Introduction to Seiberg-Witten theory

notes on reading [Ler97]

Valdo Tatitscheff
valdo.tatitscheff@math.unistra.fr
IRMA, 7 rue René Descartes, 67084 Strasbourg Cedex

June 21, 2020

Introduction

Around 1995, remarkable progress has been made to understand non-perturbative aspects of supersymmetric field and string theories, initiated by the work of Seiberg and Witten on $\mathcal{N} = 2$ SU(2) super Yang-Mills theory [SW94a, SW94b] (and Hull and Townsend heterotic - type II duality [HT95]).

Supersymmetric theories display many interesting features. For example, one of these which is oriented towards phenomenology, is that it povides a solution to the gauge hierarchy problem (the fact that the mass of the Higgs is so small compared to the Planck scale). Moreover, supersymmetry often makes computations doable, and some non-perturbative aspects of the theories under study, tractable. Were supersymmetry a mere technical tool, the characteristics it displays are physically relevant anyways. Supersymmetric theories enjoy:

- Non-renormalisation properties, that is, less violent quantum corrections,
- Holomorphicity, which allows the use of complex analysis methods [Sei94, IS97],
- More or less manifest dualities which can be studied quite directly.

In four dimensions, the maximal number of supersymmetries in a non-gravitational field theory is four, for more would yield particles of spin strictly higher than one, and thus be a super-gravity theory. Whereas $\mathcal{N} = 4$ SYM theories are completely rigid (because of self-dualities, no quantum corrections, and finiteness), $\mathcal{N} = 1$ SYM theories are presumably not exactly solvable, and only certain sub-sectors are governed by holomorphic objects. In between, $\mathcal{N} = 2$ theories are exactly solvable in the low energy limit. The low-energy effective theory is governed by a holomorphic prepotential $\mathcal{F}$ which is (perturbatively) one-loop exact. One however expects contributions from gauge instantons, but the form of these is fixed. The knowledge of the prepotential of 4d, $\mathcal{N} = 2$, SU(n) super Yang-Mills theories, possibly with matter, thus merely boils down to the knowledge of some coefficients appearing in it. Seiberg-Witten theory provides a way to compute these coefficients, which even if there is no direct proof of it, not only seems right, but has been checked successfully multiple times, and in multiple ways.

The stream of thoughts is roughly the following. The Coulomb branch of the 4d, $\mathcal{N} = 2$, super Yang-Mills theory with gauge group G and possibly matter, is parametrized by the VEVs of the scalars in the vector multiplets (and more VEVs if there is some matter), modulo the Weyl group of G, which remains unbroken. The low-energy theory at some point in the moduli space is nicely encoded in an effective prepotential, which is an analytic function away from the singularities in the moduli space. The Hessian of the prepotential is the effective complexified gauge coupling, and it forms a metric on the moduli space. One can see that the prepotential (as well as the complexified effective gauge coupling) cannot be an entire function on the moduli space, but is rather multi-valued, i.e. the section of some bundle on the moduli space. Rather than being an issue, this fact is a blessing for the following reason.

Whereas the structure of the (semi-) classical moduli space is straightforward to derive from the Lagrangian, for some values of the VEVs the original theory is strongly coupled (as soon as the VEVs are below the strong coupling scale of the theory in the cases of pure gauge, for example, since $\mathcal{N} = 2$ pure non-abelian gauge theories in 4d are assymptotically free) and the true behavior of the full and quantum theory has no reason to be similar to what one naively expects from semi-classical considerations. Singularities in the quantum moduli space arise as some state in the theory becomes massless. One can compute the monodromy of the prepotential around this singularity as soon as one knows the charges of this state. If one knows, for example, the number of singularities in the quantum moduli space, the constraint that all these monodromies glue together nicely can be enough to fully determine the charges of the special states corresponding to the singularities, and moreover, to compute the coefficients appearing in the effective prepotential locally around every singularity.

Everything can be encoded in a fibration of the quantum moduli space by complex curves called Seiberg-Witten curves. Above each singularity in the moduli space, a 1-cycle in this family of curves degenerates. Since the prepotential is the integral of some periods on the curves, its monodromies can be computed using the Picard-Lefschetz formula. Moreover, since the periods satisfy Picard-Fuchs differential equations whose solutions are known in terms of hypergeometric functions, one gets the instanton coefficients appearing in the prepotential, completely explicitely.

In what follows we focus on pure SU(2) gauge super Yang Mills theory, which corresponds to the case originally studied by Seiberg and Witten [**?**].

1 The semi-classical theory of $\mathcal{N}=2$, SU(2) Yang-Mills Theory

The UV fields of pure $\mathcal{N}=2$ SYM are $\mathcal{N}=2$ vector multiplets Ψ^i in the adjoint representation of $G=\mathrm{SU}(2)$. In $\mathcal{N}=1$ language, eahc of them contains a vector multiplet W^i_μ and chiral multiplet Φ^i. On shell, the former describes a gauge field A^i_μ and a gluino λ^i, and the latter, a fermion ψ^i and a complex scalar ϕ^i. The scalar $\phi=\phi^i\sigma_i$ (the σ_i are the Pauli matrices) enters a potential $V(\phi)=\mathrm{Tr}\big[\phi,\phi^\dagger\big]$. Along the direction $\phi=a\sigma_3$ the latter is identically zero, thus there is (classically, at least) a complex line of inequivalent vacua, parametrised by the gauge invariant function $u(a)=\mathrm{Tr}\,\phi^2=2a^2$ (which is the quadratic Casimir of SU(2).)

