CERN-TH/95-303
hepth/yymmnnn

Introduction to Seiberg-Witten Theory

W. Lerche

CERN, Geneva, Switzerland

Abstract..

We give an elementary introduction to the SW solution of $N = 2$ supersymmetric Yang-Mills theory.

CERN-TH/95-303
November 1995

1 Introduction

The last year has seen a remarkable progress in understanding non-perturbative properties of supersymmetric field and string theories. This quite dramatic development was initiated by the work of Seiberg and Witten on supersymmetric Yang-Mills theory [1], and by Hull and Townsend on heterotic-type II string equivalence [2], after interest in non-perturbative duality has been revived by the work of Sen [3]. Now, one year later, many non-perturbatively exact statements can be made about various $N=2$ supersymmetric Yang-Mills theories with and without matter [4, 5], and similar exact statements about $N=2$ string theories [6] that include gravity [7]; not even to speak about $N=4$ string theories.

It is becoming evident that the main insight is of conceptional nature and goes far beyond original expectations. The picture that seems to emerge is, essentially, that the various perturbatively defined string theories represent non-perturbatively equivalent, or dual, descriptions of one and the same fundamental theory; moreover, it seems that strings do not play a privileged role in this theory besides p-branes, which now appear on a footing quite similar to strings [8, 9]. It may well turn out, ultimately, that there is just one theory that is fully consistent at the non-perturbative level, or a very small number of such theories. Though the number of free parameters ("moduli") may a priori be very large –which would hamper predictive power–, it is clear that investigating this kind of issues is important and will shape our understanding of the very nature of grand unification.

Because a full treatment of these matters is outside the scope of these lecture notes, but some of the relevant concepts play a role already in supersymmetric Yang-Mills theory, we like to confine ourselves to a very basic discussion of the work of Seiberg and Witten, and some generalizations of it. Specifically, since a detailed review that emphasizes monopole physics and duality symmetries has appeared recently [10], we intend to present the subject here, in a complementary way, from the view point of analytic continuation and the underlying Riemann-Hilbert problem.

So, in a nutshell, what is all the excitement about that has made furor even in the mass media ? As one of the main results one may state the exact non-perturbative low energy effective Lagrangian of $N=2$ supersymmetric Yang-Mills theory with gauge group $SU(2)$; it contains, in particular, the effective, renormalized gauge coupling, $g_{\rm eff}$, and theta-angle, $\theta_{\rm eff}$:

$$\Big(\frac{\theta_{\rm eff}(a)}{\pi}+\frac{8\pi i}{{g_{\rm eff}}^2(a)}\Big) \;=\; \underbrace{\frac{8\pi i}{{g_0}^2}}_{\rm bare}+\underbrace{\frac{2i}{\pi}\log\Big[\frac{a^2}{\Lambda^2}\Big]}_{\rm one-loop}-\underbrace{\frac{i}{\pi}\sum_{\ell=1}^{\infty}c_\ell\Big(\frac{\Lambda}{a}\Big)^{4\ell}}_{\rm instanton\ corrections} \qquad (1.1)$$

Here, Λ is the scale at which the gauge coupling becomes strong, and a is the Higgs field. This effective, field dependent coupling arises by setting the renormalization scale, μ, equal to the characteristic scale of the theory, which is given by the Higgs VEV: $g_{\rm eff}(\mu) \to g_{\rm eff}(a)$. Specifically, the running of the perturbative coupling constant can be depicted as follows:

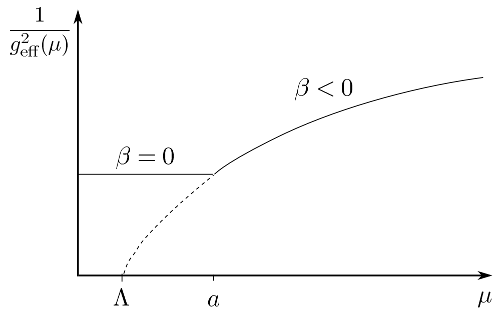


Figure 1: *At scales above the Higgs VEV a, the masses of the non-abelian gauge bosons, $W^\pm$, are negligible, and we can see the ordinary running of the coupling constant of an asymptotically free theory. At scales below a, $W^\pm$ freeze out, and we are left with just an effective $U(1)$ gauge theory with vanishing β-function.*

The general form of the full non-perturbatively corrected coupling (1.1) has been known for some time [11], where one has made use of the fact that all what can come from perturbation theory arises to one loop order only [12], and of the known amount of R-charge violation (given by 8ℓ) of the ℓ-instanton process. The unknown piece of (1.1) are the precise values of the infinitely many instanton coefficients c_ℓ, and it is the achievement of Seiberg and Witten to determine all of these coefficients explicitly. This gives infinitely many predictions for zero momentum correlators involving a and gauginos in non-trivial instanton backgrounds. Such correlators are topological and also have an interpretation in terms of Donaldson theory, which deals with topological invariants of four-manifolds. It is the easy determination of such topological quantities that has been the main reason for excitement on the mathematician's side. The fact that highly non-trivial mathematical results can be reproduced gives strong evidence that S&W's approach for solving the Yang-Mills theory is indeed correct, even though many details, like a rigorous field theoretic definition of the theory, may not yet be completely settled.

It is, however, presently not clear what lessons can ultimately be drawn for non-supersymmetric theories, like ordinary QCD. The hope is, of course, that even though supersymmetry is an essential ingredient, it is only a technical requirement that facilitates computations, and thus that the supersymmetric toy model displays the physically relevant features.

Let us thus list some typical benefits of supersymmetric field theories:

•*Non-renormalization properties:* perturbative quantum corrections are less violent; this is related to a

•*Holomorphic structure,* which leads to vacuum degeneracies, and allows to use powerful methods of complex analysis.

•*Duality symmetries* between electric and magnetic, or weak and strong coupling sectors are more or less manifest, depending on the number of supersymmetries.

The maximum number of supersymmetries is four in a globally supersymmetric theory:

•*N=4* supersymmetric Yang-Mills theory is conjectured to be self-dual [13], ie., completely invariant under the exchange of electric and magnetic sectors. However, though interesting, this theory is too simple for the present purpose of investigating quantum corrections, since there aren't any in this theory.

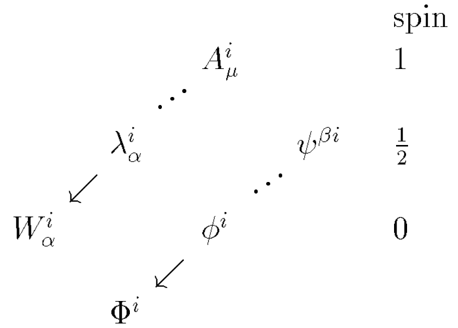
•*N=1* supersymmetric Yang-Mills theory, on the other hand, is presumably not exactly solvable, since the quantum corrections are not under full control; only certain sub-sectors of the theory are governed by holomorphic objects (like the chiral superpotential), and thus are protected from perturbative quantum corrections.

•*N=2* supersymmetric Yang-Mills theory is at the border between "trivial" and "not fully solvable", in that it is (in the low-energy limit) exactly solvable. It is governed by a holomorphic function, the "prepotential" $\mathcal{F}$, for which the perturbative quantum corrections are under control, ie., occur just to one loop order.

Having motivated why it should be fruitful to study $N=2$ Yang-Mills theory, we now turn to discuss it in more detail.

2 Semi-classical N=2 Yang-Mills theory for $G = SU(2)$

The field content for pure $N=2$ Yang-Mills theory is given by vector supermultiplets in the adjoint representation of the gauge group. For convenience, one often rewrites such multiplets in terms of $N=1$ chiral multiplets, W_α^i, Φ^i, as follows:



The bottom component, the scalar field ϕ, feels the following potential:

$$V(\phi) \;=\; \frac{1}{g^2}\,\mathrm{Tr}[\phi,\phi^\dagger]^2\ . \tag{2.1}$$

It displays a typical feature of supersymmetric theories, namely flat directions along which $V(\phi)\equiv 0$. That is, field values

$$\phi \;=\; a\,\sigma_3 \tag{2.2}$$

do not cost any energy. Of course, if $a\neq 0$, we have a spontaneous symmetry breakdown: $SU(2) \hookrightarrow U(1)$. A more suitable "order" parameter is given by the gauge invariant Casimir

$$u(a) \;=\; \mathrm{Tr}\,\phi^2 \;=\; 2a^2\ . \tag{2.3}$$

It is in particular invariant under the Weyl group of $SU(2)$, which is, physically, the discrete remnant of the gauge transformations that act within the Cartan subalgebra: $a\to -a$. The quantity u represents a good coordinate of the manifold of inequivalent vacua, which is usually called "moduli space". Since u can be any complex number, the moduli space is given by the complex plane, which may be compactified to the Riemann sphere by adding a point at infinity. The moduli space, $\mathcal{M}_c$, of the classical theory is depicted in Fig.2.

