Published by Institute of Physics Publishing for SISSA/ISAS
Received: *December 17, 2002*
Revised: *December 22, 2002*
Accepted: *December 24, 2002*

Chiral rings and anomalies in supersymmetric gauge theory

Freddy Cachazo,[a] Michael R. Douglas,[bc*] Nathan Seiberg[a] and Edward Witten[a]

[a] *School of Natural Sciences, Institute for Advanced Study Princeton NJ 08540 USA*
[b] *Department of Physics and Astronomy, Rutgers University Piscataway, NJ 08855-0849 USA*
[c] *I.H.E.S., Le Bois-Marie Bures-sur-Yvette, 91440 France*
E-mail: cachazo@ias.edu, mrd@physics.rutgers.edu, seiberg@ias.edu, witten@ias.edu

Abstract: Motivated by recent work of Dijkgraaf and Vafa, we study anomalies and the chiral ring structure in a supersymmetric $\mathrm{U}(N)$ gauge theory with an adjoint chiral superfield and an arbitrary superpotential. A certain generalization of the Konishi anomaly leads to an equation which is identical to the loop equation of a bosonic matrix model. This allows us to solve for the expectation values of the chiral operators as functions of a finite number of "integration constants." From this, we can derive the Dijkgraaf-Vafa relation of the effective superpotential to a matrix model. Some of our results are applicable to more general theories. For example, we determine the classical relations and quantum deformations of the chiral ring of $\mathcal{N}=1$ super Yang-Mills theory with $\mathrm{SU}(N)$ gauge group, showing, as one consequence, that all supersymmetric vacua of this theory have a nonzero chiral condensate.

Keywords: 1/N Expansion, Duality in Gauge Field Theories, Supersymmetry and Duality, Supersymmetric Effective Theories.

*Louis Michel Professor.

JHEP12(2002)071

Contents

1. Introduction

It is widely hoped that gauge theories with $\mathcal{N} = 1$ supersymmetry will be relevant for real world physics. Much work has been done on their dynamics. In the early and mid-1990's, holomorphy of the effective superpotential and gauge couplings were used, together with numerous other arguments, to obtain many nonperturbative results about supersymmetric dynamics. For a review, see e.g. [1]. Many such results were later obtained by a variety of constructions that depend on embedding the gauge theories in string theory. These provided systematic derivations of many results for extensive but special classes of theories. Of course such methods do not enable one to say anything about Kahler metrics, which are not governed by holomorphy.

Recently Dijkgraaf and Vafa [2] have made a striking conjecture, according to which the exact superpotential and gauge couplings for a wide class of $\mathcal{N}=1$ gauge theories can be obtained by doing perturbative computations in a closely related matrix model, in which the superpotential of the gauge theory is interpreted as an ordinary potential. Even more strikingly, these results are obtained entirely from the *planar* diagrams of this matrix model, even though no large-N limit is taken in the gauge theory. This conjecture was motivated by the earlier work [3]–[7] and a perturbative argument was given in [8].

A prototypical example for their results, which we will focus on, is the case of a $\mathrm{U}(N)$ gauge theory with $\mathcal{N}=1$ supersymmetry and a chiral superfield Φ in the adjoint representation of $\mathrm{U}(N)$. The superpotential is taken to be

$$W(\Phi)=\sum_{k=0}^{n}\frac{g_k}{k+1}\operatorname{Tr}\Phi^{k+1} \tag{1.1}$$

for some n.[1] If $W'(z)=g_n\prod_{i=1}^{n}(z-a_i)$, then by taking Φ to have eigenvalues a_i, with multiplicities N_i (which obey $\sum_i N_i=N$), one breaks $\mathrm{U}(N)$ to $G=\prod_i \mathrm{U}(N_i)$; we denote the gauge superfields of the $\mathrm{U}(N_i)$ gauge group as $W_{\alpha i}$. If the roots a_i of W' are distinct, as we will assume throughout this paper, then the chiral superfields are all massive and can be integrated out to get an effective lagrangian for the low energy gauge theory with gauge group G. This effective lagrangian is, as we explain in more detail in section 2, a function of $S_i=-\frac{1}{32\pi^2}\operatorname{Tr}W_{\alpha i}W^{\alpha i}$ and $w_{\alpha_i}=\frac{1}{4\pi}\operatorname{Tr}W_{\alpha i}$. Explicitly, the effective lagrangian has the form

$$L_{\mathrm{eff}}=\int d^2\theta\, W_{\mathrm{eff}}(S_i,w_{\alpha i},g_k)+\text{ complex conjugate }+\int d^4\theta(\ldots)\,, \tag{1.2}$$

where

$$W_{\mathrm{eff}}(S_i,w_{\alpha i},g_k)=\omega(S_i,g_k)+\sum_{l,m}t_{lm}(S_i,g_k)w_{\alpha l}w_m^{\alpha}\,. \tag{1.3}$$

In (1.2), the ellipses refer to irrelevant non-chiral interactions. Our goal, following [2], is to determine ω and t_{lm} (and to show that W_{eff} is quadratic in $w_{\alpha i}$, as has been assumed in writing (1.2)).

[1] This action is unrenormalizable if $n>3$, and hence to quantize it requires a cutoff. This is inessential for us, since the cutoff dependence does not affect the chiral quantities that we will study. Actually, by introducing additional massive superfields, one could obtain a renormalizable theory with an arbitrary effective superpotential (1.1). For example, a theory with two adjoint superfields Φ and Ψ and superpotential $W(\Phi,\Psi)=\operatorname{Tr}(m\Psi^2+\Psi\Phi^2)$ would be equivalent, after integrating out Ψ, to a theory with Φ only and a Φ^4 term in the superpotential.

JHEP12(2002)071

Since $\mathrm{U}(N_i) = \mathrm{SU}(N_i) \times \mathrm{U}(1)_i$, it is natural for many purposes (such as analyzing the infrared dynamics, where $\mathrm{SU}(N_i)$ confines and $\mathrm{U}(1)_i$ is weakly coupled) to separate the photons of $\mathrm{U}(1)_i \subset \mathrm{U}(N_i)$ from the gluons of $\mathrm{SU}(N_i) \subset \mathrm{U}(N_i)$. For this, we write

$$S_i = \widehat{S}_i - \frac{1}{2N} w_{\alpha i} w_i^\alpha \tag{1.4}$$

where $\widehat{S}_i = -\frac{1}{32\pi^2} \operatorname{Tr} \widehat{W}_{\alpha i}^2$ is constructed out of the $\mathrm{SU}(N_i)$ gauge fields $\widehat{W}_{\alpha i}$. The $\widehat{S}_i$ are believed to behave as elementary fields in the infrared, while the abelian fields $w_{\alpha i}$ are infrared-free. Although this separation is useful in understanding the dynamics, it is not very useful in computing $W_{\rm eff}$, mainly because $W_{\rm eff}$ is quadratic in $w_{\alpha i}$ if written in terms of S_i and $w_{\alpha i}$, while its dependence on $w_{\alpha i}$ is far more complicated if it is written in terms of $\widehat{S}_i$ and $w_{\alpha i}$.

As we have already explained, in $L_{\rm eff}$, S_i is a bilinear function, and $w_{\alpha i}$ a linear function, of the superspace field strengths $W_{\alpha i}$. So $L_{\rm eff}$ is a lagrangian for a gauge theory with gauge group $G = \prod_i \mathrm{U}(N_i)$. It is a nonrenormalizable lagrangian, and quantizing it requires a cutoff, but (as in the footnote above), this is inessential for us as it does not affect the chiral quantities. In this paper, our principal goal is to determine $L_{\rm eff}$ (modulo non-chiral quantities), not to quantize it or to study the $\prod_i \mathrm{U}(N_i)$ dynamics. In section 5, however, after completing the derivation of $L_{\rm eff}$, we will give some applications of it, which generally do involve gauge dynamics. The purpose of section 5 is to test and learn how to use $L_{\rm eff}$, once it has been derived.

Since it arises by integrating out only massive fields, $L_{\rm eff}$ is holomorphic in the S_i and $w_{\alpha i}$ near $S_i = 0$. A term in $L_{\rm eff}$ of given order in S_i and $w_{\alpha i}$ arises only from *perturbative* contributions in Φ with a certain number of loops. For example, an S^2 term can arise precisely in two-loop order. The number of loops for a given contribution was determined in [2], using results of [3], and will be explained in section 2. A principal result of [2] is that the perturbative contribution to $L_{\rm eff}$ with a given number of loops can be reproduced from a perturbative contribution, with the same number of loops, in an auxiliary matrix model described there. The auxiliary matrix model is non-supersymmetric and has for its ordinary potential the same function W that is the superpotential of the four-dimensional gauge theory. The original derivation of this result made use of string theory. Our goal is to provide a direct field theory derivation of the same result.

The basic technique is to compare the Konishi anomaly [9, 10] and generalizations of it to the equations of motion of the matrix model. In section 2, we describe some basic facts about the problem. We show that the general form of the effective action follows the structure of the chiral ring and from symmetry considerations. In the process, we show that the chiral ring has surprisingly tight properties. For example, all single-trace operators in the chiral ring are at most quadratic in W_α, and the operator $S = -\frac{1}{32\pi^2} \operatorname{Tr} W_\alpha W^\alpha$ obeys a classical relation $S^N = 0$, which is subject to quantum deformation. In the low energy pure $\mathrm{SU}(N)$ gauge theory, this relation, after quantum deformation, implies that all supersymmetric vacua of $\mathrm{SU}(N)$ supersymmetric gluodynamics have a chiral condensate. In sections 3 and 4, we present our derivation. We construct the generalized Konishi anomalies, which determine the quantum-corrected chiral ring, in section 3, and we compare

JHEP12(2002)071

to the matrix model in section 4. The computation depends upon a precise definition of the S_i given in section 2.7 in terms of gauge-invariant quantities of the underlying $\mathrm{U}(N)$ theory. In section 5, we discuss in detail some examples of applications of the results in some simple cases.

The result that we want to establish compares the effective superpotential of a four-dimensional gauge theory to a computation involving the "same" Feynman diagrams in a matrix model that is a (bosonic) truncation of the zero momentum sector of the four-dimensional gauge theory. So one might think that the starting point would be to compare the gauge theory to its zero momentum sector or bosonic reduction. If this strategy would work in a naive form, the effective superpotential of the four-dimensional gauge theory would be the same as the effective superpotential of the theory obtained from it by dimensional reduction to $n < 4$ dimensions. This is not so at all; as is clear both from the string-based Dijkgraaf-Vafa derivation and from the simple examples that we consider in section 2, the result that we are exploring is specifically four-dimensional.

Comparison with gauge dynamics. Though gauge dynamics is not the main goal of the present paper, it may help the reader if we summarize some of the conjectured and expected facts about the dynamics of the low energy effective gauge theory with gauge group $G = \prod_i \mathrm{U}(N_i)$ that arises if we try to quantize L_{eff}:

1. It is believed to have a mass gap and confinement, and as a result an effective description in terms of G-singlet fields.

2. For understanding the chiral vacuum states and the value of the superpotential in them, the important singlet fields are believed to be the chiral superfields $\widehat{S}_i = S_i + \frac{1}{2N_i} w_{\alpha i} w^{\alpha i})$ defined in (1.4).

3. Finally, it is believed that the relevant aspects of the dynamics of the $\widehat{S}_i$ can be understood by treating the effective action as a superpotential for elementary fields $\widehat{S}_i$ and $w_{\alpha i}$, accounting for gauge dynamics by adding to it the Veneziano-Yankielowicz superpotential [11] $\widetilde{W} = \sum_i N_i \widehat{S}_i (1 - \ln \widehat{S}_i)$, and extremizing the sum $W_{\text{eff}} = W_{\text{eff}} + \widetilde{W}$ with respect to the $\widehat{S}_i$ to determine the vacua. In particular, when this is done, the $\widehat{S}_i$ and therefore also S_i acquire vacuum expectation values, spontaneously breaking the chiral symmetries of the low energy effective gauge theory (and, depending on W, possibly spontaneously breaking some exact chiral symmetries of the underlying theory).

These statements are on a much deeper, and more difficult, level than the statements that we will explore in the present paper in computing the effective action. Our results are strictly perturbative and governed by anomalies; they are not powerful enough to imply results such as the mass gap and confinement in the gauge theory. The mass gap and confinement, in particular, are needed for the formulation in point 3, with the S_i treated as elementary fields, and other fields ignored, to make sense.

JHEP12(2002)071

It might clarify things to outline the only precise derivation of the Veneziano-Yankielowicz superpotential [11] that we know [12, 1], focusing for brevity on pure $\mathcal{N}=1$ gluodynamics with SU(N) gauge group. The tree level action of the gauge theory is

$$L_{\text{tree}} = \int d^4x d^2\theta \; 2\pi i \tau_{bare} S + \text{c.c.} \tag{1.5}$$

Logarithmic divergences of the one loop graphs force us to replace (1.5) with

$$L_{\text{tree}} = \int d^4x d^2\theta \; 3N \ln\left(\frac{\Lambda}{\Lambda_0}\right) S + \text{c.c.} \tag{1.6}$$

where Λ_0 is an ultraviolet cutoff and Λ is a finite scale which describes the theory. We can think of (1.6) as a microscopic action which describes the physics at the scale Λ_0. The long distance physics does not change when Λ_0 varies with fixed Λ. The factor of $3N$ comes from the coefficient of the one-loop beta function. Consider performing the complete path integral of the theory over the gauge fields. Nonperturbatively, a massive vacuum is generated, with a superpotential

$$W_{\text{eff}}(\Lambda) = N\Lambda^3 \,. \tag{1.7}$$

Since the well-defined instanton factor is Λ^{3N}, we prefer to write this as

$$W_{\text{eff}}(\Lambda) = N\left(\Lambda^{3N}\right)^{\frac{1}{N}} . \tag{1.8}$$

This superpotential controls the tension of BPS domain walls. The N^{th} root is related to chiral symmetry breaking. The N branches correspond to the N vacua associated with chiral symmetry breaking.

Now looking back at (1.6), we see that $3N\ln\Lambda$ couples linearly to S and thus behaves as a "source" for S. The superpotential $W_{\text{eff}}(\Lambda)$ is defined directly by the gauge theory path integral. If we want to compute an effective superpotential for S, the general recipe is to introduce S as a c-number field linearly coupled to the source $\ln\Lambda$, and perform a Legendre transform of $W_{\text{eff}}(\Lambda)$. A simple way to do that is to introduce an auxiliary field C and to consider the superpotential

$$W_{\text{eff}}(\Lambda, S, C) = NC^3 + S\ln\left(\frac{\Lambda^{3N}}{C^{3N}}\right) \tag{1.9}$$

Integrating out S using its equation of motion sets $C^{3N} = \Lambda^{3N}$ and leads to (1.8), showing that $W_{\text{eff}}(\Lambda, S, C)$ of (1.9) leads to results identical to those found from $W_{\text{eff}}(\Lambda)$ of (1.8). Integrating the auxiliary field C out of (1.9) using its equation of motion leads to the effective superpotential for S:

$$W_{\text{eff}}(S) = S\left[\ln\left(\frac{\Lambda^{3N}}{S^N}\right) + N\right] . \tag{1.10}$$

JHEP12(2002)071

This gives a clear-cut derivation and explanation of the meaning of the Veneziano-Yankielowicz superpotential [11], but to make the derivation one must already know about chiral symmetry breaking and the nonzero $W_{\rm eff}(\Lambda)$. One does not at present have a derivation in which one first computes the superpotential, proves that S can be treated as an elementary field, and then uses the superpotential to prove chiral symmetry breaking.

In this paper, we concentrate first on generating the function $W_{\rm eff}(S_i)$, understood (as we have explained above in detail) as an effective lagrangian for the low energy gauge fields. We do not (until section 5) analyze the gauge dynamics, give or assume expectation values for the S_i, treat the S_i or $\widehat{S}_i$ as elementary fields, or extremize $W_{\rm eff}(S_i)$ (or any superpotential containing it as a contribution) with respect to the S_i. These are not needed to *derive* the function $W_{\rm eff}(S_i)$. We stress that S_i cannot be introduced by a Legendre transform, as we did above for $S = \sum_i S_i$, because they do not couple to independent sources.

We thus do not claim to have an *a priori* argument that the gauge dynamics is governed by an effective lagrangian for elementary fields S_i. However, we will argue in section 4 that, if it is described by such a lagrangian, its effective superpotential must be the sum of $W_{\rm eff}(S_i)$ with a second contribution $\widetilde{W} = \sum_i N_i S_i(1 - \ln S_i)$, which (in the sense that was just described) is believed to summarize the effects of the low energy gauge dynamics. (In some cases it describes the effects of instantons of the underlying $\mathrm{U}(N)$ theory, as we explain in section 5.)

Finally, we note that Dijkgraaf and Vafa showed that this additional "gauge" part of the superpotential can be extracted from the measure of the matrix model. Regrettably, in this paper we cast little new light on this fascinating result.

2. Preliminary results

In this section, we will obtain a few important preliminary results, and give elementary arguments for why, as claimed by Dijkgraaf and Vafa, only planar diagrams contribute to the chiral effective action.

2.1 The chiral ring

Chiral operators are simply operators (such as $\mathrm{Tr}\,\Phi^k$ or S) that are annihilated by the supersymmetries $\overline{Q}_{\dot\alpha}$ of one chirality. The product of two chiral operators is also chiral. Chiral operators are usually considered modulo operators of the form $\{\overline{Q}_{\dot\alpha}, \ldots\}$. The equivalence classes can be multiplied, and form a ring called the chiral ring. A superfield whose lowest component is a chiral operator is called a chiral superfield.

Chiral operators are independent of position x, up to $\overline{Q}^{\dot\alpha}$-commutators. If $\{\overline{Q}_{\dot\alpha}, \mathcal{O}(x)\} = 0$, then

$$\frac{\partial}{\partial x^\mu}\mathcal{O}(x) = [P^\mu, \mathcal{O}(x)] = \{\overline{Q}^{\dot\alpha}, [Q^\alpha, \mathcal{O}(x)]\}. \tag{2.1}$$

This implies [13] that the expectation value of a product of chiral operators is independent of each of their positions:

$$\begin{aligned}\frac{\partial}{\partial x_1^{\alpha\dot\alpha}}\left\langle \mathcal{O}^{I_1}(x_1)\mathcal{O}^{I_2}(x_2)\cdots\right\rangle &= \left\langle \{\overline{Q}^{\dot\alpha},[Q^\alpha,\mathcal{O}^{I_1}(x_1)]\}\mathcal{O}^{I_2}(x_2)\cdots\right\rangle \\ &= -\sum_{k>1}\left\langle [Q^\alpha,\mathcal{O}^{I_1}(x_1)]\cdots\left[\overline{Q}^{\dot\alpha},\mathcal{O}^{I_k}(x_k)\right]\cdots\right\rangle \\ &= 0\,. \end{aligned}\tag{2.2}$$

Thus we can write $\langle\prod_I \mathcal{O}^I(x)\rangle = \langle\prod_I \mathcal{O}^I\rangle$ without specifying the positions x.

Using this invariance, we can take a correlation function of chiral operators at distinct points, and separate the points by an arbitrarily large distance. Cluster decomposition then implies that the correlation function factorizes [13]:

$$\langle \mathcal{O}^{I_1}(x_1)\mathcal{O}^{I_2}(x_2)\cdots\mathcal{O}^{I_n}(x_n)\rangle = \langle\mathcal{O}^{I_1}\rangle\langle\mathcal{O}^{I_2}\rangle\cdots\langle\mathcal{O}^{I_n}\rangle. \tag{2.3}$$

There are no contact terms in the expectation value of a product of chiral fields, because as we have just seen a correlation function such as $\langle\mathcal{O}^{I_1}(x_1)\mathcal{O}^{I_2}(x_2)\cdots\rangle$ is entirely independent of the positions x_i and so in particular does not have delta functions. A correlation function of chiral operators together with the upper component of a chiral superfield (for example $\langle \mathrm{Tr}\,\Phi^k\cdot\int d^2\theta\,\mathrm{Tr}\,\Phi^m\rangle$) can have contact terms.

In the theory considered here, with an adjoint superfield Φ, we can form gauge-invariant chiral superfields $\mathrm{Tr}\,\Phi^k$ for positive integer k. These are all non-trivial chiral fields. The gauge field strength W_α is likewise chiral, and though it is not gauge-invariant, it can be used to form gauge-invariant chiral superfields such as $\mathrm{Tr}\,\Phi^k W_\alpha$, $\mathrm{Tr}\,\Phi^k W_\alpha\Phi^l W_\beta$, etc. (Similar chiral fields involving W_α were important in the duality of $\mathrm{SO}(N)$ theories [14, 15].) Setting $k=l=0$, we get, in particular, chiral superfields constructed from vector multiplets only.

There is, however, a very simple fact that drastically simplifies the classification of chiral operators. If $\mathcal{O}$ is any adjoint-valued chiral superfield, we have

$$\left[\overline{Q}^{\dot\alpha}, D_{\alpha\dot\alpha}\mathcal{O}\right\} = [W_\alpha, O\}\ , \tag{2.4}$$

($D_{\alpha\dot\alpha}=\frac{D}{Dx^{\alpha\dot\alpha}}$ is the bosonic covariant derivative) using the Jacobi identity and definition of W_α plus the assumption that $\mathcal{O}$ (anti)commutes with $\overline{Q}^{\dot\alpha}$. Taking $\mathcal{O}=\Phi$, we see that in operators such as $\mathrm{Tr}\,\Phi^k W_\alpha\Phi^m W_\beta$, W_α commutes with Φ modulo $\{\overline{Q}_{\dot\alpha},\ldots\}$, so it suffices to consider only operators $\mathrm{Tr}\,\Phi^n W_\alpha W_\beta$.

Moreover, taking $\mathcal{O}=W_\beta$ in the same identity, we learn that

$$\{\overline{Q}^{\dot\alpha},[D_{\alpha\dot\alpha},W_\beta]\} = \{W_\alpha,W_\beta\}\,, \tag{2.5}$$

so in the chiral ring we can make the substitution $W_\alpha W_\beta\to -W_\beta W_\alpha$. So in any string of W's, say $W_{\alpha_1}\ldots W_{\alpha_s}$, we can assume antisymmetry in $\alpha_1,\ldots,\alpha_s$. As the α_i only take two values, we can assume $s\le 2$. So a complete list of independent single-trace chiral operators is $\mathrm{Tr}\,\Phi^k$, $\mathrm{Tr}\,\Phi^k W_\alpha$, and $\mathrm{Tr}\,\Phi^k W_\alpha W^\alpha$. This list of operators has already been studied by [16, 17].