Any non-zero a triggers the spontaneous breaking of SU(2) to U(1), however the Weyl group of SU(2) (which is the remnant of the gauge transformations after the symmetry breaking) acts as $a\mapsto -a$ and hence the gauge invariant Casimir is a better coordinate (than a itself) on the moduli space of vacua $\mathcal{M}_c$. The latter is the complex line introduced above, compactified to $\mathbb{P}^1(\mathbb{C})$.

The goal of Seiberg and Witten is to derive the effective low-energy description of the theory, obtained (outside a neighbourhood of $0\in\mathcal{M}_c$) by integrating out the massive non-abelian gauge bosons $W^\pm$, which yields an action that only involves the neutral massless W^0. The point 0 is singular since the $W^\pm$ become massless there, at least classically.

Pure $\mathcal{N}=2$ SYM in four dimensions is asymptotically free. At energies far above the scalar VEV a, the masses of the gauge bosons are neglectable, and below a, one is left with an effective U(1) theory with vanishing β function. Hence, around $\infty\in\mathcal{M}_c$, one is in the (semi-classical) weak-coupling region, while close to $0\in\mathcal{M}_c$ one certainly is at strong coupling, since there is a dynamically generated scale $\Lambda>0$ at which the perturbatively renormalised coupling blows up.

Thanks to supersymmetry, the low-effective energy lagrangian of $\mathcal{N}=2$ SU(2) SYM is completely encoded in a holomorphic prepotential $\mathcal{F}$:

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

where $A\sigma_3:=\Phi$, and with:

$$K(A,\bar{A})=\frac{\partial\mathcal{F}(A)}{\partial A}\bar{A}\quad;\quad \tau(A)=\frac{\partial^2\mathcal{F}(A)}{\partial A^2}\quad .$$

The bosonic part of this lagrangian reads:

$$\mathcal{L}=\mathrm{Im}(\tau)\{\partial a\partial\bar{a}+FF\}+\mathrm{Re}(\tau)F\tilde{F}+... \tag{1}$$

from which one sees that

$$\tau(a):=\frac{\Theta(a)}{\pi}+\frac{8\pi i}{g^2(a)}$$

is the complexified effective gauge coupling.

In [Sei88] Seiberg shows that $\mathcal{F}(A)$ is necessarily of the form:

$$\mathcal{F}(A)=\frac{1}{2}\tau_0A^2+\frac{i}{\pi}A^2\log\left[\frac{A^2}{\Lambda^2}\right]+\frac{1}{2\pi i}A^2\sum_{l=1}^{\infty}c_l\left(\frac{\Lambda}{A}\right)^{4l}\ , \tag{2}$$

where the terms in this sum are respectively the bare term, one-loop effects and instanton effects. For $A_{|\Theta=0}=a>\Lambda$, the instanton sum converges and the physics is dominated by the semi-classical one-loop effects.

As explained in the introduction, Seiberg-Witten theory gives a prediction for the value of the coefficient c_l of the instanton sum. This implies infinitely many predictions for zero momentum correlators between a and gauginos in a non-trivial instanton background, which are topological invariants of the four-manifold one is working with, and related to Donaldson invariants (see [Wit88, Wit94]).

The importance of the global properties of τ and $\mathcal{F}$. Neither τ nor $\mathcal{F}$ are well defined globally, but they are multi-valued sections of some bundle over the moduli space. Indeed, from the generic form of $\mathcal{F}$ one sees that around $\infty\in\mathcal{M}_c$, $\tau=c^{\underline{t}}+\frac{2\pi i}{\pi}\log[\frac{u}{\Lambda^2}]+$ single-valued. Because of the branch cut of the log, winding once around ∞ in $\mathcal{M}_c$ shifts τ to $\tau-4$. Since

$$\tau(a):=\frac{\Theta(a)}{\pi}+\frac{8\pi i}{g^2(a)}\ ,$$

this shift however corresponds to an irrelevant shift of the theta angle. Another way to see that τ cannot be globally defined, is that because of Eq. 1, $\mathrm{Im}(\tau)$ defines a metric on $\mathcal{M}_c$, and unitarity imposes its strict positivity everywhere on $\mathcal{M}_c$. However, $\mathrm{Im}(\tau)$ is harmonic, and hence cannot have a minimum if non-constant and globally defined.

2 The quantum moduli space

The structure of moduli space is expected to be modified by quantum effects, and the "true", quantum moduli space $\mathcal{M}_q$ can be (and actually is) different than the classical moduli space $\mathcal{M}_c$, because of these quantum corrections. While $\mathcal{M}_q$ is still parametrized by the VEV a (and thus has the same topology as $\mathcal{M}_c$) the singularities are typically expected to differ, especially in the strong coupling region where the quantum corrections are not subleading. Around ∞, the physics is dominated by the semi-classical approximation. Seiberg and Witten argue that there must be two singularities at finite values of $a \in \mathcal{M}_q$, located at $u = \pm\Lambda^2$ (Λ is still the dynamically generated scale) (and none at $0 \in \mathcal{M}_q$).

While it is difficult to derive this statement rigorously, some points make this ansatz very compelling.

- The fact that there is no singularity at $0 \in \mathcal{M}_q$ (hence, that there is no extra massless gauge field as $W^\pm$ in the full quantum theory) is consistent with the fact that there is no R-current in the theory, whereas a superconformal theory with massless gauge bosons would have one.
- The global R-symmetry acts as $u \mapsto -u$, and one can appreciate the fact that the set of singularities is invariant under this transformation.
- The corresponding $\mathcal{N} = 1$ theory (which is obtained by adding a mass term for Φ in the Lagrangian) has exactly two vacua, since it is known that the Witten index $\mathrm{Tr}(-1)^F$ is n for $\mathcal{N} = 1$, SU(n) SYM. Hence one expects only two singularities in $\mathcal{M}_q$ (at finite u's).

