UNIVERSITY OF AMSTERDAM

MSc Physics
Theoretical Physics

Master Thesis

Dualities in Gauge Theories and their Geometric Realization

---

by

S.P.G. van Leuven
5756561

May 2014
60 ECTS
February 2013-May 2014

Supervisor:
Prof.dr. E.P. Verlinde

Second Supervisor:
Prof.dr. J. de Boer

ITFA

Abstract

In this thesis, exact results in $\mathcal{N} = 2$ super Yang-Mills theories are discussed. Starting from the exact solution of pure SYM provided by Seiberg and Witten through an elliptic curve construction, generalizations are discussed such as the inclusion of hypermultiplets. Special emphasis is put on the selfdual $N_f = 4$ theory, which will play a central role in remainder. Furthermore, a brief summary is given of the M-theory construction of Witten to determine the Seiberg-Witten curves of gauge theories with arbitrary unitary product gauge groups, also dubbed quiver gauge theories. This provides the necessary introduction to understand the recent developments made by Gaiotto. Superconformal quiver theories are associated in a very precise manner to punctured Riemann surfaces $C_{n,g}$. Gaiotto's conjecture that the UV moduli space of the quiver theories equals the moduli space of the associated Riemann surface is explained in detail. Furthermore, explicit checks of his proposal are performed by comparing the boundaries of the moduli spaces for $\mathrm{T}_{n,g}[A_1]$, $g = 0, 1$. We briefly re-examine the M-theory construction in this new light, essentially explaining the reduction of the elusive 6d $(2, 0)$ theory on $C_{n,g}$. At last, the extension to A_2 and A_{N-1} quiver gauge theories is discussed.

Contents

1 Introduction **1**

2 Supersymmetry **3**

2.1 Superspace and Superfields . . . 3
2.2 Super Yang-Mills Theory . . . 5
2.3 Coupling SYM to Matter . . . 8
2.4 R-Symmetry and an Anomaly . . . 8
2.5 Exact Renormalization of the Prepotential . . . 11

3 Seiberg-Witten Theory **15**

3.1 The Classical Moduli Space . . . 15
3.2 Duality of the Super Yang-Mills Action . . . 18
3.3 Central Charge and Dual Theories . . . 20
3.4 Monodromies and the Quantum Moduli Space . . . 22
3.5 Solution of the Model . . . 27
3.6 Discussion . . . 33
3.7 Confinement of Electric Charge . . . 34
3.8 Adding Matter . . . 36
3.8.1 Global Symmetries of $N_f \leq 4$ . . . 37
3.8.2 Moduli Spaces of $N_f \leq 3$ . . . 39
3.8.3 $N_f = 4$ Theory . . . 41
3.9 Seiberg-Witten Curves from M-Theory . . . 45

4 Gaiotto Dualities **49**

4.1 Quiver Gauge Theories . . . 50
4.2 A Geometric Realization of Duality . . . 56
4.3 Mathematical Description . . . 61
4.3.1 Massless $T_{4,0}[A_1]$ . . . 62
4.3.2 Massive $T_{4,0}[A_1]$ . . . 64
4.3.3 Massless $T_{m,0}[A_1]$ . . . 67
4.3.4 Massive $T_{m,0}[A_1]$ . . . 67
4.3.5 $T_{m,1}[A_1]$ . . . 69
4.4 Summary: A Six Dimensional Construction . . . 73
4.5 Generalizations . . . 75
4.5.1 $\mathrm{T}_{(\mathrm{f}_1,\mathrm{f}_3),\mathrm{g}}[\mathrm{A}_2]$ . . . 76
4.5.2 $\mathrm{T}_{(\mathrm{f}_\mathrm{a}),\mathrm{g}}[\mathrm{A}_{\mathrm{N}-1}]$ . . . 80

5 Summary and Outlook **84**

A Electric-magnetic Duality **86**

B Supersymmetry Multiplets and Superfields Expansions **88**
B.1 Vector and Hypermultiplet 88
B.2 Expansion of Superfields 88

C Half Hypers, Bifundamentals and Trifundamentals **90**

D Explicit Expression Quadratic Differential **94**

Chapter 1

Introduction

Yang-Mills theories with $\mathcal{N} = 2$ supersymmetry present an intriguing playground to help understand the non-perturbative dynamics of non-abelian gauge theories, which arise inescapably in the IR of UV free theories. Starting with the developments of Seiberg and Witten in [59] and [60], it was realized that the low energy effective actions of a large class of spontaneously broken supersymmetric Yang-Mills theories, possibly coupled to hypermultiplets, allow a purely geometrical description in terms of (hyper)elliptic curves. The basic property of $\mathcal{N} = 2$ gauge theories which underlies the appearance of elliptic curves is the complex moduli space of vacua on which the effective action depends holomorphically. A beautiful interplay between physical assumptions and the mathematical rigidity of complex analysis fixes the effective action completely. An infinite series of non-perturbative corrections was exactly computed, standing in great contrast to increasingly difficult explicit instanton calculations.[1] Whereas Seiberg and Witten originally studied the low energy effective behaviour of $SU(2)$ theories coupled to $N_f \leq 4$ flavours, the analysis was subsequently generalized to $SU(N), SO(N), Sp(N)$ gauge groups with and without the inclusion of hypermultiplets[2][9][11][48].

The fact that the appearance of geometry seems a generic property of $\mathcal{N} = 2$ gauge theories leads to the question whether the elliptic curves are purely auxiliary or if they carry an actual physical interpretation. A natural place to search for an answer to this question is in the realm of string- or M-theory, since these theories try to encode our four dimensional world through geometric constructions in extra dimensions. And indeed, it was found by Witten in [72] that Seiberg-Witten curves can be constructed in Type IIA or M-theory using certain brane configurations. Whereas four-dimensional spacetime is wrapped by the branes, they additionally extend in extra dimensions to form the Seiberg-Witten curves. This insight extended the class of gauge theories described by complex curves enormously, primarily opening the door to arbitrary product gauge theories.[2] But not only does the M-theory construction give a straightforward recipe to construct Seiberg-Witten curves, it also naturally knows about rather non-trivial aspects of gauge charges in the theories[31][40]. By natural it is meant that the natural geometry of M-theory inadvertently explains these fundamental questions. This demonstrates the power of M-theory in its deep capability of explaining complicated four-dimensional physics in terms of only two basic objects: the M2 and M5 brane.

The M-theory construction of Witten was further scrutinized by Gaiotto, with success[29]. Gaiotto studied $\mathcal{N} = 2$ superconformal field theories and was able to determine elementary building blocks, in particular for $SU(2)$ and $SU(3)$ gauge theories but also the road towards $SU(N)$, to construct product gauge theories in the form of generalized quiver diagrams. Generalized quiver

[1] In 2002, Nekrasov devised a method to calculate instanton corrections in a more direct manner[45]. Although this method provides a check on the Seiberg-Witten solution, it will not be discussed in this thesis.

[2] Witten found the brane constructions for products of unitary gauge groups. See [1] for products of symplectic and orthogonal gauge groups.

diagrams are trivalent graphs which provide a convenient depiction of both gauge *and* flavour symmetries and the corresponding matter representations. This realization was subsequently understood to allow for a natural identification of the quiver theories with genus g Riemann surfaces with a variety of punctures, corresponding to possible flavour symmetries. A degeneration of the Riemann surface is interpreted as a weak coupling limit of the gauge theory; different degenerations of the Riemann surface correspond to different decoupling limits of the gauge theory. This suggests the conjecture made by Gaiotto: the UV moduli space of the gauge theory coincides with moduli space of these Riemann surfaces. In particular, the boundaries of the moduli spaces are matched, mathematically checked through the construction of the Seiberg-Witten curves, and surprising dual descriptions for certain quiver gauge theories are found. Theories underlying all dual descriptions are denoted by $\mathrm{T}_{f_1,\ldots,f_N,g}[A_N]$, making a clear two quivers are dual whenever the number and type of punctures and genus coincides. Although in the $SU(2)$ case the possible dual descriptions of a single theory are relatively mild, the analysis extends to higher rank gauge theories. For general $SU(N)$ quiver gauge theories, a rich zoo of building blocks exist to produce SCFTs[17]. For these higher rank theories, dual descriptions of a seemingly ordinary $SU(N)$ theory generically include non-Lagrangian isolated interacting SCFTs whose flavour symmetries are partly gauged and coupled to the rest of the quiver. This is much in the spirit of Argyres-Seiberg duality [10] and indeed, Argyres-Seiberg duality is recovered as a special case of this much larger web of dualities.

Apart from the fact the Riemann surfaces provide a useful tool to understand dualities in gauge theories, they also provide an explicit UV-IR correspondence. The distinction between the exactly marginal gauge coupling, parametrizing the UV, and the physical gauge coupling as understood from the Seiberg-Witten curve arises naturally and explains the discrepancy between the original assumptions of Seiberg-Witten of no renormalization and the explicit calculations in [44] in a beautiful geometrical way. From the M-theory construction, it becomes apparent the Seiberg-Witten curves are realized as wrapping the Riemann surface in a very particular way: the curves are represented as k-differentials living in the k^{th} symmetric power of the cotangent bundle of the Riemann surface for gauge group $SU(k)$. This additional structure is responsible for a correct identification of UV and IR parameters, and geometrically encodes how S-duality in the UV translates to IR dual descriptions.

The goal of this thesis is to provide an introduction to the geometric aspects of $\mathcal{N} = \in$ gauge theories, culminating in the discussion of Gaiotto's article. To this end, in Chapter 2 certain aspects of supersymmetry are briefly reviewed, including the low energy effective action of $\mathcal{N} = 2$ super Yang-Mills, the anomaly in the $U(1)_R$ symmetry and the precise form of the renormalized prepotential following the original argument of Seiberg[57]. In Chapter 3, the Seiberg-Witten solution to pure $SU(2)$ SYM is extensively discussed, focussing on the elliptic curve construction. As a short intermezzo, we utilize the solution to understand confinement in $\mathcal{N} = 1$ SYM, which is shown to be caused by the condensation of magnetic monopoles. We resume with the Seiberg-Witten analysis of the $SU(2)$ theory coupled to flavours, focussing on qualitative aspects such as the structure of the moduli spaces, global symmetries and, for $N_f = 4$, its self-duality properties. At last, we finish the chapter with a brief discussion of the M-theory construction of Witten, providing the general curve for an arbitrary unitary product gauge group theory. Having all the tools at hand, Chapter 4 tries to present a complete discussion of Gaiotto's '$\mathcal{N} = 2$ Dualities'. Starting with the conceptual idea on how to relate the quivers to Riemann surface, we will provide a thorough quantitative analysis on the Seiberg-Witten curves corresponding to $g = 0, 1$ A_1 quiver gauge theories. Checks of statements in the more conceptual sections are performed. We then review the six-dimensional origin of the quiver theories, and conclude with a discussion of the extension to $SU(3)$ and $SU(N)$ quiver theories.

Chapter 2

Supersymmetry

An essential ingredient in the forthcoming analysis will be supersymmetry. Supersymmetry is a theoretically proposed symmetry between bosons and fermions. It states that every fermion in a particular theory must have a boson superpartner and vice versa. In contemporary high energy physics this concept is ubiquitous. The most important reason for this is that supersymmetry naturally cures a lot of issues that appear when one renormalizes a quantum field theory. However, it has also been found to be an indispensable concept in string theories, hence the name superstring theories. Extended supersymmetry also plays an essential role in the AdS/CFT correspondence. However, because as of yet no signs have shown up of supersymmetry at experiments, in this thesis we adopt an opportunistic point of view in which we consider supersymmetry a valuable calculational tool to uncover exact results in gauge theories, whether or not supersymmetry is realized in nature. If it is not, at least we hope it teaches us lessons about structures in gauge theories which in the end are independent of supersymmetry.

In this chapter we will study the exact renormalization of the $\mathcal{N} = 2$ supersymmetric Yang-Mills action. We will find that holomorphic properties of supersymmetric theories allow us to fix the precise form of the low energy effective action. That is: we will find an exact expression for all perturbative contributions and the general form of non-perturbative contributions. Before turning to this, we briefly introduce some aspects of supersymmetry which include the notion of superspace and superfields, the $\mathcal{N} = 2$ SYM action and R-symmetry.

The primary source for the first part is [27], which contains among other things a good introduction to the superfield formalism in $\mathcal{N} = 1$ supersymmetry. The content on extended supersymmetry primarily comes from [37]. Other references include the standard works on supersymmetry [68] and [69]. The sections on R-symmetry and renormalization also borrow from the reviews [4], [33] and [63].

2.1 Superspace and Superfields

Superspace presents a mathematical formalism which allows one to write manifestly supersymmetric actions, just like Minkowski spacetime allows one to write manifestly Lorentz invariant actions. Superspace is an extension of ordinary Minkowski spacetime: next to the ordinary bosonic coordinates x^μ it also has four Grassmann valued fermionic coordinates θ_a:

$$\theta_a = \begin{pmatrix} \theta_\alpha \\ \bar{\theta}_{\dot{\alpha}} \end{pmatrix}. \tag{2.1.1}$$

By construction, this is a Majorana spinor. As such, it provides the minimal amount of independent components a spinor can have in four dimensions and therefore is the minimal fermionic extension of ordinary spacetime. In the following, we will denote the full spinor by θ_a, whereas its chiral components will be denoted as $\theta_\alpha \equiv \theta$ and $\bar{\theta}_{\dot{\alpha}} \equiv \bar{\theta}$

To gain some intuition in the abstract notion of superspace, we describe the construction of superspace from the super Poincaré group. To appreciate the construction, we first look at normal Minkowski spacetime as constructed from the Poincaré group. The Poincaré group consists of translations and Lorentz transformations and is postulated to be a global symmetry of quantum field theories. The relation between the group and spacetime is made as follows: if we associate a point x^μ in Minkowski spacetime to a translation element $\exp(ix^\mu P_\mu)$ in the Poincaré group, i.e. we choose an origin, we can reach every point in Minkowski spacetime by applying a (left) multiplication with another translation element:

$$\begin{aligned} &\exp(iy^\mu P_\mu)\cdot\exp(ix^\mu P_\mu) = \exp(i(y^\mu + x^\mu)P_\mu) \\ \Leftrightarrow\ & x^\mu \mapsto x^\mu + y^\mu . \end{aligned}$$

Since Lorentz transformations keep the origin fixed, any translation followed by a Lorentz transformation will *not* be a translation. This means that every translation determine a unique right coset of the Lorentz group, considered as a subgroup of the Poincaré group. Clearly, there are no more right cosets of Lorentz group than translations determine. Therefore, there is a one-to-one correspondence between points in Minkowski space and the right cosets of the Lorentz group. This construction allows for a more general approach to the construction of spacetime: spacetime can be defined as the set of right cosets of the Lorentz group embedded in some larger group acting on spacetime.

However, a no-go theorem of Coleman and Mandula states that any additional continuous global symmetry of a quantum field theory cannot mix in a non-trivial way with the Poincaré group, or equivalently: acts trivially on spacetime.[1] So it seems there is not 'some larger group' to act on spacetime.

Of course, the Coleman-Mandula theorem can be circumvented. This is achieved through the introduction of a graded or super Lie algebra. Generators of a graded Lie algebra satisfy commutation or *anti*commutation relations. More to the point, we introduce the super Poincaré algebra which has an extra anticommuting spinorial generator, the so-called supercharge:

$$Q_a = \begin{pmatrix} Q_\alpha \\ \bar{Q}_{\dot{\alpha}} \end{pmatrix},\ \alpha, \dot{\alpha} = 1, 2$$

It generates translations in the coordinate θ_a and mixes the bosonic coordinate x^μ with θ_a, which we will see in more detail below. The additional commutation relations of the super Poincaré algebra are given by:

$$\begin{aligned} \left[P_\mu, Q_a^A\right] &= 0 \\ \left[M_{\mu\nu}, Q_\alpha^A\right] &= -\frac{1}{2}\left(\sigma_{\mu\nu}\right)_\alpha{}^\beta Q_\beta^A \\ \left[M_{\mu\nu}, \bar{Q}_{\dot{\alpha}}^A\right] &= \frac{1}{2}\left(\bar{\sigma}_{\mu\nu}\right)_{\dot{\alpha}}{}^{\dot{\beta}} \bar{Q}_{\dot{\beta}}^A \\ \{Q_\alpha^A, \bar{Q}_{\dot{\beta}B}\} &= 2\left(\sigma^\mu\right)_{\alpha\dot{\beta}} P_\mu \delta^A{}_B \\ \{Q_\alpha^A, Q_\beta^B\} &= \epsilon_{\alpha\beta} Z^{AB} \\ \{\bar{Q}_{\dot{\alpha}}^A, \bar{Q}_{\dot{\beta}}^B\} &= \epsilon_{\dot{\alpha}\dot{\beta}}\left(Z^\dagger\right)^{AB} \end{aligned}$$

where we added the capital Latin index to denote the possibility of the introduction of more than one supercharge. Z^{AB} is called the central charge because it commutes with all other generators of the

[1]Flavour symmetries, for instance, are internal symmetries of the fields and therefore their generators fully commute with the Poincaré algebra.

algebra. It is antisymmetric, such that it necessarily vanishes for $\mathcal{N}=1$ supersymmetry, $\mathcal{N}$ denoting the number of supercharges. Detailed discussions on the construction of the supersymmetry algebra may be found in [68] and [69].

Superspace is defined as the space of right cosets of the Lorentz group, now considered as a subgroup of the super Poincaré group. In analogy with the construction of Minkowski spacetime, we associate a point in superspace with a translation element in the super Lie group:

$$\exp\left(ix^\mu P_\mu\right)\exp\left(\bar\theta^a Q_a\right) \tag{2.1.2}$$

We still have P_μ as the generator of spacetime translations. However, looking at the action of a superspace translation generated by Q, we see:

$$\begin{aligned} &\exp\left(\bar\epsilon Q\right)\cdot\exp\left(\bar\theta Q\right) = \exp\left((\bar\epsilon+\bar\theta)Q - \bar\epsilon\gamma^\mu\theta P_\mu\right) \\ \Leftrightarrow\ &(x^\mu,\theta)\mapsto(x^\mu-\bar\epsilon\gamma^\mu\theta,\theta+\epsilon), \end{aligned}$$

where the Baker-Campbell-Hausdorff formula was used. The fact that a translation of ϵ in fermionic coordinates also affects the bosonic coordinate x^μ is a sign of the mixing of bosons and fermions: supersymmetry. This is the essence of the implementation of supersymmetry by using the superspace formalism.

Let us elucidate this a bit by considering the action of a super translation on a superfield. Superfields $\Phi(x,\theta,\bar\theta)$ are defined as differentiable functions on superspace. Since the coordinate θ is Grassmann valued, the Taylor expansion of Φ in θ is finite. For instance, the highest components of the superfield could look like:

$$\sim \bar\theta^2\theta^\alpha\lambda_\alpha(x) + \theta^2\bar\theta^{\dot\alpha}\bar\lambda_{\dot\alpha}(x) + \theta^2\bar\theta^2 F(x)$$

with $F(x)$ a bosonic field and λ a fermion. In general, we will be concerned only with superfields containing (complex) scalars, spinors or vector fields.

Integrals of arbitrary functions of superfields over superspace, i.e. $\int \mathrm{d}\bar\theta^2\mathrm{d}\theta^2\mathcal{G}(\Phi_i,\ldots,W_\alpha)$, automatically provide supersymmetry invariant actions. This can be seen from the action of the superspace translation above which acts on component fields as:

$$f(x^\mu)\to f(x^\mu-\bar\epsilon\gamma^\mu\theta) = f(x^\mu)+\partial_\mu f(x^\mu)\bar\epsilon\gamma^\mu\theta$$

With a little work one can now check that the supersymmetry variation of a function of superfields contains a total derivative at its highest component. Upon integrating over all of space, the supersymmetry variation vanishes. Therefore, the superspace integration automatically provides supersymmetry invariant actions.

As a general superfield contains a large amount of component fields, we usually consider constrained superfields. The expansions of the superfields we will use are given in Appendix B. These constrained superfields are called (anti)chiral as they only depend on θ ($\bar\theta$) and a particular combination:

$$y^\mu = x^\mu \pm i\theta\sigma^\mu\bar\theta$$

Highest components of (functions of) chiral superfields are proportional to θ^2 and therefore need only to be integrated over half of superspace to render an action supersymmetric: $\int \mathrm{d}\theta^2$ or $\int \mathrm{d}\bar\theta^2$. In the following we will also call these chiral functions *holomorphic*.

2.2 Super Yang-Mills Theory

The general $\mathcal{N}=2$ SYM action is written conveniently in terms of an $\mathcal{N}=1$ vector and chiral multiplet. These multiplets are represented respectively by the field strength W^α, a chiral superfield which satisfies an additional reality condition, and Φ, a general chiral superfield. Expansions of both superfields are given in Appendix B. Together, these superfields make up an $\mathcal{N}=2$ vector multiplet, containing spins $s \leq 1$. The action is:

$$S = \int \mathrm{d}^4x \left[\mathrm{Im}\left(\frac{\tau_{\mathrm{cl}}}{8\pi} \int \mathrm{d}^2\theta\, W^\alpha W_\alpha \right) + \frac{1}{g^2} \int \mathrm{d}^2\theta\, \mathrm{d}^2\bar{\theta}\, \Phi^\dagger e^{-2V} \Phi \right] \tag{2.2.1}$$

$$= \mathrm{Im} \int \mathrm{d}^4x \frac{\tau_{\mathrm{cl}}}{4\pi} \left[\int \mathrm{d}^2\theta\, \frac{1}{2} W^\alpha W_\alpha + \int \mathrm{d}^2\theta\, \mathrm{d}^2\bar{\theta}\, \Phi^\dagger e^{-2V} \Phi \right] \tag{2.2.2}$$

with $\tau_{\mathrm{cl}} = \frac{\theta}{2\pi} + \frac{4\pi i}{g^2}$ the complexified gauge coupling. The vector superfield V is directly related to W_α, and therefore does not contain additional fields. The action, then, represents the most general renormalizable gauge invariant $\mathcal{N}=2$ action. The gauge indices are suppressed. We will only consider W^α and Φ to be in the adjoint representation of the gauge group. This means the superfields are matrix valued in a basis of the Lie algebra $\{T_a\}$ which is represented by the structure constants of the Lie algebra: $\mathcal{D}\left(T_a\right)_b^c = i f_{ab}^c$.

In this thesis, we will primarily be interested in a low energy effective action. In particular, we are interested in the Wilsonian effective action. The Wilsonian effective action is obtained by integrating out certain momentum modes between a UV cutoff Λ_{UV} and an IR cutoff μ, where the μ symbol is chosen since the IR cutoff determines the scale of the effective theory. The reason to consider the Wilsonian effective action instead of the 1PI generating functional $\Gamma(\phi)$ is the Wilsonian renormalization preserves the holomorphicity of the field operators, guaranteeing unbroken supersymmetry. Furthermore, the effective action is also guaranteed to remain a holomorphic function of the bare couplings. This allows for powerful non-renormalization theorems[12][61].

We will now derive the general form of the low energy theory for $\mathcal{N}=2$ SYM. This is achieved most easily by considering an $\mathcal{N}=2$ superspace formulation. This requires the introduction of an extra fermionic coordinate $\tilde{\theta}_a$. In this formulation the most general pure gauge and renormalizable $\mathcal{N}=2$ Yang-Mills action is:

$$S = \frac{1}{4\pi} \mathrm{Im} \left(\int \mathrm{d}x^4 \int \mathrm{d}^2\theta\, \mathrm{d}^2\tilde{\theta}\, \frac{1}{2} \tau\, \mathrm{Tr}\, \Psi^2 \right) \tag{2.2.3}$$

Here, Ψ is a chiral $\mathcal{N}=2$ superfield. Expanding in the $\tilde{\theta}$ coordinates, it reads:

$$\Psi(\tilde{y}^\mu, \theta, \tilde{\theta}) = \Phi(\tilde{y}^\mu, \theta) + \sqrt{2}\tilde{\theta}^\alpha W_\alpha(\tilde{y}^\mu, \theta) + \tilde{\theta}^2 G(\tilde{y}^\mu, \theta) \tag{2.2.4}$$

where $\tilde{y}^\mu = y^\mu + i\tilde{\theta}\sigma_\mu \bar{\tilde{\theta}}$ and $y^\mu = x^\mu + i\theta\sigma_\mu\bar{\theta}$. Furthermore, G is an auxiliary $\mathcal{N}=1$ superfield and can be written in terms of the fields Φ and V:

$$G(\tilde{y}^\mu, \theta) = \int \mathrm{d}\bar{\theta}^2 \Phi^\dagger(\tilde{y}^\mu - i\theta\sigma\mu\bar{\theta}, \bar{\theta}) e^{-2V(\tilde{y}^\mu - i\theta\sigma\mu\bar{\theta}, \theta, \bar{\theta})} \tag{2.2.5}$$

where the integral should be taken at fixed $\tilde{y}$ such that G remains a function of $\tilde{y}$ and guarantees Ψ is a chiral $\mathcal{N}=2$ superfield. The constraint on G makes sure that Ψ^2 integrated over half of the tilde superspace yields the familiar $\mathcal{N}=1$ superspace action (2.2.2), as can be checked by plugging in expression (2.2.4) into (2.2.3).

The requirement for $\mathcal{N}=2$ supersymmetry is, analogously to $\mathcal{N}=1$ supersymmetry, that functions integrated over (anti)chiral half of $\mathcal{N}=2$ superspace should be (anti)chiral. Hence, since $\mathcal{N}=2$ supersymmetry is not broken by the renormalization, the action should take the form:

$$S = \frac{1}{4\pi} \mathrm{Im} \left(\int \mathrm{d}x^4 \int \mathrm{d}^2\theta \mathrm{d}^2\tilde{\theta}\, \mathcal{F}(\Psi) \right) \tag{2.2.6}$$

The function $\mathcal{F}(\Psi)$ is holomorphic in Ψ and is called the prepotential. Its form is thus far constrained only by holomorphy. To make contact with our $\mathcal{N}=1$ formulation we may Taylor expand (2.2.6) around the superfield Φ. Since this expansion is finite due to the anticommutativity of the fermionic coordinates, we are still left with an exact form of the action. Integrating over the chiral part of tilde superspace leaves us with:

$$S = \frac{1}{4\pi}\text{Im} \int \mathrm{d}x^4 \left[\int \mathrm{d}^2\theta \, \frac{1}{2}\mathcal{F}''_{ab}(\Phi)W^{\alpha a}W^b_\alpha + \int \mathrm{d}^2\theta \, \mathrm{d}^2\bar{\theta} \, (\Phi^\dagger e^{-2V})^a \mathcal{F}'_a(\Phi)\right] \tag{2.2.7}$$

The prepotential $\mathcal{F}(\Phi)$ is now a holomorphic function of the chiral field Φ. Its derivatives with respect to the Lie algebra valued field $\Phi = \Phi^a T_a$ are denoted with a subscript, i.e. $\mathcal{F}'_a(\Phi) = \frac{\partial \mathcal{F}(\Phi)}{\partial \Phi^a}$ and $\mathcal{F}''_{ab}(\Phi) = \frac{\partial^2 \mathcal{F}(\Phi)}{\partial \Phi^a \partial \Phi^b}$. The indices are to be summed over and replace the trace. We recognize the renormalized gauge coupling as:

$$\tau^{eff}_{ab}(\Phi) = \mathcal{F}''_{ab}(\Phi) \tag{2.2.8}$$

In the next chapter we will assume all non-abelian degrees of freedom are frozen out in our low energy effective action due to a Higgs condensation. The resulting abelian action is given by:

$$S = \frac{1}{4\pi}\text{Im} \int \mathrm{d}x^4 \left[\int \mathrm{d}^2\theta \, \frac{1}{2}\frac{\partial^2 \mathcal{F}(\Phi)}{\partial \Phi^2} W^\alpha W_\alpha + \int \mathrm{d}^2\theta \, \mathrm{d}^2\bar{\theta} \, \Phi^\dagger \frac{\partial \mathcal{F}(\Phi)}{\partial \Phi}\right] \tag{2.2.9}$$

Notice that the interaction of the chiral multiplet with the vector multiplet has disappeared because the adjoint representation of an abelian group is trivial. In the next chapter this action will be determined exactly.

The superspace formalism renders the action manifestly $\mathcal{N} = 1$ supersymmetric. Traces of the full $\mathcal{N} = 2$ supersymmetry can be recognized in that the gauge kinetic term and the Kähler potential

$$K(\Phi, \Phi^\dagger) = \text{Im}\left(\Phi^\dagger \mathcal{F}'(\Phi)\right) \tag{2.2.10}$$

are both dependent on the prepotential. The name Kähler potential derives from the fact that $\mathcal{F}''(\phi)$ defines a Kähler metric, an internal metric on the space of (scalar) fields, which can be seen from the expansion of the kinetic terms of the action (2.2.9) in components:

$$\begin{aligned} S_{kin} \sim \text{Im} \int \mathrm{d}x^4 \mathcal{F}''(\phi) F_{\mu\nu}\left(F^{\mu\nu} - i\tilde{F}^{\mu\nu}\right) + \mathcal{F}''(\phi)|\partial_\mu \phi|^2 \\ + i\mathcal{F}''(\phi)\lambda\sigma^\mu \partial_\mu \bar{\lambda} - i\mathcal{F}''(\phi)\psi\sigma^\mu\partial_\mu\bar{\psi}. \end{aligned} \tag{2.2.11}$$

The Kähler metric is related to the Kähler potential as:

$$\text{Im}\,\mathcal{F}''(\phi) = \text{Im}\,\frac{\partial^2 K(\phi, \phi^\dagger)}{\partial\phi\partial(\phi)^\dagger} \tag{2.2.12}$$

By virtue of $\mathcal{N} = 2$ supersymmetry the holomorphic prepotential enters in the Kähler metric. This will be an important ingredient in the exact solution of the low energy action.

We conclude this section by noting that pure $\mathcal{N} = 2$ SYM is a supersymmetric extension of the Georgi-Glashow model. As such, we expect it to contain 't Hooft-Polyakov monopoles after a Higgs condensation. Fermion zero modes in the $\mathcal{N} = 2$ vector multiplet act as supersymmetry generators on the monopoles and provide supersymmetric partners. Furthermore, due to the Witten effect dyons of arbitrary charge must appear in the spectrum. We will come back to these statements in the following chapter. For a review of the non-supersymmetric Georgi-Glashow model and the appearance of monopoles and dyons in the spectrum, we refer the reader to [5]. For a review of the supersymmetric case and the action of zero modes on the monopoles and dyons, see [65].

2.3 Coupling SYM to Matter

Next to the $\mathcal{N}=2$ vector multiplet there is one other non-gravitational $\mathcal{N}=2$ multiplet called the hypermultiplet, containing spins $s \leq \frac{1}{2}$. Whereas the vector multiplet generalizes the notion of the gauge bosons of a non-supersymmetric quantum field theory, the hypermultiplet represents supersymmetric matter. Accordingly, the hypermultiplets appear in the fundamental representation of the gauge group. More precisely, the hypermultiplet consists of two chiral superfields $Q, \tilde{Q}$ transforming in the fundamental and antifundamental representation of the gauge group respectively. Again, expansions of the fields are given in Appendix B.

For completeness and later reference we give the part of the $\mathcal{N}=2$ $SU(2)$ SYM action which is coupled to a hypermultiplet:

$$S = \int \mathrm{d}\theta^2 \mathrm{d}\bar{\theta}^2 \left(Q^\dagger e^{-2V} Q + \tilde{Q} e^{2V} \tilde{Q}^\dagger \right) + \int \mathrm{d}\theta^2 \left(\sqrt{2} \tilde{Q} \Phi Q + m \tilde{Q} Q \right) + h.c. \tag{2.3.1}$$

Q and $\tilde{Q}$ are to be read as N dimensional vectors for gauge group $SU(N)$. The coupling between Φ and $Q, \tilde{Q}$ is required by $\mathcal{N}=2$ supersymmetry. Using gauge indices, this Yukawa-like term is written as:

$$\tilde{Q}^c (\Phi^a T_a)_c{}^b Q_b$$

Note that we use lower indices for the fundamental representation and upper for the antifundamental representation. Furthermore, the action of T_a is now to be read as ordinary matrix multiplication on the vectors. Only couplings between several fields in the adjoint gives rise to the structure constant representation, or equivalently to commutator couplings. We will come back to the hypermultiplet in Section 3.8.

2.4 R-Symmetry and an Anomaly

The super Poincaré algebra with extended supersymmetry is invariant under unitary rotations of the supercharges. This symmetry is called R-symmetry. In superspace formalism this symmetry has a very natural interpretation: it rotates the various fermionic coordinates into each other. This is the fermionic analogue of the $SO(3,1)$ group of isometries on the bosonic coordinates. In this short outset, we mainly follow the analysis of [4] and consequently the formulas given here are taken (almost) exactly from that paper. We choose to repeat them for they motivate some important conclusions on which the next chapter will be built.

The R-symmetry group in the case of $\mathcal{N}=2$ supersymmetry is $U(2)_R$. It may be decomposed into $SU(2)_R \times U(1)_R$ of which the diagonal $U(1)_R$ does not mix the θ and $\tilde{\theta}$. The $U(1)_R$ action is defined on the chiral components as:

$$\begin{aligned} \theta, \tilde{\theta} &\to e^{i\alpha} \theta, \tilde{\theta} \\ \bar{\theta}, \bar{\tilde{\theta}} &\to e^{-i\alpha} \bar{\theta}, \bar{\tilde{\theta}} \end{aligned} \tag{2.4.1}$$

Under the $SU(2)_R$ symmetry θ and $\tilde{\theta}$ form a doublet. The subgroup $U(1)_J \subset SU(2)_R$, which is generated by the diagonal Pauli matrix σ_3, acts on the doublet as:

$$\begin{pmatrix} \theta \\ \tilde{\theta} \end{pmatrix} \to \begin{pmatrix} e^{i\alpha}\theta \\ e^{-i\alpha}\tilde{\theta} \end{pmatrix} \tag{2.4.2}$$

To find the transformation properties of the various superfields, we first notice that (2.4.1) implies:

$$\mathrm{d}^2\theta \to e^{-2i\alpha} \mathrm{d}^2\theta \tag{2.4.3}$$

The differential has opposite charge with respect to the $U(1)_R$ as compared to the coordinates themselves due the fact Grassmann integrations are defined as Grassmann derivations.

For the microscopic action (2.2.2) to be invariant under the $U(1)_R$ symmetry, the above mentioned transformations on the coordinates imply transformations on the superfields. More concretely, we see that the Kähler potential part of the Lagrangian is invariant under (2.4.1) and (2.4.3) if it simultaneously transforms as:

$$K(\theta) \to K(e^{-i\alpha}\theta) \tag{2.4.4}$$

On the other hand, chiral terms such as the gauge kinetic term and the superpotential in (2.3.1) should carry an R-charge 2 to be invariant:

$$G(\theta) \to e^{2i\alpha} G(e^{-i\alpha}\theta) \tag{2.4.5}$$

To satisfy these constraints in the case of the action (2.2.2) coupled to a hypermultiplet, the superfields should transform as:

$$\begin{aligned} U(1)_R: \quad & W_\alpha(\theta) \to e^{i\alpha} W_\alpha(e^{-i\alpha}\theta) \qquad && \Phi(\theta) \to e^{2i\alpha}\Phi(e^{-i\alpha}\theta) \\ & Q(\theta) \to Q(e^{-i\alpha}\theta) && \tilde{Q}(\theta) \to \tilde{Q}(e^{-i\alpha}\theta) \end{aligned} \tag{2.4.6}$$

Note that the R-charge of Φ is in principle not constrained by the transformation of the Kähler potential but chosen such that the (real) vector field does not transform under $U(1)_R$, as we will see just below. Fixing it at 2, the superpotential term requires Q and $\tilde{Q}$ to be neutral under $U(1)_R$. Also, a bare mass term for Q and $\tilde{Q}$ explicitly breaks the $U(1)_R$ symmetry.

Using the expansions of the superfields as given in Appendix B, the component fields transform as:

$$\begin{aligned} U(1)_R: \quad & \phi \to e^{2i\alpha}\phi \qquad && q \to q \\ & \chi \to e^{i\alpha}\chi && \psi_q \to e^{-i\alpha}\psi_q \\ & \lambda \to e^{i\alpha}\lambda && \tilde{q}^\dagger \to \tilde{q}^\dagger \\ & A_\mu \to A_\mu && \psi_{\tilde{q}}^\dagger \to e^{i\alpha}\psi_{\tilde{q}}^\dagger \end{aligned} \tag{2.4.7}$$

We consider the transformation properties of $\tilde{Q}^\dagger$ instead of $\tilde{Q}$ since it appears in the $SU(2)_R$ doublet with Q, while $\tilde{Q}$ sits in an $SU(2)_R$ doublet with $Q^\dagger$.

The action of $U(1)_J$ is more easily derived by looking at the transformations of the supercharges instead of the coordinates, as a sign of the fact we are using an $\mathcal{N}=1$ language while describing an $\mathcal{N}=2$ superspace symmetry. As reviewed in Appendix B, the Weyl spinors in the vector multiplet transform as a doublet under $SU(2)_R$, whereas the scalar and the vector field transform as singlets. For the hypermultiplet, the scalars transform in the doublet and the spinors as a singlet. Hence, we find:

$$\begin{aligned} U(1)_J: \quad & W_\alpha(\theta) \to e^{i\alpha} W_\alpha(e^{-i\alpha}\theta) \qquad && \Phi(\theta) \to \Phi(e^{-i\alpha}\theta) \\ & Q(\theta) \to e^{i\alpha} Q(e^{-i\alpha}\theta) && \tilde{Q}(\theta) \to e^{i\alpha}\tilde{Q}(e^{-i\alpha}\theta) \end{aligned} \tag{2.4.8}$$

such that the components transform according to their representation:

$$\begin{aligned} U(1)_J: \quad & \phi \to \phi \qquad && q \to e^{i\alpha} q \\ & \chi \to e^{-i\alpha}\chi && \psi_q \to \psi_q \\ & \lambda \to e^{i\alpha}\lambda && \tilde{q}^\dagger \to e^{-i\alpha}\tilde{q}^\dagger \\ & A_\mu \to A_\mu && \psi_{\tilde{q}}^\dagger \to \psi_{\tilde{q}}^\dagger \end{aligned} \tag{2.4.9}$$

We can form a Dirac spinor, which transforms properly under the Lorentz group, from the two Weyl spinors in the vector multiplet and hypermultiplet respectively as[?]:

$$\psi_D^v = \begin{pmatrix} \chi \\ \lambda \end{pmatrix},\ \psi_D^h = \begin{pmatrix} \psi_q \\ \psi_{\tilde{q}}^\dagger \end{pmatrix} \tag{2.4.10}$$

From the action on the component fields, we see that $U(1)_R$ acts as a chiral $U(1)$ and that $U(1)_J$ acts as a normal global phase transformation on the Dirac spinors.

