Recherche Coopérative sur Programme nº 25

Adel Bilal

Duality in $N = 2$ Susy $SU(2)$ Yang-Mills Theory : A Pedagogical Introduction to the Work of Seiberg and Witten

Les rencontres physiciens-mathématiciens de Strasbourg - RCP25, 1997, tome 48
« Conférences de M. Bauer, A. Beauville, O. Babelon, A. Bilal, R. Stora », , exp. nº 4, p. 87-120

<http://www.numdam.org/item?id=RCP25_1997__48__87_0>

DUALITY IN $N = 2$ SUSY $SU(2)$ YANG-MILLS THEORY: A pedagogical introduction to the work of Seiberg and Witten

ADEL BILAL

CNRS - Laboratoire de Physique Théorique de l'Ecole Normale Supérieure★
24 rue Lhomond, 75231 Paris Cedex 05, France
bilal@physique.ens.fr

ABSTRACT

These are notes from introductory lectures given at the Ecole Normale in Paris and at the Strasbourg meeting dedicated to the memory of Claude Itzykson.

I review in considerable detail and in a hopefully pedagogical way the work of Seiberg and Witten on $N = 2$ supersymmetric $SU(2)$ gauge theory without extra matter. This presentation basically follows their original work, except in the last section where the low-energy effective action is obtained emphasizing more the relation between monodromies and differential equations rather than using elliptic curves.

To appear in the Proceedings of the "61. Rencontre entre Physiciens Théoriciens et Mathématiciens", Strasbourg, France, December 1995 dedicated to the memory of Claude Itzykson

---

★ unité propre du CNRS, associé à l'Ecole Normale Supérieure et l'Université Paris-Sud

Ce texte reprend une prépublication du Laboratoire de Physique Théorique de l'Ecole Normale Supérieure de Décembre 1995 - LPTENS-95/53, hep-th@xxx/9601007.

1. Introduction

Although a quite old one, the notion of duality has become most central in field and string theory during the last year and a half. The major breakthrough in field theory was the paper by Seiberg and Witten [1] considering the pure $N = 2$ supersymmetric $SU(2)$ Yang-Mills theory. This work was then generalized to other gauge groups [2,3] and to theories including extra matter fields [4]. In the same time, it became increasingly clear that dualities in string theories play a maybe even more fascinating role (for a brief review see e.g. [5] or [6]). Rather than attempting to give an overview of the situation, in the present notes I will try to give a *pedagogical* introduction to the first paper by Seiberg and Witten [1]. Several other introductions do exist [7], and I hope that the present notes complement them in a useful way.

The idea of duality probably goes back to Dirac who observed that the source-free Maxwell equations are symmetric under the exchange of the electric and magnetic fields. More precisely, the symmetry is $E \to B$, $B \to -E$, or $F_{\mu\nu} \to \tilde{F}_{\mu\nu} = \frac{1}{2}\epsilon_{\mu\nu}{}^{\rho\sigma} F_{\rho\sigma}$. (Here $\epsilon_{\mu\nu\rho\sigma}$ is the flat-space antisymmetric ϵ-tensor with $\epsilon^{0123} = +1$ and $\eta_{\mu\nu}$ has signature $(1,-1,-1,-1)$.) To maintain this symmetry in the presence of sources, Dirac introduced, somewhat ad hoc, magnetic monopoles with magnetic charges q_m in addition to the electric charges q_e, and showed that consistency of the quantum theory requires a charge quantization condition $q_m q_e = 2\pi n$ with integer n. Hence the minimal charges obey $q_m = \frac{2\pi}{q_e}$. Duality exchanges q_e and q_m, i.e. q_e and $\frac{2\pi}{q_e}$. Now recall that the electric charge q_e also is the coupling constant. So duality exchanges the coupling constant with its inverse (up to the factor of 2π), hence exchanging strong and weak coupling. This is the reason why we are so much interested in duality: the hope is to learn about strong-coupling physics from the weak-coupling physics of a dual formulation of the theory. Of course, in classical Maxwell theory we know all we may want to know, but this is no longer true in quantum electrodynamics.

Actually, quantum electrodynamics is not a good candidate for exhibiting a duality symmetry since there are no magnetic monopoles, but the latter naturally appear in spontaneously broken non-abelian gauge theories [8]. Unfortunately, electric-magnetic duality in its simplest form cannot be a symmetry of the quantum theory due to the running of the coupling constant (among other reasons). Indeed, if duality exchanges $\alpha(\Lambda) \leftrightarrow \frac{1}{\alpha(\Lambda)}$ (where $\alpha(\Lambda) = \frac{4\pi}{e^2(\Lambda)}$) at some scale Λ, in general this won't be true at another scale. This argument is avoided if the coupling does not run, i.e. if the β-function vanishes as is the case in certain ($N = 4$)

supersymmetric extensions of the Yang-Mills theory. This and other reasons led Montonen and Olive [9] to conjecture that duality might be an exact symmetry of $N = 4$ susy Yang-Mills theory. A nice review of these ideas can be found in [10].

Let me recall that a somewhat similar duality symmetry appears in the two-dimensional Ising model where it exchanges the temperature with a dual temperature, thereby exchanging high and low temperature analogous to strong and weak coupling. For the Ising model, the sole existence of the duality symmetry led to the exact determination of the critical temperature as the self-dual point, well prior to the exact solution by Onsager. One may view the existence of this self-dual point as the requirement that the dual high and low temperature regimes can be consistently "glued" together. Similarly, in the Seiberg-Witten theory, as will be explained below, duality allows us to obtain the full effective action for the light fields at any coupling (the analogue of the Ising free energy at any temperature) from knowledge of its weak-coupling limit and the behaviour at certain strong-coupling "singularities", together with a holomorphicity requirement that tells us how to patch together the different limiting regimes.

Let me give an overview of how I will proceed. $N = 2$ supersymmetry is central to the work of Seiberg and Witten and to the way duality works, so we must spend some time in the next section to review those notions of supersymmetry that we will need, including the formulation of the $N = 2$ super Yang-Mills action. In section 3, I will discuss the Wilsonian low-energy effective action corresponding to the (microscopic) $N = 2$ super Yang-Mills action for the gauge group $SU(2)$. The original $SU(2)$ gauge symmetry has been broken down to $U(1)$ by the expectation value a of the scalar field ϕ contained in the $N = 2$ multiplet, and the effective action describes the physics of the remaining massless $U(1)$ susy multiplet in terms of an a priori *unknown* function $\mathcal{F}(a)$. $N = 2$ supersymmetry constrains $\mathcal{F}$ to be a (possibly multivalued) *holomorphic* function. Different vacuum expectation values a, or rather different values of the gauge-invariant vacuum expectation value $u = \langle \mathrm{tr}\, \phi^2 \rangle$ lead to physically different theories. So u parametrizes the space of inequivalent vacua, called the moduli space.

In section 4, I will discuss how one defines the duality transformations and show that duality inverts a certain combination τ of the effective coupling constant and the effective theta angle. I will also discuss the spectrum of massive states (BPS mass formula). Let me insist that this duality is an *exact* symmetry of the abelian low-energy *effective* theory, not of the microscopic $SU(2)$ theory. This is different from the Montonen-Olive conjecture about an exact duality symmetry of a microscopic gauge theory.

In section 5, we will study the behaviour of the low-energy effective theory at certain singular points in the moduli space, i.e. at certain values of the parameter u where a magnetic monopole or a dyon (an electrically and magnetically charged state) becomes massless, leading to a singularity of the effective action. These singularities translate into certain monodromies of $a \sim \langle\phi\rangle$ and its dual partner $a_D = \frac{\partial \mathcal{F}(a)}{\partial a}$. In section 6, we put everything together, and I show how to obtain $a(u)$ and $a_D(u)$ and hence $\mathcal{F}(a)$ from the knowledge of these monodromies. Then the low energy effective action is known and the theory solved for all values of the effective coupling constant $\tau = \frac{\partial^2 \mathcal{F}(a)}{\partial a^2}$. Section 6 is the only part where the presentation does not follow the logic of Seiberg and Witten's paper, but I rather emphasize the relation between monodromies and differential equations, and obtain $a(u)$ and $a_D(u)$ as solutions of a hypergeometric equation. I then show how this fits into the reasoning of Seiberg and Witten using elliptic curves. In a concluding section 7, I mention some of the developments that followed the work of Seiberg and Witten described in these notes.

2. Some notions of supersymmetry

Clearly, I cannot give a complete discussion of the theory of $N = 2$ supersymmetry, see e.g. ref. [11]. Instead, I will introduce just as much as I believe is necessary to understand the basic features of the $N = 2$ supersymmetric Yang-Mills theory. Maybe I should stress that all of the physics will be in four dimensional Minkowski space, so the supersymmetries all refer to the standard $D = 4$ case. A standard Dirac spinor then has four complex components and transforms reducibly under the action of the (covering group $Sl(2, \mathbf{C})$ of the) Lorentz group. It is more convenient to break such a Dirac spinor into pieces each having 2 complex components and transforming irreducibly. These two-component spinors are denoted χ^α and $\bar{\chi}_{\dot{\alpha}} = (\chi^\alpha)^*$ according to their Lorentz transformation properties. Dealing with two-component spinors, one also encounters the matrices σ^μ and $\bar{\sigma}^\mu$: $(\sigma^\mu)_{\alpha\dot{\alpha}}$ being the unit matrix for $\mu = 0$ and the Pauli matrices σ^i for $\mu = i = 1,2,3$, while for $\bar{\sigma}^\mu$ one has $-\sigma^i$ instead. For completeness, I mention that one also needs the antisymmetric tensor $\epsilon^{\alpha\beta}$ with $\epsilon^{01} = +1$ and its inverse to raise and lower spinor indices. The convention for contracting indices is $\psi\chi \equiv \psi^\alpha\chi_\alpha$ and $\bar{\psi}\bar{\chi} = \bar{\psi}_{\dot{\alpha}}\bar{\chi}^{\dot{\alpha}}$.