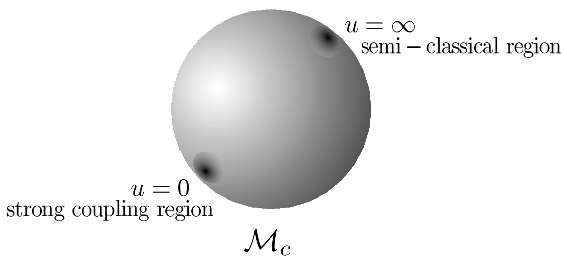


Figure 2: *The classical moduli space of* $SU(2)$ $N=2$ *gauge theory has singularities at* $u=0$ *and* ∞.

In the bulk of $\mathcal{M}_c$ there is just an unbroken $U(1)$ gauge symmetry, only at the origin it will be enhanced to $SU(2)$. What we are after is a "Wilsonian" effective lagrangian description of the theory, for any given value of u. Such an effective lagrangian can in principle be obtained by integrating out all fluctuations above some scale μ (that, as we have indicated earlier, we choose to be equal to a). In particular, we would integrate out the massive non-abelian gauge bosons $W^{\pm}$, to obtain an effective action that involves only the neutral gauge multiplet, $W^0 = (A \equiv \Phi^0, W^0_\lambda)$. It is clear that, semi-classically, this theory can possibly be meaningful only outside a neighborhood of $u = 0$, since at $u = 0$ the non-abelian gauge bosons $W^{\pm}$ will be massless, and the effective description in terms of only W^0 will not be accurate – actually, it will become non-local. This means that $u = 0$ will be a singular point on $\mathcal{M}_c$, besides the point of infinity.

Note that it is clear from Fig.1 that because of asymptotic freedom, the region near $u = \infty$ will correspond to weak coupling, so that only in this semi-classical region reliable computations can be done in perturbation theory. On the other hand, the theory will be strongly coupled near the classical $SU(2)$-enhancement point $u = 0$, so that a priori no reliable quantum statements about the theory can be made here.

It is known (just from supersymmetry) that the low energy effective lagrangian[1] is completely determined by a holomorphic prepotential $\mathcal{F}$ and must be of the form:

$$\mathcal{L} \;=\; \frac{1}{4\pi}\mathrm{Im}\Big[\int d^4\theta\, K(A,\bar{A}) + \int d^2\theta\,(\frac{1}{2}\sum\tau(A)W^\alpha W_\alpha)\Big]\,. \tag{2.4}$$

Here, $\Phi \equiv: A\,\sigma_3$, and

$$K(A,\bar{A}) \;=\; \frac{\partial\mathcal{F}(A)}{\partial A}\bar{A} \tag{2.5}$$

is the "Kähler potential" which gives a supersymmetric non-linear σ-model for the field A, and

$$\tau(A) \;=\; \frac{\partial^2\mathcal{F}(A)}{\partial^2 A}\,. \tag{2.6}$$

That is, the bosonic piece of (2.4) is, schematically,

$$\mathcal{L} \;=\; \mathrm{Im}\tau\Big\{\partial a\,\partial\bar{a} + F\cdot F\Big\} + \mathrm{Re}\tau\, F\cdot\tilde{F} \;+\ldots, \tag{2.7}$$

from which we see that

$$\tau(a) \;\equiv\; \frac{\theta(a)}{\pi} + \frac{8\pi i}{g^2(a)} \tag{2.8}$$

represents the complexified effective gauge coupling and that $\mathrm{Im}\tau$ is the σ-model metric on $\mathcal{M}_c$. Classically, $\mathcal{F}(A) = \frac{1}{2}\tau_0 A^2$, where τ_0 is the bare coupling constant. The full quantum

[1]By this we mean the piece of the effective lagrangian that is leading for vanishing momenta, ie., that contains at most two derivatives. There are of course infinitely many higher derivative terms in the full effective action; these are not governed by holomorphic quantities, and we cannot say anything about them in the present context.

prepotential will receive [12] perturbative (one-loop) and non-perturbative corrections, and must be of the form [11]:

$$\mathcal{F}(A) \;=\; \frac{1}{2}\tau_0 A^2 + \frac{i}{\pi}\,A^2 \log\Big[\frac{A^2}{\Lambda^2}\Big] + \frac{1}{2\pi i}A^2\sum_{\ell=1}^{\infty} c_\ell\Big(\frac{\Lambda}{A}\Big)^{4\ell} \qquad (2.9)$$

By taking two derivatives, $\mathcal{F}$ gives rise to the effective coupling (1.1) mentioned in the introduction. Note that indeed for large $a \equiv A|_{\theta=0}$, the instanton sum converges well, and the theory is dominated by semi-classical, one-loop physics.

A crucial insight [1] is that the global properties of the effective gauge coupling $\tau(a)$ are very important. Specifically, we know that near $u = \infty$:

$$\tau \;=\; \text{const} + \frac{2i}{\pi}\,\log\Big[\frac{u}{\Lambda^2}\Big] + \text{single-valued}\ . \qquad (2.10)$$

This implies that if we loop around $u = \infty$ in the moduli space, the logarithm will produce an extra shift of $2\pi i$ because of its branch cut, and thus:

$$\tau \;\longrightarrow\; \tau - 4\ . \qquad (2.11)$$

From (2.8) it is clear that this monodromy just corresponds to an irrelevant shift of the θ-angle, but what we learn is that τ, as well as $\mathcal{F}$, are not functions but rather multi-valued sections. Actually, the full story is more complicated than that, in that also the imaginary part, $\mathrm{Im}\tau = \frac{8\pi}{g^2}$, will be globally non-trivial.

More specifically, we see from (2.7) that $\mathrm{Im}\tau$ represents a metric on the moduli space, and the physics requirement of unitarity implies that it must be positive throughout the moduli space:

$$\mathrm{Im}\tau(u) \;>\; 0\,. \qquad (2.12)$$

It is now a simple mathematical fact that since $\mathrm{Im}\tau$ is a harmonic function (ie., $\partial\overline{\partial}\mathrm{Im}\tau = 0$), it cannot have a minimum if it is globally defined. Thus, in order not to conflict with unitarity, we learn that $\mathrm{Im}\tau$ can only locally be defined – a priori, it is defined only in the semi-classical coordinate patch near infinity, cf., (2.10). We thus conclude that the global structure of the true "quantum" moduli space, $\mathcal{M}_q$, must be very different as compared to the classical moduli space, $\mathcal{M}_c$. In particular, any situation with just two singularities must be excluded.

3 The exact quantum moduli space

The question thus arises, how many and what kind of singularities the exact quantum moduli space should have, and what the physics significance of these singularities might be. Seiberg and Witten proposed that there should be two singularities at $u = \pm\Lambda^2$, where Λ is the dynamically generated quantum scale, and that the classical singularity at the origin disappears:

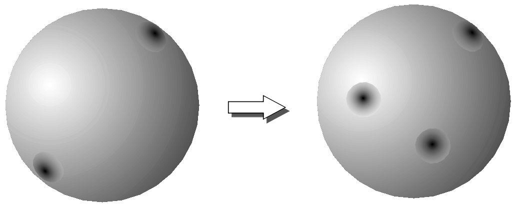


Figure 3: *The transition from the classical to the exact quantum theory involves splitting and shifting of the strong coupling singularity away from $u = 0$ to $u = \pm\Lambda^2$.*

Though this will prove to be a physically motivated and self-consistent assumption about the strong coupling behavior, it cannot, at present, be rigorously derived from first principles. But there is a whole bunch of arguments, with varying degree of rigor, why precisely the situation depicted in Fig.3. must be the correct one. For example, the absence of a singularity at $u = 0$ (which implies that there are, in the full quantum theory, no extra massless gauge fields $W^{\pm}$) is motivated by the absence of an R-current that a superconformal theory with massless gauge bosons would otherwise have [1]. Furthermore, the appearance of just two, and not $2n^2$ strong coupling singularities reflects that the corresponding $N = 1$ theory (obtained by explicitly breaking the $N = 2$ theory by a mass term for Φ) has precisely two vacua (from Witten's index, $\mathrm{Tr}(-1)^F = n$ for $SU(n)$). More mathematically speaking, the singularity structure poses, as will be explained later, a particular non-abelian monodromy problem, and apparently there is no solution for this problem for any other arrangement of singularities.

The most interesting question is clearly what the physical significance of the extra strong coupling singularities is. One expects in analogy to the classical theory, where at $u = 0$ the singularity is due to the extra massless gauge bosons $W^{\pm}$, that the strong coupling singularities in the quantum moduli space should be attributed to certain massless excitations as well. Guided by the early ideas of 't Hooft about confinement [14], Seiberg and Witten postulated that near these singularities certain 't Hooft-Polyakov monopoles must become light.