Although classically the Lagrangian has all the above mentioned R-symmetries, quantum mechanically the chiral $U(1)_R$ symmetry will be broken. This anomaly goes under various names in the literature: the chiral anomaly, the triangle anomaly (after the problematic Feynman diagram) or the Adler-Bell-Jackiw anomaly (after its finders). In short, the ABJ anomaly states that in the presence of a background electromagnetic field the current associated to chiral rotations of the fermions is not conserved quantum mechanically. This may be directly computed from a triangle (loop) diagram (and assuming local gauge invariance) or by using a trick due to Fujikawa which shows the path integral measure changes in the presence of an instanton. For an extensive discussion, see Section 7.2 of [67].

The anomaly depends on the the number of chiral (or Weyl) fermions and their representations under the gauge group. We just state the result for $\mathcal{N} = 2$ supersymmetry, which contains two Weyl fermions in the adjoint representation, with gauge group $SU(N)$ and N_f hypermultiplets in the fundamental representation of the gauge group:

$$\partial^\mu j_\mu^5 = -\frac{2N - N_f}{16\pi^2} F_{\mu\nu}^a \tilde{F}_a^{\mu\nu} \tag{2.4.11}$$

The independence of the right hand side on the coupling constant shows it is a one-loop contribution.[2] It is an old result of Adler and Bardeen in [3] that higher order diagrams do not contribute to the anomaly. It is known that the right hand side represents a total derivative, which not necessarily integrates to zero due to the possibility of non-trivial gauge configurations at ∞. See for instance [19] or [54].

The non-conservation of a current implies the effective Lagrangian changes under the associated anomalous symmetry transformation, parametrized by an angle α, by:

$$\delta\mathcal{L}_{eff} = -\frac{\alpha(2N - N_f)}{16\pi^2} F_{\mu\nu}^a \tilde{F}_a^{\mu\nu} \tag{2.4.12}$$

Denoting ν as the instanton number, or Pontryagin index, the total action changes as:

$$\begin{aligned}\Delta S &= -\int \mathrm{d}^4x \frac{\alpha(2N - N_f)}{16\pi^2} F_{\mu\nu}^a \tilde{F}_a^{\mu\nu} \\ &= -2(2N - N_f)\nu\alpha\end{aligned} \tag{2.4.13}$$

In the case of gauge group $SU(2)$ with no added flavours, the shift will amount to $\Delta S = -8\nu\alpha$. For $\nu = 1$, this corresponds to a shift in the θ angle: $\theta \to \theta - 8\alpha$. Since physics is periodic in θ, chiral rotations with $\alpha = \frac{2\pi}{8}$ are a true symmetry of the action. Therefore, we have the following breakdown of R-symmetry:

$$SU(2)_R \times U(1)_R/\mathbb{Z}_2 \longrightarrow SU(2)_R \times \mathbb{Z}_8/\mathbb{Z}_2 \tag{2.4.14}$$

The division by $\mathbb{Z}_2$ has been performed because the (anomalous) $U(1)_R$ has a non-empty intersection with $U(1)_J$: a multiplication of the two spinors with a phase $e^{i\pi}$. Furthermore, a non-zero vacuum expectation value of the Higgs scalar in the vector multiplet will break this R-symmetry even further.

[2] The fields are scaled such that tree level is at g^{-2} while one loop is at g^0.

As will be argued in more detail in Section 3.1, the correct parameter to describe the moduli space of $\mathcal{N}=2$ SYM is:

$$u = \langle \mathrm{Tr}\,\phi^2 \rangle \tag{2.4.15}$$

Because the R-charge of ϕ is 2 a non-zero expectation value of ϕ^2 transforms $\alpha = \frac{2\pi n}{8}$, n odd, as:

$$\phi^2 \longrightarrow -\phi^2. \tag{2.4.16}$$

Therefore, the remaining R-symmetry is:

$$SU(2)_R \times \mathbb{Z}_4/\mathbb{Z}_2. \tag{2.4.17}$$

The relevance of knowledge about the remaining unbroken R-symmetry group will become clear when we will discuss Seiberg-Witten theory.

We conclude this section by noting that a priori, the θ angle may be put to zero in the partition function. This is achieved by absorbing the θ term in the fermion path integral measure. This amounts to a chiral rotation of the fermion fields, i.e. a redefinition of the fermions. Because the Atiyah-Singer index theorem precisely relates the difference of chiral fermion zero modes and the instanton number, this redefinition is consistent for all instanton configurations. However, we will see in the next section that for generic values of $\langle \mathrm{Tr}\,\phi^2 \rangle$ a non-zero θ enters the β function of g and renormalizes itself due to non-perturbative effects.

2.5 Exact Renormalization of the Prepotential

The exact form of the prepotential $\mathcal{F}(\Phi)$, as introduced in Section 2.2, was determined purely from symmetry considerations by Seiberg in [57]. It turns out that are no perturbative corrections to the prepotential apart from a one-loop correction. The non-perturbative corrections originate from (multi-)instanton contributions.[3] The values of the contributions will remain unknown, but the general form of the expansion is made precise. In this section we will reproduce the reasoning of the original article [57] that led to its form.

The one-loop beta function for the gauge coupling g of an $\mathcal{N}=2$ $SU(N)$ gauge theory with N_f fundamental and antifundamental flavours reads[61]:

$$\beta_{\frac{4\pi}{g^2}}(\mu) = \mu \frac{\mathrm{d}}{\mathrm{d}\mu}\left(\frac{4\pi}{g^2(\mu)}\right) = \frac{2N - N_f}{2\pi} \tag{2.5.1}$$

Notice the theory is UV free for $2N > N_f$. We will be concerned in first instance with the pure gauge theory.

We will first argue that the one-loop beta function is in fact the exact perturbative beta function. As mentioned in the previous section the anomaly of the chiral current is a total derivative which not necessarily vanishes due to instantons. Then, however, the low energy *perturbative* action should remain invariant under the anomalous $U(1)_R$. In Section 2.2 we gave the general form of the low energy $\mathcal{N}=2$ action, which we repeat for convenience:

$$S = \frac{1}{4\pi}\mathrm{Im}\left(\int \mathrm{d}x^4 \int \mathrm{d}^2\theta \mathrm{d}^2\tilde{\theta}\; \mathcal{F}(\Psi)\right) \tag{2.5.2}$$

We will look at this theory for scales $\mu < a$, effectively leaving us with a non-interacting massless abelian theory. If the perturbative $\mathcal{F}_{\mathrm{pert}}(\Psi)$ is to be invariant under $U(1)_R$, it should carry R-charge

[3] For all possible non-perturbative corrections, we only consider instantons. In fact, these will be the leading non-perturbative corrections in the case we are considering: no unbroken non-abelian gauge groups [58].

4 from considerations mentioned in the previous section. Therefore, $\mathcal{F}_{\text{pert}}(\Psi)$ should be proportional to Ψ^2. Seiberg then considered:

$$\mathcal{F}_{\text{pert}}(\Psi) = \Psi^2 \left[a_1 + a_2 \log\left(\frac{\Psi^2}{\mu_0^2}\right)\right] \tag{2.5.3}$$

which presents the only possible terms compatible with holomorphicity and the $U(1)_R$ symmetry as we will see soon. The bare coupling is defined at the UV cutoff μ_0. The constant a_1 may be seen from the microscopic theory to equal:

$$a_1 = \frac{1}{2}\tau_{\text{cl}}$$

with $\tau_{\text{cl}} = \tau(\mu_0)$ the bare complexified gauge coupling. Let us first analyse why this leads to an action invariant under $U(1)_R$. Clearly, it is proportional to Ψ^2. However, the prepotential *does* vary under $U(1)_R$:

$$\Delta\mathcal{F}_{\text{pert}}(\Psi) = 4i\alpha a_2 \Psi^2 \tag{2.5.4}$$

From (2.5.2) we see that the action then changes proportional to:

$$\Delta S_{\text{pert}} \sim \text{Im}\left(\int \mathrm{d}x^4 \int \mathrm{d}^2\theta \mathrm{d}^2\tilde{\theta}\ i\Psi^2\right) \tag{2.5.5}$$

$$\sim \int \mathrm{d}^4x\, F^a_{\mu\nu} F^{\mu\nu}_a \tag{2.5.6}$$

Ignoring non-perturbative effects, this indeed integrates to zero. However, the Lagrangian is not invariant under the R-symmetry transformation. Comparing with (2.4.12), we can see that a_2 is some numerical factor independent of the gauge coupling. The full perturbative beta function is now given by:

$$\mathcal{F}_{\text{pert}}(\Psi) = \frac{1}{2}\tau_{\text{cl}}\Psi^2 + a_2\Psi^2 \log\left(\frac{\Psi^2}{\mu_0^2}\right) \tag{2.5.7}$$

Since a_2 is independent of g, we conclude the second term is a one-loop effect. Thus, while demanding invariance of the perturbative action under $U(1)_R$, we find that the only possible contribution to the prepotential is a one-loop correction. We conclude that the one-loop beta function (2.5.1) is the full perturbative beta function. Integrating it from some UV cutoff μ_0 down to Higgs expectation value at which the coupling stops running, we find:

$$\tau(\mu_0) - \tau(a) = \frac{2i}{\pi}\log\left(\frac{\mu_0}{a}\right)$$

$$\Leftrightarrow\ \tau(a) = \tau(\mu_0) + \frac{i}{\pi}\log\left(\frac{a^2}{\mu_0^2}\right) \tag{2.5.8}$$

We can take the two terms together to obtain:

$$\tau(a) = \frac{i}{\pi}\log\left(\frac{a^2}{\Lambda^2}\right) \tag{2.5.9}$$

where we have defined $\Lambda \equiv \mu_0\, e^{2\pi i\tau(\mu_0)/4}$. Let us remark two things about Λ:

1. At values of the Higgs expectation value $a \sim \Lambda$ the gauge coupling g diverges signalling strong coupling and perturbation theory breaks down.
2. Λ^4 is cut-off independent. Indeed, using (2.5.1) with $N = 2$ and $N_f = 0$ we find:

$$\partial_{\mu_0}\Lambda^4 = 2\pi i e^{2\pi i\tau(\mu_0)}\mu_0^3\left(\mu_0\partial_{\mu_0}\tau(\mu_0) - \frac{2i}{\pi}\right) = 0. \tag{2.5.10}$$

   For this reason, Λ^4 is called the dynamically generated scale of the theory and characterizes the scale at which the coupling becomes strong.

Now that we know in what regime we can trust the above arguments, it is interesting to look at the components of $\tau(a)$. Writing $a = a_0 e^{ir}$ with r some real number, we identify:

$$\frac{4\pi}{g(a)} \equiv \mathrm{Im}\tau(a) = \frac{1}{\pi}\log\left(\frac{|a|^2}{\Lambda^2}\right) \tag{2.5.11}$$

$$\frac{\theta}{2\pi} \equiv \mathrm{Re}\tau(a) = -\frac{2r}{\pi} \tag{2.5.12}$$

The fact that $\theta = -4r$ is non-zero for certain values of a may be traced back to the anomaly. Indeed, since a has R-charge 2 we indeed retrieve that under a $U(1)_R$ transformation on a, θ changes exactly as anticipated in the previous section. It is in this sense that for generic values of a, or better u, a non-zero θ parameter exists. Indeed, this is an artefact from the complexification of the IR cut-off.

Now we are curious to learn what happens at strong coupling. As may be seen from (2.5.12), for $a \lesssim \Lambda$ the gauge coupling becomes negative. Hence, the perturbative form of the action cannot be the final answer if our theory is to be physical and non-perturbative corrections should become important. Although these should render the expression for $\tau(a)$ positive, the cancellation of infinities seems to imply we are expanding around the wrong vacuum. In fact, in the next chapter we will find a description for the strong coupling region which behaves more appropriately.

Again, we recall the fact that in the UV we may trust the form of τ_{pert}. Non-perturbative corrections therefore should vanish at large values of a, implying the corrections should be proportional to a^{-n} for n strictly positive. As the non-perturbative corrections should respect the remaining $\mathbb{Z}_8$, acting as $\mathbb{Z}_4$ on a, on the moduli space, the corrections should be of the form:

$$\left(\frac{\Lambda}{a}\right)^{4k} = e^{\frac{8\pi^2}{g(a)^2}k} \tag{2.5.13}$$

where the equality comes from the perturbative β function, showing the typical instanton corrections. To determine precisely how these contributions modify the prepotential, we note that also the prepotential should keep alive a $\mathbb{Z}_8$ symmetry. This determines $k = 1$ contribution to be proportional to Ψ^{-2}. Hence we expect:

$$\mathcal{F}_{\mathrm{non-pert}}(\Psi) = \sum_{i=1}^{\infty} c_k \left(\frac{\Lambda}{\Psi}\right)^{4k} \Psi^2 \tag{2.5.14}$$

As argued by Seiberg, higher perturbative corrections around the instanton will be irreconcilable with the remaining R-symmetry. Anti-instanton contributions to $\mathcal{F}(\Psi)$ would lead to terms proportional to positive powers of Ψ. These blow up at weak coupling and therefore cannot be generated. However, anti-instantons do generate similar terms for $\mathcal{F}(\bar{\Psi})$.

The above determines the explicit form of the prepotential[**?**]:

$$\mathcal{F}(\Psi) = \Psi^2 \frac{i}{2\pi}\log\left(\frac{\Psi^2}{\Lambda^2}\right) + \Psi^2 \sum_{i=1}^{\infty} c_k \left(\frac{\Lambda}{\Psi}\right)^{4k}. \tag{2.5.15}$$

Taking the derivative twice with respect to Ψ^2, ignoring constant terms and evaluating Ψ on the vacuum manifold we obtain:

$$\tau(a) = \frac{i}{\pi}\log\left(\frac{a^2}{\Lambda^2}\right) + \sum_{i=1}^{\infty} c_k \left(\frac{\Lambda}{a}\right)^{4k}. \tag{2.5.16}$$

To conclude, we have seen the severe constraints imposed by supersymmetry on the form of the general low energy effective pure $SU(2)$ $\mathcal{N} = 2$ SYM action. We have solved the theory exactly in the perturbative regime. However, non-perturbatively there remain unknowns such as the instanton coefficients in the prepotential. Seiberg and Witten were able to solve for these coefficients exactly and therefore have determined the exact form of the action of this theory. Not only is this an amazing result considering the complexity of this physical system, it turns out there is a wealth of physics and mathematics hidden in it. The next section will occupy us with a thorough analysis of the solution provided by Seiberg and Witten.

Chapter 3

Seiberg-Witten Theory

A fundamental achievement in the understanding of $\mathcal{N}=2$ supersymmetric theories was made in 1994 by Seiberg and Witten. They wrote two articles, in the first of which they gave an exact solution for the low energy effective Lagrangian of pure $\mathcal{N}=2$ SYM theory with gauge group $SU(2)$.

The approach of Seiberg and Witten has been to study the moduli space of the low energy theory, parametrized by the Higgs vacuum expectation value. It turns out the structure of the moduli space is rather non-trivial yet it *is* possible to study it very precisely using the power of the holomorphic formulation while relying on a few physical assumptions. The first four sections of this chapter will occupy us with determining the precise structure of the moduli space. The most important insight of Seiberg and Witten was to note that the moduli space of the physical theory is in one-to-one correspondence with the moduli space of a certain elliptic curve. What is more, physical constraints on the gauge coupling appear naturally in the elliptic curve description. Exploiting the equivalence carefully, one can perform a relatively simple analysis on the elliptic curve and calculate to arbitrary order the prepotential of the physical theory.

We will provide a thorough discussion of the Seiberg-Witten solution to the pure low energy effective action, focussing on the solution obtained via the elliptic curve construction. Having done so, we briefly discuss confinement of electric charge through monopole condensation. The fact that Seiberg and Witten found a quantitative description of this particular process resulting in confinement, albeit of abelian charges, should be considered as an important motivation for the study of supersymmetric theories.

In the last sections, we will discuss the generalization of the results of the pure theory. First, we consider the addition of matter in the form of hypermultiplets as done originally in [60]. New interesting phenomena will appear and the discussion places the pure case in a broader perspective. At last, we will discuss an M-theory construction of Seiberg-Witten curves, which provides a straightforward recipe for the construction of curves for a large class of $\mathcal{N}=2$ gauge theories.

The literature discussing the first (and sometimes second) paper of Seiberg and Witten is vast. This chapter borrows elements from many reviews. Of particular good help, next to the original articles [59] and [60], were [4], [15], [21], [33], [35] and [50].

3.1 The Classical Moduli Space

As already anticipated and explained in the previous chapter we are interested in the low energy effective action (2.2.7) which we repeat here for convenience:

$$S=\frac{1}{4\pi}\mathrm{Im}\int \mathrm{d}x^4\left[\int \mathrm{d}^2\theta\,\frac{1}{2}\mathcal{F}_{ab}(\Phi)W^{\alpha a}W^b_\alpha+\int \mathrm{d}^2\theta\,\mathrm{d}^2\bar\theta\,(\Phi^\dagger)^a\mathcal{F}_a(\Phi)\right] \tag{3.1.1}$$

The microscopic action (2.2.2) contains a scalar potential:

$$V = \frac{1}{g^2}\,\mathrm{Tr}\left[\phi^\dagger,\,\phi\right] \tag{3.1.2}$$

The vanishing of the scalar potential determines the vacuum of the theory. We see a non-vanishing expectation value for ϕ is allowed:

$$\phi_{\mathrm{vac}} = \frac{1}{2}\begin{pmatrix} a & 0 \\ 0 & -a \end{pmatrix} \Leftrightarrow \;\phi_{\mathrm{vac}} = \frac{1}{2}a\sigma_3 \tag{3.1.3}$$

where a may be any complex number. This is called a flat direction of the potential since there exists a continuous set of vacua, the moduli space of the theory, each of which gives rise to a (distinct) theory. It is the objective of the first four sections to understand the properties of the effective theories as we move around the moduli space. Let us first naively discuss a generic theory.

Non-zero values of a break the gauge symmetry spontaneously: $SU(2) \to U(1)$. The gauge bosons A^1_μ and A^2_μ obtain a mass $M_W = |a|$ while A^3_μ remains massless.[1] Because of $\mathcal{N} = 2$ supersymmetry, also the fermions and the scalar in the vector multiplets corresponding to A^1_μ and A^2_μ become massive. Similarly, the fermions and scalars in the same multiplet as A^3_μ remain massless. Therefore, for scales $\mu < a$ the W-bosons freeze out and an abelian gauge theory remains. As already mentioned at the end of Section 2.2, this theory is non-interacting. Since non-interacting massless theories are conformal, at $\mu = a$ a fixed point of the renormalization group flow is reached. See Figure 3.1.1. We anticipate that for values small values of $a \sim \Lambda$ our Lagrangian analysis may

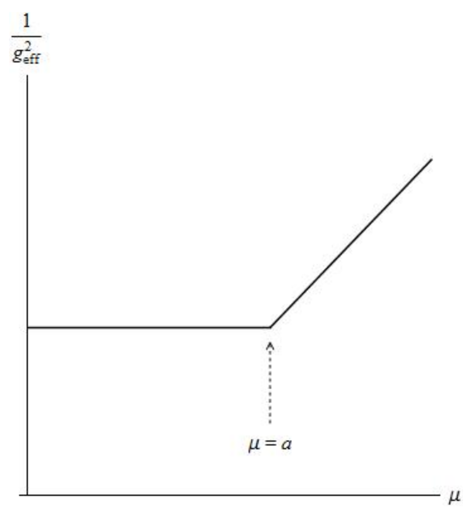

Figure 3.1.1: Schematic running of the coupling constant.

not be adequate, since the theory has been strongly coupled during the process of integrating out massive modes.

The form of the low energy effective action simplifies to:

$$S = \frac{1}{4\pi}\mathrm{Im}\int \mathrm{d}x^4\left[\left[\int \mathrm{d}^2\theta\,\frac{1}{2}\frac{\partial^2\mathcal{F}(\Phi)}{\partial\Phi^2}W^\alpha W_\alpha + \int \mathrm{d}^2\theta\,\mathrm{d}^2\bar\theta\,\Phi^\dagger\frac{\partial\mathcal{F}(\Phi)}{\partial\Phi}\right]\right. \tag{3.1.4}$$

The metric on the moduli space is given by the Kähler metric as introduced in Section 2.2:

$$g \equiv \frac{\partial^2 K(a,\bar a)}{\partial a\partial\bar a} = \mathrm{Im}\,\mathcal{F}''(a) \;\Rightarrow\; \mathrm{d}s^2 = \mathrm{Im}\,\mathcal{F}''(a)\,\mathrm{d}a\,\mathrm{d}\bar a. \tag{3.1.5}$$

[1] We use a normalization such that the W-bosons have electric charge 1.

where the Higgs expectation value plays the role of a local coordinate. In fact, a does not yet provide a gauge invariant coordinate. This is due to a discrete residual gauge symmetry, the Weyl group of $SU(2)$, which is defined as:

$$W(SU(2)) = N(U(1))/U(1)$$

with $N(U(1))$ the normalizer of the $U(1) \subset SU(2)$ subgroup. It is abstractly isomorphic to $\mathbb{Z}_2$ and acts on $U(1)$ essentially through complex conjugation. This implies its action on ϕ_{vac}:

$$\phi_{\text{vac}} \to -\phi_{\text{vac}}. \tag{3.1.6}$$

Geometrically, it may be understood as a rotation of an angle π around either the σ_1 or σ_2 direction. Clearly, a modulus which does provide a gauge invariant description of the moduli space is given by:

$$u = \langle \operatorname{Tr} \phi^2 \rangle \approx \frac{1}{2} a^2 \tag{3.1.7}$$

where the approximate sign is to remind us that we only trust our Lagrangian description at large values of u or a. The metric changes accordingly:

$$\mathrm{d}s^2 = \operatorname{Im} \mathcal{F}''(u) \frac{\mathrm{d}a}{\mathrm{d}u} \frac{\mathrm{d}\bar{a}}{\mathrm{d}\bar{u}} \mathrm{d}u \, \mathrm{d}\bar{u}. \tag{3.1.8}$$

As a short remark on notation, in the remainder of this thesis we will sometimes still write the prepotential as a function of a or Φ. The latter notation will generally be used when considering our theory above the vacuum, whereas the former is a slight abuse of notation.

From Section 2.5 we know the general the form of the prepotential:

$$\mathcal{F}''(u) = \frac{i}{\pi} \log\left(\frac{2u}{\Lambda^2}\right) + \sum_{i=1}^{\infty} c_k \left(\frac{\Lambda^2}{2u}\right)^{2k}. \tag{3.1.9}$$

Although explicit instanton calculations in [57] have shown that $c_1 \neq 0$, there is a more compelling argument why the perturbative part cannot be the only contribution to the prepotential. From a mathematical perspective, a Kähler metric is in particular a Riemannian metric. This means the metric is positive definite. From a physical point of view, the positivity of the gauge coupling, $\frac{4\pi}{g(u)^2} = \operatorname{Im} \mathcal{F}''(u)$ implies the metric should be positive. We thus have the following requirement on the metric for the theory to make sense:

$$\operatorname{Im} \mathcal{F}''(u) = \operatorname{Im} \tau(u) > 0. \tag{3.1.10}$$

Since $\tau(u)$ is holomorphic in u, the metric is harmonic. Harmonic functions satisfy a minimum and maximum principle. That is, they do not acquire a maximum or minimum at any point in the interior of the domain they are defined in, except when they are constant. Clearly, $\operatorname{Im} \tau(u)$ is not a constant function. Therefore, we see that the positivity requirement leads to a contradiction. Not giving up on the holomorphy of the prepotential, which would imply a breaking of supersymmetry, we conclude that the perturbative prepotential cannot be defined globally on the moduli space. We *do* see that at large values of $u \gg \Lambda^2$ the metric is positive and single valued:

$$\operatorname{Im} \tau(u) \sim \frac{1}{\pi} \log\left(\frac{|u|}{\Lambda^2}\right) \tag{3.1.11}$$

The fact that the gauge coupling is well behaved in this regime should not be surprising since our perturbative analysis applies and non-perturbative contributions should be unimportant.

To advance our understanding of strongly coupled regime of the moduli space any further, we need a new tool next to the already proven powerful tool of holomorphy. This turns out to be electric-magnetic duality.

3.2 Duality of the Super Yang-Mills Action

In this section we will show the low energy effective action of $\mathcal{N}=2$ SYM can be written in terms of different local fields which allow a weakly coupled description in the region of the moduli space where the original theory is strongly coupled. In Section 3.4 we will present the motivation of Seiberg and Witten that the degrees of freedom which become important at scales $u \lesssim \Lambda^2$ are the 't Hooft-Polyakov monopoles and their associated dyons. Before turning to their arguments we first show it is possible to formulate an electric-magnetically dual description of the low energy effective action rather similar to the dualization of Maxwell theory. In this case the dual description is more natural since we do not have to introduce monopoles to our theory.

Because the dualization of SYM is analogous to the Maxwell dualization presented in Appendix A, we point out the differences. The first difference is that in SYM we not only have an F^2 term, but also an $F\tilde{F}$ term. However, this term is a total derivative and therefore will not affect the Gaussian functional integration. We simply obtain:

$$\mathrm{Im}\int \tau(a)\left(F^2 + i\tilde{F}F\right) \longrightarrow \mathrm{Im}\int \tau_D(a_D)\left(F_D^2 + i\tilde{F}_D F_D\right). \tag{3.2.1}$$

where $\tau_D(a_D) \equiv \frac{-1}{\tau(a)}$. The real difference between the pure SYM and Maxwell theory is that in SYM we have to integrate out an entire $\mathcal{N}=1$ vector multiplet.[2] The integrating out of a supermultiplet is conveniently performed in $\mathcal{N}=1$ superfield formulation. The gauge kinetic part of the low energy effective action reads:

$$S = \frac{1}{4\pi}\int \mathrm{d}x^4 \left(\mathrm{Im}\int \mathrm{d}^2\theta\, \frac{1}{2}\mathcal{F}''(\Phi)W^\alpha W_\alpha\right). \tag{3.2.2}$$

To perform the duality transformation we impose the Bianchi identity $\mathrm{Im}\mathcal{D}_\alpha W^\alpha = 0$ by a Lagrange multiplier in the action:

$$\Delta S = \frac{1}{4\pi}\,\mathrm{Im}\int \mathrm{d}x^4 \left(\int \mathrm{d}^2\theta \mathrm{d}^2\bar{\theta}\, V_D\, \mathcal{D}_\alpha W^\alpha\right) \tag{3.2.3}$$

Considering the real vector superfield V_D as a dynamical object, we integrate (3.2.3) by parts. Adding this to (3.2.2), completing the square and performing the Gaussian functional integral over the unconstrained W_α, we arrive at the dual description:

$$S = \frac{1}{8\pi}\int \mathrm{d}x^4 \left(\mathrm{Im}\int \mathrm{d}^2\theta \frac{-1}{\tau(\Phi)} W_D^\alpha W_{D\,\alpha}\right) \tag{3.2.4}$$

There are some subtleties concerning this duality transformation. For a more careful analysis, see [15]. We conclude that we have achieved a strong-weak coupling duality transformation in which we have exchanged electric and magnetic variables, since V_D naturally couples to magnetic charges.[3] We have to hold our horses for the moment though, since we should also consider what happens to the chiral multiplet Φ. This may be seen by looking at the transformed gauge coupling:

$$\tau_D(\Phi_D) = \frac{-1}{\tau(\Phi)} \tag{3.2.5}$$

This equation defines Φ_D in terms of Φ. We can make this definition more concrete by using the expression of the gauge coupling in terms of the prepotential:

$$\frac{\partial \mathcal{F}_D'(\Phi_D)}{\partial \Phi_D} = -\left(\frac{\partial \mathcal{F}'(\Phi)}{\partial \Phi}\right)^{-1} \tag{3.2.6}$$

[2]The Kähler potential already is in a duality invariant form as we will see soon. This should come as no real surprise since the low energy Kähler potential does not contain gauge interactions and is neutral under the remaining $U(1)$ gauge group.

[3]Notice that this strong weak coupling transformation does not precisely invert the gauge coupling. The interpretation of strong-weak coupling duality clearly works for small values of $\theta \mod 2\pi$. When this is not the case, strong-weak duality relations will generally be more complicated.

From this we identify $\mathcal{F}'_D(\Phi_D) = -\Phi$ and $\mathcal{F}'(\Phi) = \Phi_D$. In terms of these dual variables, the Kähler potential reads:

$$\text{Im} \int \mathrm{d}^2\theta\, \mathrm{d}^2\bar{\theta}\; \bar{\Phi}\, \mathcal{F}'(\Phi) = \text{Im} \int \mathrm{d}^2\theta\, \mathrm{d}^2\bar{\theta}\; \bar{\Phi}_D\, \mathcal{F}'_D(\Phi_D) \tag{3.2.7}$$

As expected, it retains the exact same form as it had originally.[4] We conclude that the $\mathcal{N} = 2$ SYM low energy effective action is duality invariant. We summarize the duality transformations:

$$\begin{aligned} \tau(a) &\mapsto \frac{-1}{\tau(a)} \equiv \tau_D(a_D) & \qquad \Phi &= -\mathcal{F}'_D(\Phi_D) \\ W^\alpha &\leftrightarrow W^\alpha_D & \qquad \mathcal{F}'(\Phi) &= \Phi_D \end{aligned} \tag{3.2.8}$$

The reason we use different arrows and equality signs for the different lines is because these transformations are quite different. First of all, the gauge coupling really is mapped onto its inverse. The double arrow for the gauge superfields signifies the non-locality of the transformation. The chiral superfield undergoes a redefinition only.

There is another transformation on the fields, which unlike the transformation above, is a real symmetry of the action: a shift of the θ angle by 2π. This is implemented via:

$$\tau(a) \mapsto \tau(a) + 1. \tag{3.2.9}$$

Together with (3.2.8), these transformations generate the full duality group $SL(2,\mathbb{Z})$. A general $SL(2,\mathbb{Z})$ transformation then acts on τ as:

$$\tau(a) \mapsto \frac{a\tau + b}{c\tau + d}. \tag{3.2.10}$$

Using (3.2.6), we may write:

$$\tau(a) = \frac{\mathrm{d}a_D}{\mathrm{d}a} \tag{3.2.11}$$

We see that (3.2.10) implies an $SL(2,\mathbb{Z})$ action on the chiral superfields:[5]

$$\begin{pmatrix} a_D \\ a \end{pmatrix} \mapsto \begin{pmatrix} a_D(d) \\ a(d) \end{pmatrix} = \begin{pmatrix} a\, a_D + b\, a \\ c\, a_D + d\, a \end{pmatrix} \tag{3.2.12}$$

In mathematical terms, the vector field $(a_D(u), a(u))$ represents a section of a flat $SL(2,\mathbb{Z})$ bundle over the u plane. The transformation, translated to the superfields Φ and Φ_D, does not affect Kähler potential as one may check directly. Let us conclude this section with some remarks:

1. There exists an infinite number of dual forms of the low energy action, generated by $SL(2,\mathbb{Z})$ transformations on the gauge coupling while simultaneously changing the chiral superfields as in (3.2.12). It means the theory retains its precise form and therefore its precise dynamics in terms of different variables $a(d)$ and a different coupling $\tau_d(a(d))$.[6] We stress electric-magnetic duality is a property of the IR Lagrangian.

2. From (3.2.11) it is clear there is the additional freedom of adding a constant to Φ_D or Φ if we take the equation as a definition of the gauge coupling. This extension of $SL(2,\mathbb{Z})$ naturally appears when we couple the pure theory to hypermultiplets. For the pure case, this transformation is not consistent with properties of the central charge as we will see in the next section.

[4] Notice that for $w, z \in \mathbb{C}$: $\text{Im}\, \bar{w}z = -\text{Im}\, w\bar{z}$.

[5] In a particular dual representation we will use d to label the representation whereas the subscript D is used to denote the prepotential term $\frac{\partial \mathcal{F}_d(a(d))}{\partial a(d)}$.

[6] This statement perhaps seems vacuous since the theory considered is a pure non-interacting abelian gauge theory. However, we will find out that in the strongly coupled region the theory will become interacting and we do need a dual $U(1)$ gauge multiplet.

3. We can rewrite the metric on the moduli space (3.1.8) in terms of the new coordinates as:

$$\mathrm{d}s^2 = \mathrm{Im}\,\frac{\mathrm{d}a_D}{\mathrm{d}u}\frac{\mathrm{d}\bar{a}}{\mathrm{d}\bar{u}}\mathrm{d}u\,\mathrm{d}\bar{u} = -\frac{i}{2}\left(\frac{\mathrm{d}a_D}{\mathrm{d}u}\frac{\mathrm{d}\bar{a}}{\mathrm{d}\bar{u}} - \frac{\mathrm{d}a}{\mathrm{d}u}\frac{\mathrm{d}\bar{a}_D}{\mathrm{d}\bar{u}}\right)\mathrm{d}u\,\mathrm{d}\bar{u}. \tag{3.2.13}$$

This expression is manifestly duality invariant.

3.3 Central Charge and Dual Theories

In this section we will analyse the duality of the $\mathcal{N} = 2$ SYM action further. First of all, we did not mention how the collective excitations, the monopole and the dyons, appear in dual descriptions and how they relate to a certain $SL(2,\mathbb{Z})$ transformation on the vector multiplet. The designated object to study this is the central charge. As a short intermezzo, we will briefly introduce it.

The central charge was calculated for the case of $\mathcal{N} = 2$ SYM in [73] and for the microscopic theory found to be:

$$Z = a\left(n_e + \tau_{\mathrm{cl}} n_m\right) \tag{3.3.1}$$

The absolute value of the central charge determines the mass of a BPS saturated supermultiplet as:

$$m = \sqrt{2}|Z| \tag{3.3.2}$$

Although we have already seen the functions $a(u)$ and $\tau(a)$ receive perturbative and non-perturbative corrections when considering the full theory, it is believed that the masses of the fields in the low energy effective action still satisfy a BPS bound, albeit a renormalized version. One good reason to believe this statement is that a priori the superfields in the UV action (2.2.2) are massless and therefore belong to a short supersymmetry multiplet. Since the Higgs mechanism does not generate new degrees of freedom, only a redistribution of them, the massive fields should be BPS saturated. Furthermore, it is assumed that also quantum corrections do not generate enough degrees of freedom such that the effective fields will belong to long supermultiplets. Therefore, we expect all states in the full low energy theory still to be BPS saturated. The low energy version of the BPS bound reads:

$$Z = \begin{pmatrix} n_m & n_e \end{pmatrix}\begin{pmatrix} a_D \\ a \end{pmatrix} = a n_e + a_D n_m \tag{3.3.3}$$

To see why is a rather simple argument, presented in [59]. In short, when you couple a hypermultiplet to the low energy effective action, you can read of the mass of the hypermultiplet by looking at the superpotential term. For a hypermultiplet with electric quantum number n_e this turns out to be:

$$m = \sqrt{2} a n_e \tag{3.3.4}$$

From which we read off the central charge as $Z = an_e$ (up to a phase).[7] After dualizing the theory we have a magnetic chiral superfield $\Phi(d) = \Phi_D$ coupling to a magnetically charged hypermultiplet. This corresponds to a central charge: $Z = a_D n_m$. Taking these together for a general dyon leads to (3.3.3). We note that the stable dyon solutions are such that $\gcd(n_m, n_e) = \pm 1$[59].

The physical relevance of the central charge is clear. It provides us with the particle spectrum of BPS saturated states which may appear in our theory at a given value of u. For fixed u the spectrum should remain invariant under duality transformations. The only way for the central charge to remain invariant under a duality transformation is when we choose a different set of quantum numbers:

$$\begin{pmatrix} a_D(d) \\ a(d) \end{pmatrix} = M\begin{pmatrix} a_D \\ a \end{pmatrix} \Rightarrow \begin{pmatrix} n_m(d) & n_e(d) \end{pmatrix} = \begin{pmatrix} n_m & n_e \end{pmatrix} M^{-1},\; M \in SL(2,\mathbb{Z}). \tag{3.3.5}$$

[7] In our present normalization of the electric charge, an electric hypermultiplet will have $n_e = \frac{1}{2}$, but dyons will have $n_e = 1$. When we add matter, i.e. electric hypermultiplets, we will change the normalization such that all charges are integer valued. For now, we follow the conventions of [59].

Suppose now that we are interested in a local Lagrangian description of a dyon of charge (n_m, n_e) with central charge as in (3.3.3). To couple this field locally we use a dual vector multiplet to which the dyon appears with unit electric charge $(n_m(d), n_e(d)) = (0, 1)$. It may be checked that the correct $SL(2,\mathbb{Z})$ transformation is given by:

$$\begin{pmatrix} n_m(d) & n_e(d) \end{pmatrix} = \begin{pmatrix} n_m & n_e \end{pmatrix} \begin{pmatrix} \frac{1+bn_m}{n_e} & b \\ n_m & n_e \end{pmatrix}^{-1} \tag{3.3.6}$$

Here, $b \in \mathbb{Z}$ is a parameter such that the upper left entry is an integer. Note this is only possible when n_m and n_e are coprime. If $k = \gcd(n_m, n_e) > 1$, the dual charge vector will be of charge $(0, k)$. The transformation implies the following dual Higgs field:

$$\begin{pmatrix} a_D(d) \\ a(d) \end{pmatrix} = \begin{pmatrix} \frac{1+bn_m}{n_e} & b \\ n_m & n_e \end{pmatrix} \begin{pmatrix} a_D \\ a \end{pmatrix} \tag{3.3.7}$$

Lastly, the gauge coupling is given by:

$$\tau_d(a(d)) = \frac{\mathrm{d}a_D(d)}{\mathrm{d}a(d)} = \frac{\frac{1+bn_m}{n_e}\tau(a) + b}{n_m\tau(a) + n_e} \tag{3.3.8}$$

Of course, if we find an $SL(2,\mathbb{Z})$ transformation changing a strong gauge coupling into a weak gauge coupling, this procedure can be reversed to find the appropriate quantum numbers of the weakly coupled field.