2.1. Unextended supersymmetry

The simplest, unextended supersymmetry (called $N = 1$ susy in contrast with the extended $N > 1$ susys) can be represented on a variety of multiplets of fields involving bosons and fermions. One of the simplest representations involves a complex scalar field ϕ and a two-component spinor ψ_α ($\alpha = 1, 2$). They form the so-called chiral scalar multiplet. I do not write the susy transformations since we do not need them here (see [11]). To write down susy invariant Lagrangians (actions) it is convenient to assemble ϕ and ψ into a superfield. Therefore one introduces anticommuting varaiables θ^α and $\bar\theta_{\dot\alpha}$ and writes

$$\Phi = \phi(y) + \sqrt{2}\theta\psi(y) + \theta^2 F(y) \tag{2.1}$$

where $y^\mu = x^\mu + i\theta\sigma^\mu\bar\theta$ and $\theta\psi \equiv \theta^\alpha\psi_\alpha$, $\theta^2 \equiv \theta\theta \equiv \theta^\alpha\theta_\alpha = -2\theta^1\theta^2$, $\theta\sigma^\mu\bar\theta \equiv \theta^\alpha\sigma^\mu_{\alpha\dot\alpha}\bar\theta^{\dot\alpha}$. (Notice that θ^2 is used to denote $\theta\theta$ as well as the second component of θ. It should be clear which one is meant, and almost always it is $\theta\theta$.) Φ is a chiral superfield. One also needed to include a field F that will turn out to be an auxiliary field. Expanding the y-dependence (and using $\theta^1\theta^1 = \theta^2\theta^2 = 0$) one finds

$$\begin{aligned}\Phi =&\phi(x) + i\theta\sigma^\mu\bar\theta\partial_\mu\phi(x) - \frac{1}{4}\theta^2\bar\theta^2\partial^2\phi(x) + \sqrt{2}\theta\psi(x)\\ &- \frac{i}{\sqrt{2}}\theta^2(\partial_\mu\psi(x)\sigma^\mu\bar\theta) + \theta^2 F(x)\ .\end{aligned} \tag{2.2}$$

A supersymmetry invariant action then is given by the superspace integral

$$\frac{1}{4}\int \mathrm{d}^4x\, \mathrm{d}^2\theta\, \mathrm{d}^2\bar\theta\, \Phi^+\Phi\ . \tag{2.3}$$

The θ-integrations are defined such that only the term proportional to $\theta^2\bar\theta^2$ in $\Phi^+\Phi$ gives a non-vanishing result. (One has $\int \mathrm{d}^2\theta\, \mathrm{d}^2\bar\theta\, \theta^2\bar\theta^2 = 4$.) Then (2.3) becomes

$$\int \mathrm{d}^4x\ \left(\partial_\mu\phi\partial^\mu\phi^+ - i\bar\psi\bar\sigma^\mu\partial_\mu\psi + F^+F\right)\ . \tag{2.4}$$

We see that the simple $\Phi^+\Phi$-term has produced the standard kinetic terms for a complex scalar ϕ and the spinor ψ. F is an auxiliary field which can be set equal to zero by its equation of motion. Supersymmetry invariant interactions can be generated by a superpotential $\int \mathrm{d}^4x\ \left[\int \mathrm{d}^2\theta\, \mathcal{W}(\Phi) + \text{h.c.}\right]$ where $\mathcal{W}(\Phi)$ depends only on Φ and not on Φ^+.

Another supersymmetry multiplet is the vector multiplet that contains a (massles gauge) vector field A_μ and its superpartner λ_α (gaugino). They are combined together with an auxiliary field D into a superfield V as*

$$V = -\theta\sigma^\mu\bar\theta A_\mu + i\theta^2(\bar\theta\bar\lambda) - i\bar\theta^2(\theta\lambda) + \tfrac{1}{2}\theta^2\bar\theta^2 D \ . \tag{2.5}$$

We will be interested in the case of non-abelian gauge symmetry where A_μ, and hence $\lambda, \bar\lambda$ and D are in the adjoint representation: $A_\mu = A_\mu^a T_a$, $[T_a, T_b] = f_{abc}T_c$, etc. From the superfield V one defines another (spinorial) superfield W_α as

$$W = \left(-i\lambda + \theta D - i\sigma^{\mu\nu}\theta F_{\mu\nu} + \theta^2\sigma^\mu\nabla_\mu\bar\lambda\right)(y) \tag{2.6}$$

(again, $y^\mu = x^\mu + i\theta\sigma^\mu\bar\theta$), where $\sigma^{\mu\nu} = \frac{1}{4}(\sigma^\mu\bar\sigma^\nu - \sigma^\nu\bar\sigma^\mu)$, $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu - ig[A_\mu, A_\nu]$, $\nabla_\mu\lambda = \partial_\mu\lambda - ig[A_\mu, \lambda]$ and g is the gauge coupling constant. The corresponding superspace formula is

$$W_\alpha = \frac{1}{8g}\bar D^2\left(e^{2gV}D_\alpha e^{-2gV}\right) \ . \tag{2.7}$$

Here D_α and $\bar D_{\dot\alpha}$ are the superspace derivatives $\partial/\partial\theta^\alpha + i\sigma^\mu_{\alpha\dot\alpha}\bar\theta^{\dot\alpha}\partial_\mu$ and $-\partial/\partial\bar\theta^{\dot\alpha} - i\sigma^\mu_{\alpha\dot\alpha}\theta^\alpha\partial_\mu$. The supersymmetric Yang-Mills action then simply is (one has $\int \mathrm{d}^2\theta\,\theta^2 = -2$)

$$-\frac{1}{4}\int \mathrm{d}^4x\,\mathrm{d}^2\theta\,\mathrm{tr}\ W^\alpha W_\alpha = \int \mathrm{d}^4x\,\mathrm{tr}\ \left[-\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{i}{4}F_{\mu\nu}\tilde F^{\mu\nu} - i\lambda\sigma^\mu\nabla_\mu\bar\lambda + \frac{1}{2}D^2\right] \ . \tag{2.8}$$

In addition to the standard Yang-Mills term $-\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$ one has also generated a term $\frac{i}{4}F_{\mu\nu}\tilde F^{\mu\nu}$ which, after integration, gives the instanton number. It should appear in the action multiplying the θ-parameter (not to be confused with the anticommuting θ-variables of superspace!) and with a real coefficient. Hence if one introduces the complex coupling constant

$$\tau = \frac{\theta}{2\pi} + \frac{4\pi i}{g^2} \tag{2.9}$$

* Actually the form given here is the one obtained after fixing the Wess-Zumino gauge in the general real superfield V using $V \to V + \Lambda + \Lambda^+$ where Λ is a chiral superfield.

then the following *real* action precisely does what one wants:

$$
\begin{aligned}
\frac{1}{16\pi}\mathrm{Im}\left[\tau\int \mathrm{d}^4x\,\mathrm{d}^2\theta\,\mathrm{tr}\ W^\alpha W_\alpha\right] =&\frac{1}{g^2}\int \mathrm{d}^4x\,\mathrm{tr}\ \left[-\frac{1}{4}F_{\mu\nu}F^{\mu\nu}-i\lambda\sigma^\mu\nabla_\mu\bar\lambda+\frac{1}{2}D^2\right]\\
&+\frac{\theta}{32\pi^2}\int \mathrm{d}^4x\,F_{\mu\nu}\tilde F^{\mu\nu}
\end{aligned}
\tag{2.10}
$$

with the F^2-term and the instanton number conventionally normalized.

The matter field Φ can be minimally coupled to the Yang-Mills field by putting it in some representation of the gauge group, say the adjoint, and replacing (2.3) and (2.4) by

$$
\begin{aligned}
&\frac{1}{4}\int \mathrm{d}^4x\,\mathrm{d}^2\theta\,\mathrm{d}^2\bar\theta\,\mathrm{tr}\ \Phi^+e^{-2gV}\Phi\\
&=\int \mathrm{d}^4x\,\mathrm{tr}\ \left(|\nabla_\mu\phi|^2-i\bar\psi\bar\sigma^\mu\nabla_\mu\psi+F^+F-g\phi^+[D,\phi]-\sqrt{2}ig\phi^+\{\lambda,\psi\}+\sqrt{2}ig\bar\psi[\bar\lambda,\phi]\right)\ .
\end{aligned}
\tag{2.11}
$$

In addition to the appearance of the covariant derivatives ∇_μ we also see explicit couplings between ϕ, ψ and λ, D as required by supersymmetry.

2.2. The $N=2$ super Yang-Mills action

$N=2$ supersymmetry combines all of the fields ϕ,ψ and A_μ,λ into a single susy multiplet. Of course, this means that all fields must be in the same representation of the gauge group as A_μ, i.e. in the adjoint representation. This multiplet contains two spinor fields ψ and λ on equal footing. So the simplest guess for the $N=2$ super Yang-Mills action is a combination of (2.10) and (2.11) with relative coefficients such that the two kinetic terms for ψ and λ have the same coefficients. Integrating by parts one of them, we see that we have to add (2.10) and $\frac{1}{g^2}$ times (2.11). It is by no means obvious that the resulting sum has $N=2$ supersymmetry, but one can check that it does. Thus the $N=2$ super Yang-Mills action is

$$
\begin{aligned}
S&=\int \mathrm{d}^4x\left[\mathrm{Im}\left(\frac{\tau}{16\pi}\mathrm{d}^2\theta\,\mathrm{tr}\ W^\alpha W_\alpha\right)+\frac{1}{4g^2}\int \mathrm{d}^2\theta\,\mathrm{d}^2\bar\theta\,\mathrm{tr}\ \Phi^+e^{-2gV}\Phi\right]\\
&=\mathrm{Im}\ \mathrm{tr}\int \mathrm{d}^4x\,\frac{\tau}{16\pi}\left[\int \mathrm{d}^2\theta\,W^\alpha W_\alpha+\int \mathrm{d}^2\theta\,\mathrm{d}^2\bar\theta\,\Phi^+e^{-2gV}\Phi\right]\ .
\end{aligned}
\tag{2.12}
$$

Note that a non-trivial superpotential $\mathcal{W}(\Phi)$ is not allowed by $N=2$ supersymmetry.

An important point concerns the auxiliary fields in S:

$$S_{\rm aux} = \frac{1}{g^2}\int \mathrm{d}^4x \,\mathrm{tr}\, \left[\frac{1}{2}D^2 - g\phi^+[D,\phi] + F^+F\right] \ . \tag{2.13}$$

Solving the auxiliary field equations and inserting the result back into the action gives

$$S_{\rm aux} = -\int \mathrm{d}^4x \, \frac{1}{2}\mathrm{tr}\, \left([\phi^+,\phi]\right)^2 \tag{2.14}$$

which shows that the bosonic potential is $V(\phi) = \frac{1}{2}\mathrm{tr}\,\left([\phi^+,\phi]\right)^2 \geq 0$. As is well known, a ground state field configuration ϕ_0 with $V(\phi_0) > 0$ does break supersymmetry. In other words, unbroken susy requires a ground state (vacuum) with $V(\phi_0) = 0$. Note that this does not imply $\phi_0 = 0$. A sufficient and necessary condition is that ϕ_0 and ϕ_0^+ commute.

The $N = 2$ supersymmetry of (2.12) can be rendered manifest by using a $N = 2$ superspace notation. I will not go into any details and simply quote some relevant formulas. In addition to the anticommuting $\theta_\alpha, \bar\theta_{\dot\alpha}$ of $N = 1$ susy, one now needs a second set of anticommuting $\tilde\theta_\alpha, \bar{\tilde\theta}_{\dot\alpha}$. One introduces the $N = 2$ chiral superfield

$$\Psi = \Phi(\tilde y,\theta) + \sqrt{2}\tilde\theta^\alpha W_\alpha(\tilde y,\theta) + \tilde\theta^\alpha\tilde\theta_\alpha G(\tilde y,\theta) \tag{2.15}$$

where $\tilde y^\mu = x^\mu + i\theta\sigma^\mu\bar\theta + i\tilde\theta\sigma^\mu\bar{\tilde\theta} = y^\mu + i\tilde\theta\sigma^\mu\bar{\tilde\theta}$ and

$$G(\tilde y,\theta) = -\frac{1}{2}\int \mathrm{d}^2\bar\theta\, \left[\Phi(\tilde y - i\theta\sigma\bar\theta,\theta,\bar\theta)\right]^+ \exp\left[-2gV(\tilde y - i\theta\sigma\bar\theta,\theta,\bar\theta)\right] \tag{2.16}$$

with $\Phi(y,\theta)$ and $\Phi(x,\theta,\bar\theta)$ as given in (2.1) and (2.2) and $W(y,\theta)$ as given in (2.6). The $\mathrm{d}^2\bar\theta$-integration is meant to be at fixed $\tilde y$. Ψ is the $N = 2$ analogue of a chiral superfield, subject to the constraint (2.16) necessary in order to eliminate certain unphysical degrees or freedeom. The $N = 2$ superspace notation "implies" that the following action is $N = 2$ susy invariant:

$$\mathrm{Im}\left[\frac{\tau}{16\pi}\int \mathrm{d}^4x\,\mathrm{d}^2\theta\,\mathrm{d}^2\tilde\theta\,\frac{1}{2}\mathrm{tr}\ \Psi^2\right] \ . \tag{2.17}$$

Carrying out the $\mathrm{d}^2\tilde\theta$-integration yields precisely the action (2.12).

