10. *For any dimer in which the zigzag paths have winding numbers that are prime (i. e. their x and y components are relatively prime, or equivalently, they can each be sent to $(1,0)$ by an $SL_2(\mathbb{Z})$ transformation) and not all parallel and in which no zigzag path intersects itself, the set of solutions to (2.1) and (2.2) has dimension equal to the number of zigzag paths minus one.*

Proof. First we will show that the number of solutions depends only on the winding numbers of the zigzag paths. We will work with the zigzag path diagram. In this diagram, R is a function on vertices. We can unify (2.1) and (2.2) into a single equation as follows. We first define the function $\sigma_{v,f}(x)$, where v is a vertex of the zigzag path diagram and f is a face of the zigzag path diagram having v as a corner. If the two zigzag paths containing v are similarly oriented around f, then $\sigma_{v,f}(x) = x$; if they are oppositely oriented around f then $\sigma_{v,f}(x) = 1 - x$. Then (2.1) and (2.2) can be expressed as

$$\sum_{v \in f} \sigma_{v,f}(R(v)) = 2. \tag{2.4}$$

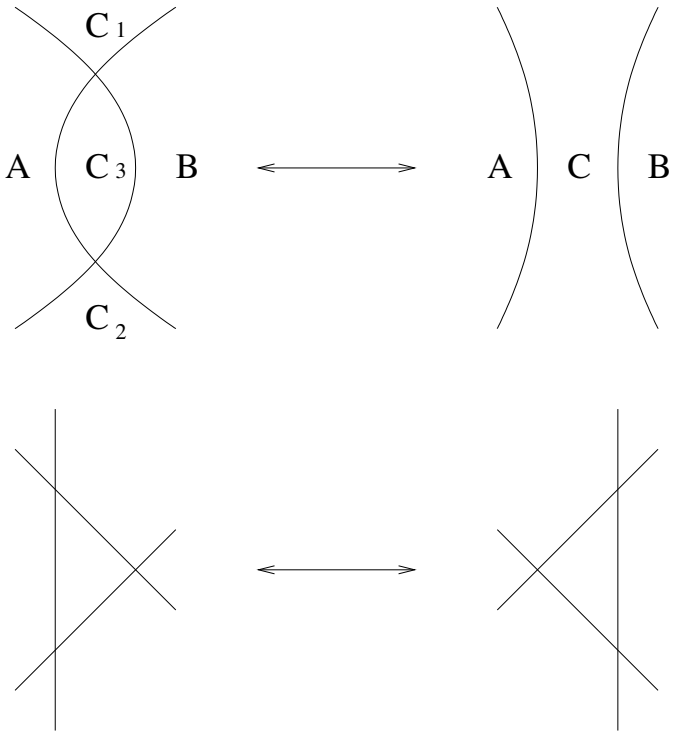


Figure 2.5: Combinatorial changes in the zigzag path diagram as a result of deformations.

(See Figure 2.4).

We can deform any zigzag path diagram with non-self-intersecting zigzag paths to any other zigzag path diagram with non-self-intersecting zigzag paths with the same winding numbers. As the diagram is deformed, it can change combinatorially in several ways: a pair of intersections between a pair of zigzag paths can be added or removed, or a zigzag path can be moved past the crossing of two other zigzag paths. Figure 2.5 illustrates these possibilities. Note that at intermediate steps, the zigzag path diagram may not correspond to a dimer, but we can still consider the set of solutions to (2.4).

First consider the case where a pair of intersections between a pair of zigzag paths is added or removed. If C_1 is not the same face as C_2, then the values of the two new crossings are constrained by the equations for C_1 and C_2 and the dimension of the set of solutions to (2.1) and (2.2) remains unchanged. If the two zigzag paths have winding numbers that are not parallel, then they must intersect somewhere else, which implies $C_1 \neq C_2$. If the winding numbers are parallel, then there must be some other zigzag path whose winding number is not parallel to either and hence must intersect both. Again $C_1 \neq C_2$.

Now consider the case where a zigzag path is moved past the crossing of two other zigzag paths. We can check that any solution to (2.4) in the first diagram is also a solution to (2.4) in the second diagram, and vice versa. So performing the move depicted in the second diagram does not change the solution set. So we have shown that the dimension of the solution space to (2.4) depends only on the winding numbers of the zigzag paths.

In Lemma 2.6.2 we will exhibit for any set of winding numbers a dimer for which the number of independent solutions to (2.1) and (2.2) is the number of zigzag paths minus one. □

Lemma 2.3.10 tells us how to solve (2.1) and (2.2) for a large class of dimers, many of which are not properly ordered. It is interesting to note that the second move shown in Figure 2.5 does not change either a or $\sum_e(1 - R(e))$. The first move also leaves a and $\sum_e(1 - R(e))$ invariant in the case where the two zigzag paths are oppositely oriented (the charges of the introduced vertices sum to two and $(R_1 - 1)^3 + (R_2 - 1)^3 = (R_1 + R_2 - 2)\left[(R_1 - 1)^2 - (R_1 - 1)(R_2 - 1) + (R_2 - 1)^2\right]$). When the zigzag path diagram corresponds to a dimer, $\sum_e(1 - R(e))$ is the number of faces in the dimer.

Proof of Theorem 2.3.9. First we will show that the δ_M are solutions to (2.2). Suppose a face f with $2n$ sides had n of those sides in a boundary perfect matching M. (A side of a face is an edge of the face along with a normal pointing into the face. If a face borders itself then the bordering edge is part of two different sides of the face. If a self-border edge is in a perfect matching, then we count two sides of f in that perfect matching.) From Corollary 2.3.8, we know that we can get from M to any other boundary perfect matching by flipping zigzag paths. Note that this operation leaves invariant the number of sides of each face in the perfect matching. Therefore every boundary perfect matching has n sides of f. So every node of f selects one of the two adjacent sides of f for all boundary perfect matchings. By Lemma 2.3.2 we know that every edge is in some corner perfect matching. So the only edges belonging to any node of f are the adjacent sides of f. Therefore, as we move along the boundary of the face we are following a zigzag path. But then we have a zigzag path with zero winding, which violates proper ordering. So the assumption that a face with $2n$ sides can have n sides in a boundary perfect matching must be false. Therefore

a face with $2n$ sides can have at most $n-1$ sides in a boundary perfect matching. Sum this inequality over all faces:

$$\sum_f \sum_{s \in f \cap M} 1 \leq \sum_f \left[\left(\sum_{s \in f} \frac{1}{2} \right) - 1 \right] \tag{2.5}$$

where f runs over faces and s runs over sidess. Now reverse the order of the sums:

$$\begin{aligned} \sum_{s \in M} \sum_{f \ni s} 1 &\leq \sum_s \sum_{f \ni s} \frac{1}{2} - F & (2.6) \\ V &\leq (2E)\left(\frac{1}{2}\right) - F & (2.7) \\ V &\leq E - F. & (2.8) \end{aligned}$$

Since we know $V = E - F$, equality must have held in each case. So (2.2) is satisfied by boundary perfect matchings.

The difference between any two boundary perfect matchings is a sum of functions δ_Z, where Z is a zigzag path and the value δ_Z alternates between 2 and -2 on Z and is zero outside of Z. The only relation obeyed by the δ_Z is that they sum to zero. So the dimension of the space of solutions to (2.1) and (2.2) that we have found equals the number of zigzag paths minus one. By Lemma 2.3.10, there can be no more solutions. □

When some of the boundary points of the toric diagram are not corners, there are many sets of perfect matchings that form a basis for the solutions to (2.1) and (2.2). We will construct a basis by associating each segment of the boundary of the toric diagram with a zigzag path, and choosing one perfect matching at each boundary point so that the difference between two consecutive perfect matchings is the zigzag path corresponding to the segment between them. Write $R = \sum_i \lambda_i \delta_{M_i}$, where M_i are the perfect matchings in the basis and the λ_i are real numbers.

Butti and Zaffaroni [45] also noted that in many cases each edge that is a positively oriented intersection of a zigzag path Z_r with another zigzag path Z_s occurs in the perfect matchings in $cc(r, s)$, the counterclockwise segment from r to s, while a negatively oriented intersection of Z_r with Z_s occurs in the perfect matchings not in $cc(r, s)$. In this case, the

value of $R-1$ for a positively oriented intersection of Z_r with Z_s is $2\left(\sum_{i\in cc(r,s)}\lambda_i\right)-1$. For a negatively oriented intersection the value of $R-1$ is $2\left(\sum_{i\notin cc(r,s)}\lambda_i\right)-1$, which equals $-\left[2\left(\sum_{i\in cc(r,s)}\lambda_i\right)-1\right]$ since $\sum_i\lambda_i=1$. So then the total contribution to $\sum_e(R-1)^3$ from the intersections of Z_r with Z_s is $(\mathbf{w}_r\wedge\mathbf{w}_s)\left[2\left(\sum_{i\in cc(r,s)}\lambda_i\right)-1\right]^3$. Hence (2.3) can be rewritten as

$$a=\frac{9N^2}{32}\left[F+\sum_{r<s}(\mathbf{w}_r\wedge\mathbf{w}_s)\left(2\left(\sum_{i\in cc(r,s)}\lambda_i\right)-1\right)^3\right]. \tag{2.9}$$

Proposition 2.3.11. *If a dimer has properly oriented nodes, then it is the case that all positively (resp. negatively) oriented intersections of Z_r with Z_s are in precisely the perfect matchings that are in $cc(r,s)$ (resp. $cc(s,r)$). Hence 2.9 holds for properly ordered dimers.*

Proof. Assume that the dimer has properly ordered nodes. As we go around the toric diagram, the perfect matching switches from containing an edge e to not containing it only if we changed the perfect matching by a zigzag path containing e. So each intersection of Z_r with Z_s occurs in either the perfect matchings in $cc(r,s)$ or the perfect matchings in its complement. From Lemma 2.3.2 we know that the positively oriented intersections are in the corners of $cc(r,s)$ and the negatively oriented intersections are not. □

A particularly nice rearrangement of (2.9) that we will find useful is [16, 46]

$$a=\frac{9N^2}{4}\sum_{ijk}\text{area}(P_iP_jP_k)\lambda_i\lambda_j\lambda_k. \tag{2.10}$$

where P_i is the point on the toric diagram corresponding to the ith perfect matching. This formula tells us that the triangle anomaly of the three symmetries with respective charges δ_{M_i}, δ_{M_j}, and δ_{M_k} is $\frac{N^2}{2}\,\text{area}(P_iP_jP_k)$. AdS-CFT predicts that the $U(1)$ symmetries of the CFT correspond to gauge symmetries in the AdS theory, and that the triangle anomalies of the CFT should equal the corresponding Chern-Simons coefficients in the AdS theory [5]. The Chern-Simons coefficients are indeed found to be $\frac{N^2}{2}\,\text{area}(P_iP_jP_k)$ [16]. So the field theory produced by a properly ordered dimer will have precisely the cubic

anomalies predicted by the AdS theory. This is strong evidence that properly ordered dimers are consistent.

2.3.6 Unitarity bound

Gauge invariant scalar operators in a four-dimensional CFT must have dimension at least one [14]. We also have the BPS bound $\Delta \geq \frac{3}{2}|R|$, where Δ is the dimension of an operator and R is its R-charge. Equality is achieved in the case of chiral primary operators [15]. So in order for the theory to be physically valid it is necessary that the gauge invariant chiral primary operators have R-charge at least $\frac{2}{3}$.

Theorem 2.3.12. *If a can be expressed in the form (2.10), then there exists an N such that in the dimer theory with N colors, each gauge invariant chiral primary operator has R-charge at least $\frac{2}{3}$. In particular properly ordered dimers have this property.*

Lemma 2.3.13. *At the point where a is locally maximized, the weight of each corner perfect matching is positive, and the weight of the other boundary perfect matchings is zero.*

Proof. This follows immediately from equation (4.2) of [45]. □

Lemma 2.3.14 (A. Kato [47])**.** *If a is given by (2.10), then at the point where a is locally maximized,*

$$\frac{\partial a}{\partial \lambda_i} = 3a. \tag{2.11}$$

Proof. We can use Lagrange multipliers to find the local maximum of a.

$$\frac{\partial a}{\partial \lambda_i} = \mu \frac{\partial}{\partial \lambda_i} \sum_j \lambda_j = \mu \tag{2.12}$$

for some constant μ. Since a is homogeneous of degree three,

$$\begin{aligned} 3a &= \sum_i \lambda_i \frac{\partial a}{\partial \lambda_i} && (2.13)\\ &= \sum_i \lambda_i \mu && (2.14)\\ &= \mu. && (2.15) \end{aligned}$$

□

Lemma 2.3.15 (A. Kato [47])**.** *At the point where a is locally maximized, each λ_i is at most $\frac{1}{3}$.*

Proof. By Lemma 2.3.14, $3\lambda_i a = \lambda_i \frac{\partial a}{\partial \lambda_i}$. Since every term of a is degree zero or one in λ_i, the right-hand side is simply the terms of a containing λ_i. We can see from (2.10) that the coefficient of each term of a is nonnegative and from Lemma 2.3.13 that each λ_i is nonnegative when a is maximized. Hence the sum of the terms of a containing λ_i is at most a. Therefore $3\lambda_i a \leq a$, so $\lambda_i \leq \frac{1}{3}$. □

Proof of Theorem 2.3.12. First consider the mesonic operators, which arise as the trace of a product of of operators corresponding to the edges around a loop of the quiver. The number of signed crossings between a loop and a perfect matching of the dimer is an affine function of the perfect matching's position in the toric diagram. If the loop has nonzero winding, then the function is not constant, and its zero locus is a line. This line can intersect the corners of the toric diagram at most twice. Therefore each loop intersects all but at most two of the corner perfect matchings. The sums of the weights of those two perfect matchings is at most $\frac{2}{3}$, and from Lemma 2.3.13 we know that the non-corner matchings have weight zero. The sum of the weights of the perfect matchings that do intersect the loop is then at least $\frac{1}{3}$. So the loop has R-charge at least $\frac{2}{3}$. The R-charge of a loop with zero winding is twice the number of intersections it has with any perfect matching. Every edge is in at least one perfect matching so this number must be positive. So a loop with zero winding has R-charge at least 2.

The theory also has baryonic operators. If the gauge groups are $SU(N)$ then these operators are the Nth exterior powers the bifundamental fields. Each edge of the dimer is contained in at least one corner perfect matching by Lemma 2.3.2, and we know from Lemma 2.3.13 that each corner of the toric diagram has a positive contribution to the R-charge. So each dimer edge has positive R-charge. For sufficiently large N, the corresponding baryonic operator will have R-charge at least $\frac{2}{3}$. □

2.4 Bounds on a

2.4.1 Bounds on a for toric theories

We can use (2.10) to establish bounds for a. In this section we let the indices ijk of the perfect matchings run over the corner perfect matchings only, since we know from Lemma 2.3.13 that the non-corner perfect matchings have weight zero.

Theorem 2.4.1. *Let N be the number of colors of each gauge group, and let K be the area of the toric diagram (which is half the number of gauge groups). Then*

$$\frac{27N^2K}{8\pi^2} < a \leq \frac{N^2K}{2}. \tag{2.16}$$

The upper bound is achieved iff the toric diagram is a triangle, and the lower bound is approached as the toric diagram approaches an ellipse.

Proof of the lower bound of Theorem 2.4.1. The polar body X_R^* of a convex polygon X with respect to the point R is defined as the set of points Q satisfying $\overrightarrow{RQ} \cdot \overrightarrow{RP} \leq 1$ for all $P \in X$. Recall that maximizing a is equivalent to minimizing the volume of a slice of the dual toric cone [45, 48]. More specifically, if $\vec{r}$ is the three-dimensional Reeb vector, then $\frac{9N^2}{8a}$ is the volume of the set of points $\vec{x}$ in the dual cone satisfying $\vec{r} \cdot \vec{x} \leq 3$. The cross section of the dual cone in the plane $\vec{r} \cdot \vec{x} = 3$ is the polar body of the toric diagram with respect to the Reeb vector (considered as a point in the plane of the toric diagram). If we call the toric diagram X, then $\frac{27N^2}{8a} = \inf_{R \in X} \operatorname{area}(X_R^*)$. Then the statement of the lower bound is equivalent to $\operatorname{area}(X) \inf_{R \in X} \operatorname{area}(X_R^*) < \pi^2$. The result $\operatorname{area}(X) \inf_{R \in X} \operatorname{area}(X_R^*) \leq \pi^2$ was proved by Blaschke [49, 50]; equality occurs in the case of an ellipse. Since the toric diagram is a polygon, it cannot be perfectly elliptical and hence equality does not hold. □

We will need to use the following results for the proof of the upper bound.

Proposition 2.4.2 (A. Kato [47])**.** *The local maximum of a is the overall maximum of a in the region $\lambda_i \geq 0, \sum_i \lambda_i = 1$.*

Proposition 2.4.3 (A. Butti and A. Zaffaroni [45]). *Let R be a point in the interior of the toric diagram. Define*

$$f_i = \frac{\text{area}(P_{i-1}P_iP_{i+1})}{\text{area}(P_{i-1}P_iR)\,\text{area}(P_iP_{i+1}R)} \tag{2.17}$$

$$S = \sum_i f_i \tag{2.18}$$

$$\lambda_i = f_i/S. \tag{2.19}$$

Then the following results hold:

$$R = \sum_i \lambda_i P_i \tag{2.20}$$

$$a = \frac{27N^2}{2S}. \tag{2.21}$$

Furthermore, when R is the Reeb vector and the λ_i are given by (2.19), a is locally maximized (over all choices of λ_i, not just those of the form (2.19)).

Proof of the upper bound of Theorem 2.4.1. We use induction on the number of corners of the toric diagram. If the toric diagram is a triangle, then a is maximized when each λ_i is $\frac{1}{3}$. So $a = \frac{9N^2}{4}K(3!)\left(\frac{1}{3}\right)^3 = \frac{N^2K}{2}$.

Assume the toric diagram has more than three corners. Let λ_i^M be the values of λ_i for which a is locally maximized. Choose a particular i and let $\lambda_i^D = 0$, $\lambda_{i+1}^D = \lambda_i^M + \lambda_{i+1}^M$, and $\lambda_j^D = \lambda_j^M$ for all other j. We will define $a^M = a|_{\lambda^M}$, $a^D = a|_{\lambda^D}$, and $\Delta a = a^D - a^M$. Since a has degree one in each individual λ_j,

$$\Delta a = \left.\frac{\partial a}{\partial \lambda_i}\right|_{\lambda^M}(-\lambda_i^M) + \left.\frac{\partial a}{\partial \lambda_{i+1}}\right|_{\lambda^M}\lambda_i^M + \left.\frac{\partial^2 a}{\partial\lambda_i\partial\lambda_{i+1}}\right|_{\lambda^M}(-\lambda_i^M)(\lambda_i^M) \tag{2.22}$$

Recall that since a is initially maximized, $\frac{\partial a}{\partial\lambda_i}|_{\lambda^M} = \frac{\partial a}{\partial\lambda_{i+1}}|_{\lambda^M}$ and hence the first two terms of (2.22) cancel. Now use (2.10) to expand the last term:

$$\Delta a = -\frac{27N^2}{2}(\lambda_i^M)^2\sum_j \lambda_j^M\,\text{area}(P_iP_{i+1}P_j). \tag{2.23}$$

Since all of the P_j are on the same side of the line P_iP_{i+1},

$$\Delta a = -\frac{27N^2}{2}(\lambda_i^M)^2 \operatorname{area}(P_iP_{i+1}R) \tag{2.24}$$

where R is the weighted center of mass of the P_j with weights λ_j^M. Now apply Proposition 2.4.3. We can write

$$\begin{aligned} \Delta a &= -\lambda_i^M \frac{27N^2 \operatorname{area}(P_{i-1}P_iP_{i+1})}{2S \operatorname{area}(P_{i-1}P_iR)} && (2.25)\\ &= -\lambda_i^M a^M \frac{\operatorname{area}(P_{i-1}P_iP_{i+1})}{\operatorname{area}(P_{i-1}P_iR)}. && (2.26) \end{aligned}$$

Since $\sum_i \lambda_i^M = 1$ and $\sum_i \operatorname{area}(P_{i-1}P_iR) = K$, there must be some i for which $\frac{\lambda_i^M}{\operatorname{area}(P_{i-1}P_iR)} \leq \frac{1}{K}$. For such an i,

$$-\frac{\Delta a}{a^M} \leq \frac{\operatorname{area}(P_{i-1}P_iP_{i+1})}{K} \tag{2.27}$$

Note that $\operatorname{area}(P_{i-1}P_iP_{i+1})$ is the amount by which K would decrease if we removed P_i from the toric diagram. Since $\lambda_i^D = 0$, the λ_j^D are a valid choice of weights for the toric diagram with P_i removed. Then

$$-\frac{\Delta a}{a^M} \leq -\frac{\Delta K}{K}. \tag{2.28}$$

Therefore $\frac{a^D}{K+\Delta K} \geq \frac{a^M}{K}$. By Proposition 2.4.2 the local maximum value of a for the new toric diagram is at least as large as a^D. We want to show that it is strictly larger, or equivalently, that λ_j^D do not locally maximize a for the new toric diagram. Recall from Lemma 2.3.14 that a is locally maximized when $\frac{\partial a}{\partial \lambda_{i+1}} = 3a$. Hence a will continue to be maximized only if $\Delta \frac{\partial a}{\partial \lambda_{i+1}} = 3\Delta a$. Once again we use the fact that a is degree one in each individual λ_j:

$$\begin{aligned} \Delta \frac{\partial a}{\partial \lambda_{i+1}} &= \left.\frac{\partial^2 a}{\partial \lambda_i \partial \lambda_{i+1}}\right|_{\lambda^M} (-\lambda_i^M) && (2.29)\\ &= \frac{\Delta a}{\lambda_i^M}. && (2.30) \end{aligned}$$

Hence a can continue to be maximized only if $\lambda_i^M = \frac{1}{3}$. But λ_{i+1}^M is positive (since we chose to let our indices enumerate corner perfect matchings only), so $\lambda_{i+1}^D = \lambda_i^M + \lambda_{i+1}^M > \frac{1}{3}$. By

Equation in [51]	(x,y,z)	$(9-n)(\alpha x^2+\beta y^2+\gamma z^2)$
(1)	$(1,1,1)$	27
(1)	$(1,1,2)$	54
(1)	$(1,2,5)$	270
(2)	$(1,1,1)$	32
(3)	$(1,1,1)$	36
(4)	$(1,2,1)$	50
(5)	$(1,1,1)$	32
(6.1)	$(1,1,1)$	27
(6.2)	$(2,1,1)$	36
(7.1)	$(2,2,1)$	32
(7.2)	$(2,1,1)$	32
(7.3)	$(3,1,1)$	36
(8.1)	$(3,3,1)$	27
(8.2)	$(4,2,1)$	32
(8.3)	$(3,2,1)$	36
(8.4)	$(5,2,1)$	50

Table 2.1: The values of $\frac{54N^2K}{a} = (9-n)(\alpha x^2+\beta y^2+\gamma z^2)$ for some of the quiver gauge theories defined in [51]. Note that the equations in [51] have infinitely many solutions (which can be seen by observing that if we fix one of x, y, z we get a form of Pell's equation), so there exist theories with arbitrarily large $\frac{54N^2K}{a}$.

Lemma 2.3.15, λ_j^D cannot be the local maximum point. By the induction hypothesis, the new $\frac{a}{K}$ is at most $\frac{1}{2}$, so the old $\frac{a}{K}$ must be smaller than $\frac{1}{2}$. □

2.4.2 Comparison to non-toric field theories

Let us consider how we might formulate a similar bound for non-toric CFTs. We need to decide how to interpret K in the non-toric case. If seems natural to replace $2N^2K$ with the sum of the squares of the numbers of colors of each gauge group.

Let's look at the values of $\frac{a}{N^2K}$ for a cone over a del Pezzo surface. Reference [51] lists some quiver gauge theories that are dual to these Calabi-Yaus. In their notation, the sum of the squares of the number of colors is $\alpha x^2+\beta y^2+\gamma z^2$. We can compute a by looking at the AdS dual theory. References [52, 53] tell us that $\frac{\pi^3}{4a}$ is the volume of the horizon, and [54] tells us that the volume of the real cone over dP_n is $\frac{\pi^3(9-n)}{27}$. So $a = \frac{27}{4(9-n)}$. So then $\frac{54N^2K}{a} = (9-n)(\alpha x^2+\beta y^2+\gamma z^2)$, and the bound (2.16) for toric theories is then equivalent to $27 \leq (9-n)(\alpha x^2+\beta y^2+\gamma z^2) < 4\pi^2$. From Table 2.4.2 we see that the toric upper bound on $\frac{54N^2K}{a}$ is not true for all quiver gauge theories. In fact, $\frac{N^2K}{a}$ can be

arbitrarily large. Equation (1) of [51] is $x^2 + y^2 + z^2 = 3xyz$. If we set $z = 1$ then we have a Pell's equation in x and y and there are infinitely many solutions. On the other hand, $27 \leq (9 - n)(\alpha x^2 + \beta y^2 + \gamma z^2)$ still holds for all of the theories considered in [51]. It would be interesting to know if the inequality holds more generally.

2.5 Merging zigzag paths

2.5.1 Deleting an edge of the dimer

Theorem 2.3.3 says that, if a dimer is properly ordered, then we can determine its toric diagram from the windings of its zigzag paths. As we mentioned in section 2.3.3, Hanany and Vegh [10] and Stienstra [12] have previously made proposals for drawing a dimer with given zigzag winding numbers, but their procedures are impractical for large dimers because of the large amount of trial and error required.

Partial resolution [40–43] has previously been suggested as a method of determining the dimer from the quiver [7, 42, 55]. It involves starting with a toric diagram whose dimer model is known and introducing Fayet-Iliopoulos terms that Higgs some of the fields and remove part of the toric diagram to create a new diagram. However, as is the case with the Fast Inverse Algorithm, the previous proposals involving partial resolution suffered from being computationally infeasible.

In this section, we will explore how certain operations on the dimer affect its zigzag paths. These operations can be interpreted as partial resolutions. We will later use these operations to construct an algorithm for drawing a properly ordered dimer with given winding numbers that requires no trial and error.

One operation that we can perform is to remove an intersection of two zigzag paths (or equivalently, delete an edge of the dimer). The operation has the effect of merging the two paths into a single path. An example is shown in figure 2.5.1. In physical terms, we are Higgsing away the edge by turning on Fayet-Iliopoulos parameters for the adjacent faces. This is an example of partial resolution of the toric singularity [40–43]. We will always merge paths that intersect just once. In the following we will sometimes assume that the windings of the paths are $(1, 0)$ and $(0, 1)$; any other case is $SL_2(\mathbb{Z})$ equivalent to this one.

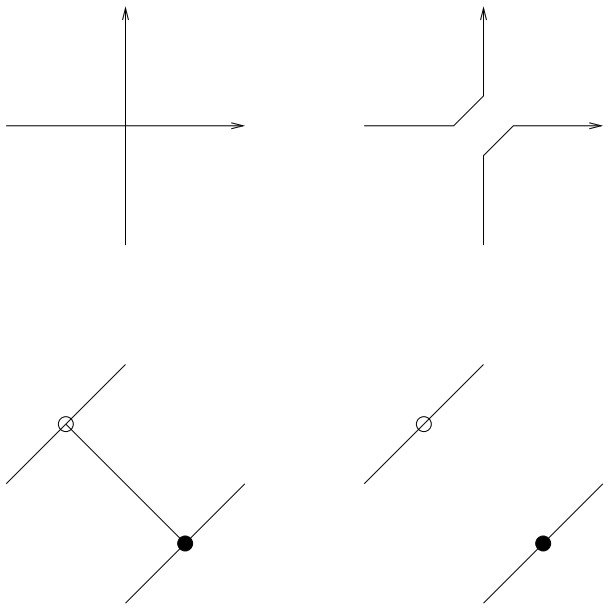


Figure 2.6: Top: Merging two zigzag paths by deleting the intersection between them. Bottom: The effect on the dimer.

2.5.2 Making multiple deletions

Suppose we want to make $n > 1$ $(1,1)$ edges from $(1,0)$ and $(0,1)$ edges. If we make them one at a time, then we would violate the proper ordering of nodes because we would have $(1,1)$ paths intersecting each other. We should instead delete all n^2 edges between the n $(1,0)$ edges and the n $(0,1)$ edges. We will refer to this procedure as Operation I.

2.5.3 Extra crossings

We mentioned in section 2.2 that the number of oriented crossings between a pair of paths is a function only on their windings. The number of unoriented crossings is greater than or equal to the absolute value of the number of oriented crossings. If equality does not hold then we say that the pair of paths has "extra crossings" . We say that a diagram has extra crossings if any pair of its paths does. There is nothing inherently wrong with extra crossings, but we may find it desirable to produce diagrams without them.

The edge deletion procedure mentioned in the previous section sometimes introduces extra crossings. An example of this is shown in Figure 2.5.3. We combine a $(1,0)$ zigzag path and a $(0,1)$ to make a $(1,1)$ zigzag path, and we also combine $(-1,0)$ and $(0,-1)$

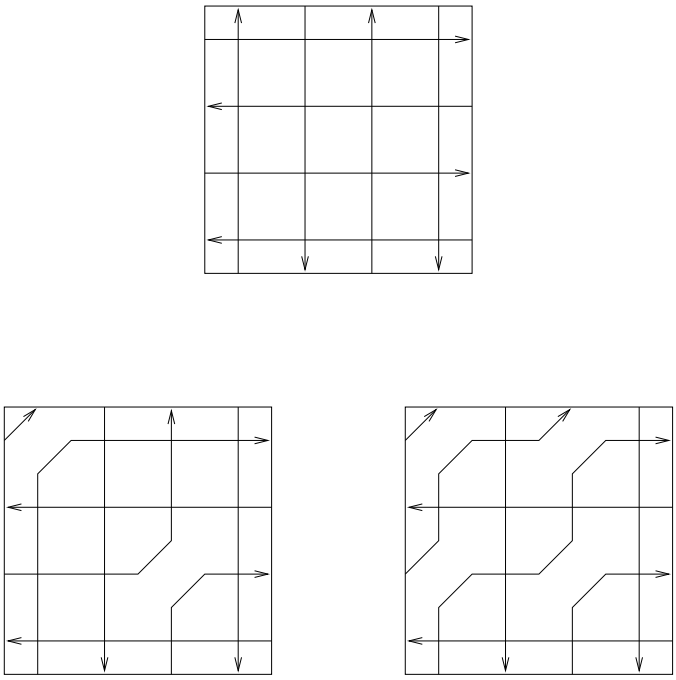


Figure 2.7: Left: An incorrect way of making two $(1, 1)$ paths. They intersect each other, which implies that the adjacent nodes are not properly ordered. Right: The correct way of making two $(1, 1)$ paths.

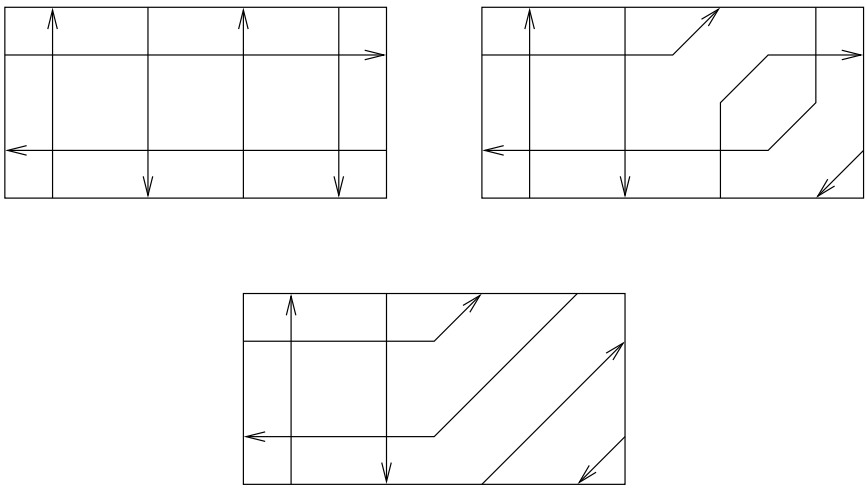


Figure 2.8: The top left diagram has no extra crossings. The top right diagram shows what happens when some zigzag paths are merged. The two diagonal paths now have extra crossings with each other. The bottom diagram shows what happens when we move the two zigzag paths past each other; they no longer intersect.

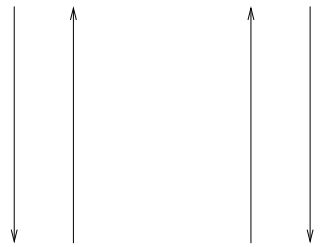


Figure 2.9: Pairs of paths that are positively and negatively oriented, respectively.

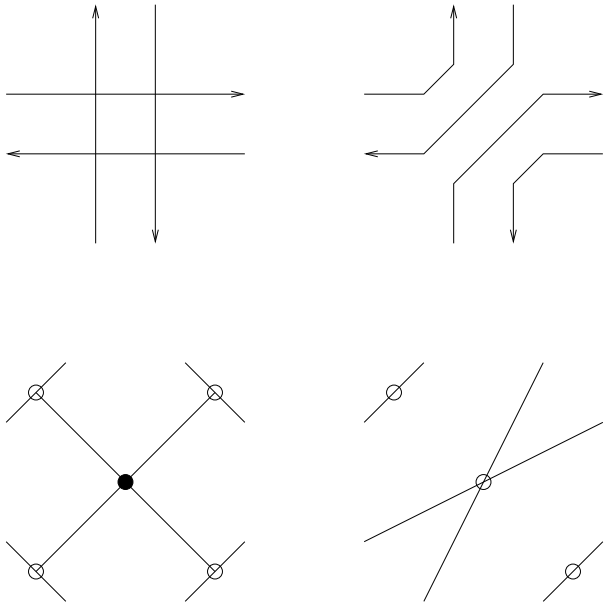


Figure 2.10: Making $(1,1)$ and $(-1,1)$ paths from horizontal and vertical paths in the zigzag path picture and the dimer picture.

paths to make a $(-1,-1)$ path. The $(1,1)$ path and $(-1,-1)$ path have a positively oriented intersection coming from the $(0,1)-(-1,0)$ intersection and a negatively oriented intersection coming from the $(1,0)-(0,-1)$ intersection. Note that we can get rid of these crossings by moving the two paths past each other. In terms of the dimer, moving the paths past each other merges the two vertices adjacent to a valence two node. Physically, we are integrating out a mass term.