As usual, one expects the singular points in $\mathcal{M}_q$ to be the locus at which some particles (which are not gauge bosons) become massless. Motivated by 't Hooft ideas on confinement [tH81], Seiberg and Witten postulate that at these singularities certain 't Hooft-Polyakov monopoles become massless.

Intermezzo on BPS states. The $\mathcal{N} = 2$, $d = 4$ super-algebra has a complex central charge $Z = U + iV$, where:

$$\{Q_{\alpha i}, \bar{Q}_{\beta,j}\} = \delta_{ij}\gamma^\mu_{\alpha\beta}P_\mu + \delta_{\alpha\beta}\epsilon_{ij}U + (\gamma_5)_{\alpha\beta}\epsilon_{ij}V \ . \tag{3}$$

The mass of every state in the theory has to exceed the absolute value of its central charge. This inequality is called BPS bound: $m^2 \geq |Z|^2$, and is saturated by a particular class of excitations, called BPS states. These preserve half of the supersymmetries. BPS states are protected against quantum corrections, in the sense that if a state indeed saturates this bound classically, then it does as well in the quantum theory.

In $\mathcal{N} = 2$ SYM theory with gauge group SU(2), the central charge is given by $Z = qa + ga_D$, where q (resp., g) is the electric (resp, magnetic) charge of the state under consideration (a_D is the magnetic dual to the Higgs field a. It belongs to the $\mathcal{N} = 2$ vector multiplet $(W_{\alpha,D}, \Phi^i_D)$ that contains the dual, magnetic, photon A^μ_D).

The study of the electric-magnetic transformations shows that $a_D = \frac{\partial}{\partial a}\mathcal{F}(a)$ (see [SW94a, SW94b]).

At $u = \pm\Lambda^2 \in \mathcal{M}_q$, one has $a \neq 0$ but $a_D = 0$, hence a hypermultiplet with charges $(g, q) = (\pm 1, 0)$ is massless. On the other hand, in the exact theory, $u = 0$ would not imply $a = 0$. Were it true, no gauge boson (which has charges $(0, \pm 2)$) would become massless at $u = 0$ (the relation $u = 2a^2$ is classical and can only hold approximatively around $\infty \in \mathcal{M}_q$, i.e in the weak-coupling region), which is something we want.

Some last remarks. In the semi-classical, "electric" region near $\infty \in \mathcal{M}_q$, the variable a is a good coordinate. One has an appropriate low-energy lagrangian, and the instanton sum converges well. However, if one tries to extend $\mathcal{F}$ to a region outside of the convergence domain of the instanton sum, Equation 2 does not really makes sense anymore. The resummation of the instanton terms yields another form of the lagrangian which converges well in another region of $\mathcal{M}_q$. Near $u = \Lambda^2$ one expects for example the following dual prepotential, which converges well near $a_D = 0$:

$$\mathcal{F}_D(a_d) = \frac{1}{2}\tau_0^D a_D^2 - \frac{i}{4\pi}a_D^2 \log\left[\frac{a_D}{\lambda}\right] - \frac{1}{2\pi i}\Lambda^2 \sum_{l=1}^{\infty} c_l^D \left(\frac{ia_D}{\Lambda}\right)^l \ . \tag{4}$$

Near $u = -\Lambda^2$, one expects something similar to what happens near $u = \Lambda^2$ because of the $u \mapsto -u$ symmetry, and it turns out that to obtain the prepotential one merely has to replace a_D in $\mathcal{F}_D(a_D)$ by $a_D - 2a$ in Equation 4.

The quantum moduli space is divided in three regions, each of them containing a singular point, and coming together with a natural coordinatem, in terms of which the prepotential is expressed. Now, one wants to determine the unknown coefficients in the prepotentials from the assumptions that govern the local physics in each patch, i.e. the coefficient of the logarithm (which encodes one-loop effects). The main point of Seiberg-Witten theory is that *patching together the local descriptions in a globally consistent way completely fixes the theory.*

3 From the monodromy to the spectral curve

The one-loop term determines the monodromy M as one goes around a singularity in $\mathcal{M}_q$. In the semi-classical region near infinity (for which one most easily grasps what is going one):

$$\begin{pmatrix} a_D(u) \\ a(u) \end{pmatrix} \simeq \begin{pmatrix} \frac{i}{\pi}\sqrt{2u}\log(u/\Lambda^2) \\ \sqrt{u/2} \end{pmatrix} , \tag{5}$$

since classically $u = 2a^2$, hence the monodromy matrix around M_∞ is:

$$M_\infty = \begin{pmatrix} -1 & 4 \\ 0 & -1 \end{pmatrix} . \tag{6}$$

On general grounds, one knows that the monodromy around a singularity where a dyon with charges (g, q) becomes massless is:

$$M^{(g,q)} = \begin{pmatrix} 1+qg & q^2 \\ -g^2 & 1-gq \end{pmatrix} \tag{7}$$

This formula can (and will) be derived later from the Picard-Lefschetz formula for vanishing cycles.

The global consistency condition which arises as one glues together the different patches of $\mathcal{M}_q$ is $M_{\Lambda^2}M_{-\Lambda^2} = M_\infty$. Up to an irrelevant conjugacy, there is a unique solution for this equation with the prescribed form of the $M_{\pm\Lambda^2}$, which is $M_{\Lambda^2} = M^{(1,0)}$ and $M_{-\Lambda^2} = M^{(1,-2)}$ (cf. [FMO+97]).