Note that the integrand in (2.17) only depends on Ψ, not on Ψ^+. More generally one can show that $N=2$ supersymmetry constrains the form of the action to be

$$\frac{1}{16\pi}\mathrm{Im}\int \mathrm{d}^4x\,\mathrm{d}^2\theta\,\mathrm{d}^2\tilde{\theta}\,\mathcal{F}(\Psi) \tag{2.18}$$

where $\mathcal{F}$, called the $N=2$ prepotential, depends only on Ψ and not on Ψ^+. This is referred to as holomorphy of the prepotential. For the $N=2$ super Yang-Mills action (2.12) or (2.17) one simply has

$$\mathcal{F}(\Psi)\equiv\mathcal{F}_{\text{class}}(\Psi)=\frac{1}{2}\,\mathrm{tr}\ \tau\Psi^2\ . \tag{2.19}$$

The quadratic dependence on Ψ is fixed by renormalisability. Below we will consider low-energy effective actions. Then the only constraint is $N=2$ susy, translated as holomorphicity of $\mathcal{F}$. In $N=1$ superspace language, the general action (2.18) reads [12]

$$\begin{aligned}\frac{1}{16\pi}\mathrm{Im}\int \mathrm{d}^4x\Big[\int \mathrm{d}^2\theta\,\mathcal{F}_{ab}(\Phi)W^{a\alpha}W^b{}_\alpha\\ +\int \mathrm{d}^2\theta\,\mathrm{d}^2\bar{\theta}\,\left(\Phi^+e^{-2gV}\right)^a\mathcal{F}_a(\Phi)\Big]\end{aligned} \tag{2.20}$$

where $\mathcal{F}_a(\Phi)=\frac{\partial\mathcal{F}(\Phi)}{\partial\Phi^a}$, $\mathcal{F}_{ab}(\Phi)=\frac{\partial^2\mathcal{F}(\Phi)}{\partial\Phi^a\partial\Phi^b}$ and where a,b are Lie algebra indices, so that $\Phi=\Phi^aT_a$, $W_\alpha=W_\alpha^aT_a$ with $\mathrm{tr}\ T_aT_b=\delta_{ab}$. This concludes our quick tour through supersymmetric Yang-Mills theories.

3. Low-energy effective action of $N=2$ susy $SU(2)$ Yang-Mills theory

Following Seiberg and Witten [1] we want to study and determine the low-energy effective action of the $N=2$ susy Yang-Mills theory with gauge group $SU(2)$. The latter theory is the microscopic theory which controls the high-energy behaviour. It is renormalisable and well-known to be asymptotically free. The low-energy effective action will turn out to be quite different.

3.1. Low-energy effective actions

There are two types of effective actions. One is the standard generating functional $\Gamma[\varphi]$ of one-particle irreducible Feynman diagrams (vertex functions). It is obtained from the standard renormalised generating functional $W[\varphi]$ of connected diagrams by a Legendre transformation. Momentum integrations in loop-diagrams are from zero up to a UV-cutoff which is taken to infinity at the end. $\Gamma[\varphi] \equiv \Gamma[\mu, \varphi]$ also depends on the scale μ used to define the renormalized vertex functions.

A quite different object is the Wilsonian effective action $S_{\rm W}[\mu, \varphi]$. It is defined as $\Gamma[\mu, \varphi]$, except that all loop-momenta are only integrated down to μ which serves as an infra-red cutoff. In theories with massive particles only, there is no big difference between $S_{\rm W}[\mu, \varphi]$ and $\Gamma[\mu, \varphi]$ (as long as μ is less than the smallest mass). When massless particles are present, as is the case for gauge theories, the situation is different. In particular, in supersymmetric gauge theories there is the so-called Konishi anomaly which can be viewed as an IR-effect. Although $S_{\rm W}[\mu, \varphi]$ depends holomorphically on μ, this is not the case for $\Gamma[\mu, \varphi]$ due to this anomaly.

3.2. The $SU(2)$ case, moduli space

What Seiberg and Witten achieved, and what will occupy the rest of these notes, is to determine the Wilsonian effective action in the case where the microscopic theory one starts with is the $SU(2)$, $N = 2$ super Yang-Mills theory (2.12) or (2.17). As explained above (see (2.14)), classically this theory has a scalar potential $V(\phi) = \frac{1}{2}{\rm tr}\ ([\phi^+, \phi])^2$. Unbroken susy requires that $V(\phi) = 0$ in the vacuum, but this still leaves the possibilities of non-vanishing ϕ with $[\phi^+, \phi] = 0$. We are interested in determining the gauge inequivalent vacua. A general ϕ is of the form $\phi(x) = \frac{1}{2}\sum_{j=1}^{3}(a_j(x) + ib_j(x))\,\sigma_j$ with real fields $a_j(x)$ and $b_j(x)$ (where I assume that not all three a_j vanish, otherwise exchange the roles of the a_j's and b_j's in the sequel). By a $SU(2)$ gauge transformation one can always arrange $a_1(x) = a_2(x) = 0$. Then $[\phi, \phi^+] = 0$ implies $b_1(x) = b_2(x) = 0$ and hence, with $a = a_3 + ib_3$, one has $\phi = \frac{1}{2}a\sigma_3$. Obviously, in the vacuum a must be a constant. Gauge transformation from the Weyl group (i.e. rotations by π around the 1- or 2-axis of $SU(2)$) can still change $a \to -a$, so a and $-a$ are gauge equivalent, too. The gauge invariant quantity describing inequivalent vacua is $\frac{1}{2}a^2$, or tr ϕ^2, which is the same, semiclassically. When quantum fluctuations are important this is no longer so. In the sequel, we will use the following definitions for a and u:

$$u = \langle \mathrm{tr}\ \phi^2 \rangle \quad , \quad \langle \phi \rangle = \tfrac{1}{2} a \sigma_3 \ . \tag{3.1}$$

The complex parameter u labels gauge inequivalent vacua. The manifold of gauge inequivalent vacua is called the moduli space $\mathcal{M}$ of the theory. Hence u is a coordinate on $\mathcal{M}$, and $\mathcal{M}$ is essentially the complex u-plane. We will see in the sequel that $\mathcal{M}$ has certain singularities, and the knowledge of the behaviour of the theory near the singularities will eventually allow the determination of the effective action S_{W}.

Clearly, for non-vanishing $\langle \phi \rangle$, the $SU(2)$ gauge symmetry is broken by the Higgs mechanism, since the ϕ-kinetic term $|\nabla_\mu \phi|^2$ generates masses for the gauge fields. With the above conventions, A^b_μ, $b = 1, 2$ become massive with masses given by $\frac{1}{2} m^2 = \frac{1}{g^2} |ga|^2$, i.e $m = \sqrt{2} a$. Similarly due to the ϕ, λ, ψ interaction terms, ψ^b, λ^b, $b = 1, 2$ become massive with the same mass as the A^b_μ, as required by supersymmetry. Obviously, A^3_μ, ψ^3 and λ^3, as well as the mode of ϕ describing the flucuation of ϕ in the σ_3-direction, remain massless. These massless modes are described by a Wilsonian low-energy effective action which has to be $N = 2$ supersymmetry invariant, since, although the gauge symmetry is broken, $SU(2) \to U(1)$, the $N = 2$ susy remains unbroken. Thus it must be of the general form (2.18) or (2.20) where the indices a, b now take only a single value ($a, b = 3$) and will be suppressed since the gauge group is $U(1)$. Also, V in (2.20) is in the adjoint representation and it is easy to see that from $e^{-2gV} = 1 - 2gV + \ldots$ only the 1 can contribute. In other words, in an abelian theory there is no self-coupling of the gauge boson and the same arguments extend to all members of the $N = 2$ susy multiplet: they do not carry electric charge. Thus for a $U(1)$-gauge theory, from (2.20) we get simply

$$\frac{1}{16\pi} \mathrm{Im} \int \mathrm{d}^4 x \left[\int \mathrm{d}^2\theta\, \mathcal{F}''(\Phi) W^\alpha W_\alpha + \int \mathrm{d}^2\theta\, \mathrm{d}^2\bar{\theta}\, \Phi^+ \mathcal{F}'(\Phi) \right] \ . \tag{3.2}$$

3.3. Metric on moduli space

Consider the second term of the effective action (3.2). In component fields this term reads

$$\frac{1}{4\pi}\mathrm{Im}\int \mathrm{d}^4x\ \left[\mathcal{F}''(\phi)|\partial_\mu\phi|^2 - i\mathcal{F}''(\phi)\psi\sigma^\mu\partial_\mu\bar{\psi} + \ldots\right] \tag{3.3}$$

where $+\ldots$ stands for non-derivative terms. Similarly, the first term in (3.2) gives

$$\frac{1}{4\pi}\mathrm{Im}\int \mathrm{d}^4x\ \left[\mathcal{F}''(\phi)\left(-\frac{1}{4}\right)F_{\mu\nu}(F^{\mu\nu} - i\tilde{F}^{\mu\nu}) - i\mathcal{F}''(\phi)\lambda\sigma^\mu\partial_\mu\bar{\lambda} + \ldots\right]\ . \tag{3.4}$$

If we think of these kinetic terms as a four dimensional sigma-model, then the $\mathcal{F}''(\phi)$ or rather $\mathrm{Im}\,\mathcal{F}''(\phi)$ that appears for all of them plays the role of a metric in field space. By the same token it defines the metric in the space of (inequivalent) vacuum configurations, i.e. the metric on moduli space. From the ϕ-kinetic term one sees that a sensible definition of the metric on the moduli space is ($\bar{a}$ denotes the complex conjugate of a)

$$\mathrm{d}s^2 = \mathrm{Im}\,\mathcal{F}''(a)\mathrm{d}a\mathrm{d}\bar{a} = \mathrm{Im}\,\tau(a)\mathrm{d}a\mathrm{d}\bar{a} \tag{3.5}$$

where $\tau(a) = \mathcal{F}''(a)$ is the effective (complexified) coupling constant in analogy with (2.19). The σ-model metric $G_{\phi\phi^+} \sim \mathrm{Im}\,\mathcal{F}''(\phi)$ has been replaced on the moduli space $\mathcal{M}$ by its expectation value in the vacuum corresponding to the given point on $\mathcal{M}$, i.e. by $\mathrm{Im}\,\mathcal{F}''(a) = \mathrm{Im}\,\tau(a)$.