We define a pair of zigzag paths to be an "opposite pair" if they have opposite winding numbers, they do not intersect, and they bound a region containing no crossings. Also, we define the orientation of an opposite pair to be positive if the area containing no crossings is to the left of an observer traveling along one of the paths, and negative if the area is on the right. (See figure 2.5.3.) We have just seen how to take a pair negatively oriented

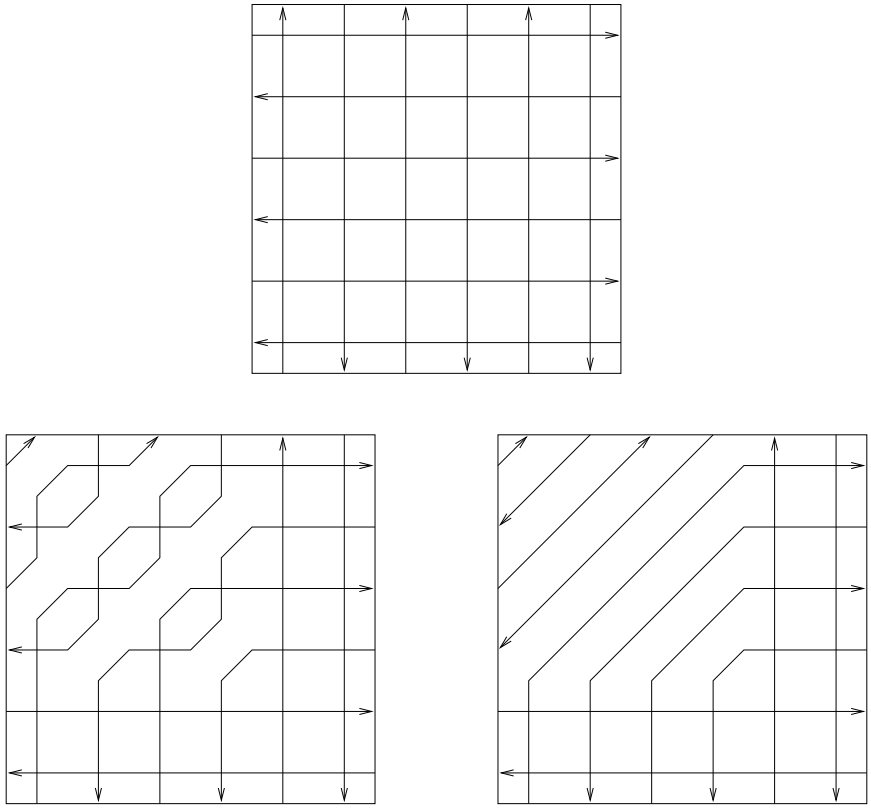


Figure 2.11: Creating two $(1,1)$ and two $(-1,-1)$ paths from horizontal and vertical paths. We first merge horizontal and vertical paths to create diagonal paths, then move the diagonal paths past each other.

horizontal paths and a pair of negatively oriented vertical paths and turn them into a pair of negatively oriented diagonal paths. Similarly we can turn a pair of positively oriented horizontal paths and a pair of positively oriented vertical paths into a pair of positively oriented diagonal paths. In terms of dimers, this operation takes a node of valence four, deletes two opposite edges, and merges the other endpoints of the two remaining edges. Figure 2.5.3 shows the operation in terms of both zigzag paths and dimers.

More generally, we can make n $(1,1)$ paths and n $(-1,-1)$ paths and get rid of their crossings. An example is given in figure 2.5.3. We have to untangle each $(1,1)$ path from each $(-1,-1)$ path. Note that all $2n$ paths must have the same orientation. We will call this procedure Operation II.

If we want to create differing numbers of $(1,1)$ and $(-1,-1)$ paths, then we run into the problem that we cannot pair them all. We will need to do something more complicated. Let m be the number of $(1,1)$ paths we want to make, and let n be the number of $(-1,-1)$ paths

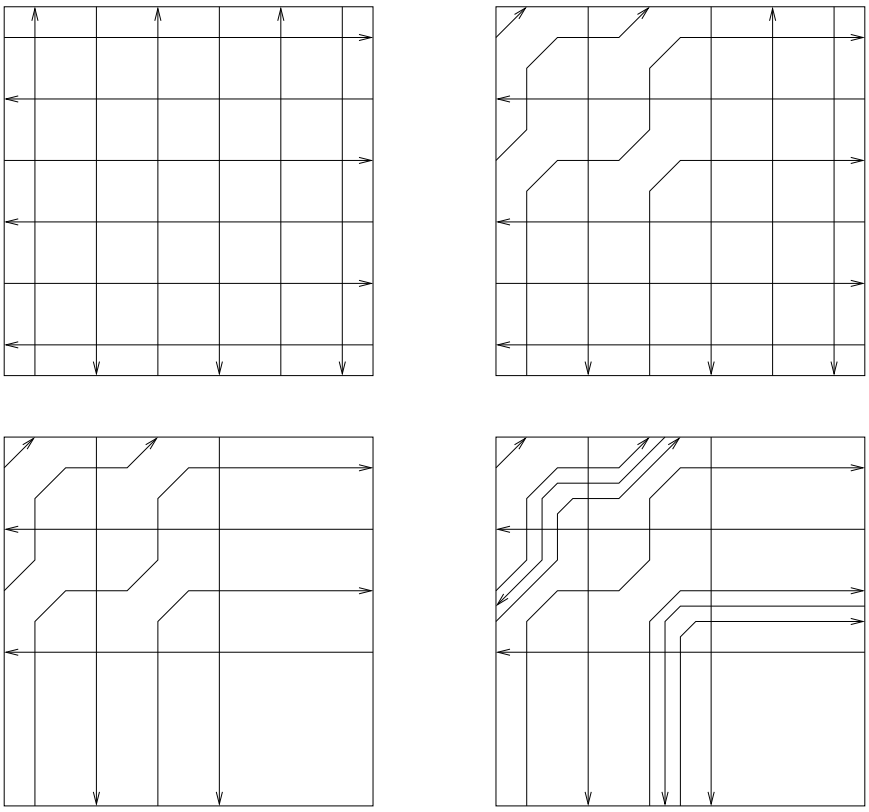


Figure 2.12: Creating three $(1,1)$ paths and one $(-1,1)$ path without introducing extra crossings. We first make two $(1,1)$ paths and then add a $(1,1)-(-1,-1)$ pair that follows one of those two paths.

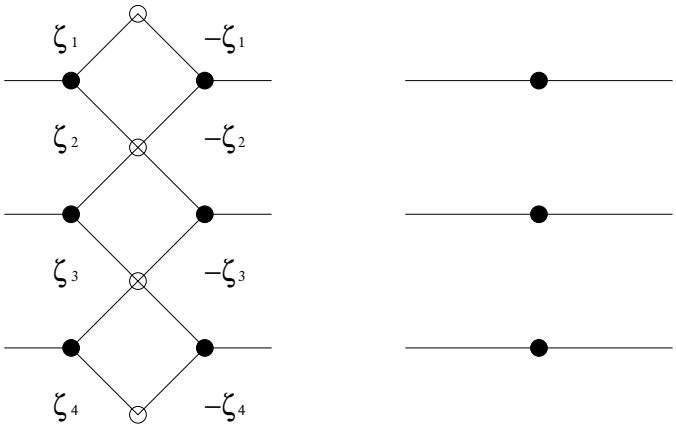


Figure 2.13: Left: a dimer with a pair of adjacent opposite zigzag paths. Right: the dimer with the paths removed. In physical terms, we are introducing performing a partial resolution by introducing Fayet-Iliopoulos parameters for the faces on either side of the diamonds [40–43]. (In particular note that the resolution of the double conifold in [43] is an example of this operation.) For each pair of faces that meet at one of the points in the middle, their FI parameters should sum to zero. All parameters on the left should have the same sign.

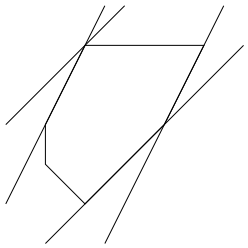


Figure 2.14: Some tangent lines to a convex polygon.

we want to make. Assume $m > n$. We first make $m - n$ $(1, 1)$ paths. Now we completely remove n pairs of adjacent $(1, 0)$ and $(-1, 0)$ paths and n pairs of adjacent $(0, 1)$ and $(0, -1)$ paths. Because the pairs are adjacent, the condition that intersection orientations alternate along a path is preserved. Now we want to insert n pairs of adjacent $(1, 1)$ and $(-1, 1)$ paths, and we want to make sure that there are no extra crossings. This can be accomplished by making them follow one of the $m - n$ already existing $(1, 1)$ paths. An example is given in figure 2.5.3. Figure 2.5.3 shows what removing or adding a pair of zigzag paths does to the dimer. This procedure will be called Operation III.

2.6 An efficient inverse algorithm

2.6.1 Description of the algorithm

In describing the algorithm we find it useful to draw toric diagrams rotated 90 degrees counterclockwise from their usual presentation. Our convention will make the algorithm easier to visualize, because it makes the windings of the zigzag paths equal to, rather than perpendicular to, the vectors of the toric diagram edges.

Let X be a toric diagram for which we would like to construct a dimer. Let Y be the smallest rectangle with horizontal and vertical sides that contains X. Since Y represents an orbifold of the conifold, we know a dimer for Y. We will modify this dimer until we get a dimer for X.

Before we begin, we need to make the following definition. A tangent line to a convex polygon P is a line ℓ such that $\ell \cap P \subseteq \partial P$ and $\ell \cap P \neq \emptyset$. Note that a convex polygon has exactly two tangent lines with a given slope.

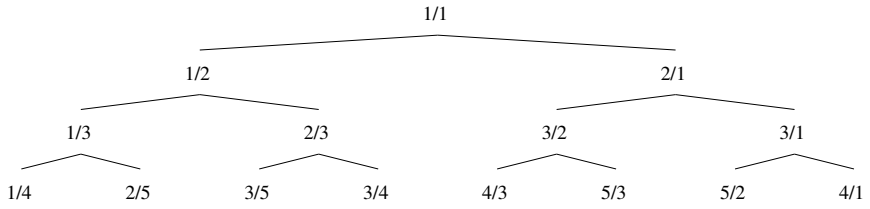


Figure 2.15: The Farey tree tells us the order in which to make zigzag paths. For example, in order to make $(3, 4)$ zigzag paths we first make $(1, 1)$ zigzag paths, then $(1, 2)$ paths, then $(2, 3)$ paths.

We begin by finding the slope one tangent lines to X and cutting Y along these lines to produce some $(1, 1)$ and $(-1, -1)$ paths. We use Operation I if the number of $(1, 1)$ or $(-1, -1)$ paths desired is zero, Operation II if the numbers are equal, and Operation III if the numbers are both nonzero and unequal. Next we want to cut along the slope $1/2$ tangent lines to X to produce $(2, 1)$ and $(-2, -1)$ paths. In fact we already know how to do this, because $SL_2(\mathbb{Z})$ equivalence reduces the problem of making $(2, 1)$ and $(-2, -1)$ paths from $(1, 0)$, $(-1, 0)$, $(1, 1)$, and $(-1, -1)$ paths to the problem of making $(1, 1)$ and $(-1, 1)$ paths from $(1, 0)$, $(-1, 0)$, $(0, 1)$ and $(0, -1)$ paths. Hence we can now cut Y along the slope $1/2$ tangent lines to X. Similarly, we can cut Y along the slope 2 tangent lines to X. After this, we can make $(3, 1)$ paths by combining $(1, 0)$ and $(2, 1)$ paths, $(3, 2)$ paths by combining $(1, 1)$ and $(2, 1)$ paths, etc. We can eventually make paths of all slopes, with the order in which we make the paths determined by the Farey tree. (See figure 2.6.1.) We can then enumerate over all negative slopes, starting with -1. When we are finished, we will have a dimer for X.

Figure 2.16 shows an example case of the algorithm.

2.6.2 Proof of the algorithm

We need to prove that we have the paths necessary to perform each step, and that the finished dimer has properly ordered nodes and has no extra crossings.

Theorem 2.6.1. *At each step of the algorithm, the following are true:*

1. *If there are m zigzag paths with winding (a, b) and n zigzag paths with winding $(-a, -b)$, then there are $\min(m, n)$ negatively oriented pairs of (a, b) and $(-a, -b)$*

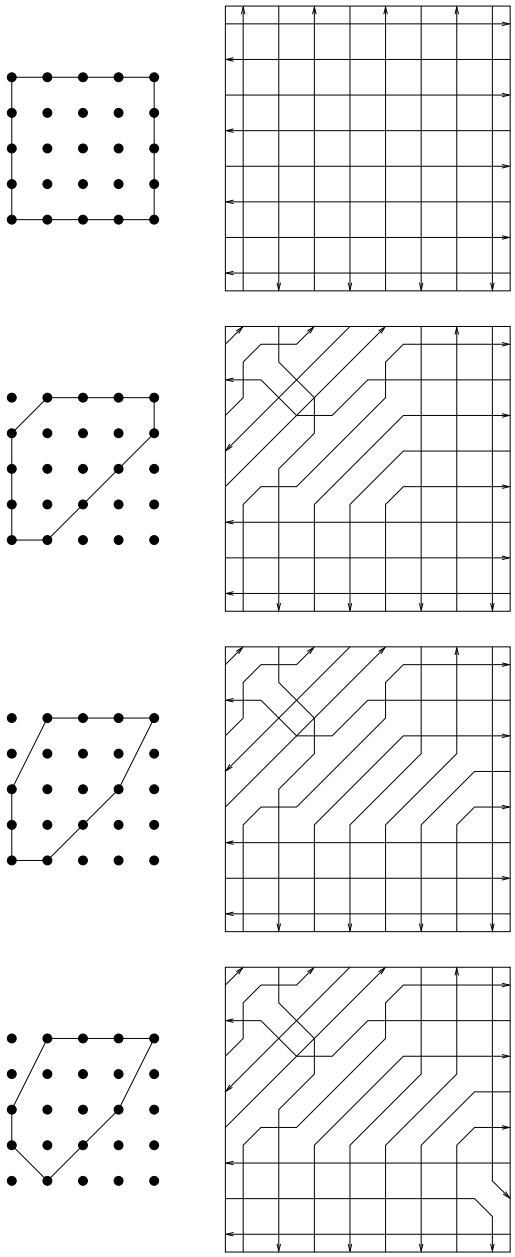


Figure 2.16: An example of the algorithm. Note that the cut made in the second diagram is the same as that of figure 2.5.3, although we have drawn it a little differently to make the spacings more equal.

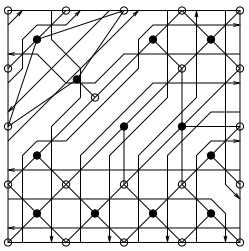
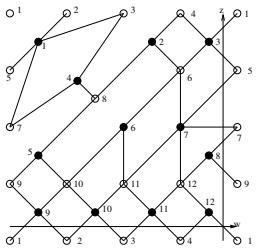


Figure 2.17: The dimer corresponding to the final zigzag path diagram in figure 2.16.



$$\begin{pmatrix}
0 & 0 & z & 0 & 0 & 0 & 0 & 0 & -w & 0 & 0 & wz \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & w & -w & 0 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & w & -w & 0 \\
0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & w & -w \\
-1 & 0 & z & 0 & 0 & 0 & z & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & -1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & -1 & 0 & 0 & z & z & 0 & 0 & 0 & 0 \\
0 & -1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -1 & 0 & 0 & z & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 1 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 1 & 1
\end{pmatrix}$$

$$\det = (w^2 - w)z^4 + (-w^4 - 37w^3 - 137w^2 - 35w - 1)z^3 + (3w^4 - 175w^3 + 146w^2 - 2w)z^2 + (-3w^4 - 40w^3 - w^2)z + w^4$$

Figure 2.18: The dimer corresponding to the final zigzag path diagram in figure 2.16 and its Kasteleyn matrix. The rows represent white nodes and the columns represent black nodes.

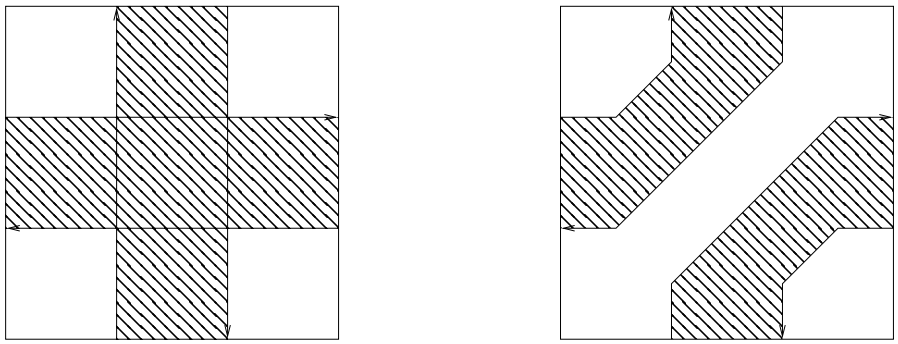


Figure 2.19: Left: We start out with two negatively oriented pairs of opposite paths. The shaded regions are free of crossings. Right: The regions formed by the merged pairs are still free of crossings.

paths. (This condition ensures that we can always perform the next step of the algorithm.)

2. *There are no extra crossings.*

3. *All nodes are properly ordered.*

Proof. It is clear that all of these conditions hold for the initial dimer. Now let's look at whether the first condition will be preserved. Operation I will preserve the condition for the winding of the paths being merged provided that we merge unpaired paths when possible. It will also satisfy the condition for the windings of the newly created paths since there are no $(-a, -b)$ paths. Operation II will preserve condition 1 for the windings of the paths being merged since it only deletes negatively oriented pairs. Figure 2.6.2 illustrates why Operation II creates negatively oriented pairs of opposite paths. For Operation III we should again merge unpaired paths when possible. It is clear that the reinserted paths form pairs, and we can make these pairs negatively oriented if we desire.

Now consider whether extra crossings are introduced. Let the windings of the paths being merged be (a, b), $(-a, -b)$, (c, d), and $(-c, -d)$, where $ad - bc = 1$. A path of winding (e, f) will have extra crossings with the new $(a + c, b + d)$ paths if $af - be$ and $cf - de$ have opposite signs. Equivalently, there will be extra crossings if f/e is between b/a and d/c. But because of the Farey fraction ordering, there are no windings (e, f) with this property. So extra crossings are not introduced.

Finally consider whether proper ordering is preserved. Again let the windings of the paths being merged be (a, b), $(-a, -b)$, (c, d), and $(-c, -d)$, $ad - bc = 1$. In Operation I, some nodes will see an (a, b) path or a (c, d) path become an $(a + c, b + d)$ path. Therefore proper ordering is preserved provided there are no windings between (a, b) and (c, d). This is always the case because of the Farey fraction ordering. In Operation II, in addition to deletion we also need to move paths past each other. Some nodes are deleted and the others remain unchanged, so proper ordering is preserved. In Operation III, the process of making the lone paths is the same as Operation I, so it preserves proper ordering. Removing pairs also preserves proper ordering. Inserting pairs of paths preserves proper ordering if each intersection between a path in the pair and another path has the same sign as their crossing number, i. e. the paths in the pair do not have extra crossings. Since we are inserting them along an existing path, they will not have extra crossings if the existing path does not have any. We have already showed that we never introduce extra crossings. □

2.6.3 Allowing extra crossings

If we want to produce diagrams with extra crossings, we can always just skip the steps for removing the extra crossings. When we want to create (a, b) and $(-a, -b)$ paths, we just perform Operation I twice. There is one potential issue in that we have always assumed that the zigzag paths that we join have just one crossing. We always join paths with oriented crossing number ± 1, but now the unoriented crossing number can be larger than the absolute value of the oriented number. But we recall that the only extra crossings we create are between paths with windings of the form (a, b) and $(-a, -b)$. We may later merge these paths with some other paths, but the extra crossings will always be between paths with oppositely signed x-coordinates and oppositely signed y-coordinates. We never merge such pairs of paths.

2.6.4 The number of independent solutions to the R-charge equations

We now exhibit the dimers required by Lemma 2.3.10.

Lemma 2.6.2. *The algorithm described in section 2.6.3 produces dimers for which the set of all solutions to equations (2.1) and (2.2) has dimension equal to the number of zigzag paths minus one.*

Proof. Our proof is by induction. Our algorithm starts with a dimer that is a diamond-shaped grid. We denote the position of an edge in the grid by (i, j). We can see (e. g. by Fourier analysis) that the general solution to (2.1) and (2.2) is $\frac{1}{2} + (-1)^i f(j) + (-1)^j g(i)$ for arbitrary functions f, g. The number of independent solutions is the number of rows plus the number of columns minus one (the minus one come from the fact that $f(j) = (-1)^j, g(j) = -(-1)^i$ produces the same solution as $f(j) = 0, g(j) = 0$), which is the number of zigzag paths minus one.

Now consider what happens when our algorithm deletes an edge of the toric diagram. If we have a solution to the equations (2.1) and (2.2) in the new dimer, we can construct a solution to the equations in the old dimer by assigning a value of zero to the deleted edge. Conversely, if we have a solution in the old dimer in which the deleted edge has value zero, then we have solution in the new dimer as well. We know that there exists a solution in the old dimer where the deleted edge is nonzero, since the deleted edge is contained in some boundary perfect matching. So deleting the edge reduces the dimension of the solution space of (2.1) and (2.2) by one, and also reduces the number of zigzag paths by one.

□

2.7 Conclusions

We showed that dimers that have the number of faces predicted by the AdS dual theory and that have valence one nodes will have many nice properties: they are "properly ordered", their cubic anomalies are in agreement with the Chern-Simons coefficients of the AdS dual, gauge-invariant chiral primary operators satisfy the unitarity bound, corner perfect matchings are unique, and zigzag path windings are in one-to-one correspondence with the (p, q)-legs of the toric diagram.

We derived some simple bounds for the cubic anomaly a in terms of the area of the toric diagram (and hence in terms of the number of gauge groups).

We provided a precise, computationally feasible algorithm for producing a dimer model for a given toric diagram based on previous partial resolution techniques and the Fast Inverse Algorithm.

Much of our work on dimer models is also applicable to the study of M2-branes. Given a dimer and a four-dimensional CFT dual to Type IIB string theory $AdS_5 \times X_5$, we can dimensionally reduce to three dimensions and add Chern-Simons terms to the gauge fields. The resulting CFT is dual to M-theory on $AdS_4 \times X_7$, where X_7 is a fibration over X_5 [29–31]. It would be interesting to see if our results could also apply to the three-dimensional dimers discussed in [56] and the orientifold dimers discussed in [57].

Since this chapter was originally published as ref. [39], Ishii and Ueda [58] devised a more general version of the algorithm.

Chapter 3

From Necklace Quivers to the F-theorem, Operator Counting, and $T(U(N))$

This chapter is an edited version of ref. [1], written in collaboration with Christoper Herzog and Silviu Pufu.

3.1 Introduction

Exact results in strongly-interacting field theories are generally rare. In supersymmetric field theories, supersymmetry places strong constraints on various properties of chiral operators, and exact results pertaining to these operators might be possible even at strong coupling. For three-dimensional superconformal theories, recent progress in finding such exact results that hold at any coupling was made in [17,25,59], where the partition function of superconformal theories on S^3 with $\mathcal{N} \geq 2$ supersymmetry, as well as the expectation values of certain BPS Wilson loops, were reduced from path integrals to finite-dimensional multi-matrix integrals. This major simplification was achieved through the localization technique developed in [18] for four-dimensional theories.

A consequence of this work is the realization that the "free energy" F defined as minus the logarithm of the path integral on S^3,

$$F = -\log |Z_{S^3}| \ , \tag{3.1}$$

with appropriate subtractions of power law divergences, might represent a good measure of the number of degrees of freedom in any field theory, supersymmetric or non-supersymmetric. One way in which F can be thought of as a measure of the effective number of degrees of freedom is the theorem first conjectured in [21] and proved in [23] that F decreases along renormalization group (RG) flows and is stationary at RG fixed points. This assertion was called the "F-theorem" in [21]. It suggests that F is be a 3-d analog of the central charge c from two-dimensional field theory, which is known to have the same monotonicity property along RG flows [60], and it would resemble the Weyl anomaly coefficient a from 4-d theories, which is also believed to decrease along RG trajectories [19]. Actually, the free energy F also resembles a in another way: just like a, F can be used to find the exact R-symmetry in the infrared (IR) by computing F as a function of a set of trial R-charges and then maximizing it [25, 26]. The analogous procedure in 4-d theories is called "a-maximization" [24].

In the context of the AdS/CFT correspondence [3–5], the free energy F can also be computed holographically from the gravity side of the correspondence. In particular, for a CFT dual to AdS_4 of radius L and effective four-dimensional Newton constant G_N, F is given by [61]

$$F = \frac{\pi L^2}{2G_N} \, . \tag{3.2}$$

It was shown in [62] that in any CFT F can also be interpreted as an entanglement entropy between a disk and its complement in the $\mathbb{R}^{2,1}$ theory. In turn, this entanglement entropy equals the holographic a_* function defined in [63, 64] that was shown to always decrease along holographic RG flows.

If the AdS_4 background mentioned above arises as a Freund-Rubin compactification $AdS_4 \times Y$ of M-theory, where Y is a seven-dimensional Sasaki-Einstein space threaded by N units of four-form flux, then the quantization of the AdS_4 radius in Planck units implies that at large N eq. (3.2) becomes [27]

$$F = N^{3/2} \sqrt{\frac{2\pi^6}{27\,\mathrm{Vol}(Y)}} + o(N^{3/2})\,. \tag{3.3}$$

Here, the volume of Y is computed with an Einstein metric that satisfies the normalization condition $R_{mn} = 6g_{mn}$. The Freund-Rubin solution $AdS_4 \times Y$ arises as the near-horizon limit of a stack of N M2-branes placed at the tip of the Calabi-Yau cone X over Y. The $N^{3/2}$ behavior of the number of degrees of freedom had been known for quite some time, as the same large N dependence appears in other quantities such as the thermal free energy that was computed in [65] more than ten years ago. A field theory explanation of this peculiar large N dependence had been lacking until recently, mostly because explicit Lagrangian descriptions of the field theories living on coincident M2-branes have been discovered only in the past few years. Starting with ABJM theory [28] that describes N M2-branes sitting at an orbifold singularity of $\mathbb{C}^4$, there are now many Chern-Simons matter $U(N)$ gauge theories that are proposed to describe the effective dynamics on N M2-branes placed at the tip of various Calabi-Yau cones (see for example [31–33, 66, 67]). However, only few of these dualities have been extensively tested. Extensive tests are difficult to perform because supergravity on $AdS_4 \times Y$ is supposed to be a good approximation to the dynamics of the CS-matter gauge theories only as one takes the gauge group ranks to infinity while keeping the CS levels fixed. In this limit, the 't Hooft coupling N/k becomes large, and there are no perturbative computations that one can perform.

That in $\mathcal{N} \geq 2$ theories one can write F exactly in terms of a matrix integral means that by evaluating this integral one can test some of these AdS_4/CFT_3 dualities that have been put forth in recent years. In particular, one can provide a field theory derivation of the $F \propto N^{3/2}$ large N dependence of the number of degrees of freedom (3.3) on N coincident M2-branes. Moreover, one can compare the coefficient of $N^{3/2}$ in (3.3) to the volume of the internal space Y that one can compute independently by integrating the square root of the

determinant of the Sasaki-Einstien metric on Y. Such comparisons were made in [68] in the case of ABJM theory, in [27,69] for a large class of theories with $\mathcal{N} = 3$ supersymmetry, and in [21,70,71] for many theories with $\mathcal{N} = 2$ supersymmetry. In [68,69] F was computed as a function of the 't Hooft coupling, while the other works focused only on the strong coupling regime and used the method developed in [27] to evaluate the multi-matrix integrals at large N and fixed CS levels in a saddle point approximation.

In this chapter, we build upon the work in [21, 27] in several ways. Our first main result consists of an infinite class of RG flows whose IR and ultraviolet (UV) fixed points preserve $\mathcal{N} = 3$ SUSY. Via the AdS/CFT correspondence, at large N such a flow is dual to a holographic RG flow between two $AdS_4 \times Y$ extrema of 11-d supergravity. That the extrema preserve $\mathcal{N} = 3$ supersymmetry means that the cones over the spaces Y are hyperkähler and, because of that, the spaces Y are called tri-Sasakian. (We will provide a more detailed exposition of hyperkähler spaces at the beginning of section 3.2.1.) The 11-d SUGRA solutions that represent M2-branes sitting at the tip of a hyperkähler cone were constructed in [72], where they were also related through string duality to a brane construction in type IIB string theory. The type IIB brane construction consists of N D3-branes filling the 0126 directions, and a sequence of (p_a, q_a) five-branes,[1] $1 \leq a \leq d$, filling the 012 directions and sitting at fixed angles in the 37-, 48-, and 59-planes. The (p, q) branes break the D3-brane stack into d segments, and the three-dimensional $U(N)^d$ CS-matter gauge theories then live on the segments. The $\mathcal{N} = 3$ RG flows we want to study correspond to removing one of the (p, q) branes or to a (p_1, q_1) brane and a (p_2, q_2) brane combining into a $(p_1 + p_2, q_1 + q_2)$ brane. We find that the F-theorem is satisfied in these examples.

Our results for the F-theorem follow most naturally from the gravity side of the AdS/CFT correspondence which provides an $SL(2, \mathbb{R})$ invariant result for F as a function of arbitrary (p_a, q_a). We provide field theory confirmation of the gravity result in the cases where $p_a = 0$ or $p_a = 1$ for each a. While the 3-d field theory is simplest when $p_a = 0$ or 1, for $p_a > 1$, ref. [73] provides a more complicated field theory description involving an

[1] We adopt the convention where by a (p, q) five-brane we mean a brane with p units of NS5 charge and q units of D5 charge.

interpolating $T(U(N))$ theory that implements a sort of local S duality. Instead of using the $T(U(N))$ theory, we derive the matrix model for $p_a > 1$ by bootstrapping from our large N results. We check that the matrix model yields the correct answer in the large N limit, and also that it is invariant under $SL(2,\mathbb{Z})$. As a bonus, we discover the matrix model of the $T(U(N))$ theory. This partition function has also been found independently by [74, 75]. The matrix model for arbitrary p_a is our second main result.

Along the way we tie several loose ends left off from [27]. In [27] the matrix model was solved explicitly only for theories on $(1, q_a)$ branes with $a \leq 4$. We provide a general solution that holds for any number of (p_a, q_a) branes. In [27] it was checked numerically in a few examples that the M-theory prediction for the volume of Y that can be extracted from (3.3) agrees with the geometric computation performed by Yee [76]. We build on Yee's work and prove that in general the two computations of $\mathrm{Vol}(Y)$ agree. In [27] it was conjectured that the volume $\mathrm{Vol}(Y)$ can be expressed in terms of a certain sum over trees. We provide a proof of this tree formula.

Our third main result, contained in Section 3.3, is a relationship between the eigenvalues in the matrix model and the number of chiral operators in the field theory. Define $\psi(r)$ to be the number of chiral operators with R-charge smaller than r for the $N = 1$ gauge theory. The authors of [54, 77] demonstrated that there is a relationship between $\psi(r)$ and the volume of the Sasaki-Einstein space in the large r limit. Given (3.3), there must also be a relationship between $\psi(r)$ and F. In fact, as we show in this chapter for the necklace theories, more precise relationships can be established between the matrix model and operator counting problems. The operators in the necklace theories also have a monopole charge m corresponding to m flux units in a diagonal subgroup of the gauge group. Thus, we may consider $\psi(r, m)$ to be the number of operators with R-charge less than r and monopole charge less than m. Let X_{ab} be a hypermultiplet transforming under the fundamental of the bth gauge group and the antifundamental of the ath gauge group. We can consider $\psi_{X_{ab}}(r, m)$ to be defined as above but now with the operator $X_{ab} = 0$. In the large N limit, the matrix model is solved by a saddle point approximation for which the eigenvalues are complex numbers $\lambda_a = N^{1/2}x + iy_a$. The eigenvalues can be parametrized by an eigenvalue density $\rho(x)$, which turns out to be the same for each gauge group, and an imaginary part

$y_a(x)$. Our two results are

$$\left.\frac{\partial^3\psi}{\partial r^2\partial m}\right|_{m=rx/\mu} = \frac{r}{\mu}\rho(x) \ , \tag{3.4}$$

$$\left.\frac{\partial^2\psi_{X_{ab}}}{\partial r\partial m}\right|_{m=rx/\mu} = \frac{r}{\mu}\rho(x)[y_b(x) - y_a(x) + R(X_{ab})] \ , \tag{3.5}$$

where $\mu = 3F/4\pi N^{3/2}$. (Here we take the liberty of replacing the operator counts, which are discrete functions, with continuous approximations.) We believe these relations will hold more generally.

3.2 Volumes of Toric tri-Sasaki Einstein Spaces

On the gravity side of the AdS/CFT duality, we have M-theory backgrounds generated by placing a stack of N M2-branes at the tip of a hyperkähler cone, with N large. In this section our aim is to compute the free energy of the M2-brane theory purely from the supergravity side of the correspondence. We start by introducing in section 3.2.1 the main ingredients in constructing the 11-d supergravity solution, namely the toric hyperkähler spaces. In section 3.2.2 we build upon the results of Yee [76] and express the volume of these spaces in terms of the volume of a certain polygon for which we will provide a field theory interpretation later on. In section 3.2.3 we comment on some field theory implications of this formula, and show explicitly that the free energy decreases along certain RG flows, in agreement with the F-theorem proposed in [21].

3.2.1 Toric Hyperkähler Cones from a Quotient Construction

We start by introducing the toric hyperkähler cones. The following discussion draws heavily from [78] and [76].

A hyperkähler manifold possesses $4n$ real dimensions and has a Riemannian metric g which is kähler with respect to 3 anti-commuting complex structures J_1, J_2, and J_3. These J_i furthermore satisfy the quaternionic relations $J_1^2 = J_2^2 = J_3^2 = J_1J_2J_3 = -1$. The simplest example of a hyperkähler manifold is the four-dimensional space of quaternions $\mathbb{H} \cong \mathbb{R}^4$, endowed with the standard line element. A single quaternion $q \in \mathbb{H} \cong \mathbb{R}^4$ can be

represented as a two-by-two complex matrix

$$q = \begin{pmatrix} u & v \\ -\bar{v} & \bar{u} \end{pmatrix} \tag{3.6}$$

parametrized in terms of two complex variables u and v. In terms of u and v the metric is $ds^2 = |du|^2 + |dv|^2$ and the three kähler forms are

$$\omega^3 = -\frac{i}{2}(du \wedge d\bar{u} + dv \wedge d\bar{v}) \ , \tag{3.7}$$

$$(\omega^1 - i\omega^2) = i(du \wedge dv) \ . \tag{3.8}$$

The ω^a transform as a triplet under the $SU(2)$ symmetry that acts as left multiplication on q.

The quaternions $\mathbb{H}$ can also be written as a $U(1)$ bundle over $\mathbb{R}^3$ where the $U(1)$ fiber shrinks to zero size at the origin of $\mathbb{R}^3$. This description comes from the uplift of the Hopf fibration from S^3 to $\mathbb{R}^4$ and makes explicit an $SO(3) \times U(1)$ subgroup of the $O(4)$ rotational symmetry of $\mathbb{R}^4$. More explicitly, if one writes

$$u = \sqrt{r}\cos(\theta/2)\, e^{-i\phi/2+i\psi} \ , \quad v = \sqrt{r}\sin(\theta/2)\, e^{-i\phi/2-i\psi} \ , \tag{3.9}$$

then the standard line element on $\mathbb{H}$ becomes

$$ds^2 = du\, d\bar{u} + dv\, d\bar{v} = \frac{1}{4}\frac{d\vec{r}^2}{r} + r\left(d\psi - \frac{1}{2}\cos(\theta)d\phi\right)^2 \ , \tag{3.10}$$

where $\vec{r} = (r\sin\theta\cos\phi, r\sin\theta\sin\phi, r\cos\theta)$ is the usual parameterization of $\mathbb{R}^3$ in terms of spherical coordinates. The coordinate $\psi \in [0, 2\pi)$ parameterizes the $U(1)$ Hopf fiber. From eq. (3.6) we see that rotations in the $U(1)$ fiber correspond to phase rotations of q.