The fundamental fields in which the monopoles and dyons appear locally in a Lagrangian description are hypermultiplets.[8] This is because magnetic monopoles are charged objects and therefore enter as matter for the (unbroken) gauge boson. Hence, at a dual magnetic or dyonic description we obtain a $U(1)$ theory coupled to a hypermultiplet.[9] This is nothing but supersymmetric QED. This theory behaves fundamentally different from the previous spontaneously broken non-abelian theory. The difference is that QED is IR free as opposed to UV free. The beta function reads[39]:

$$\mu\frac{\mathrm{d}}{\mathrm{d}\mu}g_d(\mu) = \frac{g_d(\mu)^3}{8\pi^2} \tag{3.3.9}$$

Note that, just as the perturbative beta function determined in Section 2.5 was exact at $a(u) = \infty$ due to asymptotic freedom, this beta function is exact at a point where a dyon becomes massless, i.e. $a(d)(u) = 0$, because QED is IR free. In the next section we will justify the assumption that dyons become massless in the strong coupling region of the moduli space. For now, we will just assume that there is some point u_0 in the moduli space at which a dyon becomes massless.

Complexifying the gauge coupling in the usual way (3.3.9) is equivalent to:

$$\mu\frac{\mathrm{d}}{\mathrm{d}\mu}\tau_d(\mu) = -\frac{i}{\pi} \tag{3.3.10}$$

This can be integrated to give:

$$\tau_d(a(d)) = -\frac{i}{\pi}\log\left(\frac{a(d)}{\Lambda_{\mathrm{QED}}}\right) \tag{3.3.11}$$

with Λ_{QED} some UV cutoff at which the inverse QED gauge coupling vanishes. We will take this cut-off to be the dynamically generated scale Λ which is an appropriate cut-off in the sense that

[8] As mentioned in Section 2.2, supersymmetric partners for the 't Hooft-Polyakov monopoles and dyons arise from the semiclassical quantization of fermion zero modes of the gauginos in the vector multiplet.

[9] We silently assume the monopoles are still present in our low energy effective action and could appear in a dual description. This is actually a false assumption and the observation that they have been integrated out is one of the key observations in solving the theory.

$a(d)(u_0) \ll \Lambda$, i.e. QED is weakly coupled. We may integrate this expression again to find a relation between $a_D(d)$ and $a(d)$:

$$a_D(d) = c - \frac{i}{\pi} a(d) \log\left(\frac{a(d)}{\Lambda}\right) + \frac{i}{\pi} a(d) \tag{3.3.12}$$

with c a non-zero constant. We can integrate this expression with respect to $a(d)$ to find the dual prepotential. Hence, assuming there is a point in the moduli space where a dyon becomes massless, we again are able to solve for the theory in that limit exactly (up to the constant c).

We have now set up all necessary equipment to return to the analysis of the moduli space. Summarizing, we have shown the low energy effective action of $\mathcal{N} = 2$ SYM is duality invariant. Furthermore, we have calculated the exact gauge coupling for a theory with a massless dyon. Let us now see how these tools can be used to the determine the precise moduli space of the theory.

3.4 Monodromies and the Quantum Moduli Space

Thus far, we have obtained a one-complex dimensional moduli space parametrized by the gauge invariant parameter u. This moduli space is endowed with a Kähler metric which depends on the low energy information of the physical theory and is dictated by the functions $(a_D(u), a(u))$. As already anticipated, the classical or perturbative moduli space is not expected to be the real moduli space of the theory as for one, it would imply a negative gauge coupling in the strong coupling region $u \sim \Lambda^2$. In this section, we will analyse the singular behaviour, in the form of monodromies, of the functions $(a_D(u), a(u))$. From this analysis we will be able to determine the real or quantum moduli space of the theory precisely. Before turning to this, let us first describe where the monodromies originate from.

Since the Wilsonian effective action is obtained with an IR as well as a UV cutoff, singular behaviour of the prepotential does not originate from typical UV or IR divergences. To understand where they do come from, we note the following. After the Higgs has condensed all elementary particles and collective excitations obtain a mass dictated by the BPS bound, except for the unbroken $U(1)$ vector multiplet. Having integrated out the heavy degrees of freedom, we obtain an effective Lagrangian at the scale of the Higgs expectation value $\mu = a(u)$ at which the coupling stops running. This is a divergence free Lagrangian which is holomorphic in u. In this sense, even after integrating out the modes, we may still vary u. If we vary u such that it becomes less than the scale above which everything is integrated out, we would still have a sensible theory although we could be erroneously ignoring some light degrees of freedom. However, if it turns out that for some values of the modulus u particles have become massless, the Wilsonian effective action will exhibit singular behaviour, stemming from the fact that massless particles have been integrated out.

We now turn to the singular behaviour. The asymptotic form of the section $(a_D(u), a(u))$ at $u \to \infty$ was derived in Section 2.5:

$$a_D(u) = \sqrt{2u}\frac{i}{\pi} \log\left(\frac{2u}{\Lambda^2}\right) + \frac{i\sqrt{2u}}{\pi} \tag{3.4.1}$$

$$a(u) = \sqrt{2u} \tag{3.4.2}$$

From the formula of the central charge it follows directly that for $u \to \infty$, the lightest particles are the W-bosons. All particles with a magnetic charge, like the 't Hooft-Polyakov monopoles and the dyons, will be more massive. The proper action is then indeed the non-interacting $U(1)$ gauge theory, with respect to which the electrically charged W-bosons, before having been integrated out, would describe the dominant interactions.

The fact that we are considering a correct description of the theory can also be seen from monodromy properties of the section $(a_D(u), a(u))$. That is; upon circling $u = \infty$, i.e. $u \mapsto e^{2\pi i}u$, the section transforms as:

$$a_D(u) \to -\sqrt{2u}\frac{i}{\pi}\log\left(\frac{2u}{\Lambda^2}\right) + 2\sqrt{2u} - \frac{i\sqrt{2u}}{\pi} = -a_D + 2a \quad (3.4.3)$$

$$a(u) \to -\sqrt{2u} = -a \quad (3.4.4)$$

The functions do not return to their original value; they are said to have a monodromy. The monodromy can be phrased in matrix notation as:

$$\begin{pmatrix} a_D \\ a \end{pmatrix} \mapsto \begin{pmatrix} -1 & 2 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} a_D \\ a \end{pmatrix} = M_\infty \begin{pmatrix} a_D \\ a \end{pmatrix} \quad (3.4.5)$$

Note that $M_\infty \in SL(2,\mathbb{Z})$ and acts on τ as a shift in the θ parameter. Therefore the SYM action remains invariant.

The central charge is kept invariant under the monodromy if we change the quantum numbers of our states accordingly. In fact, this is as expected and consistent with the discussion in Section 2.4, namely that the transformation $u \mapsto e^{2\pi i}u$ can be interpreted as a certain shift of the θ angle. Due

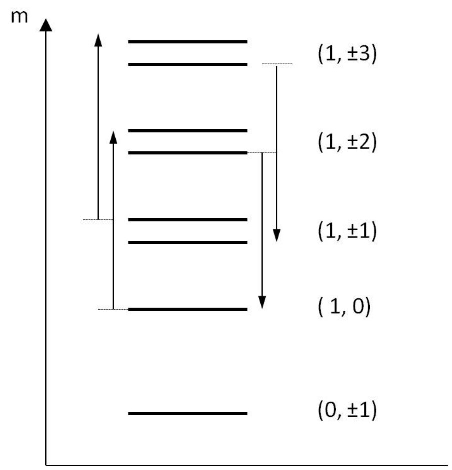

Figure 3.4.1: The spectrum of the theory for $u \to \infty$. The W-bosons are the lightest particles with degenerate masses, which are exchanged under the monodromy M_∞. The other particles are dyons of charge $(1, n)$. We suppress a negative magnetic unit charge, but its monopole and dyons may easily be seen to map exactly on the spectrum of positive magnetic unit charge. The arrows indicate the action of the monodromy on the spectrum. The monodromy effectively acts as: $(\pm 1, n) \to (\mp 1, n \pm 2)$. The spectrum as a whole remains invariant, however the quantum numbers of states of certain mass change. This figure is inspired on a figure in [50].

to the Witten effect the electric charge of dyons changes proportionally to shifts of θ angle, as was originally shown in [71]. The action of the monodromy on the spectrum is depicted in Figure 3.4.1.

An unambiguous description of the theory at $u = \infty$ is therefore only possible when the fundamental variables do not carry any magnetic charge. In this sense, the charged W-bosons[10] are again seen to provide the most convenient variables to describe the low energy theory at large u. Note that although the monodromy of $a(u)$ exchanges the W-bosons it does not change physical

[10] Or better: the free electric $U(1)$ vector multiplet, since the W-bosons are not present in the low energy description.

observables as the mass of the W-bosons. In fact, this monodromy is merely an artefact from our construction of the u plane and would disappear upon considering instead the a plane. Thus, in spite of the discussion at the beginning of the section M_∞ does not stem from the appearance of massless particles in the spectrum. It is simply a manifestation of the Witten effect.

It is well known that a monodromy at a certain point on the Riemann sphere implies the existence of at least one other singularity; otherwise one could "pull the loop over the sphere" and find a zero monodromy. We expect this additional singular point somewhere in the strong coupling region $u \sim \Lambda^2$, as the weak coupling Lagrangian analysis shows quite clearly none of the known particles become massless in the large u region. A naive guess would then be the possibility of a point u_0 where $a(u_0) = 0$, since this would correspond to a point at which the W-bosons become massless. However, we have already given an argument why such a point cannot, at least, be the only extra singularity on the moduli space. Namely, suppose it would be the only extra singularity. Then, the monodromy matrix around $u = 0$ would have to be the same as the monodromy at $u = \infty$. But this implies that $a(u) \sim \sqrt{u}$ is a good coordinate at weak *and* strong coupling in the sense it does not show any (non-trivial) monodromies. But this means the perturbative gauge coupling also provides a good metric at strong coupling. As we have discussed, this leads to the unphysical conclusion that the gauge coupling becomes negative at strong coupling.

The minimal ansatz is then that there are three singularities. From Section 2.4 we know that the unbroken subgroup $H \subset U(1)_R$ acts on u as a $\mathbb{Z}_2$ symmetry. Hence, the two singularities should be related: $u = \pm u_0$. Seiberg and Witten reasoned that these singularities are unlikely to be caused by massless W-bosons. There is a rather technical argument why this cannot be the case, which is given in Section 5.2 of the original paper [59]. We just note it seems unnatural that the W-bosons, having equal mass formulas, would become massless at different points in the moduli space. We will in fact see from the explicit solution that $a = 0$ does not belong to the quantum moduli space.

The only other particles that remain to cause the singularities are the 't Hooft-Polyakov monopoles and the dyons. At the time of the original paper, the assumption that a monopole and a dyon become massless at $u = \pm u_0$ worked out consistently and this was taken as evidence the assumption was in fact the correct one. Not much later though, it was proven in [53] that, assuming supersymmetry is not broken and the number of singularities on the moduli space is finite, the Seiberg-Witten solution is the unique solution to the low energy effective action of $\mathcal{N} = 2$ SYM.

Let us first suppose a monopole becomes massless at $u_0 = \Lambda^2$. Assuming the mass of the monopole varies continuously over the moduli space, in a neighbourhood of u_0 the monopole will be the lightest particle. Our description in terms of a free electric $U(1)$ theory will not be accurate at all. Instead, we should dualize the theory and couple it to a hypermultiplet. The right variables for a magnetic $U(1)$ vector multiplet are:

$$\begin{pmatrix} a_D(d) \\ a(d) \end{pmatrix} = \begin{pmatrix} a \\ -a_D \end{pmatrix}, \quad \tau_d(a(d)) = \frac{\mathrm{d}a_D(d)}{\mathrm{d}a(d)} = -\frac{1}{\tau(a)}. \tag{3.4.6}$$

From the beta function of QED we also obtained:

$$a_D(d) = c - \frac{i}{\pi} a(d) \log\left(\frac{a(d)}{\Lambda}\right) + \frac{i}{\pi} a(d) \tag{3.4.7}$$

Since the mass of monopole in the dual variables is given by $m = \sqrt{2}|a(d)|$, near the point $u = u_0$ at which the monopole becomes massless $a(d)$ should be of the form:[11]

$$a(d) = c(u - u_0) \tag{3.4.8}$$

From (3.4.7) and (3.4.8) we can read off the monodromy at u_0. For $(u - u_0) \mapsto e^{2\pi i}(u - u_0)$ we have:

$$\begin{pmatrix} a_D(d) \\ a(d) \end{pmatrix} \mapsto \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a_D(d) \\ a(d) \end{pmatrix} \tag{3.4.9}$$

[11]The assumption of a simple zero is the only choice consistent with the fact that the monodromies should be in $SL(2, \mathbb{Z})$ and the form of M_∞.

The monodromy in terms of the original coordinates (a_D, a) is obtained by a change of basis as dictated in this particular case by the S transformation:

$$M_{u_0} = S^{-1}\begin{pmatrix}1 & 2\\ 0 & 1\end{pmatrix}S = \begin{pmatrix}1 & 0\\ -2 & 1\end{pmatrix} \tag{3.4.10}$$

Relating to the terminology we used to interpret M_∞, we see that in this region of the moduli space a_D is the right coordinate to describe physics and we should read $a = \mathcal{F}'_D(a_D)$. Note that this depends crucially on the nonvanishing integration constant c in (3.4.7). Otherwise, all particles would have become massless at u_0 and we would have trouble to find a suitable description of the theory.

The question now is what this assumption means for the third singularity. There is an obvious way to find the monodromy at $u = -u_0$. For a total of three singularities the following equation must be satisfied:

$$M_\infty = M_{u_0}M_{-u_0} \tag{3.4.11}$$

One might expect an ordering ambiguity here. In fact, the possible orders are related by the $\mathbb{Z}_2$ symmetry. The relation is not preserved due to the non-commutativity of the matrices. This is explained by noting that the $\mathcal{Z}_2$ symmetry also acts on the base point of the monodromy. In particular, $M_\infty \to M'_\infty$ due to its dependence on this definition. Equivalently, the relation $M_\infty = M_{-u_0}M_{u_0}$ holds if the massless particle at $u = -u_0$ is now a $(1,1)$ dyon.

We leave the monodromy condition as a consistency check and exploit the residual $\mathbb{Z}_2$ symmetry to obtain the answer. Since we know the behaviour in a neighbourhood of $u = u_0$, because of the $\mathbb{Z}_2$ symmetry there should be a massless particle associated to $u = -u_0$. To probe this neighbourhood, we send: $u \mapsto e^{i\pi}u$. The dual Higgs expectation value then reads:

$$a(d) = -c(u + u_0) \tag{3.4.12}$$

What are the quantum numbers of this particle? We again use the fact that the unbroken R-symmetry rotations may be interpreted as shifts in the θ angle. Rotating u by an angle of π corresponds to rotating a by an angle $\frac{1}{2}\pi$. Remembering from Section 2.4, this corresponds to a shift in the θ angle with -2π. By the Witten effect, the spectrum will be shifted by one unit in electric charge. We conclude that the particle responsible for a singularity at $u = -u_0$ should be a dyon of charge $(1, -1)$.

To determine what kind of monodromy we expect when such a dyon becomes massless, we follow the same method that was used to determine the monodromy of the magnetic monopole in (3.4.10). In the dual description the monodromy is the same as in (3.4.9). To transform to our electric variables (a, a_D), we need the change of basis. Since the dyon is related to the original electric variables as:

$$\begin{pmatrix}a_D(d)\\ a(d)\end{pmatrix} = \begin{pmatrix}-(1+b) & b\\ 1 & -1\end{pmatrix}\begin{pmatrix}a_D\\ a\end{pmatrix}, \tag{3.4.13}$$

we find the monodromy in our original variables to be:

$$M_{-u_0} = \begin{pmatrix}-(1+b) & b\\ 1 & -1\end{pmatrix}^{-1}\begin{pmatrix}1 & 2\\ 0 & 1\end{pmatrix}\begin{pmatrix}-(1+b) & b\\ 1 & -1\end{pmatrix} = \begin{pmatrix}-1 & 2\\ -2 & 3\end{pmatrix}. \tag{3.4.14}$$

The wonderful thing is that this matrix satisfies condition (3.4.11) and therefore provides a highly non-trivial consistency check on the assumption concerning the presence of only two extra singularities on the moduli space due to magnetically charged objects.

We conclude with some remarks:

1. We have seen that under the assumption a monopole of charge $(1, 0)$ and a dyon of charge $(1, -1)$ become massless at $u = u_0$ and $u = -u_0$ respectively, the monodromy at infinity

is correctly reproduced. In principle we would expect that other dyons could also become massless, especially since the monopole really is equivalent to any dyon modulo 2π rotations in the θ angle. Since M_∞ corresponds to a shift in the θ angle of 4π, we anticipate that a dyon of charge $(1, 2k)$ becomes massless at $u = u_0$ after we performed the monodromy at $u = \infty$ k times. The monodromy associated to a general dyon is determined as:

$$M = \begin{pmatrix} \frac{1+bn_m}{n_e} & b \\ n_m & n_e \end{pmatrix}^{-1} \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{1+bn_m}{n_e} & b \\ n_m & n_e \end{pmatrix} \tag{3.4.15}$$

$$= \begin{pmatrix} 1 + 2n_m n_e & 2n_e^2 \\ -2n_m^2 & 1 - 2n_m n_e \end{pmatrix}. \tag{3.4.16}$$

Indeed, we find in complete analogy with the $(1, 0)$ monopole and the $(1, -1)$ dyon that the monodromies associated to a $(1, 2k)$ and a $(1, 2k - 1)$ dyon satisfy:

$$M_\infty = M_{u_0}^{(1,2k)} M_{-u_0}^{(1,2k-1)} \tag{3.4.17}$$

with $M_{u_0}^{(1,2k)}$ and $M_{-u_0}^{(1,2k-1)}$ related respectively to M_{u_0} and M_{-u_0} through conjugation by M_∞^k. One also finds that the condition is satisfied with dyons of negative magnetic charge or one dyon of positive and one dyon of negative magnetic charge. The condition cannot be satisfied by dyons of magnetic charge $|n_m| \geq 2$. Note that these statements are only valid if the semiclassical dyon states exist at strong coupling as well. In fact, the assumption that these states exist is false and the only states that become massless at strong coupling could be the unit charge monopole and unit charge dyons. This was pointed out in [26]. We will come back to this in the next two sections.

2. The appropriate Higgs expectation values in terms of the original variables are:

$$a(d) = a_D \tag{3.4.18}$$

$$a(d') = a_D - a \equiv \tilde{a}_D \tag{3.4.19}$$

In terms of these variables, we find two new local (asymptotic) descriptions of our theory for which the prepotentials are:

$$\mathcal{F}_D(a_D) = -\frac{a_D^2 i}{4\pi} \log\left(\frac{a_D^2}{\Lambda^2}\right) + \sum_{k=1}^{\infty} c_k^D \left(\frac{a_D}{\Lambda}\right)^k \tag{3.4.20}$$

$$\tilde{\mathcal{F}}_D(\tilde{a}_D) = -\frac{\tilde{a}_D^2 i}{4\pi} \log\left(\frac{\tilde{a}_D^2}{\Lambda^2}\right) + \sum_{k=1}^{\infty} \tilde{c}_k^D \left(\frac{\tilde{a}_D}{\Lambda}\right)^k \tag{3.4.21}$$

The qualitative differences with the prepotential at $u = \infty$ manifest themselves in an opposite sign of the logarithm, related to the fact that QED is IR free. Furthermore, the infinite sum appearing in the expressions represents corrections which are not associated to instantons, since these do not exist for this particular $U(1)$ gauge theory. They reflect corrections of massive states which have been integrated out to arrive at the particular dual theory. Note that they vanish altogether at $u = \pm u_0$, while one can check that the gauge coupling properly diverges $\tau_D, \tilde{\tau}_D \to i\infty$ at $u = \pm u_0$ respectively.

3. The monodromies generate a group $\Gamma(2) \subset SL(2, \mathbb{Z})$. The fact that a subgroup of $SL(2, \mathbb{Z})$ appears is of course not a coincidence. Since the monodromies merely reflect a bad choice of local fields, the Kähler metric, which encodes the physics on the moduli space, should be invariant under monodromies. As already argued, the metric is invariant only under $SL(2, \mathbb{Z})$.[12]

[12] Actually, the metric is invariant under $SL(2, \mathbb{R})$. However, for the quantum numbers to remain integer valued the group must be $SL(2, \mathbb{Z})$, which may be seen from the formula of the central charge.

The moduli space of the theory is given by all possible values of u: $\mathbb{C} \setminus \{\pm\Lambda^2, \infty\}$. Given a section $(a_D(d), a(d))$ at some u, there exists an entire $\Gamma(2)$ orbit of physical descriptions. The correct description is the one such that $a(d)$ is invariant under the monodromy. A crucial insight of Seiberg and Witten was to note that the thrice punctured u-plane can be thought of as the fundamental domain of $\Gamma(2)$:

$$\mathcal{M}_q \cong \mathbb{H}/\Gamma(2) \tag{3.4.22}$$

It means there is some one-to-one map $u(f)$ from $\mathbb{H}/\Gamma(2)$ to the thrice puncture u plane. The monodromies of the gauge coupling suggest: $f = \tau(u)$. This realization stands at the basis of solving the theory. The punctures on the u plane coincide with the cusps $0, 1, \infty$. One can check that the cusps are fixed points under M_{u_0}, M_{-u_0} and M_∞ respectively. The moduli space is depicted in Figure 3.4.2. The cusps correspond to the weakly coupled electric, magnetic and dyonic descriptions.

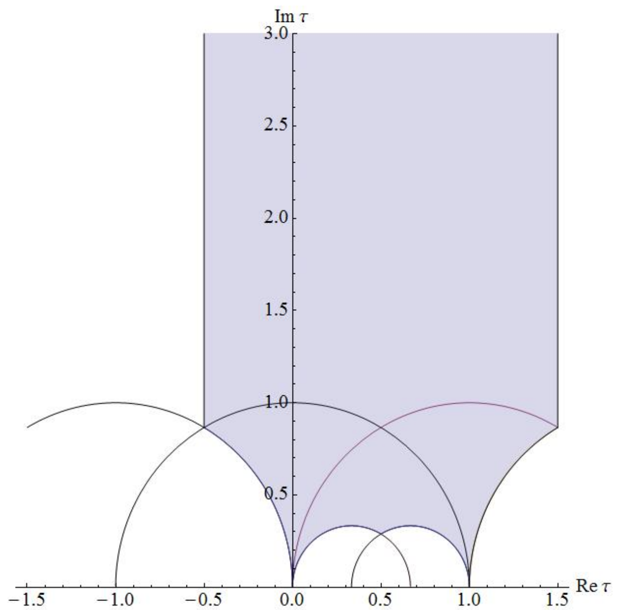

Figure 3.4.2: Fundamental domain of $\Gamma(2) \subset SL(2, \mathbb{Z})$.

As a last point, we make a remark about (non) electric-magnetic duality of the pure theory. Although the monodromies seem to imply $\Gamma(2)$ orbits of physical descriptions, one has to be aware we only know about the duality transformation relating the dyons at weak coupling. For one, no duality transformation between the W-bosons and dyons *could* exist since they make up different $\mathcal{N} = 2$ multiplets, in contrast with for instance $\mathcal{N} = 4$ SYM[47]. More strikingly, as already remarked above and to be discussed below, the strong coupling spectrum looks completely different from the weak coupling spectrum, preventing any chance on global electric-magnetic duality of the theory.

3.5 Solution of the Model

It might seem unlikely, with the information obtained thus far, to be able to calculate the instanton coefficients up to infinite order. However, this is precisely the achievement of Seiberg and Witten.

The crucial observation was to note that the quantum moduli space $\mathcal{M}_q$ of the theory coincides with the moduli space of a particular elliptic curve. Before giving the prescription for these elliptic curves, we briefly want to motivate what is the moduli space of a torus which we will see to be topologically equivalent to an elliptic curve.

First of all, a useful way to view a torus is as $\mathbb{C}/\Lambda$, with Λ a lattice spanned by two complex numbers b_1, b_2 called the periods of the torus. This corresponds to taking a primitive parallelogram in the lattice and identifying its opposite edges. Without loss of generality, we may take b_2 along the real positive axis and choose b_1 and b_2 such that: $\mathrm{Im}\,\frac{b_1}{b_2} > 0$. We may now simply rescale the lattice such that the side on the real axis has length 1. Then the torus is parametrized by one complex number, which is called the modulus of the torus:

$$\tau = \frac{b_1}{b_2}. \tag{3.5.1}$$

By construction, the imaginary part of the τ lives in the upper half plane. Two tori are equivalent if they differ by an $SL(2,\mathbb{Z})$ transformation. This is easily seen from the definition of the original lattice: $\mathbb{C} \supset \Lambda = \{nb_1 + mb_2 | n, m \in \mathbb{Z}\}$. The action of $SL(2,\mathbb{Z})$ is:

$$\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \rightarrow \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \tag{3.5.2}$$

Since $ad - bc = 1$ this action preserves the lattice, although the primitive cell may have changed. Upon scaling the lattice, the action can equivalently be expressed as:

$$\tau \rightarrow \tau' = \frac{a\tau + b}{c\tau + d} \tag{3.5.3}$$

From this perspective, the statement that tori related by such an $SL(2,\mathbb{Z})$ transformation are equivalent should either be interpreted as above or, if one wishes to take this more literal, one should properly normalize τ. Up to an irrelevant phase, a lattice spanned by a parallelogram with periods $(\frac{1}{\sqrt{\mathrm{Im}\tau}}, \frac{1}{\sqrt{\mathrm{Im}\tau}}\tau)$ is invariant under the action of $SL(2,\mathbb{Z})$ as in (3.5.3). We will always drop this normalization.

Another useful point of view from which it is clear tori related by an $SL(2,\mathbb{Z})$ transformation are equivalent is to notice $SL(2,\mathbb{Z})$ is the mapping class group of the torus. In this sense, an $SL(2,\mathbb{Z})$ transformation on the torus corresponds to a change of homology basis, just like $SL(2,\mathbb{Z})$ action on τ changes the primitive cell.

We see that the moduli space of inequivalent tori is given by a quotient of the upper half-plane: $\mathbb{H}/SL(2,\mathbb{Z})$. Indeed, this is similar to the moduli space of our theory. To make the correspondence precise, Seiberg and Witten considered the following elliptic curve:

$$y^2 = (x - \Lambda^2)(x + \Lambda^2)(x - u) \tag{3.5.4}$$

This equation defines a one-complex dimensional subspace of $\mathbb{C}^2$. This curve turns out to be a genus one Riemann surface, as may be seen in Figure 3.5.1. The components of the section $(a_D(u), a(u))$ will be determined by integrating an appropriate differential, the so-called Seiberg-Witten differential, on the a and b cycle respectively. Before we give the explicit expression, let us see why the moduli space of the theory is the same as that of the elliptic curve.

Obviously, for different values of u we obtain different elliptic curves. However, the moduli space of the elliptic curve is reduced as well by noticing some symmetries. First of all, the $\mathbb{Z}_2$ symmetry on the moduli space of the physical theory is realized on the curve as a $\mathbb{Z}_4$ symmetry on the curve, of which only a $\mathbb{Z}_2$ subgroup acts on u:

$$u \rightarrow -u$$
$$x \rightarrow -x$$
$$y \rightarrow \pm iy$$

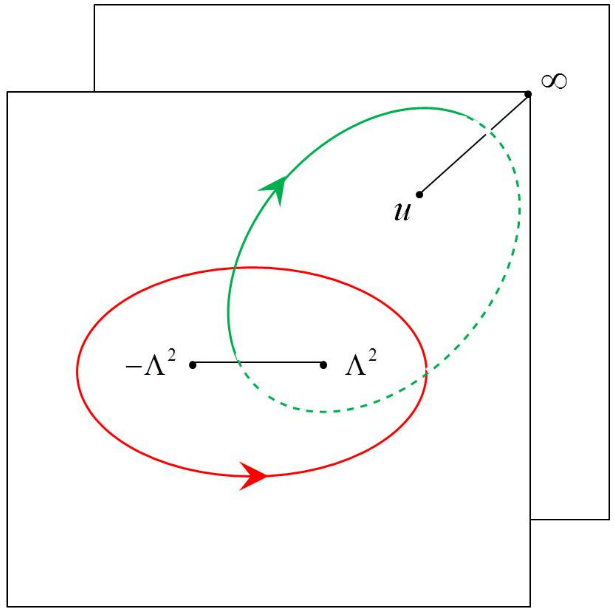

Figure 3.5.1: The elliptic curve is defined on a double cover of the x plane, compactified to two Riemann spheres. The branch are chosen between $\left[-\Lambda^2, \Lambda^2\right]$ and $[u, \infty]$. The conventional choice for an independent basis of the homology of the curve is a cycle circling the branch cut $\left[-\Lambda^2, \Lambda^2\right]$, which we call the a cycle. This cycle is defined on the upper sheet. The other cycle, the b cycle, we choose as entering the branch cut $\left[-\Lambda^2, \Lambda^2\right]$, continuing on the second sheet and returning and closing on the first sheet via the cut at $[u, \infty]$. The gluing of the two Riemann spheres along the cuts leads to a $g = 1$ Riemann surface. For definiteness, we define the a cycle to be oriented anti-clockwise while the b cycle is defined clockwise.

The transformations on x and y precisely implement the usual R-symmetry transformations of $a(u)$ and $a_D(u)$ when $u \to -u$, as we will see from the solution.[13]

As u approaches the strong coupling region, the monopole and dyon start to become light. This is realized on the elliptic curve as the degeneration of a certain cycle. In Figure 3.5.2 these vanishing cycles are depicted and readily identified with the expectation values of the corresponding magnetic and dyonic Higgs scalars. Notice that there will also be a degeneration of the curve when $u \to \infty$. However, we know from the physical theory this should not correspond to a massless particle, but rather to infinitely massive particles. The proper way to understand this pinching of the torus is to rescale $x \to \frac{x}{u}$ while simultaneously scaling $y \to \frac{y}{u^{3/2}}$ such that the curve is invariant. The strong coupling singularities then lie at $\frac{\Lambda^2}{u}$. We therefore see that in these new coordinates, in the limit $u \to \infty$, the a cycle vanishes while the b cycle diverges. Indeed, this is precisely the limit discussed in the beginning of Section 3.4 where we saw that for $u \to \infty$ the ratio $a_D/a \to \infty$.

Last but not least, we check how the monodromies of the physical theory are realized in elliptic curve construction. To check this, we analyse what happens to the a and b cycles when u circles $\pm\Lambda^2$. Let us first look at the monodromy at $u = \Lambda^2$. Notice that looping u around Λ^2 may be viewed as exchanging u and Λ^2 twice along the b cycle. Therefore, after the exchange the new b' cycle is really the same cycle as the original b cycle. However, the a cycle will be transformed.

[13] It may be seen from the abelian low energy effective action that $R_{a_D} = R_a$.

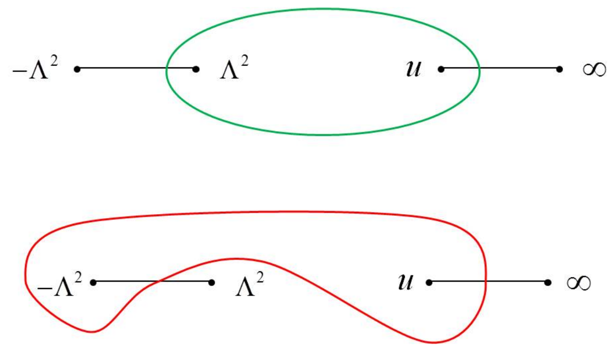

Figure 3.5.2: Vanishing cycles. In the limit that $u \to \Lambda^2$, the b cycle vanishes. In the limit that $u \to -\Lambda^2$, the $b-a$ cycle vanishes. The b cycle will correspond to a_D whereas the $b-a$ cycle will correspond to the vanishing of $\tilde{a}_D = a_D - a$. Notice that the signs appearing in front of a cycle are consequences of the orientations defined in Figure 3.5.1.

To understand how the a cycle transforms, we note there is an ambiguity when u is about to cross the branch cut. To avoid this, we should change our branches accordingly. After a rotation of π we take the new branches to be $\left[-\Lambda^2, u\right]$ and $\left[\Lambda^2, \infty\right]$. With these new branches, we canonically define the new a'. The a' cycle now encircles u and $u = -\Lambda^2$. This corresponds to an $a - b$ cycle in the original coordinates. Exchanging twice then corresponds to:

$$\begin{pmatrix} b' \\ a' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} b \\ a \end{pmatrix}. \tag{3.5.5}$$

This precisely corresponds to the monodromy of the original theory. Let us also check the monodromy at $-\Lambda^2$ in a similar manner. The cycle along we are exchanging now corresponds to $a - b$. Analogously to the previous, now the $a-b$ remains invariant, i.e. $a' - b' = a - b$. After exchanging u and $-\Lambda^2$ twice we have for the a cycle: $a' = a - 2(b-a) = 3a - 2b$. In matrix notation, we therefore have:

$$\begin{pmatrix} b' \\ a' \end{pmatrix} = \begin{pmatrix} -1 & 2 \\ -2 & 3 \end{pmatrix} \begin{pmatrix} b \\ a \end{pmatrix}. \tag{3.5.6}$$

Again, this is exactly the right monodromy. The monodromy at $u = \infty$ is automatically reproduced by applying the global constraint on the monodromies:

$$M_\infty = M_{u_0} M_{-u_0} \tag{3.5.7}$$

Hence, elliptic curves related by $\Gamma(2)$ transformations define the same elliptic curve, although the homology basis has transformed. We conclude that the moduli space of the elliptic curve precisely coincides with the moduli space of the physical theory.

We will now quantify the correspondence. First, we have seen that the b and a cycle on the elliptic curve transform exactly as $a_D(u)$ and $a(u)$ under $SL(2,\mathbb{Z})$ and have the same monodromies. In analogy with the physical theory, we can simply define an $SL(2,\mathbb{Z})$ bundle over the punctured u plane. In this case, $SL(2,\mathbb{Z})$ acts naturally as the mapping class group of the elliptic curve E_u on the abstract vector space spanned by the a and b cycle. The fiber on a particular value of the punctured u plane is then the homology group $H^1(E_u,\mathbb{C})$ of the elliptic curve. To relate both sections, we first note that we may relate the homology group $H^1(E_u,\mathbb{C})$ as a space with the space of meromorphic one-forms on E_u through the pairing:

$$\gamma \mapsto \oint_\gamma \lambda \tag{3.5.8}$$

where $\gamma \in H^1(E_u, \mathbb{C})$ and λ a meromorphic one-form on E_u. To make this relation into a bijection, we should consider the meromorphic one-forms modulo exact forms (since these will integrate to zero). Also, the forms should have vanishing residues. We demand this since the paths γ are only defined up to homotopy. If λ would have a non-zero residue at some point on the elliptic curve, the pairing would not be invariant once γ crosses a pole of λ.[14] There is an obvious differential satisfying these requirements: the unique holomorphic one-form of the elliptic curve.[15] On the torus viewed as a lattice in $\mathbb{C}$, this holomorphic form is just $\mathrm{d}z$, which naturally integrates on the two cycles to give the periods of the torus:

$$b_1 = \oint_b \mathrm{d}z \tag{3.5.9}$$

$$b_2 = \oint_a \mathrm{d}z \tag{3.5.10}$$

The equivalent differential on the elliptic curve has a different form. There are three useful ways to see that the correct holomorphic form is given by:

$$\lambda_1 = \frac{\mathrm{d}x}{y} \tag{3.5.11}$$

We briefly state how it may be derived:

1. Since the curve is described in two coordinates, yet is one dimensional, we may choose to describe its cotangent bundle in terms of x or y coordinate projection. The relation between these two parametrizations is given by the defining equation for the curve (3.5.4) (which we abbreviate for ease of notation as $f(y) = g(x)$):

$$f'(y)\mathrm{d}y = g'(x)\mathrm{d}x \tag{3.5.12}$$

If the curve has no double zeroes, $g'(x)$ has no zeroes in common with $g(x)$ and we see that we always have have a non-singular and non-vanishing differential at our disposal:

$$\omega = \frac{\mathrm{d}x}{y} \quad \text{or} \quad \omega = \frac{\mathrm{d}y}{g'(x)}, \tag{3.5.13}$$

where $y = \sqrt{g(x)}$ and similarly $g'(x)$ is considered as a function of y.

2. It is a δ function identity: consider the volume form on $\mathbb{C}^2$ $\mathrm{d}x \wedge \mathrm{d}y$. We can restrict the form to the curve by inserting an appropriate δ-function:

$$\int_{\mathbb{C}^2} \mathrm{d}x \wedge \mathrm{d}y\, \delta(y^2 - g(x)) = \int_{E_u} \frac{\mathrm{d}x}{2y} = \int_{E_u} \frac{\mathrm{d}y}{g'(x)} \tag{3.5.14}$$

Of course, the volume form is non-vanishing and holomorphic on $\mathbb{C}^2$. Upon restricting to the curve we obtain a natural holomorphic non-vanishing differential on the curve $E_u \subset \mathbb{C}^2$.

3. Alternatively, one may analyse the poles and zeroes of the numerator and the denominator in (3.5.11). They turn out to coincide and therefore cancel. See chapter 1 and 2 of [62] for more details.

[14] This condition will be relaxed when we consider the addition of flavours with non-zero masses. The residues of the differentials will then correspond to the bare masses of the flavours. Indeed, this is related to the extension of $SL(2, \mathbb{Z})$ as mentioned at the end of Section 3.2.

[15] The dimension of the space of holomorphic one-forms on a compact Riemann surface of genus g has dimension g. See for instance [41] for an introduction to complex curves and Riemann surfaces.