Solving the monodromy problem means finding multi-valued functions $a(u)$ and $a_D(u)$ displaying the required monodromies around the singularities, and such that $\tau = \partial_a a_D$ has a strictly positive imaginary part. This is a typical Riemann-Hilbert problem. There are two complementary approaches to this problem: the first one is to consider a and a_D as solutions of a differential equation with regular singular points, and the other is to view a and a_D as some period integrals related to an auxiliary *spectral surface*. The interest of the first approach is to yield explicit expressions, while the second gives a nice geometric intuition about the monodromy conditions.

It turns out that the group generated by $M_{\Lambda^2}, M_{-\Lambda^2}$ and M_∞ is the Hecke subgroup $\Gamma_0(4)$ of the modular group $\mathrm{SL}_2(\mathbb{Z})$, and thus one learns that the quantum moduli space $\mathcal{M}_q$ is $\mathbb{H}/\Gamma_0(4)$. The group $\Gamma_0(4)$ *represents the quantum symmetries of the theory*, and acts on the gauge coupling $\tau = \frac{\partial a_D(u)}{\partial a(u)} \mapsto \frac{a\tau+b}{c\tau+d}$. Since $S : \tau \mapsto -1/\tau$ does not belong to $\Gamma_0(4)$ one sees that $\mathcal{N} = 2$ SYM with gauge group SU(2) is not invariant under the usual strong-weak duality.

The spectral curve. Motivated by the appearance of the modular group, one is tempted to formulate the monodromy problem in terms of a special class of complex tori, whose moduli space is exactly $\mathcal{M}_q$. Such a family can for example be chosen as the curves defined by the equation:

$$(\Sigma) \quad : \quad y^2 = (x^2-u^2)^2 - \Lambda^4 = \prod_{i=1}^{4}(x - e_i(u,\Lambda)) . \tag{8}$$

One interprets τ as the period matrix of this family, which naturally implies that $\mathrm{Im}(\tau) > 0$ by virtue of Riemann's second bilinear relation. As such τ is the ratio $\bar{\omega}_D(u)/\bar{\omega}(u)$ of period integrals, where $\bar{\omega}_D(u) = \int_\beta \omega$ and $\bar{\omega}(u) = \int_\alpha \omega$, with ω the holomorphic differential $\omega = \frac{1}{2\pi}\frac{dx}{y(x,u)}$. The ansatz $\bar{\omega}_D(u) = \frac{\partial a_D(u)}{\partial u}$ and $\bar{\omega}(u) = \frac{\partial a(u)}{\partial u}$ would imply $\tau = \partial_a a_D$ hence one is expecting to get $a_D(u), a(u)$ and hence $\mathcal{F} = \int_a a_D(a)$ by integrating periods. One can otherwise directly write $a_D(u) = \int_\beta \lambda_{SW}$ and $a(u) = \int_\alpha \lambda_{SW}$ where $\lambda_{SW} = \frac{1}{2\pi}x^2\frac{dx}{y(x,u)}$ is the *Seiberg-Witten differential* a particular meromorphic one-form.

Vanishing cycles and BPS states. The monodromy of the periods reflects the monodromy of the 1-cycles α and β, thus one needs to study how these cycles transforms as one loops around a singularity in $\mathcal{M}_q$. The modular curves defined in Equation 8 are double covers of $\mathbb{C}$, with branching points the e_i's. The expressions of the latter are $e_1 = -\sqrt{u+\Lambda^2}$, $e_2 = -\sqrt{u-\Lambda^2}$, $e_3 = \sqrt{u-\Lambda^2}$ and $e_4 = \sqrt{u+\Lambda^2}$.

The singularities in $\mathcal{M}_q$ arise as the torus degenerates (i.e., as two e_i coalesce). This translates into the vanishing of the discriminant $\prod_{i<j}(e_i - e_j)^2 = (2\Lambda)^8(u^2 - \Lambda^4)$.

- As $u \to \Lambda^2$, e_2 and e_3 are exchanged, and the cycle β degenerates.
- As $u \to -\Lambda^2$, e_1 and e_4 are exchanged, and the cycle $\beta - 2\alpha$ degenerates.
- As $\Lambda^2/u \to 0$, e_1 and e_2 are exchanged, as well as e_3 and e_4.

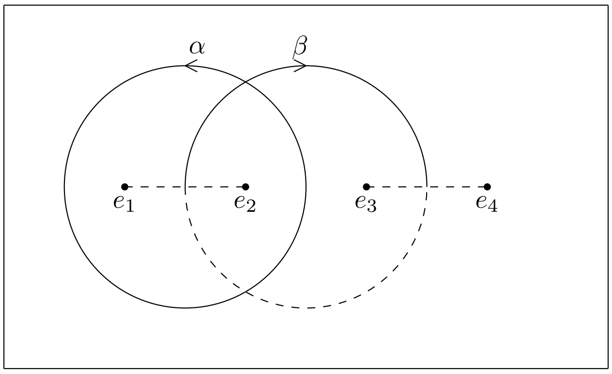

Figure 1: Configuration of the branched cover $\Sigma_1 \to \mathbb{P}^1(\mathbb{C})$ with a choice of a basis (α, β) of $H_1(\Sigma, \mathbb{Z})$

One can see that a loop around $u = \Lambda^2$ in $\mathcal{M}_q$ results in the exchange of e_2 and e_3, and therefore the cycle α becomes $\alpha - \beta$, while the cycle β remains unchanged. Looping around $u = \Lambda^2$ in $\mathcal{M}_q$ transforms β in $-\beta + 4\alpha$ and α in $-\beta + 3\alpha$. Hence on the vector $\begin{pmatrix} \beta \\ \alpha \end{pmatrix}$ the monodromy acts as:

$$M_{\Lambda^2} = M^{(1,0)} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \quad ; \quad M_{-\Lambda^2} = M^{(1,-2)} = \begin{pmatrix} -1 & 4 \\ -1 & 3 \end{pmatrix}$$

A loop around a singularity in $\mathcal{M}_q$ exchanges the branch points $e_i(u)$ along paths ν which shrink to zero as $e_i \leftrightarrow e_j$. Such paths are called *vanishing cycles* and play an important role for the properties of BPS states. Often, features of the BPS spectrum of a theory can be encoded in the homology of an appropriate auxiliary surface X - as in the theory of spectral networks [**?**].