Another example of a hyperkähler manifold is the Cartesian product of d-copies of the quaternions $\mathbb{H}^d \cong \mathbb{R}^{4d}$, also considered with the flat metric

$$ds^2 = \sum_{a=1}^{d}\left[\frac{1}{4}\frac{d\vec{r}_a^2}{r_a} + r_a\left(d\psi_a - \frac{1}{2}\cos(\theta_a)d\phi_a\right)^2\right] \ . \tag{3.11}$$

Starting with $\mathbb{H}^d$, one can construct construct a large number of further examples of toric hyperkähler spaces using the hyperkähler quotient procedure of Hitchin, Karlhede, Lindström, and Roček [79]. A toric hyperkähler manifold can be defined to be a hyperkähler quotient of $\mathbb{H}^d$ for some integer d by a $(d-n)$-dimensional subtorus $N \subset T^d$. The underlying reason that the quotient continues to be hyperkähler is that the triplet of Kähler forms (3.7) is invariant under the $U(1)$ action that sends $u \to ue^{i\alpha}$ and $v \to ve^{-i\alpha}$.

The hyperkähler quotient procedure requires the data of how $N = T^{d-n}$ sits inside T^d. The inclusion $N \subset T^d$ can be described by the short exact sequence of tori

$$0 \longrightarrow T^{d-n} \xrightarrow{i} T^d \xrightarrow{\pi} T^n \longrightarrow 0\,, \tag{3.12}$$

where we also introduced the quotient $T^n \cong T^d/T^{d-n}$, as well as the inclusion map i and the projection map π. Each torus T^k in this sequence can be identified with $\mathbb{R}^k/(2\pi\mathbb{Z}^k)$, so the data of how N sits inside T^d can be encoded in how the standard lattice $\mathbb{Z}^{d-n} \subset \mathbb{R}^{d-n}$ sits inside $\mathbb{Z}^d \subset \mathbb{R}^d$, or equivalently how $\mathbb{Z}^d \subset \mathbb{R}^d$ projects down to $\mathbb{Z}^n \subset \mathbb{R}^n$. Such a construction can be described by a short exact sequence of vector spaces

$$0 \longrightarrow \mathbb{R}^{d-n} \xrightarrow{Q} \mathbb{R}^d \xrightarrow{\beta} \mathbb{R}^n \longrightarrow 0\,, \tag{3.13}$$

which restricts to a short exact sequence of free-$\mathbb{Z}$ modules (lattices):

$$0 \longrightarrow \mathbb{Z}^{d-n} \xrightarrow{Q} \mathbb{Z}^d \xrightarrow{\beta} \mathbb{Z}^n \longrightarrow 0\,. \tag{3.14}$$

Since Q and β are linear maps, they can be represented by matrices: Q by a $d \times (d-n)$ matrix and β by an $n \times d$ matrix. That the sequence (3.13) is exact means that precisely n columns of β are linearly independent (i.e. β is surjective), that $\beta Q = 0$, and that the $d-n$ columns of Q are linearly independent (i.e. Q is injective). That (3.13) restricts to (3.14) further implies that Q and β have integer entries. The torus $N = \mathbb{R}^{d-n}/(2\pi\mathbb{Z}^{d-n})$ as we defined it has volume $(2\pi)^{d-n}$.

The hyperkähler quotient of $\mathbb{H}^d$ by the torus N is defined to be the zero locus of a set of moment maps in addition to a quotient by the torus action. The $3(d-n)$ moment maps are compactly expressed using the matrix equation

$$\begin{aligned} \mu_i &= \sum_{a=1}^{d} Q_i^a \left[q_a \sigma_3 q_a^\dagger + \vec{\lambda}_a \cdot \vec{\sigma} \right] && (3.15) \\ &= \sum_{a=1}^{d} Q_i^a \left[\begin{pmatrix} r_{a3} & r_{a1} - i r_{a2} \\ r_{a1} + i r_{a2} & -r_{a3} \end{pmatrix} + \begin{pmatrix} \lambda_{a3} & \lambda_{a1} - i\lambda_{a2} \\ \lambda_{a1} + i\lambda_{a2} & -\lambda_{a3} \end{pmatrix} \right] , && (3.16) \end{aligned}$$

where σ_1, σ_2, and σ_3 are the Pauli spin matrices. The hyperkähler quotient X is then

$$X \equiv \mathbb{H}^d /// N = \mu^{-1}(0)/N \ . \tag{3.17}$$

We are particularly interested in the case where $\vec{\lambda}_a = 0$ and the corresponding hyperkähler manifold is a cone. The base of this cone is a $4n-1$ real dimensional Riemannian manifold with positive scalar curvature and a locally free action of $SU(2)$ that descends from the $SU(2)$ rotating the three complex structures. The base of such a cone is called a tri-Sasaki manifold. Indeed, a Riemannian manifold (Y, g) is tri-Sasaki if and only if the Riemannian cone $X = (\mathbb{R}^+ \times Y, dr^2 + r^2\, g_{ij} dy^i dy^j)$ is hyperkähler. The induced metric on X from $\mathbb{H}^d$ is Ricci flat, which in turn implies that the induced metric on Y satisfies Einstein's equations with a positive cosmological constant. In other words, Y is also an Einstein manifold.

3.2.2 The Volume of the tri-Sasaki Einstein Base

Given a toric hyperkähler cone X, we would like to compute the volume of the base Y with respect to the induced metric from $\mathbb{H}^d$. A brute force approach would be to introduce the variables $\vec{t}_j$ such that

$$\vec{r}_a = \sum_j \beta_{aj} \vec{t}_j \ . \tag{3.18}$$

These constrained $\vec{r}_a$ automatically satisfy the moment map conditions. Plugging this expression into the line element (3.11) on $\mathbb{H}^d$ yields a metric on $X \times T^{d-n}$. Integrating the

square root of the determinant of the metric up to a finite radial coordinate $1 = \sum_a r_a$ and dividing by $\mathrm{Vol}(T^{d-n})$ would yield $\mathrm{Vol}(Y)/4n$.

The approach outlined above appears to be difficult. Instead, we will use a result due to Yee [76], which was proven using an elegant localization argument:

Yee's Formula. *Consider a tri-Sasaki Einstein manifold Y defined via the short exact sequence (3.14) and hyperkähler quotient construction described above. If the metric is normalized such that $R_{ij} = 2(2n-1)g_{ij}$, then the volume is*

$$\mathrm{Vol}(Y) = \frac{2^{d-n+1}\pi^{2n}}{(2n-1)!\,\mathrm{Vol}\,(N)} \int \left(\prod_{j=1}^{d-n} d\phi^j \right) \prod_{a=1}^{d} \frac{1}{1 + \left(\sum_{k=1}^{d-n} Q_k^a \phi^k \right)^2}\,. \tag{3.19}$$

This volume formula has the forgiving property of being invariant under rescaling the matrix $Q \to \lambda Q$. Under this rescaling the torus volume changes, $\mathrm{Vol}(N) \to \lambda^{n-d}\,\mathrm{Vol}(N)$. The factor λ^{n-d} is then canceled by the Jacobian introduced upon rescaling $\phi \to \phi/\lambda$. We discussed above that the columns of Q can be chosen such they form a $\mathbb{Z}$-basis of the kernel of β and $\mathrm{Vol}(N) = (2\pi)^{d-n}$. Note that if we were not so clever in our choice of Q, the volume formula would still give us the correct answer because of this scaling invariance.

So far we have assumed that N is isomorphic to T^d, but we can also choose N to be isomorphic to $N_T \times T^d$ for some finite abelian group N_T. In that case, the image of β is no longer all of $\mathbb{Z}^n$. The cokernel of β is dual to N_T. The volume of N for our choice of Q is therefore $(2\pi)^{d-n}\,|N_T| = (2\pi)^{d-n}\,|\mathrm{Coker}\,\beta|$.

We now rewrite this volume integral in a more convenient fashion. Let us assume that we chose the columns Q to form a $\mathbb{Z}$ basis of $\mathrm{Ker}(\beta)$. In this case, the volume formula simplifies:

$$\mathrm{Vol}(Y) = \frac{2\pi^{3n-d}}{(2n-1)!} \int \left(\prod_{j=1}^{d-n} d\phi^j \right) \prod_{a=1}^{d} \frac{1}{1 + \left(\sum_k Q_k^a \phi^k \right)^2}. \tag{3.20}$$

A well-known result from Fourier analysis is

$$\frac{1}{1+x^2} = \frac{1}{2} \int_{-\infty}^{\infty} dy e^{-|y| + ixy}. \tag{3.21}$$

Then $\mathrm{Vol}(Y)$ can be rewritten as

$$\mathrm{Vol}(Y) = \frac{\pi^{3n-d}}{(2n-1)! \cdot 2^{d-1}} \int \left(\prod_{j=1}^{d-n} d\phi^j \right) \left(\prod_{a=1}^{d} e^{-|y_a|} dy_a \right) \exp\left(\sum_{k,a} i Q_k^a \phi^k y_a \right). \tag{3.22}$$

We are allowed to switch the order of integration, integrating over the ϕ^j first.[2] We obtain

$$\mathrm{Vol}(Y) = \frac{\pi^{2n}}{(2n-1)! \cdot 2^{n-1}} \int \left(\prod_{a=1}^{d} e^{-|y_a|} dy_a \right) \prod_{k=1}^{d-n} \delta\left(\sum_a Q_k^a y_a \right) . \tag{3.23}$$

To integrate over the delta functions, note that we can get a basis for $\mathbb{Z}^d$ by taking the basis for $\mathrm{Ker}(\beta)$ and pullbacks of the basis for $\mathrm{Im}(\beta)$. The Jacobian for transforming from the standard basis of $\mathbb{Z}^p$ to this basis must be one, since both bases generate the same lattice. In our new coordinates, we have two kinds of variables: s_i corresponding to the columns of Q and t_i corresponding to the rows of β. The product of delta functions is just $\delta(s_1)\delta(s_2)\cdots\delta(s_{d-n})$ and can be performed straightforwardly. If the rows of β span $\mathbb{Z}^n$, the t_j can be written

$$y_a = \sum_{j=1}^{n} \beta_{aj} t_j \ . \tag{3.24}$$

The integral reduces to the following useful form:

$$\mathrm{Vol}(Y) = \frac{\pi^{2n}}{(2n-1)! \cdot 2^{n-1}} \int \left(\prod_{k=1}^{n} dt_k \right) \exp\left(- \sum_{a=1}^{d} \left| \sum_{j=1}^{n} \beta_{aj} t_j \right| \right) . \tag{3.25}$$

Note that if the rows of β do not span $\mathbb{Z}^n$, then the Jacobian will have an extra factor of $|\,\mathrm{Coker}(\beta)|$, which cancels the $|\,\mathrm{Coker}(\beta)|$ in $\mathrm{Vol}(N)$. So (3.25) holds even if the rows of β do not span $\mathbb{Z}^n$. This integral form provides us with a corollary to Yee's Formula:

Corollary 3.2.1. *i) If a column β_a is removed from β, $\mathrm{Vol}(Y)$ increases. ii) If two columns β_a and β_b of β are combined to form the new column $\beta_a + \beta_b$ of a new β' with one fewer*

[2] The integral (3.22) is not absolutely convergent. However, if we multiply the integrand by $\exp\left[-\sum_k \epsilon_k (\phi^k)^2\right]$, for some small $\epsilon_k > 0$, then the integral does converge absolutely. We can then take the limit $\epsilon_k \to 0^+$. If we integrate out the y_a then we get $\lim_{\epsilon_k \to 0^+} \frac{2\pi^{3n-d}}{(2n-1)!} \int \left(\prod_{j=1}^{d-n} d\phi^j \right) \prod_{a=1}^{d} \frac{\exp\left[-\sum_k \epsilon_k (\phi^k)^2\right]}{1+\left(\sum_k Q_k^a \phi^k\right)^2} = \frac{2\pi^{3n-d}}{(2n-1)!} \int \left(\prod_{j=1}^{d-n} d\phi^j \right) \prod_{a=1}^{d} \frac{1}{1+\left(\sum_k Q_k^a \phi^k\right)^2}$. If we integrate out the ϕ^j then we get $\lim_{\epsilon_k \to 0^+} \frac{\pi^{2n}}{(2n-1)! \cdot 2^{n-1}} \int \left(\prod_{a=1}^{d} e^{-|y_a|} dy_a \right) \prod_{k=1}^{d-n} \frac{\exp -\left(\sum_a Q_k^a y_a\right)^2 / 4\epsilon_k}{2\sqrt{\pi \epsilon_k}} = \frac{\pi^{2n}}{(2n-1)! \cdot 2^{n-1}} \int \left(\prod_{a=1}^{d} e^{-|y_a|} dy_a \right) \prod_{k=1}^{d-n} \delta\left(\sum_a Q_k^a y_a\right)$.

columns, $\mathrm{Vol}(Y)$ *either increases or stays the same. The volume remains the same if and only if the two columns are proportional,* $\beta_a = c\beta_b$ *for some* $c \in \mathbb{R}^+$.

Proof. (i) is true because the integrand of (3.25) (which is positive definite) increases when a column of β is removed. (ii) is true because the absolute value of a sum is less than or equal to the sum of the absolute values. Equality occurs only when the expressions inside the absolute values are proportional. □

The volume formula (3.25) leads to another useful corollary of Yee's Formula:

Corollary 3.2.2. *Let* $\mathcal{P} \in \mathbb{R}^n$ *be the polytope*

$$\mathcal{P} = \left\{ \vec{t} \in \mathbb{R}^n : \sum_{a=1}^{d} \left| \sum_{j=1}^{n} \beta_{aj} t_j \right| \leq 1 \right\} . \tag{3.26}$$

Let S^{4n-1} *be a* $(4n-1)$*-dimensional sphere with unit radius. The volume of the tri-Sasaki Einstein manifold satisfies the relation*

$$\frac{\mathrm{Vol}(Y)}{\mathrm{Vol}(S^{4n-1})} = \frac{n!}{2^n} \mathrm{Vol}(\mathcal{P}) . \tag{3.27}$$

Proof. Introduce the notation $\alpha = \sum_{a=1}^{d} \left| \sum_{j=1}^{n} \beta_{aj} t_j \right|$ with $\alpha > 0$ if at least one of the t_k is non-zero, and write $t_k = \alpha b_k$ for some new variables b_k. The b_k are constrained to live on the boundary of $\mathcal{P}$. Eq. (3.25) can be rewritten as

$$\mathrm{Vol}(Y) = \frac{\pi^{2n}}{(2n-1)! \cdot 2^{n-1}} \int \left(\prod_{k=1}^{n} db_k \right) \delta \left(\sum_{a=1}^{d} \left| \sum_{j=1}^{n} \beta_{aj} b_j \right| - 1 \right) \int_0^\infty d\alpha \, \alpha^{n-1} e^{-\alpha} . \tag{3.28}$$

The volume of $\mathcal{P}$ can be also written in terms of these variables:

$$\mathrm{Vol}(\mathcal{P}) = \int \left(\prod_{k=1}^{n} db_k \right) \delta \left(\sum_{a=1}^{d} \left| \sum_{j=1}^{n} \beta_{aj} b_j \right| - 1 \right) \int_0^1 d\alpha \, \alpha^{n-1} . \tag{3.29}$$

Performing the integrals in α and comparing the last two formulas, we obtain

$$\mathrm{Vol}(Y) = \frac{\pi^{2n} n!}{(2n-1)! \cdot 2^{n-1}} \mathrm{Vol}(\mathcal{P}) \,, \tag{3.30}$$

or, using the fact that the volume of the $(4n-1)$-sphere is $\mathrm{Vol}(S^{4n-1}) = 2\pi^{2n}/(2n-1)!$ the desired result follows. □

A final corollary is an explicit result for the volume of the tri-Sasaki Einstein spaces relevant for the gauge theories we discuss below:

Corollary 3.2.3. *In the case $n = 2$, choose β such that the two-vectors β_a lie in the upper half plane and order them such that $\beta_a \wedge \beta_{a+1} > 0$.*[3] *The volume of the tri-Sasaki Einstein space Y is*

$$\mathrm{Vol}(Y) = \frac{\pi^4}{6} \sum_{a=1}^{d} \frac{\gamma_{a(a+1)}}{\sigma_a \sigma_{a+1}} \tag{3.31}$$

where we have defined the quantities $\gamma_{ab} \equiv |\beta_a \wedge \beta_b|$, $\sigma_a \equiv \sum_{b=1}^{d} \gamma_{ab}$, and $\beta_{d+1} \equiv -\beta_1$.

Note that the volume of Y is independent of the sign of the β_a and their order inside β. The order of the β_a in the corollary is their order around $\mathbb{RP}^1$. (Flipping the sign of any given column of β does not change the hyperkähler quotient $\mathbb{H}///N$ because the torus N defined by (3.14) is invariant under such a sign flip. Therefore it is natural to identify β_a with $-\beta_a$ and consider the ordering of the $\beta_a = (p_a, q_a)^T$ to be the ordering of $[p_a, q_a] \in \mathbb{RP}^1$ around $\mathbb{RP}^1$.)

Proof of Corollary 3.2.3. We split the integral into various regions according to whether $\sum_j \beta_{aj} t_j$ is positive or negative. There are $2d$ such regions. Since the integral does not change if we replace each t_k with $-t_k$, we only need to consider d regions. The ordering of the β_a guarantees that any two consecutive columns determine a region boundary.

Now we choose a and compute the integral in a region bounded by the lines $u = -\sum_j \beta_{aj} t_j = 0$ and $v = \sum_j \beta_{(a+1)j} t_j = 0$. The prefactor in the integral is $\pi^4/12$, but

[3] We define $\begin{pmatrix} a \\ b \end{pmatrix} \wedge \begin{pmatrix} c \\ d \end{pmatrix} \equiv ad - bc$.

we need to multiply by two since there are two such regions. We get

$$\begin{aligned} & \frac{\pi^4}{6}\frac{1}{\gamma_{a(a+1)}}\int_0^\infty du\int_0^\infty dv\exp\left(-\sum_{b=1}^{p}\frac{\gamma_{ab}v+\gamma_{(a+1)b}u}{\gamma_{a(a+1)}}\right) \\ = \; & \frac{\pi^4}{6}\frac{\gamma_{a(a+1)}}{\left(\sum_b\gamma_{ab}\right)\left(\sum_b\gamma_{(a+1)b}\right)}. \end{aligned} \tag{3.32}$$

Summing the regions yields the volume stated in the corollary. □

Given Corollary 3.2.2, there exists an equivalent proof that involves computing the area of $\mathcal{P}$ from the definition (3.26).

The authors of [27] conjectured a formula for the volume of these $n = 2$ tri-Sasaki Einstein spaces. Their formula is interesting because it does not rely on an ordering of the β_a and makes the permutation symmetry of the columns of β manifest. We have been able to promote this conjecture to a theorem. As the techniques for the proof are not typical of the main arguments in the chapter, we include the proof as Appendix A.1.

Tree Formula. *The area of the polytope $\mathcal{P}$ described in (3.26) in the case $n = 2$ can be written*

$$\text{Area}(\mathcal{P}) = 2\frac{\sum_{(V,E)\in T}\prod_{(a,b)\in E}\gamma_{ab}}{\prod_{a=1}^{d}\sigma_a}\ . \tag{3.33}$$

where T is the set of all trees (acyclic connected graphs) with nodes $V = \{1, 2, \ldots, p\}$ and edges $E = \{(a_1, b_1), (a_2, b_2), \ldots, (a_{p-1}, b_{p-1})\}$.

3.2.3 Brane Constructions and an F-theorem

Consider the following brane construction in type IIB string theory. A stack of N coincident D3-branes spans the 01236 directions with the 6 direction periodically identified. Let there be bound states of NS5- and D5-branes. Denote the number of NS5- and D5-branes in the bound state by p_a and q_a respectively. These (p_a, q_a)-branes intersect the D3-branes at intervals around the circle and span the 012 directions. Each (p_a, q_a)-brane lies at an angle θ_a in the 37, 48, and 59 planes where $\theta_a = \arg(p_a + iq_a)$. These brane constructions are known to preserve 6 of the 32 supersymmetries of type IIB string theory [80, 81].[4]

[4]Note that a brane with charges $(-p_a, -q_a)$ is an anti-(p_q, q_a)-brane rotated by 180 degrees, which is the same as a (p_a, q_a)-brane. The (p_a, q_a) charge is most naturally defined up to this overall sign.

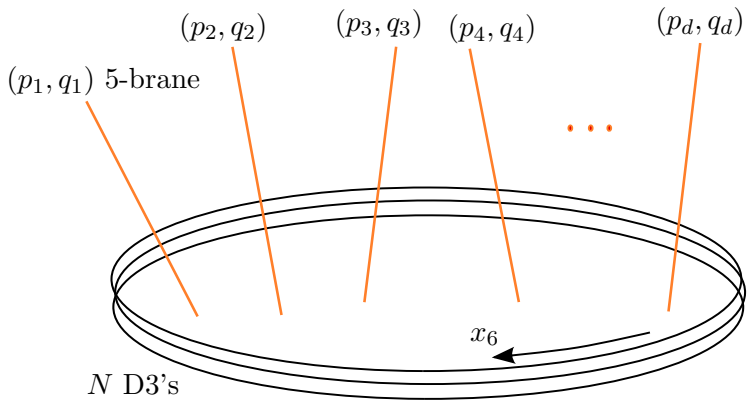


Figure 3.1: A schematic picture of the brane construction. The N D3-branes span the 0126 direction, and the (p_a, q_a) 5-branes span the 012 directions as well as the lines in the 37, 48, and 59 planes that make angles $\theta_a = \arg(p_a + iq_a)$ with the 3, 4, and 5 axes, respectively. The three-dimensional $\mathcal{N} = 3$ theories considered in this chapter live on the 012 intersection of these branes.

This brane construction cannot be described reliably within type IIB supergravity because the dilaton becomes large. A better description can be obtained after a T-duality along the 6 direction and a lift to M-theory, where the resulting configuration can be described within 11-d supergravity. The geometry depends on N. When N is small, the geometry is $\mathbb{R}^{2,1} \times X$ where X is the hyperkähler cone discussed above for $n = 2$ [72]. In the large N limit, the D3-branes produce a significant back-reaction, and close to them the geometry is $AdS_4 \times Y$ where Y is a tri-Sasaki Einstein space [66]. In both cases, the charges $(p_a, q_a) = \beta_a^T$ are the columns of β.

In the case where $p_a = 0$ or 1, this brane construction admits a simple $\mathcal{N} = 3$ supersymmetric 2+1 dimensional field theory interpretation. The $(1, q_a)$-branes, $a = 1, \ldots, d$, break the D3-branes up into d segments along the circle. For each segment, we have a $U(N)$ Chern-Simons theory at level $k_a = q_{a+1} - q_a$. The $(0, q_i)$-branes, $i = 1, \ldots, n_F$, also intersect the D3-branes. At each intersection, we have massless strings that join the D3-branes to the $(0, q_i)$-brane that correspond to flavor fields in the fundamental representation. (See figure 3.2.) These gauge theories are described in greater detail in [66, 82].

As described in the introduction, it has been conjectured that the logarithm F of the partition function of the Euclidean field theory on S^3 serves as a measure of the number of degrees of freedom in the theory, being an analog of the conformal anomaly coefficient a

of a 3+1 dimensional field theory. While the a-theorem is a conjecture that the conformal anomaly a decreases along RG flows, the conjectured F-theorem states that F should decrease along RG flows [21]. As given in (3.3), the AdS/CFT correspondence predicts that in the large N limit, $F \sim 1/\sqrt{\text{Vol}(Y)}$. Thus, along RG flows, $\text{Vol}(Y)$ should increase.

It is remarkable that for the class of theories we consider here, Corollary 3.2.1 confirms this expectation that $\text{Vol}(Y)$ should increase along RG flows. The UV of the field theory should correspond to looking at our D3-brane construction from a great distance as only high energy excitations will be able to get far from the branes, while the IR should correspond to getting very close to the D3-branes. The simplest realization of such an RG flow is to add a mass term for flavor fields corresponding to a $(0, q_i)$-brane. Such a mass term corresponds to introducing a small distance between the $(0, q_i)$-brane and the stack of D3-branes. In the UV, these flavor fields will contribute to F, while in the IR, the flavors and corresponding $(0, q_i)$-brane should be absent.

A slightly more complicated realization involves an intersecting $(0, q)$-brane and a $(1, 0)$-brane. We have a $U(N)$ vector multiplet on each side of the $(1, 0)$-brane. The $(0, q)$-brane produces q fundamental flavor fields charged under one $U(N)$ and q anti-fundamentals charged under the other. Adding real masses of the same sign for each flavor field corresponds to a "web deformation" [81] where close to the D3-branes the $(0, q)$- and $(1, 0)$-brane form a $(1, q)$-brane bound state. Far from the D3-branes, because we have fixed boundary conditions for the $(0, q)$- and $(1, 0)$-brane by specifying their angles, the two branes will remain separate. From a large distance, we will not see that the two 5-branes have combined. We will only see their asymptotic regions where they remain separated. Closer up, we will see that the branes have made a bound state. From the point of view of our volume formula, combining these two branes increases the volume and thus decreases F.

In fact, Corollary 3.2.1 and our volume formula put no restriction on the type of (p, q)-brane we remove or the type of (p, q)- and (p', q')-brane we combine to form a $(p+p', q+q')$ bound state. In all cases, the volume will increase, corresponding to a decrease in F. Unfortunately, from the field theory perspective it is not clear in general to what these more general types of RG flow correspond.

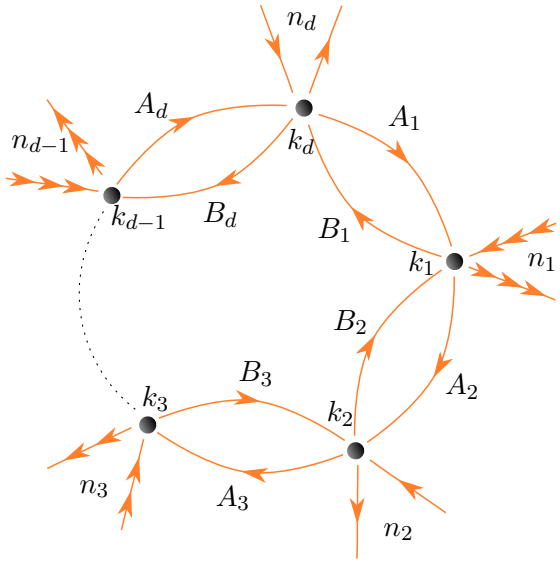


Figure 3.2: A necklace quiver gauge theory where the gauge sector consists of d $U(N)$ gauge groups with Chern-Simons coefficients k_a. The matter content consists of the bifundamental fields A_a and B_a, as well as n_a pairs of fundamentals and anti-fundamentals transforming under the ath gauge group.

3.3 Field Theory Computation of the Free Energy

3.3.1 $\mathcal{N} = 3$ Matrix Model

Consider the $\mathcal{N} = 3$ supersymmetric 2+1 dimensional Chern-Simons theories corresponding to the (p, q)-brane constructions described above where $p = 0$ or 1. Let there be d $U(N)$ gauge groups at level $k_a = q_{a+1} - q_a$, matter fields A_a and B_a in conjugate bifundamental representations of the $(a-1)$st and ath gauge group and n_a pairs of flavor fields transforming in fundamental and anti-fundamental representations of the ath gauge group, with $a = 1, \ldots, d$ (see figure 3.2). As explained in [17], the partition function for these necklace quivers localizes on configurations where the scalars σ_a in the vector multiplets are constant Hermitian matrices. Denoting the eigenvalues of σ_a by $\lambda_{a,i}$, $1 \leq i \leq N$, the partition function takes the form of the eigenvalue integral

$$Z = \int \left(\prod_{a,i} d\lambda_{a,i} \right) L_v(\{\lambda_{a,i}\}) L_m(\{\lambda_{a,i}\}) \tag{3.34}$$

where the vector multiplets contribute

$$L_v = \frac{1}{N!}\prod_{a=1}^{d}\left(\prod_{i\neq j} 2\sinh[\pi(\lambda_{a,i}-\lambda_{a,j})]\right)\exp\left(i\pi\sum_i k_a\lambda_{a,i}^2\right) \tag{3.35}$$

and the bifundamental and fundamental matter fields contribute

$$L_m = \prod_{a=1}^{d}\left(\prod_{i,j}\frac{1}{2\cosh[\pi(\lambda_{a-1,i}-\lambda_{a,j})]}\right)\left(\prod_i \frac{1}{2\cosh\pi\lambda_{a,i}}\right)^{n_a} . \tag{3.36}$$

We follow the recipe suggested in [27] for analyzing this matrix model in the large N limit. We write $\lambda = N^{1/2}x + iy$ and assume that the density of eigenvalues $\rho(x)$ is the same for each vector multiplet. To leading order in N, the matrix model for the $\mathcal{N}=3$ necklace theories involves extremizing a free energy functional of the type

$$F[\rho,\delta y_a] = \pi N^{3/2}\int \rho(x)dx\,\left[n_F|x| + 2x\sum_{a=1}^{d} q_a\delta y_a(x) + \rho(x)\sum_{a=1}^{d} f(\delta y_a(x))\right], \tag{3.37}$$

where $\delta y_a = y_{a-1} - y_a$, $n_F = \sum_a n_a$, and f is a periodic function with period one given by

$$f(t) = \frac{1}{4} - t^2 \qquad \text{when} \quad -\frac{1}{2}\leq t\leq\frac{1}{2}\,. \tag{3.38}$$

This free energy should be extremized over the set

$$\mathcal{C} = \left\{(\rho,\delta y_a) : \int dx\,\rho(x) = 1; \rho(x)\geq 0 \text{ and } \sum_{a=1}^{d}\rho(x)\delta y_a(x) = 0 \text{ a.e.}\right\}, \tag{3.39}$$

where we think of $\rho(x)$ and $\rho(x)\delta y_a(x)$ as functions defined almost everywhere (a.e.). To enforce these constraints, we introduce the Lagrange multipliers μ and $\nu(x)$:

$$\tilde{F}[\rho,\delta y_a] = F[\rho,\delta y_a] - 2\pi N^{3/2}\left[\mu\left(\int dx\,\rho(x) - 1\right) + \int dx\,\rho(x)\nu(x)\sum_{a=1}^{d}\delta y_a(x)\right] . \tag{3.40}$$

The equations of motion that follow from this action are

$$\sum_{a=1}^{d} \left[f(\delta y_a(x))\rho(x) + (q_a x - \nu(x))\ \delta y_a(x) \right] = \mu - \frac{1}{2} n_F |x| \,, \tag{3.41a}$$

$$\frac{1}{2} f'(\delta y_a(x))\rho(x) + q_a x = \nu(x) \,. \tag{3.41b}$$

The solution of these equations and the constraint $\sum_{a=1}^{d} \delta y_a = 0$ is

$$\begin{aligned} \rho(x) &= s_L(x) - s_S(x) \,, \qquad \nu(x) = -\frac{1}{2} \left[s_L(x) + s_S(x) \right] \,, \\ \delta y_a(x) &= \frac{1}{2} \frac{|s_L(x) + q_a x| - |s_S(x) + q_a x|}{s_L(x) - s_S(x)} \,, \end{aligned} \tag{3.42}$$

where we denoted by $s_L(x)$ and $s_S(x)$ the two solutions of the equation

$$\frac{1}{2} n_F |x| + \frac{1}{2} \sum_{a=1}^{d} |s(x) + q_a x| = \mu \,, \tag{3.43}$$

with $s_L(x) \geq s_S(x)$. Note that the set of s and x satisfying (3.43) defines a polygon. In fact, it defines the polygon $\mathcal{P}$ of (3.26) but rescaled by a factor of 2μ.

We show in Appendix A.2 quite generally that in the continuum limit the extremized value of the free energy F is proportional to the Lagrange multipler μ:

$$F = \frac{4\pi N^{3/2}}{3} \mu \,. \tag{3.44}$$

Thus to determine F, it suffices to find μ.

By definition, the density $\rho(x)$ should integrate to one. The solution (3.42) demonstrates that the density $\rho(x)$ is proportional to the length of a slice through $\mathcal{P}$ at constant x. Thus integrating $\rho(x)$ over x should yield a quantity proportional to the area of the polygon. We obtain

$$1 = \int dx\, \rho(x) = 4\mu^2 \,\mathrm{Vol}(\mathcal{P}) \,. \tag{3.45}$$

Assembling (3.44) and (3.45) yields $F = 2\pi N^{3/2}/3\sqrt{\mathrm{Vol}(\mathcal{P})}$. From Corollary 3.2.2 and in particular (3.30), we recover our AdS/CFT prediction (3.3).

3.3.2 Operator Counting and the Matrix Model

In this section, we relate the matrix model quantities $\rho(x)$ and $\rho(x)\delta y_a(x)$ to numbers of operators in the chiral ring that don't vanish on the geometric branch of the moduli space of the $\mathcal{N} = 3$ necklace quiver in the abelian case $N = 1$. To that end, we first characterize the chiral ring, defined to be gauge-invariant combinations of the bifundamental fields and monopole operators modulo superpotential relations. At arbitrary N, the superpotential for these $\mathcal{N} = 3$ necklace quivers when $n_F = 0$ is

$$W = \sum_{a=1}^{d} \frac{1}{k_a} \mathrm{Tr}(B_{a+1}A_{a+1} - A_aB_a)^2 \ , \tag{3.46}$$

where $k_a = q_{a+1} - q_a$. When $N = 1$, this superpotential gives rise to $d - 2$ linearly independent relations on the geometric branch of the moduli space

$$(k_a + k_{a-1})A_aB_a = k_{a-1}A_{a+1}B_{a+1} + k_aA_{a-1}B_{a-1} \ . \tag{3.47}$$

(As pointed out by [66], these relations are one set of moment map constraints in the hyperkähler quotient discussed previously.) There also exists a monopole operator T and anti-monopole operator $\tilde{T}$ that create one and minus one units of flux, respectively, through each gauge group. These monopole operators satisfy the quantum relation $T\tilde{T} = 1$ [32, 33]. The A_a and B_a fields have R-charge 1/2, guaranteeing that the R-charge of the superpotential is two. We will take the monopole operators to have zero R-charge. Gauge invariance for this $U(1)^d$ gauge theory means that the total $U(1)^d$ charge of a gauge-invariant operator constructed from A_a, B_a, and monopole fields will vanish. Let A_a (B_a) have charge $+1(-1)$ under gauge group a and charge $-1(+1)$ under gauge group $a - 1$. The monopole operators have gauge charges $\pm k_a$ under the ath gauge group. To write a gauge-invariant

operator in a compact form, it is convenient to define the operators

$$C_a^{mq_a+s} \equiv \begin{cases} A_a^{mq_a+s} & \text{if } mq_a+s>0 \\ B_a^{-mq_a-s} & \text{if } mq_a+s<0 \end{cases} \quad ; \qquad U^m \equiv \begin{cases} T^m & \text{if } x>0 \\ \tilde{T}^{-m} & \text{if } x<0 \end{cases} \ . \tag{3.48}$$

The gauge-invariant operators are

$$\mathcal{O}(m,s,i,j) = U^m C_1^{mq_1+s} C_2^{mq_2+s} \cdots C_d^{mq_d+s} (A_1B_1)^i (A_2B_2)^j \ . \tag{3.49}$$

Note we have used the superpotential relations to eliminate all but two of the (A_iB_i) factors that could potentially appear in a gauge-invariant operator.

Let us consider the set of operators of the form $\mathcal{O}(m,s,0,0)$ with R-charge less than r. The number of these operators is equal to the number of lattice points inside the polygon

$$\mathcal{P}_r = \left\{ (m,s) \in \mathbb{R}^2 : \frac{1}{2}\sum_{a=1}^{d} |s+q_a m| < r \right\} \ . \tag{3.50}$$

In the large r limit, the number of these lattice points is well approximated by Area($\mathcal{P}_r$). The polygon in Section 3.2.2 and this polygon are related via $\mathcal{P}_{1/2} = \mathcal{P}$. Corollary 3.2.2 established that Area($\mathcal{P}$) and Vol(Y) are proportional. As Area($\mathcal{P}_r$)=$4r^2$ Area($\mathcal{P}$), we have an additional relation between the number of a certain type of operator and Vol(Y).