2.2 Relations in the chiral ring

We have seen that the generators of the chiral ring are of the form $\mathrm{Tr}\,\Phi^k$, $\mathrm{Tr}\,W_\alpha \Phi^k$, $\mathrm{Tr}\,W_\alpha W^\alpha \Phi^k$. These operators are not completely independent and are subject to relations.

The first kind of relation stems from the fact that Φ is an $N \times N$ matrix. Therefore, $\mathrm{Tr}\,\Phi^k$ with $k > N$ can be expressed as a polynomial in $u_l = \mathrm{Tr}\,\Phi^l$ with $l \leq N$

$$\mathrm{Tr}\,\Phi^k = \mathcal{P}_k(u_1, \ldots, u_N) \tag{2.6}$$

In appendix A, we show that the classical relations (2.6) are modified by instantons for $k \geq 2N$. This is similar to familiar modifications due to instantons of classical relations in the two dimensional $\mathbf{CP}^N$ model [18] and in certain four dimensional $\mathcal{N} = 1$ gauge theories [19]. The general story is that to every classical relation corresponds a quantum relation, but the quantum relations may be different.

A second kind of relations follows from the tree level superpotential $W(\Phi)$. As is familiar in Wess-Zumino models, the equation of motion of Φ

$$\partial_\Phi W(\Phi) = \overline{D}_{\dot\alpha} \overline{D}^{\dot\alpha} \overline{\Phi} \tag{2.7}$$

shows that in the chiral ring $\partial_\Phi W(\Phi)_{c.r.} = 0$ (the subscript *c.r.* denotes the fact that this equation is true only in the chiral ring). In a gauge theory, we want to consider gauge-invariant chiral operators; classically, for any k, $\mathrm{Tr}\,\Phi^k \partial_\Phi W(\Phi)$ vanishes in the chiral ring. This is a nontrivial relation among the generators. In section 3, we will discuss in detail how this classical relation is modified by the Konishi anomaly [9, 10] and its generalizations.

We now turn to discuss interesting relations which are satisfied by the operator $S = -\frac{1}{32\pi^2}\,\mathrm{Tr}\,W_\alpha^2$. We first discuss the pure gauge $\mathcal{N} = 1$ theory with gauge group $\mathrm{SU}(N)$. We will comment about other gauge groups (such as $\mathrm{U}(N)$) below.

The operator S is subtle because it is a bosonic operator which is constructed out of fermionic operators. Since the gauge group has $N^2 - 1$ generators, the Lorentz index α in W_α take two different values, and S is bilinear in W_α, it follows from Fermi statistics that

$$(S^{N^2})_p = 0\,, \tag{2.8}$$

so in particular S is nilpotent. We added the subscript p to denote that this relation is valid in perturbation theory. Soon we will argue that this relation receives quantum corrections. It is important that (2.8) is true for any S which is constructed out of fermionic W_α. The latter does not have to satisfy the equations of motion.

If we are interested in the chiral ring, we can derive a more powerful result. We will show that

$$(S^N)_p = \{\overline{Q}_{\dot\alpha}, X^{\dot\alpha}\} \tag{2.9}$$

for some $X^{\dot\alpha}$, and therefore in the chiral ring

$$(S^N)_{p,c.r.} = 0 \tag{2.10}$$

JHEP12(2002)071

For SU(2), we can show (2.9) using the identity

$$\mathrm{Tr}\, ABCD = \frac{1}{2}\left(\mathrm{Tr}\, AB\,\mathrm{Tr}\, CD + \mathrm{Tr}\, DA\,\mathrm{Tr}\, BC - \mathrm{Tr}\, AC\,\mathrm{Tr}\, BD\right) \tag{2.11}$$

for any SU(2) generators A, B, C, D. Hence, allowing for Fermi statistics (which imply $\mathrm{Tr}\, W_1 W_1 = \mathrm{Tr}\, W_2 W_2 = 0$), we have

$$\mathrm{Tr}\, W_1 W_1 W_2 W_2 = \left(\mathrm{Tr}\, W_1 W_2\right)^2 . \tag{2.12}$$

The left hand side is non-chiral, as we have seen above, and the right hand side is a multiple of S^2, so we have established (2.10) for SU(2). In appendix B we extend this proof to SU(N) with any N, and we also show that $S^{N-1} \neq 0$ in the chiral ring.

If the relation $S^N = \{\overline{Q}_{\dot\alpha}, X^{\dot\alpha}\}$ were an exact quantum statement, it would follow that in any supersymmetric vacuum, $\langle S^N\rangle = 0$, and hence by factorization and cluster decomposition, also $\langle S\rangle = 0$. What kind of quantum corrections are possible in the ring relation $S^N = 0$? In perturbation theory, because of R-symmetry and dimensional analysis, there are no possible quantum corrections to this relation. Nonperturbatively, the instanton factor Λ^{3N} has the same chiral properties as S^N, so it is conceivable that instantons could modify the chiral ring relation to $S^N = \text{constant} \cdot \Lambda^{3N}$. In fact [13], instantons lead to an expectation value $\langle S^N\rangle = \Lambda^{3N}$, and therefore, they do indeed modify the classical operator relation to

$$S^N = \Lambda^{3N} + \{\overline{Q}_{\dot\alpha}, X^{\dot\alpha}\} \tag{2.13}$$

where in the chiral ring we can set the last term to zero. Equation (2.13) is an exact operator relation in the theory. It is true in all correlation functions with all operators. Also, since it is an operator equation, it is satisfied in all the vacua of the theory. The relation $S^{N^2} = 0$, which we recall is an exact relation, not just a statement in the chiral ring, must also receive instanton corrections so as to be compatible with (2.13). To be consistent with the existence of a supersymmetric vacuum in which $\langle S^N\rangle = \Lambda^{3N}$, as well as with the classical limit in which $S^{N^2} = 0$, the corrected equation must be of the form $(S^N - \Lambda^{3N})P(S^N, \Lambda^{3N}) = 0$, where P is a homogeneous polynomial of degree $N-1$ with a non-zero coefficient of $(S^N)^{N-1}$. We do not know the precise form of P.

To illustrate the power of the chiral ring relation $S^N = \Lambda^{3N}$, let us note that it implies that $\langle S\rangle^N = \langle S^N\rangle = \Lambda^{3N}$ in *all* supersymmetric vacua of the theory. Indeed, the chiral ring relation $S^N = \Lambda^{3N}$ is an exact operator relation, independent of the particular state considered. Thus, contrary to some conjectures, there does not exist a supersymmetric vacuum of the SU(N) supersymmetric gauge theory with $\langle S\rangle = 0$.

This situation is very similar to the analogous situation in the the two dimensional supersymmetric $\mathbf{CP}^{N-1}$ model as well as its generalization to Grassmannians (see [18, section 3.2]). In the $\mathbf{CP}^{N-1}$ model, a twisted chiral superfield, often called σ, obeys a classical relation $\sigma^N = 0$; this is deformed by instantons to a quantum relation $\sigma^N = e^{-I}$, where I is the instanton action. (In the *non-linear* $\mathbf{CP}^{N-1}$ model, σ is bilinear in fermions and associated with the generator of $H^2(\mathbf{CP}^{N-1})$; the classical relation $\sigma^N = 0$ then follows from Fermi statistics, as the fermions only have $2N-2$ components, and so is analogous to the classical relation $S^{N^2} = 0$ considered above.)

Role of these relations in different approaches. It is common that an operator relation which is "always true" in one formulation of a theory, being a kinematical relation that holds off-shell, follows from the equation of motion in a second formulation. In that second formulation, the given relation is not true off-shell. A familiar example is the Bianchi identity in two dual descriptions of free electrodynamics. The Bianchi identity is always true in the electric description of the theory, but it appears as an equation of motion in the magnetic description.

This analogy leads to a simple interpretation of the Veneziano-Yankielowicz superpotential $S(N + \ln(\Lambda^{3N}/S^N))$. It is valid in a description (difficult to establish rigorously) in which S is an unconstrained bosonic field. In that description, we should not use the classical equation $S^N = \{\overline{Q}_{\dot\alpha}, X^{\dot\alpha}\}$ or its quantum deformation $S^N = \Lambda^{3N} + \{\overline{Q}_{\dot\alpha}, X^{\dot\alpha}\}$ to simplify the lagrangian; such an equation should arise from the equation of motion. Indeed, the equation for the stationary points of this superpotential is $S^N = \Lambda^{3N}$ (or zero in perturbation theory) which is the relation in the chiral ring (2.13).

From here through section 4 of the paper, we will determine an effective action for the supersymmetric gauge theory with adjoint superfield Φ, understood as an action for low energy gauge fields that enter via S_i. Each term $\prod_i S_i^{k_i}$ that we generate has a clear meaning for large enough N_i. However, for fixed N_i and large enough k_i could we not simplify this lagrangian, using for example the $\mathrm{U}(N_i)$ relation $S^{N_i^2+1} = 0$? An attempt to do this would run into potentially complicated instanton corrections, because actually the S_i fields in the interactions $\prod_i S_i^{N_i}$ are smeared slightly by the propagators of the massive Φ fields; in attempting to simplify this to a local lagrangian (where one could use Fermi statistics to set $S^{N_i^2+1} = 0$), one would run into corrections from small instantons, so the attempt to simplify the lagrangian could not be separated from an analysis of gauge dynamics. Alternatively, in section 5, we use a more powerful (but less rigorous) approach in which it is assumed that the $\widehat{S}_i$ can be treated as independent classical fields. From this point of view, the ring relations for the S_i (which are written below in terms of $\widehat{S}_i$ and $w_{\alpha i}$) are not valid off-shell and should arise from the equations of motion.

Behavior for other groups. What happens for other groups? Consider an $\mathcal{N} = 1$ theory with a simple gauge group G and no chiral multiplets. The theory has a global discrete chiral symmetry $\mathbb{Z}_{2h(G)}$ with $h(G)$ the dual Coxeter number of G ($h = N$ for $\mathrm{SU}(N)$). Numerous arguments suggest that the $\mathbb{Z}_{2h}$ discrete symmetry of the system is spontaneously broken to $\mathbb{Z}_2$, and that the theory has h vacua in which $\langle S^h \rangle = c(G)\Lambda^{3h}$ with a nonzero constant $c(G)$. We conjecture that at the classical level, S obeys a relation $S^h = \{\overline{Q}_{\dot\alpha}, X^{\dot\alpha}\}$, and that instantons deform this relation to an exact operator statement

$$S^h = c(G)\Lambda^{3h} + \{\overline{Q}_{\dot\alpha}, X^{\dot\alpha}\} \tag{2.14}$$

for some $X^{\dot\alpha}$. As for $\mathrm{SU}(N)$, this statement would imply that chiral symmetry breaking occurs in all supersymmetric vacua. It would be interesting to understand how the chiral ring relation $S^h = 0$ arises at the classical level, and in particular how $h(G)$ enters.

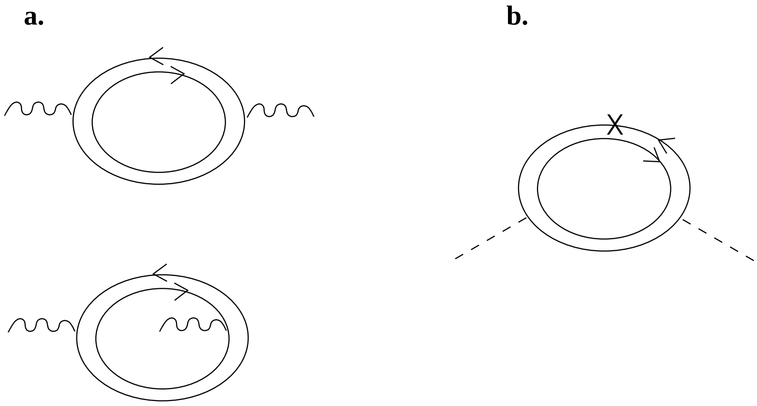

Figure 1: (a) A one-loop diagram, drawn in double line notation, with two gluons (wavy lines) inserted on the same or opposite index loops. (b) A one-loop diagram for computing $\langle \mathrm{Tr}\, \Phi^2 \rangle$ in an external field. The "$\times$" represents the operator and the dotted lines represent two external gluinos, corresponding to insertions of W_α.

Before ending this discussion, we will make a few comments about the $\mathrm{U}(N)$ gauge theory, as opposed to $\mathrm{SU}(N)$. As we explained in the introduction, here it makes sense to define (1.4) $S = \widehat{S} - \frac{1}{2N} w_\alpha w^\alpha$ where $\widehat{S}$ is the $\mathrm{SU}(N)$ part of S, and the second term is the contribution of the gauge fields of the $\mathrm{U}(1)$ part $w_\alpha = \frac{1}{4\pi} \mathrm{Tr}\, W_\alpha$. The quantum relation is $\widehat{S}^N = \Lambda^{3N} + \{\overline{Q}_{\dot\alpha}, X^{\dot\alpha}\}$. Since there are only two w_α and they are fermionic, S obeys

$$S^N = -\frac{1}{2} \widehat{S}^{N-1} w_\alpha w^\alpha + \Lambda^{3N} + \{\overline{Q}_{\dot\alpha}, X^{\dot\alpha}\} = -\frac{1}{2} S^{N-1} w_\alpha w^\alpha + \Lambda^{3N} + \{\overline{Q}_{\dot\alpha}, X^{\dot\alpha}\} \qquad (2.15)$$

In vacua where $\mathrm{U}(N)$ is broken to $\prod_i \mathrm{U}(N_i)$, each with its own $S_i = \widehat{S}_i - \frac{1}{2N_i} w_{i\alpha} w_i^\alpha$ these operators satisfy complicated relations. They follow from the equation of motion of the effective superpotential $\omega(S_i)$. What happens for $N_i = 1$? In this case, one might expect $\widehat{S}_i = 0$. This is the correct answer classically but it can be modified quantum mechanically to $\widehat{S}_i =$ constant. We will see that in more detail in section 5, where we study the case with all $N_i = 1$. In this case instantons lead to $\widehat{S}_i \neq 0$. In order to describe this case in the effective theory, one must include an independent field $\widehat{S}_i$ even for a $\mathrm{U}(1)$ factor.

2.3 First look at perturbation theory — unbroken gauge group

Now we begin our study of the theory with the adjoint chiral superfield Φ. For simplicity, we will begin with the case of unbroken gauge symmetry. So we expand around $\Phi = a$, where a is a (c-number) critical point of the function W that appears in the superpotential

$$W(\Phi) = \sum_{k=0}^{n} \frac{g_k}{k+1} \mathrm{Tr}\; \Phi^{k+1} . \qquad (2.16)$$

The mass parameter of the Φ field in expanding around this vacuum is $m = W''(a)$. We may as well assume that $a = 0$, so $g_0 = 0$ and $g_1 = m$.

Since our basic interest is in generating an effective action for the gauge fields by integrating out Φ in perturbation theory, we might begin with the one-loop diagram of figure 1(a). We have drawn this diagram in 't Hooft's double line notation, exhibiting the two index lines of the Φ propagator. As is familiar, the one-loop diagram with two external gauge fields renormalizes the gauge kinetic energy. The result can be described (modulo non-chiral operators of higher dimension) by the effective action

$$\int d^4x d^2\theta \, \ln\left(\frac{m}{\Lambda_0}\right)\left(NS + \frac{1}{2}w_\alpha w^\alpha\right) + \text{c.c.}\,, \tag{2.17}$$

where the two contributions arise from contributions where the two gluons are inserted on the same index loop or on opposite index loops. In the former case, one index trace gives a factor of N and the second a factor of $S = -\frac{1}{32\pi^2}\operatorname{Tr} W_\alpha W^\alpha$; in the second case, each gives a factor of $w_\alpha = \frac{1}{4\pi}\operatorname{Tr} W_\alpha$. (2.17) describes the standard one-loop coupling constant renormalization by the massive Φ field, with the usual logarithmic dependence on the mass. In (2.17), Λ_0 is an ultraviolet cutoff. The one-loop diagram is the only contribution to the chiral couplings that depends on Λ_0; this follows from holomorphy and supersymmetry, or alternatively from the fact that as we will see (following [2]) the multi-loop diagrams make contributions that involve higher powers of S and hence are ultraviolet convergent.

One thing that one observes immediately is that the above result is special to four dimensions. If one makes a dimensional reduction to $n < 4$ dimensions before evaluating the Feynman diagram, dimensional analysis will forbid a contribution as in (2.17), and a different result will ensue, depending on the dimension. Moreover, the m dependence in (2.16) is controlled in a familiar way by the chiral anomaly under phase rotations of the Φ field; this is a hint that also the higher order terms can be determined by using anomalies, as we will do in this paper.

One can phrase even the one loop results in a way in which the cutoff does not appear, by instead computing a first derivative of the effective superpotential with respect to the couplings. On general grounds, these first derivatives are the expectation values of the gauge invariant chiral operators. We will need this result in section 4, so we will explain it in detail. A variation of the couplings $g_k \to g_k + \delta g_k$ will produce an effective superpotential $W_{\text{eff}}(S_i, w_{\alpha i}, g_k + \delta g_k)$ given by

$$\left\langle \exp\left(-\int d^4x d^2\theta \, \sum_k \frac{\delta g_k}{k+1}\operatorname{Tr}\Phi^{k+1} - \text{c.c.}\right)\right\rangle_\Phi\,, \tag{2.18}$$

where the bracket $\langle\ \ \rangle_\Phi$ refers to the result of a path integral over the massive fields, in the presence of a (long wavelength) background gauge field characterized by variables S_i and $w_{\alpha i}$. By holomorphy and supersymmetry, this result remains valid if the coupling constants $g_k + \delta g_k$ in (2.16) and (2.18) are promoted to chiral superfields [20]. Now, take the variational derivative of (2.18) with respect to the *upper* component of δg_k. This eliminates the $d^2\theta$ integral, and produces

$$\frac{\partial W_{\text{eff}}}{\partial g_k} = \left\langle \operatorname{Tr}\frac{\Phi^{k+1}}{k+1}\right\rangle_\Phi\,. \tag{2.19}$$

Thus, if we can get an expectation value as a function of the couplings, we can integrate it to get W_{eff}, up to a constant of integration independent of the couplings.

To illustrate, let us compute the right hand side of (2.19) in the one-loop approximation for $k = 1$. The diagram that we need to evaluate is shown in figure 1(b). There are two boson propagators and one fermion propagator. The integral that we must evaluate is

$$\int \frac{d^4k}{(2\pi)^4} \left(\frac{1}{k^2 + m\overline{m}}\right)^2 \frac{\overline{m}}{k^2 + m\overline{m}} = \frac{1}{32\pi^2 m}\,. \tag{2.20}$$

This multiplies NS or $w_\alpha w^\alpha$ depending on whether the two external gluinos are inserted on equal or opposite index loops, so in the one-loop approximation, we get

$$\frac{\partial W_{\rm eff}}{\partial g_1} = \frac{NS + \frac{1}{2} w_\alpha w^\alpha}{m}\,. \tag{2.21}$$

(Recall that $g_1 = m$.)

This illustrates the process of integrating out the massive Φ fields to get a chiral function of the background gauge fields. To go farther, however, we want more general arguments.

2.4 A more detailed look at the effective action — unbroken gauge group

Next we will consider the multi-loop corrections. Let us first see what we can learn just from symmetries and dimensional analysis. The one adjoint theory has two continuous symmetries, a standard $\mathrm{U}(1)_R$ symmetry and a symmetry $\mathrm{U}(1)_\Phi$ under which the entire superfield Φ undergoes

$$\Phi \to e^{i\alpha}\Phi\,. \tag{2.22}$$

We also introduce a linear combination of these, $\mathrm{U}(1)_\theta$, which is convenient in certain arguments.

If we allow the couplings to transform non-trivially, these are also symmetries of the theory with a superpotential. The dimensions Δ and charges Q of the chiral fields and couplings are

$$\begin{array}{ccccc} & \Delta & Q_\Phi & Q_R & Q_\theta \\ \Phi & 1 & 1 & 2/3 & 0 \\ W_\alpha & 3/2 & 0 & 1 & 1 \\ g_l & 2-l & -(l+1) & \frac{2}{3}(2-l) & 2 \\ \Lambda^{2N} & 2N & 2N & 4N/3 & 0 \end{array} \tag{2.23}$$

(Although we will not use it here, we also included the gauge theory scale Λ for later reference.) Since all of these are chiral superfields, their dimensions satisfy the relation $\Delta = 3Q_R/2$.

All of these symmetries are anomalous at one loop. The $\mathrm{U}(1)_\Phi$ anomaly was first discussed by Konishi [9, 10] and (as we discuss in more detail in section 3) says that the one-loop contribution to $W_{\rm eff}(S)$ is shifted under (2.22) by a multiple of S; this can be used to recover the result already given in (2.17). On the other hand, higher loop computations are finite and the transformation (2.22) leaves these invariant (again, we discuss this point in detail in section 3). It follows that the effective action must depend on the g_k's only

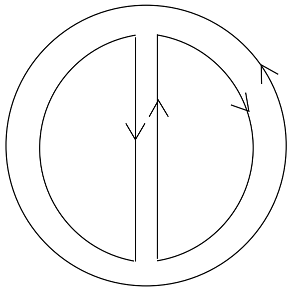

Figure 2: A planar diagram with two vertices, each of $k = 2$, three index loops — $L = 3$ — and two ordinary loops — $L' = 2$.

through the ratios $g_k/g_1^{(k+1)/2}$. Equivalently, we can write

$$\sum_k (k+1) g_k \frac{\partial}{\partial g_k} W_{\rm eff} = 0\,, \tag{2.24}$$

except at one loop.

Now let us consider dimensional analysis (by the relation $\Delta = 3Q_R/2$, R symmetry leads to the same constraint). The effective superpotential $W_{\rm eff}$ has dimension 3. So, taking into account the result in the last paragraph, $W_{\rm eff}$ must depend on the g_k's and W_α in the form[2]

$$W_{\rm eff} = W_\alpha^2 F\left(\frac{g_k W_\alpha^{k-1}}{g_1^{(k+1)/2}}\right). \tag{2.25}$$

Here we are being rather schematic, just to indicate the overall power of W_α. So, for example, W_α^2 may correspond to either S or $w_\alpha w^\alpha$. Instead of (2.25), we could write

$$\left(\sum_k (2-k) g_k \frac{\partial}{\partial g_k} + \frac{3}{2} W_\alpha \frac{\partial}{\partial W_\alpha}\right) W_{\rm eff} = 3W_{\rm eff}\,. \tag{2.26}$$

If a planar Feynman diagram has vertices of degree k_i+1, $i = 1, \ldots, s$, then the number of index loops is

$$L = 2 + \frac{1}{2}\sum_i (k_i - 1)\,. \tag{2.27}$$

This is easily proved by induction. The planar diagram sketched in figure 2 has two vertices each with $k_i = 2$; the number of index loops is 3, in accord with (2.27). Now each time we add a propagator to a planar diagram, we add one index loop and increase $k_i - 1$ by one at each end of the propagator; thus, inductively, (2.27) remains valid when a propagator is added or removed.