There is in fact a powerful tool to get a handle on soliton masses in theories with extended supersymmetry, namely the BPS-formula [15]:

$$m^2 \geq |Z|^2 , \tag{3.1}$$

[2]The number of singularities must be even to be consistent with global R-symmetry, which acts as $u \to -u$.

where Z is the central charge of the superalgebra in question. For $N=2$ supersymmetry, this formula immediately follows from unitarity ($\bar{Q}\,Q > 0$), and from the anti-commutator

$$\{Q_{\alpha i}, \overline{Q}_{\beta j}\} = \delta_{ij}\gamma^{\mu}_{\alpha\beta}P_{\mu} + \delta_{\alpha\beta}\epsilon_{ij}U + (\gamma_5)_{\alpha\beta}\epsilon_{ij}V \ , \tag{3.2}$$

where $|Z|^2 \equiv U^2 + V^2$. The important point is that the BPS bound (3.1) is saturated by a certain class of excitations, namely the "BPS-states" that obey $Q|\psi\rangle = 0$. The idea is that if a state obeys this condition semi-classically, it obeys it also in the exact quantum theory, because the number of degrees of freedom of a "short" (or "chiral") multiplet that obeys $Q|\psi\rangle = 0$ is smaller as compared to the degrees of freedom of a generic supersymmetry multiplet, and the number of degrees of freedom is supposed not to jump when switching on quantum corrections. In particular, since 't Hooft-Polyakov monopoles do satisfy the BPS bound semi-classically, they must obey it in the exact theory as well. From semi-classical considerations we can also learn that the monopoles lie in $N=2$ hypermultiplets, which have maximum spin $\frac{1}{2}$.

For $N=2$ supersymmetric Yang-Mills theories, the central charge takes the form

$$Z = q\,a + g\,a_D \ , \tag{3.3}$$

where (g, q) are the (magnetic,electric) quantum numbers of the BPS state under consideration. Above, a_D is the "magnetic dual" of the electric Higgs field a and belongs to the vector multiplet $(A_D, W_{\alpha,D})$ that contains the dual, magnetic photon, A^{μ}_{D}. By studying the electric-magnetic duality transformation, under which the ordinary electric gauge potential A^{μ} transforms into A^{μ}_{D}, it turns out [1] that in the $N=2$ Yang-Mills theory the dual variable a_D is simply given by:

$$a_D = \frac{\partial}{\partial a}\mathcal{F}(a) \ . \tag{3.4}$$

That is, the general idea is that at the singularity at $u = \Lambda^2$, one would have $a \neq 0$ but $a_D = 0$, such that (by (3.3)) a monopole hypermultiplet with charges $(g, q) = (\pm 1, 0)$ would be massless. On the other hand, one would have that $u = 0$ does not imply $a = 0$ in the exact theory, such that, in contrast to the classical theory, no gauge bosons (with charges $(0, \pm 2)$) become massless. This in particular would imply that the classical relation $u = 2a^2$ can hold only asymptotically in the weak-coupling region.

The point is to view $a_D(u)$ as a variable that is on a equal footing as $a(u)$; it just belongs to a dual gauge multiplet that couples locally to magnetically charged excitations, in the same way that a couples locally to electric excitations (such as $W^{\pm}$). A priori, it would not matter which variable we use to describe the theory, and which variable we actually use will rather depend on the region of $\mathcal{M}_q$ that we are looking at. More specifically, in the original semi-classical, "electric" region near $u = \infty$, the preferred local variable is a, and an appropriate lagrangian is given by (2.9). As mentioned above, the instanton sum converges well for large $a \simeq \sqrt{u/2}$.

However, if we try to extend $\mathcal{F}(a)$ to a region far enough away from $u = \infty$, we will leave the domain of convergence of the instanton sum, and we cannot really make any more much sense of $\mathcal{F}$. That is, in attempting to *globally extend* the effective lagrangian description outside the semi-classical coordinate patch, we face the problem of suitably analytically continuing $\mathcal{F}$. The point is that even though we cannot have a choice of $\mathcal{F}$ that would be globally valid anywhere on $\mathcal{M}_q$ (it would conflict with positivity, cf., (2.12)), we can resum the instanton terms in $\mathcal{F}$ in terms of other variables, to yield another form of the lagrangian that converges well in another region of $\mathcal{M}_q$. The reader might already have guessed that while a is the preferred variable near $u = \infty$, it is a_D that is the preferred variable in the "magnetic" strong coupling coordinate patch centered at $u = \Lambda^2$. More precisely, near $u = \Lambda^2$ we expect to have the following, dual form of the effective lagrangian:

$$\mathcal{F}_D(A_D) \;=\; \frac{1}{2}\tau_0^D {A_D}^2 - \frac{i}{4\pi}\,{A_D}^2 \log\Big[\frac{A_D}{\Lambda}\Big] - \frac{1}{2\pi i}\Lambda^2 \sum_{\ell=1}^{\infty} c_\ell^D \Big(\frac{iA_D}{\Lambda}\Big)^\ell \,. \qquad (3.5)$$

Indeed, the infinite sum converges well because at this singularity $a_D \to 0$. From the coefficient of the logarithm we see that the theory is non-asymptotically free (positive β-function), and thus weakly coupled for $a_D \to 0$ (though strongly coupled in terms of the original variable, a).

Specifically, the coefficient of the logarithm tells us that this dual theory is simply given by an abelian $U(1)$ gauge theory (contributing zero to the β-function), coupled to charged matter that is integrated out and is massless at $a_D = 0$. The coefficient of the β-function reflects that there should be a single matter field with unit charge coupling to the (dual) photon, which belongs to a $N=2$ hypermultiplet (contributing positively to the β-function). Of course this extra matter hypermultiplet is just the dual representative of the massless magnetic monopole. To the dual magnetic photon related to a_D, the monopole looks like an ordinary, elementary (local) field, in spite of that it couples to the original electric photon in a non-local way. It is this dual, abelian reformulation of the original non-abelian instanton problem what leads to substantial simplifications, especially to the mathematician's profit.

Note that the infinite sum of correction terms in (3.5) reflects the effect of integrating out infinitely many massive BPS states, and though their physical meaning is completely different, they carry the same information as the instanton terms in the original lagrangian, (2.9). Note also that the situation at the other singularity, $u = -\Lambda^2$, does not present anything new, in that it is isomorphic to the the situation at $u = \Lambda^2$ and related to it by simply replacing a_D in $\mathcal{F}_D(a_D)$ by $a_D - 2a$. The whole scheme can therefore be depicted as follows:

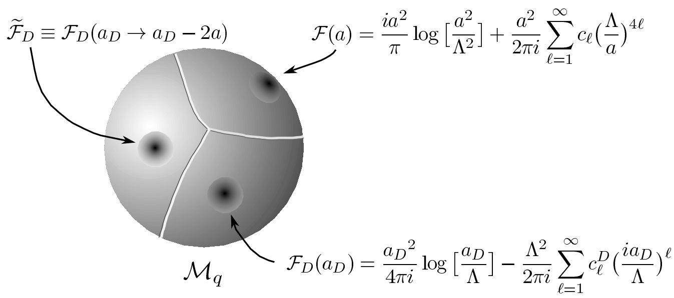


Figure 4: *The exact moduli space is covered by three distinct regions, in the center of each of which the theory is weakly coupled when choosing suitable local variables. A local effective lagrangian exists in each coordinate patch, representing a particular perturbative approximation. No local lagrangian exists that would be globally valid: the three perturbative regions involve quantities that are mutually non-local.*

The astute reader will, however, have noticed that so far nothing concrete was achieved yet – instead, we have introduced a another set of infinitely many unknowns, c^D_ℓ–, and also that we have just guessed the coefficient of the logarithm in (3.5). Indeed, this specific coefficient cannot be derived at this point, but rather is part of the assumption that a single monopole with unit charge becomes massless at $u = \Lambda^2$.

The issue is now to obtain the values of all the unknown coefficients in $\mathcal{F}$, $\mathcal{F}_D$ (2.9),(3.5) from the assumptions that govern the *local*, ie., perturbative behavior of the theory in each of the three coordinate patches in Fig.4. The local behavior is determined by the coefficients of the logarithms, which can reliably be computed in one-loop perturbation theory and obviously reflect the charge quantum numbers of the fields that are supposed to be light near a given singularity.

More precisely, the key idea is that it is the *patching together of the local data in a globally consistent way* that will completely fix the theory (up to irrelevant ambiguities like θ-shifts). That is, the logarithmic term determines the local monodromy around a given singularity that acts on the section $\binom{a_D}{a}$ as follows:

$$\begin{pmatrix} a_D(u) \\ a(u) \end{pmatrix} \longrightarrow M \begin{pmatrix} a_D(u) \\ a(u) \end{pmatrix} . \tag{3.6}$$

In particular, from our knowledge of the asymptotic behavior of $a_D(u), a(u)$ at semi-classical infinity,

$$\begin{pmatrix} a_D(u) \\ a(u) \end{pmatrix} \simeq \begin{pmatrix} \frac{i}{\pi}\sqrt{2u}\log(u/\Lambda^2)) \\ \sqrt{2u} \end{pmatrix} \tag{3.7}$$

we infer that for a loop around $u = \infty$:

$$M_\infty = \begin{pmatrix} -1 & 4 \\ 0 & -1 \end{pmatrix} . \tag{3.8}$$

As for the strong coupling singularities at $u = \pm\Lambda^2$, we choose a different strategy: we know on general grounds that the monodromy of a dyon with charges (g,q) that becomes massless at a given singularity is given by:

$$M^{(g,q)} = \begin{pmatrix} 1+qg & q^2 \\ -g^2 & 1-gq \end{pmatrix} \tag{3.9}$$

This can be seen in various ways, one of which will be explained later at the end of section 4.