Analogously to (3.5.10), the period integrals on the elliptic curve become:

$$b_1 = \oint_b \frac{\mathrm{d}x}{y} \tag{3.5.15}$$

$$b_2 = \oint_a \frac{\mathrm{d}x}{y} \tag{3.5.16}$$

The modulus of the elliptic curve is now given by:

$$\tau_u = \frac{b_1}{b_2} \tag{3.5.17}$$

which automatically satisfies: $\mathrm{Im}\,\tau_u > 0$. We now assume this holds as well for our low energy theory: $\mathrm{Im}\,\tau(u) > 0$ for all values of u. Then clearly, $\tau(u)$ also parametrizes some torus. Secondly, since $\tau(u) = \frac{\mathrm{d}a_D/\mathrm{d}u}{\mathrm{d}a/\mathrm{d}u}$ we see the identification:

$$\frac{\mathrm{d}a_D}{\mathrm{d}u} = f(u) \oint_b \frac{\mathrm{d}x}{y} \tag{3.5.18}$$

$$\frac{\mathrm{d}a}{\mathrm{d}u} = f(u) \oint_a \frac{\mathrm{d}x}{y} \tag{3.5.19}$$

makes sense as all symmetries and monodromies agree on the left and right hand side. Here, $f(u)$ is some constant which will be fixed by asymptotic matching of the identification. Accepting this identification then fully solves the theory.

It is easy to see (3.5.19) implies:

$$a_D(u) = \oint_b \lambda \tag{3.5.20}$$

$$a(u) = \oint_a \lambda \tag{3.5.21}$$

with the Seiberg-Witten differential λ, up to exact forms, given by:

$$\lambda = -2f \frac{\sqrt{x-u}\mathrm{d}x}{\sqrt{x^2-\Lambda^4}} = -2f(\frac{x\mathrm{d}x}{y} - u\frac{\mathrm{d}x}{y}) \tag{3.5.22}$$

Indeed, differentiating either side gives back equation (3.5.19). As a consistency check, we see λ has vanishing residues since its only pole is double (at ∞) and the sum of residues should vanish of a meromorphic one-form on a compact Riemann surface vanishes.[16]

Let us now state the explicit solutions for the functions $a_D(u), a(u)$:

$$a_D(u) = -2f \oint_b \frac{\sqrt{x-u}\mathrm{d}x}{\sqrt{x^2-\Lambda^4}} = -4f \int_{\Lambda^2}^{u} \frac{\sqrt{x-u}\mathrm{d}x}{\sqrt{x^2-\Lambda^4}} \tag{3.5.23}$$

$$a(u) = -2f \oint_a \frac{\sqrt{x-u}\mathrm{d}x}{\sqrt{x^2-\Lambda^4}} = -4f \int_{-\Lambda^2}^{\Lambda^2} \frac{\sqrt{x-u}\mathrm{d}x}{\sqrt{x^2-\Lambda^4}}. \tag{3.5.24}$$

In principle, these functions, together with the knowledge of appropriate dual descriptions at the cusps in the moduli space, enables us to calculate the (locally defined) prepotentials up to arbitrary order.

[16] The easiest way to see this is by determining the divisors of $x, y, \mathrm{d}x$. The divisors of $y, \mathrm{d}x$ cancel, such that a double pole at ∞ of x remains. See Section 2.3 of [62] for more details.

We can fix the constant by matching the asymptotic behaviour. Near $u = \infty$, the expression for a becomes:

$$a(u) \approx -2f\sqrt{u}\int_{-\Lambda^2}^{\Lambda^2} \frac{\sqrt{-1}\mathrm{d}x}{\sqrt{x^2-\Lambda^4}} = -2f\pi\sqrt{u}. \tag{3.5.25}$$

This the correct behaviour. We may also determine the value of f to produce $a(u) = \sqrt{2u}$:

$$f = \frac{-1}{2\sqrt{2}\pi}. \tag{3.5.26}$$

Let us also check the expression of $a_D(u)$:

$$a_D(u) = \frac{\sqrt{2u}}{\pi}\int_{\Lambda^2/u}^{1} \frac{\mathrm{d}z\sqrt{z-1}}{\sqrt{z^2-u^{-2}\Lambda^4}}. \tag{3.5.27}$$

Here we changed variables $x = uz$. We can expand the integrand in terms of the small parameter Λ^2/u. The first term corresponds to:

$$a_D(u) \approx \frac{\sqrt{2u}}{\pi}\int_{\Lambda^2/u}^{1} \frac{\mathrm{d}z\sqrt{z-1}\sqrt{z^2}}{z^2} \tag{3.5.28}$$

The formula for $a_D(u)$ diverges. Isolating the divergent term, we find the dominant contribution for $u \to \infty$:

$$a_D(u) \approx \frac{\sqrt{2u}}{\pi}\int_{\Lambda^2/u}^{1} \frac{\mathrm{d}zi}{z} = \frac{\sqrt{2u}i}{\pi}\log\left(\frac{u}{\Lambda^2}\right) \tag{3.5.29}$$

which indeed corresponds to the divergent part of $a_D(u)$ obtained in our semiclassical analysis. In principle, one can determine higher order corrections in to both $a(u)$ and $a_D(u)$. Inverting $a(u)$, one can then integrate $a_D(a)$ to obtain the prepotential $\mathcal{F}(a)$. Similarly, one can determine the corrections to the prepotentials in terms of the monopole and dyonic coordinates, reinterpreting the functions $a(u)$ and $a_D(u)$ properly, while working in the limits $u \to \pm\Lambda^2$.

However, there is a more efficient way to calculate the instanton corrections. It turns out that the solutions for $a(u), a_D(u)$ admit a hypergeometric function representation. Expanding these functions order by order, one can obtain the instanton corrections up to arbitrary order. It can also be seen that $a(u) \neq 0$ for all values of u, such that indeed no singularities appear due to massless gauge bosons. For these explicit solutions, the reader is referred to [35]. Explicit instanton calculations agree with the Seiberg-Witten solution[42][45].

3.6 Discussion

In the previous section, we have seen an explicit solution for the prepotential. We have found that the prepotential can only be interpreted as a local function on the moduli space. For these local representations, we have provided a recipe through which in principle all perturbative and non-perturbative (instanton) corrections can be determined. In this sense, the theory is solved exactly. Of course, there is the possibility of other non-perturbative contributions to the prepotential. However, as was already mentioned in a footnote in the beginning of Section 2.5, the instantons represent the dominant non-perturbative contributions.

A second point is more of a remark than a real obstacle, and is made in preparation of Section 3.8. It concerns the normalization of electric charge. For the W-bosons, the electric charge should be normalized as $n_e = 2$ for the hypermultiplets to have integer electric charge. The effects are not drastic but should be taken into account. Note that we can effectively absorb the charge via the central charge formula into a redefinition of $a' = a/2$. This leads to a new τ as well:

$\tau'(a) = 2\tau(a) = \frac{\theta(a)}{\pi} + \frac{8\pi i}{g^2(a)}$. In these conventions the W-bosons behave as particles of charge $(n_m, n_e) = (0, \pm 2)$. Also the dyons obtain an even electric charge. Further consequences are that under a 2π rotation of θ, $\tau \to \tau + 2$. This implies the monodromy at infinity will be $\tau \to \tau - 4$. Also the monodromies at $u = \pm u_0$ will change, which together will now generate $\Gamma_0(4)$. This still is a sixfold cover of the modular domain, implying equivalent monodromy properties. The curve and Seiberg-Witten differential change accordingly and are given for instance in [35]. For a comprehensive discussion of the above, see Section 4.5 of [65].

A third point concerns the spectrum of the theory. Definitely, the semiclassical spectrum consists of all possible dyons with $(n_m, n_e) = (\pm 1, n)$ with n some integer. Because of the $\Gamma(2)$ monodromies we would expect at any value of u to have an infinite tower of dyon states, related by a $\Gamma(2)$ monodromy. However, as already suggested at the end of [59] and worked out in [26], there exists a boundary between the weak and strong coupling region at which: $\mathrm{Im}\frac{a_D(u)}{a(u)} = 0$. This implies the spectrum collapses from a lattice in the complex plane to the real line. All states decay into monopoles and dyons. The $(1, 0)$ monopole and the $(1, -1)$ dyon cannot decay because they are the lightest particles in the spectrum. The curve of marginal stability passes through the monopole and dyon point since there $a_D/a = 0$ and $a_D/a = \pm 1$ respectively. The shape of the curve is almost an ellipse, depicted in Figure 3.6.1. It was convincingly argued in [25] the curve of marginal stability

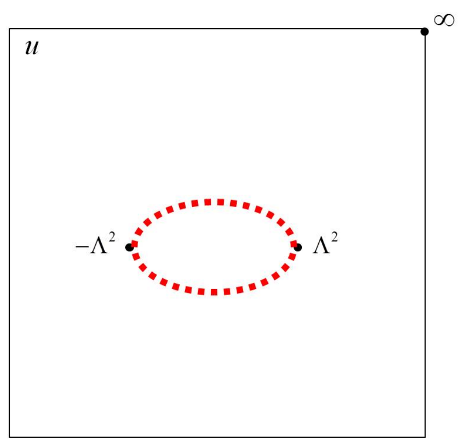

Figure 3.6.1: The moduli space of the pure theory. The singularities and the curve of marginal stability are depicted.

is simple and closed, and does not extend to infinity. This was also argued in the short note [49]. Of course, the explicit solution agrees with these observations. As already mentioned above, the existence of a curve of marginal stability is the final blow to some sort of electric-magnetic duality of the effective $\mathcal{N} = 2$ Yang-Mills action over the whole of the moduli space.

3.7 Confinement of Electric Charge

One interesting physical result of the solution of $\mathcal{N} = 2$ SYM is that for the first time a quantitative description has been found for confinement of electric charge through monopole condensation. This mechanism was conjectured to be responsible for confinement already in the late seventies by Mandelstam and made more precise by 't Hooft in [64]. Their conjecture is an electric-magnetically dual description of the Meissner effect in superconductors. In a superconducting vacuum, electrons have condensed to form Cooper pairs. These pairs are well described as a scalar field which assumes a non-zero expectation value in the vacuum. This spontaneously breaks a $U(1)$ gauge symmetry and the photons are given a mass by the Higgs mechanism inside the superconductor. However, it has been found that this theory does allow for stable magnetic flux tubes, Abrikosov vortices, through the condensate. An appropriate underlying description of these tubes is that they represent bound states of a monopole and antimonopole connected via the tube. These bound states are confined in the sense that the (magnetic) force between them increases with the distance.

Motivated by the appearance of magnetic monopoles in non-abelian gauge theories, 't Hooft conjectured that the condensation of magnetic monopoles would induce electric vortex tubes with at the ends a bound state of a quark and an antiquark. This could be a qualitative explanation of the confinement of electrically charged quarks. Seiberg and Witten found a quantitative description of this mechanism in $\mathcal{N}=2$ SYM. Although the discussion of this is not directly relevant for the remainder of this thesis, it is such an interesting result we decide to include it.

The idea is to break $\mathcal{N}=2$ SYM theory to a pure $\mathcal{N}=1$ theory. This is motivated by the fact that pure $\mathcal{N}=1$ SYM is believed to generate a mass gap at low energies.[17] Furthermore, it has two inequivalent vacua related by a $\mathbb{Z}_2$ symmetry. It is then interesting to see how the low energy $\mathcal{N}=2$ theory, whose moduli space we understand completely, flows to the pure $\mathcal{N}=1$ theory.

To break $\mathcal{N}=2$ supersymmetry explicitly, we can simply add a mass term for the chiral superfield Φ:

$$W(\Phi) = m\,\mathrm{Tr}\,\Phi^2 \tag{3.7.1}$$

In the low energy limit, Φ will freeze out and we will indeed be left with a pure $\mathcal{N}=1$ theory. Also, the operator will be identified with $U=\mathrm{Tr}\,\Phi^2$ and the effective superpotential becomes:

$$W_{\mathrm{eff}} = mU \tag{3.7.2}$$

where the vev of the scalar component of U is related in the usual way to the scalar component of Φ: $u=\langle \mathrm{Tr}\,\phi^2\rangle$. The expression for the low energy effective superpotential is exact due to a non-renormalization theorem. See [59].

Clearly, the addition of such a term to the theory will affect the vacuum degeneracy. Reasoning semi-classically, one would be tempted to argue the real vacuum is now unique and given for $a=0$ with a the vacuum expectation value of the scalar component of Φ. However, we have learned from the previous analysis that this is not what happens in the $\mathcal{N}=2$ theory. In fact, $a(u)$ does not vanish anywhere on the moduli space. Also from the $\mathcal{N}=1$ perspective, we expect a doubly degenerate vacuum with a mass gap because of its presumed confining phase. The natural place to search for the $\mathcal{N}=1$ vacua is the strong coupling region of the $\mathcal{N}=2$ theory. Indeed, this exhibits a $\mathbb{Z}_2$ symmetry. We should change our description of the $\mathcal{N}=2$ theory accordingly and take into account chiral multiplets M and $\tilde{M}$ which constitute a magnetic $\mathcal{N}=2$ hypermultiplet. The exact effective superpotential now is:

$$W_{eff} = mU(\Phi_D) + \sqrt{2}\Phi_D M\tilde{M} \tag{3.7.3}$$

Again, this is an exact expression for the superpotential. We can find the vacuum of the theory by analysing the minimum of the superpotential. Furthermore, the D-term generated by the hypermultiplet should vanish as well, requiring $|\langle M\rangle| = |\langle\tilde{M}\rangle|$. The minimum is obtained by differentiating the superpotential with respect to the fields and evaluating on the vacuum manifold. We slightly abuse notation by denoting the vacuum expectation value of the monopole superfields again with $M, \tilde{M}$:

$$a_D M = a_D\tilde{M} = 0 \tag{3.7.4}$$

$$\sqrt{2}M\tilde{M} + m\frac{\mathrm{d}u}{\mathrm{d}a_D} = 0 \tag{3.7.5}$$

[17] For a review, see Section 4 of [50].

Assuming $\frac{\mathrm{d}u}{\mathrm{d}a_D} \neq 0$, which at least near the monopole point is a reasonable assumption since $a_D(u)$ is regular, we see the equations can only be satisfied for non-vanishing values of $M, \tilde{M}$. Furthermore, a_D should vanish. This is precisely what happens at Λ^2.The vacuum expectation values of the hypermultiplets read:

$$M = \tilde{M} = \left(-mu'(0)/\sqrt{2}\right)^{1/2}. \tag{3.7.6}$$

This is determined up to $U(1)$ gauge transformations, under which M and $\tilde{M}$ transform oppositely. The non-zero vevs of the scalar components of M and $\tilde{M}$ give rise to a Higgs mechanism resulting in a massive magnetic $U(1)$ gauge boson. This is precisely when confinement of electric charge is expected to occur. We conclude that the $\mathcal{N} = 2$ theory provides an underlying physical explanation for the mass gap of the $\mathcal{N} = 1$ theory. A similar analysis could be done at the dyon singularity. The condensation of the dyon gives rise to a mass gapped theory as well, establishing the $\mathbb{Z}_2$ symmetry on the two vacua of the $\mathcal{N} = 1$ theory.

3.8 Adding Matter

The exact solution for the prepotential of pure $SU(2)$ theory has inspired a great deal of effort into the study of other, more complicated $\mathcal{N} = 2$ gauge theories. It has turned out the methods and insights of Seiberg and Witten are applicable to a huge class of $\mathcal{N} = 2$ theories, which subsequently have been solved for exactly. For instance, the curves for arbitrary rank classical gauge groups have been found. See [2] and [9] on the generalization to $SU(N)$ gauge groups and [11] for $SO(N)$ and $Sp(N)$ gauge groups. Although the curves are found using a similar analysis as the original analysis of Seiberg and Witten, it was Witten again who found a more general and insightful construction in Type IIA string theory (or M-theory) in [72]. In the next section, we will briefly mention the most important elements of his analysis. Here, we comment on an interesting aspect of the generalization to higher rank $SU(N)$ groups stemming from a richer singularity structure of the higher rank theories, without plunging into a quantitative analysis.

In the case of higher rank gauge groups, the Coulomb branch of a spontaneously broken gauge theory will be a higher dimensional manifold. For instance, a generic Higgs expectation value will spontaneously break $SU(N)$ to $U(1)^{N-1}$ corresponding to a Coulomb branch of complex dimension $N-1$. The full quantum moduli space of the low energy theory then corresponds to the moduli space of genus $g = N-1$ curves, where every torus represents a $U(1)$ theory. The period matrix of the elliptic curve represents the set of gauge couplings corresponding to the $U(1)$ subgroups. It is a matrix of dimension $N-1 \times N-1$ matrix τ_{ij} which carries the information about the ratios a and b periods:

$$\sum_{j=1}^{g} \tau_{ij} a_j = b_i \tag{3.8.1}$$

where we warn the reader a_j and b_i are separately matrices, since they depend on another index $k = 1, \ldots, g$ corresponding to a choice of holomorphic one-form on the Riemann surface. See for instance [21].

Massless particles appear generically along complex codimension one submanifolds of the Coulomb branch, since there exists a dual description in which a generic dyon is electrically charged with respect to one of the $U(1)$ gauge groups.[18] This implies one may expect intersections of these codimension one manifolds leading to theories of several $U(1)$ vector multiplets coupled to several electric, magnetic or dyonic hypermultiplets.

[18] In the case of $SU(N)$ theories, the duality group which keeps the Kähler metric invariant is $Sp(2N-2)$. Its action on τ_{ij} is precisely as $SL(2,\mathbb{Z})$ in the $SU(2)$ case. The statement is there exists some $A \in Sp(2N-2)$ which sends a generic charge vector: $\vec{n} \mapsto (k, 0, \ldots, 0)$ with $k = \gcd(n_i)$. This is a straightforward generalization of the statement for $SU(2)$ in Section 3.3.

Interestingly, it was argued in [8] that mutually *non-local* states become massless at submanifolds of the Coulomb branch, and an explicit example in the case of $SU(3)$ gauge group was provided. This insight has led to a new type of superconformal theory in which $U(1)$ vector multiplets are simultaneously coupled to massless electric and magnetic hypermultiplets.

In this section, however, we will discuss the other obvious generalization of the analysis, namely the addition of matter in the fundamental representation of the gauge group as first considered for $SU(2)$ in the 'second' paper [60] of Seiberg and Witten. For a generalization to $SU(N)$ gauge groups, see [48]. We will not discuss or give the curves corresponding to the matter included theories, since again the M-theory construction of the next section provides a much more general way of understanding these. However, some qualitative aspects of the theories coupled to matter are interesting and put the pure case in a somewhat broader perspective. For this, we mainly use the original paper by Seiberg and Witten [60], especially Sections 5, 6, 7, 8 and 10. Furthermore, elements of the reviews [46] and [65] are used.

3.8.1 Global Symmetries of $N_f \leq 4$

We will only consider UV free theories. The exact perturbative beta function for $\mathcal{N} = 2$ SYM with gauge group $SU(N)$ coupled to N_f massless flavours in the fundamental representation is given by:

$$\beta_{\frac{4\pi}{g^2}}(\mu) = \frac{2N - N_f}{2\pi} \tag{3.8.2}$$

UV free theories are seen to correspond to $N_f \leq 3$. For $N_f = 4$ the one-loop beta function vanishes. To include matter in an $\mathcal{N} = 2$ invariant way, we couple the vector multiplet to hypermultiplets as in Section 2.3. Although the moduli space is extended with a Higgs branch, we will only focus on the Coulomb branch. Much information about the structure of the Coulomb branch can be obtained by a careful analysis of the (quantum) global symmetries. Let us therefore first see what are the global symmetries of a theory coupled to N_f flavours.

First of all, we have the $U(2)_R$ symmetry. As discussed in Section 2.4, its $U(1)_R$ subgroup is anomalous. It follows from the general expression of the anomaly (2.4.11) that for gauge group $SU(2)$ and N_f fermions in the fundamental representation, the surviving R-symmetry group is:

$$SU(2)_R \times \mathbb{Z}_{8-2N_f}/\mathbb{Z}_2 \tag{3.8.3}$$

In Table 3.1, the residual R-symmetries for $1 \leq N_f \leq 4$ are shown. The $U(1)_R$ symmetry is broken

N_f	1	2	3	4
H	$\mathbb{Z}_3$	$\mathbb{Z}_2$	none	$U(1)_R$

Table 3.1: The unbroken R-symmetry group $H \subset U(1)_R$ as a function of the number of flavours added to the theory.

altogether for non-vanishing bare masses of the hypermultiplets. The other global symmetry is flavour symmetry. We repeat the superpotential of the hypermultiplet with explicit gauge indices:

$$W = \sum_{i=1}^{N_f} \sqrt{2}\tilde{Q}_i^a \, \Phi_a{}^b \, Q_{ib} + m_i \tilde{Q}_i^a \, Q_{ia}. \tag{3.8.4}$$

For non-zero and equal m_i both the superpotential and the kinetic terms of the hypermultiplets are invariant under a $U(N_f)$ flavour symmetry. For unequal masses it generically reduces to a product $U(1)^{N_f}$. Notice that Q and $\tilde{Q}$ carry opposite flavour charges with respect to the $U(1)$ flavour symmetries.

Since hypermultiplets must appear in short representations of the $\mathcal{N} = 2$ algebra, massive hypermultiplets inescapably saturate a BPS bound. The BPS bound of the pure theory has to be adjusted to account for non-vanishing bare masses. To determine the contribution of these hypermultiplets one could construct the supercurrents of the explicit supersymmetry transformations on the component fields of this flavoured $SU(2)$ super Yang-Mills theory. Then one can construct by hand the supercharges and calculate the central charges. For the pure theory we already quoted the result in Section 3.3 from [73]. A similar calculation can be done including hypermultiplets. Some details of this calculation are given in [4]. We will not perform the calculation but rather motivate the answer.

Writing the superpotential in terms of a single hypermultiplet $Q_{1,2}$ and $\tilde{Q}^{1,2}$, giving Φ its usual expectation value and expanding the superpotential (and its hermitian conjugate), we obtain:

$$\begin{aligned} W + W^\dagger = &\tilde{Q}^1(m + \sqrt{2}a)Q_1 + \tilde{Q}^2(m - \sqrt{2}a)Q_2 \\ &+ Q^{*1}(-m + \sqrt{2}a)\tilde{Q}_1^* + Q^{*2}(-m - \sqrt{2}a)\tilde{Q}_2^* \end{aligned} \tag{3.8.5}$$

The minus signs in front of the bare masses appear due to the transposition performed, which adds a minus sign because the $SU(2)$ inner product is antisymmetric. One might worry that the expression seems to imply a breaking of $SU(2)_R$ symmetry. For this, note that an $SU(2)_R$ doublet is made up from (the scalars of):

$$Q = \begin{pmatrix} Q_1 \\ Q_2 \end{pmatrix} \quad \text{and} \quad \tilde{Q}^\dagger = \begin{pmatrix} \tilde{Q}_2^* \\ \tilde{Q}_1^* \end{pmatrix} \tag{3.8.6}$$

where these are the $SU(2)$ gauge spinors and $SU(2)_R$ mixes both the upper (lower) components. See Section 2.4 for more details on the action of $\mathcal{N} = 2$ R-symmetry on the hypermultiplet. We conclude the BPS masses *do* respect the $SU(2)_R$ symmetry, as desired of course.

Then we see that the upper and lower components of the spinor have effective masses:

$$m_\pm^{\mathrm{BPS}} = |m \pm \sqrt{2}a| \tag{3.8.7}$$

For a single quark, we should therefore adjust the central charge as:

$$Z = an_e + a_D n_m + S\frac{m}{\sqrt{2}}, \tag{3.8.8}$$

where $S = \pm 1$ for $Q, \tilde{Q}^\dagger$ and $\tilde{Q}, Q^\dagger$ respectively and ensures the $SU(2)_R$ will not be broken. This is easily extended to N_f massive quarks:

$$Z = an_e + a_D n_m + \sum_{i=1}^{N_f} S_i \frac{m_i}{\sqrt{2}}, \tag{3.8.9}$$

where quark Q_i has $U(1)$ flavour charge $S_j = 1$ if $j = i$ and otherwise $S_j = 0$. This is the low energy effective BPS bound due to a non-renormalization theorem for the superpotential.

For gauge group $SU(2)$ there is a flavour symmetry enhancement possible broken by non-vanishing bare masses. Namely, pseudo-reality of the fundamental representation of $SU(2)$ implies it is isomorphic to its complex conjugate. Therefore, we may combine Q and $\tilde{Q}$ into a single $2N_f$ dimensional spinor $\widehat{Q}$ transforming in the fundamental representation [34]. For a single hypermultiplet this spinor is represented as:

$$\widehat{Q} = \begin{pmatrix} Q \\ \tilde{Q}^* \end{pmatrix} = \begin{pmatrix} Q \\ \sigma_G \tilde{Q} \end{pmatrix} \tag{3.8.10}$$

Our conventions have shifted slightly with respect to the discussion in Section 2.3; we have redefined $\tilde{Q}^t \to \tilde{Q}$. See Appendix C for the motivation and details. The action of $\sigma_G = i\sigma_2$ is on the $SU(2)$

gauge indices. Notice this action can be expressed explicitly as $(\sigma_G)_a{}^b\tilde{Q}_b$ which is equivalent to $Q^b\epsilon_{ba}$.

It is shown in the Appendix of [34] that the hypermultiplet Lagrangian can be rewritten in terms of $\widehat{Q}$ as follows:

$$\mathcal{L} = \int d\theta^2 d\bar{\theta}^2 \begin{pmatrix} Q^\dagger & \tilde{Q}^\dagger\sigma_G^\dagger \end{pmatrix} e^{-2V} \begin{pmatrix} Q \\ \sigma_G\tilde{Q} \end{pmatrix} + \int d\theta^2 \frac{1}{\sqrt{2}} \begin{pmatrix} Q^t & \tilde{Q}^t\sigma_G^t \end{pmatrix} \sigma_G\Phi \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} Q \\ \sigma_G\tilde{Q} \end{pmatrix} + h.c. \tag{3.8.11}$$

We suppress gauge indices and use vector/matrix notation for the flavour indices. The kinetic part of the Lagrangian is manifestly invariant under a $U(2)$ flavour symmetry. However, the Yukawa term only respects an $SO(2)$ symmetry. This can be seen by noticing the flavour matrix is symmetric, and only an $SO(2) \subset U(2)$ preserves a symmetric orthogonal combination of the fields.

This construction generalizes straightforwardly to an arbitrary number of flavours:[19]

$$\widehat{Q}^t = \begin{pmatrix} Q_1 & \cdots & Q_{N_f} & \sigma_G\tilde{Q}_1 & \ldots & \sigma_G\tilde{Q}_{N_f} \end{pmatrix} \tag{3.8.12}$$

The kinetic term, then, exhibits a full $U(2N_f)$ flavour symmetry. However, the Yukawa coupling only preserves an $SO(2N_f)$ subgroup. The superpotential becomes in terms of $\widehat{Q}$:

$$W = \frac{1}{\sqrt{2}}\widehat{Q}^t\sigma_G\Phi \begin{pmatrix} 0 & \mathbf{1}_{N_f} \\ \mathbf{1}_{N_f} & 0 \end{pmatrix} \widehat{Q} \tag{3.8.13}$$

This ends our discussion of the global symmetries.

3.8.2 Moduli Spaces of $N_f \leq 3$

A new feature of the flavoured theories is that new singularities appear on the moduli space at which an electric hypermultiplet becomes massless. In particular, for large bare masses $m_i \gg \Lambda$, so that we trust our semiclassical analysis, we have massless hypermultiplets whenever $a = \pm\frac{m_i}{\sqrt{2}}$. The massless quarks will cause a monodromy in the formula of a_D exactly as the formula for a general dyon dictates as in (3.3.12) (with new electric charge normalization), however with an additive constant. More precisely, near $a = a_0 = \frac{m_i}{\sqrt{2}}$ the section reads:

$$a_D \sim c - \frac{i}{2\pi}(a - a_0)\log(a - a_0) \tag{3.8.14}$$

$$a \sim a_0 \tag{3.8.15}$$

where we omit all subleading terms except the constant. Sending $(a - a_0) \to e^{2\pi i}(a - a_0)$, we find the monodromy:

$$a_D \to a_D + a - a_0 \tag{3.8.16}$$

$$a \to a \tag{3.8.17}$$

The fact that monodromies now not only mix the sections $(a_D(u), a(u))$ but also include the addition of constants $a_0 = \frac{m_i}{\sqrt{2}}$ enlarges our duality group and has some remarkable effects. A general monodromy will now act as:

$$\begin{pmatrix} \mathbf{1}_{N_f} & \emptyset \\ C & M \end{pmatrix} \begin{pmatrix} \frac{m_i}{\sqrt{2}} \\ a_D \\ a \end{pmatrix} \tag{3.8.18}$$

[19] We use the transposition t to denote transpositions on both flavour and gauge indices. Which is meant should be clear from the context.

with $M \in SL(2,\mathbb{Z})$ and C a constant integer valued 2×1 vector. Consider the new central charge written as an inner product:

$$Z = \begin{pmatrix} S_i & n_m & n_e \end{pmatrix} \begin{pmatrix} \frac{m_i}{\sqrt{2}} \\ a_D \\ a \end{pmatrix} \tag{3.8.19}$$

Since the central charge should remain invariant under a monodromy M, the charge vector should transform under M^{-1}. It can be checked from the general form of the monodromy matrix in (3.8.18) this implies that not only n_m and n_e will mix among themselves, but also S_i can receive contributions from n_m and n_e. For instance, under a particular monodromy a semiclassical monopole may obtain a non-vanishing S-charge proportional to its magnetic charge. However, as argued by Seiberg and Witten, the range of S-charges of monopoles is bounded.

This follows from the quantization of fermion zero modes of the hypermultiplet fermions. These act on the monopole as gamma matrices, themselves transforming in the vector representation of the flavour symmetry $SO(2N_f)$. Therefore, the monopoles live in a spinor representation of the flavour symmetry group $SO(2N_f)$.[20] Clearly, this is a finite representation. Therefore, there is a maximum amount of flavour charge the monopoles could have obtained from the action of fermion zero modes. This implies that for a certain amount of monodromies, the spectrum has to jump, i.e. the monopole crosses some wall of marginal stability and decays.

Somehow, solitons split up in particles with and without magnetic charge, possibly with a global flavour charge. This strange sounding phenomenon will be needed to understand the singularity structure of the theory. In fact, we have already encountered a similar phenomenon in the pure theory. There, however, the curve of marginal stability appeared at strong coupling and no global flavour charges (dis)appeared.

As in [60], we start with the analysis of $N_f = 3$ moduli space. We look at the theory for large masses $m_i \gg \Lambda$ such that we can analyse the singularity caused by the hypermultiplet semiclassically. For equal hypermultiplet bare masses, we know there is a $U(3)$ flavour symmetry. This means that the singularity appears in a 3 of $SU(3) \subset U(3)$. Flowing down to the strong coupling region, the flavours freeze out and we effectively obtain the pure theory. Thus, at strong coupling we just have the $(1,0)$ monopole and the $(1,2)$ dyon, which do not carry any flavour charge and therefore appear in the singlet 1 of $SU(3)$.

Consider now adiabatically lowering the bare masses towards the strong coupling region. For $m_i = 0$, we expect the hypers to become light near $u \sim \Lambda^2$ in analogy with the pure theory. Furthermore, the flavour symmetry is enhanced to $SO(6)$. The quarks, classically, transform in the 6 dimensional vector representation of $SO(6)$ whereas the monopole and the dyon transform in 4 dimensional spinor representation of $Spin(6) \cong SU(4)$. Furthermore, there is no residual R-symmetry on the moduli space as seen from Table 3.1. These two facts imply a big change in the singularity spectrum. Without knowing the details, the singularity spectrum can be guessed. The only way a 3 and two 1 singularities can combine into representations of $SU(4)$ without losing their global non-abelian charges is as a 4 and a 1. The minimal and coprime quantum numbers associated to these states are $(1,0)$ which is a multiplicity 4 singularity and $(2,1)$ which has multiplicity 1 respectively. For details on this, see Section 5 of [60].

Semiclassically there is no real distinction between the monopole and the dyon point: for $m \gg \Lambda$, near $u \sim \Lambda^2 = (m^3\Lambda_3)^{1/2}$, rotating the m by 2π interchanges the monopole and the dyon. Hence, there is no obvious way to determine which of the singlets has combined with the triplet. We conclude that the moduli space of $N_f = 3$ is much more intricate than that of the pure theory. Also, we have now experienced first hand the transformation of an elementary particles into solitons by letting the hypermultiplet singularities drop to the strong coupling region.

[20] Semiclassically, there are also four bosonic zero modes corresponding to the electric charge and the spatial position of the center of mass of the monopole. These are invariant under the flavour symmetry. However, a dyon with even (odd) electric charge transforms in the spinor (cospinor) representation of $SO(2N_f)$. See [60] for more details. For a brief introduction on the above, see [65].

For $N_f = 2$ a similar story holds. For large values of the bare masses, the hypers are in doublets of the $U(2)$ flavour symmetry, whereas the monopole and dyon point transform as singlets. As the bare masses approach zero, the flavour symmetry is enhanced to $SO(4) \cong SU(2) \times SU(2)$. Also, the R-symmetry has remained as a $\mathbb{Z}_2$ symmetry. The only way the doublet can combine with the singlets to correctly reproduce the singularity spectrum when turning on masses *and* respect the residual R-symmetry when the masses are zero is to split up and combine into a left and right handed doublet with the soliton singularities: $(2,1)$ and $(1,2)$ under $SO(4)$, both singularities with multiplicity 2. To transform as a doublet under the enhanced flavour symmetry, the quantum numbers of the fields are again restricted. The minimal choice corresponds to $(1,0)$ and $(1,1)$.

For $N_f = 1$ the enhanced symmetry is $SO(2)$, i.e. there are only singlets. This is consistent with the residual R-symmetry, which implies a threefold symmetry between the singularities. Due to the R-symmetry and its relation to the Witten effect, the quark singularity must have attained a magnetic charge. A minimal consistent set of quantum numbers would be $(1,0)$, $(1,1)$ and $(1,2)$.

There is a consistency check on the above mentioned arguments for the singularity structure. Namely, instead of viewing the theories separately, one could also try to obtain the singularity structure of the theories $N_f \leq 2$ by turning on one mass at a time in massless $N_f = 3$ and sending it off to infinity. This is done in Section 7 of [60] and indeed agrees with the results obtained above. For instance, we can flow from the massless $N_f = 3$ theory to $N_f = 2$ by turning on m_3. In the 4 singularity, we had 4 monopoles all of which transform in $SO(6) \cong SU(4)$. The non-zero mass lifts two of the zero modes and breaks the flavour symmetry:

$$SO(6) \longrightarrow SO(4) \times SO(2) \;\Leftrightarrow\; SU(4) \longrightarrow SU(2) \times SU(2) \times U(1) \tag{3.8.20}$$

This corresponds to the breaking of the 4 into two doublets: $(2,1,\frac{1}{2})$ and $(1,2,-\frac{1}{2})$ while the singlet singularity remains. Note that the 4 singularity has split due to the opposite charges of the doublets with respect to the $U(1)$. Since we expect a singularity at large u when we increase the mass, we associate the $SO(6)$ singlet, originally the bound state of monopoles $(2,1)$, to have become an elementary quark flowing away to infinity.

3.8.3 $N_f = 4$ Theory

Having seen the qualitative differences between the pure theory and the theory coupled to hypermultiplets, we will discuss the theory of most relevance to the next chapter: the case of $N_f = 4$. This theory is special for a number of reasons.

First of all, if the flavours are massless the one-loop beta function vanishes. In their original paper, Seiberg and Witten assumed: $\tau_{\rm UV}^{SU(2)} = \tau_{\rm IR}^{U(1)}$. This in fact turns out to be true only when $\tau_{\rm UV}^{SU(2)} = i\infty$. If not, there is a non-zero (finite) renormalization as was first shown by explicit calculation in [43]. Although the same authors suggested a new form of the renormalized prepotential in [44], they were unable to calculate the contributions. It took a while, but the contributions were calculated up to infinite order by Nekrasov in [45]. The calculation of the (finite) renormalization for massless $N_f = 4$ is written down explicitly in the appendix of [36]. The relation between the UV and IR gauge couplings is as follows:

$$(2\pi i)\tau_{\rm IR} = (2\pi i)\tau_{\rm UV} - \log 16 + \left(\frac{1}{2}q_{\rm UV} + \frac{13}{64}q_{\rm UV}^2 + \frac{23}{192}q_{\rm UV}^3 + \dots\right) \tag{3.8.21}$$

Here, $q_{\rm UV} = e^{2\pi i \tau_{\rm UV}}$. For now, we conclude that this renormalization is due to the inherent scheme dependence of renormalization which makes a definition of a UV gauge coupling ambiguous, and therefore the relation between the IR and UV is not necessarily as Seiberg and Witten assumed. We will have more to say about this in Section 4.2.

Due to the fact the renormalization is finite and has no scale dependence, for all practical purposes we may treat $\tau_{\rm UV}$ as an exactly marginal operator. This validates the original analysis of $N_f = 4$ of

Seiberg and Witten and all their conclusions, if we read their gauge coupling τ as τ_{IR}.[21]

So let us give the original conclusions on $N_f = 4$. We will write $\tau \equiv \tau_{\mathrm{IR}}$. The section retains its classical form:

$$a_D(u) = \tau a \tag{3.8.22}$$

$$a(u) = \frac{1}{2}\sqrt{2u} \tag{3.8.23}$$

These functions show no monodromy except for the monodromies at infinity and at $a = 0$ which now *is* part of the moduli space. Hence, a is a good global coordinate (except at $a = 0$) and we are able to obtain a global description for $N_f = 4$. Also we note that these solutions do not signify a jumping phenomenon, since τ is independent of a. This gives hope for self-duality of $N_f = 4$: the semiclassical BPS spectrum is protected against decay.