By taking a linear combination of (2.24) and (2.26), we get

$$\left(\sum_k (1-k) g_k \frac{\partial}{\partial g_k} + W_\alpha \frac{\partial}{\partial W_\alpha}\right) W_{\rm eff} = 2W_{\rm eff}\,. \tag{2.28}$$

[2]The one-loop diagram is again an exception, because its contribution depends on the ultraviolet cutoff Λ_0, as we see in (2.17), and this modifies the dimensional analysis, so again more care is needed to study the one-loop contribution in this way. By supersymmetry and holomorphy, Λ_0 will not enter in the higher order contributions.

(2.27) lets us replace $\sum_k(1-k)g_k\partial/\partial g_k$ by $4-2L$. It follows therefore that the power of W_α in a contribution coming from a planar Feynman diagram with L index loops is $2L-2$. Such a contribution might be a multiple of S^{L-1}, or of $S^{L-2}w_\alpha w^\alpha$. We get S^{L-1} if $L-1$ index loops each have two insertions of external gauge bosons and one has none, or $S^{L-2}w_\alpha w^\alpha$ if $L-2$ index loops each have two gauge insertions and two loops have one gauge insertion each.

The generalization of (2.27) for a diagram that is not planar but has genus g (that is, the Riemann surface made by filling in every index loop with a disc has genus g) is

$$L = 2 - 2g + \frac{1}{2}\sum_i (k_i - 1)\,. \tag{2.29}$$

This statement reflects 't Hooft's observation that increasing the genus by one removes two index loops and so multiplies an amplitude by $1/N^2$. So a diagram of genus g has a power of W_α equal to $2L+4g-2$.

At first sight, it appears that additional invariants such as Tr $W_\alpha W^\alpha W_\beta W^\beta$ will arise if there are more then two gauge insertions on the same index loop. But as we have seen, such operators are non-chiral and so can be neglected for the purposes of this paper. Also, the bilinear expression Tr $W_\alpha W_\beta$ vanishes by Fermi statistics unless the indices α and β are contracted to a Lorentz singlet. Therefore the effective superpotential can be expressed in terms of $S = -\frac{1}{32\pi^2}\,\mathrm{Tr}\,W_\alpha W^\alpha$ and $w_\alpha = \frac{1}{4\pi}$ Tr W_α only.

Since each index loop in a Feynman diagram can only contribute a factor of S, w_α, or N, depending on whether the number of gauge insertions on the loop is two, one, or zero, it now follows, as first argued by Dijkgraaf and Vafa, that only planar diagrams can contribute to the effective chiral interaction for the gauge fields. Indeed, according to (2.29), in the case of a non-planar diagram, at least one index loop would have to give a trace of a product of more than two W_α's, but these are trivial as elements of the chiral ring.

Furthermore, the interactions generated by planar diagrams are highly constrained. The only way to obey (2.29) without generating a trace with more than two W's is to have one index loop with no W and the others with two each, or two index loops with one W and again the others with two each. The possible interactions generated by a planar diagram with L index loops are hence S^{L-1} and $S^{L-2}w_\alpha w^\alpha$.

This reproduces the results of [8] and the previous papers cited in the introduction.

2.5 The case of broken gauge symmetry

Now let us briefly consider how to extend these arguments for the case of spontaneously broken gauge symmetry. Here we assume that $\mathrm{U}(N)$ is spontaneously broken to $\prod_{i=1}^n \mathrm{U}(N_i)$. We integrate out all massive fields, including the massive gauge bosons and their superpartners, to obtain an effective action for the massless superfields of the unbroken gauge group. The effective action depends now on the assortment of gauge-invariant operators $S_i = -\frac{1}{32\pi^2}\,\mathrm{Tr}\,W_{\alpha i}W^{\alpha i}$ and $w_{\alpha i} = \frac{1}{4\pi}\,\mathrm{Tr}\,W_{\alpha i}$ that appeared in the introduction. Their definitions will be considered more critically in section 2.7, but for now we can be casual.

Unlike the simpler case of unbroken gauge symmetry, here we need to integrate out massive gauge fields. Their propagator involves the complexified gauge coupling τ_{bare} or $\ln\Lambda$. But the perturbative answers must be independent of its real part, which is the theta angle, and therefore τ_{bare} cannot appear in holomorphic quantities (except for the trivial classical term). This means that the number of massive gauge field propagators must be equal to the number of vertices of three gauge fields.

The considerations of symmetries and dimensional analysis that led to (2.25) are still valid.[3] However, the W_α's that appear in (2.25) can now be of different kinds; they can be of the type $W_{\alpha i}$ for any $1 \le i \le n$. They thus can assemble themselves into different invariants such as S_i and $w_{\alpha i}$, but for the moment we will not be specific about this.

A key difference when the gauge symmetry is spontaneously broken is that the dependence on the g_k can be much more complicated. The g_k do not merely enter as factors at interaction vertices. The reason is that perturbation theory is constructed by expanding around a critical point of $W(\Phi)$ in which the matrix Φ has eigenvalues a_i (the zeroes of W') with multiplicities N_i. The a_i depend on the g_k, as do the masses of various components of Φ; and after shifting Φ to the critical point about we intend to expand, a Feynman vertex of order r is no longer merely proportional to g_{r-1} but receives contributions from g_s with $s > r-1$. Because of all this, we need a better way to determine the relation between the number of loops and the S-dependence.

Let L' be the number of ordinary loops in a Feynman diagram, as opposed to L, the number of index loops. Thus, in the diagram of figure 2, $L' = 2$ and $L = 3$. For planar diagrams, $L' = L - 1$, and in general,

$$L' = L - 1 + 2g\,. \tag{2.30}$$

In order to enumerate the loops we modify the tree level lagrangian as follows. We replace the superpotential W with $W/\hbar$, we replace the kinetic term $\mathrm{Tr}\,\Phi^\dagger e^V \Phi$ with $\frac{Z}{\hbar}\,\mathrm{Tr}\,\Phi^\dagger e^V \Phi$, and we replace the gauge kinetic term $\tau_{bare} W_\alpha^2$ with $\frac{1}{\hbar}\tau_{bare} W_\alpha^2$. Here Z is a real vector superfield, τ is a chiral superfield and $\hbar$ is taken to be a real number. The contribution $W_{\rm eff}^{L'}$ to $W_{\rm eff}$ from diagrams with L' loops is then proportional to $\hbar^{L'-1}$, so

$$\hbar\frac{\partial}{\partial\hbar} W_{\rm eff}^{L'} = (L'-1) W_{\rm eff}^{L'}\,. \tag{2.31}$$

Since Z is a real superfield, it cannot appear in the effective superpotential. We have already remarked that the perturbative superpotential is independent of τ_{bare} (except the trivial tree dependence). Therefore, simply by calculus, since $W/\hbar$ depends only on ratios $g_k/\hbar$, we have

$$\hbar\frac{\partial}{\partial\hbar} W_{\rm eff}^{L'} = -\sum_k g_k \frac{\partial}{\partial g_k} W_{\rm eff}^{L'} \tag{2.32}$$

[3]The one-loop diagram still provides an exception to (2.25) because it depends logarithmically on the cutoff Λ_0; it can be studied specially and gives, in the usual way, a linear combination of the S_i and of $w_{\alpha i} w_j^\alpha$. This is of the general form that the discussion below would suggest, though this discussion will be formulated in a way that does not quite apply at one-loop order.

JHEP12(2002)071

so

$$\sum_k g_k \frac{\partial}{\partial g_k} W_{\text{eff}}^{L'} = -(L'-1) W_{\text{eff}}^{L'} \,. \tag{2.33}$$

Using (2.30), the contribution $W_{\text{eff}}^{L,g}$ with L index loops and genus g hence obeys

$$\hbar \frac{\partial}{\partial \hbar} W_{\text{eff}}^{L,g} = -(L-2+2g) W_{\text{eff}}^{L,g} \,. \tag{2.34}$$

By combining (2.34) with (2.24) and (2.26), we learn that

$$\sum_{\alpha,i} W_{\alpha i} \frac{\partial}{\partial W_{\alpha i}} W_{\text{eff}}^{L,g} = (2L-2+4g) W_{\text{eff}}^{L,g} \,, \tag{2.35}$$

so a diagram with a given L and g makes contributions to the superpotential that is of order $2L-2+4g$ in the $W_{\alpha i}$'s. This is the same result as in the case of unbroken symmetry; we have merely given a more general derivation using (2.34) instead of (2.29). It follows from (2.35) that for $g > 0$, the overall power of $W_{\alpha i}$ is greater than $2L$, so at least one index loop contains more than two gauge insertions and gives a trace of more than two powers of $W_{\alpha i}$ for some i.

Generalizing the discussion of the unbroken group, each index loop in a Feynman diagram labeled by one of the unbroken gauge groups $\mathrm{U}(N_i)$ can only contribute a factor of S_i, $w_{\alpha i}$, or N_i, depending on whether the number of gauge insertions on the loop is two, one, or zero. It again follows that only planar diagrams can contribute to the effective chiral interaction for the gauge fields. As in the unbroken case, we learn that the possible interactions generated by a planar diagram with L index loops are S^{L-1} and $S^{L-2} w_\alpha w^\alpha$, where here each factor of S or w_α can be any of the S_i or $w_{\alpha i}$, depending on how the index loops in the diagram are labeled. In particular, the chiral effective action is quadratic in the $w_{\alpha i}$.

2.6 General form of the effective action

Making use of the last statement to constrain the dependence on the w's, the general form of the chiral effective action for the low energy gauge fields is thus

$$\int d^4x \, d^2\theta \; W_{\text{eff}}(S_k, w_{\alpha k}) \tag{2.36}$$

with

$$W_{\text{eff}}(S_k, w_{\alpha k}) = \omega(S_k) + \sum_{i,j=1}^{n} t_{ij}(S_k) w_{\alpha i} w_j^\alpha \,. \tag{2.37}$$

We now make a simple observation that drastically constrains the form of the effective action. Since all fields in the underlying $\mathrm{U}(N)$ gauge theory under discussion here are in the adjoint representation, the $\mathrm{U}(1)$ factor of $\mathrm{U}(N)$ is free. It is described completely by the simple W_α^2 classical action.

In particular, there is an exact symmetry of shifting W_α by an anticommuting c-number (times the identity $N \times N$ matrix). In the low energy theory with gauge group $\prod_i \mathrm{U}(N_i)$, this symmetry simply shifts each $W_{\alpha i}$. So the transformation of $S_i = -\frac{1}{32\pi^2} \operatorname{Tr} W_{\alpha i} W_i^\alpha$ and $w_{\alpha i} = \frac{1}{4\pi} \operatorname{Tr} W_{\alpha i}$ under $W_\alpha \to W_\alpha - 4\pi\psi_\alpha$ is

$$\begin{aligned} \delta S_i &= -\psi^\alpha w_{\alpha i} \\ \delta w_{\alpha_i} &= -N_i \psi_\alpha \,. \end{aligned} \tag{2.38}$$

It is not difficult to work out by hand the conditions imposed by the invariance (2.38) on the effective action (2.37). However, it is more illuminating to combine the operators S_i and $w_{\alpha i}$ in a superfield $\mathcal{S}_i$ which will also play a very useful role in the sequel:

$$\mathcal{S}_i = -\frac{1}{2} \operatorname{Tr} \left(\frac{1}{4\pi} W_{\alpha_i} - \psi_\alpha \right) \left(\frac{1}{4\pi} W_i^\alpha - \psi^\alpha \right) = S_i + \psi_\alpha w_i^\alpha - \psi^1 \psi^2 N_i \,. \tag{2.39}$$

In this description, the symmetry is simply generated by $\partial/\partial\psi_\alpha$. Invariance under this transformation implies that the effective action is

$$W_{\text{eff}} = \int d^2\psi \mathcal{F}_p \tag{2.40}$$

for some function $\mathcal{F}_p$. $\mathcal{F}_p$ is not uniquely determined by (2.40), since, for example, we could add to $\mathcal{F}_p$ any function of the $\mathrm{SU}(N_i)$ fields $\widehat{S}_i$ of (1.4). However, the form of (2.36) implies that we can take $\mathcal{F}_p$ to be a function only of the $\mathcal{S}_i$, and $\mathcal{F}_p$ is then uniquely determined if we also require

$$\mathcal{F}_p(\mathcal{S}_i = 0) = 0 \,. \tag{2.41}$$

The subscript p in the definition of $\mathcal{F}_p$ denotes the fact that $\int d^2\psi \mathcal{F}_p$ reproduces the chiral interactions induced by integrating out the massive fields (without the nonperturbative contribution from low energy gauge dynamics). On doing the ψ integrals, we learn that

$$W_{\text{eff}} = \omega(S_k) + t_{ij} w_{\alpha i} w_j^\alpha = \int d^2\psi \mathcal{F}_p(\mathcal{S}_k) = \sum_i N_i \frac{\partial \mathcal{F}_p(S_k)}{\partial S_i} + \frac{1}{2} \sum_{ij} \frac{\partial^2 \mathcal{F}_p(S_k)}{\partial S_i \partial S_j} w_{\alpha i} w_j^\alpha \,. \tag{2.42}$$

This is the general structure claimed by Dijkgraaf and Vafa, with $\mathcal{F}_p$ still to be determined.

By expanding (2.42) in powers of $w_{\alpha i}$ (and recalling the expansion (1.4)), we derive their gauge coupling matrix

$$\tau_{ij} = \frac{\partial^2 \mathcal{F}_p(S_k)}{\partial S_i \partial S_j} - \delta_{ij} \frac{1}{N_i} \sum_l N_l \frac{\partial^2 \mathcal{F}_p(S_k)}{\partial S_i \partial S_l} \,. \tag{2.43}$$

It is easy to see that

$$\sum_j \tau_{ij} N_j = 0 \tag{2.44}$$

which signals the decoupling of the overall $\mathrm{U}(1)$. Note that it decouples without using the equations of motion. While the overall $\mathrm{U}(1)$ decouples from the perturbative corrections, it of course participates in the classical kinetic energy of the theory, which in the present variables is $\tau \sum_i S_i$, with τ the bare coupling parameter.

JHEP12(2002)071

In many string and brane constructions, as mentioned in [2], nonrenormalizable interactions are added to the underlying $\mathrm{U}(N)$ gauge theory in such a way that the symmetry generated by $\partial/\partial\psi$ actually becomes a second, spontaneously broken supersymmetry. This fact is certainly of great interest, but will not be exploited in the present paper. We only note that as the scale of that symmetry breaking is taken to infinity, the Goldstone fermion multiplet which is the U(1) gauge multiplet becomes free. In this limit the broken generators of the $\mathcal{N} = 2$ supersymmetry are contracted. By that we mean that while the anticommutator of two unbroken supersymmetry generators is a translation generator, two broken supersymmetry generators anticommute. Indeed, our fermionic generators $\partial/\partial\psi$ anticommute with the analogous anti-chiral symmetries $\partial/\partial\psi^\dagger$ that shift $W_{\dot\alpha}$ by a constant.

Despite this relation to spontaneously broken $\mathcal{N} = 2$ supersymmetry in certain constructions, the symmetry (2.38) is present even in theories which have no obvious $\mathcal{N} = 2$ analog. For example, it would be present in any $\mathcal{N} = 1$ quiver gauge theory (i.e. with $\mathrm{U}(N)$ gauge groups and bifundamental and adjoint matter), including chiral theories.

2.7 $\mathrm{U}(N)$-invariant description of effective operators of the low energy theory

As preparation for the next section, we now wish to give a more accurate description of the effective operators S_i and $w_{\alpha i}$, $i = 1, \ldots, n$ of the low energy theory in terms of $\mathrm{U}(N)$-invariant operators of the underlying $\mathrm{U}(N)$ theory. We assume as always that the superpotential W has n distinct critical points a_i, $i = 1, \ldots, n$.

Roughly speaking, we can do this by taking the S_i to be linear combinations of $\mathrm{Tr}\ \Phi^k W_\alpha W^\alpha$ (as we have seen, the precise ordering of factors does not matter), for $k = 0, \ldots, n-1$. Similarly, one would take the $w_{\alpha i}$ to be linear combinations of $\mathrm{Tr}\ \Phi^k W_\alpha$. However, to be precise, we need to take linear combinations of these operators to project out, for example, the S_i associated with a given critical point a_i. Proceeding in this way, one would get rather complicated formulas which might appear to be subject to quantum corrections.

There is a more illuminating approach which also will be extremely helpful in the next section. We let C_i be a small contour surrounding the critical point a_i in the counterclockwise direction, and not enclosing any other critical points. And we define

$$\begin{aligned} S_i &= -\frac{1}{2\pi i}\oint_{C_i} dz \frac{1}{32\pi^2}\,\mathrm{Tr}\, W_\alpha W^\alpha \frac{1}{z-\Phi} \\ w_{\alpha_i} &= \frac{1}{2\pi i}\oint_{C_i} dz \frac{1}{4\pi}\,\mathrm{Tr}\, W_\alpha \frac{1}{z-\Phi}\,. \end{aligned} \tag{2.45}$$

It is easy to see, first of all, that in the semiclassical limit, the definitions in (2.45) give what we want. In the classical limit, to evaluate the integral, we set Φ to its vacuum value. In the vacuum, Φ is a diagonal matrix with diagonal entries $\lambda_1, \lambda_2, \ldots, \lambda_N$ (which are equal to the a_i with multiplicity N_i). If M is any matrix and M_{xy}, $x, y = 1, \ldots, N$ are the matrix elements of M in this basis, then

$$\mathrm{Tr}\, M \frac{1}{z-\Phi} = \sum_x \frac{M_{xx}}{z-\lambda_x}\,, \tag{2.46}$$

so

$$\frac{1}{2\pi i}\oint_{C_i} dz\,\mathrm{Tr}\,M\frac{1}{z-\Phi}=\sum_x\frac{1}{2\pi i}\oint_{C_i} dz\frac{M_{xx}}{z-\lambda_x}=\sum_{\lambda_x\in C_i}M_{xx}=\mathrm{Tr}\;MP_i\,. \tag{2.47}$$

Here $\lambda_x \in C_i$ means that λ_x is inside the contour C_i, and P_i is the projector onto eigenspaces of Φ corresponding to eigenvalues that are inside this contour. In the classical limit, P_i is just the projector onto the subspace in which $\Phi = a_i$. Hence the above definitions amount in the classical limit to

$$S_i=-\frac{1}{32\pi^2}\,\mathrm{Tr}\;W_\alpha W^\alpha P_i,\quad w_{\alpha i}=\frac{1}{4\pi}\,\mathrm{Tr}\,W_\alpha P_i\,. \tag{2.48}$$

This agrees with what we want.

The formula (2.47) is still valid in perturbation theory around the classical limit, except that the projection matrix P_i might undergo perturbative quantum fluctuations. One might expect that in perturbation theory, P_i in (2.48) could be expressed as its classical value plus loop corrections that would themselves be functions of the vector multiplets of the low energy theory. In this case, the objects S_i and $w_{\alpha i}$ as defined in (2.45) would not be precisely the usual functions of the massless vector multiplets of the low energy theory. We will now argue, however, that there are no such corrections. First we complete S_i and $w_{\alpha i}$ into the natural superfield

$$\mathcal{S}_i=-\frac{1}{2\pi i}\oint dz\frac{1}{2}\,\mathrm{Tr}\left(\frac{1}{4\pi}W_\alpha-\psi_\alpha\right)\left(\frac{1}{4\pi}W^\alpha-\psi^\alpha\right)\frac{1}{z-\Phi}\,. \tag{2.49}$$

So

$$\mathcal{S}_i=S_i+\psi_\alpha w_i^\alpha-\psi^1\psi^2N_i'\,, \tag{2.50}$$

where

$$N_i'=\frac{1}{2\pi i}\oint_{C_i} dz\;\mathrm{Tr}\,\frac{1}{z-\Phi}\,. \tag{2.51}$$

A priori, it might seem that N_i' is another chiral superfield, but actually, in perturbation theory N_i' is just the constant N_i. In fact, using (2.47), $N_i' = \mathrm{Tr}\,P_i$. Although the projection matrix P_i can fluctuate in perturbation theory, the dimension of the space onto which it projects *cannot* fluctuate, since perturbation theory only moves eigenvalues by a bounded amount.[4] Thus this dimension N_i' is the constant N_i, the dimension of the space on which $\Phi = a_i$. Because of this relation, we henceforth drop the prime and refer to N_i' simply as N_i.

Now we can show that in (2.48), the P_i can be simply replaced by their classical values, so that S_i and $w_{\alpha i}$ are the usual functions of the low energy $\mathrm{U}(N_i)$ gauge fields. The components of the superfield $\mathcal{S}_i$ transform nontrivially under the symmetry (2.38), but the top component of this superfield is a c-number. Therefore, the bottom component S_i is only a quadratic function of the gauge fields $W_{\alpha i}$ of the low energy theory. If fluctuations in P_i caused S_i to be a non-quadratic function of the low energy $W_{\alpha i}$, then $\partial^2 S_i/\partial\psi^2$ would be non-constant. Thus, what is defined in (2.45) is precisely the object that one would want to define as S_i in the low energy $\mathrm{U}(N_i)$ gauge theory. Likewise, the "middle"

[4]This is trivial in perturbation theory. Moreover, the reduction to planar diagrams means that perturbation theory is summable, so the rank of P_i also cannot fluctuate nonperturbatively.

component $w_{\alpha i}$ of this superfield is linear in the low energy gauge fields, so it is precisely the superspace field strength of the i^{th} U(1) in the low energy gauge theory, that is, of the center of $\mathrm{U}(N_i)$.

In particular, since the S_i and $w_{\alpha i}$ have no quantum corrections, they are functions only of the low energy gauge fields and not of the g_k. In section 4, we will want to differentiate with respect to the g_k at "fixed gauge field background"; this can be done simply by keeping S_i and $w_{\alpha i}$ fixed.

We can describe this intuitively by saying that modulo $\{\overline{Q}_{\dot\alpha},\ \}$, the fluctuations in P_i are pure gauge fluctuations, roughly since there are no invariant data associated with the choice of an N_i-dimensional subspace in $\mathrm{U}(N)$.