The global consistency condition on how to patch together the local, perturbative data is then simply

$$M_{+\Lambda^2} \cdot M_{-\Lambda^2} \;=\; M_\infty \;, \tag{3.10}$$

since we can smoothly pull the monodromy paths γ around the Riemann sphere (u_0 is an arbitrary base point):

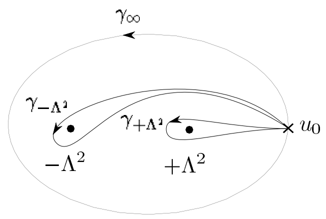


Figure 5: *Monodromy paths in the u-plane.*

One may view equation (3.10) as a condition on the possible massless spectra at $u = \pm\Lambda^2$. For matrices of the restricted form (3.9), its solution is:

$$\begin{aligned} M_{+\Lambda^2} &= M^{(1,0)} \\ M_{-\Lambda^2} &= M^{(1,-2)} \;, \end{aligned} \tag{3.11}$$

which is unique up to irrelevant conjugacy. From this we can read off the allowed (magnetic, electric) quantum numbers of the massless monopoles/dyons. They indeed give back the coefficient of the logarithmic term of $\mathcal{F}_D$ that we had anticipated in eq. (3.5).

If we would consider a situation with more than two strong coupling singularities, we would have to solve an equation like (3.10) with the corresponding product of matrices. However, a cursory investigation indicates that such equations for more than three matrices do not have any solution; this can probably be made rigorous, by taking the special (parabolic) form (3.9) of the matrices into account.

4 Solving the monodromy problem

The physics problem has now become a mathematical one, namely simply to find multi-valued functions $a(u), a_D(u)$ that display the required monodromies $M_{\pm\Lambda^2,\infty}$ around the singularities (and that lead to a coupling $\tau \equiv \partial_a a_D$ with $\text{Im}\tau > 0$). This is a classical mathematical problem, the "Riemann Hilbert" problem, which is known to have a unique[3] solution.

The RH problem can be accessed from two complementary point of views: either by considering a, a_D as solutions of a differential equation with regular singular points, or from considering a, a_D as certain "period" integrals. The latter approach, to be discussed momentarily, allows an easy implementation of the right monodromy properties, while the differential equation approach, to be considered later, is more useful for obtaining explicit expressions.

Any two of the monodromy matrices $M_{\pm\Lambda^2,\infty}$ generate the monodromy group Γ_M, which constitutes the subgroup $\Gamma_0(4)$ of the modular group $SL(2,\mathbb{Z})$ and consists of matrices of the form

$$\Gamma_0(4) \;=\; \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2,\mathbb{Z}),\ b = 0 \bmod 4 \right\} . \tag{4.1}$$

This group represents the quantum symmetries of the theory. In particular, we see that $S \equiv \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ (which acts as $\tau \to -\frac{1}{\tau}$) is *not* part of Γ_M, and this means that the theory is not self-dual (in contrast to $N=4$ Yang-Mills theory); however, other transformations do exist that relate weak and strong coupling sectors.

The quantum moduli space can therefore be viewed as

$$\mathcal{M}_q \;\cong\; H^+ \big/ \Gamma_0(4) \ , \tag{4.2}$$

where H^+ is the upper half-plane. Now, motivated by the appearance of a subgroup of the modular group (which is the group of the discontinuous reparametrizations of a torus), the basic idea is that the monodromy problem can be formulated in terms of a toroidal Riemann surface, whose moduli space is precisely $\mathcal{M}_q$ [1]. Such an elliptic curve indeed exists and can be algebraically characterized by:

$$y^2(x,u) \;=\; (x^2-u)^2 - \Lambda^4 \tag{4.3}$$

$$\equiv: \prod_{i=1}^{4} (x - e_i(u,\Lambda)) \ . \tag{4.4}$$

[3]Unique up to multiplication of $\left(\begin{smallmatrix} a_D \\ a \end{smallmatrix}\right)(u)$ by an entire function; this can however be fixed by considering the asymptotic behavior.

The point is to interpret the gauge coupling $\tau(a)$ as the period "matrix" of this torus, and this has the added bonus that manifestly $\mathrm{Im}\tau > 0$, by virtue of a mathematical theorem called "Riemann's second relation". As such τ is defined by a ratio of period integrals:

$$\tau(u) \;=\; \frac{\varpi_D(u)}{\varpi(u)}\ , \tag{4.5}$$

where

$$\varpi_D(u) \;=\; \oint_\beta \omega\ , \qquad\qquad \varpi(u) \;=\; \oint_\alpha \omega \tag{4.6}$$

with $\omega \;\equiv \frac{dx}{y(x,u)}$. Here, α, β are the canonical basis homology cycles of the torus, as shown as follows:

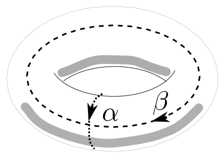


Figure 6: *Basis of one-cycles on the torus.*

From the relation $\tau = \partial_a a_D$ we thus infer that

$$\varpi_D(u) \;=\; \frac{\partial a_D(u)}{\partial u}\ , \qquad\qquad \varpi(u) \;=\; \frac{\partial a(u)}{\partial u}\ . \tag{4.7}$$

That is, the yet unknown functions $a_D(u), a(u)$, and consequently the prepotential $\mathcal{F} = \int_a a_D(a)$, are supposed to be obtained by integrations of torus periods. Note that (4.7) implies that we can also write

$$a_D(u) \;=\; \oint_\beta \lambda\ , \qquad\qquad a(u) \;=\; \oint_\alpha \lambda\ , \tag{4.8}$$

where

$$\lambda \;=\; x^2 \frac{dx}{y(x,u)} \tag{4.9}$$

(up to normalization and total derivatives) is a particular meromorphic one-form.

What needs to be shown is that the periods, derived from the specific choice of elliptic curve given in (4.4), indeed enjoy the correct monodromy properties. The periods (4.6) and (4.8) are actually largely fixed by their monodromy properties around the singularities of $\mathcal{M}_q$, and obviously just reflect the monodromy properties of the basis homology cycles, α and β. It therefore suffices to study how the basis cycles α, β of the torus transform when we loop around a given singularity.

For this, we represent the above torus in a convenient way that is well-known in the mathematical literature: we will represent it in terms of a two-sheeted cover of the branched x-plane. More precisely, denoting the four zeroes of $y^2(x,u)=0$ by

$$e_1 = -\sqrt{u+\Lambda^2}\ , \qquad e_2 = -\sqrt{u-\Lambda^2} \tag{4.10}$$
$$e_3 = \sqrt{u-\Lambda^2}\ , \qquad e_4 = \sqrt{u+\Lambda^2}\ , \tag{4.11}$$
$$\tag{4.12}$$

we can specify the torus in the following way:

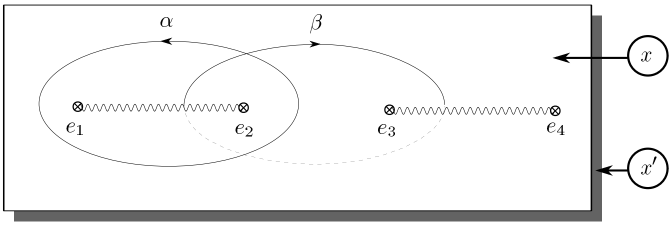


Figure 7: *Representation of the auxiliary elliptic curve in terms of a two-sheeted covering of the branched x-plane. The two sheets are meant to be glued together along the cuts that run between the branch points $e_i(u)$. Shown is our choice of homology basis, given by the cycles α, β. This picture corresponds to the choice of the basepoint $u_0 > \Lambda^2$ real.*

The singularities in the quantum moduli space arise when the torus degenerates, and this obviously happens when any two of the zeros e_i coincide. This can be expressed as the vanishing of the "discriminant"

$$\Delta_\Lambda = \prod_{i<j}^{4}(e_i-e_j)^2 = (2\Lambda)^8(u^2-\Lambda^4)\ . \tag{4.13}$$

The zeroes of Δ_Λ describe the following degenerations of the elliptic curve:

i_+) $u \to +\Lambda^2$, for which $(e_2 \to e_3)$, i.e., the cycle $\nu_{+\Lambda^2} \equiv \beta$ degenerates,

i_-) $u \to -\Lambda^2$, for which $(e_1 \to e_4)$, i.e., the cycle $\nu_{-\Lambda^2} \equiv \beta - 2\,\alpha$ degenerates,

ii) $\Lambda^2/u \to 0$, for which $(e_1 \to e_2)$ *and* $(e_3 \to e_4)$.