The massless classical theory has an enhanced flavour symmetry group $SO(8)$. Clearly, the normal vector representation of this theory is 8 dimensional, and we will denote it by 8_v. Unlike the previous cases however, the two irreducible spinor representations of $SO(8)$ are 8 dimensional as well. We will denote them as 8_s and 8_c. As already mentioned above, dyons of even electric charge transform 8_s and dyons of odd electric charge transform in 8_c.

The fact that the dimensions of the three irreducible representations match implies an operation triality which permutes the representations among each other. Seiberg and Witten conjecture that $N_f = 4$ theory is self-dual under the action of the semidirect product $\mathrm{Spin}(8) \rtimes SL(2,\mathbb{Z})$.[22] The statement is that the theory is electric-magnetically dual all over the moduli space as long as we choose an appropriate flavour symmetry representation for the fundamental fields under consideration.

Let us point out some differences with the moduli spaces of $N_f \leq 3$. The Coulomb branch of $N_f = 4$ as parametrized by u is of no significance any more to the gauge coupling, since its renormalization is independent of the Higgs expectation value. Instead, there is a simple relation between the UV moduli space and the IR given as above. Moving to the strong coupling region of the IR moduli space cannot be done by changing u as in the $N_f \leq 3$ theories. Rather, we should slightly change τ_{UV} in the UV Lagrangian and then flow to the IR. Starting at $\tau_{\mathrm{UV}} = i\infty$, if all bare masses are equal to zero, all magnetic monopoles and dyons will be much heavier than our electric variables. However, if we change τ_{UV} towards strong coupling, the roles get reversed. Since we know that the IR theories admit duality transformations we can still obtain a sensible Lagrangian in the IR, which again directly relates to the UV. Hence, although in principle we only have proof of duality in the IR, it seems the strong coupling region of the UV moduli (for which we a priori have no Lagrangian description) turns out to have a weakly coupled IR description in terms of differently charged particles. We see that the conformal theories, accompanied by the Seiberg-Witten analysis, can be used as probes for the UV moduli space. Note that this depends crucially on the fact the semi-classical spectrum does not jump, where dyons and magnetic monopole solutions provide the first evidence of a duality in the spectrum. More evidence for the conjectured $SL(2,\mathbb{Z})$ duality can be found by searching for stable (p,q) states in the spectrum, p,q coprime. Indeed, these become light whenever $\tau \sim -\frac{q}{p}$, provided we make sure the flavour symmetry representations are correctly adjusted. We will have something to say on these at the end of this section.

At last, we turn to massive $N_f = 4$ theory. Conformal invariance is broken, the gauge coupling τ *is* renormalized perturbatively and obtains an a dependence[44]. Furthermore, the flavour symmetry is generically broken to $SO(8) \to U(1)^4$, where $U(1)^4$ is associated to the Cartan of $SO(8)$.

[21] In [44] it is also argued we should reinterpret u. We refer the reader to Section 3 of their paper for an explicit expression. We just note that as long as we calculate the prepotential as a function of a, by integrating $a_D(a)$, we escape the need to worry about a redefinition of u. Whenever it appears in the following analysis, we regard it as a dummy variable with no direct relation to the modulus u we used for $N_f \leq 3$.

[22] The permutation group S_3 is isomorphic to $SL(2,\mathbb{Z}_2)$. This provides an action of $SL(2,\mathbb{Z})$ on Spin(8). Most easily, this action is seen on the Dynkin diagram. S_3 permutations of the roots of the Cartan subalgebra correspond to the different representations. See the appendix of [55] for a detailed analysis.

Triality can then be understood as an action on the mass parameters. To see this, consider the two transformations

$$\begin{aligned} m_1 &\to m_1 \\ m_2 &\to m_2 \\ m_3 &\to m_3 \\ m_4 &\to -m_4 \end{aligned} \tag{3.8.24}$$

and

$$\begin{aligned} m_1 &\to \frac{1}{2}(m_1 + m_2 + m_3 + m_4) \\ m_2 &\to \frac{1}{2}(m_1 + m_2 - m_3 - m_4) \\ m_3 &\to \frac{1}{2}(m_1 - m_2 + m_3 - m_4) \\ m_4 &\to \frac{1}{2}(m_1 - m_2 - m_3 + m_4) \end{aligned} \tag{3.8.25}$$

These transformations generate S_3. The claim is that the above transformations should be applied accompanying respectively a T or S transformation on the gauge coupling. We provide circumstantial evidence of the assertion.[23]

We start with a simple example. Consider $m_i = (m, 0, 0, 0)$. For values $m \gg \Lambda$, this theory is described near $u \sim m^2$ by a hypermultiplet coupled to an electric vector multiplet. At strong coupling, it is described by massless $N_f = 3$, i.e. a 4 $(1, 0)$ monopole or a 1 $(2, 1)$ dyon. Suppose now we perform the triality transformation (3.8.24). Clearly, the vector representation which describes the hypermultiplet is invariant nuder this. However, from the perspective of perturbed $N_f = 4$, two of the fermion zero modes are lifted by the non-vanishing mass. This breaks $SO(8) \to SO(6) \times SO(2)$ by a mass term. Replacing $m \to -m$ corresponds to exchanging the spinor and cospinor representation. Hence, this triality transformation should accompany an $SL(2, \mathbb{Z})$ transformation acting trivially on electric variables but shifts the electric charge of monopoles by an odd integer: an odd power of T.

Let us also consider (3.8.25) on our particular state, schematically depicted in Figure 3.8.1. It maps to $m'_i = (m, m, m, m)$. This new parametrization carries exactly the same global symmetries as the previous $m_i = (m, 0, 0, 0)$ since $U(4) \cong SO(6) \times U(1)$, but now the hypermultiplets transform in a 4 and the two singlets are the monopole and dyon from the pure theory instead of the single hypermultiplet and the singlet dyon. We conclude that, since non-abelian flavour changes cannot change, the monopoles of the original description are now described by massive mass degenerate hypermultiplets. The moduli spaces of these theories are identical and we say they are actually dual descriptions related by an S-duality transformation combined with a triality transformation.

A more general example is the equivalence of the global symmetries of (m, m, μ, μ) which has global symmetry $U(2) \times U(2)$ and $(m+\mu, m-\mu, 0, 0)$ which has global symmetry $SO(4) \times U(1) \times U(1)$. One can easily verify that (3.8.25) maps the two parametrizations onto each other. See Figure 3.8.2.

Of particular interest for the next chapter will be to consider a particular (maximal) flavour symmetry subgroup $SO(4) \times SO(4) \subset SO(8)$ of massless $N_f = 4$. The $SO(4)$ subgroups split up into the product product of $SO(4) \cong SU(2) \times SU(2)$. We can associate a mass parameter to each of the $SU(2)$ flavour symmetry subgroups. Notice these mass parameters should be: $\mu_{1,2} = m_1 \pm m_2$ and $\mu_{3,4} = m_3 \pm m_4$. The non-zero bare masses will break the flavour symmetry. However, we can observe what happens when performing (3.8.25) on the slightly perturbed Hamiltonian. We see that the particles with bare mass $\mu_1 = m_1 + m_2$ and $\mu_4 = m_3 - m_4$ are invariant whereas μ_2 and μ_3 are

[23] The masses will be related in a certain way to the weights of the Cartan $SU(2) \times SU(2) \times SU(2) \times SU(2) \subset SO(8)$. For an extensive treatment on the action of triality from a mathematical perspective we refer to the appendix of [55].

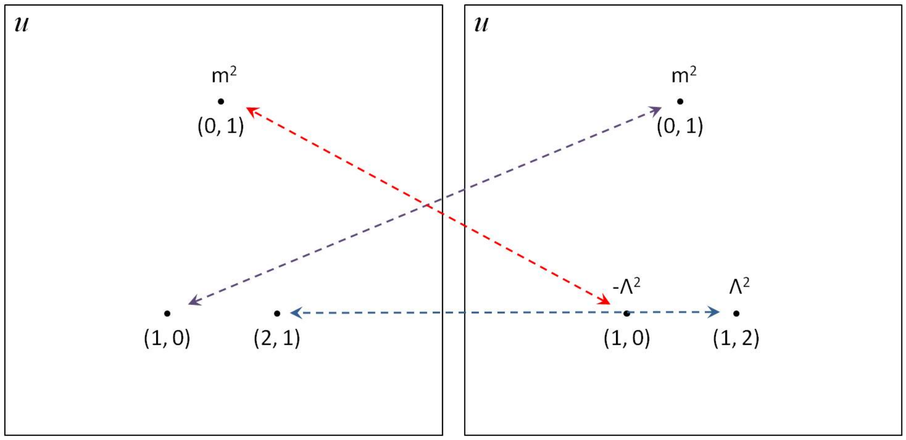

Figure 3.8.1: Singularity structure of the moduli spaces of $N_f = 4$ with $(m, 0, 0, 0)$ (left) and $N_f = 4$ with (m, m, m, m). The moduli space on the left consists of the strong coupling region of massless $N_f = 3$. Hence, the $(2, 1)$ singlet appears with multiplicity one, whereas the $(1, 0)$ singularity has multiplicity 4. At large u, there is a single $(0, 1)$ singularity, which is charged under a $U(1)$ flavour symmetry. The strong coupling region of the moduli space on the right is that of the pure theory. The monopole and dyon singularities have multiplicity 1 and live in the singlet of the global flavour symmetry while the four hypermultiplets at weak coupling appear in a $(0, 1)$ singularity with multiplicity 4. The global symmetries match: $SO(6) \times U(1) \cong U(4)$. The arrows indicate the action of (3.8.25), together with $\tau \to -\frac{1}{\tau}$

exchanged. Changing the sign of m_4 by (3.8.24) we see μ_3 and μ_4 are exchanged. Therefore, we see that triality permutes the flavour symmetry subgroups $SU(2)_i$ among each other. This will be of utmost importance in the next chapter, where we will repeat this statement a bit more carefully.

This completes our global analysis of the Seiberg-Witten solution to the theory coupled to hypermultiplets. From this point on, elliptic curves are introduced such that the moduli spaces of the elliptic curves are in exact agreement with the moduli space of the theory. These are employed to calculate the central charges and prepotential in a similar manner as the pure case.

We conclude this section with two remarks:

1. In the article [56] the existence is checked of monopole and dyon bound states of magnetic charge 2 in $\mathcal{N} = 2$ SYM for $N_f \leq 4$. It was concluded that such a state indeed exists in $N_f = 3$ for odd electric charge, and appears in the singlet of $SU(4)$ precisely as predicted by Seiberg and Witten. Furthermore, $N_f = 4$ contains bound states of magnetic charge 2 for odd and even values of the electric charge, living in the vector and singlet representation of the flavour symmetry respectively. This provides evidence for the expected self-duality of $N_f = 4$, where the states of $(2, 2k)$ are expected to be dual to W-bosons and states of $(2, 2k + 1)$ to hypermultiplets.

2. In a paper by Argyres [6] superconformal theories are embedded into higher rank asymptotically free theories. More precisely, the marginal couplings of the SCFT are regarded as the extra Coulomb branch parameter of the asymptotically free theory. It is shown that the S-duality properties of the SCFT arise from global symmetries on the Coulomb branch of the asymptotically free theory. This is a strong argument in favour of the exactly self-dual properties of an SCFT, not just of its mass spectrum.

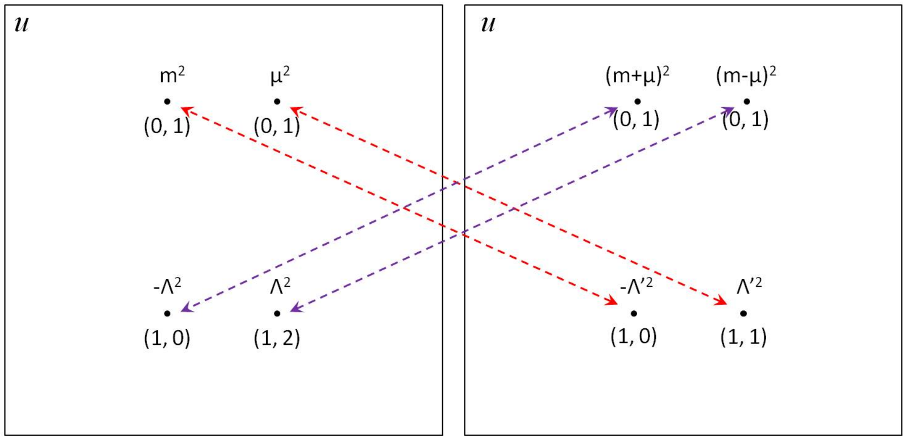

Figure 3.8.2: Singularity structure of the moduli spaces of $N_f = 4$ with (m, m, μ, μ) (left) and $N_f = 4$ with $(m + \mu, m - \mu, 0, 0)$. Notice the strong coupling region of the left moduli space is just the strong coupling region of the pure theory. The monopole and dyon singularities have multiplicity 1 and live in the singlet of the global flavour symmetry while the hypermultiplets at weak coupling live in doublets (have multiplicity 2). The strong coupling region of the moduli space on the right is that of $N_f = 2$. Now the monopole and dyon singularity appear in the doublet of the global symmetry whereas the hypermultiplets live in singlets. The global symmetries match: $U(2) \times U(2) \cong SO(4) \times U(1) \times U(1)$. Note that the strong coupling singularities lie at the scale Λ of $N_f = 0$ respectively Λ' of massless $N_f = 2$.

3.9 Seiberg-Witten Curves from M-Theory

In this section we provide a short recapitulation of the second section of [72], in which a convenient construction in Type IIA string theory and M-theory is provided to determine the Seiberg-Witten curves for a large class of product unitary gauge group theories. For generalizations to other classical Lie groups, see for instance [1].

The Seiberg-Witten curve arose in the context of four-dimensional field theory as an auxiliary geometrical object. In string theory and M-theory, the extra dimensions are employed to embed the curve in actual spacetime. This is achieved in Type IIA string theory by a configuration of D4 and NS5 branes as depicted in Figure 3.9.1. The world volumes of the branes are summarized as:

$$\begin{aligned} \text{D4}:&\ 01236 \\ \text{NS5}:&\ 012345 \end{aligned}$$

Apart from the semi-infinite D4 branes at either end of the chain, the D4 branes are of finite extent in the x_6 direction and therefore will appear macroscopically as objects with three spatial dimensions. At length scales much larger than the separation of the NS5 branes, the NS5 brane dynamics decouple and the brane configuration describes a four dimensional $\mathcal{N} = 2$ gauge theory with gauge group $\prod_{a=1}^{n} SU(k_a)$ where k_a represents the number of D4 branes suspended between the $a - 1^{\text{th}}$ and a^{th} NS5 brane.[24]

The particle spectrum appears as excitations of strings stretching between pairs of branes. For instance, the vector multiplet in the four dimensional theory can be understood as coming from

[24] One might expect a $U(k_a)$ gauge symmetry but a constraint on the positions of the D4 branes on either side of the NS5 brane freezes out an entire $U(1)$ vector multiplet.

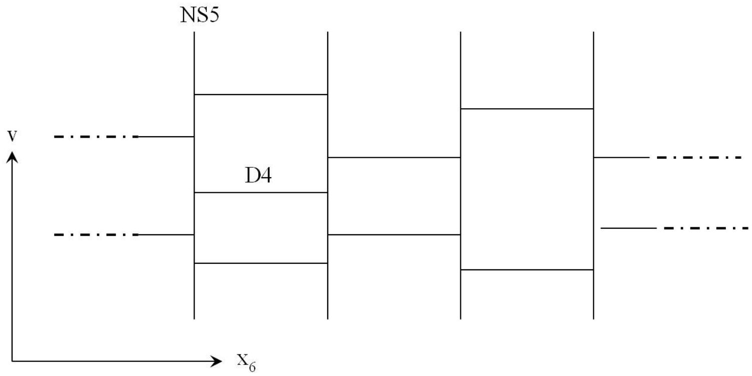

Figure 3.9.1: Type IIA construction of an $SU(3) \times SU(2) \times SU(2)$ gauge theory where the first and last gauge group are both coupled to two fundamentals. The v coordinate is a holomorphic combination of the coordinates (x_4, x_5): $v = x_4 + ix_5$.

strings stretching between D4 branes which are suspended between the same pair of NS5 branes, such that they have Chan-Paton indices associated to the $k_a \otimes \bar{k}_a$, equivalently adjoint, representation. If these D4 branes are coincident, there is a non-abelian gauge symmetry $SU(k_a)$. Generically, we will look at separated D4 branes which spontaneously break the gauge symmetry to the maximal torus. The separation between the D4 branes determines the vacuum expectation value of the Higgs scalar. The massless $U(1)_i$ vector multiplet originates from strings with both ends on the same D4.

Fundamental hypermultiplets arise from strings stretching between the semi-infinite D4 branes and the D4 branes suspended between the first or last two NS5 branes. Similarly, bifundamental hypermultiplets originate from strings stretching between two D4 branes on different sides of an NS5 brane. The (bare) masses of the hypermultiplets depend on the (average) separation of the D4 branes.

This configuration is lifted to M-theory by replacing x_6 with the holomorphic combination $s = x_6 + ix_{10}$, x_{10} being the M-theory S^1 coordinate. This lift is convenient as we will be able to describe the construction in terms of a single object, the M5 brane. Indeed, the NS5 brane and the D4 brane may lift to an M5 brane when the D4 brane is interpreted as an M5 brane wrapping the S^1. Hence, an M5 brane, albeit with a rather complicated worldvolume, could describe the entire Type IIA set-up.[25] In short, it extends in the $\mathbb{R}^{3,1}$ and is fixed on $\mathbb{R}^3$, reflecting respectively Poincaré invariance and the $SU(2)_R$ R-symmetry of the four dimensional theory. Furthermore, it spans a two dimensional surface $\Sigma \subset Q^4$ with $Q^4 = \mathbb{R}^3 \times S^1$ the remainder of spacetime.[26] This surface actually needs to be a Riemann surface for $\mathcal{N} = 2$ supersymmetry to be preserved in the four dimensional theory. For more details on these statements, see [31]. We will see that a polynomial relation between the complex (v, e^{-s}) coordinates of Q^4 will determine the surface. Indeed, this will be identified with the Seiberg-Witten curve.

Since M-theory does not contain strings as fundamental objects, one might wonder how BPS states are included in the theory. This is possible through the inclusion of minimal area M2 branes, or membranes. The following discussion summarizes qualitatively the more elaborate discussions of the inclusion of membranes in [40] and [31]. By similar considerations as above, the M2 branes

[25]The perspective more closely related to Type IIA is that of a collection of sets of k_a parallel M5 branes intersected by transverse M5 branes. Supposing all $k_a = k$ the picture one should have is that of k M5 branes wrapping a cylinder $(x_6, x_{10}) \in \mathrm{R} \times S^1$ and four dimensional spacetime, while being intersected at points $t_1, \ldots, t_n$ by n transverse M5 branes. We will employ this point of view as well, in the next chapter, as it eases the analysis of the (low energy) worldvolume theory.

[26]The given form of Q^4 holds for unitary gauge groups, the case we are interested in.

should be embedded holomorphically into Q^4 with respect to some complex structure. This complex structure should be orthogonal to the complex structure through which the $\Sigma \subset Q^4$ is embedded.[27] Note that the M2 brane in this way constitutes a worldline in $\mathbb{R}^{3,1}$ and spans a two dimensional surface in $\mathbb{R}^3 \times S^1$, while being fixed on the remaining $\mathbb{R}^3$. We will call this surface D and denote its boundary by $\partial D = \Gamma$. Assuming the winding number of Γ along S^1 is zero, we can associate the closed loop Γ to a homology class $[\gamma] \in H^1(\Sigma, \mathbb{Z})$. Indeed, the homology class will correspond to the charge of a BPS particle with respect to a particular (combination of) $U(1)$ gauge group(s).

The surface area of an M2 brane with boundary Γ is calculated as:

$$S(\Gamma) = \int_D \mathrm{d}S \geq \int_D |\mathrm{d}s \wedge \mathrm{d}v| \geq |\int_D \mathrm{d}s \wedge \mathrm{d}v| = |\int_\Gamma v(s)\mathrm{d}s| \tag{3.9.1}$$

Here, $\mathrm{d}S$ is the natural volume form on D, the first inequality follows from the fact the volume form splits up into two positive (semi-)definite forms (see Section 2 of [40] for an explicit calculation or Section 3 of [31] for a rather general treatment), the second inequality is just the triangle inequality and in the last equality Stokes' theorem was used. We write $v(s)$ because v will be determined in terms of s through the defining equation for Σ.

The two conditions are equivalent to the requirement the surface D is holomorphic with respect to a complex structure orthogonal to that of Σ and that the phase of $\mathrm{d}s \wedge \mathrm{d}v$ on D is constant. Hence, the minimization of the membrane surface respects $\mathcal{N} = 2$ supersymmetry of the four dimensional theory while the requirement of a constant phase is the geometric realization of the saturation of the BPS bound. We conclude that minimal area membranes correspond to BPS states in the four dimensional theory. The differential $v\mathrm{d}s$ is accordingly identified with the Seiberg-Witten differential and the central charge is given by:

$$Z_\Gamma = \int_\Gamma v\,\mathrm{d}s \tag{3.9.2}$$

Having found the general construction of $\mathcal{N} = 2$ gauge theories in M-theory, let us highlight a couple of physically interesting properties and provide a curve for a general product gauge theory.

The gauge coupling $\tau_a = \frac{\theta_a}{2\pi} + \frac{4\pi i}{g_a^2}$ of the gauge group between the $a-1^{\text{th}}$ and a^{th} NS5 brane can be expressed in terms of the coordinates (v, s):

$$-i\tau_a = s_a(v) - s_{a-1}(v) \tag{3.9.3}$$

Here, the v dependence of s_a is *not* determined by the curve equation. Rather, the positions of the NS5 branes depend on v due to the fact the D4 branes create dimples in the NS5 brane. Only far away from the D4 brane disturbances, the $v \to \infty$ limit, the brane diagram of Figure 3.9.1 should be taken seriously. The dependence of s_a on v is calculated to be:

$$s_a(v) = (k_{a+1} - k_a)\log v + c \tag{3.9.4}$$

representing the pulling and pushing of D4 branes on either side of the NS5 brane. Note that the NS5 brane seems only well defined in the limit $v \to \infty$ when $k_a - k_{a-1} = 0$. When this requirement is not met, the NS5 brane position $s_a(v)$ diverges. However, Witten gave this divergence a satisfying physical interpretation.

Namely, if we plug (3.9.4) into (3.9.3) we obtain the well known one-loop formula for the gauge coupling of a theory with gauge group $SU(k_a)$ coupled to $k_{a+1}+k_{a-1}$ flavours. Hence, the divergence of the NS5 brane position is identified with the usual divergence of the one-loop gauge coupling τ of an asymptotically free theory at weak coupling.

[27] If both manifolds are embedded with respect to the same complex structure, in particular their intersection will be holomorphically embedded with respect to the complex structure. This implies the intersection will be a one complex dimensional curve instead of a one real dimensional curve, which is the case we will be interested in. The fact that such an orthogonal structure indeed exists stems from the fact that Q^4 is hyper-Kähler. The argument is taken from [31].

Since $\mathrm{Im}\, s = x_{10}$ is a periodic variable, in the following we will use a single-valued form of s given by $t = e^{-s}$. In terms of t the gauge coupling becomes:

$$\tau_\alpha = \frac{1}{i} \log\left(\frac{t_a}{t_{a-1}}\right) \tag{3.9.5}$$

Notice that in this case weak coupling corresponds, as expected, to $x_6^{a-1} \to -\infty$ ($t_{a-1} \to \infty$) or $x_6^a \to \infty$ ($t_a \to 0$), i.e. infinite separation of the NS5 in the Type IIA picture. As x_6 generically has non-trivial v dependence as in (3.9.4), we note that weak coupling corresponds to $v \to \infty$ except when the theory is conformal.

As motivated in [72], a general form of the Seiberg-Witten curve is given by:

$$t^{n+1} + f_1(v)t^n + \ldots + f_n(v)t + 1 = 0 \tag{3.9.6}$$

with $\deg(f_a) = r+1$ with r the rank of the a^{th} gauge group or, equivalently, the number of D4 branes between $a-1^{\text{th}}$ and a^{th} NS5 brane. The v dependent roots of the polynomial in t are interpreted as the positions of the NS5 branes, while the roots of the f_i correspond to the positions of D4 branes between the $i-1^{\text{th}}$ and the i^{th} NS5 brane. The inclusion of semi-infinite D4 branes on the right (left) side of the brane diagram is achieved by taking for the lowest (highest) degree coefficients polynomials in v of appropriate degree. We will see explicit examples in the next chapter. The two highest degree coefficients of the polynomials f_i parametrize the UV gauge couplings and the bare masses of the bifundamentals respectively, while the other r parameters parametrize the Coulomb branch.

We conclude with a simple example of pure $SU(2)$ theory and compare the M-theory results to the original results of Seiberg and Witten. The curve is given by:

$$t^2 + (v^2 + u')t + 1 = 0 \tag{3.9.7}$$

Note a linear term in v is absent, as appropriate because of the absence of bifundamentals. Indeed, for any curve one can use the freedom to shift the coordinate v to cancel one unphysical bare mass parameter.

The curve can be rewritten more conveniently for comparison as:

$$\tilde{t}^2 = \frac{1}{4}(v^2 + u')^2 - 1 \tag{3.9.8}$$

With $\tilde{t}$ a shifted version of t and u' a Coulomb branch parameter. Rescaling $y = \Lambda^2 \tilde{t}$ and $x = \frac{\Lambda}{\sqrt{2}} v$ and redefining $u = -\frac{\Lambda}{\sqrt{2}} u'$, we obtain a more familiar form:

$$y^2 = (x^2 - u)^2 - \Lambda^4 \tag{3.9.9}$$

Although this is not the original curve Seiberg and Witten used, this form turned out to be more convenient for generalization and was first derived in [2] and [9]. It is equivalent to the curve derived originally by Seiberg and Witten in their second paper [60], where they used different conventions suited to the inclusion of hypermultiplets. To see this, one can show that for some suitable scalings of u and Λ, the roots of the curves are mapped onto each other via a Möbius transformation.

Chapter 4

Gaiotto Dualities

A great step forward in our understanding of the structure of $\mathcal{N} = 2$ gauge theories, especially how to understand and find dual descriptions of a particular SCFT, was only made recently by Davide Gaiotto in [29]. Essentially, Gaiotto has discovered an underlying structure in the Seiberg-Witten curves which turns out to be a very general feature of a large class of $\mathcal{N} = 2$ SCFTs.

In the first section we will introduce quiver gauge theories, all of type A_1, and discuss their S-duality properties qualitatively. The analysis relies on the original argument of Seiberg and Witten for S-duality of $N_f = 4$ theory and a careful analysis of the behaviour of the global symmetries under S-duality.

After this, we will introduce a geometric perspective on these quiver theories. In particular, the quiver diagrams can be understood as certain Riemann surfaces $C_{n,g}$ with n punctures and g holes. Punctures will be associated to flavour symmetries, whereas tubes will be seen to correspond to gauge groups. This picture culminates in a conjecture by Gaiotto: the moduli space of exactly marginal couplings is equal to the moduli space of these Riemann surfaces. The S-duality group is identified with the mapping class group MCG$[C_{n,g}]$ of the Riemann surface. In the process, a full renormalization of the UV gauge coupling for $N_f = 4$ is obtained using a well known geometric relation between the modular parameter of an elliptic curve and the cross ratio of a four-punctured Riemann surface.

In the third section, we will provide a detailed analysis of the quantitative form of these statements by constructing the Seiberg-Witten curves for the quiver gauge theories using methods introduced in Section 3.8. We will consider explicit expression of these curves in the $g = 0, 1$ case and shortly comment on the higher genus generalization. Along the way, we will explicitly work out examples and provide consistency checks on the statements made.

In the fourth section, we will discuss the six-dimensional construction of the quiver theories. The basic elements of this construction were already discussed in Section 3.9. However, the framework of Gaiotto puts the construction in a slightly different light, as we will see.

At last, we will discuss the generalization to higher rank gauge groups. We will focus in particular on the A_2 case, while briefly commenting on A_{N-1}.

The main source for this chapter is the original article by Gaiotto [29]. A nice and instructive, and definitely more accessible review has recently been given by Tachikawa in [65]. Also, a lecture given by Gaiotto at the PiTP summer school in 2010 provides a good introduction to the original paper. This video lecture is freely accessible at `http://www.sns.ias.edu/pitp2/`. Further references will be given along the way.

4.1 Quiver Gauge Theories

In this section, we will introduce the Sicilian quiver diagrams which we will use to represent our gauge theories. Some concepts which are essential to this construction, such as enhancement of flavour symmetry, the bifundamental and trifundamental representations and gauging flavour indices, are explained in Section 3.8 and in Appendix C respectively.

A quiver diagram provides a convenient way to represent a gauge theory with a product gauge group coupled to matter in a particular representation. We know that four dimensional (non-gravitational) theories with $\mathcal{N} = 2$ supersymmetry are described by vector multiplets and hypermultiplets. To draw a quiver diagram for $\mathcal{N} = 2$ theories, then, requires knowledge of the product gauge group and the matter representations:

$$\prod_a G_a, \quad \bigoplus_i n_i \left(R_i \oplus \bar{R}_i\right) \tag{4.1.1}$$

A Lagrangian description can be given whenever there is a weakly coupled (dual) frame. The form of this Lagrangian is essentially the one introduced in Sections 2.2 and 2.3. In this section, we will only consider products of $SU(2)$ gauge groups coupled to four flavours in the fundamental representation. An example of such a theory is described by the quiver diagram in Figure 4.1.1. The circles denote gauge groups whereas the squares represent fundamental flavours coupled to the

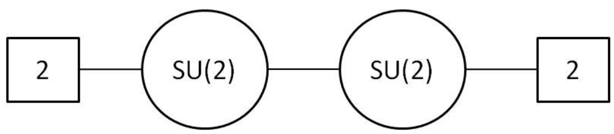

Figure 4.1.1: Quiver representation of an $SU(2) \times SU(2)$ gauge theory. Both gauge groups are coupled to two fundamental flavours and one bifundamental, effectively providing both gauge groups with four flavours.

particular gauge group. The number in the square denotes the number of flavours or, equivalently, the number of hypermultiplets. Furthermore, lines in between gauge groups denote matter in a bifundamental representation of the gauge groups. A general bifundamental of the gauge group $SU(N_1) \times SU(N_2)$ transforms in the $(N_1, \bar{N}_2) \oplus (\bar{N}_1, N_2)$. This corresponds to a single hypermultiplet in the bifundamental representation.[1]

The quiver theory in Figure 4.1.1 is superconformal. To see this, note that the bifundamental hypermultiplet effectively provides two fundamental hypermultiplets to both of the gauge groups. Furthermore, both gauge groups are supplied with an additional two fundamental flavours, such that they both see four fundamental hypermultiplets.

An important observation in the case of $SU(2)$ gauge symmetry is that a bifundamental, sitting in a real representation of the gauge group, has an enhanced flavour symmetry of $USp(2) \cong SU(2)$, while two fundamental flavours with vanishing bare mass have an enhanced flavour symmetry of $SO(4) \cong SU(2) \times SU(2)$.[2] These enhancements suggest a new, Sicilian quiver diagram in which also the flavour symmetries are manifest.[3] The original quiver diagram as depicted in Figure 4.1.1 is

[1] Conventional quiver diagrams use directed arrows to distinguish between representations and their complex conjugates. With these conventions, our single line would have to be replaced by two opposite pointing arrows representing the two chiral multiplets which make up the hypermultiplet. Keeping this in mind, our notation is unambiguous and we stick to it for its simplicity.

[2] The tensor product of two pseudoreal representations is real. This follows most easily from the fact the $SU(2) \times SU(2)$ invariant relation: $Q^*_{ij} = \epsilon_{im}\epsilon_{jn}Q^{mn}$, which provides a non-trivial reality condition. The discussion on the enhancement of flavour symmetry in Section 3.8 readily applies to real representations. In this case, the flavour matrix will be antisymmetric. Therefore the enhanced symmetry is now $USp(2) = U(2) \cap Sp(2, \mathbb{C})$. See [34].

[3] This nomenclature was proposed in [22], because the trivalent vertex bears a resemblance to a triskelion, a figure dominating the Sicilian flag. And probably the homestead of the author has something to do with it.

now represented as in Figure 4.1.2. The crucial idea is to identify the trivalent vertex as a building

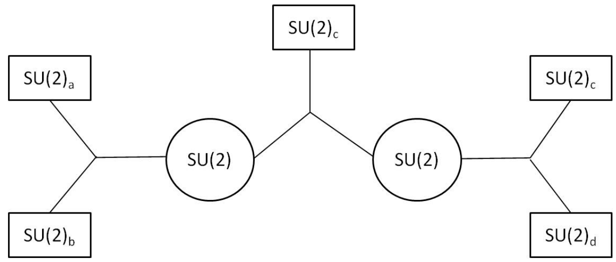

Figure 4.1.2: Sicilian quiver representation of the quiver in Figure 4.1.1. Not only the gauge symmetry, but also the flavour symmetry is explicitly shown.

block for general SCFTs with $SU(2)$ gauge groups. If one diagonally gauges the flavour symmetries of two trivalent vertices, the resulting theory is superconformal. Repeating this process produces arbitrary trivalent networks of gauge theories with exactly marginal couplings. Let us identify the distinct ways such a trivalent vertex may appear in a gauge theory. This is discussed in detail in Appendix C.

1. The first possibility is that the vertex represents a half hypermultiplet in the trifundamental representation $2_1 \otimes 2_2 \otimes 2_3$ of three gauge groups $SU(2)_1 \times SU(2)_2 \times SU(2)_3$. This effectively provides four half hypermultiplets in the fundamental representation to each of the gauge groups.

2. Taking one of the gauge couplings $\tau_1 \to i\infty$, the gauge symmetry $SU(2)_1$ becomes a global symmetry $SU(2)_a$. We will be left with two half hypermultiplets in the bifundamental representation of the remaining gauge groups: $2_a \otimes (2_2 \otimes 2_3)$. A possible mass parameter for the bifundamental corresponds to the Coulomb branch parameter of the very weakly coupled gauge group: $m_a^2 = u_1$. Its vanishing corresponds to an enhanced flavour symmetry of $USp(2)$ for the bifundamental, in correspondence with the enhanced gauge symmetry of $SU(2)$ when the Coulomb branch parameter vanishes.

3. Coupling a second gauge group very weakly gives us four half hypermultiplets in the fundamental representation of the remaining gauge group: $(2_a \otimes 2_b) \otimes 2_3$. We recognize the enhanced flavour symmetry $SU(2)_a \times SU(2)_b$ of the fundamental flavours.

4. The simplest and last interpretation is that of eight half hypermultiplets not coupled to any gauge field. It does carry a flavour symmetry $SU(2)_a \times SU(2)_b \times SU(2)_c$. We will call this object the anfundamental.

The list given here is represented in Figure 4.1.3. We presented the results in the order in which they were derived in the appendix. We will build our theories starting from the anfundamentals and gauging diagonal subgroups of pairs of $SU(2)$ flavour symmetries.

Let us see how this works for the known superconformal field theory $\mathcal{N} = 2$ $SU(2)$ with $N_f = 4$. We take two anfundamentals, both associated to three bare mass parameters whose combinations correspond to the bare masses of the eight half hypermultiplets: Q_{ijk}, Q'_{lmn}.[4] Now we weakly gauge one of the flavour groups of both anfundamentals. The bare masses associated to the gauged flavour group takes the role of the Coulomb branch parameter u of the vector multiplet. Taking the diagonal of the gauged flavour symmetry essentially glues the anfundamentals together. If the remaining bare masses are put to zero, the (manifest) flavour symmetry is $SO(4) \times SO(4)$. In fact, we know the

[4]More details can be found in Appendix C.

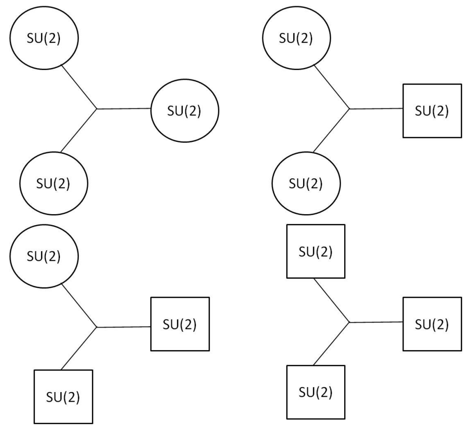

Figure 4.1.3: From the upper left corner, left to right, the Sicilian quiver diagrams corresponding to a trifundamental, a bifundamental, a fundamental and an anfundamental are shown. Theories built by gauging diagonal subgroups of the flavour symmetry of a pair of anfundamentals provide automatically SCFTs.

flavour symmetry of $N_f = 4$ to be larger: $G_f = SO(8)$. The choice to look at only this particular subgroup is, as we have seen above, the possibility to easily construct SCFTs. Furthermore, the full Cartan of $SO(8)$ is still visible in this subgroup, which is the flavour symmetry subgroup of importance for non-vanishing bare masses.

The quiver representation for $N_f = 4$ is drawn in Figure 4.1.4. As already noted in Section 3.8,

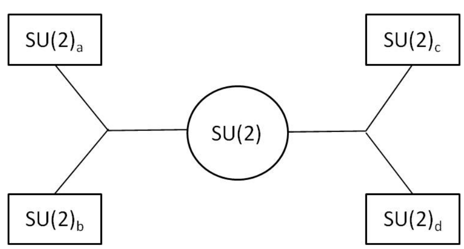

Figure 4.1.4: Quiver representation of $N_f = 4$ theory. Note the arrangement of flavour symmetry groups. This particular flavour symmetry arrangement corresponds to the vector representation 8_v.

this theory is $SL(2,\mathbb{Z})$ invariant as long as we combine duality transformations with triality. It was also shown that triality permutes the bare masses corresponding to the $SU(2)$ flavour symmetry groups. In the notation of the previous sections, we would associate $(\mu_a, \mu_b, \mu_c, \mu_d)$ to $SU(2)_a \times SU(2)_b \times SU(2)_c \times SU(2)_d$ respectively. Whereas Figure 4.1.4 represents an electric description of four hypermultiplets coupled to an (electric) vector multiplet, a suitable description for $\tau = i\infty$, in Figure 4.1.5 the monopole and dyon descriptions are shown which are suitable descriptions near $\tau = 0$ and $\tau = 1$ respectively.