Let us consider a path ν vanishing at a singularity in $\mathcal{M}_q$, and with expression $\nu = g\beta + q\alpha$ is some fixed basis (α, β) of $H^1(\Sigma, \mathbb{Z})$. If λ_{SW} behaves (nicely enough) there, then:

$$0 = \int_\nu \lambda_{SW} = g \int_\beta \lambda_{SW} + q \int_\alpha \lambda_{SW} = g a_D + q a = Z \ , \tag{9}$$

so that at this singularity there is a massless BPS state with (magnetic, electric) charges equal to (g, q). In other words, *the quantum numbers of massless state are the coordinates in homology of the vanishing cycle.* Under a change of basis in $H_1(\Sigma, \mathbb{Z})$, the charges changes as well. These tranformations encode dualities of the theory. The homological intersection number $\nu_i \circ \nu_j = g_i q_j - g_j q_i$ remains invariant. The fact that it is an integer corresponds the Dirac-Zwanziger quantization condition for the possible electric and magnetic charges.

Moreover, two states that are local one to another if and only if their corresponding vanishing cycles have zero intersection (see 't Hooft's paper Nucl. Phys. B190 (1981) 455). For example the monopole of charges $(1, 0)$, the dyon of charges $(1, -2)$ and the massive $W^\pm$ of charges $(0, -2)$ are mutually non-local and hence do not fit in a single local Lagrangian.

There is a closed formula for the monodromy around a singularity corresponding to a vanishing cycle ν:

$$\begin{array}{rcl} M_\nu : H_1(\Sigma, \mathbb{Z}) & \to & H_1(\Sigma, \mathbb{Z}) \\ \gamma & \mapsto & \gamma - (\gamma \circ \nu)\nu \end{array} \ . \tag{10}$$

It is *Picard-Lefschetz formula*, and induces in particular Equation 7.

4 BPS spectrum

We have seen evidence supporting the fact that the exact, quantum theory at hand contains two special BPS excitations of charges $(\pm 1, 0)$ and $\pm(1, -2)$, which become massless respectively at $u = \Lambda^2$ and $u = -\Lambda^2$. What about the other BPS states in the theory?

The charge labels (g, q) are ambiguous. For example, looping around $\infty \in \mathcal{M}_q$ changes the charges (g, q) to $(-g, -q - 4g)$ since:

$$M_\infty M^{(g,q)} M_\infty^{-1} = M^{(-g,-q-4g)} \ . \tag{11}$$

In [Wit79], Witten explains how the electric charge of a dyon changes as the vacuum angle Θ varies.

The monodromy can be induced by arbitrarily small loops around $\infty \in \mathcal{M}_q$, and in particular one can judiciously choose the loops so that they remain inside the weak-coupling region. Hence, beside the massive $W^\pm$ of charges $(0, \pm 2)$, we see that there are BPS dyons with charges $\pm(1, 2l)$ since under shifts $\Theta \to \Theta - 4\pi n$ for $n \in \mathbb{N}$, the BPS spectrum must be mapped to itself (as a whole).

An important point made in Seiberg and Witten's articles [SW94a, SW94b] is that the BPS spectrum at strong coupling is different than the one at weak coupling. This is because the moduli space decomposes in two different regions with different physics (strong coupling and weak coupling). These two regions are separated by a curve $\mathcal{C}$ on which most of the semi-classical BPS states can (and do) decay. It is called a *marginal line of stability*, and it is defined as:

$$\mathcal{C} = \left\{ u : \frac{a_D(u)}{a(u)} \in \mathbb{R} \right\} \ . \tag{12}$$

It is (nearly) an ellipse which goes through the singular points $u = \pm\Lambda^2$. This latter fact is not a surprise since massless BPS states satisfy $ga_D + qa = 0$ (hence $a_D/a \in \mathbb{R}$).

Because of the branch cut of the logarithm, as one follows once $\mathcal{C}$ clockwise from $u = \Lambda^2$ to itself, $(a_D/a)(u)$ varies monotonically from -2 to 2, hence one sees that there really is an ambiguity in the definition of the electric charge of the dyon (see [FB96]).

One the marginal line of stability, the lattice (or Jacobian) of the central charges $Z = ga_D + qa$ degenerates to a line. There the BPS inequality $|Z_{g_1+g_2,q_1+q_2}| \leq |Z_{g_1,q_1}| + |Z_{g_2,q_2}|$ is saturated, hence mass and charge conservation does not prohibit the BPS states to decay into monopoles and dyons anymore. For example, if $a_D = \xi a$, with $\xi \in [0, 2]$, the gauge field with charges $(g, q) = (0, 2)$ and $m_{0,2} = 2|a|$ is unstable against decay into a monopole-dyon pair with $m_{-1,2} = (2-\xi)|a|$ and $m_{1,0} = \xi|a|$. It turns out that the BPS states indeed decay exactly in this manner, and it can be shown directly thanks to the study of anti-self dual strings on Riemann surfaces.