The relation between Area($\mathcal{P}$) and an operator counting problem reveals an additional relationship between the eigenvalue density $\rho(x)$ and the chiral ring. We claim that in the large r limit the number of operators $\mathcal{O}(m,s,0,0)$ of R-charge less than r and monopole charge between $m = xr/\mu$ and $m+dm = (x+dx)r/\mu$ is

$$\frac{r^2}{\mu^2}\rho(x)dx \ . \tag{3.51}$$

From (3.42), the quantity $\rho(x)dx$ in the matrix model corresponds to the area of a constant x strip of $\mathcal{P}_\mu$ of height $s_L(x) - s_S(x)$ and width dx. Roughly speaking, the operator counting gives $\rho(x)$ a new interpretation as the number of operators $\mathcal{O}(x,s,0,0)$ of R-charge less than μ and monopole charge bounded between x and $x+dx$. As μ is of order one in the matrix

model, we should be more careful and consider first $\mathcal{P}_r$ for some large r, giving rise to the factors of r/μ in the claim.

There is a subtle relation between δy_a and the chiral ring where A_a or B_a is set to zero. In the large r limit, we claim that

$$\frac{r^2}{\mu^2}\left(\delta y_a(x)+\frac{1}{2}\right)\rho(x)dx \tag{3.52}$$

counts the number of operators of the form $\mathcal{O}(m,s,0,0)$ with R-charge less than r and with monopole charge between $m = xr/\mu$ and $m+dm = (x+dx)r/\mu$ and with $B_a = 0$. Flipping the sign of $\delta y_a(x)$ yields an equivalent expression for operators with $A_a = 0$. Note that

$$\rho(x)\left(\delta y_a(x)+\frac{1}{2}\right) \;=\; \begin{cases} 0 & q_a x < -s_L(x)\,, \\ s_L(x)+q_a x & -s_L(x) < q_a x < -s_S(x)\,, \\ s_L(x)-s_S(x) & -s_S(x) < q_a x\,. \end{cases} \tag{3.53}$$

For simplicity, let us first assume that μ is large and count the operators with R-charge less than μ. We get the correct counting when $q_a x < -s_L(x)$ and when $-s_S(x) < q_a x$ because the region $q_a x + s_S(x) > 0$ corresponds to a portion of the polygon where the $\mathcal{O}(x,s,0,0)$ contain no B_a while the region $q_a x + s_L(x) < 0$ corresponds to operators that contain no A_a. In the central region, the operators that contain no B_a satisfy the constraint $s_L(x) > s > -q_a x$. Because there is one operator per lattice point, the number of operators that contain no B_a is proportional to the difference $s_L(x) - (-q_a x)$. Similarly $\rho(x)(-\delta y_a(x) + 1/2)$ will count the number of operators with no A_a. Again, since μ is of order one in the matrix model, we should rescale our results for a polygon $\mathcal{P}_r$ in the large r limit, yielding the extra factors of r/μ in the claim (3.52) for δy_a.

For a general supersymmetric gauge theory, we will not be able to make a clean separation between all operators in the chiral ring and operators of this special form $\mathcal{O}(m,s,0,0)$. It is thus useful to reformulate these statements about $\rho(x)$ and δy_a in terms of all chiral operators. Let us introduce the function $\psi(r,m)$ which counts the number of operators with

R-charge less than r and monopole charge less than m. We claim that

$$\left.\frac{\partial^3\psi}{\partial r^2\partial m}\right|_{m=rx/\mu} = \frac{r}{\mu}\rho(x) \ . \tag{3.54}$$

Note that in the large r limit, $(\partial^2\psi/\partial r\partial m)dr\,dm$ can be interpreted as the number of operators of R-charge between r and $r+dr$ and monopole charge between m and $m+dm$. Given an operator $\mathcal{O}(m,s,0,0)$ of R-charge $\frac{1}{2}\sum_a|s+q_am| = r_0 < r$, we can form an operator of R-charge equal to r by multiplying $\mathcal{O}(m,s,0,0)$ by a factor of $(A_1B_1)^j(A_2B_2)^{r-r_0-j}$. There are precisely $r-r_0+1$ ways of forming such a factor (assuming $r-r_0$ is an integer). This multiplicity $r-r_0+1$ associated to a lattice point (m,s) of the polygon can be interpreted as the total number of operators of fixed m and s with R-charge equal to r. The difference in the number of operators with R-charge $r+1$ and r integrated over a strip at constant m now has the dual interpretation as $r\rho(\mu m/r)/\mu$ or $\partial^3\psi/\partial r^2\partial m$ (in the large r limit).

As for $\delta y_a(x)$, let us introduce $\psi_{A_a}(r,m)$ and $\psi_{B_a}(r,m)$ as the number of operators with R-charge less than r and monopole charge less than m, with $A_a = 0$ and $B_a = 0$ respectively. We claim that

$$\left.\frac{\partial^2\psi_{A_a}}{\partial r\partial m}\right|_{m=rx/\mu} = \frac{r}{\mu}\rho(x)\left(-\delta y_a(x)+\frac{1}{2}\right) \ , \tag{3.55}$$

$$\left.\frac{\partial^2\psi_{B_a}}{\partial r\partial m}\right|_{m=rx/\mu} = \frac{r}{\mu}\rho(x)\left(\delta y_a(x)+\frac{1}{2}\right) \ . \tag{3.56}$$

Note that $(\partial^2\psi_{B_a}/\partial r\partial m)dr\,dm$ can be interpreted as the number of operators of R-charge between r and $r+dr$, monopole charge between m and $m+dm$, and no B_a operators, in the large r limit. Given an operator $\mathcal{O}(m,s,0,0)$ of R-charge $r_0 < r$ that does not involve the field B_a, we can form a unique operator of R-charge equal to r by multiplying $\mathcal{O}(m,s,0,0)$ by $(A_bB_b)^{r-r_0}$ where $b \neq a$ (provided $r-r_0$ is an integer). Thus, eq. (3.52) is counting the number of operators in the chiral ring with R-charge between r and $r+dr$, magnetic charge between $m = rx/\mu$ and $m+dm = r(x+dx)/\mu$, and no B_a operators. There is a parallel argument for ψ_{A_a}.

3.3.3 Operator Counting and Volumes

Given the close relation between $\rho(x)$ and $\mathrm{Vol}(Y)$, it is not surprising that there is also a close connection between $\rho(x)$ and numbers of operators in the chiral ring. Motivated by Weyl's Law for eigenfunctions of a Laplacian on a curved manifold, the authors of [54] noticed that the number of holomorphic functions on a certain class of Calabi-Yau cones could be related to the volume of the Sasaki-Einstein manifold base. Their result holds for a Calabi-Yau cone which is a $\mathbb{C}^*$-fibration over a variety in weighted projective space with a Kähler-Einstein metric. This special case was later generalized to any complex Kähler cone by Martelli, Sparks, and Yau [77]. Note that the set of holomorphic functions on the cone is precisely the chiral ring of the $N = 1$ gauge theory and that the ring has a natural grading from the R-charge. The relation between the number of holomorphic functions and the $\mathrm{Vol}(Y)$ is

$$\mathrm{Vol}(Y) = \frac{\pi^n n}{2^{n-1}} \lim_{r\to\infty} \frac{1}{r^4} \lim_{m\to\infty} \psi(r,m) \ , \tag{3.57}$$

where $\dim(Y) = 2n - 1$. In fact the result (3.57) helps to explain not only the relation (3.54) between $\rho(x)$ and $\psi(r,m)$ but also the relations (3.55) and (3.56) between $\delta y_a(x)$ and restrictions of the chiral ring to rings where $A_a = 0$ or $B_a = 0$. The subspace $A_a = 0$ of the Calabi-Yau cone is also a complex cone with Kähler structure. The level surfaces of this complex cone will be Sasaki manifolds which satisfy (3.57) with $n = 3$. The identifications (3.54), (3.55), and (3.56) along with (3.57) imply that

$$\int \rho(x)dx = \frac{24\mu^2}{\pi^4}\mathrm{Vol}(Y) \ , \tag{3.58}$$

$$\int \rho(x)\left(-\delta y_a(x) + \frac{1}{2}\right)dx = \frac{4\mu^2}{\pi^3}\mathrm{Vol}(Y_{A_a}) \ , \tag{3.59}$$

$$\int \rho(x)\left(\delta y_a(x) + \frac{1}{2}\right)dx = \frac{4\mu^2}{\pi^3}\mathrm{Vol}(Y_{B_a}) \ . \tag{3.60}$$

The relation (3.58) we derived already, but the second two relations are seemingly new. From the free energy functional (3.37), it is clear that $\delta y_a(x)$ is an odd function of x, and thus for these necklace quivers the integral $\int \rho(x)\delta y_a(x)dx$ vanishes trivially. The remaining

integral yields the result

$$\mathrm{Vol}(Y_{A_a}) = \mathrm{Vol}(Y_{B_a}) = \frac{3}{\pi}\,\mathrm{Vol}(Y)\ , \tag{3.61}$$

a result Yee found by other means [76].

A field theoretic interpretation of these five dimensional cycles was provided by [83, 84]. In an $AdS_4 \times Y$ solution of eleven-dimensional supergravity, an M5-brane wrapping such a cycle in Y looks like a point particle in AdS_4 with a mass proportional to the volume of the cycle times the tension of the five-brane. The AdS/CFT dictionary provides a relationship between the mass of the particle and its conformal dimension. As the wrapped five-brane is supersymmetric, the conformal dimension can be related to the R-charge of the corresponding operator that creates the state. If in the geometry the five-cycle corresponds to setting $A_a = 0$, the corresponding baryonic-like operator should involve an anti-symmetric product of N copies A_a. For our purposes, the essential point is a relation between $\mathrm{Vol}(Y_{A_a})$ and the R-charge of A_a:

$$R(A_a) = \frac{\pi}{6}\frac{\mathrm{Vol}(Y_{A_a})}{\mathrm{Vol}(Y)} = \frac{1}{2}\ . \tag{3.62}$$

This point suggests a way of generalizing (3.55) and (3.56) to an arbitrary quiver gauge theory. We should replace the $1/2$ with the R-charge of the corresponding bifundamental field X_{ab} with charge $+1$ under gauge group b and charge -1 under gauge group a:

$$\left.\frac{\partial^2 \psi_{X_{ab}}}{\partial r \partial m}\right|_{m=rx/\mu} = \frac{r}{\mu}\rho(x)\left[y_b(x) - y_a(x) + R(X_{ab})\right]\ . \tag{3.63}$$

3.4 The Matrix Model for (p, q)-Branes

As discussed in section 3.2.3, the free energy for a system of d (p_a, q_a) five-branes can be computed on the gravity side by combining the M-theory prediction (3.3) with the eq. (3.27), where $n = 2$ and $(p_a, q_a) = \beta_a^T$. More explicitly, the polygon $\mathcal{P}$ takes the form

$$\mathcal{P} = \left\{(x, s) \in \mathbb{R}^2 : \sum_{a=1}^{d} |p_a s + q_a x| \leq 1\right\}\ , \tag{3.64}$$

and the volume of $\mathcal{P}$ is related to F through $F = 2\pi N^{3/2}/3\sqrt{\text{Vol}(\mathcal{P})}$. In the previous section we were able to reproduce this formula from the field theory side in the cases where $p_a = 0$ or $p_a = 1$ for each a, as these were the cases where a simple Lagrangian description of the field theory led [17] to an expression for F in terms of a matrix integral. For $p_a > 1$, as mentioned in the introduction, a Lagrangian description would involve the coupling to the $T(U(N))$ theory described in [73]. We will now use our large N intuition to figure out the matrix model at finite N.

As a first step towards obtaining such a matrix model, we note that at large N one can obtain the correct value for F by extremizing the free energy functional:

$$F[\rho, \delta y_a] = \pi N^{3/2} \int \rho(x) dx \left[2x \sum_{a=1}^{d} \frac{q_a}{p_a} \delta y_a(x) + \rho(x) \sum_{a=1}^{d} p_a f(\delta y_a(x)/p_a) \right] . \tag{3.65}$$

The saddlepoint is a generalization of the earlier eigenvalue distribution (3.42) for the $(1, q_a)$-branes:

$$\begin{aligned} \rho(x) &= s_L(x) - s_S(x) \,, \qquad \nu(x) = -\frac{1}{2} \left[s_L(x) + s_S(x) \right] \,, \\ \delta y_a(x) &= \frac{1}{2} \frac{|p_a s_L(x) + q_a x| - |p_a s_S(x) + q_a x|}{s_L(x) - s_S(x)} \,, \end{aligned} \tag{3.66}$$

where $s_L(x) > s_S(x)$ are the two solutions of $\frac{1}{2} \sum_{a=1}^{d} |p_a s + q_a x| = \mu$.

The equation (3.65) seems ill behaved in the limit $p_a \to 0$, but in fact it is not. What happens is that $\delta y_a(x)$ will saturate very quickly to $\pm p_a/2$ as we move away from $x = 0$. Thus the function $f(\delta y_a/p_a)$ will vanish, and $2x q_a \delta y_a/p_a$ can be replaced with $q_a|x|$. Before, we identified the number of flavors with the sum $n_F = \sum_{p_a=0} q_a$, and in this way we see how a term of the form $n_F|x|$ will appear in the free energy (3.65).

To give the reader a better sense of how $SL(2,\mathbb{Z})$ acts on the polygon, we show $\mathcal{P}$ for ABJM theory and some of its $SL(2,\mathbb{Z})$ transforms in figure 3.3. While all three polygons have the same volume (as they must given Corollary 3.2.3), the eigenvalue distributions can look quite different. In figure 3.3a, $\rho(x)$ is a constant function for $-1/k < x < 1/k$. In figure 3.3b, $\rho(x)$ has two piecewise linear regions for $-1 < x < 1$. Finally, in figure 3.3c, $\rho(x)$ has one constant region and two linear regions in the interval $-1 + 1/k < x < 1 - 1/k$.

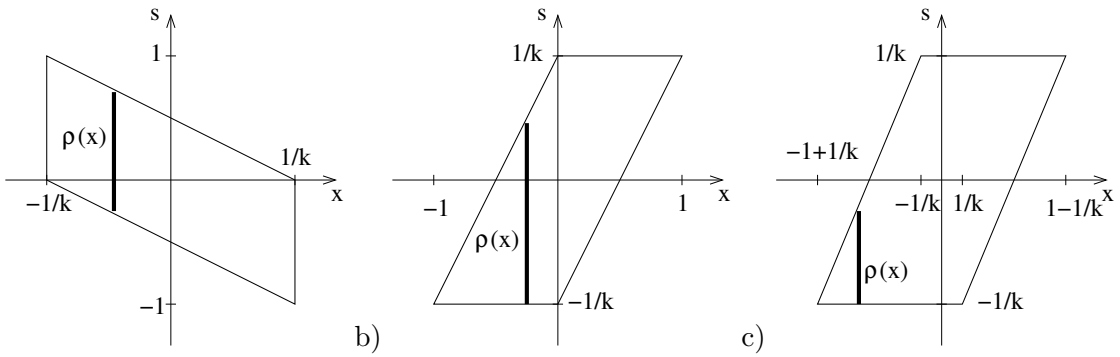


Figure 3.3: a) The polygon for ABJM theory which can be built from a $(1,0)$ and $(1,k)$ brane; b) the S-dual configuration involving a $(0,1)$ and $(-k,1)$ brane; c) an $SL(2,\mathbb{Z})$ transform to a $(1,1)$ and $(1-k,1)$ brane.

3.4.1 (p,q)-branes at finite N

One can go a little further and conjecture a finite N analog of the matrix integral (3.34) that reduces to (3.65) in the large N limit. If we move the sinh and $N!$ factors from (3.35) to (3.36), then the contribution from a $(1,q_a)$-brane to the partition function is given by

$$\begin{aligned}\frac{1}{N!}\exp\left[\pi i q_a\left(\sum_i \lambda_{a-1,i}^2-\sum_i\lambda_{a,i}^2\right)\right]\\ \times\frac{\prod_{i<j}2\sinh\pi(\lambda_{a-1,i}-\lambda_{a-1,j})\prod_{i<j}2\sinh\pi(\lambda_{a,i}-\lambda_{a,j})}{\prod_{i,j}2\cosh\pi(\lambda_{a-1,i}-\lambda_{a,j})}\,.\end{aligned}\tag{3.67}$$

We argue that the generalization to (p_a,q_a) branes is given by

$$\begin{aligned}L_{(p_a,q_a)}(\lambda_{a-1},\lambda_a)\equiv&\frac{1}{N!}|p_a|^{-N}\exp\left[\pi i\frac{q_a}{p_a}\left(\sum_i\lambda_{a-1,i}^2-\sum_i\lambda_{a,i}^2\right)\right]\\ &\times\frac{\prod_{i<j}2\sinh\frac{\pi}{p_a}(\lambda_{a-1,i}-\lambda_{a-1,j})\prod_{i<j}2\sinh\frac{\pi}{p_a}(\lambda_{a,i}-\lambda_{a,j})}{\prod_{i,j}2\cosh\frac{\pi}{p_a}(\lambda_{a-1,i}-\lambda_{a,j})}\,,\end{aligned}\tag{3.68}$$

where this formula is correct only when (p_a,q_a) are relatively prime. An (np_a,nq_a) five-brane, with (p_a,q_a) relatively prime, should be thought of as n (p_a,q_a) five-branes. Eq. (3.68) is based largely on the structure of (3.65). The term involving the hyperbolic cosine in (3.67) gives rise to the second term in (3.65) in the case $p_a=1$. As the imaginary parts of the eigenvalues appear with a factor of p_a in (3.65), at finite N, we should divide the eigenvalue differences $\lambda_{a-1,i}-\lambda_{a,j}$ by p_a. The exponential term in (3.67) gives rise to the first term in (3.65) in the case $p_a=1$. As a result at finite N we should replace the coefficient q_a with

q_a/p_a in (3.67). This replacement suggests that the field theory on (p_a, q_a)-branes in some sense can be thought of as having Chern-Simons couplings $\frac{1}{4\pi}(q_{a+1}/p_{a+1} - q_a/p_a)\int \mathrm{tr}\, A_a \wedge dA_a$. (Such an observation was also made in [73, 80].) Recall that the factor $\exp[\pi i(q_{a+1} - q_a)\sum_i \lambda_{a,i}^2]$ in the original matrix model comes from the classical contribution of the Chern-Simons term $\frac{1}{4\pi}(q_{a+1} - q_a)\int \mathrm{tr}\, A_a \wedge dA_a$ and its supersymmetric completion.

The remaining factors of p_a are required for (3.68) to be invariant under S-duality as we will now see. As an added bonus, by studying the action of S-duality, we will be able to deduce the partition function for the $T(U(N))$ theory of [73].

Our (p, q)-brane construction exists in type IIB string theory which is well known to be invariant under the action of $SL(2, \mathbb{Z})$. One of the generators of $SL(2, \mathbb{Z})$ is S-duality which we define to map a (p, q)-brane to a $(-q, p)$-brane. The work of Gaiotto and Witten [73] suggests that we also should be able to realize S-duality locally, on one (p, q)-brane at a time. The fact that S-duality squares to minus one suggests that we may be able to realize it as a Fourier transform acting on (3.68).

For simplicity, we will restrict to the case where the ranks of the gauge groups are equal to N. Before introducing the Fourier transform, we make use of the identity

$$\begin{aligned}&\frac{\prod_{i<j}\sinh(x_j - x_i)\prod_{i<j}\sinh(y_j - y_i)}{\prod_{i,j}\cosh(x_i - y_j)}\\ &= \det\begin{pmatrix} \mathrm{sech}(x_1 - y_1) & \mathrm{sech}(x_2 - y_1) & \dots & \mathrm{sech}(x_n - y_1) \\ \mathrm{sech}(x_1 - y_2) & \mathrm{sech}(x_2 - y_2) & \dots & \mathrm{sech}(x_n - y_2) \\ \vdots & \vdots & dots & \dots \\ \mathrm{sech}(x_1 - y_N) & \mathrm{sech}(x_2 - y_N) & \dots & \mathrm{sech}(x_N - y_N) \end{pmatrix}\end{aligned} \tag{3.69}$$

previously considered in [85]. Given this identity, we can write (3.68) as a sum over permutations

$$L_{(p,q)}(\lambda, \sigma) = \frac{1}{N!}\sum_{\rho\in S_N}(-1)^\rho \frac{1}{|p|}\prod_j \exp\left[\pi i \frac{q}{p}(\lambda_j^2 - \sigma_{\rho(j)}^2)\right]\frac{1}{2\cosh\frac{\pi}{p}(\lambda_j - \sigma_{\rho(j)})}\ . \tag{3.70}$$

We claim that a local S-duality is implemented by the following Fourier transform:

$$L_{(-q,p)}(\mu,\nu) = \int e^{2\pi i\mu\cdot\lambda} L_{(p,q)}(\lambda,\sigma) e^{-2\pi i\nu\cdot\sigma} d^N\lambda\, d^N\sigma \ . \tag{3.71}$$

To demonstrate this claim, we isolate the integrals over λ_j and $\sigma_{\rho(j)}$:

$$I = \frac{1}{|p|}\int d\lambda_j d\sigma_{\rho(j)} e^{\pi i\left(\frac{q}{p}(\lambda_j^2-\sigma_{\rho(j)}^2)+2(\mu_j\lambda_j-\nu_{\rho(j)}\sigma_{\rho(j)})\right)} \frac{1}{\cosh\frac{\pi}{p}(\lambda_j-\sigma_{\rho(j)})} \ . \tag{3.72}$$

With the change of variables $x_\pm = \lambda_j \pm \sigma_{\rho(j)}$, this integral is straightforward to perform:

$$\begin{aligned} I &= \frac{1}{2|p|}\int dx_+\, dx_-\, e^{\pi i\left(\frac{q}{p}x_+x_- + x_+(\mu_j-\nu_{\rho(j)}) + x_-(\mu_j+\nu_{\rho(j)})\right)} \operatorname{sech}\frac{\pi x_-}{p} \\ &= \frac{1}{|p|}\int dx_-\,\delta\left(\frac{q}{p}x_- + \mu_j - \nu_{\rho(j)}\right) e^{\pi i x_-(\mu_j+\nu_{\rho(j)})} \operatorname{sech}\frac{\pi x_-}{p} \\ &= \frac{1}{|q|}\exp\left[-\pi i\frac{p}{q}(\mu_j^2-\nu_{\rho(j)}^2)\right]\frac{1}{\cosh\frac{\pi}{q}(\mu_j-\nu_{\rho(j)})} \ . \end{aligned} \tag{3.73}$$

Taking the product over the eigenvalues and averaging over permutations yields (3.71).

This local S-duality composes in a nice way. Consider applying similar Fourier transforms to neighboring (p,q)-branes:

$$\begin{aligned} &\int L_{(-q,p)}(\lambda,\mu) L_{(-q',p')}(\mu,\nu) d^N\mu \\ &= \int e^{2\pi i\lambda\cdot\tilde\lambda} L_{(p,q)}(\tilde\lambda,\mu_1) e^{2\pi i\mu\cdot(\mu_2-\mu_1)} L_{(p',q')}(\mu_2,\tilde\nu) e^{-2\pi i\tilde\nu\cdot\nu} d^N\tilde\lambda\, d^N\mu_1\, d^N\mu\, d^N\mu_2\, d^N\tilde\nu \\ &= \int e^{2\pi i\lambda\cdot\tilde\lambda} L_{(p,q)}(\tilde\lambda,\tilde\mu) L_{(p',q')}(\tilde\mu,\tilde\nu) e^{-2\pi i\tilde\nu\cdot\nu} d^N\tilde\lambda\, d^N\tilde\nu\, d^N\tilde\mu \ . \end{aligned} \tag{3.74}$$

Thus if we apply a local S-duality to each (p_a,q_a)-brane in the necklace, the factors of $e^{2\pi i\mu\cdot\lambda}$ cancel out and the resulting partition function is invariant under a global action of S-duality.

We would like to give a better interpretation of this group action. Consider acting on a single (p,q)-brane with this local S-duality:

$$\int L_{(-q,p)}(\lambda,\mu) L_{(p',q')}(\mu,\nu) d^N\mu = \int e^{2\pi i\lambda\cdot\tilde\lambda} L_{(p,q)}(\tilde\lambda,\tilde\mu) e^{-2\pi i\mu\cdot\tilde\mu} L_{(p',q')}(\mu,\nu) d^N\tilde\lambda\, d^N\tilde\mu\, d^N\mu \ . \tag{3.75}$$

One way of interpreting the $e^{-2\pi i\mu\cdot\tilde{\mu}}$ is to posit that some new object has been inserted between the (p,q)-brane and the (p',q')-brane that implements a local S-duality. This object contributes to the partition function an amount

$$L_S(\mu,\tilde{\mu}) \equiv e^{-2\pi i\mu\cdot\tilde{\mu}} \ . \tag{3.76}$$

Similarly, right before the (p,q) brane we introduced another object that undoes the local S-duality:

$$L_{S^{-1}}(\lambda,\tilde{\lambda}) \equiv e^{2\pi i\lambda\cdot\tilde{\lambda}} \, . \tag{3.77}$$

Let us see how the $T(U(N))$ theory arises. So far, we have been thinking of the (p,q)-branes as the building blocks out of which we construct the partition function. Alternately, we can decompose the partition function into the contributions from the D3-brane segments and associated $U(N)$ gauge groups. From the D3-brane point of view, the object $L_S(\mu,\tilde{\mu})$ implementing S-duality splits a D3-brane segment into two regions, each characterized by a $U(N)$ gauge theory, one with eigenvalues μ and one with eigenvalues $\tilde{\mu}$. The $L_{(p,q)}$ and $L_{(p',q')}$ factors do not have enough factors of hyperbolic sine to describe two $U(N)$ gauge theories. These factors are simple to recover if we say the new object has a partition function

$$\tilde{L}_S(\mu,\tilde{\mu}) = N!\, e^{-2\pi i\mu\cdot\tilde{\mu}} \prod_{i<j} \frac{1}{4\sinh\pi(\mu_i-\mu_j)\sinh\pi(\tilde{\mu}_i-\tilde{\mu}_j)} \ . \tag{3.78}$$

This object has a natural interpretation as the partition function of the $T(U(N))$ theory.[5] This partition function has been found independently by [74].

At this point, it is clear that we should be able to realize any element of $SL(2,\mathbb{Z})$ acting locally on our necklace theories. The group $SL(2,\mathbb{Z})$ has two generators: S, which we discussed above, and T. We define T to send $q \to q+p$ and leave p invariant. If we think of (p,q) as a two-component column vector on which $SL(2,\mathbb{Z})$ acts in the fundamental

[5] Actually, the $T(U(N))$ partition function needs to be antisymmetrized with respect to permuting μ. But since $L_{(p,q)}(\mu,\lambda)$ is already antisymmetric in μ, we can get away with not antisymmetrizing L_S.

representation, then S and T are the two-by-two matrices

$$S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \qquad T = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}. \tag{3.79}$$

They satisfy $S^2 = -1$ and $(ST)^3 = 1$. To find the action of T on the matrix model, one can see from (3.70) that

$$L_{(p,q+p)}(\lambda, \sigma) = e^{\pi i \lambda \cdot \lambda - i\theta} L_{(p,q)}(\lambda, \sigma) e^{-\pi i \sigma \cdot \sigma + i\theta}, \tag{3.80}$$

where θ is a phase to be determined. Similarly to L_S, one can therefore define

$$L_T(\sigma) \equiv e^{-\pi i \sigma \cdot \sigma + i\theta}, \qquad L_{T^{-1}}(\lambda) \equiv e^{\pi i \lambda \cdot \lambda - i\theta}, \tag{3.81}$$

so one can describe the contribution of a $(p, q+p)$ five-brane to the matrix model integrand as $L_{T^{-1}}(\lambda) L_{(p,q)}(\lambda, \sigma) L_T(\sigma)$, where the factor of $L_T(\sigma)$ corresponds to the local action of a T-transformation and $L_{T^{-1}}(\lambda)$ corresponds to the local action of T^{-1}.

We have defined the effect of the S and T generators on the matrix model so that S requires us to have two distinct sets of eigenvalues μ and $\tilde{\mu}$ in the two regions separated by an "S-boundary", while T doesn't, acting just by multiplication of the eigenvalues σ in the region containing the "T-boundary" by $L_T(\sigma)$. We could have said, however, that a T-boundary also requires two distinct sets of eigenvalues σ and $\tilde{\sigma}$ around it, in which case we should have described its contribution as $L_T(\sigma)\delta(\sigma - \tilde{\sigma})$. By the same logic, it follows that the identity element should be also viewed as a Dirac delta function $L_1(\lambda, \sigma) = \delta(\lambda - \sigma)$. The operator equal to minus the identity is $L_{-1}(\lambda, \sigma) = \delta(\lambda + \sigma)$. We can verify by explicit computation that

$$L_{SS^{-1}}(\mu, \lambda) = \int e^{2\pi i \sigma \cdot (\lambda - \mu)} d^N \sigma = \delta(\mu - \lambda) = L_1(\mu, \lambda), \tag{3.82}$$

and similarly that $L_{TT^{-1}} = L_1$, $L_{S^2} = L_{-1}$ and $L_{(ST)^3} = L_1$ provided $3\theta = \pi N/4$.

We should mention that the expressions (3.76), (3.77), and (3.81) can also be justified by analyzing the case $N = 1$, where the action of $SL(2,\mathbb{Z})$ on Chern-Simons theories was described in [73, 86]. Indeed, as explained in [86], the T generator just shifts the CS level by one unit, so the action changes by $-\frac{1}{4\pi}\int A \wedge dA$ plus its supersymmetric completion. The classical contribution to the partition function from the scalar λ in the $\mathcal{N} = 2$ vector multiplet then gives $L_T(\lambda)$. The action of S on a CS theory with some gauge field $\tilde{A}$ consists of introducing another dynamical gauge field A that couples to the topological current $*d\tilde{A}$. This coupling takes the form of an off-diagonal Chern-Simons term $-\frac{1}{2\pi}\int A \wedge d\tilde{A}$ plus its supersymmetric completion. If μ and $\tilde{\mu}$ are the scalars in the corresponding vector multiplets, the classical contribution to the partition function from this off-diagonal Chern-Simons term is precisely given by (3.76).

3.5 Discussion

Our main results are additional evidence presented in Section 3.2.3 for the F-theorem conjecture, the relations (3.4) and (3.5) between numbers of chiral operators and eigenvalue distributions, and a conjectured form (3.68) of the matrix model corresponding to a (p_a, q_a)-brane construction in type IIB string theory. Each of these results requires some brief discussion.

We would like to investigate further our proposed matrix model (3.68). In particular, it would be interesting to see how (3.68) transforms under Seiberg duality. One statement of the s-rule [87] is that a theory breaks supersymmetry for which there exists a Seiberg duality that produces a gauge group with a negative rank. Given the matrix model's status as a type of supersymmetric index, one expects that the partition function should vanish for theories that violate the s-rule [88]. Willett and Yaakov [89] have shown that the matrix model is invariant under Seiberg duality in the case $p_a = 1$.

Regarding the F-theorem, we have not constructed any explicit RG flows, either on the gauge theory side or, via the AdS/CFT correspondence, on the gravity side. Instead, we have posited the existence of reasonable seeming flows, and we have examined F at the IR and UV fixed points. For example, by adding a mass to fundamental flavors, one should

be able to flow from a theory with a $(0, q)$-brane in the UV to one without it in the IR. The corresponding volume of the tri-Sasaki Einstein manifold will increase, leading to a decrease in F. Similarly, we can consider an RG flow where a (p, q)-brane forms a bound state with a (p', q')-brane. Under such a flow, F will also decrease. Given the result of refs. [63, 64] described in the introduction, it seems likely that any gravity dual of an RG flow will obey the F-theorem. One way of interpreting our results, given that our flows also obey the F-theorem, is that it may be possible to realize these flows as solutions of eleven-dimensional supergravity.

Given that the operator counting relations (3.4) and (3.5) can be defined for essentially any KWY matrix model and corresponding superconformal field theory, one wonders if they hold more generally. In the next chapter, we will investigate the large N limit of the KWY matrix models for a number of other superconformal field theories, and we will find that indeed these relations are always satisfied. We will look at necklace quivers with additional adjoint and fundamental fields. We will look at a couple of non-necklace quivers, for example a $\mathbb{Z}_2 \times \mathbb{Z}_2$ orbifold theory. In order to look at theories with $\mathcal{N} = 2$ supersymmetry, we have to generalize the KWY matrix model to allow for arbitrary R-charges [25]. It turns out that in the theories we study, relations (3.4) and (3.5) are valid not just for the correct R-charges but for any R-charges compatible with the marginality of the superpotential.

One constraint in these investigations is that for chiral theories, the KWY matrix model does not seem to have a large N limit that is compatible with a dual eleven-dimensional supergravity description [21]. It will be interesting to see if the relations (3.4) and (3.5) can give any insight into how the matrix model might be modified to allow for such a limit.

After this work appeared as [1], Mariño and Putrov [35] gave an interpretation of the necklace quiver matrix model as a Fermi gas. It turns out that curves of constant energy E in the phase space of this gas are approximated by the polygons (3.26) when E is large.

Chapter 4

Operator Counting and Eigenvalue Distributions for 3D Supersymmetric Gauge Theories

This chapter is an edited version of ref. [2], written in collaboration with Christopher Herzog and Silviu Pufu.

4.1 Introduction

For those interested in superconformal gauge theories in three dimensions, the matrix model of Kapustin, Willett, and Yaakov [17] provides a powerful tool. Using this matrix model, one can compute the partition function and the expectation values of supersymmetric Wilson loops on a three sphere, even when the gauge theory is strongly interacting. The matrix model was derived through a localization procedure [18] that obscures the connection between matrix model quantities and microscopic degrees of freedom in the gauge theory. Given the success of the matrix model in post-dicting the $N^{3/2}$ large N scaling of the free energy[1] of maximally supersymmetric $SU(N)$ Yang-Mills theory at its infrared fixed point [90], it is a worthwhile goal to try to figure out the map between matrix model and

[1] By "free energy" we mean minus the logarithm of the path integral on S^3, with an appropriate subtraction of UV divergences.

gauge theory quantities in greater detail. In the previous chapter, we made some progress in understanding the relation between the eigenvalue distributions in the matrix model and the chiral ring of the supersymmetric gauge theory for the so-called necklace quivers, and we conjectured this relation would hold more generally. In this chapter, we work out further examples of field theories that obey this conjecture.