The meaningfulness of the couplings t_{ij} in (1.2) depends on having correctly normalized the field strengths $w_{\alpha i}$; in the absence of a preferred normalization, by making arbitrary field redefinitions one could always set $t_{ij} = \delta_{ij}$. Since the low energy gauge theory contains particles charged under $\mathrm{U}(1)^n$ (the W bosons of broken gauge symmetry), the preferred normalization is the one in which these particles have integer charge in each U(1) factor. Again, this is clearly true of (2.45) in the semiclassical limit, and in the form (2.48) is clearly true in general.

These properties are not preserved under general superfield redefinitions $f_j(\mathcal{S}_i)$. By analogy with conventional $\mathcal{N} = 2$ supersymmetry, in which one speaks of the preferred $\mathcal{N} = 2$ superfields for which $\mathrm{U}(1)^n$ charge quantization is manifest as defining "special coordinates," we can also use the term special coordinates for the variables S_i we have discussed here.

3. The generalized Konishi anomaly

The Konishi anomaly is an anomaly for the (superfield) current[5]

$$J = \mathrm{Tr}\,\overline{\Phi} e^{\mathrm{ad}\,V}\Phi\,, \tag{3.1}$$

which generates the infinitesimal rescaling of the chiral field,

$$\delta\Phi = \epsilon\Phi\,. \tag{3.2}$$

It can be computed by any of the standard techniques: point splitting, Pauli-Villars regularization (since our model is non-chiral), anomalous variation of the functional measure (at one loop) or simply by computing Feynman diagrams [10, 21]. The result is a superfield generalization of the familiar $\mathrm{U}(1) \times \mathrm{SU}(N)^2$ mixed chiral anomaly for the fermionic component of Φ; in the theory with zero superpotential,[6]

$$\overline{D}^2 J = \frac{1}{32\pi^2}\mathrm{tr}_{adj}\ (\mathrm{ad}\,W_\alpha)(\mathrm{ad}\,W^\alpha) \tag{3.3}$$

where the trace is taken in the adjoint representation.

[5]Here $\mathrm{ad}\,V$ signifies the adjoint representation, i.e. $(\mathrm{ad}\,V\ \Phi)^i{}_j = V^i{}_k\Phi^k{}_j - \Phi^i{}_k V^k{}_j$.

[6]In a general supersymmetric theory, we have to express such an identity in terms of commutators with $\overline{Q}_{\dot\alpha}$ In the present theory, we use the existence of a superspace formalism and write the identity in terms of $\overline{D}_{\dot\alpha}$. Note that when a superspace formalism exists, $\overline{Q}_{\dot\alpha}$ and $\overline{D}_{\dot\alpha}$ are conjugate (by $\exp(\overline{\theta}\theta\partial/\partial x)$, so the chiral ring defined in terms of operators annihilated by $\overline{Q}_{\dot\alpha}$ is isomorphic to that defined using $\overline{D}_{\dot\alpha}$.

Evaluating this trace and adding the classical variation present in the theory with superpotential, we obtain

$$\overline{D}^2 J = \operatorname{Tr}\Phi \frac{\partial W(\Phi)}{\partial \Phi} + \frac{N}{16\pi^2} \operatorname{Tr} W_\alpha W^\alpha - \frac{1}{16\pi^2} \operatorname{Tr} W_\alpha \operatorname{Tr} W^\alpha. \tag{3.4}$$

One way to see why this combination of traces appears here is to check that the diagonal U(1) subgroup of the gauge group decouples.

We now take the expectation value of this equation. Since the divergence $\overline{D}^2 J$ is a $\overline{Q}^{\dot\alpha}$-commutator, it must have zero expectation value in a supersymmetric vacuum. Furthermore, we can use (2.3) and $\langle \operatorname{Tr} W_\alpha \rangle = 0$ to see that the last term is zero. Thus, we infer that

$$\left\langle \operatorname{Tr}\Phi \frac{\partial W(\Phi)}{\partial \Phi} \right\rangle = -2\frac{N}{32\pi^2} \left\langle \operatorname{Tr} W_\alpha W^\alpha \right\rangle .$$

Using (2.19), this can also be formulated in terms of a constraint on W_{eff}. The left hand side reproduces (2.24), so we find that

$$\sum_{k\geq 0} (k+1) g_k \frac{\partial}{\partial g_k} W_{\text{eff}} = 2NS \tag{3.5}$$

where (in the notations of section 2)

$$S = \sum_i S_i = -\frac{1}{32\pi^2} \langle \operatorname{Tr} W_\alpha W^\alpha \rangle.$$

This has the solution (2.17), and the general solution is (2.17) plus a solution of the homogeneous equation (2.24).

In general, this anomaly receives higher loop contributions, which are renormalization scheme dependent and somewhat complicated. However, one can see without detailed computation that these contributions can all be written as non-chiral local functionals.

To see this, we consider possible chiral operator corrections to (3.4), and enforce symmetry under the $\mathrm{U}(1)_\Phi \times \mathrm{U}(1)_\theta$ of (2.23). An anomaly must have charges $(0,2)$ under $\mathrm{U}(1)_\Phi \times \mathrm{U}(1)_\theta$. Furthermore the correction must vanish for $g_k = 0$, so inverse powers of the couplings cannot appear. Referring to (2.23), we see that the only expressions with the right charges are W_α^2 and $g_k \Phi^{k+1}$, with no additional dependence on couplings. These are the terms already evaluated in (3.4), and thus no further corrections are possible.[7] This

[7] The situation in the full gauge theory is less clear. Including nonperturbative effects, one might generate corrections such as $g_{2Nl+k-1} \Lambda^{2Nl} \operatorname{Tr} \Phi^k$. Such terms would not affect the perturbative analysis of W_{eff} in section 4, but would affect the study of gauge dynamics in section 5. In particular, this simple form of the Ward identities depends on making the proper definition of the operators $\operatorname{Tr} \Phi^k$ for $k \geq N$, which has interesting aspects discussed in section 5 and appendix A. We suspect, but will not prove here, that such nonperturbative corrections can be excluded (after properly defining the operators) by considering the algebra of chiral transformations $\delta\Phi = \sum_n \epsilon_n \Phi^{n+1}$ (as is well known in matrix theory, this algebra is a partial Virasoro algebra), showing that it can undergo no quantum corrections, and using this algebra to constrain the anomalies, along the lines of the Wess-Zumino consistency condition for anomalies.

JHEP12(2002)071

implies that, for purposes of computing the chiral ring and effective superpotential, the result (3.4) is exact.[8]

Thus we have an exact constraint on the functional $W_{\rm eff}$. Readers familiar with matrix models will notice the close similarity between (3.5) and the L_0 Virasoro constraint of the one bosonic matrix model (this was also pointed out in [22]). Indeed, the similarity is not a coincidence as the matrix model L_0 constraint has a very parallel origin: it is a Ward identity for the matrix variation $\delta M = \epsilon M$. In the matrix model, one derives further useful constraints (which we recall in section 4.2) from the variation $\delta M = \epsilon M^{n+1}$.

This similarity suggests that we try to get more constraints on the functional $W_{\rm eff}(g_n, S_i)$ by considering possible anomalies for the currents

$$J_n = {\rm Tr}\, \overline{\Phi} e^{{\rm ad}\, V} \Phi^{n+1}\,, \tag{3.6}$$

which generate the variations

$$\delta\Phi = \epsilon_n \Phi^{n+1}\,. \tag{3.7}$$

The hope would be that these would provide an infinite system of equations analogous to (3.5), which could determine the infinite series of expectation values $\langle {\rm Tr}\, \Phi^n \rangle$.

It is easy to find the generalization of the classical term in (3.4), by applying the variations (3.7) to the action. This produces

$$\overline{D}^2 J_n = {\rm Tr}\, \Phi^{n+1} \frac{\partial W(\Phi)}{\partial \Phi} + \text{anomaly} + \overline{D}(\ldots)\,. \tag{3.8}$$

Following the same procedure which led to (3.5), we obtain

$$\sum_{k\geq 0} (k+n+1) g_k \frac{\partial}{\partial g_{k+n}} W_{\rm eff} = \text{anomaly terms}\,. \tag{3.9}$$

These are independent constraints on $W_{\rm eff}$, motivating us to continue and compute the anomaly terms. However, before doing this, we should recognize that we did not yet consider the most general variation in the chiral ring. This would have been

$$\delta\Phi = f(\Phi, W_\alpha) \tag{3.10}$$

for a general holomorphic function $f(\Phi, W_\alpha)$. It is no harder to compute the anomaly for the corresponding current

$$J_f = {\rm Tr}\, \overline{\Phi} e^{{\rm ad}\, V} f(\Phi, W_\alpha)\,, \tag{3.11}$$

so let us do that.

[8] We can also reach this conclusion by quite a different argument. Pauli-Villars regularization of the Φ kinetic energy improves the convergence of all diagrams containing Φ fields with more than one loop, without modifying chiral quantities. So anomalies obtained by integrating out Φ must arise only at one-loop order.

3.1 Computation of the generalized anomaly

We want to compute

$$\overline{D}^2\langle J_f\rangle = \overline{D}^2\langle \mathrm{Tr}\,\overline{\Phi}e^{\mathrm{ad}\,V}f(\Phi,W_\alpha)\rangle\,. \tag{3.12}$$

Let us first do this at zeroth order in the couplings g_k (except that we assume a mass term); we will then argue that holomorphy precludes corrections depending on these couplings.

At zero coupling, the one loop contributions to (3.12) come from graphs involving $\overline{\Phi}$ and a single Φ in $f(\Phi,W_\alpha)$. In any one of these graphs, the other appearances of Φ and W_α in f are simply spectators. In other words, given

$$A_{ij,kl} \equiv \overline{D}^2\langle\overline{\Phi}_{ij}e^{\mathrm{ad}\,V}\Phi_{kl}\rangle\,,$$

the generalized anomaly at one loop is

$$\overline{D}^2\langle \mathrm{Tr}\,\overline{\Phi}e^{\mathrm{ad}\,V}f(\Phi,W_\alpha)\rangle = \sum_{ijkl} A_{ij,kl}\frac{\partial f(\Phi,W_\alpha)_{ji}}{\partial\Phi_{kl}}\,. \tag{3.13}$$

The index structure on the right hand side should be clear on considering the following example,

$$\frac{\partial}{\partial\Phi_{kl}}\left(\Phi^{m+1}\right)_{ji} = \sum_{s=0}^{m}(\Phi^s)_{jk}(\Phi^{m-s})_{li}\,.$$

In general, the expression (3.13) might need to be covariantized using e^V and e^{-V}. However, we are only interested in the chiral part of the anomaly, for which this is not necessary.

In fact, the computation of $A_{ij,kl}$ is the same as the computation of (3.3), with the only difference being that we do not take the trace. Thus

$$\begin{aligned} A_{ij,kl} &= \frac{1}{32\pi^2}\left[(W_\alpha W^\alpha)_{il}\delta_{jk} + (W_\alpha W^\alpha)_{jk}\delta_{il} - 2(W_\alpha)_{il}(W^\alpha)_{jk}\right] \\ &\equiv \frac{1}{32\pi^2}\left[W_\alpha,[W^\alpha,e_{lk}]\right]_{ij} \end{aligned}$$

where e_{lk} is the basis matrix with the single non-zero entry $(e_{lk})_{ij} = \delta_{il}\delta_{jk}$.

Substituting this in (3.13) and adding the classical variation, we obtain the final result

$$\overline{D}^2 J_f = \mathrm{Tr}\,f(\Phi,W_\alpha)\frac{\partial W(\Phi)}{\partial\Phi} + \frac{1}{32\pi^2}\sum_{i,j}\left[W_\alpha,\left[W^\alpha,\frac{\partial f}{\partial\Phi_{ij}}\right]\right]_{ji}\,. \tag{3.14}$$

Finally, we argue that this result cannot receive perturbative (or nonperturbative) corrections in the coupling. The argument starts off in the same way as for the standard Konishi anomaly, using $\mathrm{U}(1)_\Phi\times\mathrm{U}(1)_\theta$ symmetry to conclude that the only allowed terms are those involving the same powers of the fields and the couplings as already appear in (3.14). The same sort of analysis that we used in section 2 to show that particular interactions only arise from diagrams with a particular number of loops (for example, S^2 only arises from two-loop diagrams) then shows that the terms proportional or not proportional to W_α^2 in (3.14) can only arise from one-loop or tree level diagrams. But the tree level contribution is classical, and we have already evaluated the one-loop contribution to arrive at (3.14).

Taking expectation values, we obtain the Ward identities

$$\left\langle \operatorname{Tr} f(\Phi, W_\alpha) \frac{\partial W}{\partial \Phi} \right\rangle = -\frac{1}{32\pi^2} \left\langle \sum_{i,j} \left(\left[W_\alpha, \left[W^\alpha, \frac{\partial f(\Phi, W_\alpha)}{\partial \Phi_{ij}} \right] \right] \right)_{ij} \right\rangle . \tag{3.15}$$

Functional measure description. In the next section, we will use these Ward identities to solve for W_{eff}. They have a close formal similarity with those of the one matrix model, and a good way to see the reason for this is to rederive them using the technique of anomalous variation of the functional measure. This technique can be criticized on the grounds that it is not obvious how to extend it beyond one loop, and this is why we instead gave the more conventional perturbative argument for (3.14). Now that we have the result in hand, we make a brief interlude to explain it from this point of view.

A simple way to derive classical Ward identities such as (3.8) is to perform the variation under the functional integral: we write

$$\int [D\Phi] \operatorname{Tr} \left(f(\Phi) \frac{\partial}{\partial \Phi} \right) \exp \left(-\int d^2\theta W(\Phi) - \text{c.c.} \right) = -\int d^2\theta \left\langle \operatorname{Tr} f(\Phi) \frac{\partial W}{\partial \Phi} \right\rangle . \tag{3.16}$$

In the classical theory, this would vanish by the equations of motion. However, in the quantum theory, the functional measure might not be invariant under such a variation. Formally, we might try to evaluate this term by integrating by parts under the functional integral (3.16). This would produce the anomalous Ward identity

$$\left\langle \operatorname{Tr} f(\Phi) \frac{\partial W}{\partial \Phi} \right\rangle = \left\langle \sum_{i,j} \frac{\partial f(\Phi)_{ij}}{\partial \Phi_{ij}} \right\rangle . \tag{3.17}$$

Such a term might be present in bosonic field theory, and is clearly present in zero dimensional field theory (i.e. the bosonic one matrix integral), as integration by parts is obviously valid. In fact it is simply the jacobian for an infinitesimal transformation $\delta\Phi = f$,

$$\log J = \operatorname{tr} \log \left(1 + \frac{\delta f(\Phi)}{\delta \Phi} \right) , \tag{3.18}$$

which expresses the change in functional measure from Φ to $\Phi + \delta\Phi$. Thus the Ward identities (3.17), which can serve as a starting point for solving the matrix model (and which we will discuss further in section 4.2), in fact express this "anomalous" or "quantum" variation of the measure.

One can make the same computation in supersymmetric theory, by computing the product of jacobians for the individual component fields. As is well-known (this is used, for example, in the discussion of the Nicolai map [23]), the result is $J = 1$, by cancellations between the jacobians of the bosonic and fermionic components. Consistent with this, the anomalous term in (3.17) is not present in (3.15).

However, to properly compute the variation of the functional measure in field theory, one must regulate the trace appearing in (3.18). This is how the Konishi anomaly was computed in [10]. It is straightforward to generalize their computation by taking

$$\begin{aligned} \log J &= \epsilon \ \operatorname{Str} e^{-tL} \frac{\delta f(\Phi, W_\alpha)}{\delta \Phi} \\ &= \epsilon \left(\operatorname{Tr}_\phi \ e^{-t(D_\mu)^2} \frac{\delta f(\Phi, W_\alpha)}{\delta \Phi} - \operatorname{Tr}_\psi \ e^{-t\not{D}^2} \frac{\delta f(\Phi, W_\alpha)}{\delta \Phi} + \operatorname{Tr}_F \ e^{-t(D_\mu)^2} \frac{\delta f(\Phi, W_\alpha)}{\delta \Phi} \right) , \end{aligned} \tag{3.19}$$

where $L = \overline{D}^2 e^{-2V} D^2 e^{2V}$ is an appropriate superspace wave operator, to obtain precisely the result (3.15).

More general quiver theories. Without going into details, let us mention that all of the arguments we just gave generalize straightforwardly to a general gauge theory with a product of classical gauge groups, bifundamental and adjoint matter. An important point here is that the identities (2.4) and (2.5), which freed us from the need to specify the operator ordering of W_α insertions, now free us from the need to specify which gauge group each W_α lives in (in any given ordering of Φ^a and W_α's, this is fixed by gauge invariance). One still needs to keep track of the ordering of the various chiral multiplets, call them Φ^a.

One can then use a general holomorphic variation

$$\delta\Phi^a = f^a(\Phi, W_\alpha)$$

in the arguments following (3.10), to derive Ward identities very similar to (3.15). These are analogs of standard Ward identities known for multi-matrix models [24] which generalize the "factorized loop equations" of large-N Yang-Mills theory [25].

4. Solution of the adjoint theory

We now have the ingredients we need to find $W_{\text{eff}}(S, g)$. Readers familiar with matrix models will recognize that much of the formalism used in the following discussion was directly borrowed from that theory. However, we do not assume familiarity with matrix models, nor do we rely on any assumed relation between the problems in making the following arguments.

We also stress that the arguments in this section generally do not assume the reduction to planar diagrams which we argued for in section 2, but will lead to an independent derivation of this claim. Indeed, the arguments of section 3 and this section do not use any diagrammatic expansion and are completely non-perturbative.

We are still considering the one adjoint theory with superpotential (1.1). We assemble all of its chiral operators into a generating function,

$$\mathcal{R}(z, \psi) = -\frac{1}{2} \operatorname{Tr} \left(\frac{1}{4\pi} W_\alpha - \psi_\alpha \right)^2 \frac{1}{z - \Phi} . \tag{4.1}$$

We simply regard this as a function whose expansion in powers of $1/z$ and ψ generates the chiral ring. Its components in the ψ expansion are

$$\mathcal{R}(z, \psi) = R(z) + \psi_\alpha w^\alpha(z) - \psi^1 \psi^2 T(z) \tag{4.2}$$

with

$$\begin{aligned} T(z) &= \sum_{k \geq 0} z^{-1-k} \operatorname{Tr} \Phi^k = \operatorname{Tr} \frac{1}{z - \Phi} ; \\ w_\alpha(z) &= \frac{1}{4\pi} \operatorname{Tr} W_\alpha \frac{1}{z - \Phi} ; \\ R(z) &= -\frac{1}{32\pi^2} \operatorname{Tr} W_\alpha W^\alpha \frac{1}{z - \Phi} . \end{aligned} \tag{4.3}$$

We will also use the same symbols to denote the vacuum expectation values of these operators; the meaning should be clear from the context. Finally, we write

$$\mathcal{R}(z,\psi)_{ij} = -\frac{1}{2}\left(\left(\frac{1}{4\pi}W_\alpha - \psi_\alpha\right)^2 \frac{1}{z-\Phi}\right)_{ij} \tag{4.4}$$

to denote the corresponding matrix expressions (before taking the trace).

Note that the expansion in powers of ψ runs "backwards" compared to a typical superspace formalism: the lowest component is $R(z)$, the gaugino condensate, while operators constructed only from powers of Φ are upper components. Because of this, it will turn out that $R(z)$ is the basic quantity from which the behavior of the others can be inferred.

Basic loop equations for gauge theory. Let us start by deriving Ward identities for the lowest component $R(z)$ in (4.2), i.e. with $\psi = 0$. By then taking non-zero ψ, we will obtain the Ward identities following from the most general possible variation which is a function of chiral operators.

So, we start with

$$\delta\Phi_{ij} = f_{ij}(\Phi, W_\alpha) = -\frac{1}{32\pi^2}\left(\frac{W_\alpha W^\alpha}{z-\Phi}\right)_{ij}. \tag{4.5}$$

Substituting into (3.15), one finds the Ward identity

$$\left\langle -\frac{1}{32\pi^2}\sum_{i,j}\left[W_\alpha, \left[W^\alpha, \frac{\partial}{\partial\Phi_{ij}}\frac{W_\alpha W^\alpha}{z-\Phi}\right]\right]_{ij}\right\rangle = \left\langle \mathrm{Tr}\left(\frac{W'(\Phi)W_\alpha W^\alpha}{z-\Phi}\right)\right\rangle. \tag{4.6}$$

We next note the identity

$$\sum_{i,j}\left[\chi_1, \left[\chi_2, \frac{\partial}{\partial\Phi_{ij}}\frac{\chi_1\chi_2}{z-\Phi}\right]\right]_{ij} = \left(\mathrm{Tr}\,\frac{\chi_1\chi_2}{z-\Phi}\right)^2 \tag{4.7}$$

which is valid if $\chi_1^2 = \chi_2^2 = 0$ and the chiral operators Φ and χ_α commute, as can be verified by elementary algebra.

Applying this identity with $\chi_\alpha = W_\alpha$, we find

$$\langle R(z)\ R(z)\rangle = \left\langle \mathrm{Tr}\left(W'(\Phi)R(z)\right)\right\rangle.$$

We now come to the key point. Expectation values of products of gauge invariant chiral operators factorize, as expressed in (2.3). This allows us to rewrite (4.6) as

$$\langle R(z)\rangle^2 = \left\langle \mathrm{Tr}\left(W'(\Phi)R(z)\right)\right\rangle.$$

Thus, both sides have been expressed purely in terms of the vacuum expectation values $\langle \mathrm{Tr}\, W_\alpha W^\alpha \Phi^k\rangle$, allowing us to write a closed equation for $R(z)$.

This point can be made more explicit by rewriting the right hand side in the following way. We start by noting the identity

$$\begin{aligned}\mathrm{Tr}\,\frac{W'(\Phi)W_\alpha W^\alpha}{z-\Phi} &= \mathrm{Tr}\,\frac{(W'(\Phi) - W'(z) + W'(z))W_\alpha W^\alpha}{z-\Phi} \\ &= W'(z)\mathrm{Tr}\,\frac{W_\alpha W^\alpha}{z-\Phi} + \frac{1}{4}f(z)\end{aligned} \tag{4.8}$$

with

$$f(z) = 4 \ \mathrm{Tr}\, \frac{(W'(\Phi) - W'(z))W_\alpha W^\alpha}{z - \Phi}\,.$$

(The factor of 4 is introduced for later convenience.) The function $f(z)$ is analytic and, for polynomial W, is a polynomial in z of degree $n - 1 = \deg W' - 1$, simply because $W'(\Phi) - W'(z)$ vanishes at $z - \Phi = 0$. In terms of this function, we can write

$$R(z)^2 = W'(z)R(z) + \frac{1}{4}f(z) \tag{4.9}$$

This statement holds as a statement in the chiral ring, and (therefore) also holds if the operators $R(z)$ and $f(z)$ are replaced by their values in a supersymmetric vacuum.