Is is now easy to see that a loop $\gamma_{+\Lambda^2}$ around the singularity at $u=\Lambda^2$ makes e_2 and e_3 rotate around each other, so that the cycle α gets transformed into $\alpha-\beta$, as can be seen from Fig.8. This means that on the basis vector $\binom{\beta}{\alpha}$, the monodromy action looks

$$\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \equiv M^{(1,0)} = M_{+\Lambda^2}\ . \tag{4.14}$$

Similarly, from Fig.8 one can see that the monodromy around $u=-\Lambda^2$ is given by

$$\begin{pmatrix} -1 & 4 \\ -1 & 3 \end{pmatrix} \equiv M^{(1,-2)} = M_{-\Lambda^2}\ . \tag{4.15}$$

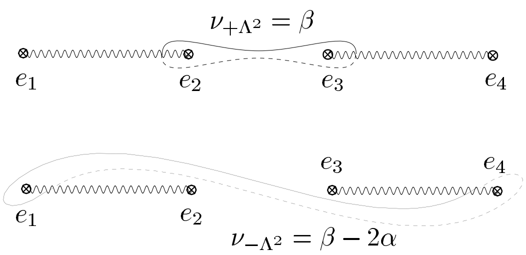


Figure 8: *Vanishing cycles on the torus that shrink to zero as one moves towards a degeneration point.*

To obtain the monodromy around $\Lambda^2/u \to 0$, one can compactify the $u-$plane to P^1, as we did before, and get the monodromy at infinity from the global relation $M_\infty = M_{+\Lambda^2} M_{-\Lambda^2}$ (cf., Fig.5.).

We thus have reproduced the monodromy matrices associated with the exact quantum moduli space directly from the the elliptic curve (4.4), and what this means is that the integrated torus periods $a_D(u), a(u)$ defined by (4.7) must indeed have the requisite monodromy properties. However, before we are going to explicitly determine these functions in the next section, let us say some more words on the general logic of what we have just been doing.

We have seen in Fig.8 that when we loop around a singularity in $\mathcal{M}_q$, the branch points $e_i(u)$ exchange along certain paths, ν, which shrink to zero as $e_i \to e_j$. Such paths are called "vanishing cycles" and play, in a quite general context, an important role for the properties of BPS states. Indeed many features of a BPS spectrum can directly be studied in terms of the singular homology of an auxiliary Riemann surface (or more generally, a K_3 surface or a Calabi-Yau threefold, depending on the specific physical model under consideration).

More concretely, assume that a path vanishes at a given singularity that has the following expansion in terms of given basis cycles:

$$\nu \ = \ g\,\beta + q\,\alpha \ . \tag{4.16}$$

Then obviously, assuming that λ does not blow up, we have

$$0 \ = \ \oint_\nu \lambda \ = \ g \oint_\beta \lambda + q \oint_\alpha \lambda \ = \ g\,a_D + q\,a \ \equiv \ Z \ , \tag{4.17}$$

so that we have at the given singularity a massless BPS state with (magnetic,electric) charges equal to (g, q). That is, we can simply read off the quantum numbers of massless states from the coordinates of the vanishing cycle ! Obviously, under a change of homology basis, the charges change as well, but this is nothing but a duality rotation. What remains invariant is the intersection number

$$\nu_i \circ \nu_j \;=\; \nu^t \cdot \Omega \cdot \nu \;= g_i q_j - g_j q_i \;\in\; \mathbb{Z}\;, \tag{4.18}$$

where $\circ$ is the intersection product of one-cycles and Ω is the symplectic (skew-symmetric) intersection metric for the basis cycles. Note that this represents the well-known Dirac-Zwanziger quantization condition for the possible electric and magnetic charges, and we see that it satisfied by construction. The vanishing of the r.h.s. of (4.18) is required for two states to be local with respect to each other [14, 16]. This means that only states that are related to non-intersecting vanishing cycles are mutually local. In our example, the monopole with charges $(1,0)$, the dyon with charges $(1,-2)$ and the (massive) gauge boson W^+ with charges $(0,2)$ are all mutually non-local, and thus cannot be simultaneously represented in a local effective lagrangian.

Furthermore, there is a closed formula for the monodromy around a given singularity associated with a vanishing cycle ν: the monodromy action on any given cycle, $\gamma \in H_1(X,\mathbb{Z})$, is directly determined in terms of this vanishing cycle by means of the "Picard-Lefshetz" formula [17]:

$$M_\nu: \quad \gamma \;\longrightarrow\; \gamma - (\gamma\circ\nu)\,\nu\;. \tag{4.19}$$

This implies that for a vanishing cycle of the form (4.16), the monodromy matrix is precisely as given in (3.9), as promised.

5 Picard-Fuchs equations

In order to obtain the effective action explicitly, one needs to evaluate the period integrals (4.6). However, instead of directly computing the integrals, one may use the fact that the periods form a system of solutions of the Picard-Fuchs equation associated with the curve (4.4). One then has to evaluate the integrals only in leading order, just to determine the correct linear combinations of the solutions.

Concretely, in order to derive the PF equations, let us first write the defining relation of the curve (4.4) in a homogenous form:

$$W(x,y,z,u) \;\equiv\; (x^2 - u\,z^2)^2 - z^4 - y^2 \;=\; 0\;, \tag{5.1}$$

where we have set $\Lambda = 1$. We also introduce the following integrals over certain globally defined one-forms:

$$\Omega_1 \;=\; \oint_\gamma \frac{1}{W}\,d\omega \tag{5.2}$$

$$\Omega_2 \;=\; \oint_\gamma \frac{x^2 z^2}{W^2}\,d\omega\;, \tag{5.3}$$

where γ is a one-cycle that winds around the surface $W=0$, and $d\omega$ is a volume form. Then one easily finds:

$$
\begin{aligned}
\frac{\partial}{\partial u}\Omega_1 &= \oint_\gamma \frac{2z^2(x^2-u\,z^2)}{W^2}d\omega & (5.4)\\
&= -\frac{2}{(u^2-1)}\Omega_2 - \oint_\gamma \frac{u\,z}{2(u^2-1)}\frac{\partial_z W}{W^2}d\omega\ , & (5.5)
\end{aligned}
$$

where we have used in the second line the following expansion into "ring elements and vanishing relations":

$$
2z^2(x^2-u\,z^2)\ \equiv\ -\frac{2}{(u^2-1)}x^2z^2 - \frac{u}{2(u^2-1)}z\partial_z W\ . \tag{5.6}
$$

Integrating by parts we can cancel W in the second term to get

$$
\frac{\partial}{\partial u}\Omega_1 \ =\ -\frac{2}{(u^2-1)}\Omega_2 - \frac{u}{2(u^2-1)}\Omega_1\ . \tag{5.7}
$$

We can repeat a similar game for Ω_2, and obtain, after multiple partial integrations and expansions similar to (5.6), the following differential identity:

$$
\begin{aligned}
\frac{\partial}{\partial u}\Omega_2 &= \oint_\gamma x\,z^4\frac{\partial_x W}{W^3}d\omega \\
&= \frac{1}{8(u^2-1)}\Omega_1 + \frac{u}{2(u^2-1)}\Omega_2\ . & (5.8)
\end{aligned}
$$

We now can eliminate Ω_2 from (5.5) and (5.8) to obtain a differential equation for the fundamental period: $\mathcal{L}\Omega_1=0$, with $\mathcal{L}=(\Lambda^4-u^2)\partial_u^2-2u\partial_u-\frac{1}{4}$. This Picard-Fuchs equation is supposed to be satisfied by all the periods, in particular by $(\varpi_D(u),\varpi(u))\equiv(\partial_u a_D,\partial_u a)$. In terms of the variable $\alpha = u^2\ over\Lambda^4$, the PF differential operator turns into $(\theta_\alpha=\alpha\partial_\alpha)$

$$
\mathcal{L}=\theta_\alpha(\theta_\alpha-\frac{1}{2})-\alpha(\theta_\alpha+\frac{1}{4})^2\ , \tag{5.9}
$$

which constitutes the hypergeometric system $F(\frac{1}{4},\frac{1}{4};\frac{1}{2};\alpha)$.

It is also possible to derive a second order differential equation for the section (a_D,a) directly. In fact, one easily verifies that $\mathcal{L}\partial_u=\partial_u\tilde{\mathcal{L}}$ with

$$
\tilde{\mathcal{L}}=\theta_\alpha(\theta_\alpha-\frac{1}{2})-\alpha(\theta_\alpha-\frac{1}{4})^2\ , \tag{5.10}
$$

and this forms the hypergeometric system $F(-\frac{1}{4},-\frac{1}{4};\frac{1}{2},\alpha)$. One may also verify directly that $\tilde{\mathcal{L}}\cdot\oint\lambda=0$.