Let us begin by recalling the relation noticed in the previous chapter between the eigenvalue distribution in the matrix model and the chiral ring for the necklace quiver gauge theories. These field theories have $\mathcal{N} = 3$ supersymmetry (SUSY), gauge group $U(N)^d$, and associated Chern-Simons levels k_a, $a = 1, \ldots, d$, such that $\sum_a k_a = 0$. The matter sector consists of the bifundamental fields $X_{a,a+1}$ and $X_{a+1,a}$ that connect the gauge groups together into a circle (see figure 3.2). The localization procedure [17] reduces the partition function to an integral over d constant $N \times N$ matrices σ_a, where σ_a is the real scalar that belongs to the same $\mathcal{N} = 2$ multiplet as the gauge connection. In the large N limit, the matrix integral can be evaluated in the saddle point approximation. As was shown in [27], at the saddle point, the real parts of the eigenvalues $\lambda_j^{(a)}$ of σ_a grow as $N^{1/2}$ while their imaginary parts stay of order one as N is taken to infinity. In addition, to leading order in N the real parts of the eigenvalues are the same for each gauge group. Therefore, in order to find the saddle point one can consider the large N expansion

$$\lambda_j^{(a)} = N^{1/2} x_j + i y_{a,j} + \ldots \,. \tag{4.1}$$

As one takes $N \to \infty$, the x_j and $y_{a,j}$ become dense, and one can pass to a continuum description by defining the distributions

$$\rho(x) = \lim_{N\to\infty} \frac{1}{N} \sum_{j=1}^{N} \delta(x - x_j)\,, \qquad \rho(x) y_a(x) = \lim_{N\to\infty} \frac{1}{N} \sum_{j=1}^{N} y_{a,j} \delta(x - x_j)\,. \tag{4.2}$$

The saddle point is then found by extremizing a free energy functional $F[\rho, y_a]$ under the assumption that ρ is a density, namely that $\rho(x) \geq 0$ and $\int dx\, \rho(x) = 1$. It is convenient to enforce the latter constraint with a Lagrange multiplier μ that will appear in the formulae presented below. In general, $F[\rho, y_a]$ may be a non-local functional because the eigenvalues

could interact with one another through long-range forces, and if this is the case the saddle point equations are usually hard to solve. The key insight in solving the saddle point equations in [27] was that, luckily, in the continuum limit (4.2) the ansatz (4.1) leads to a *local* expression for $F[\rho, y_a]$ due to the cancellation of long-range forces. By solving the saddle point equations, it was shown in [27] that the distributions $\rho(x)$ and $\rho(x)[y_a(x) - y_b(x)]$ can be identified for any a and b with piecewise linear functions with compact support. While the free energy F can be calculated by evaluating the functional $F[\rho, y_a]$ on the saddle point configuration, it is also possible to calculate F by noticing that $F[\rho, y_a]$ satisfies a virial theorem that gives $F = 4\pi\mu N^{3/2}/3$.

The chiral ring of the necklace quiver gauge theories consists of gauge invariant products of the $X_{a,a+1}$ and $X_{a+1,a}$ fields and monopole operators modulo superpotential and monopole relations. While one can define monopole operators that turn on any number of flux units through each $U(N)$ gauge group, at large N the only relevant ones are the so-called "diagonal monopole operators" that turn on the same number of units of flux through the diagonal $U(1)$ subgroup of each $U(N)$ gauge group. Operators in the chiral ring therefore have an associated R-charge r and a (diagonal) monopole charge m. We can also introduce the the function $\psi_{X_{ab}}(r, m)$ that counts in the same way operators that don't vanish when the bifundamental field X_{ab} is set to zero.[2] In the previous chapter, we found the following relation between the saddle point eigenvalue distribution and the chiral ring:

$$\left.\frac{\partial^3\psi}{\partial r^2\partial m}\right|_{m=rx/\mu} = \frac{r}{\mu}\rho(x) \ , \tag{4.3a}$$

$$\left.\frac{\partial^2\psi_{X_{ab}}}{\partial r\partial m}\right|_{m=rx/\mu} = \frac{r}{\mu}\rho(x)[y_b(x) - y_a(x) + R(X_{ab})] \ . \tag{4.3b}$$

In other words, the matrix model eigenvalue density $\rho(x)$ and the quantity $\rho(x)[y_b(x) - y_a(x) + R(X_{ab})]$, which as mentioned above are linear functions of x, should be interpreted as derivatives of numbers of operators whose monopole charge to R-charge ratio is given by x/μ.

One of the goals of this chapter will be to provide further evidence for the conjectures (4.3) in superconformal theories with gravity duals that preserve only $\mathcal{N} = 2$ supersymmetry

[2] In our conventions, X_{ab} transforms under the $(\overline{\mathbf{N}}_a, \mathbf{N}_b)$ representation of $U(N_a) \times U(N_b)$.

as opposed to the $\mathcal{N} = 3$ SUSY of the necklace quivers studied in chapter 3. In an $\mathcal{N} = 2$ theory, the $U(1)$ R-symmetry can mix with other Abelian flavor symmetries, so the matter fields can have R-charges different from the canonical free-field value $1/2$. The generalization of the Kapustin-Willett-Yaakov matrix model to non-canonical R-charges was worked out in [25, 59]. Furthermore, since the $U(1)_R$ symmetry can now mix with other Abelian flavor symmetries, it was conjectured in [25] and proved in [26] that the correct R-symmetry in the IR can be found by extremizing the free energy F as a function of all trial R-charges that are consistent with the marginality of the superpotential.[3]

We find that eqs. (4.3) are satisfied for more general quiver gauge theories where the bi-fundamental matter multiplets are non-chiral, meaning that they come in pairs of conjugate representations of the gauge group. In the first half of section 4.4, we examine the necklace quiver gauge theories, this time with an arbitrary R-charge assignment consistent with the marginality of the superpotential. In the second half of section 4.4 and appendix B, we examine theories where we add flavor (meaning $\mathcal{N} = 2$ matter multiplets that transform in the fundamental or anti-fundamental representation of one of the gauge groups) to the maximally SUSY $\mathcal{N} = 8$ theory and to the $\mathcal{N} = 6$ ABJM theory of [28]. Lastly, in appendix B.2, we consider a theory that shares the same quiver with its $(3 + 1)$-dimensional cousin that has a $\mathbb{C}^3/\mathbb{Z}_2 \times \mathbb{Z}_2$ moduli space (see figure B.1). In all of these examples, eqs. (4.3) are satisfied for any choice of trial R-charges.

Another goal of this chapter is to relate the conjectures (4.3) to the observation made in [21,70] that, as checked in a number of examples, the relation (3.3) between the free energy and the volume of the internal space holds for any trial R-charges, and not just the ones that extremize F. That this relation[4] holds for any trial R-charges is surprising because only for the critical R-charges does there exist a known 11-d supergravity background $AdS_4 \times Y$. For non-critical R-charges, measured geometrically in terms of the volume of some corresponding five-cycles of Y, one can still identify a class of Sasakian metrics on Y and compute their volume. The volume $\mathrm{Vol}(Y)$ is a function of the Reeb vector of Y, which

[3] It was suggested in [21] that F might be a good measure of the number of degrees of freedom even in non-supersymmetric field theories. See also [91].

[4] A similar relation between the anomaly coefficient a computed with a set of trial R-charges and the volume of a 5-d Sasakian space Y is known to hold in theories with AdS_5 duals [92, 93].

parameterizes the way the $U(1)_R$ symmetry sits within the isometry group of Y. We show in section 4.3 that (4.3a) holds for some choice of trial R-charges if and only if eq. (3.3) holds for the same choice of trial R-charges for the matter fields and a range of R-charges for the monopole operators. We also show an analogous result that relates (3.5) to the volumes of five-dimensional sub-manifolds of Y. For a gauge invariant operator constructed from a closed loop of bifundamental fields X_{ab}, it must be true that $\sum_{X_{ab}}(y_a - y_b) = 0$. Given (4.3), there is a geometric version of this sum that must also vanish. The last part of section 4.3 explains why.

There are previously recognized difficulties, involving cancellation of long-range forces, in using the matrix model to study the large N limit of theories with chiral bifundamental fields [21]. We do not surmount these difficulties, but we investigate in section 4.5 what (4.3a) and (4.3b) predict for a theory with a moduli space that is a fibration over $\mathbb{C}^3/\mathbb{Z}_3$ (see figure 4.1). We also study a field theory that was conjectured to be dual to $AdS_4 \times Q^{2,2,2}/\mathbb{Z}_k$ in appendix B.4 (see figure B.3).

Some additional material relevant to this chapter is collected in the appendices. Appendix A.3 proves that the critical R-charges maximize F for the necklace quivers. Appendix B.3 reviews how to count gauge invariant operators for an Abelian gauge theory with a toric branch of its moduli space.

4.2 Matrix models at non-critical R-charges

4.2.1 Review of the large N limit

To understand what it means to consider non-canonical (or non-critical) R-charges, let us introduce some of the ideas developed recently in refs. [21, 25, 59]. Building on the work of [17], refs. [25,59] used localization to reduce the path integral of any $\mathcal{N} = 2$ Chern-Simons matter on S^3 to a matrix integral. By a Chern-Simons-matter theory we mean a theory constructed from some number d of $\mathcal{N} = 2$ vector multiplets with gauge groups G_a ($a = 1, \ldots, d$) and Chern-Simons kinetic terms $i\pi k_a \int \mathrm{tr}\, A_a \wedge dA_a$ + supersymmetric completion, as well as any number of $\mathcal{N} = 2$ chiral superfields transforming in representations R_i of the total gauge group $G = \prod_{a=1}^{d} G_a$. As mentioned in the introduction, one difference between

theories with $\mathcal{N} = 2$ supersymmetry and theories with more supersymmetry is that the R-charges Δ_i of the chiral fields at the IR superconformal fixed point are not fixed at the free field values $\Delta_i = 1/2$, so the free energy will generically depend on these R-charges. In fact, it was proposed in [25] that a prescription for finding the correct R-charges in the IR is to calculate the free energy F as a function of all possible R-charge assignments consistent with the marginality of the superpotential and to extremize F over the set of all such assignments.

Let us focus on the case where all gauge groups are $U(N)$ and index the gauge groups by $a = 1, \ldots, d$. Generalizing the techniques developed in [27], the authors of [21] used the saddle point approximation to evaluate the path integral on S^3 for a class of $\mathcal{N} = 2$ Chern-Simons-matter theories at large N that satisfy the following five conditions:

1. The CS levels sum to zero: $\sum_{a=1}^{d} k_a = 0$.

2. Any matter field X transforms either in the $\mathbf{N}_a$, or $\overline{\mathbf{N}}_b$, or $(\mathbf{N}_a, \overline{\mathbf{N}}_b)$ representation for some a and b.

3. The total number of fundamental fields equals the total number of anti-fundamental fields.

4. For any bifundamental field X transforming in $(\mathbf{N}_a, \overline{\mathbf{N}}_b)$, there exists another bifundamental field $\tilde{X}$ transforming in the conjugate representation $(\mathbf{N}_b, \overline{\mathbf{N}}_a)$.

5. For each gauge group a we have

$$\sum_{X \text{ in } (\mathbf{N}_a, \overline{\mathbf{N}}_b)} (R[X] - 1) + \sum_{\tilde{X} \text{ in } (\mathbf{N}_b, \overline{\mathbf{N}}_a)} \Big(R[\tilde{X}] - 1 \Big) = -2 \,. \tag{4.4}$$

This last condition is sufficient to guarantee the vanishing of the long-range forces on the eigenvalues in the saddle point approximation. Interestingly, this condition has appeared before in the context of superconformal $(3 + 1)$-dimensional gauge theories. The condition (4.4) would imply that the NSVZ beta function of gauge group a vanishes [70]. For quiver gauge theories with a toric moduli space, bifundamental fields appear in exactly two terms

in the superpotential. Thus, if we sum (4.4) over a, we find the condition that

$$(\#\text{ of gauge groups}) - (\#\text{ of bifundamentals}) + (\#\text{ of superpotential terms}) = 0\ . \quad (4.5)$$

In other words the quiver may give a triangulation of a torus where the faces of the triangulation are superpotential terms [8].

If these five conditions are satisfied, one can take the $N \to \infty$ limit as described in eqs. (4.1) and (4.2) in the introduction. The free energy is the extremum of the free energy functional

$$\begin{aligned} F[\rho(x), y_a(x)] &= 2\pi N^{3/2} \int dx\, x\rho(x) \left(\sum_{a=1}^{d} k_a y_a(x) + \Delta_m \right) \\ &+ 2\pi N^{3/2} \int dx\ |x|\, \rho(x) \left[\sum_{X \text{ in } \mathbf{N}_a} \left(\frac{1 - R[X]}{2} - \frac{1}{2} y_a(x) \right) + \sum_{X \text{ in } \overline{\mathbf{N}}_b} \left(\frac{1 - R[X]}{2} + \frac{1}{2} y_b(x) \right) \right] \\ &+ \frac{\pi N^{3/2}}{3} \int dx\, \rho(x)^2 \sum_{X \text{ in } (\mathbf{N}_a, \overline{\mathbf{N}}_b)} (\delta y_{ab}(x) + R[X])(\delta y_{ab}(x) + R[X] - 1)(\delta y_{ab}(x) + R[X] - 2)\,, \end{aligned} \quad (4.6)$$

where $\delta y_{ab}(x) \equiv y_a(x) - y_b(x)$. This formula was derived assuming the bifundamental fields satisfy $0 \leq R[X] + \delta y_{ab}(x) \leq 2$. Extra care must be taken when $R[X] + \delta y_{ab} = 0$ or 2 because in these cases the discrete nature of the eigenvalues becomes important, and the equation of motion derived from varying (4.6) with respect to $\delta y_{ab}(x)$ need not hold.

Generically, the functional 4.6 has many flat directions. The following d of them play an important role in this thesis because they correspond to changing the R-charges of the matter fields by linear combinations of the gauge charges with respect to the diagonal

$U(1) \in U(N)_a$:

$$\begin{aligned}
y_a(x)\text{:} &\qquad y_a(x) \to y_a(x) - \delta^{(a)}\,, \\
\text{chiral superfield } X \text{ in } \mathbf{N}_a\text{:} &\qquad R[X] \to R[X] + \delta^{(a)}\,, \\
\text{chiral superfield } X \text{ in } \overline{\mathbf{N}}_b\text{:} &\qquad R[X] \to R[X] - \delta^{(b)}\,, \\
\text{chiral superfield } X \text{ in } (\mathbf{N}_a, \overline{\mathbf{N}}_b)\text{:} &\qquad R[X] \to R[X] + \delta^{(a)} - \delta^{(b)}\,, \\
\Delta_m\text{:} &\qquad \Delta_m \to \Delta_m + \sum_a k_a \delta^{(a)}\,.
\end{aligned} \tag{4.7}$$

See [21] for a more detailed discussion of these flat directions and their AdS/CFT interpretation.

The Δ_m appearing in (4.6) is the bare R-charge of the "diagonal" monopole operator $T^{(1)}$. A monopole operator $T_a^{(q_a)}$ turns on q_a units of $\mathrm{tr}\, F_a$ flux through a two-sphere surrounding the insertion point. Diagonal monopole operators $T^{(m)}$ turn on the same number m of $\mathrm{tr}\, F_a$ flux units in each gauge group. At large N, only the diagonal monopole operators are important.

We will usually impose the constraint $\int dx\, \rho(x) = 1$ by introducing a Lagrange multiplier μ and defining the functional

$$\tilde{F}[\rho, y_a, \mu] = F[\rho, y_a] - 2\pi N^{3/2} \mu \left(\int dx\, \rho(x) - 1 \right) . \tag{4.8}$$

This functional should be extremized with respect to $\rho(x)$, $y_a(x)$, and μ.

4.2.2 Flavored theories

In all gauge theories that we examine in this chapter the fundamental and anti-fundamental fields q_α and $\tilde{q}_\alpha$ appear in the superpotential as

$$\delta W = \sum_\alpha \mathrm{tr}\, [q_\alpha \mathcal{O}_\alpha \tilde{q}_\alpha] \,, \tag{4.9}$$

where $\mathcal{O}_\alpha$ are polynomials in the bifundamental fields. It was conjectured in [32, 33] that if this is case then the diagonal monopole operators $T^{(m)}$ satisfy the following OPE:

$$T^{(m)}T^{(-m)} = \left(\prod_\alpha \mathcal{O}_\alpha\right)^{|m|} . \tag{4.10}$$

This OPE was conjectured in part because a parity anomaly argument shows that the monopole operators have gauge charges

$$g_a[T^{(m)}] = mk_a + \frac{|m|}{2}\sum_\alpha g_a[\mathcal{O}_\alpha] \tag{4.11}$$

with respect to the diagonal $U(1) \subset U(N)_a$, and R-charges

$$R[T^{(m)}] = m\Delta_m + \frac{|m|}{2}\sum_\alpha R[\mathcal{O}_\alpha] . \tag{4.12}$$

Using the fact that each term in (4.9) must be gauge-invariant and have R-charge two, we have $R[q_\alpha] + R[\tilde{q}_\alpha] + R[\mathcal{O}_\alpha] = 2$ and $g_a[q_\alpha] + g_a[\tilde{q}_\alpha] + g_a[\mathcal{O}_\alpha] = 0$ for any a. One can use these relations to eliminate the sum over the flavor fields in (4.6):

$$\begin{aligned} F[\rho(x), y_a(x)] = {} & 2\pi N^{3/2}\int dx\ |x|\,\rho(x)\left(R[T^{(\operatorname{sgn} x)}] + \sum_{a=1}^{d} y_a(x) g_a[T^{(\operatorname{sgn} x)}]\right) \\ & + \frac{\pi N^{3/2}}{3}\int dx\,\rho(x)^2 \sum_{X \text{ in } (\mathbf{N}_a, \overline{\mathbf{N}}_b)} (\delta y_{ab}(x) + R[X])(\delta y_{ab}(x) + R[X] - 1)(\delta y_{ab}(x) + R[X] - 2) . \end{aligned} \tag{4.13}$$

4.3 An Equivalent form of our conjecture

4.3.1 Eigenvalue density and volumes of Sasakian spaces

We can relate the conjecture (4.3) to the observation that eq. (3.3) holds for any trial R-charges. In particular, we prove the following result: In a CS-matter theory dual to $AdS_4 \times Y$ fix a set of matter R-charges $R[X]$ and a bare monopole charge Δ_m so that the conformal dimensions of all gauge-invariant operators satisfy the unitarity bound. Let

$\rho(x)$, μ, and $\psi(r,m)$ be as defined in the introduction, and let's assume $\rho(x)$ has compact support. The following two statements are equivalent:

1. The conjecture (4.3a) holds for the given R-charges $R[X]$ and bare monopole charge Δ_m.

2. For any δ in a small enough neighborhood of zero, we have

$$\lim_{N\to\infty} \frac{2\pi^6 N^3}{27F^2} = \mathrm{Vol}(Y,\delta)\,, \tag{4.14}$$

   where the free energy F of the CS-matter theory and the volume $\mathrm{Vol}(Y,\delta)$ of the internal space Y are both computed assuming that the matter R-charges are $R[X]$ and the bare monopole charge is $\Delta_m + \delta$.

For notational convenience, let's denote the LHS of eq. (4.14) by $\mathrm{Vol}_m(Y,\delta)$ and let's introduce the rescaled matrix model quantities:

$$\hat{x} = \frac{x}{\mu}\,, \qquad \hat{\rho}(\hat{x}) = \frac{\rho(x)}{\mu}\,, \qquad \hat{y}_a(\hat{x}) = y_a(x)\,. \tag{4.15}$$

The equivalence between (1) and (2) follows from the following two equations:

$$\mathrm{Vol}_m(Y,\delta) = \frac{\pi^4}{24}\int d\hat{x}\,\frac{\hat{\rho}(\hat{x})}{(1+\hat{x}\delta)^3}\,, \tag{4.16}$$

$$\mathrm{Vol}(Y,\delta) = \frac{\pi^4}{24}\int d\hat{x}\,\frac{\lim_{r\to\infty}\psi^{(2,1)}(r,r\hat{x})/r}{(1+\hat{x}\delta)^3}\,, \tag{4.17}$$

which we prove in sections 4.3.2 and 4.3.3, respectively.

Assuming the eqs. (4.16) and (4.17) to be true, it is clear that the statement (1) implies (2). That (2) implies (1) follows from the fact that knowing $\mathrm{Vol}_m(Y,\delta)$ for δ in a small neighborhood of zero, one can reconstruct $\hat{\rho}(\hat{x})$, and analogously, from $\mathrm{Vol}(Y,\delta)$ one can reconstruct $\lim_{r\to\infty}\psi^{(2,1)}(r,r\hat{x})/r$. Indeed, one can extend $\mathrm{Vol}_m(Y,\delta)$ to any complex δ as an analytic function with singularities. We assume that $\hat{\rho}$ is supported on $[\hat{x}_-,\hat{x}_+]$ for some $\hat{x}_- < 0 < \hat{x}_+$. We see from eq. (4.16) that the integral converges absolutely if

$\delta \in (-1/\hat{x}_+, -1/\hat{x}_-)$ or $\delta \notin \mathbb{R}$, so $\mathrm{Vol}_m(Y,\delta)$ can only have singularities on $(-\infty, -1/\hat{x}_+] \cup [-1/\hat{x}_-, \infty)$.

To relate the singularities of $\mathrm{Vol}_m(Y,\delta)$ to $\hat{\rho}(\hat{x})$ we can perform two integrations by parts in (4.25)

$$\mathrm{Vol}_m(Y,\delta) = \frac{\pi^4}{48\delta^3} \int d\hat{x}\, \frac{\hat{\rho}''(\hat{x})}{\hat{x} + \frac{1}{\delta}} \tag{4.18}$$

for any $\delta \in \mathbb{C} \setminus ((-\infty, -1/\hat{x}_+] \cup [-1/\hat{x}_-, \infty))$. Generically, eq. (4.18) shows that $\mathrm{Vol}_m(Y,\delta)$ has two branch cuts, one on $(-\infty, -1/\hat{x}_+]$ and one on $[-1/\hat{x}_-, \infty)$. From the discontinuities of $\mathrm{Vol}_m(Y,\delta)$ one can read off $\hat{\rho}''(-1/\delta)$. Simple poles of $\mathrm{Vol}_m(Y,\delta)$ at $\delta = -1/\hat{x}'$ correspond to contributions to $\hat{\rho}''(\hat{x})$ proportional to $\delta(\hat{x} - \hat{x}')$; second order poles of $\mathrm{Vol}_m(Y,\delta)$ at $\delta = -1/\hat{x}'$ correspond to $\delta'(\hat{x} - \hat{x}')$, etc. From the singularities of the analytic continuation of $\mathrm{Vol}_m(Y,\delta)$ one can therefore reconstruct uniquely $\hat{\rho}''(\hat{x})$, and hence $\hat{\rho}(\hat{x})$, and similarly for $\mathrm{Vol}(Y,\delta)$ and $\lim_{r\to\infty} \psi^{(2,1)}(r, r\hat{x})/r$. If $\mathrm{Vol}_m(Y,\delta)$ and $\mathrm{Vol}(Y,\delta)$ agree on an open set, then (1) holds.

In our examples, $\mathrm{Vol}_m(Y,\delta)$ is a rational function of δ with poles of order at most three, so $\hat{\rho}(\hat{x})$ is piecewise linear and it may have delta-functions. From the location and residues of the poles one can first reconstruct $\hat{\rho}''(\hat{x})$, and then $\hat{\rho}(\hat{x})$ by integrating $\hat{\rho}''(\hat{x})$ twice. To perform this reconstruction starting with $\mathrm{Vol}_m(Y,\delta)$, one first decomposes $\mathrm{Vol}_m(Y,\delta)$ into partial fractions, and then identifies the terms in $\hat{\rho}''(\hat{x})$ that give those partial fractions: if, for example,

$$\mathrm{Vol}_m(Y,\delta) = \frac{\pi^4}{48\delta^2} \sum_i \frac{a_i}{1 + \hat{x}_i\delta} - \frac{\pi^4}{48\delta} \sum_i \frac{b_i}{(1 + \hat{x}_i\delta)^2} \tag{4.19}$$

for some $\hat{x}_i$, then

$$\hat{\rho}''(\hat{x}) = \sum_i a_i \delta(\hat{x} - \hat{x}_i) + \sum_i b_i \delta'(\hat{x} - \hat{x}_i)\,. \tag{4.20}$$

4.3.2 Matrix model dependence on δ

In this subsection we prove the result (4.16). As we have seen in the previous section, the matrix model generally takes the form

$$\begin{aligned}\tilde{F}[\rho, y_a, \mu] = &\int dx\, \rho(x)^2 f(y_a(x)) - \int dx\, \rho(x) V(x, y_a(x)) \\ &+ 2\pi N^{3/2} \int dx\ |x|\, \rho(x) R[T^{(\mathrm{sgn}\, x)}] - 2\pi N^{3/2} \mu \left(\int dx\, \rho(x) - 1 \right) ,\end{aligned} \tag{4.21}$$

for some functions f and V. While the explicit form of these function is given in (4.13), their precise form doesn't matter. The only property of V that we will use is that it is homogeneous of degree one in x, namely $V(\lambda x, y_a(x)) = \lambda V(x, y_a(x))$ for any $\lambda > 0$. With the rescaling (4.15), one can write $\tilde{F}$ as

$$\begin{aligned}\tilde{F}[\hat{\rho}, \hat{y}_a, \mu] = -2\pi N^{3/2}\mu + \mu^3 \int d\hat{x}\, \hat{x}^2 \Bigg[&\frac{\hat{\rho}(\hat{x})^2}{\hat{x}^2} f(\hat{y}_a(\hat{x})) - \frac{\hat{\rho}(\hat{x})}{\hat{x}} \frac{V(\hat{x}, \hat{y}_a(\hat{x}))}{\hat{x}} \\ &+ 2\pi N^{3/2} \frac{\hat{\rho}(\hat{x})}{|\hat{x}|} \left(R[T^{(\mathrm{sgn}\, \hat{x})}] - \frac{1}{|\hat{x}|} \right) \Bigg] .\end{aligned} \tag{4.22}$$

The rescaling (4.15) is useful because now the equations of motion for $\hat{\rho}$ and $\hat{y}_a$ do not involve μ. One can first solve these equations, and then μ can be found by integrating $\hat{\rho}$: the normalization condition $\int dx\, \rho(x) = 1$ becomes

$$\int d\hat{x}\, \hat{\rho}(\hat{x}) = \frac{1}{\mu^2} . \tag{4.23}$$

We now see that if we extremized (4.22) in the case where the monople R-charges were $R[T^{(\pm 1)}]$, we could obtain the saddle point when they are $R[T^{(\pm 1)}] \pm \delta^{(\pm 1)}$ through the transformation:

$$\begin{aligned} \frac{\hat{\rho}_\delta(\hat{x}_\delta)}{\hat{x}_\delta} &= \frac{\hat{\rho}(\hat{x})}{\hat{x}} , & \frac{1}{\hat{x}_\delta} &= \frac{1}{\hat{x}} + \delta^{(\mathrm{sgn}\, \hat{x})} , \\ \hat{y}_{a,\delta}(\hat{x}_\delta) &= \hat{y}_a(\hat{x}) , & R[T_\delta^{(\pm 1)}] &= R[T^{(\pm 1)}] \pm \delta^{(\pm 1)} . \end{aligned} \tag{4.24}$$

Indeed, the equations of motion for $\hat{\rho}$ and $\hat{y}_a$ are obtained by extremizing the expression in the square brackets in (4.22), and this expression is invariant under (4.24). Given that

$\hat{\rho}$ has compact support, the transformations (4.24) make sense only when $\delta^{(\pm)}$ are small enough.

For simplicity, from now on let's restrict ourselves to the case $\delta^{(+1)} = \delta^{(-1)} = \delta$, even though one can make similar arguments for the case where $\delta^{(+1)}$ and $\delta^{(-1)}$ are arbitrary or satisfy a different relation. In chapter 3, we showed that $F = 4\pi N^{3/2}\mu/3$, which implies that

$$\mathrm{Vol}_m(Y,\delta) = \frac{\pi^4}{24\mu_\delta^2} = \frac{\pi^4}{24}\int d\hat{x}_\delta\, \hat{\rho}_\delta(\hat{x}_\delta) = \frac{\pi^4}{24}\int d\hat{x}\, \frac{\hat{\rho}(\hat{x})}{(1+\hat{x}\delta)^3}\,. \tag{4.25}$$

4.3.3 Operator counting dependence on δ

We now prove the result (4.17). Let A be the chiral ring associated to the superconformal field theory dual to $AdS_4 \times Y$ in the Abelian case $N = 1$. A is also a vector space over $\mathbb{C}$ that is graded by the R-charge and monopole charge, meaning that one can define a basis of operators with well-defined R-charge and monopole charge. Let $A_{m,r}$ be the vector subspace of elements of A with monopole charge m and R-charge r. We introduce the Hilbert-Poincaré series

$$f(t,u) = \sum_{m,r} \dim(A_{m,r}) t^r u^m\,. \tag{4.26}$$

Since the Abelian moduli space of the gauge theory is the Calabi-Yau cone over Y one can view the operators in the chiral ring as holomorphic functions on this cone. Martelli, Sparks, and Yau [77] show that

$$\mathrm{Vol}(Y,\delta) = \frac{\pi^4}{48}\lim_{t\to 1}(1-t)^4 f(t,t^\delta)\,. \tag{4.27}$$

One can compute the Hilbert-Poincaré series for Y in terms of $\psi(r,m)$, the number of operators with R-charge at most r and monopole charge at most m. Approximating ψ by a continuous function of homogeneous degree four, the definition (4.26) gives

$$f(t,u) \approx \int dr\, dm\, \psi^{(1,1)}(r,m) t^r u^m\,. \tag{4.28}$$

Since $1-t \approx -\ln t$ for $t \approx 1$, we can use (4.27) and (4.28) to write $\mathrm{Vol}(Y,\delta)$ as

$$\begin{aligned}\mathrm{Vol}(Y,\delta) &\approx \frac{\pi^4}{48}(\ln t)^4 \int dr\, dm\, \psi^{(1,1)}(r,m) t^{r+m\delta} \\ &= -\frac{\pi^4}{48}(\ln t)^3 \int dr\, dm\, \psi^{(2,1)}(r,m) t^{r+m\delta} \\ &= -\frac{\pi^4}{48}(\ln t)^3 \int dr\, d\hat{x}\, r\psi^{(2,1)}(r,r\hat{x}) t^{r(1+\hat{x}\delta)} \\ &= \frac{\pi^4}{24} \int d\hat{x}\, \frac{\psi^{(2,1)}(r,r\hat{x})/r}{(1+\hat{x}\delta)^3}\,,\end{aligned} \tag{4.29}$$

where in the second line we integrated by parts once, and in the third line we defined $m = r\hat{x}$.

4.3.4 Matrix model and volumes of five-cycles

For any gauge invariant operator X, we should have [94] $R[X] = \pi\,\mathrm{Vol}(\Sigma_X)/6\,\mathrm{Vol}(Y)$, where by Σ_X we denoted the 5-d submanifold of Y defined by the equation $X = 0$. Using $\mathrm{Vol}(Y) = (\pi^4/24)\int d\hat{x}\,\hat{\rho}(\hat{x})$, one can rewrite this equation as

$$\mathrm{Vol}(\Sigma_X) = \frac{\pi^3}{4} \int d\hat{x}\, \hat{\rho}(\hat{x}) R[X]\,. \tag{4.30}$$

For an operator X that is not gauge invariant, such as a bifundamental field that transforms in $(\mathbf{N}_a, \overline{\mathbf{N}}_b)$, $R[X]$ is not invariant under baryonic symmetries (4.7), but $R[X]+\hat{y}_a(\hat{x})-\hat{y}_b(\hat{x})$ is. So we suspect that

$$\mathrm{Vol}(\Sigma_X) = \frac{\pi^3}{4} \int d\hat{x}\, \hat{\rho}(\hat{x}) (R[X] + \hat{y}_a(\hat{x}) - \hat{y}_b(\hat{x}))\,. \tag{4.31}$$

We can think of this relation as a conjecture and prove the following result: If X is a chiral operator transforming in $(\mathbf{N}_a, \overline{\mathbf{N}}_b)$, then for δ in a neighborhood of zero, let

$$\mathrm{Vol}_m(\Sigma_X,\delta) = \frac{\pi^3}{4} \int d\hat{x}_\delta\, \hat{\rho}_\delta(\hat{x}_\delta) (R[X] + \hat{y}_{a,\delta}(\hat{x}_\delta) - \hat{y}_{b,\delta}(\hat{x}_\delta))\,. \tag{4.32}$$

The following two statements are equivalent:

1. The conjecture (3.5) holds for the given R-charges $R[X]$ and bare monopole charge Δ_m.

2. For any δ in a small enough neighborhood of zero, we have

$$\mathrm{Vol}_m(\Sigma_X, \delta) = \mathrm{Vol}(\Sigma_X, \delta)\,, \tag{4.33}$$

   where the volume $\mathrm{Vol}(\Sigma_X, \delta)$ is computed with the induced Sasakian metric on Y that corresponds to the matter R-charges $R[X]$ and the bare monopole charge $\Delta_m + \delta$.

The proof of this result is similar to that of the equivalence between (1) and (2) we discussed above, so we skip most of the details. Using (4.24), one can check that

$$\mathrm{Vol}_m(\Sigma_X, \delta) = \frac{\pi^3}{4} \int d\hat{x}\, \frac{\hat{\rho}(\hat{x})(R[X] + \hat{y}_a(\hat{x}) - \hat{y}_b(\hat{x}))}{(1 + \hat{x}\delta)^3}\,. \tag{4.34}$$

Defining $f_X(t, u)$ to be the Hilbert-Poincaré series for the ring of chiral operators obtained from the chiral ring by setting $X = 0$, and using the Martelli, Sparks, and Yau result [77]

$$\mathrm{Vol}(\Sigma_X, \delta) = \frac{\pi^3}{8} \lim_{t\to 1} (1-t)^3 f_X(t, t^\delta)\,, \tag{4.35}$$

one can show as in section 4.3.3 that

$$\mathrm{Vol}(\Sigma_X, \delta) = \frac{\pi^3}{4} \int d\hat{x} \frac{\lim_{r\to\infty} \psi_X^{(1,1)}(r, r\hat{x})/r}{(1 + \hat{x}\delta)^3}\,. \tag{4.36}$$

Here, $\psi_X(r, m)$ denotes the number of chiral ring operators with $X = 0$, R-charge at most r, and monopole charge at most m, and can be approximated by a smooth function of homogeneous degree three. By an argument analogous to the one in section 4.3.1 it follows that the statements (1) and (2) are equivalent.

4.3.5 A Consistency condition

Note that gauge invariant operators in the quiver that have no monopole charge are constructed from closed paths of bifundamental fields $\mathcal{O}_\alpha$. A consequence of our conjecture

(4.3b) is then that for a gauge invariant operator $X = \prod_\alpha \mathcal{O}_\alpha$ with no monopole charge the following sum vanishes:

$$\sum_\alpha \left[\frac{\partial^2 \psi_{\mathcal{O}_\alpha}}{\partial r \partial m} - R[\mathcal{O}_\alpha] \frac{\partial^3 \psi}{\partial r^2 \partial m} \right] = 0 \ . \tag{4.37}$$

Of course $\sum_\alpha R[\mathcal{O}_\alpha] = R[X]$, and we can simplify this expression:

$$\sum_\alpha \frac{\partial^2 \psi_{\mathcal{O}_\alpha}}{\partial r \partial m} = R[X] \frac{\partial^3 \psi}{\partial r^2 \partial m} \ . \tag{4.38}$$

We would like to show why (4.38) must hold from geometric considerations alone.[5]

The number of gauge invariant operators of fixed R-charge r and monopole charge m that do not contain the operator X is approximately

$$\psi_X^{(1,1)}(r,m) \approx \psi^{(1,1)}(r,m) - \psi^{(1,1)}(r - R[x], m) \approx R[X] \psi^{(2,1)}(r,m) \tag{4.39}$$

when $r \gg R[X]$ is large.