The question now is whether the description of the effective action in terms of the fields Φ, W and the function $\mathcal{F}$ is appropriate for all vacua, i.e. for all value of u, i.e. on all of moduli space. In particular the kinetic terms (3.3), (3.4), or what is the same, the metric (3.5) on moduli space should be positive definite, translating into $\mathrm{Im}\,\tau(a) > 0$. However, a simple argument shows that this cannot be the case: since $\mathcal{F}(a)$ is holomorphic, $\mathrm{Im}\,\tau(a) = \mathrm{Im}\,\frac{\partial^2\mathcal{F}(a)}{\partial a^2}$ is a harmonic function and as such it cannot have a minimum, and hence (on the compactified complex plane) it cannot obey $\mathrm{Im}\,\tau(a) > 0$ everywhere (unless it is a constant as in the classical case). The way out is to allow for different local descriptions: the coordinates $a, \bar{a}$ and the function $\mathcal{F}(a)$ are appropriate only in a certain region of $\mathcal{M}$. When a singular point with $\mathrm{Im}\,\tau(a) \to 0$ is approached one has to use a different set of coordinates $\hat{a}$ in which $\mathrm{Im}\,\hat{\tau}(\hat{a})$ is non-singular (and non-vanishing). This is possible provided the singularity of the metric (3.5) is only a coordinate singularity, i.e. the kinetic terms of the effective action are not intrinsically singular, which will be the case.

3.4. Asymptotic freedom and the one-loop formula

Classically the function $\mathcal{F}$ is given by $\frac{1}{2}\tau_{\text{class}}\Psi^2$. The one-loop contribution has been determined in [13]. The combined tree-level and one-loop result is

$$\mathcal{F}_{\text{pert}}(\Psi) = \frac{i}{2\pi}\Psi^2 \ln \frac{\Psi^2}{\Lambda^2} \ . \tag{3.6}$$

Here Λ^2 is some combination of μ^2 and numerical factors chosen so as to fix the normalisation of $\mathcal{F}_{\text{pert}}$. Note that due to non-renormalisation theorems for $N = 2$ susy there are no corrections from two or more loops to the Wilsonian effective action S_{W} and (3.6) is the full perturbative result. There are however non-perturbative corrections that will be determined below.

For very large a the dominant contribution when computing S_{W} from the microscopic $SU(2)$ gauge theory comes from regions of large momenta ($p \sim a$) where the microscopic theory is asymptotically free. Thus, as $a \to \infty$ the effective coupling constant goes to zero, and the perturbative expression (3.6) for $\mathcal{F}$ becomes an excellent approximation. Also $u \sim \frac{1}{2}a^2$ in this limit.* Thus

$$\begin{aligned} \mathcal{F}(a) &\sim \frac{i}{2\pi}a^2 \ln \frac{a^2}{\Lambda^2} \\ \tau(a) &\sim \frac{i}{\pi}\left(\ln \frac{a^2}{\Lambda^2} + 3\right) \end{aligned} \qquad \text{as } u \to \infty \ . \tag{3.7}$$

Note that due to the logarithm appearing at one-loop, $\tau(a)$ is a multi-valued function of $a^2 \sim 2u$. Its imaginary part, however, $\operatorname{Im}\tau(a) \sim \frac{1}{\pi}\ln\frac{|a|^2}{\Lambda^2}$ is single-valued and positive (for $a^2 \to \infty$).

---

* One can check from the explicit solution in section 6 that one indeed has $\frac{1}{2}a^2 - u = \mathcal{O}(1/u)$ as $u \to \infty$.

4. Duality

As already noted, a and $\bar{a}$ do provide local coordinates on the moduli space $\mathcal{M}$ for the region of large u. This means that in this region Φ (and Φ^+) and W^α are appropriate fields to describe the low-energy effective action. As also noted, this description cannot be valid globally, since $\mathrm{Im}\,\mathcal{F}''(a)$, being a harmonic function, must vanish somewhere, unless it is a constant - which it is not. Duality will provide a different set of (dual) fields Φ_D and W_{D}^α that provide an appropriate description for a different region of the moduli space.

4.1. Duality transformation

Consider the form (3.2) of the effective action. Define a field dual to Φ by

$$\Phi_D = \mathcal{F}'(\Phi) \tag{4.1}$$

and a function $\mathcal{F}_D(\Phi_D)$ dual to $\mathcal{F}(\Phi)$ by

$$\mathcal{F}_D'(\Phi_D) = -\Phi \tag{4.2}$$

where, of course, $\mathcal{F}_D'(\Phi_D)$ means $\mathrm{d}\mathcal{F}_D(\Phi_D)/\mathrm{d}\Phi_D$. These duality transformations simply constitute a Legendre transformation† $\mathcal{F}_D(\Phi_D) = \mathcal{F}(\Phi) - \Phi\Phi_D$ with Φ_D defined as in (4.1). Equation (4.2) then is the standard inverse relation that follows from the Legendre transform. Using these relations, the second term in the action (3.2) can be written as

$$\begin{aligned}\mathrm{Im}\int \mathrm{d}^4x\,\mathrm{d}^2\theta\,\mathrm{d}^2\bar{\theta}\,\Phi^+\mathcal{F}'(\Phi) &= \mathrm{Im}\int \mathrm{d}^4x\,\mathrm{d}^2\theta\,\mathrm{d}^2\bar{\theta}\,\left(-\mathcal{F}_D'(\Phi_D)\right)^+\Phi_D \\ &= \mathrm{Im}\int \mathrm{d}^4x\,\mathrm{d}^2\theta\,\mathrm{d}^2\bar{\theta}\,\Phi_D^+\mathcal{F}_D'(\Phi_D)\,.\end{aligned} \tag{4.3}$$

We see that this second term in the effective action (3.2) is invariant under the duality transformation (4.1), (4.2).

† This was pointed out to me by Frank Ferrari.

The reader will recognise the similarity of (4.1), (4.2) with a canonical transformation. Indeed $\mathcal{F}'(\Phi) = \partial\mathcal{F}/\partial\Phi$ ressembles a (complex) momentum (remember that the effective action is $\sim \int d^4x\, d^2\theta\, d^2\tilde{\theta}\, \mathcal{F}(\Phi)$, eq. (2.18)), so that the second term in (3.2) is like $\mathrm{Im} \int q^* p = \frac{i}{2} \int (p^* q - q^* p)$ and the duality transformation is $q_D = p$ and $p_D = -q$ which clearly is a canonical transformation. It is well-known that canonical transformations preserve the phase-space measure. As a consequence, if the functional integral is formulated as a phase-space integral ($\sim \int \mathcal{D}\Phi\mathcal{D}\Pi \exp[\int \dot{\Phi}\Pi - H]$), under appropriate conditions, the Jacobian for the integration measure is unity for canonical transformations. The present duality transformation is a particularly simple canonical transformation and we expect the Jacobian to be one.

Next, consider the $\mathcal{F}''(\Phi)W^\alpha W_\alpha$-term in the effective action (3.2). While the duality transformation (4.1), (4.2) on Φ is local, this will not be the case for the transformation of W^α. Recall that W contains the $U(1)$ field strength $F_{\mu\nu}$, cf. eq. (2.6). This $F_{\mu\nu}$ is not arbitrary but of the form $\partial_\mu A_\nu - \partial_\nu A_\mu$ for some A_μ. This can be translated into the Bianchi identity $\frac{1}{2}\epsilon^{\mu\nu\rho\sigma}\partial_\nu F_{\rho\sigma} \equiv \partial_\nu \tilde{F}^{\mu\nu} = 0$. The corresponding constraint in superspace is $\mathrm{Im}\,(D_\alpha W^\alpha) = 0$ where D_α is the same superspace derivative as in (2.7). This constraint is a consequence of the abelian version of the expression (2.7) of W in terms of V. In the functional integral one has the choice of integrating over V only, or over W^α and imposing the constraint $\mathrm{Im}\,(D_\alpha W^\alpha) = 0$ by a real Lagrange multiplier superfield which we call V_D:

$$
\begin{aligned}
&\int \mathcal{D}V \exp\left[\frac{i}{16\pi}\mathrm{Im} \int d^4x\, d^2\theta\, \mathcal{F}''(\Phi) W^\alpha W_\alpha\right] \\
&\simeq \int \mathcal{D}W\mathcal{D}V_D \exp\left[\frac{i}{16\pi}\mathrm{Im} \int d^4x \left(\int d^2\theta\, \mathcal{F}''(\Phi) W^\alpha W_\alpha + \frac{1}{2}\int d^2\theta\, d^2\bar{\theta}\, V_D D_\alpha W^\alpha\right)\right] .
\end{aligned}
\tag{4.4}
$$

Observe that

$$
\begin{aligned}
\int d^2\theta\, d^2\bar{\theta}\, V_D D_\alpha W^\alpha &= -\int d^2\theta\, d^2\bar{\theta}\, D_\alpha V_D W^\alpha = +\int d^2\theta\, \bar{D}^2 (D_\alpha V_D W^\alpha) \\
&= \int d^2\theta\, (\bar{D}^2 D_\alpha V_D) W^\alpha = -4 \int d^2\theta\, (W_D)_\alpha W^\alpha
\end{aligned}
\tag{4.5}
$$

where we used $\bar{D}_{\dot{\beta}} W^\alpha = 0$ and where the dual W_D is defined from V_D in analogy with the abelian version of (2.7) as $(W_D)_\alpha = -\frac{1}{4}\bar{D}^2 D_\alpha V_D$. Then one can do the functional integral over W and one obtains

$$\int \mathcal{D}V_D \exp\left[\frac{i}{16\pi}\mathrm{Im}\int \mathrm{d}^4x\,\mathrm{d}^2\theta\left(-\frac{1}{\mathcal{F}''(\Phi)}W_D^\alpha W_{D\alpha}\right)\right] . \tag{4.6}$$

This reexpresses the ($N = 1$) supersymmetrized Yang-Mills action in terms of a dual Yang-Mills action with the effective coupling $\tau(a) = \mathcal{F}''(a)$ replaced by $-\frac{1}{\tau(a)}$. Recall that $\tau(a) = \frac{\theta(a)}{2\pi} + \frac{4\pi i}{g^2(a)}$, so that $\tau \to -\frac{1}{\tau}$ generalizes the inversion of the coupling constant discussed in the introduction. Also, it can be shown that W_D actually describes the electromagnetic dual $F_{\mu\nu} \to \tilde{F}_{\mu\nu}$, so that the manipulations leading to (4.6) constitute a duality transformation that generalizes the old electromagnetic duality of Montonen and Olive (cf. the introduction). Expressing the $-\frac{1}{\mathcal{F}''(\Phi)}$ in terms of Φ_D one sees from (4.2) that $\mathcal{F}_D''(\Phi_D) = -\frac{\mathrm{d}\Phi}{\mathrm{d}\Phi_D} = -\frac{1}{\mathcal{F}''(\Phi)}$ so that

$$-\frac{1}{\tau(a)} = \tau_D(a_D) . \tag{4.7}$$

The whole action (3.2) can then equivalently be written as

$$\frac{1}{16\pi}\mathrm{Im}\int \mathrm{d}^4x\left[\int \mathrm{d}^2\theta\,\mathcal{F}_D''(\Phi_D)W_D^\alpha W_{D\alpha} + \int \mathrm{d}^2\theta\,\mathrm{d}^2\bar{\theta}\,\Phi_D^+\mathcal{F}_D'(\Phi_D)\right] . \tag{4.8}$$