It is also proved in [FB96] that the only stable BPS states in $\mathcal{M}_q^{\text{strong}}$ are precisely the monopole and the dyon of charge $\pm(1, -2)$, and that the semi-classical stable BPS states in $\mathcal{M}_q^{\text{weak}}$ consist exactly of the states mentioned above.

5 Picard-Fuchs and the coefficients of the effective lagrangian

Our main goal is still to obtain the effective action explicitly. Moreover, instead of brute-forcing the computation of the period integrals, one can benefit from the fact that these periods form a system of solutions of the Picard-Fuchs equation associated with the curve Σ. First let us homogenise Eq. 8 to:

$$W(x, y, z, u) = (x^2 - uz^2)^2 - z^4 - y^2z^2 = 0 \ , \tag{13}$$

in which equation we have set $\Lambda = 1$. Now let:

$$\Omega_1 = \int_\gamma \frac{1}{W}\bar{\omega} \quad ; \quad \Omega_2 = \int_\gamma \frac{x^2z^2}{W^2}\bar{\omega} \quad , \tag{14}$$

where γ denotes a generic two-cycle winding around the surface $\{W = 0\}$, and $\bar{\omega}$, a volume form on $\mathbb{P}^2$. One easily computes that:

$$\begin{cases} \dfrac{\partial}{\partial u}\Omega_1 = -\dfrac{2}{u^2-1}\Omega_2 - \dfrac{u}{2(u^2-1)}\Omega_1 \\ \dfrac{\partial}{\partial u}\Omega_2 = \dfrac{1}{8(u^2-1)}\Omega_1 + \dfrac{u}{2(u^2-1)}\Omega_2 \end{cases} \ , \tag{15}$$

which implies a Picard-Fuchs equation $\mathcal{L}_{PF}\Omega_1 = 0$ with Picard-Fuchs operator $\mathcal{L}_{PF} = (\Lambda^4 - u^2)\partial_u^2 - 2u\partial_u - \frac{1}{4}$.

Any period is a solution of this Picard-Fuchs equation, and in particular $(\partial_u a_D(u), \partial_u a(u))$ is a solution. Let $\alpha = \frac{u^2}{\Lambda^4}$, and $\theta_\alpha = \alpha\partial_\alpha$. The Picard-Fuchs operator now is:

$$\mathcal{L} = \theta_\alpha(\theta_\alpha - \frac{1}{2}) - \alpha(\theta_\alpha + \frac{1}{4})^2 \ , \tag{16}$$

which is a *hypergeometric system* of type $(1/4, 1/4; 1/2)$. See [KLT96] for another (direct) derivation of a differential equation for (a_D, a).

One can also write $\mathcal{L}\partial_u = \partial_u\tilde{\mathcal{L}}$ with $\tilde{\mathcal{L}} = \theta_\alpha(\theta_\alpha - \frac{1}{2}) - \alpha(\theta_\alpha - \frac{1}{4})^2$. The sections $a_D(u)$ and $a(u)$ are in the kernel of this operator $\tilde{\mathcal{L}}$, which is the differential equation of a hypergeometric system of type $(-1/4, -1/4, 1/2)$. For $|u| > |\Lambda|$, a basis of solutions to this equation is given by:

$$w_0(u) = \frac{\sqrt{u}}{\Lambda}\sum c(n)\left(\frac{\Lambda^4}{u^2}\right)^n \quad ; \quad w_1(u) = w_0(u)\log\left(\frac{\Lambda^4}{u^2}\right) + \frac{\sqrt{u}}{\Lambda}\sum d(n)\left(\frac{\Lambda^4}{u^2}\right)^n \quad , \tag{17}$$

for some functions c and d of n, whose expression can be found in Leerche, hep-th/9611190.

Matching the asymptotic expansions of the period integrals, one finds

$$a(u) = \frac{\Lambda}{\sqrt{2}}w_0(u), \quad a_D(u) = -\frac{i\Lambda}{\sqrt{2}\pi}[w_1(u) + (4 - 6\log(2))w_0(u)] \ , \tag{18}$$

hence in terms of hypergeometric functions:

$$\begin{cases} a_D(\alpha) = \dfrac{i}{4}\Lambda(\alpha-1)_2F_1(3/4,3/4,2;1-\alpha) \\ a(\alpha) = \dfrac{1}{\sqrt{2}}\Lambda(\alpha^{1/4})_2F_1(-1/4,1/4,1;1/\alpha) \end{cases} . \tag{19}$$

Inverting $a(u)$ (as a series) for large a/Λ yields:

$$\frac{u(a)}{\Lambda^2} = 2\left(\frac{a}{\Lambda}\right)^2 + \frac{1}{16}\left(\frac{a}{\Lambda}\right)^2 + \frac{5}{4096}\left(\frac{a}{\Lambda}\right)^6 + O\left(\left(\frac{a}{\Lambda}\right)^{10}\right)$$

After inserting this series into $a_D(u)$ one merely needs to integrate to find:

$$\mathcal{F}(a) = \frac{ia^2}{2\pi}\left(2\log\frac{a^2}{\Lambda^2} - 6 + 8\log 2 - \sum_{l=1}^{\infty} c_l \left(\frac{\Lambda}{a}\right)^{4l}\right) .$$

It has the form we want, and moreover we now have got the values of the coefficients of the instanton sum:

l	1	2	3	4	5	6
c_l	$\frac{1}{2^5}$	$\frac{5}{2^{14}}$	$\frac{3}{2^{18}}$	$\frac{1469}{2^{31}}$	$\frac{4471}{5\cdot 2^{34}}$	$\frac{40397}{2^{43}}$

Figure 2: The first instanton coefficients

One can treat the dual magnetic semi-classical regime in an analogous way.