We can also use (4.35) to express the operator counts in terms of volumes. For coordinates x and y on a compact space, the volume of the set of points $xy = 0$ is the union of the set of points where $x = 0$ with the set of points where $y = 0$. That the volumes are additive implies the identity

$$\mathrm{Vol}(\Sigma_X, \delta) = \sum_\alpha \mathrm{Vol}(\Sigma_{\mathcal{O}_\alpha}, \delta) \ . \tag{4.40}$$

From this identity at large r and the results in the earlier part of this section, we have

$$\psi_X^{(1,1)}(r,m) = \sum_\alpha \psi_{\mathcal{O}_\alpha}^{(1,1)}(r,m) \ . \tag{4.41}$$

Combining (4.39) with (4.41) yields (4.38).

[5]For a Calabi-Yau four fold, at large r and m we approximate $\psi(r,m)$ and $\psi_{\mathcal{O}_\alpha}(r,m)$ by homogeneous polynomials of degree four and three respectively.

4.4 Theories with non-chiral bifundamental fields

4.4.1 $\mathcal{N} = 2$ deformations of the necklace quivers and matrix model

Our first field theory example consists of deformations of the necklace quiver gauge theories whose (undeformed) matrix models were also studied in [27] and in chapter 3. In $\mathcal{N} = 2$ notation, the field content of the necklace quiver theories consists of d vector multiplets with Chern-Simons kinetic terms and coefficients k_a, and chiral multiplets A_a and B_a connecting the gauge groups into a necklace (see figure 3.2). The superpotential

$$W = \sum_{a=1}^{d} \frac{1}{k_a} \operatorname{tr}(B_{a+1}A_{a+1} - A_a B_a)^2 \tag{4.42}$$

preserves $\mathcal{N} = 3$ supersymmetry. For any given k_a satisfying $\sum_{a=1}^{d} k_a = 0$, the field theory is dual to $AdS_4 \times Y$ where Y is a tri-Sasakian space, which is by definition the base of a hyperkähler cone [66].

While $\mathcal{N} = 3$ SUSY restricts the R-charges of A_a and B_a to be $1/2$, in this section we examine what happens if we make more general R-charge assignments for the A_a and B_a fields that break $\mathcal{N} = 3$ down to $\mathcal{N} = 2$. These R-charge assignments are required to preserve the marginality of the superpotential (4.42). This condition implies that for generic values of the CS levels k_a, namely if there are no cancellations between the various terms in (4.42), we must have $R[B_a] = 1 - R[A_a]$. The matrix model free energy functional is in this case

$$\begin{aligned} \tilde{F}[\rho, y_a, \mu] = {} & 2\pi N^{3/2} \int dx\, \rho x \sum_{a=1}^{d} q_a \delta y_a + 2\pi \Delta_m N^{3/2} \int dx\, \rho x \\ & - \pi N^{3/2} \int dx\, \rho^2 \sum_{a=1}^{d} (\delta y_a - R[A_a])\,(\delta y_a + R[B_a]) - 2\pi N^{3/2} \mu \left(\int dx\, \rho - 1 \right), \end{aligned} \tag{4.43}$$

where $\delta y_a = y_{a-1} - y_a$, $k_a = q_{a+1} - q_a$. As per the discussion after eq. (4.6), the equations of motion for δy_a following from (4.43) hold only when $-R[B_a] < \delta y_a < R[A_a]$. It is possible to have $\delta y_a = R[A_a]$ or $\delta y_a = -R[B_a]$, but in that case we should not impose the equation of motion for that particular δy_a.

Using the rescalings (4.15), one can write the solution of the equations of motion following from (4.43) as

$$\hat{\rho}(\hat{x}) = s_L(\hat{x}) - s_S(\hat{x})\,, \quad \delta\hat{y}_a(\hat{x}) = \frac{R[A_a] - R[B_a]}{2} + \frac{1}{2}\frac{|s_L(\hat{x}) + q_a\hat{x}| - |s_S(\hat{x}) + q_a\hat{x}|}{s_L(\hat{x}) - s_S(\hat{x})}\,, \tag{4.44}$$

where $s_L(\hat{x}) \geq s_S(\hat{x})$ are the two solutions of the equation

$$s(\hat{x})c_1 + \hat{x}c_2 + \sum_{a=1}^{d} |s(\hat{x}) + \hat{x}q_a| = 2\,, \tag{4.45}$$

with

$$c_1 \equiv \sum_{a=1}^{d} \left(R[A_a] - R[B_a]\right)\,, \qquad c_2 \equiv 2\Delta_m + \sum_{a=1}^{d} q_a \left(R[A_a] - R[B_a]\right)\,. \tag{4.46}$$

The constraint imposed by varying $\tilde{F}$ with respect to μ is $\int d\hat{x}\,\hat{\rho}(\hat{x}) = 1/\mu^2$.

We have encountered a solution of this type in chapter 3 in the case where $R[A_a] = R[B_a] = 1/2$ and $\Delta_m = 0$. As in chapter 3, one can think of eq. (4.45) as defining the boundary of a polygon

$$\mathcal{P} = \left\{ (\hat{x}, s) \in \mathbb{R}^2 : sc_1 + \hat{x}c_2 + \sum_{a=1}^{d} |s + \hat{x}q_a| \leq 2 \right\}\,. \tag{4.47}$$

The quantity $\hat{\rho}(\hat{x}) = s_L(\hat{x}) - s_S(\hat{x})$ can then be interpreted as the thickness of a constant $\hat{x}$ slice $\mathcal{P}_{\hat{x}}$ through this polygon, $\hat{\rho}(\hat{x}) = \text{Length}(\mathcal{P}_{\hat{x}})$. Consequently, $\int d\hat{x}\,\hat{\rho}(\hat{x}) = \text{Area}(\mathcal{P})$ and

$$\text{Vol}_m(Y) = \frac{\pi^4}{24\mu^2} = \frac{\pi^4}{24}\,\text{Area}(\mathcal{P})\,. \tag{4.48}$$

See appendix A.3 for a proof that the $\mathcal{N} = 3$ R-charge assignments minimize $\text{Vol}_m(Y)$ or, equivalently, maximize F. Just like $\hat{\rho}(\hat{x})$, the quantities $\hat{\rho}(\hat{x})\delta\hat{y}_a(\hat{x})$ can also be given

geometrical interpretations:

$$\begin{aligned}\hat{\rho}(\hat{x})(\delta\hat{y}_a(\hat{x}) + R[B_a]) &= \text{Length}\left(\mathcal{P}_{\hat{x}} \cap \{s + q_a\hat{x} \geq 0\}\right) , \\ \hat{\rho}(\hat{x})(-\delta\hat{y}_a(\hat{x}) + R[A_a]) &= \text{Length}\left(\mathcal{P}_{\hat{x}} \cap \{s + q_a\hat{x} \leq 0\}\right) .\end{aligned} \tag{4.49}$$

The equations above were written in a way that makes manifest the invariance under the flat directions exhibited in (4.7). Indeed, while in writing the free energy functional (4.43) we assumed $R[A_a]$ and Δ_m to be independent, we see that the eigenvalue density $\hat{\rho}(\hat{x})$ and the quantities appearing on the LHS of (4.49) depend non-trivially only on the linear combinations c_1 and c_2 that were defined in (4.46). These are the only linear combinations of $R[A_a]$ and Δ_m that are invariant under all symmetries in (4.7). The reason why we were able to find two such linear combinations at all is that the spaces Y have generically two $U(1)$ isometries that commute with $U(1)_R$.

4.4.2 Operator counting for necklace quivers

We now relate the matrix model quantities $\hat{\rho}(\hat{x})$ and $\hat{\rho}(\hat{x})\delta\hat{y}_a$ from the previous section to numbers of operators in the chiral ring of the gauge theory when $N = 1$. In the previous chapter we provided such a relation in the case $R[A_a] = R[B_a] = 1/2$ and $\Delta_m = 0$, and the argument holds, with minor modifications, for the more general R-charge assignments considered in this chapter. As explained in chapter 3, gauge invariant operators can be constructed out of the bifundamental fields A_a and B_a and the diagonal monopole operators $T^{(m)}$, and they are

$$\begin{aligned}\mathcal{O}(m, s, i, j) &= T^{(m)} C_1^{mq_1+s} C_2^{mq_2+s} \cdots C_d^{mq_d+s} (A_1 B_1)^i (A_2 B_2)^j , \\ C_a^{mq_a+s} &\equiv \begin{cases} A_a^{mq_a+s} & \text{if } mq_a + s > 0 \\ B_a^{-mq_a-s} & \text{if } mq_a + s < 0 \end{cases} .\end{aligned} \tag{4.50}$$

The labels m and s run over all integers, while i and j should be nonnegative integers.

Let $\psi(r, m)$ ($\psi_0(r, m)$) be the number of operators $\mathcal{O}(m, s, i, j)$ ($\mathcal{O}(m, s, 0, 0)$) with R-charge at most r and monopole charge at most m. In chapter 3, we showed that at large r and m we have $\psi^{(2,0)}(r, m) \approx \psi_0(r, m)$. This relation holds for the more general R-charge

assignments too because the only assumption needed to prove it was $R[A_1B_1] = R[A_2B_2] = 1$, which we still assume. A simple computation yields

$$R[\mathcal{O}(m,s,0,0)] = m\Delta_m + \sum_k R[C_k]\,|mq_a + s| = \frac{1}{2}\left[sc_1 + mc_2 + \sum_{a=1}^{d}|s + mq_a|\right], \quad (4.51)$$

where c_1 and c_2 are as defined in (4.46). Using this formula one can check, as in chapter 3, that $\psi^{(2,1)}(r, r\hat{x})/r \approx \psi_0^{(0,1)}(r, r\hat{x})/r$ is indeed given by the length of the slice $\mathcal{P}_{\hat{x}}$ through $\mathcal{P}$. We have therefore verified explicitly eq. (4.3a) for the necklace quivers at non-critical R-charges.

Let $\psi_{X_a}(r,m)$ be the number of chiral operators with R-charge at most r and monopole charge at most m that are nonzero when $X_a = 0$. As in chapter 3, we have that $\psi_{X_a}^{(1,0)}(r,m)$ equals the number of operators of the form $\mathcal{O}(m,s,0,0)$ with R-charge at most r and monopole charge at most m with the extra constraint that $mq_a + s \leq 0$ if $X_a = A_a$ and $mq_a + s \geq 0$ if $X_a = B_a$. As argued in chapter 3, these extra constraints imply that when r is large $\psi_{X_a}^{(1,1)}(r, r\hat{x})/r$ is given by the length of the intersection between the slice $\mathcal{P}_{\hat{x}}$ and the half-plane $s + q_a\hat{x} \geq 0$ if $X_a = B_a$ or $s + q_a\hat{x} \leq 0$ if $X_a = A_a$. Comparing with eq. (4.49) we see that the necklace quivers at arbitrary R-charges also obey our second conjecture (3.5).

4.4.3 Flavored necklace quivers

The discussion in the previous two subsections can be generalized by including flavor fields that interact with the existing matter fields through the superpotential

$$\delta W \sim \sum_{a=1}^{d} \mathrm{tr}\left[\sum_{j=1}^{n_a} \tilde{q}_j^{(a)} A_a q_j^{(a)} + \sum_{j=1}^{m_a} Q_j^{(a)} B_a \tilde{Q}_j^{(a)}\right]. \quad (4.52)$$

Given that the A_a transform in $(\overline{\mathbf{N}}_{a-1}, \mathbf{N}_a)$ and the B_a transform in the conjugate representation $(\mathbf{N}_{a-1}, \overline{\mathbf{N}}_a)$, for eq. (4.52) to make sense we must take $q_j^{(a)}$, $\tilde{q}_j^{(a)}$, $Q_j^{(a)}$, and $\tilde{Q}_j^{(a)}$ to transform in $\mathbf{N}_{a-1}$, $\overline{\mathbf{N}}_a$, $\overline{\mathbf{N}}_{a-1}$, and $\mathbf{N}_a$, respectively.

We discussed a superpotential of this form at the end of section 4.2, where we found that the effect of including the flavor fields was that the CS levels k_a and Δ_m of the unflavored model were replaced by $(\mathrm{sgn}\,x)g_a[T^{(\mathrm{sgn}\,x)}]$ and $(\mathrm{sgn}\,x)R[T^{(\mathrm{sgn}\,x)}]$, respectively. Eqs. (4.11)

and (4.12) applied to our flavored necklace quivers give

$$\begin{aligned} k_a \to (\operatorname{sgn} x) g_a[T^{(\operatorname{sgn} x)}] &= k_a + \frac{\operatorname{sgn} x}{2} \left(n_a - m_a - n_{a+1} + m_{a+1}\right) , \\ \Delta_m \to (\operatorname{sgn} x) R[T^{(\operatorname{sgn} x)}] &= \Delta_m + \frac{\operatorname{sgn} x}{2} \sum_a \left(n_a R[A_a] + m_a R[B_a]\right) . \end{aligned} \tag{4.53}$$

From $k_a = q_{a+1} - q_a$ we further have

$$q_a \to q_a - \frac{\operatorname{sgn} x}{2} \left(n_a - m_a\right) . \tag{4.54}$$

We believe that all the formulas presented in the previous two subsections continue to hold for the flavored theory if one makes the above three replacements. In particular, the relation between the matrix model quantities and operator counting we conjectured in eq. (4.3) continues to hold, and the volume of the 7-d space Y is still proportional to the area of a polygon $\mathcal{P}$ of the type (4.47).

4.4.4 Flavored $\mathcal{N} = 8$ theory and its matrix model

We broaden our scope of examples and verify (4.3) for maximally supersymmetric Yang-Mills theory to which we add flavor. The theory has one gauge group and three adjoint fields X_i, $i = 1, 2, 3$ coupled to $n_1 + n_2 + n_3$ pairs of fundamental fields through the superpotential

$$W \sim \operatorname{tr} \left[X_1 [X_2, X_3] + \sum_{j=1}^{n_1} q_j^{(1)} X_1 \tilde{q}_j^{(1)} + \sum_{j=1}^{n_2} q_j^{(2)} X_2 \tilde{q}_j^{(2)} + \sum_{j=1}^{n_3} q_j^{(3)} X_3 \tilde{q}_j^{(3)} \right] . \tag{4.55}$$

The corresponding matrix model was solved in [21] in the large N limit. We review their solution for $\rho(x)$. In the next subsection, we will compare $\rho(x)$ with the distribution of operators in the chiral ring and show that (4.3a) holds. In this case, eq. (4.10) takes the form $T^{(1)} T^{(-1)} = X_1^{n_1} X_2^{n_2} X_3^{n_3}$ [32,33]. To keep the notation concise, we define $\Delta_i \equiv R[X_i]$, $\Delta \equiv R[T^{(1)}]$, and $\tilde{\Delta} \equiv R[T^{(-1)}]$. The matrix model free energy functional is then

$$\tilde{F}[\rho] = \pi N^{3/2} \left[\int dx \rho \left(\Delta_1 \Delta_2 \Delta_3 \rho + (\Delta + \tilde{\Delta}) |x| + (\Delta - \tilde{\Delta}) x \right) - 2\mu \left(\int dx \, \rho - 1 \right) \right] . \tag{4.56}$$

As before, we define the hatted quantities (4.15). The eigenvalue density $\hat{\rho}(\hat{x})$ is

$$\hat{\rho}(\hat{x}) = \begin{cases} \frac{1-\hat{x}\Delta}{\Delta_1\Delta_2\Delta_3} & \text{if } 0 < \hat{x} < \frac{1}{\Delta}\,, \\ \frac{1+\hat{x}\tilde{\Delta}}{\Delta_1\Delta_2\Delta_3} & \text{if } -\frac{1}{\tilde{\Delta}} < \hat{x} < 0\,, \\ 0 & \text{otherwise}\,, \end{cases} \tag{4.57}$$

which agrees with (4.8) of [21].

4.4.5 Operator counting in flavored $\mathcal{N} = 8$ theory

The gauge-invariant operators built out of diagonal monopole operators and adjoint fields in this theory are $\mathrm{tr}[T^{(m)}X_1^{a_1}X_2^{a_2}X_3^{a_3}]$. The R-charges of these operators are

$$R[T^{(m)}X_1^{a_1}X_2^{a_2}X_3^{a_3}] = \begin{cases} m\Delta + \sum_{i=1}^3 a_i\Delta_i & m \geq 0\,, \\ -m\tilde{\Delta} + \sum_{i=1}^3 a_i\Delta_i & m < 0\,. \end{cases} \tag{4.58}$$

Let $\psi(r, m)$ be the number of operators with R-charge smaller than r and monopole charge smaller than m. To match with $\rho(x)$, we want to calculate $\partial^3\psi/\partial r^2\partial m$ at large r. It is easiest to start by calculating the derivative $\partial\psi/\partial m$ which equals the number of operators with R-charge smaller than r and monopole charge equal to m. For $m > 0$, at large r the number of operators $\mathrm{tr}[T^{(m)}X_1^{a_1}X_2^{a_2}X_3^{a_3}]$ is approximately equal to the volume of a tetrahedron with sides of length $(r - m\Delta)/\Delta_i$; similarly, for $m < 0$, the number of operators is equal to the volume of a tetrahedron with sides of length $(r + m\tilde{\Delta})/\Delta_i$. We thus have

$$\frac{\partial\psi}{\partial m} = \begin{cases} \frac{(r-m\Delta)^3}{6\Delta_1\Delta_2\Delta_3} & \text{if } 0 < m < \frac{r}{\Delta}\,, \\ \frac{(r+m\tilde{\Delta})^3}{6\Delta_1\Delta_2\Delta_3} & \text{if } -\frac{r}{\tilde{\Delta}} < m < 0\,, \\ 0 & \text{otherwise}\,. \end{cases} \tag{4.59}$$

Taking two derivatives with respect to r, we find agreement with (4.57) and confirmation of the conjecture (4.3a).

4.4.6 Other examples

We presented flavored $\mathcal{N} = 8$ in the main text because of its simplicity. One disadvantage of this example is that it possesses a single $U(N)$ factor and so we could not compute a δy and check (4.3b). To remedy this problem, in appendix B we consider two more complicated examples. The first of these is ABJM Chern-Simons theory (a theory with two gauge groups) [28] to which we add flavor. The second example has four gauge groups (see figure B.1). When a four-dimensional gauge theory has the field content of this second example, the Abelian moduli space is a $\mathbb{Z}_2 \times \mathbb{Z}_2$ orbifold of $\mathbb{C}^3$. Thus, with some abuse of notation, we refer to this second example as the $\mathbb{Z}_2 \times \mathbb{Z}_2$ orbifold theory.

The verification of (4.3) requires on the one hand calculating $\rho(x)$ and $\delta y(x)$ using the large N limit of the matrix model (4.6) and on the other counting operators in the chiral ring. We have two methods at our disposal for this counting. One may count the operators directly as we did above. Because the moduli space is toric for these last three examples, the direct approach has some generic features which we review in appendix B.3. In section 4.3, we presented an indirect counting method that involved calculating $\mathrm{Vol}(Y, \delta)$ (4.29) and $\mathrm{Vol}(\Sigma_X, \delta)$ (4.36) as a function of $\rho(x)$ and $\delta y(x)$.

4.5 Theories with chiral bifundamental fields

4.5.1 Noncancellation of long-range forces

As noted in [21], the functional (4.6) does not appear to describe the large N limit of gauge theories with chiral bifundamental fields. To derive (4.6), it was assumed that the long-range forces on the eigenvalues cancel. But for theories with chiral bifundamentals, there is no such cancellation.

The long-range forces at issue come from the interactions between the eigenvalues, both within a vector multiplet and between vector multiplets connected by a bifundamental field

X_{ab} [21]:

$$
\begin{aligned}
F^{(a)}_{i,\text{self}} &= \sum_{j\neq i} \coth \pi(\lambda^{(a)}_i - \lambda^{(a)}_j) \ , \\
F^{(a,b)}_{i,\text{inter}} &= \sum_j \left[\frac{R[X_{ab}] - 1 - i(\lambda^{(b)}_i - \lambda^{(a)}_j)}{2}\right] \coth \pi \left(\lambda^{(b)}_i - \lambda^{(a)}_j - i(1 - R[X_{ab}])\right) \ , \\
F^{(b,a)}_{i,\text{inter}} &= \sum_j \left[\frac{R[X_{ba}] - 1 + i(\lambda^{(b)}_i - \lambda^{(a)}_j)}{2}\right] \coth \pi \left(\lambda^{(b)}_i - \lambda^{(a)}_j + i(1 - R[X_{ba}])\right) \ .
\end{aligned} \tag{4.60}
$$

If $\left|\lambda^{(a)}_i - \lambda^{(b)}_j\right| \gg 1$, then we may approximate $\coth x \approx \operatorname{sgn} \Re x$. The long-range forces are the forces (4.60) with coth replaced by $\operatorname{sgn} \Re$. For theories with non-chiral bifundamentals and equal ranks, the long-range forces cancel out when $\Re\lambda^{(a)}_i = \Re\lambda^{(b)}_i$ for all i, a, b and (4.4) is satisfied. In general, the long-range forces on $\lambda^{(a)}_i$ cancel out only when

$$
\sum_b (R[X_{ab}] - 1 + y_{b,j}) + \sum_b (R[X_{ba}] - 1 - y_{b,j}) = -2 \ . \tag{4.61}
$$

Thus the free energy functional (4.6) is correct for theories with chiral bifundamentals only if the $y_a(x)$ satisfy some constraints.

4.5.2 Operator counting for the $\mathbb{C}^3/\mathbb{Z}_3$ theory

To investigate what the matrix model for a chiral theory should give in the large N limit, we study the $U(N)^3$ Chern-Simons theory described by the quiver in figure 4.1. Let the Chern-Simons coefficients be (k_1, k_2, k_3) such that $k_1 + k_2 + k_3 = 0$. We will assume $k_1 > 0$, $k_2 < 0$, $k_3 < 0$. The moduli space is a Kähler quotient of $\mathbb{C}^5$ with weights $(\frac{1}{3}(k_+ + k_-), \frac{1}{3}(k_+ + k_-), \frac{1}{3}(k_+ + k_-), -k_+, -k_-)$, where we define $k_- = k_1 - k_2$ and $k_+ = k_1 - k_3$.

There is a superpotential of the form

$$
W \sim \operatorname{tr}\left[\epsilon_{ijk} A_{31,k} A_{23,j} A_{12,i}\right] \ , \tag{4.62}
$$

and a monopole relation $T^{(1)}T^{(-1)} = 1$. We let $R[A_{ij,1}] = \Delta_x$, $R[A_{ij,2}] = \Delta_y$, $R[A_{ij,3}] = \Delta_z$, with $\Delta_x + \Delta_y + \Delta_z = 2$ as any other choice of R-charges may be transformed into this choice by a transformation of the form (4.7). We denote $R[T^{(1)}] = -R[T^{(-1)}] = \Delta$.

The gauge invariant operators have the form

$$T^{(m)} \prod_{i=1}^{3} \prod_{j=1}^{3} (A_{i(i+1),j})^{n_{i(i+1),j}} \,. \tag{4.63}$$

To be gauge invariant, for $m \geq 0$ we must impose $\sum_j n_{12,j} = mk_1 + s$, $\sum_j n_{23,j} = mk_1 + mk_2 + s$ and $\sum_j n_{31,j} = s$ and for $m < 0$ we must impose $\sum_j n_{23,j} = mk_2 + s$, $\sum_j n_{31,j} = mk_2 + mk_3 + s$, and $\sum_j n_{12,j} = s$. Given the R-charge assignments, it is convenient to introduce $n_j = \sum_i n_{i(i+1),j}$. Each gauge invariant operator corresponds to a quadruple (n_1, n_2, n_3, m) such that $\sum_j n_j = mk_{\mathrm{sgn}(m)} + 3s$ and m is bounded between $-\sum_j n_j/k_-$ and $\sum_j n_j/k_+$.

Given the description of the gauge invariant operators, it is now a straightforward task to count them by either the direct method described in appendix B.3 or the indirect method described in section 4.3. For $\Delta_x \geq \Delta_y \geq \Delta_z$ a piecewise expression for $\hat{\rho}(\hat{x})$ is:

$$\hat{\rho}(\hat{x}) = \begin{cases} 0\,, & \hat{x} \leq -\frac{1}{k_-\Delta_z - \Delta}\,, \\ \frac{1+(k_-\Delta_z-\Delta)\hat{x}}{3\Delta_z(\Delta_x-\Delta_z)(\Delta_y-\Delta_z)}\,, & -\frac{1}{k_-\Delta_z-\Delta} \leq \hat{x} \leq -\frac{1}{k_-\Delta_y-\Delta}\,, \\ \frac{(\Delta_x-\Delta_y-\Delta_z)(1-\Delta\hat{x})-\Delta_y\Delta_z k_-\hat{x}}{3(\Delta_x-\Delta_y)(\Delta_x-\Delta_z)\Delta_y\Delta_z}\,, & -\frac{1}{k_-\Delta_y-\Delta} \leq \hat{x} \leq -\frac{1}{k_-\Delta_x-\Delta}\,, \\ \frac{1-\Delta\hat{x}}{3\Delta_x\Delta_y\Delta_z}\,, & -\frac{1}{k_-\Delta_x-\Delta} \leq \hat{x} \leq \frac{1}{k_+\Delta_x+\Delta}\,, \\ \frac{(\Delta_x-\Delta_y-\Delta_z)(1-\Delta\hat{x})+\Delta_y\Delta_z k_+\hat{x}}{3(\Delta_x-\Delta_y)(\Delta_x-\Delta_z)\Delta_y\Delta_z}\,, & \frac{1}{k_+\Delta_x+\Delta} \leq \hat{x} \leq \frac{1}{k_+\Delta_y+\Delta}\,, \\ \frac{1-(k_+\Delta_z+\Delta)\hat{x}}{3\Delta_z(\Delta_x-\Delta_z)(\Delta_y-\Delta_z)}\,, & \frac{1}{k_+\Delta_y+\Delta} \leq \hat{x} \leq \frac{1}{k_+\Delta_z+\Delta}\,, \\ 0\,, & \frac{1}{k_+\Delta_z+\Delta} \leq \hat{x}\,. \end{cases} \tag{4.64}$$

We note three odd things about (4.64): 1) If $\Delta_x = \Delta_y = \Delta_z$, $\hat{\rho}$ has a delta function at $-\frac{1}{k_-\Delta_x-\Delta}$ and $\frac{1}{k_+\Delta_x+\Delta}$. 2) In contrast to nonchiral examples, $\hat{\rho}(\hat{x})$ while still piecewise linear is no longer a convex function of $\hat{x}$. 3) The matrix model (4.6) gives the same result for $\hat{\rho}$ in the central region despite the fact that the long range forces do not cancel. (In other regions and for $\delta\hat{y}_{ab}$, the matrix model results are different.)

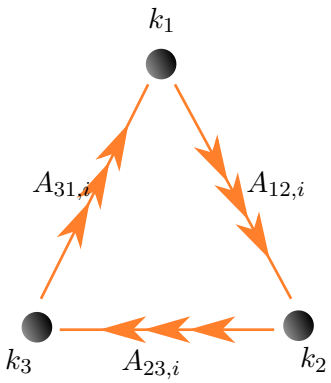


Figure 4.1: The quiver for the $\mathbb{C}^3/\mathbb{Z}_3$ theory. When the CS levels are $(2k,-k,-k)$ this field theory is believed to be dual to $AdS_4 \times M^{1,1,1}/\mathbb{Z}_k$.

Now we set $A_{23,1}$ to zero. The nonzero operators are those with no $A_{i(i+1),1}$ fields. As a piecewise function:

$$\hat{y}_3(\hat{x}) - \hat{y}_2(\hat{x}) = \begin{cases} -\Delta_z\,, & -\frac{1}{k_-\Delta_z-\Delta} \le \hat{x} \le -\frac{1}{k_-\Delta_y-\Delta}\,, \\ \frac{\Delta_y\Delta_z(1+(k_-\Delta_x-\Delta)\hat{x})}{(\Delta_x-\Delta_y-\Delta_z)(1-\Delta\hat{x})-\Delta_y\Delta_z k_-\hat{x}}\,, & -\frac{1}{k_-\Delta_y-\Delta} \le \hat{x} \le -\frac{1}{k_-\Delta_x-\Delta}\,, \\ 0\,, & -\frac{1}{k_-\Delta_x-\Delta} \le \hat{x} \le \frac{1}{k_+\Delta_x+\Delta}\,, \\ \frac{\Delta_y\Delta_z(1-(k_+\Delta_x+\Delta)\hat{x})}{(\Delta_x-\Delta_y-\Delta_z)(1-\Delta\hat{x})+\Delta_y\Delta_z k_+\hat{x}}\,, & \frac{1}{k_+\Delta_x+\Delta} \le \hat{x} \le \frac{1}{k_+\Delta_y+\Delta}\,, \\ -\Delta_z\,, & \frac{1}{k_+\Delta_y+\Delta} \le \hat{x} \le \frac{1}{k_+\Delta_z+\Delta}\,. \end{cases} \tag{4.65}$$

Finally, we set $A_{31,1}$ to zero. The nonzero operators are those with no $A_{i(i+1),1}$'s, and those with $m \ge 0, n_x + n_y + n_z = k_+ m$. As a piecewise function:

$$\hat{y}_1(\hat{x}) - \hat{y}_3(\hat{x}) = \begin{cases} -\Delta_z\,, & -\frac{1}{k_-\Delta_z-\Delta} \le \hat{x} \le -\frac{1}{k_-\Delta_y-\Delta}\,, \\ \frac{\Delta_y\Delta_z(1+(k_-\Delta_x-\Delta)\hat{x})}{(\Delta_x-\Delta_y-\Delta_z)(1-\Delta\hat{x})-\Delta_y\Delta_z k_-\hat{x}}\,, & -\frac{1}{k_-\Delta_y-\Delta} \le \hat{x} \le -\frac{1}{k_-\Delta_x-\Delta}\,, \\ 0\,, & -\frac{1}{k_-\Delta_x-\Delta} \le \hat{x} \le \frac{1}{k_+\Delta_x+\Delta}\,, \\ -2\frac{\Delta_y\Delta_z(1-(k_+\Delta_x+\Delta)\hat{x})}{(\Delta_x-\Delta_y-\Delta_z)(1-\Delta\hat{x})+\Delta_y\Delta_z k_+\hat{x}}\,, & \frac{1}{k_+\Delta_x+\Delta} \le \hat{x} \le \frac{1}{k_+\Delta_y+\Delta}\,, \\ 2\Delta_z\,, & \frac{1}{k_+\Delta_y+\Delta} \le \hat{x} \le \frac{1}{k_+\Delta_z+\Delta}\,. \end{cases} \tag{4.66}$$

The result for $\hat{y}_2(\hat{x}) - \hat{y}_1(\hat{x})$ follows by taking the difference of (4.65) and (4.66). We have checked that the operator counts where we set each of the remaining seven bifundamental fields to zero in turn yield the same results for the differences in the $\hat{y}$'s.

4.5.3 Particular case: the cone over $M^{1,1,1}/\mathbb{Z}_k$

Consider the case where the internal space Y is $M^{1,1,1}/\mathbb{Z}_k$. It was proposed in [67, 95, 96] that the dual field theory is the one in figure 4.1 with CS levels $k_1 = 2k$ and $k_2 = k_3 = -k$, so $k_+ = k_- = 3k$. As a function of the trial R-charges, the volume of Y is

$$\mathrm{Vol}(Y) = \frac{\pi^4}{24}\int d\hat{x}\,\hat{\rho}(\hat{x}) = \frac{3k^3\pi^4\left(\Delta^2 + 9k^2(\Delta_x\Delta_y + \Delta_x\Delta_z + \Delta_y\Delta_z\right)}{8\left(9k^2\Delta_x^2 - \Delta^2\right)\left(9k^2\Delta_y^2 - \Delta^2\right)\left(9k^2\Delta_z^2 - \Delta^2\right)}\,. \tag{4.67}$$

Under the constraint $\Delta_x + \Delta_y + \Delta_z = 2$, this expression is maximized for $\Delta_x = \Delta_y = \Delta_z = 2/3$ and $\Delta = 0$, and the maximum is $9\pi^4/(128k)$, which is the volume of $M^{1,1,1}/\mathbb{Z}_k$ [94].

For the critical R-charges, our predicted eigenvalue density is

$$\begin{aligned}
\hat{\rho}(\hat{x}) &= \frac{9}{8}\theta\left(\frac{1}{2k} - |\hat{x}|\right) + \frac{9}{32k}\delta\left(\hat{x} + \frac{1}{2k}\right) + \frac{9}{32k}\delta\left(\hat{x} - \frac{1}{2k}\right)\,,\\
\hat{\rho}(\hat{x})\left(\hat{y}_3(\hat{x}) - \hat{y}_2(\hat{x})\right) &= -\frac{3}{16k}\delta\left(\hat{x} + \frac{1}{2k}\right) - \frac{3}{16k}\delta\left(\hat{x} - \frac{1}{2k}\right)\,,\\
\hat{\rho}(\hat{x})\left(\hat{y}_2(\hat{x}) - \hat{y}_1(\hat{x})\right) &= -\frac{3}{16k}\delta\left(\hat{x} + \frac{1}{2k}\right) + \frac{3}{8k}\delta\left(\hat{x} - \frac{1}{2k}\right)\,.
\end{aligned} \tag{4.68}$$

The volumes of the five-cycles corresponding to the bifundamental fields are

$$\begin{aligned}
\mathrm{Vol}(\Sigma_{A_{23,a}}) &= \frac{\pi^3}{4}\int dx\,\hat{\rho}(\hat{x})\left(\hat{y}_3(\hat{x}) - \hat{y}_2(\hat{x}) + \frac{2}{3}\right) = \frac{3\pi^3}{16k}\,,\\
\mathrm{Vol}(\Sigma_{A_{12,a}}) &= \frac{\pi^3}{4}\int dx\,\hat{\rho}(\hat{x})\left(\hat{y}_2(\hat{x}) - \hat{y}_1(\hat{x}) + \frac{2}{3}\right) = \frac{21\pi^3}{64k}\,,\\
\mathrm{Vol}(\Sigma_{A_{31,a}}) &= \frac{\pi^3}{4}\int dx\,\hat{\rho}(\hat{x})\left(\hat{y}_1(\hat{x}) - \hat{y}_3(\hat{x}) + \frac{2}{3}\right) = \frac{21\pi^3}{64k}\,.
\end{aligned} \tag{4.69}$$

Let us understand how these volumes are related to the volumes of the divisors computed in [94]. The cone over $M^{1,1,1}$ is a Kähler quotient of $\mathbb{C}^5$ by a $U(1)$ that acts with weights $(2, 2, 2, -3, -3)$ on the coordinates $(u_1, u_2, u_3, v_1, v_2)$ parameterizing $\mathbb{C}^5$. The $\mathbb{Z}_k$ orbifold used to produce the quiver in figure 4.1 acts by the identification $(v_1, v_2) \sim (v_1 e^{2\pi i/k}, v_2 e^{-2\pi i/k})$ leaving the u_i coordinates untouched. It is natural to identify $A_{23,a}$ with u_a, $A_{12,a}$ with $u_a v_1$, and $A_{31,a}$ with $u_a v_2$. Using the explicit metric on $M^{1,1,1}$ the authors of [94] calculated the volumes of the five-cycles corresponding to either $u_a = 0$ or

$v_b = 0$ in $M^{1,1,1}$ to be

$$\mathrm{Vol}(\Sigma_{u_a}) = \frac{3\pi^3}{16}\,, \qquad \mathrm{Vol}(\Sigma_{v_b}) = \frac{9\pi^3}{64}\,. \tag{4.70}$$

We see that these equations are consistent with (4.69): we have $k\,\mathrm{Vol}(\Sigma_{A_{23,a}}) = \mathrm{Vol}(\Sigma_{u_a})$ as well as $k\,\mathrm{Vol}(\Sigma_{A_{12,a}}) = \mathrm{Vol}(\Sigma_{u_a}) + \mathrm{Vol}(\Sigma_{v_1})$ and $k\,\mathrm{Vol}(\Sigma_{A_{31,a}}) = \mathrm{Vol}(\Sigma_{u_a}) + \mathrm{Vol}(\Sigma_{v_2})$. The factor of k in these formulas comes from the fact that the cycles whose volumes are given in (4.69) belong to a $\mathbb{Z}_k$ orbifold of $M^{1,1,1}$.