4.2. The duality group

To discuss the full group of duality transformations of the action it is most convenient to write it as

$$\frac{1}{16\pi}\mathrm{Im}\int \mathrm{d}^4x\,\mathrm{d}^2\theta\,\frac{\mathrm{d}\Phi_D}{\mathrm{d}\Phi}W^\alpha W_\alpha + \frac{1}{32i\pi}\int \mathrm{d}^4x\,\mathrm{d}^2\theta\,\mathrm{d}^2\bar{\theta}\,\left(\Phi^+\Phi_D - \Phi_D^+\Phi\right) . \tag{4.9}$$

While we have shown in the previous subsection that there is a duality symmetry

$$\begin{pmatrix}\Phi_D\\ \Phi\end{pmatrix} \to \begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}\begin{pmatrix}\Phi_D\\ \Phi\end{pmatrix} , \tag{4.10}$$

the form (4.9) shows that there also is a symmetry

$$\begin{pmatrix}\Phi_D\\ \Phi\end{pmatrix} \to \begin{pmatrix}1 & b\\ 0 & 1\end{pmatrix}\begin{pmatrix}\Phi_D\\ \Phi\end{pmatrix} \quad , \quad b \in \mathbf{Z} . \tag{4.11}$$

Indeed, the second term in (4.9) remains invariant since b is real, while the first term in (4.9) gets shifted by

$$\frac{b}{16\pi}\mathrm{Im}\int \mathrm{d}^4x\,\mathrm{d}^2\theta\,W^\alpha W_\alpha = -\frac{b}{16\pi}\int \mathrm{d}^4x\,F_{\mu\nu}\tilde{F}^{\mu\nu} = -2\pi b\nu \tag{4.12}$$

where $\nu \in \mathbf{Z}$ is the instanton number. Since the action appears as e^{iS} in the functional integral, two actions differing only by $2\pi\mathbf{Z}$ are equivalent, and we conclude that (4.11) with integer b is a symmetry of the effective action. The transformations (4.10) and (4.11) together generate the group $Sl(2,\mathbf{Z})$. This is the group of duality symmetries.

Note that the metric (3.5) on moduli space can be written as

$$\mathrm{d}s^2 = \mathrm{Im}\,(\mathrm{d}a_D\mathrm{d}\bar{a}) = \frac{i}{2}(\mathrm{d}a\mathrm{d}\bar{a}_D - \mathrm{d}a_D\mathrm{d}\bar{a}) \tag{4.13}$$

where $\langle\phi_D\rangle = \frac{1}{2}a_D\sigma_3$ and $a_D = \partial\mathcal{F}(a)/\partial a$, and that this metric obviously also is invariant under the duality group $Sl(2,\mathbf{Z})$

4.3. Monopoles, dyons and the BPS mass spectrum

At this point, I will have to add a couple of ingredients without much further justification and refer the reader to the literature for more details.

In a spontaneously broken gauge theory as the one we are considering, typically there are solitons (static, finite-energy solutions of the equations of motion) that carry magnetic charge and behave like non-singular magnetic monopoles [8] (for a pedagogical treatment, see [14]). The duality transformation (4.10) constructed above exchanges electric and magnetic degrees of freedom, hence electrically charged states, as would be described by hypermultiplets of our $N = 2$ supersymmetric version, with magnetic monopoles.

In $N = 2$ susy theories there are two types of multiplets: small (or short) ones (4 helicity states) and large (or long) ones (16 helicity states). Massless states must be in short multiplets, while massive states are in short ones if they satisfy $m^2 = 2|Z|^2$, Z being the central charge of the $N = 2$ susy algebra, or in long ones if $m^2 > 2|Z|^2$ [15]. The states that become massive by the Higgs mechanism must be in short multiplets since they were before the symmetry breaking (if one imagines turning on the scalar field expectation value), and the Higgs mechanism cannot generate the missing $16 - 4 = 12$ helicity states. For purely electrically charged states one has $Z = an_e$ where n_e is the (integer) electric charge. Duality then implies that a purely magnetically charged state has $Z = a_D n_m$ where n_m is the (integer) magnetic charge. A state with both types of charge, called a dyon, has $Z = an_e + a_D n_m$ since the central charge is additive. All this applies to states in short multiplets, so-called BPS-states. The mass formula for these states then is

$$m^2 = 2|Z|^2 \quad , \quad Z = (n_m, n_e) \begin{pmatrix} a_D \\ a \end{pmatrix} . \tag{4.14}$$

It is clear that under a $Sl(2, \mathbf{Z})$ transformation $M = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in Sl(2, \mathbf{Z})$ acting on $\begin{pmatrix} a_D \\ a \end{pmatrix}$, the charge vector gets transformed to $(n_m, n_e)M = (n'_m, n'_e)$ which are again integer charges. In particular, one sees again at the level of the charges that the transformation (4.10) exchanges purely electrically charged states with purely magnetically charged ones. It can be shown [16,10,1] that precisely those BPS states are stable for which n_m and n_e are relatively prime, i.e. for stable states $(n_m, n_e) \neq (qm, qn)$ for integer m, n and $q \neq \pm 1$.

5. Singularities and Monodromy

In this section we will study the behaviour of $a(u)$ and $a_D(u)$ as u varies on the moduli space $\mathcal{M}$. Particularly useful information will be obtained from their behaviour as u is taken around a closed contour. If the contour does not encircle certain singular points to be determined below, $a(u)$ and $a_D(u)$ will return to their initial values once u has completed its contour. However, if the u-contour goes around these singular points, $a(u)$ and $a_D(u)$ do not return to their initial values but rather to certain linear combinations thereof: one has a non-trivial monodromy for the multi-valued functions $a(u)$ and $a_D(u)$.

5.1. The monodromy at infinity

This is immediately clear from the behaviour near $u = \infty$. As already explained in section 3.4, as $u \to \infty$, due to asymptotic freedom, the perturbative expression for $\mathcal{F}(a)$ is valid and one has from (3.6) for $a_D = \partial\mathcal{F}(a)/\partial a$

$$a_D(u) = \frac{i}{\pi} a \left(\ln \frac{a^2}{\Lambda^2} + 1 \right) \quad , \quad u \to \infty . \tag{5.1}$$

Now take u around a counterclockwise contour of very large radius in the complex u-plane, often simply written as $u \to e^{2\pi i}u$. This is equivalent to having u encircle the point at ∞ on the Riemann sphere in a *clockwise* sense. In any case, since $u = \frac{1}{2}a^2$ (for $u \to \infty$) one has $a \to -a$ and

$$a_D \to \frac{i}{\pi}(-a)\left(\ln \frac{e^{2\pi i}a^2}{\Lambda^2} + 1\right) = -a_D + 2a \tag{5.2}$$

or

$$\begin{pmatrix} a_D(u) \\ a(u) \end{pmatrix} \to M_\infty \begin{pmatrix} a_D(u) \\ a(u) \end{pmatrix} \quad , \quad M_\infty = \begin{pmatrix} -1 & 2 \\ 0 & -1 \end{pmatrix} . \tag{5.3}$$

Clearly, $u = \infty$ is a branch point of $a_D(u) \sim \frac{i}{\pi}\sqrt{2u}\left(\ln \frac{u}{\Lambda^2} + 1\right)$. This is why this point is referred to as a singularity of the moduli space.

5.2. How many singularities?

Can $u = \infty$ be the only singular point? Since a branch cut has to start and end somewhere, there must be at least one other singular point. Following Seiberg and Witten, I will argue that one actually needs three singular points at least. To see why two cannot work, let's suppose for a moment that there are only two singularities and show that this leads to a contradiction.

Before doing so, let me note that there is an important so-called $U(1)_R$-symmetry in the classical theory that takes $\phi \to e^{2i\alpha}\phi$, $W \to e^{i\alpha}W$, $\theta \to e^{i\alpha}\theta$, $\bar\theta \to e^{i\alpha}\bar\theta$, thus $\mathrm{d}^2\theta \to e^{-2i\alpha}\mathrm{d}^2\theta$, $\mathrm{d}^2\bar\theta \to e^{-2i\alpha}\mathrm{d}^2\bar\theta$ and hence $\Psi \to e^{2i\alpha}\Psi$, so that the classical action (2.12) or (2.17) is invariant under this global symmetry. More generallly, the action (2.18) will be invariant if $\mathcal{F}(\Psi) \to e^{4i\alpha}\mathcal{F}(\Psi)$. This symmetry is broken by the one-loop correction and also by instanton contributions. The latter give corrections to $\mathcal{F}$ of the form $\Psi^2 \sum_{k=1}^\infty c_k \left(\Lambda^2/\Psi^2\right)^{2k}$, and hence are invariant only for $\left(e^{4i\alpha}\right)^{2k} = 1$, i.e. $\alpha = \frac{2\pi n}{8}$, $n \in \mathbf{Z}$. Hence instantons break the $U(1)_R$-symmetry to a dicrete $\mathbf{Z}_8$. The one-loop corrections behave as $\frac{i}{2\pi}\Psi^2 \ln \frac{\Psi^2}{\Lambda^2} \to e^{4i\alpha}\left(\frac{i}{2\pi}\Psi^2 \ln \frac{\Psi^2}{\Lambda^2} - \frac{2\alpha}{\pi}\Psi^2\right)$. As in the paragraph before eq. (4.12) one shows that this only changes the action by $2\pi\nu\left(\frac{4\alpha}{\pi}\right)$ where ν is integer, so that again this change is irrelevant as long as $\frac{4\alpha}{\pi} = n$ or $\alpha = \frac{2\pi n}{8}$. Under this $\mathbf{Z}_8$-symmetry, $\phi \to e^{i\pi n/2}\phi$, i.e. for odd n one has $\phi^2 \to -\phi^2$. The non-vanishing expectation value $u = \langle \mathrm{tr}\ \phi^2 \rangle$ breaks this $\mathbf{Z}_8$ further to $\mathbf{Z}_4$. Hence for a given vacuum, i.e. a given point on moduli space there is only a $\mathbf{Z}_4$-symmetry left from the $U(1)_R$-symmetry. However, on the manifold of all possible vacua, i.e. on $\mathcal{M}$, one has still the full $\mathbf{Z}_8$-symmetry, taking u to $-u$.

Due to this global symmetry $u \to -u$, singularities of $\mathcal{M}$ should come in pairs: for each singularity at $u = u_0$ there is another one at $u = -u_0$. The only fixed points of $u \to -u$ are $u = \infty$ and $u = 0$. We have already seen that $u = \infty$ is a singular point of $\mathcal{M}$. So if there are only two singularities the other must be the fixed point $u = 0$.

If there are only two singularities, at $u = \infty$ and $u = 0$, then by contour deformation ("pulling the contour over the back of the sphere")* one sees that the monodromy around 0 (in a counterclockwise sense) is the same as the above monodromy around ∞: $M_0 = M_\infty$. But then a^2 is not affected by any monodromy and hence is a good global coordinate, so one can take $u = \frac{1}{2}a^2$ on all of $\mathcal{M}$, and furthermore one must have

$$\begin{aligned} a_D &= \frac{i}{\pi} a \left(\ln \frac{a^2}{\Lambda^2} + 1 \right) + g(a) \\ a &= \sqrt{2u} \end{aligned} \tag{5.4}$$

where $g(a)$ is some entire function of a^2. This implies that

$$\tau = \frac{\mathrm{d}a_D}{\mathrm{d}a} = \frac{i}{\pi} \left(\ln \frac{a^2}{\Lambda^2} + 3 \right) + \frac{\mathrm{d}g}{\mathrm{d}a} \,. \tag{5.5}$$

The function g being entire, $\mathrm{Im}\, \frac{\mathrm{d}g}{\mathrm{d}a}$ cannot have a minimum (unless constant) and it is clear that $\mathrm{Im}\, \tau$ cannot be positive everywhere. As already emphasized, this means that a (or rather a^2) cannot be a good global coordinate and (5.4) cannot hold globally. Hence, two singularities only cannot work.