We now proceed to build more complicated quiver gauge theories. At least in the limit where all gauge groups but one are very weakly coupled we should retrieve the familiar S-duality of the $N_f = 4$ theory. However, it is a priori not clear whether the full theory also has an S-duality acting on it. Naively, one could guess the moduli space of an $SU(2) \times SU(2)$ theory, where both gauge groups are coupled to four fundamental flavours, to be the Cartesian product of the separate moduli

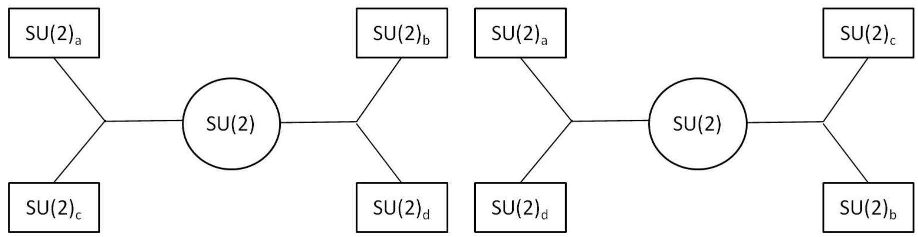

Figure 4.1.5: Quiver diagram for $N_f = 4$ theory. The flavour symmetry arrangements correspond to the spinor representation 8_s and the cospinor representation 8_c respectively.

spaces: $\mathcal{M} = \mathbb{H}/SL(2,\mathbb{Z}) \times \mathbb{H}/SL(2,\mathbb{Z})$. As already found in [7] by Argyres this turns out to be too naive a guess. The S-duality groups of the separate $SU(2)$ factors do not commute. However, there does exist an S-duality group for the product gauge theory as was originally found in the same paper.[5]

We can demonstrate the non-commutativity of the S-duality groups rather easily by using the formulation of Gaiotto. Let us assume the gauge coupling of the second gauge group is very weak, i.e. $\tau_2 = i\infty$. Then we may effectively view it as decoupled, leaving us with a single $SU(2)$ gauge group with flavour symmetry $SU(2)_a \times SU(2)_b \times SU(2)_c \times SU(2)_2$. The representation of the remaining matter is thus: $2_1 \otimes (2_a \otimes 2_b \oplus 2_c \otimes 2_2)$. Keeping the gauge coupling τ_2 very weak, we may bring τ_1 to the strongly coupled region. To arrive at a weakly coupled dual description at the cusps $\tau = 0, 1$ we should permute the flavour symmetry groups. This is just the $N_f = 4$ triality, leaving us with the representations $2_1 \otimes (2_a \otimes 2_c \oplus 2_b \oplus 2_2)$ and $2_1 \otimes (2_a \otimes 2_2 \oplus 2_c \otimes 2_b)$.

We could now imagine doing the same while keeping the first gauge coupling very weak. In Figure 4.1.6 part of the S-dualities are depicted. Note that every quiver diagram corresponds to some weakly coupled cusp in the moduli space. Therefore, at each dual description we can safely perform an S-duality at either node to obtain a new weakly coupled description of the theory. One can verify that the naive expectation of commuting S-duality groups turns out to be false: $1 \circ 2 \neq 2 \circ 1$. However, the weakly coupled theories *are* related by an S-duality transformation, as long as we do not pay attention to the gauge group labelling. This can be justified in the sense that a gauge symmetry does not represent a physical global symmetry of the theory but rather eliminates unphysical degrees of freedom. We conclude it is impossible to keep track of the gauge group while moving around the UV moduli space.

Performing all possible permutations on the masses associated to the $SU(2)^5$ flavour symmetry is clearly a closed process. This would correspond to $5 \cdot \binom{4}{2} = 30$ different weakly coupled descriptions of a single theory. Notice that we do not count purely T-dual theories since these theories do not represent a different physical description but rather comprise the Witten effect. Not surprisingly, yet a rather beautiful mathematical fact is that this number of vertices precisely corresponds to an icosidodecahedron, which is an icosahedron with truncated vertices such that the closed figure consists of twenty identical triangles and twelve identical pentagons. This polyhedron represents the boundary of the UV moduli space.

We see that the naive assumption of the moduli space to be a product of two upper half planes was far too simple, which would only have corresponded to a total of $3 \cdot 3 = 9$ weakly coupled descriptions. The original conclusion of Argyres in [7] about the S-dual properties of this theory is that twenty distinct weakly coupled descriptions exist. This is precisely the number of vertices in a dodecahedron, consisting of twelve pentagonal faces. If combined with triality, this turns into an

[5] The methods Argyres used were quite different from Gaiotto's construction. Therefore, comparing the statements would provide a consistency check on either construction. We will see below that their conclusions agree.

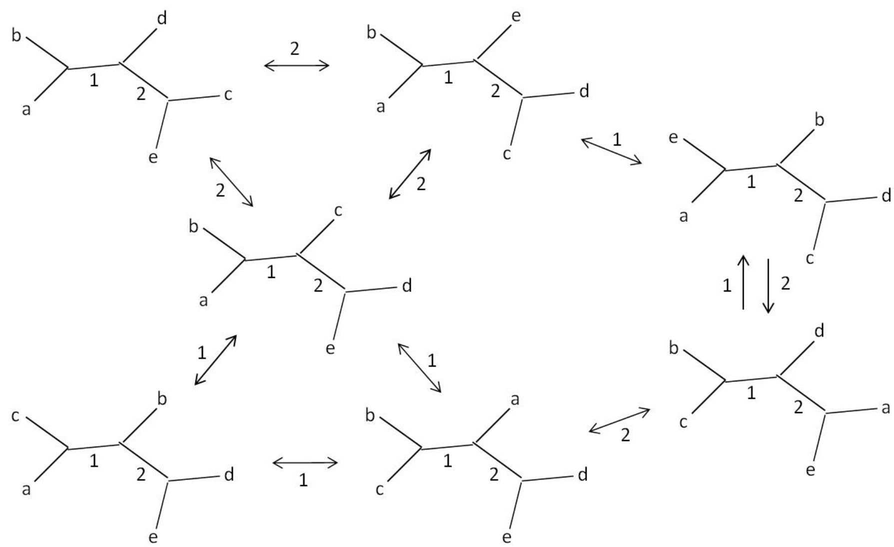

Figure 4.1.6: S-dualities of $SU(2)_1 \times SU(2)_2$. In the middle left side is the theory with two weakly coupled gauge groups. Going up or down, we identify the action of triality on either one of the gauge groups, while performing S or ST transformation on the gauge coupling. The pentagon on the right side shows the non-commutativity of the separate S-duality groups of the gauge groups.

icosidodecahedron, by truncating vertices, in accordance with Gaiotto's construction. All in all, we conclude Gaiotto's method provides a simple strategy to exhibit (all) S-dual descriptions of a single theory.

Before turning to the geometric interpretation, we consider an additional example since, if Gaiotto's construction is correct, it predicts an interesting new aspect of duality. This happens when we consider three gauge groups: $SU(2)_1 \times SU(2)_2 \times SU(2)_3$. We have represented it in Figure 4.1.7. This diagram is labelled to correspond to weak coupling of all three gauge couplings:

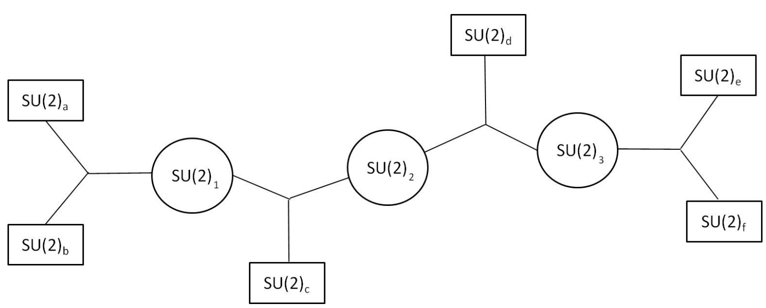

Figure 4.1.7: Quiver diagram of the superconformal $SU(2)_1 \times SU(2)_2 \times SU(2)_3$ theory.

$(\tau_1, \tau_2, \tau_3) = (i\infty, i\infty, i\infty)$. We now play a similar game as in the previous quiver. First we hold τ_2 arbitrarily weak such that we effectively have two separate $N_f = 4$ theories with representations $2_1 \otimes (2_a \otimes 2_b \oplus 2_c \otimes 2_2)$ and $2_3 \otimes (2_d \otimes 2_2 \oplus 2_e \otimes 2_f)$. Performing S-duality transformations at both edges permutes the mass parameters μ_a, μ_b, μ_c and μ_d, μ_e, μ_f among themselves. This is similar to the previous example. However, we may also imagine coupling both $SU(2)_1$ and $SU(2)_2$ arbitrarily weak. The representations of the hypermultiplets are then: $2_2 \otimes (2_1 \otimes 2_c \oplus 2_d \otimes 2_3)$. Performing an electric-magnetic, or in this case an electric-dyon exchange, we obtain a fully gauged trifundamental: $2_2 \otimes (2_1 \otimes 2_3 \oplus 2_d \otimes 2_c)$ when reinstating the gauge groups. This corresponds to the quiver depicted in Figure 4.1.8.

Although the total flavour symmetry does not change, we have unexpectedly obtained very different matter representations in a certain corner of the UV moduli space. This is a beautiful example suggesting a more general perspective on duality. Triality and $SL(2,\mathbb{Z})$ transformations do not only permute mass parameters with different couplings in the IR, also matter representations may have their flavour symmetry exchanged for an extra gauge symmetry.

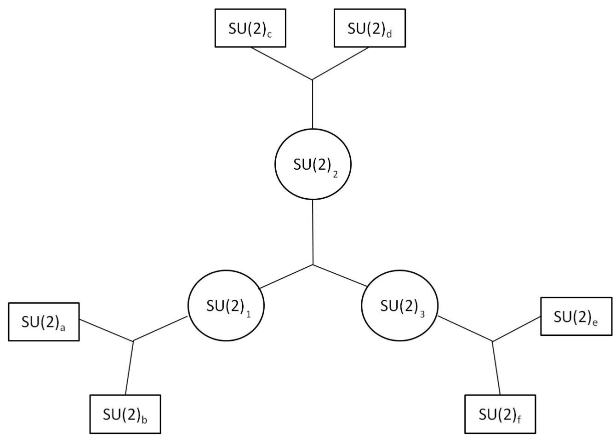

Figure 4.1.8: Quiver representation of $SU(2)_1 \times SU(2)_2 \times SU(2)_3$ theory. Notice that at this cusp in the moduli space, instead of two bifundamentals and 4 fundamentals, we have one trifundamental and six fundamental flavours.

The above mentioned peculiarities are part of an exhaustive list of dualities that can appear in product gauge theories with gauge group $SU(2)$. From this point on, we can consider any particular linear quiver as depicted in or even quivers with loops. There is a simple formula to calculate the dimension of the UV moduli space, in the $SU(2)$ case also equal to the dimension of the Coulomb branch. We denote the theory associated to the quiver $\mathrm{T}_{n,g}[A_1]$ where n denotes the number of external legs ($SU(2)$ flavour symmetries) and g the number of loops in the quiver, then:

$$\dim \mathcal{M}(\mathrm{T}_{n,g}[A_1]) = n - 3 + 3g \tag{4.1.2}$$

To obtain such a theory, one has to glue together $n-2+2g$ anfundamentals. Gaiotto has conjectured that *every* superconformal $SU(2)$ quiver theory with g loops and n flavour groups is some weakly coupled description of a *single* theory $\mathrm{T}_{n,g}[A_1]$. Furthermore, the space of exactly marginal couplings of these theories, or the UV moduli space, corresponds to the moduli space of a Riemann surface with n punctures and g holes. We will make this correspondence more precise in the following section.

We conclude this section by noting that the degeneration of theory $\mathrm{T}_{n,g}[A_1]$ can occur in two distinct ways: $\mathrm{T}_{n,g}[A_1] \to \mathrm{T}_{n+2,g-1}[A_1]$ in the case of breaking of a loop or $\mathrm{T}_{n,g}[A_1] \to \mathrm{T}_{n'+1,g'}[A_1] \times \mathrm{T}_{n-n'+1,g-g'}[A_1]$ in the case of breaking of a linear piece. The cusp at which all gauge groups are weakly coupled results in a degeneration of the theory into $n-2+2g$ anfundamentals. We will analyse these degenerations quantitatively in Section 4.3.

4.2 A Geometric Realization of Duality

In this section the relation between the quiver theories and Riemann surfaces will be explained. We will introduce some necessary mathematical concepts relating to moduli spaces of Riemann surfaces and elliptic functions. Due to the low dimensionality of these geometrical objects, we can approach the subject in a conceptual (and visual) manner. However, the formal theory behind these concepts is vast. The reader is referred to the literature for a solid mathematical understanding. For a review on Riemann surfaces (and algebraic curves) and their differential geometry of particular use and not too assuming were [41] and [24]. These two references also give a good review on the Riemann-Roch theorem, which will be used as well in this section. A pedagogical review on the mapping class group can be found in Part 1 of [23]. Part 2 of the same book describes in detail the Riemann moduli space, Teichmüller space, quadratic and Beltrami differentials and the interconnections between these concepts. See [18] for a classical discussion on elliptic and modular functions. For the cutting and sewing of Riemann surfaces, see for instance Section 9.3 of [51].

As mentioned at the end of the previous section, we can associate the tuple (n, g) to a quiver theory. The number of flavour symmetries is denoted by n, while g corresponds to the number of loops in the quiver. If we fatten up the lines making up the quiver, we obtain a genus g Riemann surface with n punctures. The anfundamentals are identified with thrice punctured Riemann spheres, a limit of a pairs of pants in which the boundary components shrink to a point. See Figure 4.2.1. Gauging diagonal flavour symmetries of anfundamentals is now interpreted as the sewing of thrice

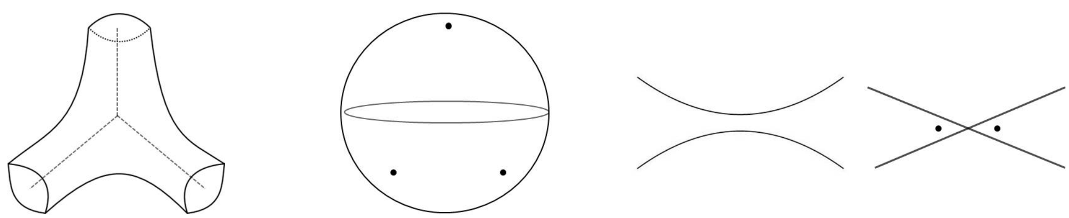

Figure 4.2.1: On the left a pairs of pants and a thrice punctured sphere are depicted. These represent the geometrical analogue of the anfundamentals and can be glued together to form more general Riemann surfaces. On the right, a tube is depicted, whose geodesic is of length $l \sim |q|$. When $|q| \to 0$ the tube pinches and two disjoint pieces of the original Riemann surface are left, both with an extra puncture.

punctured spheres. More precisely, we take on two thrice punctured spheres local coordinates z_1, z_2 near a puncture, cut out a disc around each puncture and sew the two surfaces together through the sewing equation $z_1 z_2 = q$. Here, q is a complex parameter whose absolute value defines the radius of the cut-out discs and its phase corresponds to a possible twist. For instance, applying this procedure to two thrice punctured spheres leads to a four punctured sphere. Furthermore, it follows that if $|q| \to 0$ the four punctured sphere degenerates to two thrice punctured spheres.

Sometimes a degenerating Riemann surface is said to consist of very long tubes connecting the thrice punctured spheres. This picture is equivalent to the picture above; consider the regions around the punctures of the thrice punctured spheres to be $|z_1| < 1$ and $z_2 < 1$ in which the the cut-out discs of radius q lies. The annulus $\{|q| < |z_i| < 1\}$ represents the regions which are to be identified according to the sewing equation $z_1 z_2 = q$. We can perform the coordinate transformation $z_1 = e^{-iw}$ and $z_2 = qe^{iw}$ such that in terms of w the annulus is mapped to $\{\log |q| < \operatorname{Im} w < 0\}$ and $w \simeq w + 2\pi$. In terms of w then, the tube connecting the spheres becomes a cylinder whose length diverges in the $|q| \to 0$ limit. This is just the conformal, or holomorphic, map from a disc to a semi-infinite cylinder via the complex logarithm.

Yet another point of view which we will frequently use is that the degeneration of a surface can be interpreted as the collision of a number of punctures. This is a consequence of the fact that after the sewing the coordinates of punctures, except for three which are fixed by conformal transformations, depend on q. The limit in which $|q| \to 0$ can for instance correspond to the collision of a number of punctures with coordinates $z_i \sim q$ with $z = 0$.

From the considerations above, it it clear that the complex structure of the surface depends on q. This determines q as a so-called Fenchel-Nielsen coordinate on Teichmüller space, which is the moduli space of marked Riemann surfaces up to $SL(2,\mathbb{C})$ transformations. The complex dimension of this space is:[6]

$$\dim \mathcal{T}_{n,g} = n - 3 + 3g \tag{4.2.1}$$

This dimension corresponds precisely to the dimension of the UV moduli space of the quiver theories with n external legs and g loops. This is no coincidence: just as in the construction of quiver theories from anfundamentals, a Riemann surface $C_{n,g}$ can be constructed from $n-2+2g$ thrice punctured spheres by connecting them with $n-3+3g$ tubes. To each tube is related a complex parameter q_i, a Fenchel-Nielsen coordinate, whose absolute value determines the radius of the tube and whose phase determines the twist. These tube parameters exhaust all free parameters: if we disconnect the tubes, we will be left with $n-2+2g$ thrice punctured spheres, which are rigid and therefore carry no free parameters. Hence, the Teichmüller space of $C_{n,g}$ is of complex dimension $n-3+3g$. The sewing parameters will be related to the exactly marginal gauge couplings of the quiver theory, where a degenerating Riemann surface will correspond to the decoupling limit of a gauge group in the quiver.

We can now make Gaiotto's conjecture more precise: the moduli space of the theory $\mathrm{T}_{n,g}[A_1]$, the parameter space of exactly marginal (UV) gauge couplings, equals the Riemann moduli space $\mathcal{M}_{n,g}$ which is related to Teichmüller space as:

$$\mathcal{M}_{n,g} = \mathcal{T}_{n,g}/\mathrm{MCG}(C_{n,g}) \tag{4.2.2}$$

with $\mathrm{MCG}(C_{n,g})$ the mapping class group of the Riemann surface which is interpreted as the group of S-dualities of the theory: it permutes punctures and acts on closed curves or arcs on the Riemann surface through Dehn twists. As discussed, the first evidence for this conjecture is the fact that the dimensions match. In the next section we will provide quantitative evidence for the conjecture. However, let us first investigate a couple of simple examples qualitatively.

We start, as usual, with $N_f = 4$ theory. The moduli space of this theory is shown in Figure 4.2.2. The shaded domain on the left shows the combined $SL(2,\mathbb{Z})$ and triality operations, exhibiting full self-duality, while the other shows the moduli space without the triality operation, exhibiting the three weakly coupled cusps. We will give two independent arguments why these are also the moduli spaces of the four punctured sphere $C_{4,0}$ with equivalent or inequivalent punctures respectively. The first argument will not use the formula (4.2.2). The reason we do mention it is because through it, we obtain a full (finite) renormalization of the 'exactly marginal' gauge coupling.

As shown in Figure 4.2.3, there are three possible degenerations of a four-punctured sphere. Through the cutting and sewing operations q appears as a coordinate of the fourth point on the sphere. For $q \to 0, 1, \infty$ we obtain the three possible degenerations of the four punctured sphere into two thrice punctured spheres. We have performed coordinate transformations such that a degeneration always corresponds to the collision of a point with $z = 0$. We have $q \to 0$, $q' = 1 - q \to 0$ or $q'' = \frac{1}{q} \to 0$ corresponding to a weakly coupled electric, magnetic and dyonic description respectively. The relation with the quiver is obvious. Taking a decoupling limit in the quiver corresponds to the degeneration of the Riemann surface. It seems natural to relate: $q(d) = e^{2\pi i \tau(d)}$ with d denoting a particular dual frame. In fact, we will find this relation from the M-theory construction in the next section. We warn the reader that this relation only holds at weak coupling.

[6]This formula is defined for all Riemann surfaces except $C_{1,0}, C_{2,0}, C_{3,0}, C_{0,1}$. Our quiver theories automatically fulfil these requirements.

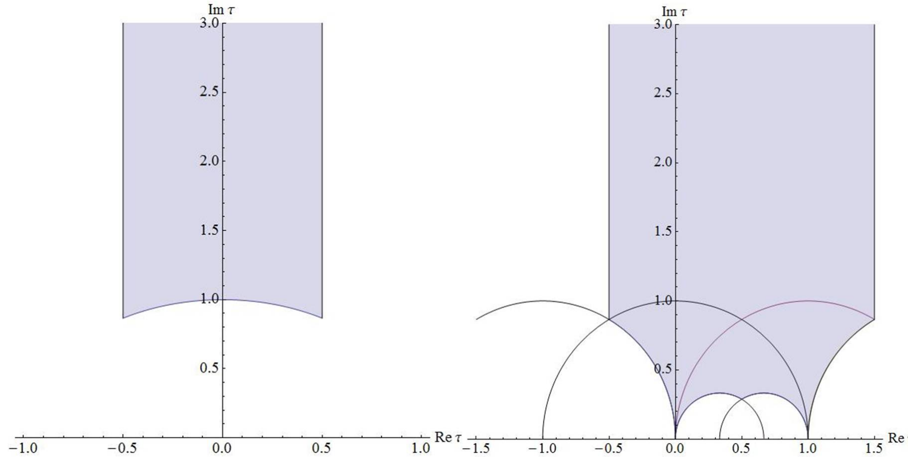

Figure 4.2.2: Fundamental domains of $SL(2,\mathbb{Z})$ and $\Gamma(2)$. The three cusps in $\mathbb{H}/SL(2,\mathbb{Z})$ represent the electric, magnetic and dyonic weakly coupled descriptions.

A conformal invariant is given by the cross ratio of the punctures. We can take the cross ratio to be:

$$q = (q, 1; 0, \infty) \tag{4.2.3}$$

This will be the primary object of interest in the next section, as it provides a convenient parametrization of the gauge couplings independent of $SL(2,\mathbb{C})$ transformations on the punctures. Note that it is defined on the Riemann sphere. Therefore, we can use it as a convenient object only for $g = 0$ quivers.

One can still consider S-duality transformations on this object, however its implementation changes slightly. Instead of $SL(2,\mathbb{C})$ coordinate transformations one now permutes the punctures to obtain a dual description. For instance, permuting 1 and 0 or 0 and ∞ can be seen to correspond to the magnetic and dyonic description and accordingly. These two permutations generate the anharmonic group, isomorphic to S_3. Let us recall the appropriate dual descriptions for the monopole and dyon in terms of the IR gauge coupling τ and compare them with the appropriate, i.e. $q(d) \to 0$, cross ratios:

$$S: \; \tau \to -\frac{1}{\tau} \quad \Rightarrow \quad q \to 1 - q \tag{4.2.4}$$

$$STS: \; \tau \to \frac{\tau}{1-\tau} \quad \Rightarrow \quad q \to \frac{1}{q} \tag{4.2.5}$$

If we take this implication seriously, the above behaviour shows q is invariant under a $\Gamma(2)$ action on τ. There exists a natural function of τ with the exact same behaviour that q exhibits: the modular $\lambda(\tau)$ function. In fact, $\lambda(\tau)$ provides a biholomorphic map between $q \in \widehat{\mathbb{C}} \backslash \{0, 1, \infty\}$ and $\tau \in \mathbb{H}/\Gamma(2)$. Therefore, the moduli space of the Riemann surface is in one-to-one correspondence with the moduli space of the gauge coupling τ. We conclude we have found an alternative, geometric realization of S-duality on the gauge coupling through the anharmonic group action on the cross ratio. Note that if we were to take the punctures equivalent, i.e. forget about triality, we would take analogously: $q = j(\tau)$ which is invariant under a full $SL(2,\mathbb{Z})$ action on τ. Then, $j(\tau)$ would be a modulus of the Riemann surface and indeed provides a biholomorphic map between $q \in \mathbb{C}$ and $\mathbb{H}/SL(2,\mathbb{Z})$.

In the next section we will argue that the Seiberg-Witten curve of this theory is given as a double cover of the Riemann surface. The identification $q = \lambda(\tau)$ appears naturally in this context

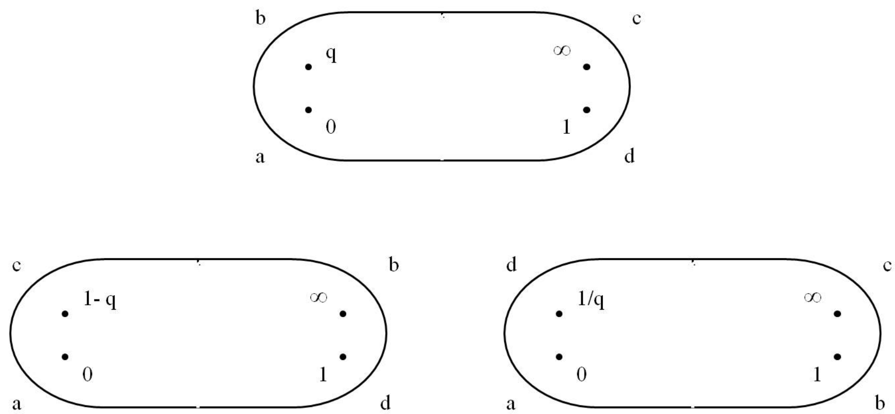

Figure 4.2.3: The three distinct ways in which the Riemann surface may degenerate. We have performed $SL(2,\mathbb{C})$ transformations such that in the coordinates at hand, the sewing parameter always collides with 0. The four points are labelled in the same way as the original quiver: while the upper diagram describes the electric description, the left diagram describes a weakly coupled dyon of $n_e = 2k$ and the right diagram a weakly coupled dyon of $n_e = 2k+1$.

through a well-known relation between the cross ratio of punctures on a Riemann surface and the cross ratio of the Weierstrass half-periods of the elliptic curve considered as double cover of the same surface. This is briefly mentioned in Section 2.1 of [32]. As will be elucidated in the next section, this identification relates a UV parameter q, the cross ratio of the punctures of $C_{n,0}$, with an IR parameter τ, the modulus of the Seiberg-Witten curve. At least for massless $N_f = 4$, we will see that this purely geometric IR-UV relation encompasses the full finite renormalization alluded on in Section 3.8. It can be directly compared to Nekrasov's explicit calculation in [45], and for example in the appendix of [36] found to be equivalent.

The more general calculation of the moduli space of a Riemann surface can be done with (4.2.2). For the case of $N_f = 4$ the Teichmüller space $\mathcal{T}_{4,0} \cong \mathbb{H}$. The mapping class group of a four-punctured sphere is $PSL(2,\mathbb{Z})$ if we take the punctures to be equivalent and $\Gamma(2)/\pm\mathbf{1}$ if we take the punctures to be inequivalent.[7] The distinction stems from the fact that the mapping class group may permute punctures. If the punctures are equivalent, these permutations are irrelevant. However, if the punctures are inequivalent the a and b cycles transform non-trivially under permutations and only a $\Gamma(2)$ democracy survives, much alike the pure Seiberg-Witten monodromies. More generally, it can be shown that for $C'_{n,0}$ and $C_{n,0}$ the Riemann surfaces with inequivalent and equivalent punctures respectively,

$$\mathrm{MCG}(C_{n,0})/\mathrm{MCG}(C'_{n,0}) \cong S_n$$

Note that in the case of $n = 4$ we found the permutation group is in fact $SL(2,\mathbb{Z})/\Gamma(2) \cong S_3$. This anomaly is explained by the isomorphism $S_4/\mathbb{V} \cong S_3$ with $\mathbb{V}$ the Klein group. The division by the Klein group follows from the fact that the cross ratio is invariant under this group. For $n > 4$ there will also be symmetries but they will not constitute a normal subgroup.

Another example for which we can easily see the conjecture holds is the case of $\mathcal{N} = 4$ SYM, or equivalently $\mathcal{N} = 2$ SYM coupled to one hypermultiplet in the adjoint representation. The corresponding quiver and Riemann surface are depicted in Figure 4.2.4. We first note that unlike

[7] See Section 2.2.5 of [23] for a detailed derivation of the mapping class group of the four-punctured sphere. Note we interpret $C_{4,0} = C_{0,1}/\mathbb{Z}_2$, with $\mathbb{Z}_2$ denoting the hyperelliptic involution of the torus. Indeed, this action fixes four points, corresponding to the punctures of $C_{4,0}$.

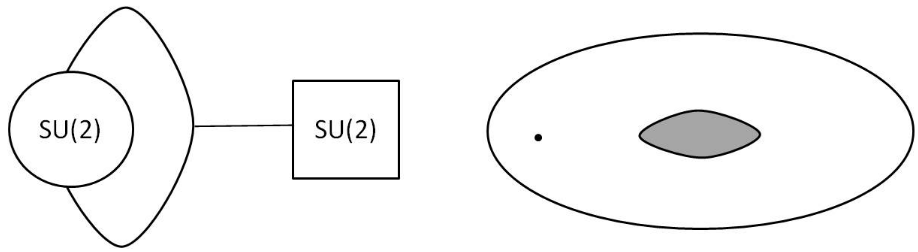

Figure 4.2.4: Quiver diagram for $\mathcal{N} = 4$ SYM and the corresponding Riemann surface. The remaining flavour symmetry show it is a real representation, as required for the adjoint representation.

the massless Seiberg-Witten curves[60], the Riemann surface (Gaiotto curve) is clearly different for $\mathcal{N} = 4$ and $N_f = 4$ theory. This is already a nice result, although it demonstrates only a tiny part of the power of Gaiotto's construction as we will see in later sections.

The quiver is constructed by taking one anfundamental and gauging a diagonal subgroup of two flavour symmetries of the same anfundamental: $SU(2)_a \times SU(2)_b \times SU(2)_c \to SU(2)_a \times SU(2)_1$. The original anfundamental representation of the flavour symmetry groups changes accordingly:

$$2_a \otimes 2_b \otimes 2_c \to 2_a \otimes (3_1 \oplus 1_1)\,.$$

Hence, we obtain one hypermultiplet in the adjoint of $SU(2)_1$, while another hypermultiplet appears in the singlet and is completely decoupled. Effectively, this then is $\mathcal{N} = 2^*$ theory: $\mathcal{N} = 4$ SYM perturbed with a bare mass term associated to the adjoint hypermultiplet. It is well known that the Lagrangian corresponding to this theory has an $SU(4)_R$ R-symmetry. The presence of this symmetry is not manifest in Gaiotto's construction but can be derived from the $SU(2)_R$ and flavour symmetry. See for instance Section 2.1.2 of [65] or a generic review on AdS/CFT.

$\mathcal{N} = 4$ SYM is expected to be fully $SL(2,\mathbb{Z})$ invariant without the need for triality. Therefore, its moduli space, characterized by the familiar gauge coupling τ, is given by $\mathbb{H}/SL(2,\mathbb{Z})$. The Riemann surface corresponding to this theory is a once punctured torus. The once punctured torus and an ordinary torus are very much alike. In particular, their Teichmüller spaces and mapping class groups coincide.[8] The Riemann moduli space of the torus is: $\mathcal{M}_{0,1} = \mathbb{H}/SL(2,\mathbb{Z})$. Hence, we see that this example works out consistently.

Now that we have seen two illustrations of the conjecture of Gaiotto, let us delve into the equivalence of the moduli spaces a little more by comparing the exactly marginal deformations of the gauge theory with deformations of the Riemann surface. An exactly marginal deformation on the $\mathcal{N} = 2$ theory is implemented by:

$$\mathcal{L} \to \mathcal{L} + \delta\tau_i \int \mathrm{d}^2\theta\, \mathrm{d}^2\tilde{\theta}\ \mathrm{Tr}\,\Psi_i^2 + c.c. \tag{4.2.6}$$

with Ψ_i denoting the $\mathcal{N} = 2$ vector multiplet associated to the i^{th} gauge group. Equivalently, an exactly marginal deformation corresponds to the insertion of the operator $\mathrm{Tr}\,\Psi_i^2$ into a correlation function. Hence, if one considers marginal deformations of the gauge theory as above, arbitrary correlators in the gauge theory should be adjusted as:

$$\langle X \rangle \to \langle \mathrm{Tr}\,\Psi_i^2 X \rangle \tag{4.2.7}$$

[8]The mapping class group requires a base point. For $C_{1,1}$ we can choose this point at the puncture. It is easy to see the simple closed curves of $C_{1,1}$ precisely match those of $C_{0,1}$, and therefore their mapping class groups are equivalent.

We will identify the prepotential term with $u_i = \langle \mathrm{Tr}\, \Psi_i^2 \rangle$. Hence, an insertion of the Coulomb branch parameter in a correlator is related to the deformations of the gauge theory. We will come back to this statement at the end of this section. However, first we very briefly discuss the deformations of a Riemann surface.

Deforming a marked Riemann surface means a change of position in Teichmüller space $\mathcal{T}_{n,g}$. Two Riemann surfaces which differ by a conformal transformation define the same point in Teichmüller space. A non-trivial deformations of a Riemann surface is achieved by the use of a quasiconformal transformation, which maps circles to ellipses. Quasiconformal maps, in fact, play a primary role in the infinitesimal deformation of a Riemann surface, and are therefore closely related to the tangent space of $\mathcal{T}_{n,g}$. More precisely, sections of the tangent space of $\mathcal{T}_{n,g}$ are represented by $(-1,1)$ differentials on the Riemann surface. These are called Beltrami differentials and have the following form:

$$\mu \equiv \mu(z)\frac{\mathrm{d}\bar{z}}{\mathrm{d}z} = \frac{f_{\bar{z}}\mathrm{d}\bar{z}}{f_z \mathrm{d}z} \tag{4.2.8}$$

with f a quasiconformal map. Note that $f_{\bar{z}}$ should be non-vanishing in order to constitute a non-trivial deformation. This is in accordance with the well known relation between holomorphic and conformal maps on Riemann surfaces.For the calculation of a finite deformation the natural candidate that appears is a quadratic or $(2,0)$ differential, locally represented as $\phi = \phi(z)\mathrm{d}z^2$ where $\phi(z)$ has simple poles at the punctures. These are elements of the cotangent bundle of Teichmüller space. The Riemann-Roch theorem can be used to find that indeed the dimension of the space of quadratic differentials coincides with the dimension of Teichmüller space. An explicit formula for this dimension, as well as the dimension of k-differentials with prescribed poles, can be found in [17].

On the Riemann surface, the quadratic differential contracts with Beltrami differential to a volume form on the Riemann surface:

$$\langle \mu, \phi \rangle = \int_\Sigma \phi(z)\frac{f_{\bar{z}}}{f_z}|\mathrm{d}z|^2 \tag{4.2.9}$$

As we have seen, S-duality has a geometric realization on the Riemann surface. The information encoded in $\langle \mu, \phi \rangle$ is therefore manifestly S-duality invariant. It will turn out that the Seiberg-Witten curves are naturally represented by quadratic differentials with simple poles at the punctures. There exists a particular form of the Beltrami differential on a collar (a long tube), found in [74] and used in this context in [16].[9] The Coulomb branch parameters will be packaged by these quadratic differentials in such a way that when a quadratic differentials is integrated against a Beltrami differential, the Coulomb branch parameter comes out.

This completes the connection between the gauge theory deformations and the deformations of the Riemann surface. In the previous references more details can be found, although the first describes it in a non-physical context, while the second describes in context of the AGT relation. In the latter, the energy momentum tensor of the CFT will be identified with a quadratic differential on the gauge theory side.

4.3 Mathematical Description

In this section we will quantify the claims made in the previous sections by constructing the Seiberg-Witten curves for the quiver theories. Through an unorthodox reparametrization of the curves, an underlying structure of the curve is found: the Gaiotto curve. We start by revisiting the $N_f = 4$ massless and massive theory, show the curves as constructed in M-theory and comment on properties.

[9] Actually, these references describe the Beltrami differential containing a factor z, while the quadratic differentials contain a factor z^{-2}. We redefine the quadratic differential to contain only a factor z^{-1}, a simple pole, while having the Beltrami differential independent of z. This modification should be harmless; for one, the inner product is insensitive to it.

Next to the fact that the S-duality properties of this theory are used to deduce S-duality properties of all other (A_1) quiver theories, its construction also exhibits much of the details we will encounter in more general quiver theories. We follow the original article of Gaiotto [29], while also borrowing elements from Section 9 of the review [65].

4.3.1 Massless $T_{4,0}[A_1]$

The curve for massless $N_f = 4$ is given in complex coordinates $(v, t) \in \mathbb{C} \times C_{4,0}$:

$$v^2t^2 + c_1(v^2 - u)t + c_2v^2 = 0 \tag{4.3.1}$$

A more convenient form for our purposes is obtained by collecting powers of v:

$$(t-1)(t-q)v^2 = ut \tag{4.3.2}$$

In this parametrization, the four punctures of $C_{4,0}$ will be seen to correspond to $t = 0, q, 1, \infty$. Notice t and v are rescaled to eliminate one combination of the constants c_1 and c_2. The other combination, parametrized by q, cannot be scaled away and is a dimensionless modulus of the theory. This is a manifestation of the fact we are considering a conformal field theory. The UV gauge coupling is determined by q. More precisely, from Section 3.9 we have:

$$\tau \equiv \frac{\theta}{\pi} + \frac{8\pi i}{g^2} = \frac{1}{i\pi}\log q \tag{4.3.3}$$

where we have adjusted the normalization by a factor of π for later convenience. Notice that weak coupling corresponds to $q \to 0$. Let us rewrite the curve in terms of $x = \frac{v}{t}$:

$$t(t-1)(t-q)x^2 = u \tag{4.3.4}$$

The corresponding Seiberg-Witten differential is given by:

$$\begin{aligned}\lambda &= v(t)\frac{\mathrm{d}t}{t}\\ &= \frac{\sqrt{u}}{\sqrt{t(t-1)(t-q)}}\mathrm{d}t\\ &= x\mathrm{d}t\end{aligned} \tag{4.3.5}$$

For the time being, we have chosen a positive root of v. Up to some multiplicative factor

$$\frac{\partial\lambda}{\partial u} = \frac{\mathrm{d}t}{\sqrt{ut(t-1)(t-q)}} \tag{4.3.6}$$

is a holomorphic differential. This is most easily seen by introducing the auxiliary curve:

$$y^2 = ut(t-1)(t-q) \tag{4.3.7}$$

This is a more familiar form of the elliptic curve and we indeed see (4.3.6) corresponds to the unique holomorphic differential on the curve.[10] Note that in this case, already (4.3.5) is a holomorphic differential with respect to the curve since it differs only by a rescaling from $\partial_u\lambda$. This is a consequence of conformal invariance, which implies:

$$\tau_{U(1)} = \frac{\partial a_D}{\partial a} = \frac{a_D}{a} = \tau \tag{4.3.8}$$

[10] The appearance of the auxiliary curve may seem unnatural at this point. In the following, when we discuss more general quivers, the origin of the auxiliary curve and its role will become clear. For now, it should be regarded as a convenient curve encapturing the modular parameter of the Seiberg-Witten curve.