Another way to understand this result is to note that

$$\mathrm{Tr}\, \frac{W'(\Phi)W_\alpha W^\alpha}{z - \Phi} \sim O\left(\frac{1}{z}\right)$$

for large z, while $f(z)$ as defined above is polynomial. Thus we can write

$$-\frac{1}{32\pi^2}\mathrm{Tr}\, \frac{W'(\Phi)W_\alpha W^\alpha}{z - \Phi} = [W'(z)R(z)]_-$$

where the notation $[F(z)]_-$ means to drop the non-negative powers in a Laurent expansion in z, i.e.

$$[z^k]_- = \begin{cases} z^k & \text{for } k < 0 \\ 0 & \text{for } k \geq 0 \end{cases}\,.$$

Using this notation, we can write the Ward identity as

$$R(z)^2 = [W'(z)R(z)]_-\,. \tag{4.10}$$

In other words, the role of $f(z)$ in (4.9) is just to cancel the polynomial part of the right hand side.

Readers familiar with matrix models will recognize (4.9) and (4.10) as the standard loop equations for the one bosonic matrix model. If the resolvent of the matrix field M is defined by

$$R_M(z) = \frac{g_m}{\widehat{N}}\left\langle \mathrm{Tr}\, \frac{1}{z - M}\right\rangle, \tag{4.11}$$

then it obeys precisely the above Ward identities, as we recall more fully in section 4.2. Thus, we will soon be in a position to prove a precise relation between gauge theory expectation values $\langle \ldots \rangle_{g.t.}$ and matrix model expectation values $\langle \ldots \rangle_{m.m.}$,

$$-\frac{1}{32\pi^2}\left\langle \mathrm{Tr}\, \frac{W_\alpha W^\alpha}{z - \Phi}\right\rangle_{g.t.} = \frac{g_m}{\widehat{N}}\left\langle \mathrm{Tr}\, \frac{1}{z - M}\right\rangle_{m.m.}.$$

Let us continue, however, without assuming any *a priori* relationship, and complete the computation of W_{eff} purely in terms of gauge theory.

General loop equations for gauge theory. The general system of Ward identities is no harder to derive. We apply the general variation (3.10) with

$$\delta\Phi_{ij} = f_{ij}(\Phi, W_\alpha) = \mathcal{R}(z,\psi)_{ij}\,. \tag{4.12}$$

Substituting into (3.15), one finds the Ward identity

$$\left\langle \frac{1}{4(4\pi)^2}\sum_{i,j}\left[W_\alpha,\left[W^\alpha,\frac{\partial}{\partial\Phi_{ij}}\frac{(\frac{1}{4\pi}W_\alpha-\psi_\alpha)^2}{z-\Phi}\right]\right]_{ij}\right\rangle = \left\langle \mathrm{Tr}\left(W'(\Phi)\mathcal{R}(z,\psi)\right)\right\rangle. \tag{4.13}$$

Now, one can substitute $\mathrm{ad}(\frac{1}{4\pi}W_\alpha-\psi_\alpha)$ for $\mathrm{ad}\,\frac{1}{4\pi}W_\alpha$ in this expression, as $\mathrm{ad}\,\mathbf{1}=0$. One can then apply (4.7) with $\chi_\alpha=\frac{1}{4\pi}W_\alpha-\psi_\alpha$, and gauge theory factorization (2.3), exactly as before, to find

$$\mathcal{R}(z,\psi)^2 = \mathrm{Tr}\left(W'(\Phi)\mathcal{R}(z,\psi)\right).$$

Note that the variable ψ does not appear explicitly in this result. This expresses the decoupling of the diagonal U(1), and thus was clear *a priori*. Again, this statement holds as a statement about the chiral ring, or as a statement about expectation values.

Finally, we can apply the same identity (4.8), now using a degree $n-1$ polynomial $f(z,\psi)$, to obtain

$$\mathcal{R}(z,\psi)^2 = W'(z)\mathcal{R}(z,\psi)+\frac{1}{4}f(z,\psi)\,. \tag{4.14}$$

This equation contains (4.9) and higher components which will allow us to determine the expectation values of the other operators in the chiral ring. Expanding (4.14) in powers of ψ, and writing

$$f(z,\psi) = f(z)+\psi_\alpha\rho^\alpha(z)-\psi_1\psi_2 c(z), \tag{4.15}$$

we obtain the component equations

$$\begin{aligned} R^2(z) &= W'(z)R(z)+\frac{1}{4}f(z)\,,\\ 2R(z)w_\alpha(z) &= W'(z)w_\alpha(z)+\frac{1}{4}\rho_\alpha(z)\,,\\ 2R(z)T(z)+w_\alpha(z)w^\alpha(z) &= W'(z)T(z)+\frac{1}{4}c(z)\,. \end{aligned} \tag{4.16}$$

This system of equations is the complete set of independent Ward identities which can be derived using the generalized Konishi anomaly.[9] It can be used in various ways. For example, by expanding in powers of $1/z$, one can derive a system of recursion relations determining all vevs in terms of those with $k\le n$. One can also expand $\mathcal{R}$ in powers of the couplings, to derive Schwinger-Dyson equations which recursively build up planar diagrams from subdiagrams.

[9]There are further equations which can be derived by considering correlation functions of (3.11) with other chiral operators, but these are equivalent to derivatives of (4.14) with respect to the couplings.

4.1 General solution and derivation of $\mathcal{F}$

Since the equation (4.14) is quadratic in $\mathcal{R}$, it is trivial to write its general solution:

$$2\mathcal{R}(z,\psi) = W'(z) - \sqrt{W'(z)^2 + f(z,\psi)}\,. \tag{4.17}$$

The sign of the square root is determined by the asymptotics $\mathcal{R}(z,\psi) \sim 1/z$ at large z.

The result determines all expectation values of chiral operators in terms of the $2n$ bosonic and $2n$ fermionic coefficients of $f(z,\psi)$. Furthermore, to obtain $W_{\rm eff}$, we need merely integrate one of these expectation values with respect to its coupling. Thus we are close to having the solution, if we can interpret these parameters.

The meaning of the solution is perhaps easier to understand by considering the component expansion (4.16). The ψ^0 term is a quadratic equation for $R(z)$, whose solution is determined by the n coefficients f_i in the expansion

$$f(z) = \sum_{i=0}^{n-1} f_i z^i\,.$$

$R(z)$ is not a single-valued function of z, but rather is single valued on a Riemann surface branched over the z plane; call this Σ. Since $W'(z)^2 + f(z)$ is a polynomial of order $2n$, Σ is a Riemann surface of genus $n-1$.

The functions $w_\alpha(z)$ and $T(z)$ are then derived from $R(z)$ by solving linear equations involving the higher ψ components in $f(z,\psi)$. In particular, the gauge theory expectation values

$$\left\langle \mathrm{Tr}\,\Phi^{n+1}\right\rangle$$

are encoded in the expansion of the component $T(z)$ that is quadratic in ψ, and depend on all of the parameters in f.

As we discussed in section 2, there is a related set of $2n+2n$ parameters, the components of $\mathcal{S}_i$ defined in (2.49),

$$\begin{aligned}\mathcal{S}_i &= S_i + \psi_\alpha w_i^\alpha - \psi^1\psi^2 N_i\\ &= \frac{1}{2\pi i}\oint_{C_i} dz\,\mathcal{R}(z,\psi)\,.\end{aligned} \tag{4.18}$$

These are different from the coefficients in the expansion of $f(z,\psi)$, but the two sets of parameters are simply related. Substituting (4.17) in (4.18) and considering the $\psi^\alpha = 0$ component, we get

$$S_i = -\frac{1}{4\pi i}\oint_{C_i} dz\sqrt{W'(z)^2 + f(z)}\,. \tag{4.19}$$

(The polynomial $W'(z)$, being non-singular, does not contribute to the contour integral.) We expanded $f(z)$ as $f(z) = \sum_i f_i z^i$. Differentiating with respect to the coefficients f_i, one finds

$$\frac{\partial S_i}{\partial f_j} = -\frac{1}{8\pi i}\oint_{C_i} dz\frac{z^j}{\sqrt{W'(z)^2 + f(z)}}\,. \tag{4.20}$$

The integrands in (4.20) for $0 \le j \le n-2$ provide a complete set of $n-1$ holomorphic one-forms on the genus $g = n-1$ Riemann surface Σ, while the n cycles C_i provide a (once redundant) basis of A cycles on Σ. Thus, by taking an $(n-1)\times(n-1)$ submatrix of (4.20), we have a period matrix, which for a generic surface Σ this matrix will have non-vanishing determinant.

Finally, by examining the $\mathcal{O}(1/z)$ term of $R(z)$, the remaining variable f_{n-1} can be identified with $S=\sum_i S_i$, the ψ^0 component of an integral (4.18) around a contour surrounding all the cuts. More precisely, one finds

$$R(z) \sim \frac{S}{z} = -\frac{1}{z} f_{n-1} \frac{n!}{4(\partial^{n+1}W/\partial z^{n+1})} = -\frac{1}{z}\frac{f_{n-1}}{4g_n}\,.$$

Combining this relation with (4.20), one finds that the coordinate transformation $S_i \to S_i'(S_j)$ is generically non-singular. Thus, in general either set of parameters, S_i or f_j, could be used to parameterize the family of surfaces Σ which appear in the solutions of (4.14).

As we discussed in section 2, the natural fields which specify a gauge field background in a Φ-independent way are the $\mathcal{S}_i$. Thus we can regard the solution (4.17) as a function of the couplings g_k, but at fixed $\mathcal{S}_i$, as giving us the complete set of expectation values of chiral operators in the background specified by $\mathcal{S}_i$. We could then obtain the effective superpotential $W_{\rm eff}$ by integrating any of these expectation values with respect to its coupling (using (2.19)), up to a coupling-independent constant of integration.

Rather than derive $W_{\rm eff}$ directly, we can get the complete effective action by first deriving the function $\mathcal{F}_p$ of section 2. As we discussed there, the ψ shift symmetry tells us that

$$W_{\rm eff} = \int d^2\psi\ \mathcal{F}_p(\mathcal{S}_i, g_k)$$

for some function $\mathcal{F}_p(\mathcal{S}_i, g_k)$. Using (2.19), we have that

$$\begin{aligned} -\frac{1}{2(k+1)}\int d^2\psi \left\langle \mathrm{Tr}(\frac{1}{4\pi}W_\alpha - \psi_\alpha)^2\Phi^{k+1}\right\rangle_\Phi &= \frac{1}{k+1}\left\langle \mathrm{Tr}\,\Phi^{k+1}\right\rangle_\Phi \\ &= \int d^2\psi \frac{\partial \mathcal{F}_p(\mathcal{S}_i, g_j)}{\partial g_k}\,. \end{aligned} \qquad (4.21)$$

This equation does not provide enough information to determine $\partial\mathcal{F}_p/\partial g_k$, since there are functions (depending only on the $\widehat{S}_i$ introduced in (1.4), for example) that are annihilated by $\int d^2\psi$. However, all we are planning to do with $\mathcal{F}_p$, or its derivative $\partial\mathcal{F}_p/\partial g_k$, is to act with $\int d^2\psi d^2\theta$ to compute the effective action, or its derivative with respect to g_k. As we discussed in introducing (2.40), $\mathcal{F}_p$ is uniquely determined if we require it to be a function only of the $\mathcal{S}_i$, and to vanish at $\mathcal{S}_i=0$. In that case, $\partial\mathcal{F}_p/\partial g_k$ will have the same properties. The unique $\mathcal{F}_p$ that has those properties and obeys (4.21) also satisfies

$$\frac{\partial \mathcal{F}_p}{\partial g_k} = -\frac{1}{2(k+1)}\left\langle \mathrm{Tr}(\frac{1}{4\pi}W_\alpha - \psi_\alpha)^2\Phi^{k+1}\right\rangle_\Phi \qquad (4.22)$$

By integrating over ψ, (4.22) clearly implies (4.21). (4.22) implies the auxiliary conditions on $\mathcal{F}_p$, since from section 2, we know that the right hand side of (4.22) is a function only of $\mathcal{S}_i$ and vanishes at $\mathcal{S}_i=0$. That (4.21) plus the auxiliary conditions on $\mathcal{F}_p$ imply (4.22) should be clear: no nonconstant function of the $\mathcal{S}_i$ is annihilated by $\int d^2\psi$, so adding such a function to the right hand side of (4.22) would spoil (4.21).

JHEP12(2002)071

Finally, we can complete our argument. The right hand side of (4.22) is determined as a function of $\mathcal{S}_i$ by (4.17). By integrating with respect to the g_k, $\mathcal{F}_p$ is determined from (4.22) up to a function independent of the g_k. By scaling the g_k, or equivalently scaling the superpotential $W \to W/\hbar$ with $\hbar$ small, we can reduce the evaluation of $\mathcal{F}_p$ to a one-loop contribution, which can be computed explicitly. So finally we have shown that the anomalies plus explicit evaluation of a one-loop contribution suffice to determine W_{eff}.

In summary, using gauge theory arguments, we have derived a procedure which determines $\mathcal{F}_p$ up to a g_k-independent, but possibly S_i-dependent, contribution. The subsequent gauge theory functional integral can also produce such terms, so if we are discussing the solution of the full gauge theory, the problem of determining this contribution is not really separable from the problem of gauge dynamics.

4.2 Comparison with the bosonic matrix model

Our arguments so far did not assume any relationship between gauge theory and matrix models, and did not assume that the gauge theory reduced to summing planar diagrams. Indeed, given the non-perturbative validity of (3.14), the arguments are valid non-perturbatively.

Nevertheless, the final result is most simply described by the relation to the matrix model. We very briefly review the matrix model to explain this point. For a more comprehensive review, see [26].

The bosonic one matrix model is simply the integral over an $\widehat{N} \times \widehat{N}$ hermitian matrix M, with action $(\widehat{N}/g_m)\mathrm{tr}\, W(M)$. ($\widehat{N}$ is different from N of the gauge theory, and will be taken to infinity, while N is held fixed.) To make the action dimensionless, we introduced a dimensionful parameter g_m with dimension 3 (like $W(\Phi)$). W in general could be any smooth function with reasonable properties, but for present purposes, it is the same function that serves as the superpotential $W(\Phi)$ of the four-dimensional gauge theory. The free energy $F_{m.m.}$ of the matrix theory is defined by

$$\exp\left(-\frac{\widehat{N}^2}{g_m^2} F_{m.m.}\right) = \int d^{\widehat{N}^2} M \ \exp\left(-\frac{\widehat{N}}{g_m}\,\mathrm{tr}\, W(M)\right). \tag{4.23}$$

This model admits an 't Hooft large $\widehat{N}$ limit, which is reached by taking $\widehat{N} \to \infty$ at fixed $W(M)$ and g_m. As is well known, in this limit, the perturbative expansion of the free energy $F_{m.m.}$ reduces to planar diagrams. One also has many nonperturbative techniques for computing the exact free energy $F_{m.m.}$, as first explained in [27].

One of the many ways to see this is to derive loop equations, which are Ward identities derived by considering the variations

$$\delta M = \epsilon M^{n+1}$$

or equivalently by using the identity

$$0 = \int d^{\widehat{N}^2} M \ \mathrm{Tr}\left(\frac{\partial}{\partial M}\, M^n\right) \exp\left(-\frac{\widehat{N}}{g_m}\,\mathrm{tr}\, W(M)\right). \tag{4.24}$$

These Ward identities are most simply formulated in terms of the matrix model resolvent. For this, one takes a generating function of the identities (4.24), and writes

$$0 = \int d^{\widehat{N}^2} M \ \mathrm{Tr}\left(\frac{\partial}{\partial M}\,\frac{1}{z-M}\right)\exp\left(-\frac{\widehat{N}}{g_m}\,\mathrm{tr}\,W(M)\right). \tag{4.25}$$

Evaluating this expression, one learns that

$$\left(\frac{g_m}{\widehat{N}}\right)^2\left\langle\left(\mathrm{Tr}\,\frac{1}{z-M}\right)^2\right\rangle = \frac{g_m}{\widehat{N}}\left\langle\mathrm{Tr}\,\frac{W'(M)}{z-M}\right\rangle. \tag{4.26}$$

If now we define the matrix model resolvent,

$$R_m(z) = \frac{g_m}{\widehat{N}}\left\langle\mathrm{Tr}\,\frac{1}{z-M}\right\rangle, \tag{4.27}$$

then from (4.26), exactly as in section 3, we deduce that

$$\left\langle R_m(z)^2\right\rangle = \left\langle W'(z)R_m(z)\right\rangle + \frac{1}{4}f_m(z)\,, \tag{4.28}$$

where $f_m(z)$ is a polynomial of degree $n-1$. Equivalently, we deduce that

$$\left\langle R_m(z)^2\right\rangle = \left\langle\,[W'(z)R_m(z)]_-\,\right\rangle\,. \tag{4.29}$$

If now we take $\widehat{N}\to\infty$, then correlation functions factor in the matrix model,

$$\left\langle R_m(z)^2\right\rangle = \left\langle R_m(z)\right\rangle^2. \tag{4.30}$$

In the gauge theory, the analogous factorization was justified using the properties of chiral operators, without a large-N limit. This lets us rewrite (4.28) in the form

$$R_m(z)^2 = W'(z)R_m(z) + \frac{1}{4}f_m(z)\,, \tag{4.31}$$

where now the expectation value of $R_m(z)$ is understood. Here we recognize an equation (the first equation in (4.16)) that we have seen in analyzing the gauge theory.

In the matrix model, as explained by Dijkgraaf and Vafa, the choice of f_m corresponds to the choice of how to distribute $\widehat{N}$ eigenvalues of the matrix M among the n critical points of W. Since the gauge theory object $R(z)$ obeys the same equation as the matrix model resolvent $R_m(z)$, they will be equal if $f = f_m$, or equivalently if the fields S_i of the gauge theory equal the analogous objects in the matrix model. In other words, they are equal if

$$S_i = \frac{1}{2\pi i}\oint_{C_i} R_m(z)dz = \frac{1}{2\pi i}\oint_{C_i}\frac{g_m}{\widehat{N}}\left\langle\mathrm{Tr}\,\frac{1}{z-M}\right\rangle = \frac{g_m\widehat{N}_i}{\widehat{N}}\,, \tag{4.32}$$

where $\widehat{N}_i$ is the number of eigenvalues of M near the i^{th} critical point. Dijkgraaf and Vafa express this relation as $S_i = g_s\widehat{N}_i$, where $g_s = g_m/\widehat{N}$. In the gauge theory, $R(z)$ depends on complex variables S_i, which cannot be varied independently if we want them to be of the form $S_i = g_m \widehat{N}_i/\widehat{N}$. Nevertheless, the holomorphic function $R(z)$ is determined by its values for S_i that are of this form, and in this sense the matrix model determines $R(z)$.

The derivative of the matrix model free energy with respect to the couplings is

$$\frac{\partial F_{m.m.}}{\partial g_k} = \left\langle \frac{\mathrm{Tr}\ M^{k+1}}{k+1} \right\rangle, \tag{4.33}$$

as one learns by directly differentiating the definition (4.23). So $R_m(z)$ is the generating function for $\partial F_{m.m.}/\partial g_k$, just as the gauge theory object $R(z)$ is the generating function for $\partial \mathcal{F}/\partial g_k$ at $\psi = 0$. So for gauge theory and matrix model parameters related by (4.32), we have

$$F_{m.m.}(S_i, g_k) = \mathcal{F}_p(S_i, g_k)|_{\psi=0} + \mathcal{H}(S_i), \tag{4.34}$$

where $\mathcal{H}(S_i)$ is independent of g_k. This is the relation between the two models as formulated by Dijkgraaf and Vafa, who further claim that the left-hand side (or the right hand side including $\mathcal{H}$) is the full gauge theory effective superpotential, including the effects of gauge dynamics, in a description in which the S_i can be treated as elementary fields.

Despite the similarities in the gauge theory and matrix model derivations, there are certainly some differences. As we have already noted, the crucial factorization of correlation functions is justified in the gauge theory using the properties of chiral operators, and in the matrix model using a large $\widehat{N}$ limit. Another difference is that in the gauge theory, the variables are naturally complex, but the parameters and eigenvalues of M naturally take real values. Furthermore, N_i and some other natural gauge theory quantities have no direct matrix model analogs.

Now that we have derived the relation to the matrix model, we are free to use this where it simplifies the arguments, to get concrete expressions for $\mathcal{F}(\mathcal{S}, g)$, and so forth, as we will do in section 5.

The deepest point where the matrix model produces a stronger result than our gauge theory considerations so far is in predicting the coupling-independent factor in $\mathcal{F}(S, g)$. It is easy to compute this, in the matrix model, by taking special values for the couplings. For example, for the case of unbroken gauge symmetry, if we take the superpotential $W = m\Phi^2/2$, we have, following Dijkgraaf and Vafa,

$$\begin{aligned}
\exp\left(-\frac{\widehat{N}^2 F_{m.m.}(S,m)}{g_m^2}\right) &= \int \frac{d^{\widehat{N}^2} M}{\mu^{\widehat{N}^2}} \exp\left(-\frac{\widehat{N} m}{2 g_m} \mathrm{Tr}\, M^2\right) \\
&= \left(\frac{2\pi g_m}{\widehat{N} m \mu^2}\right)^{\widehat{N}^2/2}
\end{aligned}$$

where the constant μ with dimensions of mass was introduced in order to keep the measure dimensionless. From (4.32), $S = g_m$ so

$$F_{m.m.}(S,m) = \frac{1}{2} S^2 \log\left(\frac{m\mu^2 \widehat{N}}{2\pi S}\right).$$

Using (4.34) and inserting this in (2.42), the m dependence agrees with (2.17), which is the simplest check of the general claim. More impressively, this determines the integration constant as well. Identifying $\widehat{N}\mu^2/2\pi$ with the gauge theory cutoff as $e^{3/2}\Lambda_0^2$ (we absorbed the $1/\Lambda_0^N$ from (2.17) here) and again using (2.42), we find precisely (1.10).

4.3 Determining the full effective superpotential

There are at least two ways one can interpret the final result. One, which we have emphasized in the introduction as being conceptually more straightforward, is to consider it as an effective lagrangian to be used in the low energy gauge dynamics. Our arguments seem adequate for proving the result with this conservative interpretation.