The next simplest choice is to try 3 singularities. Due to the $u \to -u$ symmetry, these 3 singularities are at ∞, u_0 and $-u_0$ for some $u_0 \neq 0$. In particular, $u = 0$ is no longer a singularity of the quantum moduli space. To get a singularity also at $u = 0$ one would need at least four singularities at $\infty, u_0, -u_0$ and 0. As discussed later, this is not possible, and more generally, exactly 3 singularities seems to be the only consistent possibility.

So there is no singularity at $u = 0$ in the quantum moduli space $\mathcal{M}$. Classically, however, one precisely expects that $u = 0$ should be a singular point, since classically $u = \frac{1}{2}a^2$, hence

$\star$ It is well-known from complex analysis that monodromies are associated with contours around branch points. The precise from of the contour does not matter, and it can be deformed as long as it does not meet another branch point. Our singularities precisely are the branch points of $a(u)$ or $a_D(u)$.

$a = 0$ at this point, and then there is no Higgs mechanism any more. Thus all (elementary) massive states, i.e. the gauge bosons A^1_μ, A^2_μ and their susy partners $\psi^1, \psi^2, \lambda^1, \lambda^2$ become massless. Thus the description of the lights fields in terms of the previous Wilsonian effective action should break down, inducing a singularity on the moduli space. As already stressed, this is the clasical picture. While $a \to \infty$ leads to asymptotic freedom and the microscopic $SU(2)$ theory is weakly coupled, as $a \to 0$ one goes to a strong coupling regime where the classical reasoning has no validity any more, and $u \neq \frac{1}{2}a^2$. By the BPS mass formula (4.14) massless gauge bosons still are possible at $a = 0$, but this does no longer correspond to $u = 0$.

So where has the singularity due to massless gauge bosons at $a = 0$ moved to? One might be tempted to think that $a = 0$ now corresponds to the singularities at $u = \pm u_0$, but this is not the case as I will show in a moment. The answer is that the point $a = 0$ no longer belongs to the quantum moduli space (at least not to the component connected to $u = \infty$ which is the only thing one considers). This can be seen explicitly from the form of the solution for $a(u)$ given in the next section.

5.3. THE STRONG COUPLING SINGULARITIES

Let's now concentrate on the case of three singularities at $u = \infty, u_0$ and $-u_0$. What is the interpretation of the (strong-coupling) singularities at finite $u = \pm u_0$? One might first try to consider that they are still due to the gauge bosons becoming massless. However, as Seiberg and Witten point out, massless gauge bosons would imply an asymptotically conformally invariant theory in the infrared limit and conformal invariance implies $u = \langle \mathrm{tr}\ \phi^2 \rangle = 0$ unless tr ϕ^2 has dimension zero and hence would be the unity operator - which it is not. So the singularities at $u = \pm u_0$ ($\neq 0$) do not correspond to massless gauge bosons.

There are no other elementary $N = 2$ multiplets in our theory. The next thing to try is to consider collective excitations - solitons, like the magnetic monopoles or dyons. Let's first study what happens if a magnetic monopole of unit magnetic charge becomes massless. From the BPS mass formula (4.14), the mass of the magnetic monopole is

$$m^2 = 2|a_D|^2 \tag{5.6}$$

and hence vanishes at $a_D = 0$. We will see that this produces one of the two stron-coupling singularities. So call u_0 the value of u at whiche a_D vanishes. Magnetic monopoles are described by hypermultiplets M of $N = 2$ susy that couple locally to the dual fields Φ_D and W_D, just as electrically charged "electrons" would be described by hypermultiplets that couple locally to Φ and W. So in the dual description we have Φ_D, W_D and M, and, near u_0, $a_D \sim \langle\Phi_D\rangle$ is small. This theory is exactly $N = 2$ susy QED with very light electrons (and a subscript D on every quantity). The latter theory is not asymptotically free, but has a β-function given by

$$\mu \frac{\mathrm{d}}{\mathrm{d}\mu} g_D = \frac{g_D^3}{8\pi} \tag{5.7}$$

where g_D is the coupling constant. But the scale μ is proportional to a_D and $\frac{4\pi i}{g_D^2(a_D)}$ is τ_D for $\theta_D = 0$ (of course, super QED, unless embedded into a larger gauge group, does not allow for a non-vanishing theta angle). One concludes that for $u \approx u_0$ or $a_D \approx 0$

$$a_D \frac{\mathrm{d}}{\mathrm{d}a_D} \tau_D = -\frac{i}{\pi} \ \Rightarrow \ \tau_D = -\frac{i}{\pi} \ln a_D \ . \tag{5.8}$$

Since $\tau_D = \frac{\mathrm{d}(-a)}{\mathrm{d}a_D}$ this can be integrated to give

$$a \approx a_0 + \frac{i}{\pi} a_D \ln a_D \qquad (u \approx u_0) \tag{5.9}$$

where we dropped a subleading term $-\frac{i}{\pi} a_D$. Now, a_D should be a good coordinate in the vicinity of u_0, hence depend linearly* on u. One concludes

$$\begin{aligned} a_D &\approx c_0(u - u_0) \\ a &\approx a_0 + \frac{i}{\pi} c_0 (u - u_0) \ln(u - u_0) \ . \end{aligned} \tag{5.10}$$

From these expressions one immediately reads the monodromy as u turns around u_0 counterclockwise, $u - u_0 \to e^{2\pi i}(u - u_0)$:

$$\begin{pmatrix} a_D \\ a \end{pmatrix} \to \begin{pmatrix} a_D \\ a - 2a_D \end{pmatrix} = M_{u_0} \begin{pmatrix} a_D \\ a \end{pmatrix} \quad , \quad M_{u_0} = \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix} \ . \tag{5.11}$$

To obtain the monodromy matrix at $u = -u_0$ it is enough to observe that the contour around $u = \infty$ is equivalent to a counterclockwise contour of very large radius in the complex

$\star$ One might want to try a more general dependence like $a_D \approx c_0(u - u_0)^k$ with $k > 0$. This leads to a monodromy in $Sl(2, \mathbf{Z})$ only for integer k. The factorisation condition below, together with the form of $M(n_m, n_e)$ also given below, then imply that $k = 1$ is the only possibility.

plane. This contour can be deformed into a contour encircling u_0 and a contour encircling $-u_0$, both counterclockwise. It follows the factorisation condition on the monodromy matrices†

$$M_\infty = M_{u_0} M_{-u_0} \tag{5.12}$$

and hence

$$M_{-u_0} = \begin{pmatrix} -1 & 2 \\ -2 & 3 \end{pmatrix} . \tag{5.13}$$

What is the interpretation of this singularity at $u = -u_0$? To discover this, consider the behaviour under monodromy of the BPS mass formula $m^2 = 2|Z|^2$ with Z given by (4.14), i.e. $Z = (n_m, n_e)\begin{pmatrix} a_D \\ a \end{pmatrix}$. The monodromy transformation $\begin{pmatrix} a_D \\ a \end{pmatrix} \to M \begin{pmatrix} a_D \\ a \end{pmatrix}$ can be interpreted as changing the magnetic and electric quantum numbers as

$$(n_m, n_e) \to (n_m, n_e) M \ . \tag{5.14}$$

The state of vanishing mass responsible for a singularity should be invariant under the monodromy, and hence be a left eigenvector of M with unit eigenvalue. This is clearly so for the magnetic monopole: $(1, 0)$ is a left eigenvector of $\begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix}$ with unit eigenvalue. This simply reflects that $m^2 = 2|a_D|^2$ is invariant under (5.11). Similarly, the left eigenvector of (5.13) with unit eigenvalue is $(n_m, n_e) = (1, -1)$ This is a dyon. Thus the sigularity at $-u_0$ is interpreted as being due to a $(1, -1)$ dyon becoming massless.

More generally, (n_m, n_e) is the left eigenvector with unit eigenvalue‡ of

$$M(n_m, n_e) = \begin{pmatrix} 1 + 2n_m n_e & 2n_e^2 \\ -2n_m^2 & 1 - 2n_m n_e \end{pmatrix} \tag{5.15}$$

which is the monodromy matrix that should appear for any singularity due to a massless dyon with charges (n_m, n_e). Note that M_∞ as given in (5.3) is not of this form, since it does not correspond to a hypermultiplet becoming massless.

† There is an ambiguity concerning the ordering of M_{u_0} and M_{-u_0} which will be resolved below.

‡ Of course, the same is true for any (qn_m, qn_e) with $q \in \mathbf{Z}$, but according to the discussion in section 4.3 on the stability of BPS states, states with $q \neq \pm 1$ are not stable.

One notices that the relation (5.12) does not look invariant under $u \to -u$, i.e $u_0 \to -u_0$ since M_{u_0} and M_{-u_0} do not commute. The apparent contradiction with the $\mathbf{Z}_2$-symmetry is resolved by the following remark. The precise definition of the composition of two monodromies as in (5.12) requires a choice of base-point $u = P$ (just as in the definition of homotopy groups). Using a different base-point, namely $u = -P$, leads to

$$M_\infty = M_{-u_0} M_{u_0} \tag{5.16}$$

instead. Then one would obtain $M_{-u_0} = \begin{pmatrix} 3 & 2 \\ -2 & -1 \end{pmatrix}$, and comparing with (5.15), this would be interpreted as due to a $(1,1)$ dyon. Thus the $\mathbf{Z}_2$-symmetry $u \to -u$ on the quantum moduli space also acts on the base-point P, hence exchanging (5.12) and (5.16). At the same time it exchanges the $(1,-1)$ dyon with the $(1,1)$ dyon.

Does this mean that the $(1,1)$ and $(1,-1)$ dyons play a privileged role? Actually not. If one first turns k times around ∞, then around u_0, and then k times around ∞ in the opposite sense, the corresponding monodromy is $M_\infty^{-k} M_{u_0} M_\infty^k = \begin{pmatrix} 1-4k & 8k^2 \\ -2 & 1+4k \end{pmatrix} = M(1,-2k)$ and similarly $M_\infty^{-k} M_{-u_0} M_\infty^k = \begin{pmatrix} -1-4k & 2+8k+8k^2 \\ -2 & 3+4k \end{pmatrix} = M(1,-1-2k)$. So one sees that these monodromies correspond to dyons with $n_m = 1$ and any $n_e \in \mathbf{Z}$ becoming massless. Similarly one has e.g. $M_{u_0}^k M_{-u_0} M_{u_0}^{-k} = M(1-2k,-1)$, etc.

Let's come back to the question of how many singularities there are. Suppose there are p strong coupling singularities at $u_1, u_2, \ldots u_p$ in addition to the one-loop perturbative singularity at $u = \infty$. Then one has a factorisation analogous to (5.12):

$$M_\infty = M_{u_1} M_{u_2} \ldots M_{u_p} \tag{5.17}$$

with $M_{u_i} = M(n_m^{(i)}, n_e^{(i)})$ of the form (5.15). It thus becomes a problem of number theory to find out whether, for given p, there exist solutions to (5.17) with integer $n_m^{(i)}$ and $n_e^{(i)}$. For several low values of $p > 2$ it has been checked [2] that there are no such solutions, and it seems likely that the same is true for all $p > 2$.