The solutions of $\tilde{\mathcal{L}}\,(a_D(u),a(u))=0$ in terms of hypergeometric functions, and their analytic continuation over the complex plane, are of course well known. For $|u|>|\Lambda|$ a system of solutions to the Picard-Fuchs equations is given by w_0 and w_1 with

$$w_0(u) = \frac{\sqrt{u}}{\Lambda}\sum c(n)(\frac{\Lambda^4}{u^2})^n\,, \qquad c(n) = \frac{(\frac{1}{4})_n(-\frac{1}{4})_n}{(1)_n^2} \tag{5.11}$$

$$w_1(u) = w_0(u)\log(\frac{\Lambda^4}{u^2}) + \frac{\sqrt{u}}{\Lambda}\sum d(n)(\frac{\Lambda^4}{u^2})^n\,, \tag{5.12}$$

where

$$d(n) = c(n)(2(\psi(1)-\psi(n+1)) + \psi(n+\frac{1}{4}) - \psi(\frac{1}{4}) + \psi(n-\frac{1}{4}) - \psi(-\frac{1}{4})) \tag{5.13}$$

and where $(a)_m \equiv \Gamma(a+m)/\Gamma(a)$ is the Pochhammer symbol. Matching the asymptotic expansions of the period integrals one finds

$$a(u) = \frac{\Lambda}{\sqrt{2}}w_0(u)\,, \qquad a_D(u) = -\frac{i\Lambda}{\sqrt{2}\pi}(w_1(u) + (4-6\log(2))w_0(u)), \tag{5.14}$$

which transform under counter-clockwise continuation of u along γ_∞ (c.f., Fig.5) precisely as in (3.8). These expansions correspond to particular linear combinations of hypergeometric functions, the most concise form of which are

$$a_D(\alpha) = \frac{i}{4}\Lambda(\alpha-1)\,{}_2F_1\big(\frac{3}{4},\frac{3}{4},2;1-\alpha\big) \tag{5.15}$$

$$a(\alpha) = \frac{1}{1+i}\Lambda(1-\alpha)^{1/4}\,{}_2F_1\big(-\frac{1}{4},\frac{3}{4},1;\frac{1}{1-\alpha}\big)\,. \tag{5.16}$$

From these expressions, the prepotential in the semi-classical regime near infinity in the moduli space can readily be computed to any given order. Inverting $a(u)$ as series for large a/Λ yields for the first few terms $\frac{u(a)}{\Lambda^2} = 2\left(\frac{a}{\Lambda}\right)^2 + \frac{1}{16}\left(\frac{\Lambda}{a}\right)^2 + \frac{5}{4096}\left(\frac{\Lambda}{a}\right)^6 + O(\left(\frac{\Lambda}{a}\right)^{10})$. After inserting this into $a_D(u)$, one obtains $\mathcal{F}$ by integration w.r.t. a as follows:

$$\mathcal{F}(a) = \frac{i\,a^2}{2\pi}\left(2\log\frac{a^2}{\Lambda^2} - 6 + 8\log 2 - \sum_{\ell=1}^{\infty} c_\ell\left(\frac{\Lambda}{a}\right)^{4\ell}\right)\,. \tag{5.17}$$

It has indeed the form advertised in (2.9). Specifically, the first few terms of the instanton expansion are:

ℓ	1	2	3	4	5	6	7	8
c_ℓ	$\frac{1}{2^5}$	$\frac{5}{2^{14}}$	$\frac{3}{2^{18}}$	$\frac{1469}{2^{31}}$	$\frac{4471}{2^{34}\cdot 5}$	$\frac{40397}{2^{43}}$	$\frac{441325}{2^{47}\cdot 7}$	$\frac{866589165}{2^{64}}$

One can treat the dual magnetic semi-classical regime is an analogous way. Near the point $u=\Lambda^2$ where the monopole becomes massless, we introduce $z=(u-\Lambda^2)/(2\Lambda^2)$ and rewrite the Picard-Fuchs operator as

$$\mathcal{L} = z(\theta_z - \frac{1}{2})^2 + \theta_z(\theta_z - 1) \tag{5.18}$$

At $z = 0$, the indices are 0 and 1, and we have again one power series

$$w_0(z) = \Lambda^2 \sum c(n) z^{n+1}, \qquad c(n) = (-1)^n \frac{(\frac{1}{2})_n^2}{(1)_n (2)_n} \tag{5.19}$$

and a logarithmic solution

$$w_1(z) = w_0(z) \log(z) + \sum d(n) z^{n+1} - 4 \ , \tag{5.20}$$

with

$$d(n) = c(n) \Big(2(\psi(n + \frac{1}{2}) - \psi(\frac{1}{2})) + \psi(1) - \psi(n+1) + \psi(2) - \psi(n+2) \Big) \ . \tag{5.21}$$

For small z one can easily evaluate the lowest order expansion for the integrals (5.16) and thereby determine the analytic continuation of the solutions from the weak coupling to the strong coupling domain:

$$a_D \ = \ 2 \int_{e_2}^{e_3} \lambda \ = \ i\Lambda(z + \ldots) \ = \ i\Lambda w_0(z) \tag{5.22}$$

$$a \ = \ 2 \int_{e_1}^{e_2} \lambda \ = \ \frac{\Lambda}{2\pi}(4 + z(1 + 4\log(2)) - z\log(z) + \ldots) \tag{5.23}$$

$$= \ -\frac{\Lambda}{2\pi}(w_1(z) - (1 + \log(2)) w_0(z)) \, . \tag{5.24}$$

This exhibits the monodromy of (4.14) along the path $\gamma_{+\Lambda^2}$. Inverting $a_D(z)$ yields $z(a_D) = -2\tilde{a}_D + \frac{1}{4}\tilde{a}_D^2 + \frac{1}{32}\tilde{a}_D^3 + \mathcal{O}(\tilde{a}_D^4)$, with $\tilde{a}_D \equiv i a_D/\Lambda$. After inserting this into $a(z)$ we integrate w.r.t. a_D and obtain the dual prepotential $\mathcal{F}_D$ as follows:

$$\mathcal{F}_D(a_D) = \frac{i\Lambda^2}{2\pi} \left(\tilde{a}_D^2 \log\Big[-\frac{i}{4}\sqrt{\tilde{a}_D} \Big] + \sum_{\ell=1}^{\infty} c_\ell^D \, \tilde{a}_D^\ell \right) , \tag{5.25}$$

where the lowest threshold correction coefficients c_ℓ^D are

ℓ	1	2	3	4	5	6	7	8
c_ℓ^D	4	$-\frac{3}{4}$	$\frac{1}{2^4}$	$\frac{5}{2^9}$	$\frac{11}{2^{12}}$	$\frac{63}{2^{16}}$	$\frac{527}{2^{18} \cdot 5}$	$\frac{3129}{2^{24}}$

They reflect properties of the massive BPS spectrum near $u = \Lambda^2$.

6 Generalization to $SU(n)$: classical theory

The above construction can be generalized to include extra matter fields [1], and also to other gauge groups [4, 5]; an approach via integrable systems was presented in [19, 18].

We will here just outline some of the group theoretical aspects for $G = SU(n)$, and present the discussion in a particular way that follows [20]: namely by starting with the classical theory. Indeed many interesting features appear in a simplified fashion already at the classical level, and some of these features play an important role in the generalization to string theory. We will only briefly explain how the quantum theory fits into the general picture, and refer the reader for more details to the literature.

Just like as for $G = SU(2)$, the scalar superfield component ϕ labels a continuous family of inequivalent ground states that constitutes the classical moduli space, $\mathcal{M}_c$. One can always rotate ϕ into the Cartan sub-algebra, $\phi = \sum_{k=1}^{n-1} a_k H_k$, with $H_k = E_{k,k} - E_{k+1,k+1}$, $(E_{k,l})_{i,j} = \delta_{ik}\delta_{jl}$. For generic eigenvalues of ϕ, the $SU(n)$ gauge symmetry is broken to the maximal torus $U(1)^{n-1}$, whereas if some eigenvalues coincide, some larger, non-abelian group $H \subseteq G$ remains unbroken. Precisely which gauge bosons are massless for a given background $a = \{a_k\}$, can easily be read off from the central charge formula,

$$Z_q(a) \;=\; q \cdot a \;, \qquad \text{with} \quad m^2(q) \;=\; |Z_q|^2 \;, \tag{6.1}$$

where we take here for the charge vectors q the roots $\alpha \in \Lambda_R(G)$ in Dynkin basis.

The Cartan sub-algebra variables a_k are not gauge invariant and in particular not invariant under discrete Weyl transformations. Therefore, one introduces other variables for parametrizing the classical moduli space, which are given by the Weyl invariant Casimirs $u_k(a)$. These variables parametrize the Cartan sub-algebra modulo the Weyl group, ie, $\{u_k\} \cong \mathbb{C}^{n-1}/S(n)$, and can be obtained by a Miura transformation:

$$\prod_{i=1}^{n} (x - Z_{\lambda_i}(a)) \;=\; x^n - \sum_{l=0}^{n-2} u_{l+2}(a)\, x^{n-2-l} \;\equiv\; W_{A_{n-1}}(x,u) \;. \tag{6.2}$$

Here, λ_i are the weights of the n-dimensional fundamental representation, and $W_{A_{n-1}}(x,u)$ is nothing but the "simple singularity" [17] associated with $SU(n)$, with

$$u_k(a) \;=\; (-1)^{k+1} \sum_{j_1 \neq \ldots \neq j_k} Z_{\lambda_{j_1}} Z_{\lambda_{j_2}} \ldots Z_{\lambda_{j_k}}(a) \;. \tag{6.3}$$

These symmetric polynomials are manifestly invariant under the Weyl group $S(n)$, which acts by permutation of the weights λ_i.