A Vacuum Θ-angle

In a quantum gauge theory with gauge group G (in Hamiltonian formulation) the wave function is a functional of the gauge connection A, and the matter fields generically denoted ϕ. Gauge invariance leads to a first-class constraint (of Gauss type). In flat space-time $\mathbb{R}^{1,3}$, the gauge transformations are smooth functions

$$U : \mathbb{R}^3_{\text{space}} \to G$$

which approach 1 at infinity. If a gauge transformation is homotopic to the identity map, it is said to be *small*, and *big* otherwise. It is a fact that if any simple Lie group G satisfies $\pi_3(G) = \mathbb{Z}$. For a small transformation, it is true that $\Psi[UAU^{-1} - (dU)U^{-1}, U\phi] = \Psi[A,\phi]$. However, this is not the case for big gauge transformations, and the Hilbert space might (and does) decomposes into superselection sectors labelled by a phase Θ called vacuum angle. The latter is defined in such a way that for U a (big) gauge transformation with index 1, the states in the sector satisfy $\Psi[UAU^{-1} - (dU)U^{-1}, U\phi] = e^{i\Theta}\Psi[A,\phi]$. The Θ term $(\Theta/32\pi^2)F\tilde{F}$ does not break the gauge invariance (classically at least) and can be introduced in the Lagrangian. It is topological, and gauge instantons typically provide backgrounds with non-zero theta angle. The latter can be related to some kinds of CP violation in non-abelian gauge theories.

B The effective lagrangian of $\mathcal{N} = 2$ SYM

Mostly based on [Sei88].

Supersymmetric gauge theories respect certain perturbative non-renormalisation theorems. For example, the superpotential in $\mathcal{N} = 1$ theories is not renormalised perturbatively, the $\mathcal{N} = 2$ beta functions are one-loop exact, and the $\mathcal{N} = 4$ theories are finite. These theorems are only perturbative though, and one may wonder whether they are be violated by some instanton effects. The holomorphic sector of $\mathcal{N} = 2$ theories is described by a prepotential $\mathcal{F}$ in terms of which:

$$\mathcal{L} = \int \mathrm{d}^4\theta \mathcal{F}(\Psi^a) + c.c\ , \tag{20}$$

which shows that $\mathcal{F}$ has to be *analytic* in the gauge superfields Ψ^a. The latter contain (in $\mathcal{N} = 1$ formalism) a vector multiplet W^a and a chiral multiplet Φ^a, and since in $\mathcal{N} = 1$ langage:

$$\mathcal{L} = \int \mathrm{d}^2\theta \mathrm{d}^2\bar{\theta}\frac{1}{2}K(\Phi^a, \Phi^{\dagger a}, W^a) + \int \mathrm{d}^2\theta f_{bc}(\Phi^a)W^b W^c + c.c\ , \tag{21}$$

one sees that the Kähler potential is given by $2\partial_a \mathcal{F}(e^W)_{ab}\Phi^{\dagger b} + c.c$ and f_{ab}, by $f_{ab} = \frac{1}{2}\partial_a\partial_b\mathcal{F}$. It is clear that the theory has flat directions, and that a potential for the scalars ϕ^a can only arise from the D-terms. A potential for the chiral superfields Φ^a cannot be generated perturbatively or non-perturbatively because it is not compatible with $\mathcal{N} = 2$ SUSY.

A minimal theory would correspond to the choice $\mathcal{F} = (1/8g^2)(\Psi^a)^2$. Its R-symmetry group is SU(2) $\times$ U(1), where the SU(2) factors rotates the two θ's, and the U(1) multiplies them by a phase. The R-symmetry suffers an anomaly in non-trivial instanton backgrounds:

$$\partial_\mu j^\mu = (g^2 N_c/8\pi^2)F\bar{F} \ , \tag{22}$$

(where N_c comes from the gauge group which is here asumed to be SU(N_c)) and is spontaneously broken to $\mathbb{Z}_{4N_c}$.

When $N_c = 2$, the low-energy theory (after the Higgs mechanism induced by the VEV a), contains only one multiplet Ψ which includes the gauge boson for the unbroken direction. There are no non-abelian interactions, hence the infrared behaviour is smooth and one can construct the effective action for the superfield. This latter is still $\mathcal{N} = 2$ supersymmetric, hence the effective lagrangian can also be expressed as $\mathcal{L}_{\text{eff}} = \int \mathrm{d}^4\theta\mathcal{F}_{\text{eff}} + c.c.$

Perturbation theory. Invariance under the U(1) symmetry restrict the form of $\mathcal{F}_{\text{eff}}^{\text{pert}}$ to:

$$\mathcal{F}_{\text{eff}}^{\text{pert}} = (1/8g^2)\Psi^2[A_1 + A_2\log(\Psi^2/\Lambda^2)] \ , \tag{23}$$

where A_1 and A_2 are two constants and Λ is the scale at which g is defined. Changing the value of A_1 changes Λ, and one can choose some value of Λ such that $A_1 = 1$.

Although the effective action is U(1)-invariant, for $A_2 \neq 0$ the effective lagrangian is not, and under a transformation $\alpha \in$ U(1) it gets shifted by $\alpha(A_2/g^2)F\tilde{F}$. Comparing this with Eq. 22, we see that $A_2 = g^2/4\pi^2$. Hence the second term in Eq. 23 is a one-loop effect.

As one writes the effective action in components one finds

$$\mathcal{F}_{\text{eff}}^{\text{pert}} = \left(\frac{1}{8g^2}\right)\left[1 + \frac{3g^2}{4\pi^2} + \left(\frac{g^2}{4\pi^2}\right)\log\left(\frac{\phi^2}{\Lambda^2}\right)\right] \ , \tag{24}$$

from which one can derive the effective gauge coupling.