For those interested in another simple example of a theory with chiral bifundamental fields, we describe our predictions for a theory with the cone over $Q^{2,2,2}$ as its Abelian moduli space in appendix B.4.

4.5.4 Missing operators

There is a difference between the matrix model and operator counting that manifests itself in chiral theories. The matrix model depends explicitly on the bifundamental fields, and a δy saturates when it reaches minus the R-charge of a bifundamental field. In the absence of flavors, the saturation of the δy is responsible for all of the corners in ρ and $\rho\delta y$. In the $\mathbb{C}^3/\mathbb{Z}_3$ example, ρ has a corner at $\hat{x} = \frac{1}{k_+\Delta_x+\Delta}$. We might expect that there exists some bifundamental field $A_{ij,k}$ so that $\delta y + R[A_{ij,k}]$ becomes zero at $\hat{x} = \frac{1}{k_+\Delta_x+\Delta}$, or equivalently that $\psi^{(1,1)}_{A_{ij,k}}(r, r\hat{x})$ becomes zero at $\hat{x} = \frac{1}{k_+\Delta_x+\Delta}$. There is no such field. However, if we consider the density $\psi^{(1,1)}_{(A_{31,1},A_{31,2},A_{31,3})}$ of operators when we set $A_{31,1} = A_{31,2} = A_{31,3} = 0$, then this density does become zero at $\hat{x} = \frac{1}{k_+\Delta_x+\Delta}$. So it appears to be important to allow arbitrary sets of bifundamental fields to be set to zero. A more geometric way of saying this is that the important objects in the operator counting formula are not the bifundamental fields but rather five-cycles in the Sasaki-Einstein manifold. In the $\mathbb{C}^3/\mathbb{Z}_3$ theory, there seems to be no operator constructed from bifundamental fields that corresponds to a five-brane wrapping the cycle $A_{31,1} = A_{31,2} = A_{31,3} = 0$.[6] We might say that we are missing some operators. We note that the problem could be resolved if we added an operator $A_{31,1}/A_{23,1}$,

[6] Unlike setting $A_{23,1} = A_{23,2} = A_{23,3} = 0$ where there are no non-vanishing operators, when we set $A_{31,1} = A_{31,2} = A_{31,3} = 0$, the number of non-zero operators $\psi_{(A_{31,1},A_{31,2},A_{31,3})}$ scales as r^3, indicating the presence of a 5-cycle in the geometry.

since the cycle $A_{31,1} = 0$ is the sum of the cycles $A_{23,1} = 0$ and $A_{31,1} = A_{31,2} = A_{31,3} = 0$. The problem never arises in non-chiral non-flavored theories because these theories do have an operator for every cycle.

The flavored $\mathcal{N} = 8$ and flavored ABJM model also have missing operators. At $x = 0$, there is a corner in the solutions that does not correspond to any δy_{ab} saturating at the R-charge of some bifundamental field X. Instead, the corner comes from the q fields. From the operator counting perspective, this corner can be explained by the fact that $\psi_T^{(1,1)}$ becomes zero at $\hat{x} = 0$.

Bibliography

[1] D. R. Gulotta, C. P. Herzog, and S. S. Pufu, *From Necklace Quivers to the F-theorem, Operator Counting, and T(U(N))*, *JHEP* **1112** (2011) 077, [arXiv:1105.2817].

[2] D. R. Gulotta, C. P. Herzog, and S. S. Pufu, *Operator Counting and Eigenvalue Distributions for 3D Supersymmetric Gauge Theories*, *JHEP* **1111** (2011) 149, [arXiv:1106.5484].

[3] J. M. Maldacena, *The large N limit of superconformal field theories and supergravity*, *Adv. Theor. Math. Phys.* **2** (1998) 231–252, [hep-th/9711200].

[4] S. S. Gubser, I. R. Klebanov, and A. M. Polyakov, *Gauge theory correlators from non-critical string theory*, *Phys. Lett.* **B428** (1998) 105–114, [hep-th/9802109].

[5] E. Witten, *Anti-de Sitter space and holography*, *Adv. Theor. Math. Phys.* **2** (1998) 253–291, [hep-th/9802150].

[6] P. Kasteleyn, *Graph theory and crystal physics*, in *Graph theory and Theoretical Physics* (F. Harary, ed.), pp. 43–110. Academic Press, London, 1967.

[7] A. Hanany and K. D. Kennaway, *Dimer models and toric diagrams*, hep-th/0503149.

[8] S. Franco, A. Hanany, K. D. Kennaway, D. Vegh, and B. Wecht, *Brane dimers and quiver gauge theories*, *JHEP* **0601** (2006) 096, [hep-th/0504110].

[9] S. Franco, A. Hanany, D. Martelli, J. Sparks, D. Vegh, and B. Wecht, *Gauge theories from toric geometry and brane tilings*, *JHEP* **01** (2006) 128, [hep-th/0505211].

[10] A. Hanany and D. Vegh, *Quivers, tilings, branes and rhombi*, *JHEP* **10** (2007) 029, [hep-th/0511063].

[11] B. Feng, Y.-H. He, K. D. Kennaway, and C. Vafa, *Dimer models from mirror symmetry and quivering amoebae*, *Adv.Theor.Math.Phys.* **12** (2008) 489–545, [hep-th/0511287].

[12] J. Stienstra, *Hypergeometric Systems in two Variables, Quivers, Dimers and Dessins d'Enfants*, 0711.0464.

[13] S. Benvenuti, S. Franco, A. Hanany, D. Martelli, and J. Sparks, *An infinite family of superconformal quiver gauge theories with Sasaki-Einstein duals*, *JHEP* **06** (2005) 064, [hep-th/0411264].

[14] G. Mack, *All Unitary Ray Representations of the Conformal Group $SU(2,2)$ with Positive Energy*, *Commun. Math. Phys.* **55** (1977) 1–28.

[15] V. K. Dobrev and V. B. Petkova, *All positive energy unitary irreducible representations of extended conformal superysmmetry*, *Phys. Lett.* **162B** (1985) 127–132.

[16] S. Benvenuti, L. A. Pando Zayas, and Y. Tachikawa, *Triangle anomalies from Einstein manifolds*, *Adv. Theor. Math. Phys.* **10** (2006) 395–432, [hep-th/0601054].

[17] A. Kapustin, B. Willett, and I. Yaakov, *Exact Results for Wilson Loops in Superconformal Chern- Simons Theories with Matter*, *JHEP* **03** (2010) 089, [arXiv:0909.4559].

[18] V. Pestun, *Localization of gauge theory on a four-sphere and supersymmetric Wilson loops*, *Commun.Math.Phys.* **313** (2012) 71–129, [arXiv:0712.2824].

[19] J. L. Cardy, *Is There a c Theorem in Four-Dimensions?*, *Phys.Lett.* **B215** (1988) 749–752.

[20] Z. Komargodski and A. Schwimmer, *On Renormalization Group Flows in Four Dimensions*, *JHEP* **1112** (2011) 099, [arXiv:1107.3987].

[21] D. L. Jafferis, I. R. Klebanov, S. S. Pufu, and B. R. Safdi, *Towards the F-Theorem: N=2 Field Theories on the Three-Sphere*, *JHEP* **1106** (2011) 102, [arXiv:1103.1181].

[22] A. Amariti and M. Siani, *Z-extremization and F-theorem in Chern-Simons matter theories*, *JHEP* **1110** (2011) 016, [arXiv:1105.0933].

[23] H. Casini and M. Huerta, *On the RG running of the entanglement entropy of a circle*, *Phys.Rev.* **D85** (2012) 125016, [arXiv:1202.5650].

[24] K. A. Intriligator and B. Wecht, *The Exact superconformal R symmetry maximizes a*, *Nucl.Phys.* **B667** (2003) 183–200, [hep-th/0304128].

[25] D. L. Jafferis, *The Exact Superconformal R-Symmetry Extremizes Z*, *JHEP* **1205** (2012) 159, [arXiv:1012.3210].

[26] C. Closset, T. T. Dumitrescu, G. Festuccia, Z. Komargodski, and N. Seiberg, *Contact Terms, Unitarity, and F-Maximization in Three-Dimensional Superconformal Theories*, arXiv:1205.4142.

[27] C. P. Herzog, I. R. Klebanov, S. S. Pufu, and T. Tesileanu, *Multi-Matrix Models and Tri-Sasaki Einstein Spaces*, *Phys. Rev.* **D83** (2011) 046001, [arXiv:1011.5487].

[28] O. Aharony, O. Bergman, D. L. Jafferis, and J. Maldacena, $\mathcal{N} = 6$ *superconformal Chern-Simons-matter theories, M2-branes and their gravity duals*, *JHEP* **10** (2008) 091, [arXiv:0806.1218].

[29] K. Ueda and M. Yamazaki, *Toric Calabi-Yau four-folds dual to Chern-Simons-matter theories*, *JHEP* **0812** (2008) 045, [arXiv:0808.3768].

[30] Y. Imamura and K. Kimura, *Quiver Chern-Simons theories and crystals*, *JHEP* **0810** (2008) 114, [arXiv:0808.4155].

[31] M. Aganagic, *A Stringy Origin of M2 Brane Chern-Simons Theories*, *Nucl.Phys.* **B835** (2010) 1–28, [arXiv:0905.3415].

[32] F. Benini, C. Closset, and S. Cremonesi, *Chiral flavors and M2-branes at toric CY4 singularities*, *JHEP* **1002** (2010) 036, [arXiv:0911.4127].

[33] D. L. Jafferis, *Quantum corrections to $\mathcal{N} = 2$ Chern-Simons theories with flavor and their AdS4 duals*, arXiv:0911.4324.

[34] F. Benini, C. Closset, and S. Cremonesi, *Comments on 3D Seiberg-Like Dualities*, *JHEP* **10** (2011) 075, [arXiv:1108.5373].

[35] M. Marino and P. Putrov, *ABJM theory as a Fermi gas*, *J.Stat.Mech.* **1203** (2012) P03001, [arXiv:1110.4066].

[36] M. Marino and P. Putrov, *Interacting fermions and N=2 Chern-Simons-matter theories*, arXiv:1206.6346.

[37] D. R. Gulotta, J. Ang, and C. P. Herzog, *Matrix Models for Supersymmetric Chern-Simons Theories with an ADE Classification*, *JHEP* **1201** (2012) 132, [arXiv:1111.1744].

[38] D. R. Gulotta, C. P. Herzog, and T. Nishioka, *The ABCDEF's of Matrix Models for Supersymmetric Chern-Simons Theories*, *JHEP* **1204** (2012) 138, [arXiv:1201.6360].

[39] D. R. Gulotta, *Properly ordered dimers, R-charges, and an efficient inverse algorithm*, *JHEP* **0810** (2008) 014, [arXiv:0807.3012].

[40] M. R. Douglas, B. R. Greene, and D. R. Morrison, *Orbifold resolution by D-branes*, *Nucl. Phys.* **B506** (1997) 84–106, [hep-th/9704151].

[41] D. R. Morrison and M. R. Plesser, *Non-spherical horizons. I*, *Adv. Theor. Math. Phys.* **3** (1999) 1–81, [hep-th/9810201].

[42] B. Feng, A. Hanany, and Y.-H. He, *D-brane gauge theories from toric singularities and toric duality*, *Nucl. Phys.* **B595** (2001) 165–200, [hep-th/0003085].

[43] I. Garcia-Etxebarria, F. Saad, and A. M. Uranga, *Quiver gauge theories at resolved and deformed singularities using dimers*, *JHEP* **06** (2006) 055, [hep-th/0603108].

[44] W. Fulton, *Introduction to Toric Varieties*. Princeton University Press, 1993.

[45] A. Butti and A. Zaffaroni, *R-charges from toric diagrams and the equivalence of a-maximization and Z-minimization*, *JHEP* **11** (2005) 019, [hep-th/0506232].

[46] S. Lee and S.-J. Rey, *Comments on anomalies and charges of toric-quiver duals*, *JHEP* **03** (2006) 068, [hep-th/0601223].

[47] A. Kato, *Zonotopes and four-dimensional superconformal field theories*, *JHEP* **06** (2007) 037, [hep-th/0610266].

[48] D. Martelli, J. Sparks, and S.-T. Yau, *The Geometric dual of a-maximisation for Toric Sasaki-Einstein manifolds*, *Commun.Math.Phys.* **268** (2006) 39–65, [hep-th/0503183].

[49] W. Blaschke, *Über Affine Geometrie VII: Neue Extremeigenschaften von Ellipse und Ellipsoid*, *Ber. Verh. Sächs. Akad. Wiss. Leipzig Math. Phys. Kl.* **69** (1917) 306–318.

[50] C. M. Petty, *Affine isoperimetric problems*, *Ann New York Acad Sci* **440** (1985) 113–127.

[51] B. V. Karpov and N. D. Yu, *Three-block Exceptional Collections over del Pezzo Surfaces*, *Izv. Ross. Akad. Nauk Ser. Mat.* **62** (1998) 429–463, [alg-geom/9703027].

[52] S. S. Gubser, *Einstein manifolds and conformal field theories*, *Phys. Rev.* **D59** (1999) 025006, [hep-th/9807164].

[53] M. Henningson and K. Skenderis, *The holographic Weyl anomaly*, *JHEP* **07** (1998) 023, [hep-th/9806087].

[54] A. Bergman and C. P. Herzog, *The volume of some non-spherical horizons and the AdS/CFT correspondence*, *JHEP* **01** (2002) 030, [hep-th/0108020].

[55] B. Feng, S. Franco, A. Hanany, and Y.-H. He, *Symmetries of toric duality*, *JHEP* **12** (2002) 076, [hep-th/0205144].

[56] S. Lee, *Superconformal field theories from crystal lattices*, *Phys. Rev.* **D75** (2007) 101901, [hep-th/0610204].

[57] S. Franco, A. Hanany, D. Krefl, J. Park, A. M. Uranga, and D. Vegh, *Dimers and Orientifolds*, *JHEP* **09** (2007) 075, [arXiv:0707.0298].

[58] A. Ishii and K. Ueda, *Dimer models and the special McKay correspondence*, arXiv:0905.0059.

[59] N. Hama, K. Hosomichi, and S. Lee, *Notes on SUSY Gauge Theories on Three-Sphere*, *JHEP* **1103** (2011) 127, [arXiv:1012.3512].

[60] A. B. Zamolodchikov, *Irreversibility of the Flux of the Renormalization Group in a 2D Field Theory*, *JETP Lett.* **43** (1986) 730–732.

[61] R. Emparan, C. V. Johnson, and R. C. Myers, *Surface terms as counterterms in the AdS/CFT correspondence*, *Phys. Rev.* **D60** (1999) 104001, [hep-th/9903238].

[62] H. Casini, M. Huerta, and R. C. Myers, *Towards a derivation of holographic entanglement entropy*, *JHEP* **1105** (2011) 036, [arXiv:1102.0440].

[63] R. C. Myers and A. Sinha, *Holographic c-theorems in arbitrary dimensions*, *JHEP* **1101** (2011) 125, [arXiv:1011.5819].

[64] R. C. Myers and A. Sinha, *Seeing a c-theorem with holography*, *Phys.Rev.* **D82** (2010) 046006, [arXiv:1006.1263].

[65] I. R. Klebanov and A. A. Tseytlin, *Entropy of near extremal black p-branes*, *Nucl.Phys.* **B475** (1996) 164–178, [hep-th/9604089].

[66] D. L. Jafferis and A. Tomasiello, *A Simple class of N=3 gauge/gravity duals*, *JHEP* **0810** (2008) 101, [arXiv:0808.0864].

[67] S. Franco, I. R. Klebanov, and D. Rodriguez-Gomez, *M2-branes on Orbifolds of the Cone over $Q^{1,1,1}$*, *JHEP* **08** (2009) 033, [arXiv:0903.3231].

[68] M. Marino and P. Putrov, *Exact Results in ABJM Theory from Topological Strings*, *JHEP* **1006** (2010) 011, [arXiv:0912.3074].

[69] R. C. Santamaria, M. Marino, and P. Putrov, *Unquenched flavor and tropical geometry in strongly coupled Chern-Simons-matter theories*, *JHEP* **1110** (2011) 139, [arXiv:1011.6281].

[70] D. Martelli and J. Sparks, *The large N limit of quiver matrix models and Sasaki-Einstein manifolds*, *Phys.Rev.* **D84** (2011) 046008, [arXiv:1102.5289].

[71] S. Cheon, D. Gang, S. Kim, and J. Park, *Refined test of AdS4/CFT3 correspondence for N=2,3 theories*, *JHEP* **1105** (2011) 027, [arXiv:1102.4273].

[72] J. P. Gauntlett, G. W. Gibbons, G. Papadopoulos, and P. K. Townsend, *Hyper-Kaehler manifolds and multiply intersecting branes*, *Nucl. Phys.* **B500** (1997) 133–162, [hep-th/9702202].

[73] D. Gaiotto and E. Witten, *S-Duality of Boundary Conditions In N=4 Super Yang-Mills Theory*, *Adv.Theor.Math.Phys.* **13** (2009) 721–896, [arXiv:0807.3720].

[74] S. Benvenuti and S. Pasquetti, *3D-partition functions on the sphere: exact evaluation and mirror symmetry*, *JHEP* **1205** (2012) 099, [arXiv:1105.2551].

[75] T. Nishioka, Y. Tachikawa, and M. Yamazaki, *3d Partition Function as Overlap of Wavefunctions*, *JHEP* **1108** (2011) 003, [arXiv:1105.4390].

[76] H.-U. Yee, *AdS/CFT with Tri-Sasakian Manifolds*, *Nucl.Phys.* **B774** (2007) 232–255, [hep-th/0612002].

[77] D. Martelli, J. Sparks, and S.-T. Yau, *Sasaki-Einstein manifolds and volume minimisation*, *Commun.Math.Phys.* **280** (2008) 611–673, [hep-th/0603021].

[78] R. Bielawski and A. S. Dancer, *The Geometry and topology of toric hyperkähler manifolds*, *Comm. Anal. Geom.* **8** (2000) 727–760.

[79] N. J. Hitchin, A. Karlhede, U. Lindström, and M. Roček, *Hyperkähler metrics and supersymmetry*, *Commun. Math. Phys.* **108** (1987) 535–589.

[80] T. Kitao, K. Ohta, and N. Ohta, *Three-dimensional gauge dynamics from brane configurations with (p,q)-fivebrane*, *Nucl. Phys.* **B539** (1999) 79–106, [hep-th/9808111].

[81] O. Bergman, A. Hanany, A. Karch, and B. Kol, *Branes and supersymmetry breaking in three-dimensional gauge theories*, *JHEP* **9910** (1999) 036, [hep-th/9908075].

[82] Y. Imamura and K. Kimura, *On the moduli space of elliptic Maxwell-Chern-Simons theories*, *Prog.Theor.Phys.* **120** (2008) 509–523, [arXiv:0806.3727].

[83] S. S. Gubser and I. R. Klebanov, *Baryons and domain walls in an N=1 superconformal gauge theory*, *Phys.Rev.* **D58** (1998) 125025, [hep-th/9808075].

[84] D. Berenstein, C. P. Herzog, and I. R. Klebanov, *Baryon spectra and AdS /CFT correspondence*, *JHEP* **0206** (2002) 047, [hep-th/0202150].

[85] A. Kapustin, B. Willett, and I. Yaakov, *Nonperturbative Tests of Three-Dimensional Dualities*, *JHEP* **10** (2010) 013, [arXiv:1003.5694].

[86] E. Witten, *SL(2,Z) action on three-dimensional conformal field theories with Abelian symmetry*, hep-th/0307041.

[87] A. Hanany and E. Witten, *Type IIB superstrings, BPS monopoles, and three-dimensional gauge dynamics*, *Nucl.Phys.* **B492** (1997) 152–190, [hep-th/9611230].

[88] A. Kapustin, B. Willett, and I. Yaakov, *Tests of Seiberg-like Duality in Three Dimensions*, arXiv:1012.4021.

[89] B. Willett and I. Yaakov, $\mathcal{N} = 2$ *Dualities and Z Extremization in Three Dimensions*, arXiv:1104.0487.

[90] N. Drukker, M. Marino, and P. Putrov, *From weak to strong coupling in ABJM theory*, *Commun.Math.Phys.* **306** (2011) 511–563, [arXiv:1007.3837].

[91] I. R. Klebanov, S. S. Pufu, and B. R. Safdi, *F-Theorem without Supersymmetry*, *JHEP* **1110** (2011) 038, [arXiv:1105.4598].

[92] A. Butti and A. Zaffaroni, *From toric geometry to quiver gauge theory: The Equivalence of a-maximization and Z-minimization*, *Fortsch.Phys.* **54** (2006) 309–316, [hep-th/0512240].

[93] R. Eager, *Equivalence of A-Maximization and Volume Minimization*, arXiv:1011.1809.

[94] D. Fabbri, P. Fre', L. Gualtieri, C. Reina, A. Tomasiello, *et. al.*, *3-D superconformal theories from Sasakian seven manifolds: New nontrivial evidences for AdS_4 / CFT_3*, *Nucl.Phys.* **B577** (2000) 547–608, [hep-th/9907219].

[95] D. Martelli and J. Sparks, *Moduli spaces of Chern-Simons quiver gauge theories and AdS(4)/CFT(3)*, *Phys.Rev.* **D78** (2008) 126005, [arXiv:0808.0912].

[96] A. Hanany and A. Zaffaroni, *Tilings, Chern-Simons Theories and M2 Branes*, *JHEP* **0810** (2008) 111, [arXiv:0808.1244].

[97] A. Hanany, D. Vegh, and A. Zaffaroni, *Brane Tilings and M2 Branes*, *JHEP* **0903** (2009) 012, [arXiv:0809.1440].

Appendix A

Necklace Quivers

A.1 Proof of the tree formula

We prove the Tree Formula by assuming that the columns β_a are ordered in the way described in Corollary 3.2.3. Thus, the proof reduces to showing that the following equation holds

$$\sum_{a=1}^{d} \frac{\gamma_{a(a+1)}}{\sigma_a \sigma_{a+1}} = 2 \frac{\sum_{(V,E)\in T} \prod_{(a,b)\in E} \gamma_{ab}}{\prod_{a=1}^{d} \sigma_a} \ . \tag{A.1}$$

We have the "Ptolemy relations"

$$\gamma_{ab}\gamma_{cd} = \gamma_{ac}\gamma_{bd} + \gamma_{ad}\gamma_{bc} \qquad \text{if} \qquad ab//cd. \tag{A.2}$$

The notation // means a and b separate c and d. We can use the relations to show

$$\gamma_{(a-1)(a+1)}\sigma_a + 2\gamma_{(a-1)a}\gamma_{a(a+1)} = \gamma_{(a-1)a}\sigma_{a+1} + \gamma_{a(a+1)}\sigma_{a-1}. \tag{A.3}$$

Our starting point is the Kirchhoff matrix-tree theorem which gives a relation between the absolute value of a certain determinant and the sum over trees. In particular, consider

the matrix Q^* where

$$Q^* = \begin{pmatrix} \sigma_2 & -\gamma_{23} & -\gamma_{24} & \dots & -\gamma_{2d} \\ -\gamma_{32} & \sigma_3 & -\gamma_{34} & \dots & -\gamma_{3d} \\ -\gamma_{42} & -\gamma_{43} & \sigma_4 & \dots & -\gamma_{4d} \\ \vdots & \vdots & \vdots & dots & \vdots \\ -\gamma_{d2} & -\gamma_{d3} & -\gamma_{d4} & \dots & \sigma_d \end{pmatrix} . \tag{A.4}$$

The matrix-tree theorem states that

$$\det Q^* = \sum_{(V,E)\in T} \prod_{(a,b)\in E} \gamma_{ab} \, . \tag{A.5}$$

We observe that if we take $\gamma_{(a+1)a}$ times the $(a-2)$nd row minus $\gamma_{(a-1)(a+1)}$ times the $(a-1)$st row plus $\gamma_{a(a-1)}$ times the ath row, then most of the entries will cancel out. So we have $AQ^* = B$ where

$$A = \begin{pmatrix} \gamma_{13}+\gamma_{23} & -\gamma_{12}+\gamma_{23} & \gamma_{23} & \dots & \gamma_{23} \\ -\gamma_{34} & \gamma_{24} & -\gamma_{23} & \dots & 0 \\ 0 & -\gamma_{45} & \gamma_{35} & \dots & 0 \\ \vdots & \vdots & \vdots & dots & \vdots \\ \gamma_{(d-1)d} & \gamma_{(d-1)d} & \gamma_{(d-1)d} & \dots & \gamma_{(d-1)1}+\gamma_{(d-1)d} \end{pmatrix} \tag{A.6}$$

$$B = \begin{pmatrix} \gamma_{12}\sigma_3+\gamma_{23}\sigma_1 & -\gamma_{12}\sigma_3 & 0 & \dots & 0 \\ -\gamma_{34}\sigma_2 & \gamma_{23}\sigma_4+\gamma_{34}\sigma_2 & -\gamma_{23}\sigma_4 & \dots & 0 \\ 0 & -\gamma_{45}\sigma_3 & \gamma_{34}\sigma_5+\gamma_{45}\sigma_3 & \dots & 0 \\ \vdots & \vdots & \vdots & dots & \vdots \\ 0 & 0 & 0 & \dots & \gamma_{(d-1)d}\sigma_1+\gamma_{d1}\sigma_{d-1} \end{pmatrix} \tag{A.7}$$

In constructing A we used the fact that the missing row of Q is minus the sum of all of the other rows. We used (A.2) and (A.3) to simplify B. Tri-diagonal matrices satisfy a three

term recurrence relation. If we let B_a be the matrix consisting of the first a rows of B, then

$$\det B_a = (\gamma_{a(a+1)}\sigma_{a+2} + \gamma_{(a+1)(a+2)}\sigma_a)\det B_{a-1} - \gamma_{(a+1)(a+2)}\gamma_{(a-1)a}\sigma_a\sigma_{a+1}\det B_{a-2}. \quad \text{(A.8)}$$

We can show by induction that

$$\det(B_a) = \left(\prod_{b=2}^{a}\gamma_{b(b+1)}\right)\left(\prod_{b=1}^{a+2}\sigma_b\right)\left(\sum_{b=1}^{a+1}\frac{\gamma_{b(b+1)}}{\sigma_b\sigma_{b+1}}\right) \quad \text{(A.9)}$$

and in particular

$$\det(B) = \det(B_{d-1}) = \frac{\sigma_1}{\gamma_{12}\gamma_{d1}}\left(\prod_{a=1}^{d}\sigma_a\gamma_{a(a+1)}\right)\left(\sum_{a=1}^{d}\frac{\gamma_{a(a+1)}}{\sigma_a\sigma_{a+1}}\right). \quad \text{(A.10)}$$

Therefore we want to show that $\det(A) = 2\sigma_1\prod_{a=2}^{d-1}\gamma_{a(a+1)}$.

We let A' be A with the γ_{23} in the top row and the $\gamma_{(d-1)d}$ in the bottom row removed. Let A'_{ab} be the matrix containing the ath through bth rows and columns of A'. Write the determinant of A as

$$\det(A) = (A'_1 + \gamma_{23}u)\wedge A'_2\wedge\ldots\wedge A'_{d-2}\wedge(A'_{d-1} + \gamma_{(d-1)d}u)$$

where $u = (1, 1, \ldots, 1)$ and A'_a are the rows of A'. Expanding out the anti-symmetric product, we find that

$$\begin{aligned}\det(A) &= \det(A') + \gamma_{23}u\wedge A'_2\wedge\ldots\wedge A'_{d-1} + \gamma_{(d-1)d}A'_1\wedge\ldots\wedge A'_{d-2}\wedge u \\ &= \det(A') + \sum_{a=1}^{d}\left[\gamma_{23}\left(\prod_{b=1}^{a-1}\gamma_{(b+2)(b+3)}\right)\det A'_{(a+1)(d-1)}\right. \\ &\quad \left.+\gamma_{d(d-1)}\left(\prod_{b=a+2}^{d}\gamma_{(b-1)(b-2)}\right)\det A'_{1(a-1)}\right]. \quad \text{(A.11)}\end{aligned}$$

Using the three term recurrence relation for the determinant of a tridiagonal matrix, we can show by induction that

$$\begin{aligned} \det A'_{1a} &= \gamma_{1(a+2)} \prod_{b=2}^{a} \gamma_{b(b+1)} \ , \\ \det A'_{a(d-1)} &= \gamma_{1a} \prod_{b=a}^{d-2} \gamma_{(b+1)(b+2)} \ , \\ \det A' &= 0 \, , \end{aligned} \tag{A.12}$$

where last equation follows from the second one by setting $a = 1$ and using $\gamma_{11} = 0$. Therefore

$$\begin{aligned} \det(A) &= 2 \left(\sum_a \gamma_{1a} \right) \left(\prod_{a=2}^{d-1} \gamma_{a(a+1)} \right) \\ &= 2\sigma_1 \prod_{a=2}^{d-1} \gamma_{a(a+1)} \end{aligned} \tag{A.13}$$

as expected.

A.2 A virial theorem for matrix models

To leading order in N, the matrix model for superconformal field theories with M-theory duals involves extremizing a free energy functional of the type

$$F[\rho, y_a] = \int dx \, \rho(x)^2 f(y_a(x)) - \int dx \, \rho(x) V(x, y_a(x)) \, , \tag{A.14}$$

for some functions f and V. In all examples, V is homogeneous of degree 1 in x, namely it satisfies

$$V = x \frac{\partial V}{\partial x} \, . \tag{A.15}$$

The free energy functional (A.14) should be extremized under the constraint that ρ is a density normalized so that $\int dx \, \rho(x) = 1$. Such a constraint can be implemented with a

Lagrange multiplier μ by defining a new functional

$$\tilde{F}[\rho, y_a] = F[\rho, y_a] - 2\pi N^{3/2}\mu\left(\int dx\,\rho(x) - 1\right) \tag{A.16}$$

and varying (A.16) instead of (A.14) with respect to ρ and y_a. If we denote by F the on-shell value of $F[\rho, y_a]$ (or of $\tilde{F}[\rho, y_a]$), we will now show that

$$F = \frac{4\pi N^{3/2}}{3}\mu\,, \tag{A.17}$$

regardless of which particular matrix model we're solving provided that eq. (A.15) is obeyed.

A.2.1 A slick proof

The equations of motion following from (A.16) are

$$\begin{aligned} 2\rho(x)f(y_a(x)) - V(x, y_a(x)) &= 2\pi N^{3/2}\mu\,, \\ \rho(x)^2\partial_a f(y_a(x)) - \rho(x)\partial_a V(x, y_a(x)) &= 0\,, \end{aligned} \tag{A.18}$$

Differentiating the first equation with respect to x, multiplying it by $\rho(x)$, and using the second equation we get

$$\rho(x)\left[2\rho'(x)f(y_a(x)) + \rho(x)^2\sum_a \partial_a f(y_a(x))y_a'(x)\right] = \rho(x)\frac{\partial}{\partial x}V(x, y_a(x))\,. \tag{A.19}$$

To prove (A.17), consider the function

$$G(x) = x\rho(x)^2 f(y_a(x))\,. \tag{A.20}$$

Let's calculate the derivative of this function with respect to x:

$$\begin{aligned} G'(x) &= \rho(x)^2 f(y_a(x)) + 2x\rho(x)\rho'(x)f(y_a(x)) + x\rho(x)^2\sum_a \partial_a f(y_a(x))y_a'(x) \\ &= \rho(x)^2 f(y_a(x)) + x\rho(x)\frac{\partial}{\partial x}V(x, y_a(x)) \\ &= \rho(x)^2 f(y_a(x)) + \rho(x)V(x, y_a(x))\,, \end{aligned} \tag{A.21}$$

where in the second line we used (A.19) and in the third line we used (A.15). The function $G(x)$ vanishes at $\pm\infty$, so $\int dx\, G'(x) = 0$, which from (A.21) implies that on-shell we have

$$\int dx\, \rho(x)^2 f(y_a(x)) = -\int dx\, \rho(x) V(x, y_a(x))\,. \tag{A.22}$$

From (A.14) we have then that

$$F = 2\int dx\, \rho(x)^2 f(y_a(x))\,. \tag{A.23}$$

Multiplying the first equation in (A.18) by $\rho(x)$ and integrating in x we get

$$2\pi N^{3/2}\mu = 3\int dx\, \rho(x)^2 f(y_a(x))\,. \tag{A.24}$$

Taking the ratio of the last two equations one obtains (A.17).

A.2.2 A more enlightening proof

The proof given above is really that of a virial theorem. To put the virial theorem in a more familiar form, let's define the cumulative distribution $t(x) = \int_{-\infty}^{x} dx' \rho(x') \in [0,1]$ and express (A.14) as the action

$$S[x(t), y_a(t)] = \int dt\, L\,, \qquad L \equiv \frac{f(y_a)}{\dot{x}} - V(x, y_a)\,, \tag{A.25}$$

where x and y_a should be thought of as functions of t. To go to a Hamiltonian formulation, we introduce the momentum conjugate to x:

$$p_x \equiv \frac{\partial L}{\partial \dot{x}} = -\frac{f(y_a)}{\dot{x}^2}\,. \tag{A.26}$$

The Hamiltonian is

$$H \equiv p_x \dot{x} - L = \mathrm{KE} + \mathrm{PE}\,, \qquad \mathrm{KE} \equiv -2\frac{f(y_a)}{\dot{x}} = -2 f(y_a)^{1/2}\sqrt{-p_x}\,, \qquad \mathrm{PE} \equiv V(x, y_a)\,. \tag{A.27}$$

In deriving a virial theorem for this Hamiltonian, one can just ignore the dependence on y_a because the y_a are non-dynamical. Since the kinetic energy KE is homogeneous of degree $1/2$ in p_x and the potential energy is homogeneous of degree 1 in x, we have

$$\frac{1}{2}\langle \mathrm{KE}\rangle = \langle \mathrm{PE}\rangle\,, \tag{A.28}$$

where $\langle\cdots\rangle$ means $\int_0^1 dt\,(\cdots)$. From (A.25) one sees that

$$F \equiv S_{\text{on-shell}} = -\frac{1}{2}\langle \mathrm{KE}\rangle - \langle \mathrm{PE}\rangle = -\langle \mathrm{KE}\rangle\,. \tag{A.29}$$

From the first equation in (A.18) we see that $H = -2\pi N^{3/2}\mu$, so

$$\frac{N^{3/2}}{2\pi}\mu = -\langle H\rangle = -\langle \mathrm{KE}\rangle - \langle \mathrm{PE}\rangle = -\frac{3}{2}\langle \mathrm{KE}\rangle\,. \tag{A.30}$$

Taking the ratio of (A.29) and (A.30) one again obtains (A.17).

A.3 F-Maximization for the necklace quivers

We would like to show that to leading order in N the free energy of the necklace quivers with arbitrary R-charges studied in section 4.4.1 is maximized when $R[A_a] = R[B_a] = 1/2$ and $\Delta_m = 0$. We can only show this if the gauge groups are $SU(N)$. In the $U(N)^d$ case, the symmetries (4.7) imply that the free energy has flat directions, but we can nevertheless show that the free energy is maximized when the invariant combinations c_1 and c_2 defined in eq. (4.46) are set to zero. The critical R-charges correspond to the case where there is $\mathcal{N} = 3$ supersymmetry as opposed to just $\mathcal{N} = 2$.