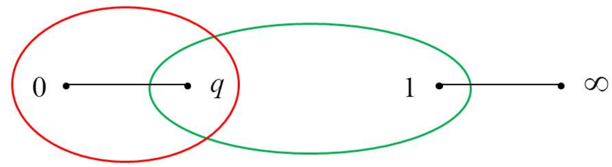

Figure 4.3.1: The canonical basis of a and b cycles on the auxiliary curve (4.3.7). We define the orientations of the cycles such that $\mathrm{Im}\tau > 0$.

The trivial u dependence is special to massless $N_f = 4$. Upon generalization, we will always derive an auxiliary curve from the Seiberg-Witten curve which encodes all low energy physics and depends on all Coulomb branch and mass parameters. The $\partial_{u_i}\lambda$ provide a basis of holomorphic differentials, whose periods determine the IR gauge couplings.

Let us briefly derive (4.3.3) from the auxiliary curve. Defining the a and b cycle on this curve in the usual way, see Figure 4.3.1, we find for $q \to 0$:

$$b_1 = \oint_a \frac{\mathrm{d}t}{\sqrt{ut(t-1)(t-q)}} \sim \oint_{t=0} \frac{\mathrm{d}t}{t\sqrt{u(t-1)}} = \frac{2\pi i}{i\sqrt{u}} \tag{4.3.9}$$

$$b_2 = \oint_b \frac{\mathrm{d}t}{\sqrt{ut(t-1)(t-q)}} = 2\int_1^q \frac{\mathrm{d}t}{\sqrt{ut(t-1)(t-q)}} \sim \frac{2}{i\sqrt{u}}\log q \tag{4.3.10}$$

We can read of the complex structure of the Seiberg-Witten curve:

$$\tau_{U(1)} = \frac{b_2}{b_1} = \frac{1}{\pi i}\log q \tag{4.3.11}$$

This is in accordance with (4.3.3). We conclude that the Seiberg-Witten curve at weak coupling implies no renormalization and find $\tau_{U(1)} = \tau$, and we have $q = e^{i\pi\tau}$ justifying the relation between the sewing parameter and the UV gauge coupling in Section 4.2. However, as already mentioned in previous sections, for more generic values of τ there is a finite renormalization. The exact renormalization is obtained by interpreting q as the cross ratio of the punctures on the base manifold $C_{4,0}$ and relating it in a natural way to the modulus of its double covering, the elliptic curve, by $q = \lambda(\tau_{U(1)})$. This relation can be inverted to obtain the full renormalization of the low energy coupling. Part of the expansion can be found at the end of Section 3.8.

We want to stress that $q = e^{i\pi\tau}$ is only a UV parameter. As such, it cannot be defined unambiguously due to the inherent scheme dependence of renormalization. We understand it as an object merely parametrizing the space of exactly marginal gauge couplings. The Seiberg-Witten modulus is the true physical gauge coupling. Its relation to this particular UV parameter is just a convenient scheme which has a geometric interpretation. In the massless case, we obtain a direct relation between the UV and IR gauge coupling. However, this relation will hold in the massive case only when $u \gg m_i$, such that the gauge coupling has long stopped running before the masses become important. In this case, the geometric relation provides a finite renormalization between two UV couplings obtained through different schemes: the UV coupling as read from the Seiberg-Witten curve at large u, $\tau = \lim_{u\to\infty}\tau_{U(1)}$, and the UV coupling as calculated from the Riemann surface cross ratio. More about this point can be found in Section 9.2 of [65] and references therein.

From the previous section, we have learned a twofold geometric realization of S-duality. Either, one may permute punctures and interpret the cross ratio as parametrizing the gauge coupling. Another equivalent perspective was provided as the change of coordinates and the interpretation of the coordinates of the punctures as parametrizing the gauge couplings. Both perspectives are useful.

The latter perspective is implemented on the Seiberg-Witten curve as:

$$
\begin{aligned}
(t,x) &\to (1-t,x)\\
(t,x) &\to \left(\frac{1}{t},-t^2x\right)
\end{aligned}
\tag{4.3.12}
$$

Under these transformations we obtain similar curves as in (4.3.4) with transformed coordinates of the punctures: $(q,1;0,\infty) \to (1-q,0;1,\infty)$ and $(q,1;0,\infty) \to (1/q,1;\infty,0)$. Note that the Seiberg-Witten differential retains its canonical form under these transformations.[11]

Gaiotto performs one more coordinate transformation, a general $SL(2,\mathbb{C})$ transformation on $t \to \frac{at+b}{ct+d}$, to bring the puncture at infinity to a finite value. Under this transformation we also have: $x \to (ct+d)^2x$. The cross ratio of the points remains invariant. The curve becomes:

$$
x^2 = \frac{u}{\Delta_4(t)} \tag{4.3.13}
$$

This is not the most convenient form to perform explicit calculations. However, we will also use this notation, for two reasons: the general form does allow one to keep track of the general structure of the curve. Furthermore, it provides a great ease of notation in the more general quiver theories. The subscript of the polynomial in t denotes its degree. The Seiberg-Witten differential is given by: $\lambda = x\mathrm{d}t$. This is the simplest example of the generic form of curves we will find for $\mathrm{T}_{n,g}[A_1]$ theories: $x^2 = \phi_2(t)$ with Seiberg-Witten differential $\lambda = \sqrt{\phi_2}\mathrm{d}t$.

From the discussion in Section 4.2, we know the natural representative of sections of the cotangent bundle of Teichmüller space are quadratic differentials on the Riemann surface with simple poles at the punctures. Indeed, (4.3.13) provides us with the natural candidate:

$$
\lambda^2 = \frac{u}{\Delta_4(t)}\mathrm{d}t^2 \tag{4.3.14}
$$

The Seiberg-Witten curves are reinterpreted as sections of the symmetric square of the cotangent bundle $T^*C_{4,0}$ of a Riemann surface, or equivalently as sections of the cotangent bundle of Teichmüller space where (x,t) specify the coordinates in the fiber and the base manifold respectively. Note that this is consistent with the original coordinates $(v,t) \in \mathbb{C}\times C_{4,0}$ in the trivial bundle, which however does not reflect the extra structure of the cotangent bundle. This additional structure is responsible for the better handle on the UV-IR relation, already experienced in examples above, through the geometric realization of S-duality which tells us how the Seiberg-Witten curves change when we choose a different coordinate frame on the base surface.

4.3.2 Massive $T_{4,0}[A_1]$

Before generalizing this construction to arbitrary quiver gauge theories, let us also look at the curve with non-zero bare masses:

$$
(v-m_1)(v-m_2)t^2 + c_1(v^2-u)t + c_2(v-m_3)(v-m_4) = 0 \tag{4.3.15}
$$

As in the massless case, we collect powers of v:

$$
\begin{aligned}
(a\tilde{t}^2+\tilde{t}+1)v^2 &= \left(a(m_1+m_2)\tilde{t}^2+m_3+m_4\right)v - \left(am_1m_2\tilde{t}^2-u\tilde{t}+m_3m_4\right)\\
a(\tilde{t}-t_+)(\tilde{t}-t_-)v^2 &= M_2(\tilde{t})v - U_2(\tilde{t})
\end{aligned}
\tag{4.3.16}
$$

[11] One may observe that x behaves as a modular form of weight 2. This is nothing new: modular forms were also employed as coefficients in the Seiberg-Witten curve of $N_f = 4$ in the original article [60]. We have now simply integrated them within the coordinates.

where we have rescaled $\tilde{t} = \frac{c_1}{c_2} t$, $a = \frac{c_2}{c_1^2}$ which will be related to the cross ratio and in the second line we abbreviated notation according to [29]. Note that all mass parameters are described by $M_2(\tilde{t})$. Furthermore, $U_2(\tilde{t})$ contains all Coulomb branch parameters and, at the highest and lowest power of $\tilde{t}$ contains bare mass parameters of the fundamental hypermultiplets. As will be justified in more detail below when considering the general case, the full flavour symmetry of the theory becomes manifest in the curve only when we shift away the linear term in v. We obtain:

$$v^2 = \frac{-a(\tilde{t}-t_+)(\tilde{t}-t_-)U_2(\tilde{t}) + \frac{1}{4}M_2^2(\tilde{t})}{a^2(\tilde{t}-t_+)^2(\tilde{t}-t_-)^2} \tag{4.3.17}$$

Under this shift, the Seiberg-Witten differential becomes:

$$\lambda_\pm = \pm\frac{\sqrt{-a(\tilde{t}-t_+)(\tilde{t}-t_-)U_2(\tilde{t}) + \frac{1}{4}M_2^2(\tilde{t})}}{a(\tilde{t}-t_+)(\tilde{t}-t_-)}\frac{\mathrm{d}\tilde{t}}{\tilde{t}} + \frac{M_2(\tilde{t})}{2a(\tilde{t}-t_+)(\tilde{t}-t_-)}\frac{\mathrm{d}\tilde{t}}{\tilde{t}} \tag{4.3.18}$$

This differential has simple poles at $\tilde{t} = 0, \infty, t_+, t_-$. It is not difficult to check that the residues will correspond to linear combinations of the mass parameters and do not depend on the Coulomb branch parameter u, as desired. Moreover, the second term in the differential contributes to the residues in a manner precisely as the first square root term and is independent of the Coulomb branch parameter. Therefore, concerning the low energy theory, the second term only constitutes a shift in flavour charge proportional to the number number of cycles, i.e. the gauge charges. Through a redefinition of the flavour charges we can put this term to zero. We are left with the following differential:

$$\lambda_\pm = \pm\frac{\sqrt{-a(\tilde{t}-t_+)(\tilde{t}-t_-)U_2(\tilde{t}) + \frac{1}{4}M_2^2(\tilde{t})}}{a(\tilde{t}-t_+)(\tilde{t}-t_-)}\frac{\mathrm{d}\tilde{t}}{\tilde{t}} \tag{4.3.19}$$

In the limit of $a \to 0$ the residues of this differential have a clear interpretation:

$$\begin{aligned}
\mathrm{Res}_{t=0}\,\lambda_\pm &= \pm\frac{m_3 - m_4}{2} \\
\mathrm{Res}_{t=\infty}\,\lambda_\pm &= \pm\frac{m_1 - m_2}{2} \\
\mathrm{Res}_{t=t_-}\,\lambda_\pm &= \pm\frac{m_1 + m_2}{2} + \mathcal{O}(a) \\
\mathrm{Res}_{t=t_+}\,\lambda_\pm &= \pm\frac{m_3 + m_4}{2} + \mathcal{O}(a)
\end{aligned} \tag{4.3.20}$$

These are the Cartan weights of the $U(1)^4 \subset SO(8)$ flavour symmetry subgroup, corresponding to the masses of the eight half hypermultiplets. See Section 3.8. We note that the reason the linear combinations of masses at $t = 0, \infty$ do not receive corrections is because the separate contributions to the two masses precisely cancel. The limit $a \to 0$ corresponds to weak coupling. To see this, let us rewrite the differential:

$$\lambda^2 = \frac{P_4(t)}{(t-q)^2(t-1)^2}\frac{\mathrm{d}t^2}{t^2} \tag{4.3.21}$$

where $q = \frac{t_+}{t_-}$, we absorbed all constants in $P_4(t)$ and we again rescaled all $\tilde{t}$ and denote the rescaled coordinate again with t.[12] By looking at the formula for q, it is indeed easy to see that for $a \to 0$ $q(a) \sim a$ justifying the assumption. The residues of this new differential will be unchanged, but the locations of the residues at t_+ and t_- are changed to q and 1 respectively. Furthermore, we see that the true mass parameters of the hypermultiplets, the residues of the Seiberg-Witten differential, are not precisely equal to the bare mass parameters but receive a finite renormalization through an

[12] An explicit expression for the differential is given in Appendix D.

expansion in q. Again, this is a manifestation of the now familiar ambiguity in UV parameters and the true low energy physical parameters.

S-duality acts on the massive curve in the same way as in the massless case ((4.3.12)). In the above formulation, it is also directly clear how S-duality acts on the mass parameters. It permutes the poles, which effectively permutes the residues as can be clearly seen from (4.3.20). We conclude this particular realization of S-duality automatically acts as triality on the Cartan weights of $SO(8)$.

Low energy information like the gauge coupling and Coulomb branch parameters is encoded in the auxiliary curve:

$$\begin{aligned} \tilde{y}^2 &= -a(\tilde{t}-t_+)(\tilde{t}-t_-)U_2(\tilde{t}) + \frac{1}{4}M_2^2(\tilde{t}) \\ \text{or } \; y^2 &= P_4(t) \end{aligned} \tag{4.3.22}$$

The appearance of the auxiliary curve is more natural now: it simply represents the numerator of the full Seiberg-Witten curve, which indeed is solely responsible for the low energy gauge coupling since its zeroes determine the branch cuts. Furthermore, we see that in the massless $N_f = 4$ theory, the coincidence of the punctures on the Riemann surface with the branch points of the auxiliary curve is, in fact, a coincidence: in general the zeroes of the numerator are not equal to the zeroes of denominator. Again,

$$\partial_u \lambda \sim \frac{\mathrm{d}t}{\sqrt{-a(\tilde{t}-t_+)(\tilde{t}-t_-)U_2(\tilde{t}) + \frac{1}{4}M_2^2(t)}} \tag{4.3.23}$$

corresponds to the, in this case only, holomorphic differential on the Seiberg-Witten curve and is manifestly holomorphic with respect to the auxiliary curve.

We conclude our analysis with a canonical form of the Seiberg-Witten curve for the massive theory:

$$\lambda^2 = \frac{P_4^{(0)}(t)}{t^2(t-1)^2(t-q)^2}\mathrm{d}t^2 + \frac{c(a)u}{t(t-1)(t-q)}\mathrm{d}t^2 \tag{4.3.24}$$

where $P_4^{(0)}(t)$ contains only mass parameters and the cross ratio while the simple pole part depends only on the Coulomb branch parameter. Through this particular parametrization, the massless is curve is manifestly regained upon putting all mass parameters to zero. The constant is given by: $c(a) = \frac{1}{a\, t_-}$. For the explicit expression of $P_4^{(0)}(t)$, see Appendix D.

Using the explicit expression, one can take the decoupling limit $a \to 0$ which shows the differential correctly degenerates into a differential of the form:

$$\lambda^2 = \frac{P_2(t)}{t^2(t-1)^2}\mathrm{d}t^2 \tag{4.3.25}$$

with residues:

$$\mathrm{Res}_{t=0}\, \lambda_\pm = \pm\sqrt{u} \tag{4.3.26}$$

$$\mathrm{Res}_{t=1}\, \lambda_\pm = \pm\frac{m_1+m_2}{2} \tag{4.3.27}$$

$$\mathrm{Res}_{t=\infty}\, \lambda_\pm = \pm\frac{m_1-m_2}{2} \tag{4.3.28}$$

Hence, we see that the Coulomb branch parameter of the decoupled gauge group plays, as desired, the role of the mass parameter of the new flavour symmetry. Note that the surviving differential, describing a thrice punctured sphere, is the 'left' side of the quiver. We could have also rescaled $t \to qt$, such that in the $q \to 0$ limit, the residues would have been the Coulomb branch parameter and the $m_{3,4}$ part.

4.3.3 Massless $T_{m,0}[A_1]$

From this point on, we can relatively easily generalize to arbitrary (linear) quiver theories. A massless linear quiver describing a gauge theory with n separate $SU(2)$ gauge groups has Seiberg-Witten curve:

$$v^2 t^{n+1} + c_1(v^2 - u_1)t^n + \ldots + c_n(v^2 - u_n)t + c_{n+1}v^2 = 0 \tag{4.3.29}$$

with $(v,t) \in \mathbb{C} \times C_{n,0}$. In analogy with the $N_f = 4$ case, we collect powers of v:

$$a_{n+1} \prod_{a=0}^{n} (\tilde{t} - t_a) v^2 = U_{n-1}(\tilde{t})\tilde{t} \tag{4.3.30}$$

with $U_{n-1}(t) = u_n + a_2 c_{n-1} u_{n-1} \tilde{t} + \ldots + a_n c_1 u_1 \tilde{t}^{n-1}$. We again rescaled $\tilde{t} = \frac{c_n}{c_{n+1}} t$ and have defined $a_i = \frac{c_{n+1}^{i-1}}{c_n^i}$. The t_a represent the asymptotic ($v \to \infty$) positions of the NS5 branes. They also determine the UV gauge coupling as:

$$\tau_a = \frac{1}{i\pi} \log\left(\frac{t_{a+1}}{t_a}\right) \tag{4.3.31}$$

We can put $t_0 = 1$ upon a further rescaling of $\tilde{t}$. The sewing parameters are related to the t_a as:

$$t_a = 1 \cdot q_1 \cdots q_a \tag{4.3.32}$$

Indeed, this corresponds (in the standard dual frame) to the identification: $q_a = e^{i\pi\tau_a}$. We adjust our coordinates similarly as in the $N_f = 4$ case and perform a last $SL(2,\mathbb{C})$ transformation to obtain:

$$\lambda^2 = \frac{U_{n-1}(t)}{\Delta_{n+3}(t)} \mathrm{d}t^2 \tag{4.3.33}$$

$$y^2 = U_{n-1}(t)\Delta_{n+3}(t) \tag{4.3.34}$$

The n cross ratios q_i of the $n+3$ roots of $\Delta_{n+3}(t)$ parametrize the n UV gauge couplings as above. The remaining constant a_{n+1} is absorbed in Δ. The moduli space is equal to the moduli space of an n punctured sphere $\mathcal{M}_{n,0}$ with S-duality group $\mathrm{MCG}(C_{n,0})$. Again, the cross ratios $q_i = e^{2\pi i \tau_{i\,UV}}$ are related to the UV couplings and generically do not coincide with the Seiberg-Witten gauge couplings.

4.3.4 Massive $T_{m,0}[A_1]$

Let us add mass parameters in a similar manner as in the example of $N_f = 4$:

$$a_{n+1} \prod_{a=0}^{n} (\tilde{t} - t_a) v^2 = M_{n+1}(\tilde{t}) v + U_{n+1}(\tilde{t}) \tag{4.3.35}$$

The explicit expressions for the polynomials are:

$$M_{n+1}(\tilde{t}) = a_{n+1}(m_1 + m_2)\tilde{t}^{n+1} + a_{n-1} c_2 m_3 \tilde{t}^{n-1} + \ldots + m_{n+1}\tilde{t} + m_{n+2} + m_{n+3} \tag{4.3.36}$$

$$U_{n+1}(t) = -a_{n+1} m_1 m_2 \tilde{t}^{n+1} + a_n c_1 u_1 \tilde{t}^n + \ldots + a_2 c_{n-1} u_{n-1} \tilde{t}^2 + u_n \tilde{t} - m_{n+2} m_{n+3} \tag{4.3.37}$$

Note that, although the $n+2$ coefficients of M_{n+1} and the highest and lowest power of U_{n+1} in principle are all mass parameters, we used the freedom to shift $v \to v + v_0$ to reduce to the physical $n+3$ mass parameters, corresponding to the number of external legs for an n gauge group linear quiver.

Let us briefly justify the remark made in the massive $N_f = 4$ case that the Seiberg-Witten differential corresponding to the unshifted curve does not exhibit its full flavour symmetry. The Seiberg-Witten differential is easily calculated as:

$$\lambda_\pm = \frac{M_{n+1} \pm \sqrt{M_{n+1}^2 + 4a_{n+1}\prod_{a=0}^n(\tilde{t}-t_a)U_{n+1}}}{2a_{n+1}\prod_{a=0}^n(\tilde{t}-t_a)}\frac{\mathrm{d}\tilde{t}}{\tilde{t}} \tag{4.3.38}$$

For $\tilde{t} = t_1, \ldots, t_{n-1}$ only λ_+ has a non-vanishing residue, obscuring the Cartan $U(1) \subset SU(2)$ flavour symmetry subgroup of the bifundamental. The residues at $t = 0, \infty, t_0, t_n$, corresponding to the masses of the fundamental hypermultiplets, obscure the real $SO(4)$ flavour symmetry in a similar manner.

To cure this, we perform a similar shift as in the $N_f = 4$ case to rid ourselves of the linear term in v. Sending $v \to v + \frac{M_{n+1}}{2a_{n+1}\prod_{a=0}^n(\tilde{t}-t_a)}$ the quadratic differential reads:

$$\lambda^2 = \frac{a_{n+1}\prod_{a=0}^n(\tilde{t}-t_a)U_{n+1} + \frac{1}{4}M_{n+1}^2}{\left(a_{n+1}\prod_{a=0}^n(\tilde{t}-t_a)\right)^2}\frac{\mathrm{d}\tilde{t}^2}{\tilde{t}^2} = \frac{P_{2n+2}(t)}{(t-1)^2(t-t_1)^2\cdots(t-t_n)^2}\frac{\mathrm{d}t^2}{t^2} \tag{4.3.39}$$

$$\to \frac{P_{2n+2}(t)}{\Delta_{n+3}^2(t)}\mathrm{d}t^2 \tag{4.3.40}$$

where we redefined the flavour charges to get the differential in its canonical form. Furthermore, in the second equality we absorbed all constants into P_{2n+2}, rescaled $\tilde{t} \to t$ such that the $t_a = 1 \cdot q_1 \cdots q_a$. Notice that P_{2n+2} depends on $2n+3$ parameters corresponding to $n+3$ mass parameters and n Coulomb branch parameters. Let us recover the massless curve by defining: $P_{2n+2}(z) = P_{2n+2}^{(0)}(z) + \Delta_{n+3}^2(z)U_{n-1}(z)$ where for all mass parameters vanishing, $P_{2n+2}^{(0)}(z) = 0$.

The auxiliary curve is given by:

$$y^2 = a_{n+1}\prod_{a=0}^n(\tilde{t}-t_a)U_{n+1} + \frac{1}{4}M_{n+1}^2 \to P_{2n+2}(t) \tag{4.3.41}$$

which defines a double cover of an $n+3$ punctured sphere of genus $g = n$, corresponding to the number of gauge groups. A basis of holomorphic differentials on this curve is given by:

$$\partial_{u_i}\lambda_\pm \sim \pm\frac{\tilde{t}^{n-i}\mathrm{d}t}{\sqrt{a_{n+1}\prod_{a=0}^n(\tilde{t}-t_a)U_{n+1}(t) + \frac{1}{4}M_{n+1}^2(t)}} = \pm\frac{\tilde{t}^{n-i}\mathrm{d}t}{y} \tag{4.3.42}$$

This basis is n dimensional.[13]

We have already explicitly seen the Seiberg-Witten curve for $N_f = 4$ theory degenerates properly into the quadratic differentials associated to thrice punctured spheres and have all desired properties. Here, we want to extend this analysis to a general linear quiver. To this end, let us look at a general degeneration of the quiver, say at the i^{th} node of a linear quiver with n gauge groups. First of all, for n gauge groups, we have $n+3$ punctures. Let us assume we are in a dual frame in which the punctures $t_0, \ldots, t_{i-1}$ are colliding with $t = 0$. Indeed, this dual frame is obtained by a standard sewing process which has the consequence all coordinates of punctures for $a \leq i-1$ are dependent on q_i.[14] We can choose our gauge couplings associated to the punctures such that:

$$t_a < |q_i| \qquad \text{for } a \leq i-1 \tag{4.3.43}$$

$$t_a > 1 \qquad \text{for } a > i-1 \tag{4.3.44}$$

[13]The holomorphicity on the zeroes of y follows from similar reasons as provided in Section 3.5. The fact the differential is holomorphic also at $t = \infty$ follows from the minimal value of the difference in degree of numerator and denominator: $n+1-(n-i) \geq 2$ for all $i \in \{1, \ldots, n\}$.

[14]Here we break the convention that $t_0 = 1$.

Taking $|q_i| \to 0$ implies the following behaviour of Δ:

$$\Delta^2_{n+3} \to t^{2(i+1)} \Delta^2_{n-i+2} \tag{4.3.45}$$

Now, during the degeneration process, at all times we want the residues of the colliding punctures to remain finite since they correspond to physical masses. Let us look a bit more closely at an explicit form of Δ, from which we can clearly see the behaviour of the residues. For definiteness, we look near some point t_a while taking the limit that all the punctures t_j, $j \leq i-1$ are colliding:

$$\lambda^2 \sim \frac{P_{2n+2}(t_a)}{t_a^{2i}(t-t_a)^2(t_a-t_i)^2 \cdots (t_a-t_n)^2} \mathrm{d}t^2 \tag{4.3.46}$$

Since the limit implies $t_a \to 0$, we can infer from the physical consideration of finite masses how the numerator should behave:

$$P_{2n+2}(t) \to t^{2i} P_{2n-2i+2}(t) \tag{4.3.47}$$

Although this behaviour is qualitatively inferred, we have shown earlier the explicit expression for $N_f = 4$ indeed exhibits this behaviour. Although explicit expressions for these curves tend to get rather complicated very quickly, in principle one can perform checks as we did in the previous with suitable software. One might also ponder about an inductive proof by reducing every Seiberg-Witten curve to the $N_f = 4$ curve. Looking at the explicit expression for the curve associated to an $SU(2) \times SU(2)$ theory and following its degeneration towards $N_f = 4$ should provide some insight. As the deadlines are closing in, the author himself cannot bring up the time to look at it before handing in the thesis.

The behaviour of P_{2n+2} shows the auxiliary curve pinches, since its discriminant vanishes. In particular, the i^{th} b cycle diverges, corresponding to the divergence of BPS masses of solitons charged with respect to the decoupling gauge group. This is just the phenomenon we have already seen in the pure theory, namely that at weak coupling the soliton masses grow proportional to $m \sim -\log|q_i| \sim \tau_i$ when $|q_i| \to 0$ (cf. Section 4.2). The Seiberg-Witten curve has become:

$$\lambda^2 = \frac{P_{2n-2i+2}(t)}{\Delta^2_{n-i+3}} \mathrm{d}t^2 \tag{4.3.48}$$

This indeed corresponds to a quiver with $n-i$ gauge groups, as expected.

Sending all gauge couplings to very weak coupling, we obtain a collection of $n+1$ thrice punctured spheres described by:

$$x^2 = \frac{P_2(z)}{\Delta^2_3(z)} \tag{4.3.49}$$

This curve depends on mass parameters for the half hypers only, which depending on its origin in the original quiver might correspond to Coulomb branch parameters of decoupled gauge groups or mass parameters associated to flavour symmetry groups.

4.3.5 $T_{m,1}[A_1]$

The extension to quiver theories of $g = 1$ is implemented using a construction of Witten given in Section 4 of [72]. We will not describe the construction in detail but rather state the result. At the end, we will also briefly comment on the extension to arbitrary genus quivers.

The Seiberg-Witten curve is now given in coordinates $(v, s) \in \mathbb{C} \times C_{n,1}$. Previously, we used t to parametrize the punctured sphere, since it was single-valued in the M-theory x_{10} coordinate. Since now both x_6 and x_{10} are periodic, we will have to consider doubly periodic functions of s to construct Seiberg-Witten curves:

$$v^2 = f_2(s) \tag{4.3.50}$$

with Seiberg-Witten differential vds. The function $f_2(s)$ has n simple poles on the torus and is constructed from Weierstrass elliptic functions $\wp(s,\tau)$.[15] For instance, one can construct doubly periodic functions with simple poles by taking $f_2 \sim \frac{\wp'}{\wp}$.

As the dimension of Teichmüller space for $C_{n,1}$ has dimension n, one might be worried about the number of free parameters in $f_2(s)$. Since the sum of residues of a meromorphic function on a compact Riemann surface should vanish, it seems that $f_2(s)$ only depends on $n-1$ independent parameters.[16] This problem is solved by understanding the Seiberg-Witten curve as a quadratic differential, as we are used to. The residues of a quadratic differential at its simple poles are not constrained and therefore $f_2(s)$ truly has n independent parameters. For an example of a $g=1$ quiver, see Figure 4.3.2.

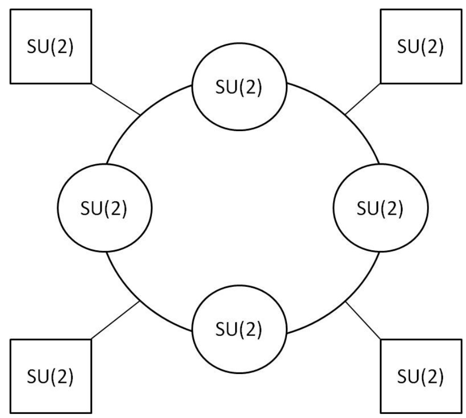

Figure 4.3.2: Quiver diagram for $SU(2)^4$ theory. The UV gauge couplings are given as $\tau_a \sim s_{a+1} - s_a$. Note that the sum corresponds to the modular parameter of the torus: $|\tau_1| + |\tau_2| + |\tau_3| + |\tau_4| = |\tau|$.

Including mass parameters involves, as usual, a linear term in v:

$$v^2 = f_1(s)v + f_2(s) \tag{4.3.51}$$

with $f_1(s)$ again doubly periodic function with simple poles at the n points, whose residues correspond to hypermultiplet mass parameters.[17] We make the usual shift to rid ourselves of the linear term in v to obtain again:

$$v^2 = f_2(s) - \frac{f_1(s)^2}{4} \tag{4.3.52}$$

$$x^2 = \phi_2(z) \tag{4.3.53}$$

where we have written the second equality to demonstrate the reappearance of the canonical form of the Seiberg-Witten curve, even for $g=1$ quivers. Thus, $\phi_2(z)$ is a quadratic differential on the

[15] The function field of doubly periodic functions has been proved to be $\mathbb{C}(\wp,\wp')$. See Section 3.5 of [18].

[16] For some analytical properties of elliptic functions, see Chapter 20 of [70].

[17] Again, one might worry about the fact that the residues of $f_1(s)$ sum to zero, leaving $n-1$ physical mass parameters instead of the desired n independent parameters. This worry is in fact justified for higher rank theories. In that case, the problem is cured by taking $(v,s) \in E$ where E represents a non-trivial $\mathbb{C}$ (the v plane) bundle over $C_{n,1}$ instead of the above mentioned spacetime. This more general construction allows the possibility the masses to sum to any constant, leaving n independent mass parameters. For the precise construction of the particular bundle, see [72]. In the case at hand, we have no need for this construction since we can shift the linear term which makes the mass parameters into coefficients of double poles. The coefficients of double poles on a compact Riemann surface are of course unconstrained.

torus with n double poles on the punctures with coefficients the mass parameters. Again (x, z) naturally live in $T^*C_{n,1}$. Choosing an origin, $\phi_2(z)$ can be split into a part with double poles, depending on mass parameters only, and a part with simple poles, corresponding to the Coulomb branch parameters. Let us explicitly write this canonical form of a $g = 1$ quiver:

$$x^2 = \sum_{i=1}^{n} \left\{ m_i^2 \wp(z - z_i, \tau) + u_i \frac{\wp'(z - z_i, \tau)}{\wp(z - z_i, \tau) + \frac{1}{3}\pi^2} \right\} \tag{4.3.54}$$

with τ the modular parameter of $C_{n,1}$. The addition of the factor $\frac{1}{3}\pi^2$ in the denominator is needed for a proper degeneration of the curve to a linear quiver, as we will see below. Note that it does not affect the locations of the original poles, but does introduce new poles[**?**]. We are somewhat breaking conventions here by taking a sum of double poles instead of a product as we did with the linear quivers. Of course, these conventions yield equivalent differentials. However, one should recognize that the u_i now actually parametrize linear combinations of the Coulomb branch parameters instead and hence the label i should not be mistaken as labelling a gauge group.

The degenerations of an n punctured torus are twofold. It can pinch of its handle and degenerate into an $n+2$ punctured sphere. Alternatively, punctures may collide, leaving a punctured torus and a punctured sphere. When the handle of the torus pinches, one effectively loses the periodicity in the x_6 direction. Hence, we expect the Weierstrass elliptic function to degenerate into a singly periodic function of x_{10}. Furthermore, the Seiberg-Witten differential is expected to have attained two new poles corresponding to the new flavour symmetry. Let us sketch this degeneration by considering an explicit formula for the Weierstrass elliptic function:

$$\wp(z, \tau) = \frac{1}{z^2} + \sum_{n^2+m^2 \neq 0} \left\{ \frac{1}{(z + n + m\tau)^2} - \frac{1}{(n + m\tau)^2} \right\} \tag{4.3.55}$$

For $\tau \to \infty$, it may be seen that the function reduces as:[18]

$$m_i^2 \wp(z - z_i, \tau) \mathrm{d}z^2 \to m_i^2 \left[\frac{\pi^2}{\sin^2(\pi(z - z_i))} - \frac{\pi^2}{3} \right] \mathrm{d}z^2 \tag{4.3.56}$$

$$= -\frac{m_i^2}{4} \left[\frac{\pi^2}{\left(e^{\frac{\pi(s-s_i)}{2}} - e^{-\frac{\pi(s-s_i)}{2}}\right)^2} - \frac{\pi^2}{3} \right] \mathrm{d}s^2 \tag{4.3.57}$$

$$= -\frac{m_i^2}{4} \left[\frac{\pi^2 e^{\pi s_i} t}{(t - e^{\pi s_i})^2} - \frac{\pi^2}{3} \right] \frac{\mathrm{d}t^2}{t^2} \tag{4.3.58}$$

$$\sim m_i^2 \left[\frac{\pi^2 t}{(t - t_i)^2} - \frac{\pi^2}{3} \right] \frac{\mathrm{d}t^2}{t^2} \tag{4.3.59}$$

Notice we have identified $z = -\frac{i}{2}s$, where we used the factor of 2 to make sure the differential only has one pole as $x_{10} \to x_{10} + 2\pi$. Hence, we see that the $\wp$ functions degenerate appropriately to single-valued functions on the punctured sphere with double poles at t_i and $t = 0, \infty$. Note that the residues for $t = 0$ and $t = \infty$ are equal, as they should be since they correspond to the Coulomb branch parameter of the decoupling gauge group.

[18] For instance, see the comments at the end of Chapter 3 of [18].

A similar analysis holds for the $\wp'/\wp$, where we find:

$$\frac{\wp'(z-z_i,\tau)}{\wp(z-z_i,\tau)+\frac{1}{3}\pi^2}\mathrm{d}z^2 \to \frac{-2\pi\cos(\pi(z-z_i))}{\sin(\pi(z-z_i))}\mathrm{d}z^2 \tag{4.3.60}$$

$$\sim \frac{e^{2\pi iz}+e^{2\pi iz_i}}{e^{2\pi iz}-e^{2\pi iz_i}}\mathrm{d}z^2 \tag{4.3.61}$$

$$\sim \frac{t+e^{\pi s_i}}{t-e^{\pi s_i}}\frac{\mathrm{d}t^2}{t^2} \tag{4.3.62}$$

where we forget about some constants multiplying the entire differential. The simple poles at $z = z_i$ carry over in a similar manner to simple poles at $t = t_i$. The double poles at $t = 0, \infty$ make the Coulomb branch parameter associated to the decoupling gauge group into a real mas parameter, as desired. We therefore see that the curve for $(n, 1)$ properly degenerates to a curve for $(n + 2, 0)$.

For the other degeneration, the collision of m punctures, it is now easy to see that the Seiberg-Witten curve degenerates appropriately. Looking at (4.3.54), we see that for m punctures colliding, m of the Weierstrass elliptic functions coincide. Hence, the Seiberg-Witten curve describes a $C_{n-m+1,1}$ theory where the new mass parameter is a sum of the old mass parameters $m = m_1^2 + \ldots + m_m^2$ and new Coulomb branch parameter $u = u_1 + \ldots + u_m$. Notice that such a limit of colliding punctures represents a strong coupling limit of the original quiver, i.e. $\tau_1, \ldots, \tau_{m-1} \to 0$ and corresponds to a weakly coupled quiver in an appropriate dual frame. See Figure 4.3.3 for an example, depicted in the form of trivalent graphs.

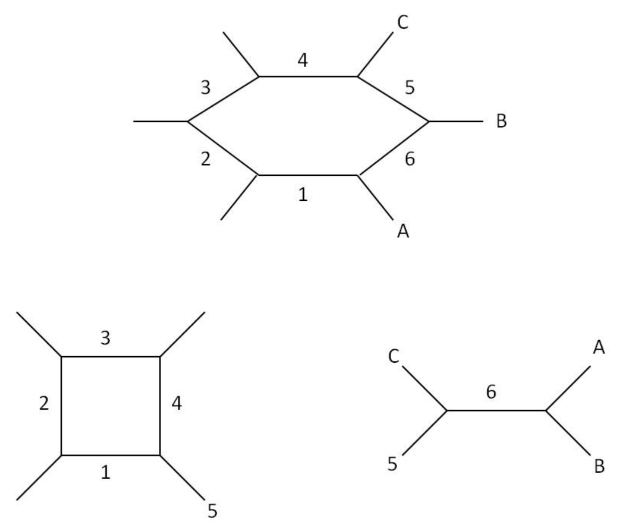

Figure 4.3.3: Two perspectives on the degeneration of a $C_{6,1} \to C_{4,1} \times C_{4,0}$. Gauge groups are indicated by numbers and the (relevant) flavour symmetries by letters. This degeneration scenario can be reached by elementary S-duality moves: first, we very weakly couple all gauge groups but 6 such that we can perform an S-duality around this node exchanging the legs 5 and A. Then we perform a similar move at node 5, now exchanging 1 and C. Reinstating all gauge groups, one can check that a trifundamental has appeared in $2_1 \otimes 2_4 \otimes 2_5$, where gauge group 5 connects the $C_{4,1}$ and $C_{4,0}$ theory. Decoupling 5 leaves us with the expected degeneration.