$\mathcal{N} = 1$ theories are less rigid, and allow field redefinitions yielding fields that have simple U(1) transformations, however these redefinitions are singular and the argument of above does not apply. This is related to the so-called *multiplet of anomalies*. See also the Adler-Bardeen theorem.

Non-perturbative effects While in the perturbative study it seems one can use the U(1) symmetry to constraint the form of the prepotential, this symmetry is broken to $\mathbb{Z}_8$ in non-trivial instanton backgrounds. At the level of one instanton, this U(1) symmetry is expected to be violated by eight units, in which case the contribution of one-instanton to $\mathcal{F}$ has to be of the form $a_1\exp(-8\pi^2/g^2)\Lambda^4/\Phi^2$. It turns out that a_1 cannot depend on g. Symmetry would allow one anti-instanton contribution in Ψ^6/Λ^4, however this contribution would blow-up at weak coupling, hence has to be discarded, since instantons are exponentially suppressed at weak coupling.

The beta functions also depends on the instantons, which thus violate the corresponding perturbative non-renormalisation theorems.

Theta term A theta term $(\Theta/32\pi^2)F\tilde{F}$ can in principle be added to the $\mathcal{N} = 2$ lagrangian, however it can be shifted to zero with help of the U(1) symmetry. However, non-trivial CP violation can be induced by the expectation value of the scalar of the vector multiplet. One then defines Θ to be the coefficient of $(1/32\pi^2)F\tilde{F}$ in the low energy effective action in a given vacuum. The beta functions are given by:

$$\beta_g(g,\Theta) = -(1/4\pi^2)g^3 + 48a_1\cos\Theta g^3\exp(-8\pi^2/g^2) + O(\exp(-16\pi^2/g^2))$$

$$\beta_\Theta(g,\Theta) = -768\pi^2 a_1\sin\Theta\exp(-8\pi^2/g^2) + O(\exp(-16\pi^2/g^2))$$

This analysis can be extended to multi-instantons, and other gauge groups, possibly with matter. For SU(N_c) gauge groups, for instance, the effective prepotential is:

$$\mathcal{F}_{\text{eff}} = \Psi^2(1/8g^2) + (N_c/64\pi^2)\log(\Psi^2/\Lambda^2) + \sum_{l=1}^{\infty} a_l\exp(-8\pi^2 l/g^2)(\Lambda^2/\Psi^2)^{lN_c}) \tag{25}$$

There is a non-renormalisation theorem in perturbation theory around any instanton configuration since the coefficients a_l are independent of g. However it is not know at the time of writing (1988) whether the sum of Eq. 25 converges, or even if some effects weaker than instantons also need to be taken into account.

Eventually, $\mathcal{N} = 4$ theories in $4d$ are so rigid that these non-perturbative corrections (which arise generically for $\mathcal{N} = 2$ theories) are absent. Hence the non-renormalisation theorems asserting that $\mathcal{N} = 4$ theories are finite, hold perturbatively as well as non-perturbatively.

References

[FB96] Frank Ferrari and Adel Bilal. The Strong coupling spectrum of the Seiberg-Witten theory. *Nucl. Phys. B*, 469:387–402, 1996.

[FMO+97] R. Flume, M. Magro, L. O'Raifeartaigh, I. Sachs, and O. Schnetz. Uniqueness of the Seiberg-Witten effective Lagrangian. *Nucl. Phys. B*, 494:331–345, 1997.

[HT95] C.M. Hull and P.K. Townsend. Unity of superstring dualities. *Nucl. Phys. B*, 438:109–137, 1995.

[IS97] Kenneth A. Intriligator and N. Seiberg. Lectures on supersymmetric gauge theories and electric-magnetic duality. *Subnucl. Ser.*, 34:237–299, 1997.

[KLT96] A. Klemm, W. Lerche, and S. Theisen. Nonperturbative effective actions of N=2 supersymmetric gauge theories. *Int. J. Mod. Phys. A*, 11:1929–1974, 1996.

[Ler97] W. Lerche. Introduction to Seiberg-Witten theory and its stringy origin. *Nucl. Phys. B Proc. Suppl.*, 55:83–117, 1997.

[Sei88] Nathan Seiberg. Supersymmetry and Nonperturbative beta Functions. *Phys. Lett. B*, 206:75–80, 1988.

[Sei94] Nathan Seiberg. The Power of holomorphy: Exact results in 4-D SUSY field theories. In *Particles, Strings, and Cosmology (PASCOS 94)*, pages 0357–369, 5 1994.

[SW94a] N. Seiberg and Edward Witten. Electric - magnetic duality, monopole condensation, and confinement in N=2 supersymmetric Yang-Mills theory. *Nucl. Phys. B*, 426:19–52, 1994. [Erratum: Nucl.Phys.B 430, 485–486 (1994)].

[SW94b] N. Seiberg and Edward Witten. Monopoles, duality and chiral symmetry breaking in N=2 supersymmetric QCD. *Nucl. Phys. B*, 431:484–550, 1994.

[tH81] Gerard 't Hooft. Topology of the Gauge Condition and New Confinement Phases in Nonabelian Gauge Theories. *Nucl. Phys. B*, 190:455–478, 1981.

[Wit79] Edward Witten. Dyons of Charge e theta/2 pi. *Phys. Lett. B*, 86:283–287, 1979.

[Wit88] Edward Witten. Topological Quantum Field Theory. *Commun. Math. Phys.*, 117:353, 1988.

[Wit94] Edward Witten. Monopoles and four manifolds. *Math. Res. Lett.*, 1:769–796, 1994.