The essential ingredient of the proof is the observation that the polygon $\mathcal{P}$, which depends on $\vec{c}$, is the polar dual of a polygon $\mathcal{Q}$ that does not depend on $\vec{c}$ about the unit circle centered at $(-\vec{c}/2)$. Let $\vec{\beta_a} = (1, q_a)$ and $\vec{c} = (c_1, c_2)$ be vectors in $\mathbb{R}^2$. The

polygon $\mathcal{Q}$ is the Minkowski sum

$$\mathcal{Q} = \left\{ \sum_{a=1}^{d} u_a \vec{\beta}_a \in \mathbb{R}^2 : u_a \in (-1/2, 1/2) \right\} \tag{A.31}$$

of the vectors $\vec{\beta}_a$. Indeed, one can rewrite $\mathcal{P}$ as the intersection of half-planes

$$\mathcal{P} = \left\{ \vec{t} \in \mathbb{R}^2 : \frac{1}{2}\vec{t} \cdot \vec{c} + \sum_{a=1}^{d} \vec{t} \cdot \left(u_a \vec{\beta}_a \right) \leq 1, \forall u_a \in (-1/2, 1/2) \right\} . \tag{A.32}$$

The boundaries of these half-planes are precisely the polar duals of the points in $\mathcal{Q}$ about the unit circle centered at $(-\vec{c}/2)$.

Let $\vec{v}_i$ be the vertices of $\mathcal{Q}$ ordered so that the line segment between $\vec{v}_i$ and $\vec{v}_{i+1}$ is part of the boundary of $\mathcal{Q}$. The line passing through $\vec{v}_i$ and $\vec{v}_{i+1}$ is polar dual to a vertex $\vec{w}_{i,i+1}$ of $\mathcal{P}$. Polar duality implies $\vec{w}_{i,i+1} \cdot (\vec{v}_i + \vec{c}/2) = \vec{w}_{i,i+1} \cdot (\vec{v}_{i+1} + \vec{c}/2) = 1$, so

$$\vec{w}_{i,i+1} = \frac{*(\vec{v}_{i+1} - \vec{v}_i)}{(*(\vec{v}_{i+1} + \vec{c}/2)) \cdot (\vec{v}_i + \vec{c}/2)} , \tag{A.33}$$

where $*$ denotes the Hodge dual in $\mathbb{R}^2$. By splitting $\mathcal{P}$ into triangles we can write the area of $\mathcal{P}$ as

$$\text{Area}(\mathcal{P}) = \sum_i \text{Area}(\vec{w}_{i-1,i}, \vec{w}_{i,i+1}, 0) = \sum_i \frac{1}{2} \left| \vec{w}_{i-1,i} \cdot (*\vec{w}_{i,i+1}) \right| , \tag{A.34}$$

where we denoted the area of a triangle whose vertices are given by the vectors $\vec{\alpha}$, $\vec{\beta}$, and $\vec{\gamma}$ by $\text{Area}(\vec{\alpha}, \vec{\beta}, \vec{\gamma})$. Using eq. (A.33), eq. (A.34) becomes

$$\text{Area}(\mathcal{P}) = \frac{1}{4} \sum_i \frac{\text{Area}(\vec{v}_{i-1}, \vec{v}_i, \vec{v}_{i+1})}{\text{Area}(\vec{v}_i, \vec{v}_{i-1}, -\vec{c}/2)\, \text{Area}(\vec{v}_{i+1}, \vec{v}_i, -\vec{c}/2)} . \tag{A.35}$$

As long as $-\vec{c}/2$ belongs to the interior of $\mathcal{Q}$, the Hessian matrix of each term in this sum, seen as a function of $\vec{c}$, is positive definite, so $\text{Area}(\mathcal{P})$ is a convex function of $\vec{c}$. (To compute the Hessian it is easiest to work in a coordinate system where $\vec{c}$ is parametrized by the distance from two neighboring sides of the polygon to $-\vec{c}/2$.)

In our case $\mathcal{Q}$ is symmetric about the origin as can be easily seen from eq. (A.31). Consequently, Area($\mathcal{P}$) is an even function of $\vec{c}$, and we have just shown that it is also convex. It follows that Area($\mathcal{P}$) is minimized for $\vec{c} = 0$. Equivalently, the free energy is maximized when $\vec{c} = 0$. Using the F-maximization conjecture of [25], we have thus shown that the correct R-charges in the necklace quivers with superpotential (4.42) satisfy $c_1 = c_2 = 0$. That's all one can say about the $U(N)^d$ theory. If the gauge groups are instead $SU(N)$, the tracelessness constraints $\int dx\, \rho(x)\delta y_a(x) = 0$ imply (when $\vec{c} = 0$)

$$\int dx\, \rho(x)\delta y_a(x) = \frac{R[B_a] - R[A_a]}{2} = 0\,, \tag{A.36}$$

so $R[A_a] = R[B_a] = 1/2$. From $c_2 = 0$ we also get $\Delta_m = 0$.

Appendix B

Further examples of operator counting

For notational convenience we set $T^{(1)} = T$ and $T^{(-1)} = \tilde{T}$.

B.1 Flavored ABJM theory

We consider the flavored ABJM model with the superpotential

$$W \sim \mathrm{tr}\left[\epsilon^{ij}\epsilon^{kl}A_iB_kA_jB_l + \sum_{j=1}^{n_{a1}} q_j^{(1)}A_1\tilde{q}_j^{(1)} + \sum_{j=1}^{n_{a2}} q_j^{(2)}A_2\tilde{q}_j^{(2)} + \sum_{j=1}^{n_{b1}} Q_j^{(1)}B_1\tilde{Q}_j^{(1)} + \sum_{j=1}^{n_{b2}} Q_j^{(2)}B_2\tilde{Q}_j^{(2)}\right]. \tag{B.1}$$

When $N = 1$, the superpotential is supplemented by the relation (4.10) which in this case is $T\tilde{T} = A_1^{n_{a1}}A_2^{n_{a2}}B_1^{n_{b1}}B_2^{n_{b2}}$ [32, 33]. The corresponding matrix model was solved in the large N limit in [21]. Our strategy is the same as for the flavored $\mathcal{N} = 8$ theory. In this section, we will review the solution for $\rho(x)$ and $\delta y \equiv y_1 - y_2$. In the next section, we will compare these results with the distribution of operators in the chiral ring.

We define $R[A_i] \equiv \Delta_{A_i}$, $R[B_i] \equiv \Delta_{B_i}$, $\Delta \equiv R[T]$, and $\tilde{\Delta} \equiv R[\tilde{T}]$. Without loss of generality, we will assume that $\Delta_{A_2} < \Delta_{A_1}$ and $\Delta_{B_2} < \Delta_{B_1}$. To keep the notation concise,

we also define

$$\begin{aligned} k_\pm &\equiv k \pm \frac{1}{2}\left(n_{a1}+n_{a2}-n_{b1}-n_{b2}\right) , \\ \Delta_2 &\equiv \Delta_{A_1}\Delta_{A_2}-\Delta_{B_1}\Delta_{B_2} , \\ \Delta_3 &\equiv \Delta_{A_1}\Delta_{A_2}(\Delta_{B_1}+\Delta_{B_2})+\Delta_{B_1}\Delta_{B_2}(\Delta_{A_1}+\Delta_{A_2}) . \end{aligned} \tag{B.2}$$

Taking the marginality constraints on the R-charges into account, in the large N limit, the matrix model free energy functional is

$$\begin{aligned} \frac{\tilde{F}[\rho,\delta y]}{2\pi N^{3/2}} = \int dx\,\rho\Bigg[&\frac{1}{2}(k_+ + k_-)x\,\delta y - \rho\left((\delta y)^2+\Delta_2\,\delta y-\frac{1}{2}\Delta_3\right)+\frac{1}{2}(\Delta-\tilde{\Delta})x \\ &+\frac{1}{2}\,|x|\left(\Delta+\tilde{\Delta}+(k_+-k_-)\,\delta y\right)\Bigg]-\mu\left(\int dx\,\rho-1\right) . \end{aligned} \tag{B.3}$$

The eigenvalue density has four regions:

$$-\frac{1}{R[\tilde{T}A_2^{k_-}]}<\hat{x}<-\frac{1}{R[\tilde{T}A_1^{k_-}]}: \qquad \hat{\rho}=\frac{1+\hat{x}R[\tilde{T}A_2^{k_-}]}{\Delta_3+2\Delta_{A_2}\Delta_2-2\Delta_{A_2}^2} , \quad \delta\hat{y}=-\Delta_{A_2} ; \tag{B.4}$$

$$-\frac{1}{R[\tilde{T}A_1^{k_-}]}<\hat{x}<0: \qquad \hat{\rho}=\frac{2+2\hat{x}\tilde{\Delta}+\hat{x}k_-\Delta_2}{\Delta_2^2+2\Delta_3} ,$$

$$\delta\hat{y}=\frac{k_-\hat{x}\Delta_3-(1+\hat{x}\tilde{\Delta})\Delta_2}{2+2\hat{x}\tilde{\Delta}+\hat{x}k_-\Delta_2} ; \tag{B.5}$$

$$0<\hat{x}<\frac{1}{R[TB_1^{k_+}]}: \qquad \hat{\rho}=\frac{2-2\hat{x}\Delta+\hat{x}k_+\Delta_2}{\Delta_2^2+2\Delta_3} ,$$

$$\delta\hat{y}=\frac{k_+\hat{x}\Delta_3-(1-\hat{x}\Delta)\Delta_2}{2-2\hat{x}\Delta+\hat{x}k_+\Delta_2} ; \tag{B.6}$$

$$\frac{1}{R[TB_1^{k_+}]}<\hat{x}<\frac{1}{R[TB_2^{k_+}]}: \qquad \hat{\rho}=\frac{1-\hat{x}R[TB_2^{k_+}]}{\Delta_3-2\Delta_{B_2}\Delta_2-2\Delta_{B_2}^2} , \quad \delta\hat{y}=\Delta_{B_2} ; \tag{B.7}$$

As in (4.15), we have introduced the rescaled variables $x=\hat{x}\mu$ and $\rho(x)=\hat{\rho}(\hat{x})\mu$.

Operator counting

There are operators containing $\tilde{T}^{-m}$ for $m<0$ and operators containing T^m for $m>0$. They take the form $T^m A_1^{\alpha_1}A_2^{\alpha_2}B_1^{\beta_1}B_2^{\beta_2}$ and $\tilde{T}^{-m}A_1^{\alpha_1}A_2^{\alpha_2}B_1^{\beta_1}B_2^{\beta_2}$, where gauge invariance demands $\alpha_1+\alpha_2-\beta_1-\beta_2=-mk_\pm$. If we wanted to count operators that don't vanish when, for example, $A_1=0$, then we just set $\alpha_1=0$.

We counted the operators using a slightly modified version of the method outlined in appendix B.3. Having written the operators in terms of both T and $\tilde{T}$, it is simpler to use two different coordinate systems on the cone C, one when $m > 0$ and one when $m < 0$. The coordinate systems are related by (4.10). The operator counts reproduce (B.4), (B.5), (B.6), and (B.7) via our conjecture (4.3).

Here are some of the details for the calculation of $\hat{\rho}(\hat{x})$ when $m > 0$. The density of operators is given by

$$\begin{aligned}\frac{\partial^2\psi}{\partial r\partial m} = &\int d\alpha_1\, d\alpha_2\, d\beta_1\, d\beta_2\, \delta(\alpha_1+\alpha_2-\beta_1-\beta_2+mk_+) \\ &\times \delta(r - m\Delta - \alpha_1\Delta_{A_1} - \alpha_2\Delta_{A_2} - \beta_1\Delta_{B_1} - \beta_2\Delta_{B_2})\ .\end{aligned} \tag{B.8}$$

This integral gives the area of a slice of a tetrahedron. The slice is either a triangle or a quadrilateral (which may be regarded as a triangle with another triangle cut out). We find for $m \geq 0$

$$\frac{\partial^2\psi}{\partial r\partial m} = \sum_{j=1}^{2} \frac{(r - mR[TB_j^{k_+}])^2\,\theta(r - mR[TB_j^{k_+}])}{2(\Delta_3 - 2\Delta_{B_j}\Delta_2 - 2\Delta_{B_j}^2)}\ . \tag{B.9}$$

Taking a derivative of this expression with respect to r yields (B.6) and (B.7).

B.2 $\mathbb{C}^3/(\mathbb{Z}_2 \times \mathbb{Z}_2)$ theory

Let's examine the field theory in figure B.1. It has four gauge groups with CS levels k_a, $a = 1, \ldots, 4$, and twelve bifundamental fields A_{ab} transforming in $(\overline{\mathbf{N}}_a, \mathbf{N}_b)$, one for every ordered pair (a, b) with $a \neq b$. The superpotential is

$$W = \mathrm{tr}\left[\sum_{a=1}^{4} \epsilon_{abcd} A_{db} A_{cd} A_{bc}\right]\ . \tag{B.10}$$

The superpotential relations are supplemented by the monopole OPE (4.10) $T\tilde{T} = 1$. We define $R[A_{ab}] \equiv R_{ab}$ and $R[T] = -R[\tilde{T}] \equiv \Delta$.

The superpotential contains eight distinct terms that impose the relations $R_{ab} + R_{bc} + R_{ca} = 2$ for any triplet (a, b, c) of pairwise distinct gauge groups. These eight equations

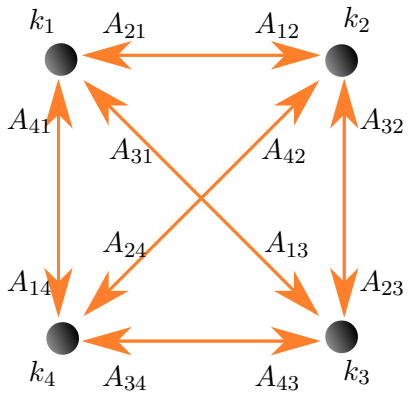


Figure B.1: The quiver for $\mathbb{C}^3/(\mathbb{Z}_2 \times \mathbb{Z}_2)$. There are four $U(N)$ gauge groups with Chern-Simons coefficients k_a. The matter content consists of the 12 bifundamental fields A_{ab} for $a \neq b$, transforming under the fundamental of the bth gauge group and the antifundamental of the ath gauge group.

imply the long-range force cancellation (4.4). Only seven of these equations are linearly independent, leaving five independent R-charges out of the twelve R_{ab}.

Even though for given k_a the matrix model depends on 6 R-charges (Δ and the five linearly independent R_{ab}), the dependence on three of these parameters is trivial because of the flat directions (4.7). We can use these symmetries to reduce the number of independent R-charges to three: Δ_x, Δ_y and Δ where we pick

$$\begin{aligned} R_{12} = R_{21} = R_{34} = R_{43} &= \Delta_x \,, \\ R_{23} = R_{32} = R_{41} = R_{14} &= \Delta_y \,, \\ R_{13} = R_{31} = R_{24} = R_{42} &= 2 - \Delta_x - \Delta_y \equiv \Delta_z \,. \end{aligned} \tag{B.11}$$

The matrix model is then

$$\begin{aligned} F[\rho, y_a] = 2\pi N^{3/2} \int dx\, \rho x \sum_{a=1}^{d} k_a y_a + 2\pi \Delta N^{3/2} \int dx\, \rho x \\ + \frac{\pi N^{3/2}}{2} \int dx\, \rho^2 \sum_{(a,b,c)} (y_b - y_a + R_{ab})\,(y_c - y_b + R_{bc})\,(y_a - y_c + R_{ca}) \ . \end{aligned} \tag{B.12}$$

For simplicity, let's focus on the case $k_1 = -k_2 = k_3 = -k_4 = k > 0$ and take $\Delta_y \geq \Delta_x$. The saddle point eigenvalue distribution splits into three regions where ρ is linear:

$$\begin{aligned}
&\frac{1}{\Delta - 2k\Delta_x} < \hat{x} < \frac{1}{\Delta - 2k\Delta_y}: \qquad \hat{\rho}(\hat{x}) = \frac{1 - \hat{x}(\Delta - 2k\Delta_x)}{4\Delta_x(\Delta_y - \Delta_x)\Delta_z}\,, \qquad \hat{y}_1 - \hat{y}_2 = -\Delta_x\,, \\
&\frac{1}{\Delta - 2k\Delta_y} < \hat{x} < \frac{1}{\Delta + 2k\Delta_y}: \qquad \hat{\rho}(\hat{x}) = \frac{1 - \hat{x}\Delta}{4\Delta_x\Delta_y\Delta_z}\,, \qquad \hat{y}_1 - \hat{y}_2 = \frac{2k\hat{x}\Delta_x\Delta_y}{1 - \hat{x}\Delta}\,, \\
&\frac{1}{\Delta + 2k\Delta_y} < \hat{x} < \frac{1}{\Delta + 2k\Delta_x}: \qquad \hat{\rho}(\hat{x}) = \frac{1 - \hat{x}(\Delta + 2k\Delta_x)}{4\Delta_x(\Delta_y - \Delta_x)\Delta_z}\,, \qquad \hat{y}_1 - \hat{y}_2 = \Delta_x\,.
\end{aligned} \tag{B.13}$$

In all three regions, $\hat{y}_1 = \hat{y}_3$ and $\hat{y}_2 = \hat{y}_4$.

Operator counting

Without the monopole operators, the ring of functions A_{ab} modulo superpotential relations is the ring of functions on $\mathbb{C}^3/(\mathbb{Z}_2 \times \mathbb{Z}_2)$. This ring consists of polynomials in x, y, z with the constraint that the numbers of x, y, z in each term must be either all even or all odd. We call $A_{12}, A_{21}, A_{34}, A_{43}$ "x fields", $A_{14}, A_{41}, A_{23}, A_{32}$ "y fields", and $A_{13}, A_{31}, A_{24}, A_{42}$ "z fields". We can get a gauge invariant operator by taking a combination of two x fields (e.g. $A_{12}A_{21}$), two y fields ($A_{13}A_{31}$), two z fields ($A_{14}A_{41}$), or one of each type of field ($A_{12}A_{23}A_{31}$). The gauge invariant operators with $m = 0$ are those with an even number of each of x, y, z, or an odd number of each of x, y, z.

Adding back the monopole operators yields a ring of functions on a four-dimensional cone. An electric charge of $(1, -1, 1, -1)$ from T can be cancelled out by two x's ($A_{12}A_{34}$) or two y's ($A_{14}A_{32}$), but not by z's. So, if we have an operator of the schematic form $T^m x^{n_x} y^{n_y} z^{n_z}$ for $m > 0$ and $\tilde{T}^{-m} x^{n_x} y^{n_y} z^{n_z}$ for $m < 0$, then we have the constraint $n_x + n_y \geq 2|m|$. The operator density is then

$$\frac{\partial^2 \psi}{\partial r \partial m} = \frac{1}{4} \int dn_x\, dn_y\, dn_z\, \theta(n_x + n_y - 2k|m|)\delta(r - \Delta_x n_x - \Delta_y n_y - \Delta_z n_z - \Delta m)\,. \tag{B.14}$$

The factor of $\frac{1}{4}$ comes from the constraint that the numbers of x, y, z must be all even or all odd.

Performing the integral over n_z introduces an overall factor of $1/\Delta_z$. The remaining integral reduces to the area of a polygonal region satisfying the constraints $n_y > 0$, $n_x > 0$,

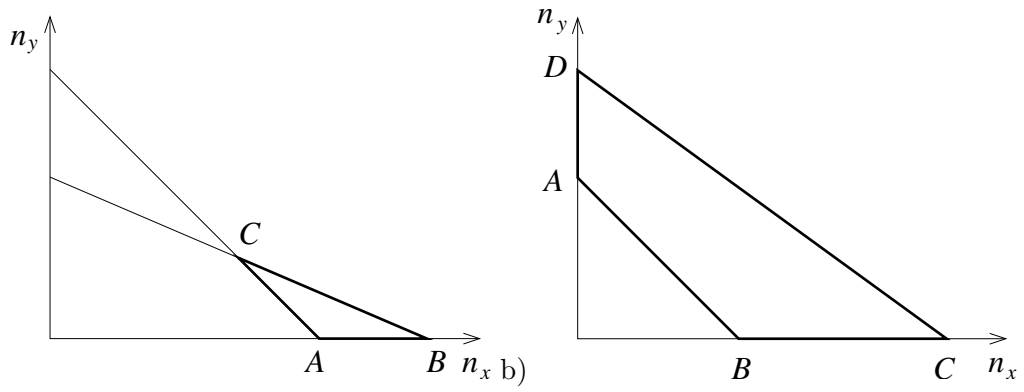


Figure B.2: The area of the polygonal regions ABC and $ABCD$ is proportional to $\partial^2\psi/\partial r\partial m$ for the $\mathbb{C}^3/\mathbb{Z}_2\times\mathbb{Z}_2$ quiver: a) $r-\Delta m-2k|m|\Delta_y<0$; b) $r-\Delta m-2k|m|\Delta_y>0$

$n_x+n_y>2k|m|$, and $\Delta_x n_x+\Delta_y n_y<r-\Delta m$. For small $|m|$, the polygonal region is a quadrilateral while for large $|m|$, the region is a triangle (see figure B.2). Assuming that $\Delta_y>\Delta_x$, we find

$$\frac{\partial^2\psi}{\partial r\partial m}=\begin{cases}\frac{1}{8\Delta_z}\frac{(r-\Delta m-2k|m|\Delta_x)^2}{\Delta_x(\Delta_y-\Delta_x)^2} & \text{if } r-\Delta m-2k|m|\Delta_y<0\ ,\\ \frac{1}{8\Delta_z}\left[\frac{(r-\Delta m)^2}{\Delta_x\Delta_y}-(2k|m|)^2\right] & \text{if } r-\Delta m-2k|m|\Delta_y>0\ .\end{cases} \tag{B.15}$$

Taking an additional derivative with respect to r, we can easily check that this formula agrees with (B.13).

Now, in order to compute $\hat{y}_1(\hat{x})-\hat{y}_4(\hat{x})$, we count gauge invariant operators with A_{14} set to zero. Because of the superpotential relations, all operators with a z are set to zero. The factor of 1/4 remains the same because now we may only consider operators with even numbers of x and y fields. The expression for $\partial\psi_{14}/\partial m$ is given by the area of the same polygonal region that governs $\partial^2\psi/\partial r\partial m$, but we lose the factor of Δ_z because we drop the integral over n_z:

$$\frac{\partial^2\psi_{14}}{\partial r\partial m}=\Delta_z\frac{\partial^3\psi}{\partial r^2\partial m}\ . \tag{B.16}$$

Therefore, we have $\hat{\rho}(\hat{x})(\Delta_z+\hat{y}_1(\hat{x})-\hat{y}_4(\hat{x}))=\hat{\rho}(\hat{x})\Delta_z$, and hence $\hat{y}_1(\hat{x})=\hat{y}_4(\hat{x})$. A similar calculation shows $\hat{y}_2(\hat{x})=\hat{y}_3(\hat{x})$.

Finally we count the operators with A_{12} set to zero. Most operators with an x will become zero. However, fields containing only T, A_{21}, A_{43}, A_{23}, A_{41}, and the z fields are

not set to zero by the superpotential relations. So the nonzero fields are those with $n_x = 0$ and an even number of y and z fields, or $m \geq 0$ and $n_x + n_y = 2km$ with an even number of z fields. After a little work, we find

$$\frac{\partial^2 \psi_{12}}{\partial r \partial m} = \begin{cases} 0 & \text{if } m/r < -(2k\Delta_y - \Delta)^{-1} \\ \frac{r - \Delta m + 2km\Delta_y}{4\Delta_y \Delta_z} & \text{if } -(2k\Delta_y - \Delta)^{-1} < m/r < (2k\Delta_y + \Delta)^{-1} \\ \frac{r - \Delta m - 2km\Delta_x}{2\Delta_z(\Delta_y - \Delta_x)} & \text{if } m/r > (2k\Delta_y + \Delta)^{-1} \end{cases} \tag{B.17}$$

This result matches $\hat{y}_1(\hat{x}) - \hat{y}_2(\hat{x})$ computed from (B.13).

B.3 Toric varieties in general

By toric moduli space we mean more specifically that the moduli space for the Abelian gauge theory is an eight-dimensional toric Calabi-Yau cone V. That V is toric means it is a T^4 torus fibration over a four-dimensional rational polyhedral cone C. This polyhedral cone is the set of points satisfying

$$C = \{y \in \mathbb{R}^4 : y \cdot v_a \geq 0\} \ , \tag{B.18}$$

where $v_a \in \mathbb{Z}^4$, $a = 1, \ldots, n$, are inward pointing vectors normal to the faces F_a of the cone:

$$F_a = \{y \in C : y \cdot v_a = 0\} \ . \tag{B.19}$$

The fact that V is Calabi-Yau implies that the end-points of the vectors v_a lie in a common hyperplane $\mathbb{R}^3$.

One convenient aspect of this construction is that lattice points in C correspond to operators in the chiral ring of the Chern-Simons theory. The coordinates of a lattice point are the $U(1)$ global charges of the operator. The vector b that measures the R-charge is often called the Reeb vector where the R-charge is then $r = y \cdot b$. The vectors v_a correspond to other global charges, $q_a = y \cdot v_a$, and we can introduce additional charges as well. In the

gauge theories considered in this chapter, the monopole charge m played an important role. Let us introduce t as the vector that measures monopole charge.

We introduced previously the function $\psi(r, m)$ as the number of operators with R-charge less than r and monopole charge less than m. From the toric perspective, this function in the large r and m limit is the volume of a four-dimensional polytope:

$$C_{r,m} = C \cap \{y \cdot b \leq r\} \cap \{y \cdot t \leq m\} \ , \tag{B.20}$$

where $\psi(r, m) = \mathrm{Vol}(C_{r,m})$.

We would like to understand geometrically how to compute derivatives of $\psi(r, m)$. The value of $\psi(r, m)$ is a four-dimensional integral we can write as

$$\psi(r, m) = \int_{C_{r,m}} d^4y \ . \tag{B.21}$$

To take a derivative of ψ with respect to r, we can rotate the coordinate system so that one of the y's points in the direction of b and replace d^4y with $d^3y\, dr/|b|$ where $|b|$ is the Jacobian factor from the change of variables. The derivative is then related to the three-dimensional volume of the polyhedron

$$D_{r,m} = C \cap \{y \cdot b = r\} \cap \{y \cdot t \leq m\} \tag{B.22}$$

where $\partial\psi/\partial r = \mathrm{Vol}(D_{r,m})/|b|$.[1]

Similarly, we can visualize $\partial^2\psi/\partial r\partial m$ as the area of a two-dimensional polygon $P_{r,m}$:

$$P_{r,m} = C \cap \{y \cdot b = r\} \cap \{y \cdot t = m\} \ . \tag{B.23}$$

[1]This last expression may seem strange because the right hand side seems to depend on a metric while the left hand side depends only on a volume form on C. Interpreting $\mathrm{Vol}(D_{r,m})$ as a three form instead of a number, we could rewrite this expression in a manifestly metric independent way: $(\partial\psi/\partial r)t = \star\,\mathrm{Vol}(D_{r,m})$.

Now we rotate our coordinate system so that two of the y's lie in the plane spanned by b and t. The Jacobian factor is $|t \wedge b| = \sqrt{t^2 b^2 - (t \cdot b)^2}$. Geometrically, the second partial is

$$\frac{\partial^2 \psi}{\partial r \partial m} = \frac{\text{Area}(P_{r,m})}{|t \wedge b|} \ . \tag{B.24}$$

The function $\psi_X(r, m)$ has a toric interpretation as well. In the examples we considered, X corresponds to an integer linear combination of the v_a. Let us consider the simple case where X_a corresponds to a single v_a. Operators with no X_a are contained in the face $F_a \subset C$. This fact suggests a relation between $\psi_{X_a}(r, m)$ and a generalization of $\psi(r, m)$ involving a third charge q_a, $\psi(r, m, q_a)$. In particular, it is true that

$$\psi_{X_a}(r, m) = \psi^{(0,0,1)}(r, m, 0) \ . \tag{B.25}$$

Operators with no X_a and fixed m and r lie along a line $L_{a,m,r} \subset F_a$:

$$L_{a,m,r} = F_a \cap \{y \cdot b = r\} \cap \{y \cdot t = m\} \ . \tag{B.26}$$

Generalizing the argument used to derive (B.24) to one more charge, we find

$$\frac{\partial \psi^2_{X_a}}{\partial r \partial m} = \psi^{(1,1,1)}(r, m, 0) = \frac{\text{Length}(L_{a,m,r})}{|t \wedge b \wedge v_a|} \ . \tag{B.27}$$

Eqs. (B.24) and (B.27) provide a convenient starting point for counting chiral operators in the examples in the text.

B.4 The Cone over $Q^{2,2,2}/\mathbb{Z}_k$

As another example with chiral bifundamental fields, we can examine the square quiver in figure B.3 with CS levels $(k, k, -k, -k)$ and matter fields A_i, B_i, C_i, and D_i, with $i = 1, 2$. With the superpotential is

$$W \sim \text{tr}\left[\epsilon^{ij}\epsilon^{kl} D_i C_k B_j A_l\right] \tag{B.28}$$

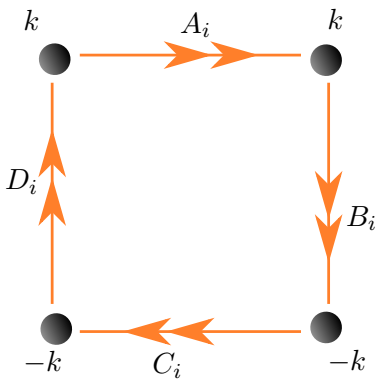


Figure B.3: Quiver gauge theory believed to be dual to $AdS_4 \times Q^{2,2,2}/\mathbb{Z}_k$.

this quiver is thought to be dual to $AdS_4 \times Q^{2,2,2}/\mathbb{Z}_k$ [67, 97]. The quiver has two flavor $SU(2)$ symmetries, one under which A_i and C_i transform as doublets, and one under which B_i and D_i transform as doublets, so one expects the R-charges of the fields belonging to the same edge of the quiver to be equal when F is maximized. Using the flat directions (4.7) and taking into account the marginality of the superpotential (B.28), one can then set the R-charges of all the bifundamental fields equal to $1/2$ and $\Delta_m = 0$. With this choice one can go through the operator counting exercise in the Abelian theory and predict that

$$\begin{aligned}
\hat{\rho}(\hat{x}) &= \theta\left(\frac{1}{2k} - |\hat{x}|\right) + \frac{1}{4k}\delta\left(\frac{1}{2k} + \hat{x}\right) + \frac{1}{4k}\delta\left(\frac{1}{2k} - \hat{x}\right) , \\
\hat{\rho}(\hat{x})\left(\hat{y}_2(\hat{x}) - \hat{y}_1(\hat{x})\right) &= -\frac{1}{8k}\delta\left(\frac{1}{2k} + \hat{x}\right) - \frac{1}{8k}\delta\left(\frac{1}{2k} - \hat{x}\right) , \\
\hat{\rho}(\hat{x})\left(\hat{y}_3(\hat{x}) - \hat{y}_2(\hat{x})\right) &= \frac{3}{8k}\delta\left(\frac{1}{2k} + \hat{x}\right) - \frac{1}{8k}\delta\left(\frac{1}{2k} - \hat{x}\right) , \\
\hat{\rho}(\hat{x})\left(\hat{y}_4(\hat{x}) - \hat{y}_3(\hat{x})\right) &= -\frac{1}{8k}\delta\left(\frac{1}{2k} + \hat{x}\right) - \frac{1}{8k}\delta\left(\frac{1}{2k} - \hat{x}\right) .
\end{aligned} \tag{B.29}$$

As a consistency check, one can compute the volumes

$$\begin{aligned}
\mathrm{Vol}(Y) &= \frac{\pi^4}{24}\int d\hat{x}\,\hat{\rho}(\hat{x}) = \frac{\pi^4}{16k}\,,\\
\mathrm{Vol}(\Sigma_{A_i}) &= \frac{\pi^3}{4}\int d\hat{x}\,\hat{\rho}(\hat{x})\left(\hat{y}_2(\hat{x}) - \hat{y}_1(\hat{x}) + \frac{1}{2}\right) = \frac{\pi^3}{8k}\,,\\
\mathrm{Vol}(\Sigma_{B_i}) &= \frac{\pi^3}{4}\int d\hat{x}\,\hat{\rho}(\hat{x})\left(\hat{y}_3(\hat{x}) - \hat{y}_2(\hat{x}) + \frac{1}{2}\right) = \frac{\pi^3}{4k}\,,\\
\mathrm{Vol}(\Sigma_{C_i}) &= \frac{\pi^3}{4}\int d\hat{x}\,\hat{\rho}(\hat{x})\left(\hat{y}_4(\hat{x}) - \hat{y}_3(\hat{x}) + \frac{1}{2}\right) = \frac{\pi^3}{8k}\,,\\
\mathrm{Vol}(\Sigma_{D_i}) &= \frac{\pi^3}{4}\int d\hat{x}\,\hat{\rho}(\hat{x})\left(\hat{y}_1(\hat{x}) - \hat{y}_4(\hat{x}) + \frac{1}{2}\right) = \frac{\pi^3}{4k}\,.
\end{aligned} \tag{B.30}$$

Since $\mathrm{Vol}(Q^{2,2,2}) = \pi^4/16$ [94], we see that $\mathrm{Vol}(Y)$ matches that of a $\mathbb{Z}_k$ orbifold of $Q^{2,2,2}$. As for $M^{1,1,1}$, we can relate the volumes of the five-cycles in (B.30) to those computed in [94]. The cone over $Q^{2,2,2}$ is a $U(1)^2$ Kähler quotient of $\mathbb{C}^6$ with weights $(1,1,-1,-1,0,0)$ and $(1,1,0,0,-1,-1)$, together with a $\mathbb{Z}_2$ quotient that flips the sign of (a_1,a_2). If we denote the coordinates in $\mathbb{C}^6$ by $(a_1,a_2,b_1,b_2,c_1,c_2)$, we have [94]

$$\mathrm{Vol}(Q^{2,2,2}) = \frac{\pi^4}{16}\,, \qquad \mathrm{Vol}(\Sigma_{a_i}) = \mathrm{Vol}(\Sigma_{b_i}) = \mathrm{Vol}(\Sigma_{c_i}) = \frac{\pi^3}{8}\,. \tag{B.31}$$

One can think of the $\mathbb{Z}_k$ orbifold as acting on c_i with opposite phases, so it is natural to interpret A_i and C_i as correponding to a_i, B_i as corresponding to b_ic_1, and D_i as corresponding to b_ic_2. Indeed $k\,\mathrm{Vol}(\Sigma_{A_i}) = \mathrm{Vol}(\Sigma_{C_i}) = \mathrm{Vol}(\Sigma_{a_i})$, $k\,\mathrm{Vol}(\Sigma_{B_i}) = \mathrm{Vol}(\Sigma_{b_i}) + \mathrm{Vol}(\Sigma_{c_1})$, and $k\,\mathrm{Vol}(\Sigma_{D_i}) = \mathrm{Vol}(\Sigma_{b_i}) + \mathrm{Vol}(\Sigma_{c_2})$, the factor of k appearing because the volumes (B.30) are computed in a $\mathbb{Z}_k$ orbifold of $Q^{2,2,2}$.