A more ambitious interpretation of the final result is to regard it as an ingredient in a complete "effective superpotential," obtained by integrating out both matter and gauge fields in a description in which the $\widehat{S}_i$ are treated as elementary fields. Despite our caveats in the introduction, one can certainly make a very simple suggestion for how to determine this. It is that the Ward identities and derivation we just discussed apply to the full gauge theory functional integral.[10] After all, these identities did not explicitly depend on the mysterious quantities S_i and w_i; rather those quantities emerged as undetermined parameters in the solution.

If we accept this idea, then the remaining ambiguity in the full effective superpotential is a coupling-independent function of the variables S_i denoted $\mathcal{H}(S_i)$ in (4.34). Since it is coupling independent, we can compute it in any convenient limit. In particular, we can consider the limit $W \gg \Lambda_i^3$ in which the gauge coupling is negligible at the scale of gauge symmetry breaking. In this limit, the gauge dynamics reduces to that of $\prod_i \mathrm{U}(N_i)$ super Yang-Mills theory. And we know the contribution of this gauge theory to the effective superpotential — it is the sum of Veneziano-Yankielowicz superpotentials (1.10) for each factor

$$\begin{aligned} \mathcal{H}(S_i) &= \sum_i N_i H(S_i) \\ \frac{\partial H(S_i)}{\partial S_i} &= S_i(1 - \ln S_i) \end{aligned} \tag{4.35}$$

Thus, we might conclude that the complete effective superpotential has the constant of integration fixed

$$W_{\text{eff}}(S_i, g_k) = 2\pi i\tau \sum_i S_i + \sum_i N_i \frac{\partial \mathcal{F}(S_i, g_k)}{\partial S_i} + \frac{1}{2}\sum_{ij} \frac{\partial^2 \mathcal{F}(S_i, g_k)}{\partial S_i \partial S_j} w_{\alpha i} w_j^\alpha \,. \tag{4.36}$$

Furthermore, since the shift (2.38) is a symmetry of the full gauge theory, the full effective action must take the form (2.42) with

$$\mathcal{F}(S_i) = F_{m.m.}(S_i)\,, \tag{4.37}$$

[10] As discussed in section 3, the nonrenormalization arguments so far leave room for Λ-dependent corrections. We suggested there that this might not be the case.

with even the coupling independent S_i dependence reproduced, as conjectured by Dijkgraaf and Vafa. We believe this conjecture is correct and the observations we just made are certainly suggestive.

We point out that since $\mathcal{H}$ is written as a sum of terms each depending only on a single $S_i = \widehat{S}_i - \frac{1}{2N} w_{\alpha i} w_i^\alpha$, its contribution to the effective superpotential depends only on the $\mathrm{SU}(N_i)$ field $\widehat{S}_i$.

There are (at least) two more points one would want to understand to feel that one indeed had a complete proof of the conjecture. One is to explain why the final step in finding the physical solution of the gauge theory is to minimize this effective superpotential. Now on general grounds, if one can convince oneself that the variables S_i and $w_{i\alpha}$ are fields in an effective lagrangian, this would clearly be a correct procedure. Now their origin was the parameters determining a particular solution of the Ward identities (4.14), and one might argue that, since one can imagine configurations in which this choice varies in space and time, the parameters must be fields.

It would be more satisfying to have a better microscopic understanding of the parameters to justify this argument. This brings us back to the other point mentioned in the introduction — we have no *a priori* argument that the solution should be representable as an effective lagrangian depending on such fields at all, and we know of very similar gauge theory problems in which this is not the case.

So the best we can say with these arguments is that assuming that the solution of the gauge theory takes the general form (1.2), we have demonstrated that the fields and effective lagrangian are as in (4.36).

Relation to special geometry. In section 2 we noted that the fields S_i, which as we discussed can be regarded as coordinates on the space of gauge backgrounds, satisfy many the defining properties of the "special coordinates" which appear in $\mathcal{N}=2$ supersymmetry (e.g. see [28] for a discussion). In particular they include gauge field strengths (the $w_{\alpha i}$) normalized so that the charge quantization condition is independent of S_i.

Another observation of [6] (stemming from previous work on geometric engineering of these models) is that the matrix model free energy $F_{m.m.}(S, g)$, and thus the gauge theory quantity $F(\mathcal{S}, g)$, can be regarded as an $\mathcal{N}=2$ prepotential. Furthermore, the relations of special geometry can be seen to follow from matrix model relations.

In particular, the matrix model/loop equation definition (4.32) of the quantities S_i, as periods of non-intersecting cycles C_i of the Riemann surface Σ. Special geometry includes a conjugate relation for the periods of a set of non-intersecting conjugate cycles B_j whose intersection form satisfies $\langle C_i, B_j \rangle = \delta_{ij}$. This relation is

$$\frac{\partial F}{\partial S_i} = \oint_{B_i} R(z) dz \,. \tag{4.38}$$

In the matrix model, one can interpret this formula as the energy required to move an eigenvalue from the i'th cut to infinity. This striking interpretation suggests that the matrix model eigenvalues are more than just a convenient technical device. Although the special geometry relations can be proven in many other ways We do not know a direct gauge theory argument for this relation.

We end this section by commenting on the situation when some of the N_i vanish. In this case the corresponding S_i do not exist. This means that the Riemann surface degenerates and the cycles associated with these N_i are pinched. In section 5 we will discuss a simple example in which the Riemann surface degenerates to a sphere.

5. Examples

We now grant that supersymmetric vacua of the one adjoint theory are critical points of the effective superpotential (4.36), and study this solution explicitly. We consider two particular cases in which we have other information to compare with. In the first case the $\mathrm{U}(N)$ gauge group is maximally Higgsed and the low energy gauge group is $\mathrm{U}(1)^N$. In the second case the $\mathrm{U}(N)$ gauge group is maximally confining and the low energy gauge group is $\mathrm{U}(1)$. Of particular interest is to see how effects usually computed using instantons can emerge from the procedure of minimizing this effective superpotential.

5.1 Maximally higgsed vacua

In this subsection we consider the maximally Higgsed vacua in which the $\mathrm{U}(N)$ gauge symmetry is broken to $\mathrm{U}(1)^N$. As shown in [5], these vacua are obtained when the tree level superpotential is of degree $N+1$; i.e. $n = N$. The vacuum with $\langle\Phi\rangle$ with eigenvalues ϕ_i is obtained when

$$W'(z) = g_N \prod_i (z - \phi_i) \tag{5.1}$$

Since there is no strong dynamics in the IR, we expect all nonperturbative phenomena to be calculable using instantons. These are instantons in the broken part of the group and since they cannot grow in the IR they lead to finite and therefore meaningful contributions to the functional integral. As the parameters in the superpotential are varied the low energy degrees of freedom are the photons of $\mathrm{U}(1)^N$ and at some special points there are also massless monopoles (there can also be Argyres-Douglas points [29, 30]). It is important that for all values of the parameters the number of photons remains N and there is no confinement. Therefore, there is never any strong IR dynamics.

Let us consider the expectation value of $t_k = -\frac{1}{32\pi^2}\operatorname{Tr} W_\alpha^2 \Phi^k$. Since the light degrees of freedom are the N photons and at special points also some massless monopoles, there is no strong dynamics and no IR divergences. Therefore, there cannot be any singularities as a function of the parameters in the superpotential g_k, and they can be treated in perturbation theory. Therefore the contribution of s instantons is of the form

$$\sum_{r_1, r_2 \dots r_n} a_{r_1, r_2 \dots r_n} \Lambda^{2Ns} g_0^{r_0} g_1^{r_1} \cdots g_n^{r_n} \tag{5.2}$$

with coefficients $a_{r_1,r_2\dots r_n}$, which are independent of the various coupling constants. Further constraints follow from the $\mathrm{U}(1)_\Phi \times \mathrm{U}(1)_\theta$ symmetry of (2.23) and holomorphy. The t_k have $\mathrm{U}(1)_\Phi \times \mathrm{U}(1)_\theta$ charges $(k, 2)$. This leads to the selection rules $2Ns - \sum_l r_l(l+1) = k$ and $\sum_l 2r_l = 2$. The only solution for $k \le N-1$ is $s = 1$, $l = N$, $k = N-1$. Therefore,

$$\langle t_k \rangle = \begin{cases} 0 & 0 \le k \le N-2 \\ c(N) g_N \Lambda^{2N} & k = N-1 \end{cases} \tag{5.3}$$

and $\langle t_{N-1}\rangle$ is given exactly by the contribution of one instanton, and it is first order in g_N. The coefficient $c(N)$ can be computed explicitly, but we will not do it here. Note that in particular, $\langle S\rangle = \langle t_0\rangle = 0$.

Let us compare these conclusions with the matrix model. To find the matrix model curve from the $\mathcal{N}=2$ curve, which is $y^2 = P_N(x)^2 - 4\Lambda^{2N}$, we are supposed to factor the right hand side as $P_N(x)^2 - 4\Lambda^{2N} = Q(x)^2((W'(x))^2 + f))$, where the double roots are contained in $Q(x)^2$. In the present case, as the unbroken group has rank N, there are no double roots, so $Q = 1/g_N$, $W' = g_N P_N$, and $f = -4g_n^2\Lambda^{2N}$. So for $0 \leq k \leq N-1$, we find

$$\begin{aligned}\langle t_k\rangle &= -\frac{1}{4\pi i}\oint_{\mathcal{C}} z^k y(z)dz = \frac{g_N^2\Lambda^{2N}}{2\pi i}\oint_{\mathcal{C}}\frac{z^k}{W'(z)}dz \\ &= \begin{cases} 0 & 0\leq k\leq N-2 \\ g_N\Lambda^{2N} & k = N-1\end{cases}\end{aligned} \tag{5.4}$$

which agrees with the expression of $\langle t_k\rangle$ found in (5.3) for the constant $c(N) = 1$. In deriving (5.4), we made the expansion $y = \sqrt{(W')^2 - 4\Lambda^{2N}} = W' - 2\Lambda^{2N}/W' + \ldots$; in this expansion, only the term Λ^{2N}/W' contributes to the integral.

We mentioned in section 2.2 that also for $N_i = 1$ we should still use $S_i = \widehat{S}_i - \frac{1}{2}w_{\alpha i}w^{\alpha i}$ with a scalar field $\widehat{S}_i$. Classically $\widehat{S}_i = 0$ but quantum mechanically it can be nonzero. Using contour integrals we see that $\langle S_i\rangle = \langle\widehat{S}_i\rangle$ are given by a power series in Λ^{2N}; i.e. they are given by an instanton sum. This is a satisfying result because in this case there is no strong IR dynamics and we do not expect "gluino condensation in the unbroken SU(1)." Instead the nonzero $\langle S_i\rangle$ arise from nonperturbative high energy effects; i.e. instantons.

In appendix A we further study these vacua of maximally broken gauge symmetry and use our Ward identities to determine the instanton corrections to $\langle \mathrm{Tr}\,\Phi^k\rangle$ in the $\mathcal{N}=2$ theory. These can be interpreted as a nonperturbative deformation of the $\mathcal{N}=2$ chiral ring.

5.2 Unbroken gauge group

In this subsection we consider the special vacua in which the U(N) gauge group is unbroken, and at very low energy only the photon of U(1) $\subset$ U(N) is massless. These vacua were also discussed recently in [31, 32].

Strong gauge coupling analysis. Following [4] we start our discussion by neglecting the tree level superpotential

$$W = \sum_{p=0}^{n+1}\frac{g_p}{p+1}\,\mathrm{Tr}\,\Phi^{p+1} \tag{5.5}$$

Now the theory has $\mathcal{N}=2$ supersymmetry. The vacua we are interested in originate from points in the $\mathcal{N}=2$ moduli space with $N-1$ massless monopoles. Turning on the tree level superpotential (5.5) leads to the condensation of these monopoles and leaves only a single massless photon. At these points in the moduli space the hyper-elliptic curve which determines the low energy dynamics factorizes as [33]

$$y^2 = P_N(z)^2 - 4\Lambda^{2N} = \left((z-z_0)^2 - 4\Lambda^2\right)H_{N-1}^2(z-z_0)\,; \qquad P_N(z) = \det(z-\Phi) \tag{5.6}$$

The constant z_0 is present because unlike [33], we consider the gauge group $\mathrm{U}(N)$ rather than $\mathrm{SU}(N)$ and therefore the matrix Φ is not traceless. This factorization problem is solved by Chebyshev polynomials and determines the locations of these points in the moduli space of the $\mathcal{N}=2$ theory at points where the eigenvalues of Φ are

$$\phi_j = z_0 + \widehat{\phi}_j = z_0 + 2\Lambda\cos\frac{\pi(j-1/2)}{N}; \qquad j=1,\ldots,N \tag{5.7}$$

or related to it by $\Lambda \to e^{2\pi i/2N}\Lambda$ (this leads to N such points). Here $\widehat{\Phi}$ is a traceless matrix in the adjoint representation of $\mathrm{SU}(N)$. Naively the expectation value $\langle \mathrm{Tr}\,\widehat{\Phi}^k\rangle$ is given by the sum

$$\sum_{i=1}^{N}\left(2\Lambda\cos\frac{\pi(i-1/2)}{N}\right)^k = \\ N\Lambda^k\begin{cases} 0 & k \text{ odd} \\ \binom{k}{\frac{k}{2}} & k \text{ even, } 0\le k<2N \\ \binom{k}{\frac{k}{2}} - 2\binom{k}{\frac{k}{2}+N} & k \text{ even, } 2N\le k<4N \\ \binom{k}{\frac{k}{2}} - 2\binom{k}{\frac{k}{2}+N} + 2\binom{k}{\frac{k}{2}+2N} & k \text{ even, } 4N\le k<6N \end{cases} \tag{5.8}$$

However, in appendix A we argue that $\langle \mathrm{Tr}\,\widehat{\Phi}^k\rangle$ is corrected by instantons and is not given by the sum (5.8). Instead

$$\langle \mathrm{Tr}\,\Phi^k\rangle = \frac{1}{2\pi i}\oint_C z^k \frac{P_N'(z)}{\sqrt{P_N(z)^2-4\Lambda^{2N}}}dz \tag{5.9}$$

with P_N of (5.6), which leads to

$$\langle \mathrm{Tr}\,\widehat{\Phi}^k\rangle = \begin{cases} 0 & k \text{ odd} \\ N\Lambda^k\binom{k}{\frac{k}{2}} & k \text{ even} \end{cases} \tag{5.10}$$

for all k.

Using (5.10) we find an effective superpotential for z_0

$$W_{\text{eff}}(z_0) = \langle W\rangle = N\sum_{l=0}^{\left[\frac{n+1}{2}\right]}\frac{\Lambda^{2l}}{(l!)^2}W^{(2l)}(z_0) \tag{5.11}$$

The field S can be integrated in by performing a Legendre transform with respect to $2N\log\Lambda$. One way to do that is, as in (1.9) by considering

$$W_{\text{eff}}(z_0, C, S) = N\sum_{l=0}^{\left[\frac{n+1}{2}\right]}\frac{C^{2l}}{(l!)^2}W^{(2l)}(z_0) + 2NS\log\frac{\Lambda}{C} \tag{5.12}$$

Integrating S and C out of $W_{\rm eff}(z_0,C,S)$ we recover $W_{\rm eff}(z_0)$ of (5.11). Alternatively, we can integrate out C to find $W_{\rm eff}(z_0,S)$. It is obtained from $W_{\rm eff}(z_0,C,S)$ by substituting C which solves

$$0 = \frac{C}{2N}\partial_C W_{\rm eff}(z_0,C,S) = \sum_{l=0}^{\left[\frac{n+1}{2}\right]} \frac{lC^{2l}}{(l!)^2} W^{(2l)}(z_0) - S \tag{5.13}$$

This equation can be interpreted as the Konishi anomaly (3.4) for the rotation of the field $\widehat{\Phi}$ by a phase. We can also integrate out z_0 to find $W_{\rm eff}(S)$ by solving its equation of motion

$$\partial_{z_0} W_{\rm eff}(z_0,C,S) = N \sum_{l=0}^{\left[\frac{n}{2}\right]} \frac{C^{2l}}{(l!)^2} W^{(2l+1)}(z_0) = 0 \tag{5.14}$$

and substituting its solution in $W_{\rm eff}(z_0,C,S)$.

It is interesting that the complexity of $W_{\rm eff}(S)$ as opposed to the simplicity of $W_{\rm eff}(z_0)$ arises because S was integrated in and z_0 was integrated out.

Weak gauge coupling analysis. We now study this vacuum in the limit of weak gauge coupling. This is the approach we used throughout most of this paper. Rather than first neglecting the tree level superpotential (5.5) and adding it later, we first neglect the gauge interactions. In the vacuum we are interested in we expand around Φ which is proportional to the unit matrix, $\Phi = z_0$. The traceless components of Φ acquire a mass $m(\widehat{\Phi}) = W''(z_0)$, and can be integrated out. We choose not to integrate out $\mathrm{Tr}\,\Phi = Nz_0$, even though its mass is also $W''(z_0)$. The remaining light degrees of freedom are the $\mathrm{U}(N)$ gauge fields. Their tree level effective lagrangian is

$$\begin{aligned} \frac{-i\tau}{16\pi}\,\mathrm{Tr}\,W_\alpha^2 &= 2\pi i\tau S \\ \tau &= \frac{\theta}{2\pi} + i\frac{4\pi}{g_{YM}^2} \end{aligned} \tag{5.15}$$

At one loop order $2\pi i\tau$ is replaced with

$$N\log\Lambda(z_0)^3 = N\log\left(\Lambda^2 W''(z_0)\right) \tag{5.16}$$

which is derived by the threshold matching with the masses of the massive particles.

Consider the effective lagrangian at energies above $\Lambda(z_0)$ but below $W''(z_0)$, which is obtained by integrating out the adjoint massive fields $\widehat{\Phi}$. The leading order F-terms are

$$NW(z_0) + 3NS\log\Lambda(z_0) + \text{higher order terms} \tag{5.17}$$

with $\Lambda(z_0)$ of (5.16). The higher order terms include higher powers of $W_\alpha W^\alpha$. The discussion of section 2 shows that the F terms are only functions of S; i.e. there are no terms with a trace of more than two W_α, or terms with different contractions of the Lorentz indices. Furthermore, terms of order S^k arise only from diagrams with k loops.

Integrating z_0 out of (5.17) leads to more higher order terms in S. They correspond to diagrams in the high energy theory which are not 1PI with respect to z_0. Note that these terms satisfy the nonrenormalization theorem of section 2 and generate S^k by diagrams with k loops.

As in the rest of the paper we integrate out the strong SU(N) dynamics by adding to (5.17) the Veneziano-Yankielowicz term $S(1 - \log S)$. We find the full effective superpotential of z_0 and S

$$W_{\text{eff}}(z_0, S) = N\left[W(z_0) + S\left[\log\left(\frac{\Lambda^2 W''(z_0)}{S}\right) + 1\right] + \text{higher order terms}\right] \tag{5.18}$$

We can compare this expression with the strong gauge coupling analysis we did earlier in this subsection. If we neglect the higher order terms in (5.18) and integrate out S, we find

$$\frac{1}{N}W_{\text{eff}}(z_0) = W(z_0) + \Lambda^2 W''(z_0) + \mathcal{O}(\Lambda^4) \tag{5.19}$$

which agrees with (5.11). We can interpret it as follows. The first term is due to the tree level superpotential. The second term $\Lambda^2 W''(z_0) = \Lambda(z_0)^3$ is due to gluino condensation in the low energy unbroken group. Higher order terms in (5.11) can then be interpreted as powers of gluino condensation.

It follows from (5.11) that if the tree level superpotential is cubic, there are no higher order corrections in (5.19), and therefore there are no higher order corrections in (5.18). This means that in this case (5.18) is given exactly by the one loop approximation. If we integrate z_0 out of (5.18) to find an effective lagrangian for S, the term S^k arises from diagrams with k loops, which are connected by z_0 lines and are not 1PI. We conclude that in this case of a cubic superpotential with a maximally confining group only a very small subset of the k loop diagrams contribute to the term S^k.

Matrix model computations. In this special case the analysis of the matrix model simplifies considerably. As mentioned at the end of section 4, when some of the N_i vanish the Riemann surface degenerates. In our case of only one nonzero N_i it degenerates to a sphere. Only one of the n cuts of the curve

$$y^2 = W'(z)^2 + f(z) \tag{5.20}$$

exists and the other $n-1$ cuts shrink to zero size; i.e. f splits only one of the double zeros of $W'(z)^2$ and shifts the others. Therefore, equation (5.20) must factorize as

$$y^2 = W'(z)^2 + f(z) = \left((z - z_0)^2 - 4C^2\right) Q(z)^2 \tag{5.21}$$

with Q a polynomial of degree $n-1$. z_0 and C will soon be identified with the objects denoted by the same symbols above. It is easy to find the polynomial Q as

$$Q(z) = \left[\frac{W'(z)}{\sqrt{(z - z_0)^2 - 4C^2}}\right]_+ \tag{5.22}$$

We first solve

$$W'(z)^2 = \left((z-z_0)^2 - 4C^2\right)\widehat{Q}(z)^2 \tag{5.23}$$

for $\widehat{Q}(z)$ and expand it around z_0

$$\widehat{Q}(z) = \frac{W'(z)}{\sqrt{(z-z_0)^2-4C^2}} = \sum_{k=0}^{n}\sum_{l=0}^{\infty}\frac{C^{2l}}{k!}\binom{2l}{l}W^{(k+1)}(z_0)(z-z_0)^{k-2l-1} \tag{5.24}$$

where we used $\frac{1}{\sqrt{1-4x}} = \sum_{l=0}^{\infty}\binom{2l}{l}x^l$. Q in (5.21) should be a polynomial and therefore

$$Q(z) = \sum_{k=1}^{n}\sum_{l=0}^{\left[\frac{k-1}{2}\right]}\frac{C^{2l}}{k!}\binom{2l}{l}W^{(k+1)}(z_0)(z-z_0)^{k-2l-1} \tag{5.25}$$

Substituting this Q in (5.21) it is easy to see (using (5.24)) that it is satisfied with f a polynomial of degree n. The order z^n term in equation (5.21) is cancelled when the coefficient of $(z-z_0)^{-1}$ in $\widehat{Q}$ vanishes. Therefore, equation (5.21) can be satisfied only when

$$\sum_{l=0}^{\left[\frac{n}{2}\right]}\frac{C^{2l}}{(l!)^2}W^{(2l+1)}(z_0) = 0 \tag{5.26}$$

Furthermore, the coefficient of z^{n-1} in f_{n-1} is determined by the coefficient of $(z-z_0)^{-2}$ in $\widehat{Q}$

$$f_{n-1} = -\frac{4W^{(n+1)}}{n!}\sum_{l=1}^{\left[\frac{n+1}{2}\right]}\frac{lC^{2l}}{(l!)^2}W^{(2l)}(z_0) \tag{5.27}$$

We can now compare our field theory results with this discussion. The relation of z_0 and C (5.26) is identical to the equation of motion of z_0 (5.14). S is determined in terms of f_{n-1} [5]

$$S = -\frac{n!}{4W^{(n+1)}}f_{n-1} = \sum_{l=1}^{\left[\frac{n+1}{2}\right]}\frac{lC^{2l}}{(l!)^2}W^{(2l)}(z_0) \tag{5.28}$$

which is the same as the equation of motion of C (5.13). This confirms the identification of z_0 and C with the objects denoted by the same symbols above.