6. The solution : determination of the low-energy effective action

So far we have seen that $a_D(u)$ and $a(u)$ are single-valued except for the monodromies around ∞, u_0 and $-u_0$. As is well-known from complex analysis, this means that $a_D(u)$ and $a(u)$ are really multi-valued functions with branch cuts, the branch points being ∞, u_0 and $-u_0$. A typical example is $f(u) = \sqrt{u}F(a,b,c;u)$, where F is the hypergeometric function. The latter has a branch cut from 1 to ∞. Similarly, $\sqrt{u}$ has a branch cut from 0 to ∞ (usually taken along the negative real axis), so that $f(u)$ has two branch cuts joining the three singular points $0, 1$ and ∞. When u goes around any of these singular points there is a non-trivial monodromy between $f(u)$ and one other function $g(u) = u^d F(a', b', c'; u)$. The three monodromy matrices are in (almost) one-to-one correspondence with the pair of functions $f(u)$ and $g(u)$.

In the physical problem at hand one knows the monodromies, namely

$$M_\infty = \begin{pmatrix} -1 & 2 \\ 0 & -1 \end{pmatrix} \quad , \quad M_{u_0} = \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix} \quad , \quad M_{-u_0} = \begin{pmatrix} -1 & 2 \\ -2 & 3 \end{pmatrix} \tag{6.1}$$

and one wants to determine the corresponding functions $a_D(u)$ and $a(u)$. As will be explained, the monodromies fix $a_D(u)$ and $a(u)$ up to normalisation, which will be determined from the known asymptotics (5.1) at infinity.

The precise location of u_0 depends on the renormalisation conditions which can be chosen such that $u_0 = 1$ [1]. Assuming this choice in the sequel will simplify somewhat the equations. If one wants to keep u_0, essentially all one has to do is to replace $u \pm 1$ by $\frac{u \pm u_0}{u_0} = \frac{u}{u_0} \pm 1$.

6.1. The differential equation approach

Monodromies typically arise from differential equations with periodic coefficients. This is well-known in solid-state physics where one considers a Schrödinger equation with a periodic potential*

$$\left[-\frac{\mathrm{d}^2}{\mathrm{d}x^2} + V(x)\right]\psi(x) = 0 \quad , \quad V(x + 2\pi) = V(x) \; . \tag{6.2}$$

There are two independent solutions $\psi_1(x)$ and $\psi_2(x)$. One wants to compare solutions at x and at $x + 2\pi$. Since, due to the periodicity of the potential V, the differential equation at

* The constant energy has been included into the potential, and the mass has been normalised to $\frac{1}{2}$.

$x + 2\pi$ is exactly the same as at x, the set of solutions must be the same. In other words, $\psi_1(x + 2\pi)$ and $\psi_2(x + 2\pi)$ must be linear combinations of $\psi_1(x)$ and $\psi_2(x)$:

$$\begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} (x + 2\pi) = M \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} (x) \tag{6.3}$$

where M is a (constant) monodromy matrix.

The same situation arises for differential equations in the complex plane with meromorphic coefficients. Consider again the Schrödinger-type equation

$$\left[-\frac{\mathrm{d}^2}{\mathrm{d}z^2} + V(z) \right] \psi(z) = 0 \tag{6.4}$$

with meromorphic $V(z)$, having poles at $z_1, \ldots z_p$ and (in general) also at ∞. The periodicity of the previous example is now replaced by the single-valuedness of $V(z)$ as z goes around any of the poles of V (with $z - z_i$ corresponding roughly to e^{ix}). So, as z goes once around any one of the z_i, the differential equation (6.4) does not change. So by the same argument as above, the two solutions $\psi_1(z)$ and $\psi_2(z)$, when continued along the path surrounding z_i must again be linear combinations of $\psi_1(z)$ and $\psi_2(z)$:

$$\begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} \left(z + e^{2\pi i}(z - z_i) \right) = M_i \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} (z) \tag{6.5}$$

with a constant 2×2-monodromy matrix M_i for each of the poles of V. Of course, one again has the factorisation condition (5.17) for M_∞. It is well-known, that non-trivial constant monodromies correspond to poles of V that are at most of second order. In the language of differential equations, (6.4) then only has *regular* singular points.

In our physical problem, the *two* multivalued functions $a_D(z)$ and $a(z)$ have 3 singularities with non-trivial monodromies at $-1, +1$ and ∞. Hence they must be solutions of a second-order differential equation (6.4) with the potential V having (at most) second-order poles precisely at these points. The general form of this potential is†

$$V(z) = -\frac{1}{4} \left[\frac{1 - \lambda_1^2}{(z+1)^2} + \frac{1 - \lambda_2^2}{(z-1)^2} - \frac{1 - \lambda_1^2 - \lambda_2^2 + \lambda_3^2}{(z+1)(z-1)} \right] \tag{6.6}$$

with double poles at $-1, +1$ and ∞. The corresponding residues are $-\frac{1}{4}(1 - \lambda_1^2)$, $-\frac{1}{4}(1 - \lambda_2^2)$

† Additional terms in V that naively look like first-order poles ($\sim \frac{1}{z-1}$ or $\frac{1}{z+1}$) cannot appear since they correspond to third-order poles at $z = \infty$.

and $-\frac{1}{4}(1-\lambda_3^2)$. Without loss of generality, I assume $\lambda_i \geq 0$. The corresponding differential equation (6.4) is well-known in the mathematical literature (see e.g. [17]) since it can be transformed into the hypergeometric differential equation. It has appeared, among others, in the study of the (classical) Liouville three-point function and the determination of constant curvature metrics on Riemann surfaces [18]. The transformation to the standard hypergeometric equation is readily performed by setting

$$\psi(z) = (z+1)^{\frac{1}{2}(1-\lambda_1)}(z-1)^{\frac{1}{2}(1-\lambda_2)} f\left(\frac{z+1}{2}\right) . \tag{6.7}$$

One then finds that f satisfies the hypergeometric differential equation

$$x(1-x)f''(x) + [c-(a+b+1)x]f'(x) - abf(x) = 0 \tag{6.8}$$

with

$$\begin{aligned} a &= \tfrac{1}{2}(1-\lambda_1-\lambda_2+\lambda_3) \\ b &= \tfrac{1}{2}(1-\lambda_1-\lambda_2-\lambda_3) \\ c &= 1-\lambda_1 \ . \end{aligned} \tag{6.9}$$

The solutions of the hypergeometric equation (6.8) can be written in many different ways due to the various identities between the hypergeometric function $F(a,b,c;x)$ and products with powers, e.g. $(1-x)^{c-a-b}F(c-a,c-b,c;x)$, etc. A convenient choice for the two independent solutions is the following [17]

$$\begin{aligned} f_1(x) &= (-x)^{-a}F(a,a+1-c,a+1-b;\frac{1}{x}) \\ f_2(x) &= (1-x)^{c-a-b}F(c-a,c-b,c+1-a-b;1-x) \ . \end{aligned} \tag{6.10}$$

f_1 and f_2 correspond to Kummer's solutions denoted u_3 and u_6. The choice of f_1 and f_2 is motivated by the fact that f_1 has simple monodromy properties around $x=\infty$ (i.e. $z=\infty$) and f_2 has simple monodromy properties around $x=1$ (i.e. $z=1$), so they are good candidates to be identified with $a(z)$ and $a_D(z)$.

One can extract a great deal of information from the asymptotic forms of $a_D(z)$ and $a(z)$. As $z \to \infty$ one has $V(z) \sim -\frac{1}{4}\frac{1-\lambda_3^2}{z^2}$, so that the two independent solutions behave asymptotically as $z^{\frac{1}{2}(1 \pm \lambda_3)}$ if $\lambda_3 \neq 0$, and as $\sqrt{z}$ and $\sqrt{z}\ln z$ if $\lambda_3 = 0$. Comparing with (5.4) (with $u \to z$) we see that the latter case is realised. Similarly, with $\lambda_3 = 0$, as $z \to 1$, one has $V(z) \sim -\frac{1}{4}\left(\frac{1-\lambda_2^2}{(z-1)^2} - \frac{1-\lambda_1^2-\lambda_2^2}{2(z-1)}\right)$, where I have kept the subleading term. From the logarithmic asymptotics (5.10) one then concludes $\lambda_2 = 1$ (and from the subleading term also $-\frac{\lambda_1^2}{8} = \frac{i}{\pi}\frac{c_0}{a_0}$). The $\mathbf{Z}_2$-symmetry ($z \to -z$) on the moduli space then implies that, as $z \to -1$, the potential V does not have a double pole either, so that also $\lambda_1 = 1$. Hence we conclude

$$\lambda_1 = \lambda_2 = 1 \ , \quad \lambda_3 = 0 \ \Rightarrow \ V(z) = -\frac{1}{4}\frac{1}{(z+1)(z-1)} \tag{6.11}$$

and $a = b = -\frac{1}{2}$, $c = 0$. Thus from (6.7) one has $\psi_{1,2}(z) = f_{1,2}\left(\frac{z+1}{2}\right)$. One can then verify, using the formulas in ref. [17] (and denoting the argument again by u rather than z) that the two solutions

$$\begin{aligned} a_D(u) &= i\psi_2(u) = i\frac{u-1}{2}F\left(\frac{1}{2},\frac{1}{2},2;\frac{1-u}{2}\right) \\ a(u) &= -2i\psi_1(u) = \sqrt{2}(u+1)^{\frac{1}{2}}F\left(-\frac{1}{2},\frac{1}{2},1;\frac{2}{u+1}\right) \end{aligned} \tag{6.12}$$

indeed have the required monodromies (6.1), as well as the correct asymptotics.

It might look as if we have not used the monodromy properties to determine a_D and a and that they have been determined only from the asymptotics. This is not entirely true, of course. The very fact that there are non-trivial monodromies only at $\infty, +1$ and -1 implied that a_D and a must satisfy the second-order differential equation (6.4) with the potential (6.6). To determine the λ_i we then used the asymptotics of a_D and a. But this is (almost) the same as using the monodromies since the latter were obtained from the asymptotics.

Using the integral representation [17] of the hypergeometric function, the solution (6.12) can be nicely rewritten as

$$\begin{aligned} a_D(u) &= \frac{\sqrt{2}}{\pi}\int_1^u \frac{\mathrm{d}x\,\sqrt{x-u}}{\sqrt{x^2-1}} \\ a(u) &= \frac{\sqrt{2}}{\pi}\int_{-1}^1 \frac{\mathrm{d}x\,\sqrt{x-u}}{\sqrt{x^2-1}} \ . \end{aligned} \tag{6.13}$$

One can invert the second equation (6.12) to obtain $u(a)$ and insert the result into $a_D(u)$ to obtain $a_D(a)$. Integrating with respect to a yields $\mathcal{F}(a)$ and hence the low-energy effective action. I should stress that this expression for $\mathcal{F}(a)$ is not globally valid but only on a certain portion of the moduli space. Different analytic continuations must be used on other portions.

6.2. The approach using elliptic curves

In their paper, Seiberg and Witten do not use the differential equation approach just described, but rather introduce an auxiliary construction: a certain elliptic curve by means of which two functions with the correct monodromy properties are constructed. I will not go into details here, but simply sketch this approach.