From the above we know that whenever $Z_{\lambda_i}(a) = Z_{\lambda_j}(a)$ for some i and j, there are, classically, extra massless non-abelian gauge bosons, since $Z_\alpha = 0$ for some root α. For such backgrounds the effective action becomes singular. The classical moduli space is thus given by the space of Weyl invariant deformations modulo such singular regions: $\mathcal{M}_0 = \{u_k\} \backslash \Sigma_0$. Here, $\Sigma_0 \equiv \{u_k : \Delta_0(u_k) = 0\}$ is the zero locus of the discriminant

$$\Delta_0(u) \;=\; \prod_{i<j}^{n} (Z_{\lambda_i}(u) - Z_{\lambda_j}(u))^2 \;=\; \prod_{\substack{\text{positive}\\ \text{roots } \alpha}} (Z_\alpha)^2(u) \;, \tag{6.4}$$

of the simple singularity (6.2). We schematically depicted the singular loci Σ_0 for $n = 2, 3, 4$ in Fig.9.

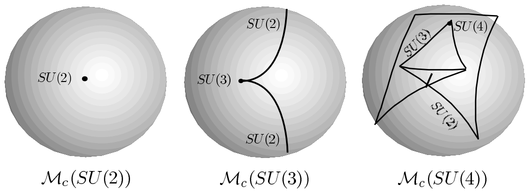


Figure 9: *Singular loci Σ_0 in the classical moduli spaces $\mathcal{M}_0$ of pure $SU(n)$ $N = 2$ Yang-Mills theory. They are nothing but the bifurcation sets of the type A_{n-1} simple singularities, and reflect all possible symmetry breaking patterns in a gauge invariant way (for $SU(3)$ and $SU(4)$ we show only the real parts). The picture for $SU(4)$ is known in singularity theory as the "swallowtail".*

The discriminant loci Σ_0 are generally given by intersecting hypersurfaces of complex codimension one. On each such surface one has $Z_\alpha = 0$ for some pair of roots $\pm\alpha$, so that there is an unbroken $SU(2)$. Furthermore, since $Z_\alpha = 0$ is a fixed point of the Weyl transformation r_α, the Weyl group action is singular on these surfaces. On the intersections of these surfaces one has, correspondingly, larger unbroken gauge groups. All planes together intersect in just one point, namely in the origin, where the gauge group $SU(n)$ is fully restored. Thus, what we learn is that all possible classical symmetry breaking patterns are encoded in the discriminants of $W_{A_{n-1}}(x, u)$.

In previous sections we have seen that $SU(2)$ quantum Yang-Mills theory is characterized by an auxiliary elliptic curve. In a more general context, one may it view as a "spectral", or "level" manifold. The relationship between BPS states and cycles on an auxiliary manifold seems actually to be generic; indeed one may introduce a similar concept here and characterize BPS states (the non-abelian gauge bosons) of classical $SU(n)$ $N=2$ Yang-Mills theory by some auxiliary manifold X. This level manifold X is zero dimensional and simply given by the following set of points:

$$X \;:\; \{\, x \,:\, W_{A_{n-1}}(x, u) = 0 \,\} \;=\; \{\, Z_{\lambda_i}(u) \,\} \,. \qquad (6.5)$$

It is singular if any two of the $Z_{\lambda_i}(u)$ coincide, and indeed, the vanishing cycles are just given by the differences: $\nu_\alpha = Z_{\lambda_i} - Z_{\lambda_j} = Z_\alpha$, ie., by the central charges associated with the non-abelian gauge bosons. It is indeed well-known [17] that ν_α generate the root lattice: $H_0(X, \mathbb{Z}) \cong \Lambda_R$. We depicted the level surface for $G = SU(3)$ in Fig.10.

Pictures like Fig.10 have a very concrete group theoretical meaning. In fact, if we choose as here a special region in the moduli space where only the top Casimir, u_n, is non-zero, the picture becomes Z_n symmetric and actually a certain projection of the weights λ_i. More precisely, a solution vector ξ to the equations: $u_k(\xi) = 0$ ($k = 2, ..., n-1$), $u_n(\xi) = 1$

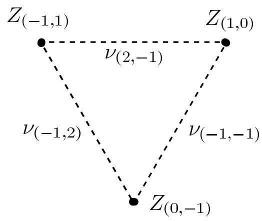


Figure 10: *Level manifold for classical $SU(3)$ Yang-Mills theory, given by points in the x-plane that form a weight diagram. The dashed lines are the vanishing cycles associated with non-abelian gauge bosons (having corresponding quantum numbers, here in Dynkin basis). The masses are proportional to the lengths of the lines and thus vanish if the cycles collapse. This kind of pictures also has an interpretation in terms of D-branes.*

is indeed known [24] to be a projection vector that projects the $n-1$-dimensional weight space on the Coxeter eigenspace with eigenvalue $e^{2\pi i/n}$ (further group theoretical aspects were discussed in [23]). It thus follows trivially that the level surface (6.5)

$$\{ Z_{\lambda_i}(x, u_k = 0, u_n = 1) \} \equiv \{ \lambda_i \cdot \xi \} \tag{6.6}$$

consists precisely of the projected weight vectors.

We thus see a close connection between the vanishing homology of X and $SU(n)$ weight space. Indeed, the intersection numbers of the vanishing cycles are just given by the inner products between root vectors, $\nu_{\alpha_i} \circ \nu_{\alpha_j} = \langle \alpha_i, \alpha_j \rangle$ (self-intersections counting $+2$), and the Picard-Lefshetz formula (4.19) coincides in this case with the well-known formula for Weyl reflections, with matrix representation: $M_{\alpha_i} = \mathbf{1} - \alpha_i \otimes w_i$ (where w_i are the fundamental weights). Most of the above considerations apply more or less directly to the other simply laced Lie groups of type D and E, for which simple simple singularities of the corresponding type are relevant [17, 5].

Note that the corresponding "classical" situation arises in string theory when one compactifies a type IIA string on $K3$; the euclidean $SU(n)$ weight space is then effectively replaced by the lorentzian lattice of vanishing two-cycles $\nu_i \in H_2(K_3, \mathbb{Z})$ (which is isomorphic to the Narain lattice $\Gamma_{20,4}$ of the dual heterotic string formulation[2, 21, 22]). Locally, near a point of $SU(n)$ enhanced gauge symmetry, the $K3$ surface has a singularity of type A_{n-1} and looks:

$$X \ : \ \{ W_{A_{n-1}}(x, u) + z^2 + w^2 = \ 0 \} \ . \tag{6.7}$$

We thus see some sort of universality at work, in that only the local neighborhood of the singularity is relevant. Indeed, the local vanishing homology of the ALE space (6.7) can be represented by the same kind of pictures as Fig.10 [22]. The dashed lines will however in this case not correspond to vanishing 0-cycles, but to vanishing two-cycles; locally, $H_2(K_3, \mathbb{Z}) \to H_2(ALE, \mathbb{Z}) \cong \Lambda_R(SU(n))$.

Actually, one can give in this context pictures like Fig.10 a very concrete physical interpretation: in an appropriate dual formulation of the type IIA string theory compactified on $K3$, the dots simply represent locations of 5-D-branes and the dashed lines (the vanishing cycles) open strings [22]. It seems indeed to be a quite general rule that level surfaces, with vanishing cycles sitting on them, represent in some dual way a physical arrangement of extended physical objects.

7 Quantum $SU(n)$ gauge theory

We now turn to the quantum version of the $N=2$ Yang-Mills theories, where the issue is to construct curves X whose moduli spaces $\mathcal{M}_\Lambda$ give the supposed quantum moduli spaces. We have seen that the classical theories are characterized by simple singularities, so we may expect that the quantum versions should also have something to do with them. Indeed, for $G = SU(n)$ the appropriate manifolds were found in [4] and are given by

$$X: \; y^2 \; = \; \Big(W_{A_{n-1}}(x,u_i)\Big)^2 - \Lambda^{2n} \quad , \tag{7.1}$$

which corresponds to special genus $g = n-1$ hyperelliptic curves. Above, Λ is the dynamically generated quantum scale.

Since y^2 factors into $W_{A_{n-1}} \pm \Lambda^n$, the situation is in some respect like two copies of the classical theory, with the top Casimir u_n shifted by $\pm\Lambda^n$. Specifically, the "quantum" discriminant, whose zero locus Σ_Λ gives the singularities in the quantum moduli space $\mathcal{M}_q$, is easily seen to factorize as follows:

$$\Delta_\Lambda(u_k,\Lambda) \;\; \equiv \;\; \prod_{i<j}(Z^+_{\lambda_i} - Z^+_{\lambda_j})^2(Z^-_{\lambda_i} - Z^-_{\lambda_j})^2 \; = \; \text{const.}\, \Lambda^{2n^2}\delta_+\,\delta_- \; , \tag{7.2}$$

$$\delta_\pm(u_k,\Lambda) \;\; = \;\; \Delta_0(u_2,...,u_{n-1},u_n \pm \Lambda^n) \; , \tag{7.3}$$

is the shifted classical discriminant (6.4). Thus, Σ_Λ consists of two copies of the classical singular locus Σ_0, shifted by $\pm\Lambda^n$ in the u_n direction. Obviously, for $\Lambda \to 0$, the classical moduli space is recovered: $\Sigma_\Lambda \to \Sigma_0$. That is, when the quantum corrections are switched on, a single isolated branch of Σ_0 (associated with massless gauge bosons of a particular $SU(2)$ subgroup) splits into two branches of Σ_Λ (reflecting two massless Seiberg-Witten dyons related to this $SU(2)$). For $G = SU(3)$, this is depicted in Fig.11.