As we reviewed at the end of section 4, the effective superpotential is determined by the period integral (4.38)

$$-2\pi i\Pi = -\int_{z_0+2C}^{\Lambda_0} y dz \tag{5.29}$$

In appendix C we perform this integral and find

$$\begin{aligned} -2\pi i\Pi &= -\int_{z_0+2C}^{\Lambda_0} y dz \\ &= -W(\Lambda_0) - 2S\log\left(\frac{\Lambda_0}{C}\right) + \sum_{k=0}^{\left[\frac{n+1}{2}\right]}\frac{C^{2k}}{(k!)^2}W^{(2k)}(z_0) + \mathcal{O}\left(\frac{1}{\Lambda_0}\right) \end{aligned} \tag{5.30}$$

JHEP12(2002)071

The first term is independent of the fields and can be ignored. The second and third terms (after renormalizing the bare gauge coupling) are as in (5.12). We conclude that as $\Lambda \to \infty$ we recover the superpotential (5.12) with $C(z_0)$ determined by (5.26) and $z_0(S)$ determined by (5.28).

Using the explicit expression for y it is straightforward to compute

$$\begin{aligned}\langle t_m \rangle &= -\frac{1}{32\pi^2}\left\langle \mathrm{Tr}\, W_\alpha^2 \Phi^m \right\rangle = -\frac{1}{4\pi i}\oint_{\mathcal{C}} z^m y dz = -\frac{1}{4\pi i}\oint_{\mathcal{C}} z^m \sqrt{(z-z_0)^2 - 4C^2}Q(z)dz \\ &= \sum_{\substack{k=1\\ k\ odd}}^{n} \sum_{r=0}^{\left[\frac{m}{2}\right]} \frac{2m!}{(2r+k+1)(r!)^2\left(\left(\frac{k-1}{2}\right)!\right)^2(m-2r)!} z_0^{m-2r} W^{(k+1)}(z_0) C^{2r+k+1} + \\ &\quad + \sum_{\substack{k=1\\ k\ even}}^{n} \sum_{r=0}^{\left[\frac{m-1}{2}\right]} \frac{2m!}{(2r+k+2)(\frac{k}{2})!(\frac{k-2}{2})!r!(r+1)!(m-2r-1)!} \times \\ &\quad \times z_0^{m-2r-1} W^{(k+1)}(z_0) C^{2r+k+2}\end{aligned} \tag{5.31}$$

Using this expression for $\langle t_m \rangle$ and the expectation values

$$\begin{aligned}\langle \mathrm{Tr}\,\Phi^k \rangle &= N \sum_{p=0}^{\left[\frac{k}{2}\right]} \frac{k!}{(k-2p)!p!^2} C^{2p} z_0^{k-2p} \\ \langle \mathrm{Tr}\,\Phi^{m+1} W'(\Phi) \rangle &= N \sum_{k=0}^{m+1} \sum_{\substack{l=0\\ k+l \text{ even}}}^{n} \frac{1}{l!} \binom{k+l}{(k+l)/2} \binom{m+1}{k} z_0^{m+1-k} W^{(l+1)}(z_0) C^{k+l}\end{aligned} \tag{5.32}$$

it is straightforward to explicitly check the generalized anomaly identity

$$\langle \mathrm{Tr}\,\Phi^{m+1} W'(\Phi) \rangle = 2\sum_{l=0}^{m} \langle \mathrm{Tr}\,\Phi^{m-l} \rangle \langle t_l \rangle \tag{5.33}$$

or equivalently

$$\left\langle \mathrm{Tr}\,\frac{W'(\Phi)}{z-\Phi} \right\rangle = 2\left\langle \mathrm{Tr}\,\frac{1}{z-\Phi} \right\rangle \left\langle -\frac{1}{32\pi^2}\,\mathrm{Tr}\,\frac{W_\alpha^2}{z-\Phi} \right\rangle . \tag{5.34}$$

Instead of checking (5.33) or (5.34) explicitly, we can proceed as follows. We use the matrix model loop equation

$$\left(-\frac{1}{4\pi i}\oint_{\mathcal{C}} \frac{y(z)}{z-x} dz\right)^2 = -\frac{1}{4\pi i}\oint_{\mathcal{C}} \frac{W'(z)y(z)}{z-x} dz \tag{5.35}$$

where $\mathcal{C}$ is a large circle at infinity and our identification

$$-\frac{1}{32\pi^2}\left\langle \mathrm{Tr}\,\frac{W_\alpha^2}{z-\Phi} \right\rangle = -\frac{1}{4\pi i}\oint_{\mathcal{C}} \frac{y(x)}{z-x} dx \tag{5.36}$$

to show that the values obtained in the field theory analysis for $\langle \mathrm{Tr}\,\Phi^k \rangle$ (5.32) or $\left\langle \mathrm{Tr}\,\frac{1}{z-\Phi} \right\rangle$ satisfy (5.34).

JHEP12(2002)071

Equations (5.32) can be written as[11]

$$\begin{aligned}\langle \mathrm{Tr}\,\Phi^k\rangle &= \frac{N}{2\pi i}\oint_{\mathcal{C}} \frac{z^k}{\sqrt{(z-z_0)^2-4C^2}}dz \\ \langle \mathrm{Tr}\,\Phi^k W'(\Phi)\rangle &= \frac{N}{2\pi i}\oint_{\mathcal{C}} \frac{z^k W'(z)}{\sqrt{(z-z_0)^2-4C^2}}dz\end{aligned} \tag{5.37}$$

or equivalently

$$\begin{aligned}\left\langle \mathrm{Tr}\,\frac{1}{z-\Phi}\right\rangle &= \frac{N}{2\pi i}\oint_{\mathcal{C}} \frac{1}{(z-x)\sqrt{(x-z_0)^2-4C^2}}dx \\ \left\langle \mathrm{Tr}\,\frac{W'(\Phi)}{z-\Phi}\right\rangle &= \frac{N}{2\pi i}\oint_{\mathcal{C}} \frac{W'(x)}{(z-x)\sqrt{(x-z_0)^2-4C^2}}dx\end{aligned} \tag{5.38}$$

In appendix C we show that for the y of (5.21),

$$\frac{\partial y}{\partial S} = -\frac{2}{\sqrt{(z-z_0)^2-4C^2}}\,. \tag{5.39}$$

and therefore (5.37) can be written as

$$\begin{aligned}\left\langle \mathrm{Tr}\,\frac{1}{z-\Phi}\right\rangle &= -\frac{N}{\pi i}\frac{\partial}{\partial S}\oint_{\mathcal{C}} \frac{y(x)}{z-x}dx \\ \left\langle \mathrm{Tr}\,\frac{W'(\Phi)}{z-\Phi}\right\rangle &= -\frac{N}{\pi i}\frac{\partial}{\partial S}\oint_{\mathcal{C}} \frac{W'(x)y(x)}{z-x}dx\end{aligned} \tag{5.40}$$

The derivative of (5.35) with respect to S is

$$2\left(\frac{1}{4\pi i}\frac{\partial}{\partial S}\oint_{\mathcal{C}}\frac{y(z)}{z-x}dz\right)\frac{1}{4\pi i}\oint_{\mathcal{C}}\frac{y(z)}{z-x}dz = -\frac{1}{4\pi i}\frac{\partial}{\partial S}\oint_{\mathcal{C}}\frac{W'(z)y(z)}{z-x}dz \tag{5.41}$$

Now equations (5.36)(5.40) leads to the desired result (5.34).

6. Conclusions

Following the ideas of Dijkgraaf and Vafa, and specifically their observation that the superpotential in $\mathcal{N}=1$ gauge theory arises entirely from the planar diagrams of an associated bosonic matrix model, we were able to derive a complete solution for the general $\mathcal{N}=1$ gauge theory with a single adjoint chiral multiplet, confirming their conjectured solution.

Besides confirming their results, we have provided an explanation for the conjecture, which allowed us to make a nonperturbative proof, and see clearly how far it should generalize. The reduction to planar diagrams is analogous to the explanation of the reduction of perturbation theory to planar diagrams in the 't Hooft large-N limit, but it is essentially different. In particular it happens for finite N.

[11] From differentiating (5.6), H_{N-1} divides $P_N P'_N$. However, (5.6) shows that P_N and H_{N-1} cannot have the same root and therefore, P'_N has a factor of H_{N-1}. Since they are polynomials of the same degree, they are proportional to each other. The proportionality factor can be determined by matching the highest power in the polynomial, and therefore $P'_N = NH_{N-1}$. Finally, the equations in appendix A show that we need $\frac{P'_N}{\sqrt{P_N^2-4\Lambda^{2N}}} = \frac{N}{\sqrt{(z-z_0)^2-4C^2}}$.

We can identify four essential ingredients in this proof: the use of the complete chiral ring, the generalized Konishi anomaly, factorization for vacuum expectation values of chiral operators, and the correct gauge-invariant identification of the gaugino condensates. Let us discuss these points in turn.

First of all, we made essential use of the complete chiral ring in the gauge theory, which we showed in the single adjoint theory consists of $\mathrm{Tr}\,\Phi^n W_\alpha^k$ for all n and all $k \leq 2$. This explicit description of the chiral ring shows up in some previous work (for example on the AdS/CFT correspondence [34]) but was not systematically exploited in solving the theories.

One reason for this is that in U(N) gauge theory at any finite N, one expects that all of the operators with $k > N$ can be written in terms of a finite $\mathcal{O}(N)$ subset of them, using operator relations which are simple deformations of those in the classical theory. We illustrated this point and its power by using it to show that in pure super Yang-Mills theory, every vacuum has a non-zero gaugino condensate. One might think that knowing these finite N relations would be essential in solving the theory, but in the end we were able to do it using only a subset of relations which survive the large-N limit.

The description of the chiral ring is simplified by making use of a nonlinearly realized $\mathcal{N} = 2$ supersymmetry, which is simply the shift of the gaugino for the decoupled diagonal U(1) subgroup of the gauge group. This symmetry leads to many parallels between the subsequent discussion and the discussion of gauge theory with $\mathcal{N} = 2$ supersymmetry. On first sight, these parallels might have led one to think that the Dijkgraaf-Vafa conjectures would only apply to theories (such as our example) with an explicitly broken $\mathcal{N} = 2$ supersymmetry. However, this particular nonlinearly realized $\mathcal{N} = 2$ is present in any gauge theory with a decoupled U(1). Indeed, our proof of the conjecture applies to a very large class of $\mathcal{N} = 1$ gauge theories, including chiral theories.

Our second ingredient was a generalization of the Konishi anomaly, an anomaly in the symmetry $\delta\Phi = \epsilon\Phi$, to the general reparameterization of the chiral ring $\delta\Phi = f(\Phi, W_\alpha)$. The analogous algebra of reparameterizations of a bosonic matrix integral leads to the Virasoro constraints and is thus central both to many methods for its solution and to its connection to topological string theory. In gauge theory, these considerations lead to a $\mathcal{N} = 2$ super-Virasoro algebra of constraints.[12]

Perhaps the simplest consequence of the generalized Konishi anomaly is a closed set of equations for expectation values of the "generalized gaugino condensate," the operator $\mathrm{Tr}\,W_\alpha^2/(z - \Phi)$. This leads to the identification of this operator with the matrix model resolvent.

Our third ingredient was the factorization of gauge-invariant expectation values for chiral operators. This is perhaps the most striking point where the discussion for the matrix model and the gauge theory, while formally parallel, in fact rest on rather different physics.

[12]Making this point explicit requires introducing couplings for all of the chiral operators. We did not need this for the rest of our arguments, and in fact giving finite values to these couplings is problematic because then the kinetic term of the scalar fields is not positive definite, but it is a nice formal way to understand the constraints.

Let us explain this point. In both cases, the key observation is factorization of correlation functions of gauge-invariant operators. This allows rewriting the Schwinger-Dyson equations, which generally relate n-point functions to $n+1$-point functions, as a closed system of equations for the one-point functions. The resulting equations can be expanded in the couplings and seen to generate planar perturbation theory. However, if their derivation is valid nonperturbatively, then so are the equations. In this case, one wants a more conceptual description of the simplification which does not rely on planarity of Feynman diagrams for its formulation.

In the large-N limit, the deeper reason for factorization is the existence of a "master field," a single field configuration which dominates the functional integral in this limit. It can be determined by solving a corrected "equation of motion," derived from an effective action which is the sum of the classical action with a "quantum entropy" term which is the logarithm of the volume of the gauge orbit of a configuration. Both terms scale as $\mathcal{O}(N^2)$ and thus one can think of the master field as a saddle point dominating the functional integral. This description is nonperturbative and can be used to exhibit many phenomena (such as certain large-N transitions) which have no diagrammatic interpretation.

In supersymmetric gauge theory, the deeper reason for factorization is the position independence of chiral operators combined with cluster decomposition. There is no obvious sense in which a single field configuration dominates the functional integral. Instead, the correlation function of chiral operators is given by a single state, the vacuum, in all possible channels. Yet the formal conclusions are the same, a system of factorized Ward identities which determine all expectation values of chiral operators.

Finally, the last ingredient in the proof was to identify the undetermined parameters in the general solution of the Ward identities, expressed in a correct set of coordinates, as fields in an effective lagrangian. This is one of the points at which previous attempts to understand generalizations of the Veneziano-Yankielowicz effective lagrangian ran into difficulties, so let us also explain this point further.

It was long realized that in simple limits (e.g. when the underlying gauge symmetry is broken to a product of nonabelian subgroups at very high energy and thus very weak couping), the physics of the theory under discussion would be well described by a sum of a Veneziano-Yankielowicz effective superpotential in each unbroken subgroup, each depending on a gaugino condensate for its subgroup. The problem with making sense of this observation more generally is that there is no obvious way to make this definition gauge invariant, as quantum fluctuations do not preserve the simple picture of subgroups sitting in the complete gauge group in a fixed way. Because of this, it is not *a priori* obvious that these variables have any meaning in the strong coupling regime.

The only gauge invariant way to pick out subgroups of the overall gauge group is to use the expectation values of the matter fields to do it. So, one can try to write an operator such as "$\mathrm{Tr}\, W_\alpha^2 \delta(\Phi - a_i)$" to pick out the gaugino condensate associated to the vacuum $W'(a_i) = 0$. However, this definition would also appear to behave in a very complicated way under adding quantum fluctuations.

The resolution of this problem is rather striking. Because of holomorphy, one can replace the $\delta(\Phi - a_i)$ of the previous definition with a contour integral $\oint dz/(z-\Phi)$ where the contour surrounds the classical vacuum a_i. Note that this object involves all powers of Φ. This is manifestly independent of infinitesimal quantum fluctuations of Φ and thus provides a good definition to all orders in perturbation theory. However it is not *a priori* obvious that it provides a good nonperturbative definition.

Of course, in the explicit solution, these poles turn into cuts at finite coupling, so by enlarging the contours to surround the cuts one does obtain a good nonperturbative definition. However this point could not have been taken for granted but actually rests on the previous discovery that computations in the chiral ring reduce to planar diagrams. Let us explain this point.

An important general feature of the sum over planar diagrams is that, while the total number of diagrams at a given order V in the couplings g (and thus weighed by g^V) grows as $V!$, the number of planar diagrams only grows exponentially in V. Furthermore, since we are integrating out massive fields and obtaining cutoff-independent results each diagram makes a finite contribution. This means that, in contrast to the typical behavior of perturbation theory, the perturbative expansion for the effective matter superpotential has a finite radius of convergence.

This makes it sensible to claim that this expansion converges to the exact result. Holomorphy precludes nonperturbative corrections because they have an essential singularity at the origin. Therefore the expansion must converge to the exact result.

Thus, perturbation theory is much more powerful than one might have expected, because of the reduction to planar diagrams combined with holomorphy. Indeed, the most striking aspect of the Dijkgraaf-Vafa conjecture is the possibility to derive effective actions which from other points of view involve infinite sums over instanton corrections, without any need to explicitly discuss instantons. On the other hand, even if one can derive correct results by summing perturbation theory, it is very useful to have a nondiagrammatic framework and derivation to properly justify and understand the results.

There are a number of fairly clear questions coming out of this work which we believe could be answered in the near future. A particularly important question is to better understand the chiral ring and the relations which hold at finite N. We discussed the exact quantum relations in some cases; having a simple and general description would significantly clarify the theory.

Our arguments generalize straightforwardly to arbitrary $\mathrm{U}(N)$ quiver theories, with the inhomogeneous $\mathcal{N}=2$ supersymmetry, loop equations, and so forth, including the general conjectured relation between these theories and the corresponding multi-matrix models. Although the list of matrix models we know how to solve exactly is rather limited and will probably stay that way, one can of course get perturbative results. There is also a general mathematical framework called free probability theory [35] which allows discussing these models nonperturbatively and may eventually be useful in this context.

It will be interesting to generalize these arguments to the other classical gauge groups. Since we do not rely on large N, it may be that even exceptional groups can be treated.

JHEP12(2002)071

Although we feel we have provided a more precise connection between supersymmetric gauge theory and matrix models than given in previous work, we have by no means understood all of the connections or even all of the aspects of Dijkgraaf and Vafa's work, from the point of view of gauge theory. We mentioned two particularly striking points in section 4; the fact that the matrix model naturally reproduces the complete effective superpotential coming from gauge theory, and the deeper aspects of the connection to $\mathcal{N}=2$ supersymmetry and special geometry. We believe the first point will obtain a purely gauge theoretic explanation, perhaps by considering additional anomalies.

Regarding the second point, given the way in which the solution for the adjoint theory is expressed using periods and a Riemann surface, one can make a simple proof of the relation to special geometry (expressed for example in (4.38)), but in itself this is not a very physical argument. A more physical argument for this relation in this example uses the relation to $\mathcal{N}=2$ super Yang-Mills and holomorphy. If this were the only correct argument, the relation might not be expected to hold in more general $\mathcal{N}=1$ theories. On the other hand, clearly many more $\mathcal{N}=1$ theories, including chiral theories, can be obtained by wrapping branes in Calabi-Yau compactification of string theory, a situation in which one might think special geometry would appear, so it could be that there are more general arguments.

Perhaps the largest question of which this question is a small part is to what extent this structure can be generalized to describe all of the $\mathcal{N}=1$ supersymmetric compactifications of string theory. Although the present rate of progress may be grounds for optimism, clearly we have much farther to go.

Acknowledgments

This work was supported in part by DOE grant #DE-FG02-96ER40959 to Rutgers, and DOE grant #DE-FG02-90ER40542 and NSF grant #NSF-PHY-0070928 to IAS.

A. Comments on the chiral ring of pure gauge $\mathcal{N}=2$ theories

Following [5] in this appendix we learn about the $\mathcal{N}=2$ theory by first perturbing it by a superpotential and then turning off the perturbation.

Let us recall some well known facts about the classical chiral ring of $\mathcal{N}=2$ theories. The Coulomb branch of the $\mathrm{U}(N)$ gauge theory with $\mathcal{N}=2$ supersymmetry can be parametrized by the eigenvalues ϕ_i of the $N\times N$ matrix Φ_{cl} (modulo permutations). Equivalently, in terms of the quantum field Φ we can take

$$\langle u_k\rangle = \langle \mathrm{Tr}\,\Phi^k\rangle = \mathrm{Tr}\,\Phi_{cl}^k = \sum_{i=1}^{N}\phi_i^k\,; \qquad k=1,\ldots,N\,. \tag{A.1}$$

as good coordinates on the Coulomb branch.

Classically it is clear that u_k generate the chiral ring of this theory. To see that, recall that Φ_{cl} is an $N \times N$ matrix and therefore for $l > N$ we can express the classical elements of the ring $\mathrm{Tr}\,\Phi_{cl}^l$ as polynomials in the classical generators $\mathrm{Tr}\,\Phi_{cl}^k$ with $k = 1, \ldots, N$

$$\mathrm{Tr}\,\Phi_{cl}^l = \mathcal{P}_l(\mathrm{Tr}\,\Phi_{cl}, \mathrm{Tr}\,\Phi_{cl}^2, \ldots, \mathrm{Tr}\,\Phi_{cl}^N) \tag{A.2}$$

These relations can be nicely summarized by introducing the characteristic polynomial of Φ_{cl}, $P_N(z, \Phi_{cl}) = \det\,(z - \Phi_{cl})$, which can be expressed as an explicit polynomial in $\mathrm{Tr}\ \Phi_{cl}^k$, $k \leq N$. In order not to clutter the equations, we will suppress the second argument Φ_{cl} in P_N. Then,

$$\mathrm{Tr}\,\frac{1}{z - \Phi_{cl}} = \frac{P_N'(z)}{P_N(z)}\,. \tag{A.3}$$

The desired relations can be found by expanding (A.3) in powers of $1/z$ or equivalently

$$\mathrm{Tr}\,\Phi_{cl}^l = \frac{1}{2\pi i}\oint_{\mathcal{C}} z^l\,\mathrm{Tr}\,\frac{1}{z - \Phi_{cl}}dz = \oint_{\mathcal{C}} z^l \frac{P_N'(z)}{P_N(z)}dz \tag{A.4}$$

where $\mathcal{C}$ is a large contour around $z = \infty$.

We expect the relations (A.2) to be modified by instantons. In order to determine these modifications we follow [5] and deform the $\mathcal{N} = 2$ theory to $\mathcal{N} = 1$ by adding a superpotential of degree $N + 1$, i.e. $n = N$. We take $W'(x) = g_N P_N(x)$ with $P_N(x) = \det(x - \Phi_{cl})$. As in section 5.1, we expand around the vacuum where the $\mathrm{U}(N)$ group is broken to $\mathrm{U}(1)^N$. This vacuum is uniquely characterized by $\langle u_k \rangle$ expressed in terms of Φ_{cl} as in (A.1).