To motivate their construction *a posteriori*, we notice the following: from the integral representation (6.13) it is natural to consider the complex x-plane. More precisely, the integrand has square-root branch cuts with branch points at $+1, -1, u$ and ∞. The two branch cuts can be taken to run from -1 to $+1$ and from u to ∞. The Riemann surface of the integrand is two-sheeted with the two sheets connected through the cuts. If one adds the point at infinity to each of the two sheets, the topology of the Riemann surface is that of two spheres connected by two tubes (the cuts), i.e. a torus. So one sees that the Riemann surface of the integrand in (6.13) has genus one. This is the elliptic curve considered by Seiberg and Witten.

As is well-known, on a torus there are two independent non-trivial closed paths (cycles). One cycle (γ_2) can be taken to go once around the cut $(-1, 1)$, and the other cycle (γ_1) to go from 1 to u on the first sheet and back from u to 1 on the second sheet. The solutions $a_D(u)$ and $a(u)$ in (6.13) are precisely the integrals of some suitable differential λ along the two cycles γ_1 and γ_2:

$$a_D = \oint_{\gamma_1} \lambda \quad , \quad a = \oint_{\gamma_2} \lambda \quad , \quad \lambda = \frac{\sqrt{2}}{2\pi} \frac{\sqrt{x-u}}{\sqrt{x^2-1}} \mathrm{d}x \ . \tag{6.14}$$

These integrals are called period integrals. They are known to satisfy a second-order differential equation, the so-called Picard-Fuchs equation, that is nothing else than our Schrödinger-type equation (6.4) with V given by (6.11).

How do the monodromies appear in this formalism? As u goes once around $+1, -1$ or ∞, the cycles γ_1, γ_2 are changed into linear combinations of themselves with integer coefficients:

$$\begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix} \to M \begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix} \quad , \quad M \in Sl(2, \mathbf{Z}) \ . \tag{6.15}$$

This immediately implies

$$\begin{pmatrix} a_D \\ a \end{pmatrix} \rightarrow M \begin{pmatrix} a_D \\ a \end{pmatrix} \tag{6.16}$$

with the same M as in (6.15). The advantage here is that one automatically gets monodromies with *integer* coefficients. The other advantage is that

$$\tau(u) = \frac{\mathrm{d}a_D/\mathrm{d}u}{\mathrm{d}a/\mathrm{d}u} \tag{6.17}$$

can be easily seen to be the τ-parameter describing the complex structure of the torus, and as such is garanteed to satisfy

$$\mathrm{Im}\,\tau(u) > 0 \tag{6.18}$$

which was the requirement for positivity of the metric on moduli space.

To motivate the appearance of the genus-one elliptic curve (i.e. the torus) *a priori* - without knowing the solution (6.13) from the differential equation approach - Seiberg and Witten remark that the three monodromies are all very special: they do not generate all of $Sl(2,\mathbf{Z})$ but only a certain subgroup $\Gamma(2)$ of matrices in $Sl(2,\mathbf{Z})$ congruent to 1 modulo 2. Furthermore, they remark that the u-plane with punctures at $1,-1,\infty$ can be thought of as the quotient of the upper half plane H by $\Gamma(2)$, and that $H/\Gamma(2)$ naturally parametrizes (i.e. is the moduli space of) elliptic curves described by

$$y^2 = (x^2-1)(x-u) \ . \tag{6.19}$$

Equation (6.19) corresponds to the genus-one Riemann surface discussed above, and it is then natural to introduce the cycles γ_1, γ_2 and the differential λ from (6.13). The rest of the argument then goes as I just exposed.

7. Conclusions and outlook

In these notes, I have given a rather detailed, and hopefully pedagogical introduction to the work of Seiberg and Witten [1]. We have seen realised a version of electric-magnetic duality accompanied by a duality transformation on the expectation value of the scalar (Higgs) field, $a \leftrightarrow a_D$. There is a manifold of inequivalent vacua, the moduli space $\mathcal{M}$, corresponding to different Higgs expectation values. The duality relates strong coupling regions in $\mathcal{M}$ to the perturbative region of large a where the effective low-energy action is known asymptotically in terms of $\mathcal{F}$. Thus duality allows us to determine the latter also at strong coupling. The holomorphicity condition from $N = 2$ supersymmetry then puts such strong constraints on $\mathcal{F}(a)$, or equivalently on $a_D(u)$ and $a(u)$ that the full functions can be determined solely from their asymptotic behaviour at the strong and weak coupling singularities of $\mathcal{M}$.

There are a couple of questions one might ask, like what is the profound reason for the appearance of elliptic curves, or of the differential equation. It is intriguing to note that the latter with the potential (6.6) appears in conformal field theories as the null vector decoupling equation. It is satisfied by certain chiral conformal four-point correlation functions

$$\langle \mathcal{V}_{\Delta_4}(\infty)\mathcal{V}_{\Delta_3}(1)\mathcal{V}_{\Delta_2}(z)\mathcal{V}_{\Delta_1}(0)\rangle$$

where the $\mathcal{V}_\Delta$ are chiral vertex operators and where the conformal dimensions Δ_j are determined in terms of the $\lambda_1, \lambda_2, \lambda_3$. Whether this is a pure coincidence or has some deeper meaning does not seem to be clear at the moment.

Also, several generalisations of the pure $SU(2)$ Yang-Mills theory exposed here have been studied. One is to add matter hypermultiplets [4], another is to consider pure Yang-Mills theory but for gauge groups different from $SU(2)$ [2,3], or to allow for different gauge groups as well as matter [19]. Here let me only note that for the pure $SU(3)$ theory, solving the condition $[\phi, \phi^+] = 0$ leads to $\phi = a_1 H_1 + a_2 H_2$ where H_i are the two Cartan generators of $SU(3)$, so that one has a two-complex dimensional moduli space, parametrized by a_1, a_2 or rather by $u = \langle \mathrm{tr}\ \phi^2\rangle$ and $v = \langle \mathrm{tr}\ \phi^3\rangle$. The duals are $a_{Di} = \frac{\partial \mathcal{F}}{\partial a_i}$, $i = 1, 2$. The monodromies in moduli space (i.e. the (u, v)-space) then act on the four-component object $(a_{D1}(u, v), a_{D2}(u, v), a_1(u, v), a_2(u, v))$. They can be reproduced from period integrals of some hyperelliptic curve [2]. The corresponding (Picard-Fuchs) differential equations are two-partial differential equations in u and v [2] with solutions given by Appel functions [17] that generalise the hypergeometric function to two variables.

Last, but not least, I should mention that similar duality ideas in string theory have led to yet another explosion of this domain of theoretical physics. A particular nice link with the field theory discussed here has been made in [20] where the field theoretic duality is related to string dualities.

Acknowledgements:

It is a pleasure to thank all participants of the duality study group at the Ecole Normale for discussions, Luis Alvarez-Gaumé for making his notes [7] available to me well before publication, and Marc Rosso for giving me the opportunity to present this lecture at the Strasbourg meeting.

REFERENCES

1. N. Seiberg and E. Witten, *Electric-magnetic duality, monopole condensation, and confinement in $N = 2$ supersymmetric Yang-Mills theory*, Nucl. Phys. **B426** (1994) 19, `hep-th/9407087`.

2. A. Klemm, W. Lerche, S. Yankielowicz and S. Theisen, *Simple singularities and $N = 2$ supersymmetric Yang-Mills theory*, Phys. Lett. **B344** (1995) 169, `hep-th/9411048`;
A. Klemm, W. Lerche and S. Theisen, *Nonperturbative effective actions of $N = 2$ supersymmetric gauge theories*, CERN preprint CERN-TH/95-104, `hep-th/9505150`.

3. P.C. Argyres and A.E. Faraggi, *The vacuum structure and spectrum of $N = 2$ supersymmetric $SU(N)$ gauge theory*, Phys. Rev. Lett. **74** (1995) 3931, `hep-th/9411057`;
U.H. Danielsson and B. Sundborg, *The moduli space and monodromies of $N = 2$ supersymmetric $SO(2r+1)$ Yang-Mills theory*, Phys. Lett. **B358** (1995) 273, `hep-th/9504102`;
A. Brandhuber and K. Landsteiner, *On the monodromies of $N = 2$ supersymmetric Yang-Mills theory with gauge group $SO(2N)$*, Phys. Lett. **B358** (1995) 73, `hep-th/9507008`.

4. N. Seiberg and E. Witten, *Monopoles, duality and chiral symmetry breaking in $N = 2$ supersymmetric QCD*, Nucl. Phys. **B431** (1994) 484, `hep-th/9408099`;
P.C. Argyres and A.D. Shapere, *The vacuum structure of $N = 2$ super QCD with classical gauge groups*, Rutgers preprint RU-95-61, `hep-th/9509175`.

5. J. Polchinski, *Recent results in string duality*, ITP Santa Barbara preprint, hep-th/9511157.

6. A. Strominger, *Black hole condensation and duality in string theory*, Santa Barbara preprint UCSBTH-95-29, hep-th/9510207.

7. C. Gomez, Electric-magnetic duality and effective field theories, Madrid preprint FTUAM 95/36, hep-th/9510023;
L. Alvarez-Gaumé, notes of lectures given at CERN in the fall of 1994, to be published.

8. G. 't Hooft, Nucl. Phys. **B79** (1974) 276;
A.M. Polyakov, JETP Letters **20** (1974) 194.

9. C. Montonen and D. Olive, *Magnetic monopoles as gauge particles?*, Phys. Lett. **72B** (1977) 117.

10. D. Olive, *Exact electromagnetic duality*, Swansea preprint SWAT/94-95/81, hep-th/9508089.

11. M.F. Sohnius, *Introducing supersymmetry*, Phys. Rep. **128** (1985) 39;
P. Fayet and S. Ferrara, *Supersymmetry*, Phys. Rep. **32C** (1977) 1.

12. J. Gates, *Superspace formulation of new non-linear sigma models*, Nucl. Phys. **B238** (1984) 349.

13. N. Seiberg, *Supersymmetry and non-perturbative beta functions*, Phys. Lett. **B206** (1988) 75.

14. S. Coleman, *Classical lumps and their quantum descendants*, in: Aspects of Symmetry, Cambridge University Press, 1985

15. E. Witten and D. Olive, *Supersymmetry algebras that include topological charges*, Phys. Lett. **78B** (1978) 97.

16. A. Sen, *Dyon-monopole bound states, self-dual harmonic forms on the multi-monopole moduli space, and* $Sl(2,\mathbf{Z})$ *invariance in string theory*, Phys. Lett. **329B** (1994) 217, hep-th/9402032.

17. A. Erdelyi et al, *Higher Transcendental Functions*, Vol 1, McGraw-Hill, New York, 1953.

18. A. Bilal and J.-L. Gervais, *Exact quantum three-point function of Liouville highest weight states*, Nucl. Phys. **B305** (1988) 33.

19. U.H. Danielsson and B. Sundborg, *Exceptional Equivalences in $N = 2$ supersymmetric Yang-Mills theory*, Uppsala preprint USITP-95-12, `hep-th/9511180`.

20. S. Kachru, A. Klemm, W. Lerche, P. Mayr and C. Vafa, *Nonperturbative results on the point particle limit of N=2 heterotic string compactifications*, CERN/Harvard preprint CERN-TH/95-231, HUTP-95/A032, `hep-th/9508155`.