Moreover, the points Z_{λ_i} of the classical level surface (6.5) split as follows,

$$Z_{\lambda_i}(u) \;\to\; Z^\pm_{\lambda_i}(u,\Lambda) \equiv \; Z_{\lambda_i}(u_2,,...,u_{n-1},u_n \pm \Lambda^n) \; , \tag{7.4}$$

and become the $2n$ branch points of the Riemann surface (7.1). The curve can accordingly be represented by the two-sheeted x-plane with cuts running between pairs $Z^+_{\lambda_i}$ and $Z^-_{\lambda_i}$. See Fig.12 for an example.

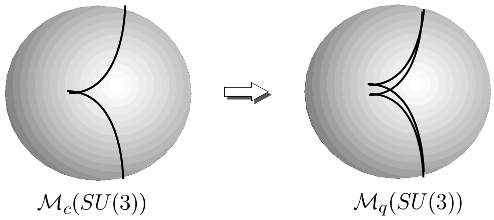


Figure 11: *When switching to the exact quantum theory, the classical singular locus splits into two quantum singularities that are associated with massless dyons; this is completely analogous to Fig.3. The distance is governed by the quantum scale* Λ.

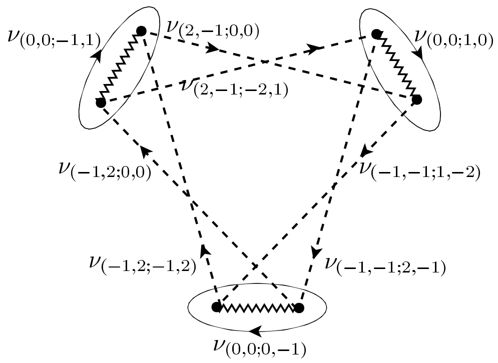


Figure 12: *The level manifold of quantum* $SU(3)$ *Yang-Mills theory is given by a genus two Riemann surface, which is represented here as a two-sheeted cover of the* x*-plane. It may be viewed as the quantum version of the classical, zero dimensional level surface of Fig.10, whose points transmute into pairs of branch points. The dashed lines represent the vanishing cycles (on the upper sheet) that are associated with the six branches of the singular locus* $\Sigma_\Lambda(SU(3))$*; these give rise to six kinds of massless dyons. The quantum numbers refer to* (g,q) *(where* g,q *are weight vectors in Dynkin basis), and can easily be determined from the weight space projection.*

Just like for the classical level surfaces, the vanishing cycles of the Riemann surfaces (7.1) have a concrete group theoretical meaning. Not only can one determine the quantum numbers of the massless dyons by just expanding the vanishing cycles in some appropriate symplectic basis, one finds that one can even more directly associate the cycles in the branched x-plane with projections of roots and weights.

More precisely, Fig.12 can be thought of as a quantum deformation of the classical level surface of Fig.10, whose points, associated with projected weight vectors λ_i, turn into branch cuts (whose length is governed by the quantum scale, Λ). In fact, one obtains two, slightly rotated copies of the weight diagram. A basis of cycles can be chosen such that the coordinates of the "electric", α-type of cycles are given by precisely the weight vectors λ_i. Moreover, the classical cycles of Fig.10 turn into pairs of "magnetic", β-type of cycles, and we can immediately read off the electric and magnetic quantum numbers of the massless dyons: both electric and magnetic charges are given by root vectors.

References..

1 N. Seiberg and E. Witten, Nucl. Phys. B426 (1994) 19, `hep-th/9407087`; Nucl. Phys. B431 (1994) 484, `hep-th/9408099`.

2 C. Hull and P. Townsend, Nucl. Phys. B438 (1995) 109, `hep-th/9410167`.

3 A. Sen, *Int. J. Mod. Phys.* **A9** (1994) 3007.

4 A. Klemm, W. Lerche, S. Theisen, and S. Yankielowicz, Phys. Lett. B B344 (1995) 169, `hep-th/9411048`;
P. Argyres and A. Faraggi, Phys. Rev. Lett. **74** (1995) 3931, `hep-th/9411057`;
M. Douglas and S. Shenker, Nucl. Phys. B447 (1995) 271, `hep-th/9503163`;
P. Argyres and M. Douglas, Nucl. Phys. B448 (1995) 93, `hep-th/9505062`;
A. Klemm, W. Lerche and S. Theisen, *Nonperturbative Effective Actions of N=2 Supersymmetric Gauge Theories*, preprint CERN-TH/95/104, LMU-TPW 95-7, `hep-th/9505150`.

5 U. Danielsson and B. Sundborg, Phys. Lett. B B358 (1995) 273, `hep-th/9504102`;
A. Hanany and Y. Oz, Nucl. Phys. B452 (1995) 283, `hep-th/9505075`;
P. Argyres, M. Plesser and A. Shapere, Phys. Rev. Lett. 75 (1995) 1699, `hep-th/9505100`;
A. Brandhuber and K. Landsteiner, Phys. Lett. B B358 (1995) 73, `hep-th/9507008`;
P. Argyres and A. Shapere, *The Vacuum Structure of N=2 SuperQCD with Classical Gauge Groups*, preprint RU-95-61, UK/95-14, `hep-th/9509175`;
A. Hanany, *On the Quantum Moduli Space of N=2 Supersymmetric Gauge Theories*, preprint IASSNS-HEP-95/76, `hep-th/9509176`.

6 A. Strominger, Nucl. Phys. B451 (1995) 96, `hep-th/9504090`;
B. Greene, D. Morrison and A. Strominger, Nucl. Phys. B451 (1995) 109, `hep-th/9504145`;
S. Kachru and C. Vafa, Nucl. Phys. B450 (1995) 69, `hep-th/9506024`;
S. Ferrara, J. Harvey, A. Strominger, and C. Vafa, Phys. Lett. B 361 (1995) 59, hep-th/9505162, `hep-th/9505162`;
C. Vafa and E. Witten, *Dual String Pairs With N=1 And N=2 Supersymmetry In Four Dimensions*, preprint HUTP-95/A023, IASSNS-HEP-95-58, `hep-th/9507050`;
S. Kachru, A. Klemm, W. Lerche, P. Mayr and C. Vafa, Nucl. Phys. B459 (1996) 537, `hep-th/9508155`.

7 A. Ceresole, R. D'Auria and S. Ferrara, Phys. Lett. B B339 (1994) 71, `hep-th/9408036`;
A. Ceresole, R. D'Auria, S. Ferrara and A. Van Proeyen, Nucl. Phys. B444 (1995) 92, `hep-th/9502072`;
M. Billo', A. Ceresole, R. D'Auria, S. Ferrara, P. Fre', T. Regge, P. Soriani, A. Van Proeyen, *A Search for Non-Perturbative Dualities of Local N=2 Yang–Mills Theories from Calabi–Yau Threefolds*, preprint SISSA 64/95/EP, POLFIS-TH 07/95, CERN-TH 95/140, IFUM 508/FT, KUL-TF-95/18, UCLA/95/TEP/19, `hep-th/9506075`.

8 P.K. Townsend. *p-Brane Democracy*, preprint R/95/34, `hep-th/9507048`.

9 K.&M. Becker and A. Strominger, *Fivebranes, Membranes and Non-Perturbative String Theory*, `hep-th/9507158`.

10 C. Gomez and R. Hernandez, *Electric-Magnetic Duality and Effective Field Theories*, preprint FTUAM 95/36, `hep-th/9510023`.

11 N. Seiberg, Phys. Lett. B 206(1988) 75.

12 V. Novikov, M. Schifman, A. Vainstein, M Voloshin and V. Zakharov, Nucl. Phys. B229 (1983) 394; V. Novikov, M. Schifman, A. Vainstein and V. Zakharov, Nucl. Phys. B229 (1983) 381, 407; M. Schifman, A. Vainstein and V. Zakharov, Phys. Lett. B 166(1986) 329.

13 C. Montonen and D. Olive, Phys. Lett. B 72 (1977) 117.

14 G. 't Hooft, Nucl. Phys. B190 (1981) 455.

15 D. Olive and E. Witten, Phys. Lett. B 78 (1978) 97.

16 P. Argyres and M. Douglas, as in ref. [4].

17 See e.g., V. Arnold, A. Gusein-Zade and A. Varchenko, *Singularities of Differentiable Maps I, II*, Birkhäuser 1985.

18 E. Martinec and N. Warner, *Integrable systems and supersymmetric gauge theory*, preprint EFI-95-61, USC-95/025, `hep-th/9509161`.

19 A.Gorsky, I.Krichever, A.Marshakov, A.Mironov and A.Morozov,Phys. Lett. B 355 (1995) 466.

20 A. Klemm, W. Lerche and S. Theisen, last reference in [4].

21 E. Witten, Nucl. Phys. B443 (1995) 85, `hep-th/9503124`.

22 M. Bershadsky, V. Sadov and C. Vafa, *D-Strings on D-Manifolds*, preprint HUTP-95/A035, IASSNS-HEP-95-77, hep-th/9510225.

23 W. Lerche and N.P. Warner, Nucl. Phys. B358 (1991) 571.

24 B. Kostant, Am. J. Math. 81 (1959) 973.