We can solve the theory, as in section 4, using the Ward identities (4.16) (with w_α set to zero as we will be considering expectation values)

$$\begin{aligned} R(z)^2 &= g_N P_N(z) R(z) + \frac{1}{4} f(z) \\ 2T(z)R(z) &= g_N P_N(z) T(z) + \frac{1}{4} c(z)\,. \end{aligned} \tag{A.5}$$

The solution of these equations is

$$\begin{aligned} R(z) &= \frac{1}{2}\left(g_N P_N(z) - \sqrt{g_N^2 P_N^2(z) + f(z)}\right) \\ T(z) &= -\frac{c(z)}{4\sqrt{g_N^2 P_N^2(z) + f(z)}}\,. \end{aligned} \tag{A.6}$$

The polynomial $c(z)$ depends only on $\langle \mathrm{Tr}\,\Phi^k \rangle = \sum_i \phi_i^k$ for $k = 1, \ldots, N-1$

$$c(z) = 4\left\langle \mathrm{Tr}\,\frac{W'(\Phi) - W'(z)}{z - \Phi}\right\rangle = -4W''(z) = -4g_N P_N'(z) \tag{A.7}$$

In a vacuum with unbroken $\mathrm{U}(1)^N$ and $W' = g_N P_N$, $f(z) = -4g_N^2 \Lambda^{2N}$, as we explained in deriving (5.4). Therefore, the quantum modified version of (A.3) is

$$T(z) = \left\langle \mathrm{Tr}\,\frac{1}{z - \Phi}\right\rangle = \frac{P_N'(z)}{\sqrt{P_N^2(z) - 4\Lambda^{2N}}}\,. \tag{A.8}$$

Notice that the g_N dependence drops out of the equation and therefore we can take it to zero. We conclude that (A.8) is satisfied in the pure $\mathcal{N}=2$ theory.

Using (A.8) we can write,

$$\langle \mathrm{Tr}\, \Phi^l \rangle = \frac{1}{2\pi i}\oint_C z^l \frac{P_N'(z)}{\sqrt{P_N^2(z)-4\Lambda^{2N}}}dz = \sum_{m=0}^{\left[\frac{2N}{l}\right]} \binom{2m}{m} \Lambda^{2Nm} \frac{1}{2\pi i}\oint_C z^l \frac{P_N'(z)}{P_N(z)^{2m+1}}dz \quad (A.9)$$

The $m=0$ term is the classical formula (A.4), and the other terms are generated by m instantons. This equation was conjectured in [36] and has been used and explored in [37, 38].

Once we establish this expectation value we can derive the modifications of the relations in the chiral ring. We replace (A.2) with

$$\mathrm{Tr}\, \Phi^l = \mathcal{Q}_l(\mathrm{Tr}\, \Phi, \mathrm{Tr}\, \Phi^2, \ldots, \mathrm{Tr}\, \Phi^N, \Lambda^{2N}) \quad (A.10)$$

Let us compute the expectation value of this equation

$$\langle \mathrm{Tr}\, \Phi^l \rangle = \mathcal{Q}_l(\langle \mathrm{Tr}\, \Phi\rangle, \langle \mathrm{Tr}\, \Phi^2\rangle, \ldots, \langle \mathrm{Tr}\, \Phi^N\rangle, \Lambda^{2N}) = \mathcal{Q}_l(\mathrm{Tr}\, \Phi_{cl}, \mathrm{Tr}\, \Phi_{cl}^2, \ldots, \mathrm{Tr}\, \Phi_{cl}^N, \Lambda^{2N}) \quad (A.11)$$

Comparing with (A.9) we find the polynomials $\mathcal{Q}_l$. The important point is that once we establish that the ring is modified as in (A.10), we can use this equation not only in expectation values of the one point function but also in all correlation functions.

Before closing this appendix, we comment on the $\mathcal{N}=1$ chiral operators $\mathrm{Tr}\, W_\alpha \Phi^k$ and $\mathrm{Tr}\, W_\alpha W^\alpha \Phi^k$. They are related by the $\mathcal{N}=2$ algebra to the chiral operators $\mathrm{Tr}\, \Phi^k$. Hence, although they are $\mathcal{N}=1$ chiral, they are not chiral with respect to the $\mathcal{N}=2$ symmetry. Therefore, they play no role in the chiral ring of $\mathcal{N}=2$ theories.

B. The classical chiral ring in SU(N) with $\mathcal{N}=1$

In this appendix, filling in some details from section 2.2, we establish the classical ring relation $S^N=0$ for SU(N) gluodynamics, and we show that $S^{N-1}\neq 0$.

We start with the following identity, which holds for any $N\times N$ matrix M:

$$\det M = \frac{1}{N!}(\mathrm{Tr}\, M)^N + \cdots . \quad (B.1)$$

Here the ellipses are a sum of products of traces of M, with each term in the sum being proportional to at least one factor of $\mathrm{Tr}\, M^k$ with some $k>1$. (To prove this relation, note that if $\lambda_1,\ldots,\lambda_N$ are the eigenvalues of M, then $\det M = \lambda_1\lambda_2\cdots\lambda_N$, while $(\mathrm{Tr}\, M)^N = N!\lambda_1\lambda_2\cdots\lambda_N+\cdots$, the ellipses being a sum of terms proportional to some λ_j^k with $k>1$. Those latter terms contribute the terms represented by ellipses in (B.1).) Now, set $M = W_1W_2$. The ellipses in (B.1) are non-chiral terms, and $(\mathrm{Tr}\, M)^N$ is a multiple of S^N, so if we can show that $\det M$ is non-chiral, it follows that S^N vanishes in the chiral ring. To show that $\det M$ is non-chiral, we write first

$$\det M = (-1)^{N(N-1)/2}\epsilon_{i_1 i_2\cdots i_N}\epsilon^{k_1k_2\cdots k_N} W^{i_1}{}_{j_1} W^{i_2}{}_{j_2}\cdots W^{i_N}{}_{j_N} \widetilde{W}^{j_1}{}_{k_1}\widetilde{W}^{j_2}{}_{k_2}\cdots \widetilde{W}^{j_N}{}_{k_N}\, , \quad (B.2)$$

where to minimize the number of indices, we write W and $\widetilde{W}$ for W_1 and W_2. We will focus on the factor

$$\epsilon_{i_1 i_2 \cdots i_N} W^{i_1}{}_{j_1} W^{i_2}{}_{j_2} \cdots W^{i_N}{}_{j_N} \,. \tag{B.3}$$

This is symmetric in $j_1, \ldots, j_N$, so we may as well set those indices to a common value, say N, replacing (B.3) with

$$\epsilon_{i_1 i_2 \cdots i_N} W^{i_1}{}_{N} W^{i_2}{}_{N} \cdots W^{i_N}{}_{N} \,. \tag{B.4}$$

We will show that (B.4) is a multiple of the non-chiral quantity

$$\begin{aligned}
\epsilon_{i_1 i_2 \cdots i_{N-1} N} W^{i_1}{}_{N} W^{i_2}{}_{N} \cdots W^{i_{N-2}}{}_{N} &\left\{\overline{Q}^{\dot\alpha}, [D_{1,\dot\alpha}, W^{i_{N-1}}{}_{N}]\right\} = \\
&= \epsilon_{i_1 i_2 \cdots i_{N-1} N} W^{i_1}{}_{N} W^{i_2}{}_{N} \cdots W^{i_{N-2}}{}_{N} \{W, W\}^{i_{N-1}}{}_{N} \\
&= \epsilon_{i_1 i_2 \cdots i_{N-1} N} W^{i_1}{}_{N} W^{i_1}{}_{N} \cdots W^{i_{N-2}}{}_{N} W^{i_{N-1}}{}_{x} W^{x}{}_{N}
\end{aligned} \tag{B.5}$$

where $D_{1,\dot\alpha}$ is the bosonic covariant derivative $D/Dx^{\alpha\dot\alpha}$ with $\alpha = 1$. Our goal is now to show that this quantity is a non-zero multiple of the quantity in (B.4). If we set $x = N$, we do get such a multiple. If we set x to be one of $i_1, \ldots, i_{N-2}$, we get an expression that vanishes by Fermi statistics. And finally, if we set $x = i_{N-1}$, we get $W^{i_{N-1}}{}_{i_{N-1}} = \mathrm{Tr}\ W = 0$, since W takes values in the Lie algebra of $\mathrm{SU}(N)$. A more precise way to express this argument is to use Fermi statistics to write

$$W^{i_1}{}_{N} W^{i_2}{}_{N} \cdots W^{i_{N-2}}_{N} W^{x}{}_{N} = \frac{1}{(N-1)!} \epsilon^{i_1 \cdots i_{N-2} x y} \epsilon_{m_1 m_2 \cdots m_{N-1} y} W^{m_1}{}_{N} W^{m_2}{}_{N} \cdots W^{m_{N-1}}{}_{N} \,. \tag{B.6}$$

Inserting this in (B.5), one then expresses $\epsilon_{i_1 i_2 \cdots i_{N-1} N} \epsilon^{i_1 \cdots i_{N-2} x y}$ as a multiple of $\delta^x_{i_{N-1}} \delta^y_N - \delta^y_{i_{N-1}} \delta^x_N$. The first term then gives a multiple of $\mathrm{Tr}\ W = 0$, and the second gives the desired multiple of (B.4).

To complete the picture, we should also prove that $S^{N-1} \neq 0$ in the chiral ring; in other words, we should prove that for any $X^{\dot\alpha}$,

$$S^{N-1} \neq \{\overline{Q}_{\dot\alpha}, X^{\dot\alpha}\} \,. \tag{B.7}$$

Any expression of the form $\{\overline{Q}_{\dot\alpha}, X^{\dot\alpha}\}$ that can be written as a polynomial in W_α is proportional to an anticommutator $\{W_\alpha, W_\beta\}$, since the action of $\overline{Q}_{\dot\alpha}$ on a bosonic covariant derivative generates $\{W_\alpha, \ldots\}$, and acting on anything else there is no way to generate an expression involving only W_α. (Conversely, we have seen that any polynomial in W proportional to $\{W_\alpha, W_\beta\}$ is of the form $\{\overline{Q}_{\dot\alpha}, X^{\dot\alpha}\}$.) If, therefore, we can find a gauge field configuration in which $\{W_\alpha, W_\beta\} = 0$ for all α, β, and in which $S^{N-1} \neq 0$, this will establish that $S^{N-1} \neq \{\overline{Q}_{\dot\alpha}, X^{\dot\alpha}\}$ for any $X^{\dot\alpha}$. We simply take an abelian background connection for which W_α is diagonal, $W_\alpha = \mathrm{diag}(\psi^1_\alpha, \psi^2_\alpha, \ldots, \psi^N_\alpha)$; here ψ^i_α are anticommuting c-numbers obeying

$$\sum_i \psi^i_\alpha = 0 \,, \tag{B.8}$$

but otherwise unconstrained. So $S = -\frac{1}{32\pi^2} \sum_{i=1}^N \psi^i_\alpha \psi^{\alpha\, i}$, from which it follows that for this configuration $S^{N-1} \neq 0$, though (B.8) implies that this configuration has $S^N = 0$.

C. Computing the period integral

The purpose of this appendix is to compute the period integral in the case of a maximally confining gauge group. We need to calculate

$$\begin{aligned} -2\pi i\Pi(z_0,\Lambda_0) &= -\int_{z_0+2C}^{\Lambda_0} y dz \\ y^2 &= W'(z)^2 + f(z) = \left((z-z_0)^2 - 4C^2\right) Q(z)^2 \end{aligned} \tag{C.1}$$

with C a solution of

$$\sum_{l=0}^{[\frac{n}{2}]} \frac{C^{2l}}{(l!)^2} W^{(2l+1)}(z_0) = 0 \tag{C.2}$$

and Q given by (5.25).

We are interested in the limit of $\Pi(z_0,\Lambda_0)$ as $\Lambda_0 \to \infty$. The asymptotic behavior of y is

$$y = W'(z) + \frac{1}{2}\frac{f_{n-1} n!}{W^{(n+1)}}\frac{1}{z} + \mathcal{O}\left(\frac{1}{z^2}\right) \tag{C.3}$$

with

$$f_{n-1} = -4\frac{W^{(n+1)}}{n!} S = -4\frac{W^{(n+1)}}{n!} \sum_{l=1}^{[\frac{n+1}{2}]} \frac{lC^{2l}}{(l!)^2} W^{(2l)}(z_0)\,. \tag{C.4}$$

Therefore by integrating (C.3),

$$-2\pi i\Pi(z_0,\Lambda_0) = -W(\Lambda_0) + 2S\log\Lambda_0 + F(z_0) + \mathcal{O}\left(\frac{1}{\Lambda_0}\right). \tag{C.5}$$

We should determine the Λ_0 independent function $F(z_0)$.

Let us differentiate (C.5) with respect to S. In doing that we should remember that z_0 is a function of S through the relation (C.4).

$$2t(S) = -2\pi i\frac{\partial}{\partial S}\Pi(z_0,\Lambda_0) = 2\log\Lambda_0 + \frac{\partial}{\partial S}F(z_0) + \mathcal{O}\left(\frac{1}{\Lambda_0}\right) \tag{C.6}$$

Although we will not need it in the derivation, we comment that this $t(S)$ is the coupling denoted by $t(S)$ in (1.3).

We can also compute $2t(S) = -2\pi i\frac{\partial}{\partial S}\Pi(z_0,\Lambda_0)$ from the integral expression,

$$2t(S) = -2\pi i\frac{\partial}{\partial S}\Pi(z_0,\Lambda_0) = -\int_{z_0+2C}^{\Lambda_0}\frac{\partial}{\partial S} y dz \tag{C.7}$$

where the term from the derivative acting on the lower bound of the integral is zero since $y(z_0+2C) = 0$ as can be seen in (C.1).

Let us determine the meromorphic function $\frac{\partial}{\partial S}y(z)$ by its behavior at its singular points.

$$\frac{\partial y}{\partial S} \to \begin{cases} -\frac{2}{z} & \text{for } z\to\infty \\ \frac{c_1}{\sqrt{z-z_0+2C}} & \text{for } z\to z_0-2C \\ \frac{c_2}{\sqrt{z-z_0-2C}} & \text{for } z\to z_0+2C. \end{cases}$$

with c_i two non zero constants. The unique meromorphic function with this behavior is,

$$\frac{\partial y}{\partial S} = -\frac{2}{\sqrt{(z-z_0)^2 - 4C^2}}$$

The integral (C.7) is elementary and is given by,

$$2t(S) = 2\log\left(\frac{\Lambda_0 - z_0 + \sqrt{(\Lambda_0 - z_0)^2 - 4C^2}}{2C}\right) = 2\log\frac{\Lambda_0}{C} + \mathcal{O}\left(\frac{1}{\Lambda_0}\right)$$

Comparing this result with (C.6) we get,

$$\frac{\partial}{\partial S}F(z_0) = -2\log C(z_0) \tag{C.8}$$

It is not difficult to check using (C.2) that

$$F(z_0) = -2S\log C(z_0) + \sum_{k=0}^{\left[\frac{n+1}{2}\right]} \frac{C^{2k}}{(k!)^2} W^{(2k)}(z_0) \tag{C.9}$$

satisfies (C.8).

Combining the results into (C.5), we get our final answer,

$$I(z_0, \Lambda_0) = -W(\Lambda_0) + 2S\log\frac{\Lambda_0}{C} + \sum_{k=0}^{\left[\frac{n+1}{2}\right]} \frac{C^{2k}}{(k!)^2} W^{(2k)}(z_0) + \mathcal{O}\left(\frac{1}{\Lambda_0}\right).$$

References

[1] K.A. Intriligator and N. Seiberg, *Lectures on supersymmetric gauge theories and electric-magnetic duality*, *Nucl. Phys.* **45BC** *(Proc. Suppl.)* (1996) 1 [hep-th/9509066].

[2] R. Dijkgraaf and C. Vafa, *A perturbative window into non-perturbative physics*, hep-th/0208048.

[3] M. Bershadsky, S. Cecotti, H. Ooguri and C. Vafa, *Kodaira-Spencer theory of gravity and exact results for quantum string amplitudes*, *Commun. Math. Phys.* **165** (1994) 311 [hep-th/9309140].

[4] F. Cachazo, K.A. Intriligator and C. Vafa, *A large-N duality via a geometric transition*, *Nucl. Phys.* **B 603** (2001) 3 [hep-th/0103067].

[5] F. Cachazo and C. Vafa, $N = 1$ *and* $N = 2$ *geometry from fluxes*, hep-th/0206017.

[6] R. Dijkgraaf and C. Vafa, *Matrix models, topological strings and supersymmetric gauge theories*, *Nucl. Phys.* **B 644** (2002) 3 [hep-th/0206255].

[7] R. Dijkgraaf and C. Vafa, *On geometry and matrix models*, *Nucl. Phys.* **B 644** (2002) 21 [hep-th/0207106].

[8] R. Dijkgraaf, M.T. Grisaru, C.S. Lam, C. Vafa and D. Zanon, *Perturbative computation of glueball superpotentials*, hep-th/0211017.

JHEP12(2002)071

[9] K. Konishi, *Anomalous supersymmetry transformation of some composite operators in sqcd*, *Phys. Lett.* **B 135** (1984) 439.

[10] K.-i. Konishi and K.-i. Shizuya, *Functional integral approach to chiral anomalies in supersymmetric gauge theories*, *Nuovo Cim.* **A90** (1985) 111.

[11] G. Veneziano and S. Yankielowicz, *An effective lagrangian for the pure $N = 1$ supersymmetric Yang-Mills theory*, *Phys. Lett.* **B 113** (1982) 231.

[12] K.A. Intriligator, R.G. Leigh and N. Seiberg, *Exact superpotentials in four-dimensions*, *Phys. Rev.* **D 50** (1994) 1092 [hep-th/9403198].

[13] V.A. Novikov, M.A. Shifman, A.I. Vainshtein and V.I. Zakharov, *Instanton effects in supersymmetric theories*, *Nucl. Phys.* **B 229** (1983) 407.

[14] N. Seiberg, *Electric-magnetic duality in supersymmetric nonabelian gauge theories*, *Nucl. Phys.* **B 435** (1995) 129 [hep-th/9411149].

[15] K.A. Intriligator and N. Seiberg, *Duality, monopoles, dyons, confinement and oblique confinement in supersymmetric* $\mathrm{SO}(N(c))$ *gauge theories*, *Nucl. Phys.* **B 444** (1995) 125 [hep-th/9503179].

[16] O. Aharony, unpublished.

[17] A. Ceresole, G. Dall'Agata, R. D'Auria and S. Ferrara, *Spectrum of type-IIB supergravity on $AdS_5 \times T(11)$: predictions on $N = 1$ SCFT's*, *Phys. Rev.* **D 61** (2000) 066001 [hep-th/9905226].

[18] E. Witten, *The verlinde algebra and the cohomology of the grassmannian*, hep-th/9312104; and in *Quantum fields and strings: a course for mathematicians*, P. Deligne et. al. eds. American Mathematical Society, 1999, vol. 2, pp. 1338–1339.

[19] N. Seiberg, *Exact results on the space of vacua of four-dimensional SUSY gauge theories*, *Phys. Rev.* **D 49** (1994) 6857 [hep-th/9402044].

[20] N. Seiberg, *Naturalness versus supersymmetric nonrenormalization theorems*, *Phys. Lett.* **B 318** (1993) 469 [hep-ph/9309335].

[21] S.J. Gates, M.T. Grisaru, M. Roček and W. Siegel, *Superspace, or one thousand and one lessons in supersymmetry*, *Front. Phys.* **58** (1983) 1–548 [hep-th/0108200].

[22] A. Gorsky, *Konishi anomaly and $N = 1$ effective superpotentials from matrix models*, hep-th/0210281.

[23] H. Nicolai, *On a new characterization of scalar supersymmetric theories*, *Phys. Lett.* **B 89** (1980) 341.

[24] M. Staudacher, *Combinatorial solution of the two matrix model*, *Phys. Lett.* **B 305** (1993) 332 [hep-th/9301038].

[25] A.A. Migdal, *Loop equations and 1/n expansion*, *Phys. Rept.* **102** (1983) 199.

[26] P. Ginsparg and G.W. Moore, *Lectures on 2-D gravity and 2-D string theory*, hep-th/9304011.

[27] E. Brezin, C. Itzykson, G. Parisi and J.B. Zuber, *Planar diagrams*, *Commun. Math. Phys.* **59** (1978) 35.

JHEP12(2002)071

[28] N. Seiberg and E. Witten, *Electric-magnetic duality, monopole condensation and confinement in $N = 2$ supersymmetric Yang-Mills theory*, *Nucl. Phys.* **B 426** (1994) 19 [hep-th/9407087]; erratum *ibid.* **430** (1994) 485.

[29] P.C. Argyres and M.R. Douglas, *New phenomena in* SU(3) *supersymmetric gauge theory*, *Nucl. Phys.* **B 448** (1995) 93 [hep-th/9505062].

[30] P.C. Argyres, M. Ronen Plesser, N. Seiberg and E. Witten, *New $N = 2$ superconformal field theories in four dimensions*, *Nucl. Phys.* **B 461** (1996) 71 [hep-th/9511154].

[31] F. Ferrari, *On exact superpotentials in confining vacua*, *Nucl. Phys.* **B 648** (2003) 161 [hep-th/0210135].

[32] H. Fuji and Y. Ookouchi, *Comments on effective superpotentials via matrix models*, *J. High Energy Phys.* **12** (2002) 067 [hep-th/0210148].

[33] M.R. Douglas and S.H. Shenker, *Dynamics of* SU(N) *supersymmetric gauge theory*, *Nucl. Phys.* **B 447** (1995) 271 [hep-th/9503163].

[34] O. Aharony, S.S. Gubser, J.M. Maldacena, H. Ooguri and Y. Oz, *Large-N field theories, string theory and gravity*, *Phys. Rept.* **323** (2000) 183 [hep-th/9905111].

[35] *Free probability theory*, D. Voiculescu ed., pp. 21–40, AMS, 1997.

[36] R. Dijkgraaf, S. Gukov, V.A. Kazakov and C. Vafa, *Perturbative analysis of gauged matrix models*, hep-th/0210238.

[37] R. Gopakumar, *$N = 1$ theories and a geometric master field*, hep-th/0211100.

[38] S.G. Naculich, H.J. Schnitzer and N. Wyllard, *The $N = 2$* U(N) *gauge theory prepotential and periods from a perturbative matrix model calculation*, hep-th/0211123.