The other perspective on the degeneration of the $C_{n,1}$ is that one obtains a $C_{m+1,0}$ theory. To see the Seiberg-Witten curve adjusts appropriately to a curve for this theory as well, we note that from this point of view the $n-(m+1)$ punctures combine with the punctures of a pinching torus. The curve degenerates properly as can be seen from a combination of the degeneration of the Weierstrass elliptic functions as above and the further degeneration of a linear quiver as described in the previous section.

To conclude this section, we note that the extension to arbitrary genus theories requires meromorphic functions on general (compact) Riemann surfaces. These are provided by Riemann theta functions, which are direct generalizations of the Weierstrass elliptic function to Riemann surfaces of arbitrary genus.

4.4 Summary: A Six Dimensional Construction

We have seen that the $\mathrm{T_{n,g}[A_1]}$ theories described in the previous, and higher rank theories as well, can be constructed in M-theory. In Section 3.9 M5 branes were used to construct the Seiberg-Witten curves, while extending in $3+1$ dimensions. Minimal area M2 branes whose boundaries end on non-trivial cycles of the Seiberg-Witten curve were identified with BPS excitations of the four dimensional theory. In this section, we take a closer look at this M-theory construction inspired by the results of Gaiotto. We will highlight a few essential aspects of the higher dimensional construction. A much more complete review is presented in [20]. A detailed analysis of the 6d $\mathcal{N}=(2,0)$ theory is given in Sections 2 and 3 of [13]. An insightful and unassuming discussion of the construction is given in Section 6 of [65]. Further references to the vast literature on the present subject are given in the three references above.

The 6d $\mathcal{N}=(2,0)$ A_1 arises from the low energy dynamics of two M5 branes extending in $\mathbb{R}^{5,1}$ while being parallel in the transverse $\mathbb{R}^5$.[19] The notation for the number of supersymmetries reflects the number of chiral and antichiral supercharges respectively. A theory whose supersymmetry transformations consist only of chiral supercharges is possible in six dimensions, as opposed to four dimensions, since spinors and their charge conjugates are of the same chirality.[20]

The presence of the M5 branes breaks the global symmetry to $SO(10,1)\to SO(5,1)\times SO(5)_R$, where our notation already suggests the transverse $SO(5)$ symmetry is interpreted as the R-symmetry of the theory. The supercharges transform in spinor representations of both symmetry groups, which are the fundamental representations of the groups $SU^*(4)\cong SO(5,1)$ and $USp(4)\cong SO(5)$. We denote the supercharges accordingly as $Q_{\alpha a}$ with $\alpha=1,\dots,4$ a fundamental $SU^*(4)$ index and $a=1,\dots,4$ a vector $USp(4)$ index. Therefore, $Q_{\alpha a}$ has sixteen complex components. Although a normal Majorana condition is not possible in six dimensions because the charge conjugation matrix satisfies $(C_\alpha^{\ \beta})^* C_\beta^{\ \gamma}=-\delta_\alpha^\gamma$, due to the pseudoreal nature of the R-symmetry representation it is possible to impose a symplectic Majorana condition. The symplectic Majorana constraint reads:

$$(Q_{\alpha a})^*=\Omega^{ab}C_\alpha^{\ \beta}Q_{\beta b} \tag{4.4.1}$$

with Ω^{ab} the invariant antisymmetric tensor of $USp(4)$, which satisfies $\Omega^{ab}\Omega_{bc}=\delta_c^a$. Consistency of the condition requires an additional constraint on Ω:

$$(\Omega^{ab})^*=-\Omega_{ab} \tag{4.4.2}$$

The symplectic Majorana condition halves the number of supercharges such that we now have sixteen real supercharges, the maximum number of supersymmetries in a *gauge* theory. This corresponds to a six dimensional $\mathcal{N}=(2,0)$ with (symplectic) Majorana-Weyl supercharges.[21]

[19] An A_n theory is realized by taking $n+1$ parallel M5 branes.

[20] See Appendix B of [52] or the first section of the chapter "Supergravity, Brane Dynamics and String Duality" in [46] for detailed analyses.

[21] We used a top-down approach in deriving the supercharges of the 6d theory. We could have also used a bottom-up approach, starting with two six-dimensional Weyl supercharges[**?**]. The R-symmetry group of the $(\in,\prime)$ theory, $U(2)$, would then have been enhanced to $USp(4)$ by the symplectic Majorana constraint one can impose on charge conjugates. Note that this enhancement *is* possible because the supercharges and their charge conjugates are of the same chirality.

The sixteen supercharges, acting on the vacuum, generate sixteen states comprising the so-called tensor multiplet. It can be seen to consist of five scalars, four Weyl fermions and a spin 1 particle corresponding to an antisymmetric tensor boson $B_{\mu\nu}$ whose field strength is a self-dual three form. See Table 4.1 for the representations they form under the helicity group $SO(4) \cong SU(2) \times SU(2)$ and the R-symmetry $SO(5) \cong USp(4)$. Note that although the field strength corresponding to the

tensor boson	$(3,1;1)$
Weyl fermion	$(2,1;4)$
scalars	$(1,1;5)$

Table 4.1: Representations of the field content of the 6d $(2,0)$ theory under the helicity group $SO(4) \subset SO(5,1)$ and the R-symmetry group $SO(5) \cong USp(4)$. Note that the number of fermionic and bosonic degrees of freedom matches.

tensor boson has ten independent components, the boson transforms only under an $SU(2) \subset SO(4)$ helicity subgroup corresponding to three physical polarizations. The five scalars are understood as the degrees of freedom associated to the position of the M5 brane in the transverse $\mathbb{R}^5$, while the fermions arise as Goldstinos associated to the broken supercharges of 11-dimensional M-theory[**?**]. Upon dimensional reduction from six to four dimensions, the six dimensional self-dual field strength will be related to the four dimensional field strength as:

$$H_{5\mu\nu} = F_{\mu\nu},\ H_{6\mu\nu} = \tilde{F}_{\mu\nu} \tag{4.4.3}$$

Due to self-duality of H only one $U(1)$ field strength arises in the dimensional reduction.

To arrive at the four dimensional theory characterized by the Gaiotto and Seiberg-Witten curve, the following dimensional reduction was proposed in [29]. We compactify the theory on $C_{n,g}$, a Riemann surface describing the particular quiver theory we are interested in. Of the global symmetries, only a subgroup remains:

$$\begin{aligned} & SO(3,1) \times SO(2)_C \times SO(3)_R \times SO(2)_R \\ \cong\ & SU(2) \times SU(2) \times SO(2)_C \times SU(2)_R \times SO(2)_R \end{aligned}$$

Note that the splitting of the R-symmetry is by choice for reasons which will become clear below. Under this decomposition the supercharges transform as:

$$\left((2,1)_{\frac{1}{2}} \oplus (1,2)_{-\frac{1}{2}}\right) \otimes \left(2_{\frac{1}{2}} \oplus 2_{-\frac{1}{2}}\right)_+$$

where we have decomposed all groups in $SU(2)$ and $SO(2)$ groups. The subscripts denote the charges of the representations with respect to the $SO(2)$ groups.

We want to twist the theory as to obtain supercharges on four-dimensional spacetime which are invariant under $SO(2)_C$ holonomies of the Riemann surface. This is achieved by twisting the holonomy group as $SO(2)'_C = \Delta\,(SO(2)_C \times SO(2)_R)$, leading to the global symmetry: $SO(3,1) \times SO(3)_R \times SO(2)'_C$. The supercharges now transform as:

$$(2,1,2)_1 \oplus (2,1,2)_0 \oplus (1,2,2)_0 \oplus (1,2,2)_{-1} \tag{4.4.4}$$

The supercharges in the representations with charge 0 with respect to the twisted $SO(2)'_C$ holonomy algebra are labelled by two $SU(2)$ indices: $Q_{\alpha A}, \bar{Q}^{\dot{\alpha}}{}_A$. Here, $\alpha, \dot{\alpha}, A = 1,2$ are the indices corresponding to the chiral and antichiral part of the Poincaré group and the R-symmetry respectively. This indeed provides us with the eight real supercharges of four-dimensional $\mathcal{N} = 2$ supersymmetry.

From the twisted superalgebra, it can be seen that chiral operators in the gauge theory correspond to holomorphic sections on the Riemann surface. This is indeed the relation between the $(2,0)$

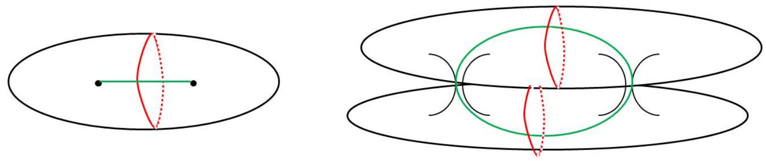

Figure 4.4.1: Lifts of simple arcs or closed curves on the Riemann surface $C_{n,g}$ to the Seiberg-Witten curve. The points at which the arc end are the branch points of Σ, and should not be mistaken for punctures of the Riemann surface. Indeed, although the UV data depends only on the Gaiotto curve, the BPS masses as obtained from the SW curve depend also on the zeroes of $\phi_2(z)$. Degenerations of the SW curve are correlated with the degenerations of the Riemann surface, as explained in Section 4.3.

differentials and the Coulomb branch parameters u_i we described in the previous sections. The $SO(2)_R$ charge of the Coulomb branch parameter becomes the holonomy charge of the quadratic differential after the twist. For more details on this, the reader is referred to [20] and [29].

From the M-theory perspective, the four dimensional A_1 quiver theories are described by two M5 branes extending in $\mathbb{R}^{3,1}$. Furthermore, they wrap a Riemann surface of genus g while being intersected at n points $t = t_a$ by transverse M5 branes, wrapping $\mathbb{C}^*$ and $\mathbb{R}^{3,1}$. This gives the Seiberg-Witten curve in $(x, z) \in T^*C_{n,g}$ as a cover of a punctured Riemann surface $C_{n,g}$.[22] At low energy the transverse M5 branes decouple and are visible only as codimension two defects on the Riemann surface. The curve describing the low energy theory, the positions of the M5 branes, is given by:

$$x^2 = \phi_2(z) \tag{4.4.5}$$

with appropriate poles at the punctures. Furthermore, it is assumed that the four-dimensional limit of the reduction depends only on the complex structure of the Riemann surface. That is, the metric on the surface to which the six dimensional theory *is* sensitive, is of no relevance to the four dimensional dynamics.

M2 branes are suspended between the parallel M5 branes which project onto the Riemann surface as strings. This leaves the possibility for two possible types of strings: arcs between punctures or simple closed curves. On the Seiberg-Witten curve these curves lift to a closed loop or two closed loops respectively. These should represent M2 brane boundaries, such that we find the allowed topologies for the M2 brane is either a disc or a cylinder (annulus). See Figure 4.4.1. The disc represents BPS states from hypermultiplets whereas the cylinder topology naturally describes W-bosons. This was first noted in [40] and [31].

This concludes our discussion of the M-theory in construction in Gaiotto's light. Many more technical details associated to this construction and translations from the 6d to 4d theory can be found in [20].

4.5 Generalizations

So far, we have extensively discussed Gaiotto dualities and their origin in the case of $SU(2)$ gauge groups. In this section we will discuss the generalization to gauge groups $SU(3)$ and $SU(N)$. We will focus on qualitative and quantitative very much alike the first three sections of this chapter.

[22] The assumption here is that the reduction of the M5 branes on the cylinder is equivalent to the reduction on the sphere with two extra punctures. In other words, the punctures at $t = 0, \infty$ on the sphere are equivalent to the punctures created by the transverse M5 branes.

However, the analysis will be considerably reduced to the mere highlighting of structures appearing. For full details, we refer to the original of Gaiotto [29].

4.5.1 $\mathrm{T}_{(\mathrm{f}_1,\mathrm{f}_3),\mathrm{g}}[\mathrm{A}_2]$

The simplest SCFT with a rank 2 gauge group is the theory with $SU(3)$ gauge group coupled to six flavours in the (anti)fundamental representation, i.e. six hypermultiplets each constituting a $3 \oplus \bar{3}$. The theory has a $U(6)$ flavour symmetry which does not enhance because representations of $SU(N)$ for $N \geq 3$ are complex.

This theory does not share the S-duality properties of $SU(2)$ $N_f = 4$. Instead of being fully $SL(2,\mathbb{Z})$ invariant, it is only invariant under a subgroup $\Gamma^0(2) \subset SL(2,\mathbb{Z})$ as found in [48]. This implies the existence of an infinite coupling point in the moduli space, which a priori is not related to some dual description á la $N_f = 4$. However, only recently Argyres and Seiberg found a surprising interpretation of the infinite coupling point in [10]. They showed evidence that at the infinite coupling point, a weakly coupled $SU(2)$ gauge theory emerges. This gauge group is coupled to one fundamental hypermultiplet and an isolated rank 1 SCFT with E_6 flavour symmetry by gauging the $SU(2)$ part of the $SU(2) \times SU(6) \subset E_6$ maximal subgroup.[23] Note that the global symmetries of the theories match: $U(6) \cong SU(6) \times SO(2)$. For more details about the evidence for this duality, we refer to the original article [10]. This new type of duality is called Argyres-Seiberg duality.

Gaiotto's insight was to identify, as in the $SU(2)$ case, building blocks to build arbitrarily complicated A_2 superconformal field theories and relate them again to certain punctured Riemann surfaces. Although we will first use Argyres-Seiberg duality to motivate the construction, the identification of the quiver theories with certain Riemann surfaces will in fact allow us to derive Argyres-Seiberg duality from the more familiar A_1 S-duality.

To identify the building blocks for A_2 theories, we start with $SU(3)^3$ gauge theory with flavour symmetry $U(1)^4 \times SU(3)^2$ as depicted in Figure 4.5.1.[24] Due to the fact bifundamentals live in a $(3,\bar{3})_1 \oplus (\bar{3},3)_{-1}$ representation, they provide three hypermultiplets to either of the gauge groups they are coupled to. Together with the fundamentals then, all gauge couplings are exactly marginal.

Keeping the gauge couplings at the first and last node very weak, we can imagine bringing the middle gauge coupling towards strong coupling. We then perform an Argyres-Seiberg duality and arrive in a new weakly coupled theory. In the dual diagram, we have split the $SU(6)$ maximal flavour symmetry subgroup into a subgroup $SU(3)\times SU(3) \subset SU(6)$ and identify these with the very weakly coupled gauge groups of the original theory. Note that the $U(1)$ subgroups of the bifundamentals in the original quiver now appear as (a combination of) the $SO(2)$ flavour symmetry of the single fundamental and the commutant $U(1)$ of the $SU(3)\times SU(3) \subset SU(6)$. Taking the $SU(2)$ and $SU(3)$ couplings to zero, the single fundamental decouples and we obtain an object with a naive flavour symmetry of $SU(3) \times SU(3) \times SU(2) \times U(1)$. That is, we now have a representation:

$$(3,\bar{3},2)_1 \oplus (\bar{3},3,2)_{-1}$$

We would have expected to obtain the isolated rank 1 theory with E_6 flavour symmetry. The misstep resides in the fact we have only considered a flavour symmetry subgroup, while forgetting about the $2 \otimes 20$ additional generators of E_6.

It turns out that the $SU(2) \times U(1)$ combines with those additional generators, that are singlets under the $SU(3) \times SU(3)$, to form an additional $SU(3)$ symmetry. The group $SU(3)_a \times SU(3)_b \times SU(3)_c$ is indeed contained in E_6, as can most easily seen from the (extended) Dynkin diagram in Figure 4.5.4. The matter representations under this subgroup read:

$$(3_a, 3_b, 3_c) \oplus (\bar{3}_a, \bar{3}_b, \bar{3}_c)$$

[23] The missing generators of E_6 can be represented in a $2 \otimes 20$ representation of $SU(2) \times SU(6)$. Indeed, the number of generators matches: $78 = 3 + 35 + 40$.

[24] N fundamental flavours carry an $SU(N) \times U(1)$ symmetry, while bifundamentals only carry a $U(1)$ flavour symmetry in the case of $SU(N)$, $N \geq 3$ gauge groups.

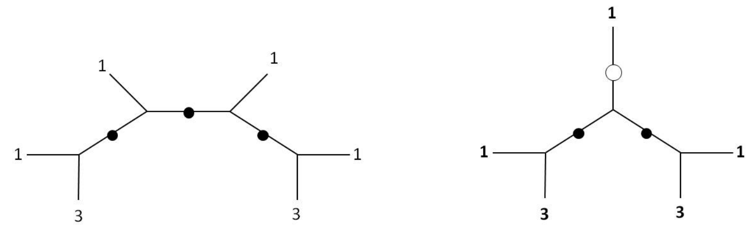

Figure 4.5.1: Generalized quiver diagram of $SU(3)^3$ theory and its Argyres-Seiberg dual theory. The filled dots represent $SU(3)$ gauge groups whereas open dots represent the gauged $SU(2) \subset SU(3)$. Furthermore, we use the number 1 and 3 to represent a $U(1)$ and $SU(3)$ flavour symmetry respectively. One has to keep in mind an extra $U(1)$ flavour symmetry in the gauged $SU(2)$, corresponding to the commutant $U(1)$ of the $SU(3) \times SU(3) \subset SU(6)$.

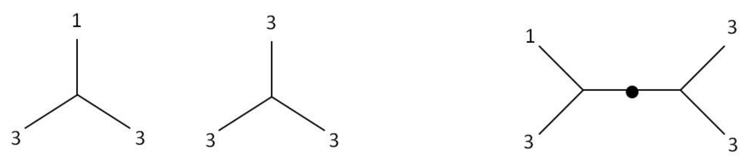

Figure 4.5.2: Building blocks for $\mathrm{T}_{(\mathrm{f}_1,\mathrm{f}_3),\mathrm{g}}[\mathrm{A}_2]$ theories. The 1 and 3 stand for a $U(1)$ and $SU(3)$ flavour symmetry respectively. Next to the gauging of diagonal $SU(3)$ flavour symmetries, one can also gauge any $SU(2) \subset SU(3)$ and couple it to a single fundamental hypermultiplet. On the right, a naively non-conformal theory is shown. It is conformal, since the trivalent vertex on the right is already the isolated SCFT which provides only three flavours to the gauge group it is coupled to.

Weakly gauging two of the $SU(3)$ flavour symmetries of the E_6 theory gives back the right superconformal quiver in Figure 4.5.1. We conclude that gauging an $SU(3)$ flavour symmetry of the E_6 theory provides three fundamental hypermultiplets to the new gauge group, precisely as does the weak gauging of the flavour symmetry of a bifundamental block.[25] Hence, we can use the E_6 theory and the bifundamental block to build superconformal field theories.

The two quivers associated to the bifundamental block and the E_6 theory are shown in Figure 4.5.2. The crucial observation is that all the three $SU(3) \subset E_6$ subgroups are identical, again manifest from the extended Dynkin diagram in Figure 4.5.4. By gauging *any* flavour $SU(3)$ of an E_6 with either another E_6 or a bifundamental block, the resulting theory will be conformal. Gauging *any* $SU(2)$ subgroup of an $SU(3)$ flavour symmetry and coupling it to a single fundamental will keep the quiver conformal as well.

This insight opens up a huge zoo of new superconformal theories, and many duality relations among them. In fact, Gaiotto has conjectured a simple classification to determine all dual theories.[26] One simply counts the number of $U(1)$ flavour symmetries, f_1, and $SU(3)$ flavour symmetries, f_3. As usual, the loops in the quiver are denoted by g. The claim is that all possible topologies of quiver diagrams characterized by these three numbers are weakly coupled descriptions of a single theory

[25]These statements are checked by explicit calculations. See [17] and references therein.

[26]This conjecture is checked in his paper [29] by determining the Seiberg-Witten curve of a particular theory using the M-theory construction, and checking whether the dual theories match all the possible cusps/degenerations of the curve, as performed in the previous section for the A_1 case.

$T_{(f_1,f_3),g}[A_2]$. The space of exactly marginal couplings of the theory coincides with the moduli space of a Riemann surface with *two* types of punctures: $\mathcal{M}_{(f_1,f_3),g}$. S-duality acts on the punctures as a permutation group on equal type punctures.

We call the punctures associated to $U(1)$ flavour symmetries type I punctures while the punctures associated to $SU(3)$ flavour symmetries the type III punctures. Note that the E_6 theory is represented by a thrice type III punctured sphere, while the bifundamental block is represented by a sphere with two type III punctures and one type I puncture. Gauging diagonal subgroups amounts to the cutting and sewing procedure described in Section 4.2 of type III punctures, except when one gauges an $SU(2) \subset SU(3)$ and couples it to a single fundamental. In this case, the gauging amounts to the splitting of an f_3 puncture into two f_1 punctures.

Degenerations are twofold:

1. An $SU(3)$ gauge group becomes very weakly coupled. It is not difficult to see, for instance from Figure 4.5.1, that this implies the collision of a certain number of type I punctures with one Type III puncture. These punctures merge into a single type III puncture associated to the new $SU(3)$ flavour symmetry. Alternatively, from a general quiver one can always gauge an $SU(2) \subset SU(3)$ of all flavour symmetries such that the resulting Riemann surface contains only type I punctures. Indeed, this amounts to the replacement of all type III punctures by two type I punctures, as above. Then, a weakly coupled $SU(3)$ corresponds to the collision of n type I punctures for $n \geq 3$.

2. The special case of the collision of two type I punctures corresponds to a very weakly coupled $SU(2)$ gauge group in the quiver.[27] Indeed, this is just the observation from Section 4.2 that the collision of two type I punctures corresponds to a very weakly coupled $SU(2)$ gauge group. In the decoupling limit of the $SU(2)$ group, an $SU(3)$ flavour symmetry is left, whereas the sphere that splits off has two type I punctures plus an 'irregular'[17] puncture corresponding to the decoupled $SU(2) \subset SU(3)$.

3. At last, there may be collisions between two (or more) type III punctures. In some particular dual frame, this could correspond to an $SU(3)$ gauge group becoming very weakly coupled and an E_6 theory splits off.

A remarkable consequence of the Riemann surface interpretation is that Argyres-Seiberg duality derives from S-duality of $SU(2)$ gauge groups. Namely, consider the $N_f = 6$ theory. It is represented by the Riemann surface in Figure 4.5.3. The collision of a type I with a type III puncture is easily identified with the $\tau \to i\infty$ limit. However, the strong coupling point of the $SU(3)$ is reached by colliding the two type I punctures. As noted above, the collision of two type I punctures implies an $SU(2)$ gauge group is becoming very weakly coupled. But this $SU(2)$ gauge group is not manifest in the original quiver. Instead, a weakly coupled $SU(2)$ gauge group emerges when the original $SU(3)$ is strongly coupled. Moreover, upon decoupling the $SU(2)$ a thrice type III punctured Riemann surface appears.[28] See Figure 4.5.3. The two lower type III punctures represent the $SU(3) \times SU(3) \subset SU(6)$ full $N_f = 6$ flavour symmetry. The third type III puncture is associated to another $SU(3)$ flavour symmetry. However, these three punctures are equivalent. This means that any pair of the remaining $SU(3)$ flavour groups are in fact part of an $SU(6)$ flavour symmetry. The only way for these pairings to be possible is in fact when the full flavour symmetry is E_6. This is most clear when one looks at the (extended) Dynkin diagram. See Figure 4.5.4. We conclude that S-duality of $SU(2)$ and the equivalence of $SU(3)$ punctures directly implies a weakly coupled dual description of the infinite

[27] Type I punctures are indistinguishable from the $SU(2)$ punctures, understood as the fact that the structure of punctures is in direct correspondence with the size of the Cartan subalgebra. We will have more to about this in the next section.

[28] The reason why a new type III puncture arises is explained by a careful analysis of degenerations of quivers with superconformal 'tails'. We will come back to this in the next section. One may also consult sections 12.2 and 12.3 of [65].

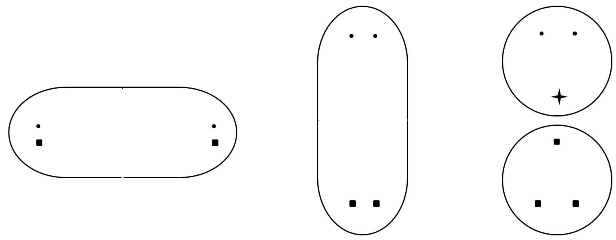

Figure 4.5.3: Two weak coupling limits of the $N_f = 6$ theory interpreted as collisions of punctures on the Riemann surface $C_{(2,2),0}$. The dots represent the $U(1)$ flavour symmetries, whereas the squares the $SU(3)$ flavour symmetries. On the right, the full degeneration of the middle Riemann surface is shown. The star represents the irregular puncture corresponding to a decoupled $SU(2) \subset SU(3)$.

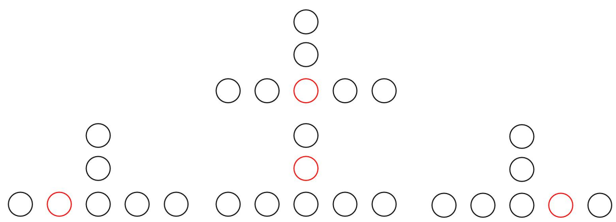

Figure 4.5.4: Extended Dynkin diagrams for E_6. Deleting any cell provides the Dynkin diagram of a maximal subgroup. On the top, the maximal $SU(3)^3$ subgroup is manifest. On the bottom, we see that any pair of $SU(3)$'s can be embedded in an $SU(6)$ while leaving one $SU(2)$.

coupling point of $N_f = 6$ theory which has E_6 flavour symmetry, of which an $SU(2)$ subgroup is gauged. This is in full agreement with the original arguments of Argyres in Seiberg.

Similar games can be played with more general quiver theories to understand the emergence of E_7 and E_8 theories as dual descriptions of simple linear quivers consisting only of A type gauge groups and fundamental hypermultiplets, confirming the conjecture of Seiberg and Argyres in [10]. Gaiotto showed already in his paper [29] how to find the E_7 theory. See also [66] of Tachikawa for a particular clear explanation how the $E_{6,7,8}$ theories appear as weakly coupled descriptions of ordinary linear quivers.

We briefly mention some quantitative aspects of the $\mathrm{T}_{(\mathrm{f}_1,\mathrm{f}_3),0}[\mathrm{A}_2]$ theories. The generic form of the Seiberg-Witten curve is given by:

$$x^3 = \phi_2(t)x + \phi_3(t) \tag{4.5.1}$$

As usual, the Seiberg-Witten differential is $\lambda = x\mathrm{d}t$, $\phi_2(t)\mathrm{d}t^2$ is a quadratic differential and $\phi_3(t)\mathrm{d}t^3$ cubic differential on the Riemann surface $C_{(f_1,f_3),0}$. In this sense, the Seiberg-Witten curve is parametrized as a threefold cover Σ of $T^*C_{(f_1,f_3),0}$. Explicit expressions for these differentials in the massless case are:

$$\phi_2(t) = \frac{U^{(2)}_{f_1+f_3-4}(t)}{\Delta_{f_1}(t)\Delta_{f_3}(t)} \tag{4.5.2}$$

$$\phi_3(t) = \frac{U^{(3)}_{f_1+2f_3-6}(t)}{\Delta_{f_1}(t)\Delta^2_{f_3}(t)} \tag{4.5.3}$$

while in the massive case we in general have:

$$\phi_2(t) = \frac{P^{(2)}_{2f_1+2f_3-4}(t)}{\Delta^2_{f_1}(t)\Delta^2_{f_3}(t)} \quad (4.5.4)$$

$$\phi_3(t) = \frac{P^{(3)}_{3f_1+3f_3-6}(t)}{\Delta^3_{f_1}(t)\Delta^3_{f_3}(t)} \quad (4.5.5)$$

Let us comment briefly on the form. First of all, the $(2,3)$ superscripts on the U polynomials denote the Coulomb branch parameters associated to the rank 2 gauge theory: $u_i^2 = \langle \Phi_i^2 \rangle$ and $u_i^3 = \langle \Phi_i^3 \rangle$.[29] The subscripts of these polynomials simply denote the degree and can be traced back to the number of $SU(2)$ and $SU(3)$ gauge groups and the number of E_6 theories in the quiver.

More interestingly, we see that the quadratic differential does not 'see' the special punctures; it has simple poles at all punctures. The cubic differential does distinguish between the two types, and has double poles at the type III punctures. Note that this does not lead to a non-vanishing residue of the Seiberg-Witten differential. However, it does make sure that when an $SU(3)$ group decouples, the resulting flavour symmetry has two independent mass parameters, as required by the Cartan.

If an $SU(2)$ subgroup of all $SU(3)$ flavour symmetry groups is gauged, we can simply replace all type III punctures by two type I punctures and the cubic *will* have simple poles all the same. Notice then that the gauge coupling moduli space will be the moduli space of a Riemann surface with equal punctures: i.e. our familiar $\mathcal{M}_{f_1+f_3,0}$. From this it is quite clear the UV moduli spaces carry in fact little information about the theory in question, as a complicated theory consisting of isolated E_6 rank 1 theories and $SU(3)$ and $SU(2)$ gauge groups in fact has the exact same moduli space as a 'simple' linear A_1 theory. The additional information about the possible fixtures and cylinders appearing whenever the Riemann surface degenerates is encoded in the Seiberg-Witten curves. This interplay of the UV and IR is beautifully contained in the cotangent bundle of the Riemann surface.

This concludes the essence of the quantitative analysis of the $\mathrm{T}_{(f_1,f_3),g}[A_2]$ theories. There are many subtleties we have overstepped, and we have not considered higher genus quivers. For this, the reader is referred to the original article [29].

4.5.2 $\mathrm{T}_{(\mathrm{f_a}),\mathrm{g}}[\mathrm{A_{N-1}}]$

In this section, we briefly mention some aspects of the analysis of more general quivers with A type gauge groups. Quite surprisingly, Gaiotto has found out much about the structure of the moduli spaces of these theories and the possible dual frames by straightforwardly generalizing his approach to $A_{1,2}$ quivers. The mere fact that this generalization seems to work demonstrates the extreme power of the Gaiotto curves.

Since the Seiberg-Witten curves are already fully available for the case we consider, linear quivers of unitary gauge groups, let us point out what is new in the analysis. Again, new information is contained in the structure of punctures on the Gaiotto curve. Let us first state the curve for generalized $SU(N)$ theories:

$$F(x,z) = x^N - \sum_{i=2}^{N} x^{N-i}\phi_i(z) \quad (4.5.6)$$

This curve is an N fold cover of the Riemann surface, on which the Seiberg-Witten differential $\lambda = x\mathrm{d}z$ is defined. The $\phi_i(z)\mathrm{d}z^i$ are degree i differentials which have poles at the n punctures of the Riemann surface $C_{n,g}$. For a massless curve, these poles are restricted to be of order $p_i < i$. Inserting mass deformations gives the differentials poles of order $p_i = i$. The residue of the Seiberg-Witten differential at a particular puncture will be some linear combination of the mass deformations

[29] It is again a straightforward application of the Riemann-Roch theorem to find that the dimension of the space of quadratic respectively cubic differentials equals the number of dimension 2 and 3 operators parametrizing the Coulomb branch of the gauge theory. For the explicit formula, see [17].

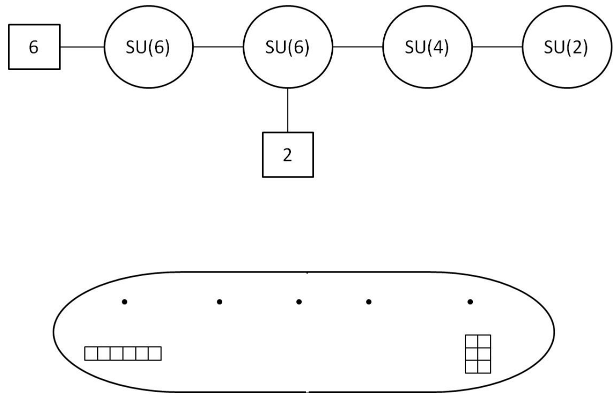

Figure 4.5.5: A superconformal $SU(6)$ quiver with tails. The left tail describes a full puncture with $SU(6)$ flavour symmetry, described by the associated Young tableau. The right tail corresponds to the partition $6 = 2 + 2 + 2$ producing a Young tableau with an $SU(2)$ symmetry, in accordance with the insertion of the extra two fundamental flavours.

at that puncture. We note that for an N-fold cover, λ in general has N residues which have to sum to zero. Indeed, this gives $N - 1$ independent mass parameters in accordance with the rank (Cartan weights) of an $SU(N)$ flavour symmetry. This is also reflected in the absence of an x^{N-1} term in the Seiberg-Witten curve. The explicit form of the ϕ_i is given by appropriate generalizations of (4.5.5).

We have already seen in the A_2 case one has to deal with different types of punctures. In the $SU(N)$ case, we will generalize the idea of a particular puncture on a Riemann surface. Consider a mass deformed curve. Its Seiberg-Witten differential has n simple poles at the punctures. On Σ, at every puncture it has a residue at all N sheets giving a maximum of $N - 1$ independent residues. If this is the case, one has flavour symmetry group $SU(N)$. However, some of these residues may be equal signifying a flavour symmetry with smaller rank.

For the quiver theories we consider, there is a simple way to make manifest the correct flavour symmetry at a particular puncture. A general A type superconformal quiver is a linear quiver with n gauge groups:

$$SU(k_1) \times \cdots \times SU(k_i) \times SU(N)^j \times SU(k_{i+j+1}) \times \cdots \times SU(k_n)$$

For this quiver to be conformal at every node the equation $2k_l = k_{l-1}k_{l+1} + m_l$ should hold, where m_l are possibly additional fundamentals and $k_0 = k_{n+1} = 0$. Furthermore, without loss of generality we take $k_i, k_{i+j+1} < N$. The middle piece is a generic $SU(N)$ quiver, whereas the two ends are the tails. The ranks of the gauge groups in the tails must monotonically decrease towards the ends for the quiver to remain conformal.

Information about the flavour symmetry of this quiver is contained in the differences $s_l = k_l - k_{l-1}$ for $0 \leq l \leq i+1$ and $s'_l = k_l - k_{l+1}$ for $i + j \leq l \leq n + 1$. Indeed, s_l counts how many fundamentals have to be added at each node l for the quiver to be conformal. Restricting the story to one tail, we note that $\sum_{l=0}^{i+1} s_l = N$, hence s_l provides a partition of N. Furthermore, s_l can be seen to be monotonically increasing. We can therefore naturally associate a Young diagram to the partition made by s_l, where s_l determines the width of a row and we build the Young diagram top-down. The symmetry associated to such a Young diagram is $G = S\left(\prod_l U(n_l)\right)$ with n_l the number of columns with equal height. The S removes a diagonal $U(1)$. This construction precisely determines the correct flavour symmetry for the punctures and is illustrated in Figure 4.5.5.

The construction can be applied to any linear quiver of unitary gauge groups. The corresponding

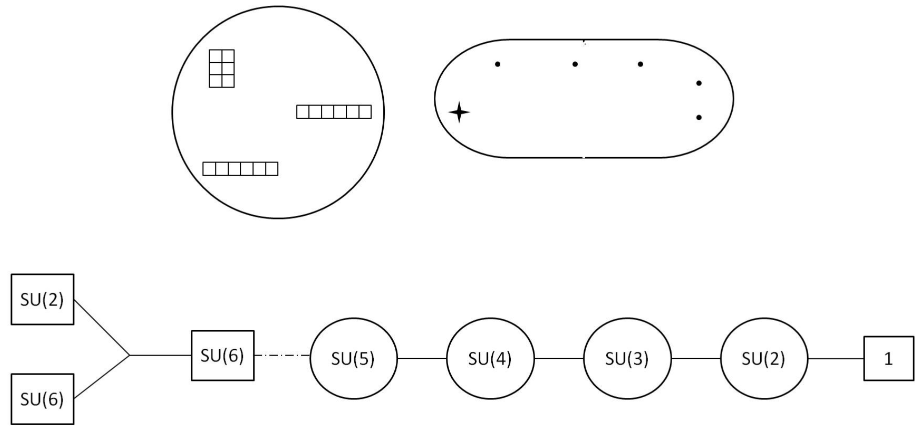

Figure 4.5.6: Dual weakly coupled description of the strong coupling point of Figure 4.5.5. The collision of five basic punctures implies an $SU(5)$ tail is weakly coupled. Decoupling it leaves a puncture of partition $6 = (6-5)+1+\ldots+1$; a full puncture, implying it was weakly gauged in an $SU(6)$ flavour symmetry group. The global symmetries match: $U(6)\times U(2)\times U(1)^3 \cong SU(6)\times SU(2)\times U(1)^4\times U(1)$ where the last $U(1)$ factor stems from the ungauged $U(1)\subset SU(6)$

Riemann surface will have two generic punctures accompanied by two additional simple punctures. Furthermore, the number of bifundamentals determines the number of additional simple punctures.

The moduli spaces of these theories are rich and can be conveniently studied by looking at collisions of various types of punctures on the Riemann surface. For instance in the case of an $SU(N)$ quiver theory with generic tails, the collision of $m < N-1$ punctures corresponds to the weak coupling limit of an $SU(m)$ gauge group coupled to a superconformal tail consisting only at the end of single hypermultiplet in the fundamental of $SU(2)$. This leaves behind a puncture of $SU(m)$ flavour symmetry, corresponding to the partition $N = (N-m)+1+\ldots+1$. This happens in Figure 4.5.5 when two or four simple punctures collide.

More interesting is when $m = N-1$ punctures collide. This corresponds to some strong coupling limit of an $SU(N)$ group. In the Figure 4.5.6, the total number of simple punctures is 5. The collision of all five punctures produces a sphere with two full punctures and the $SU(2)$ Young tableau. However, the collision of five punctures we also know to correspond to a weakly coupled superconformal $SU(5)$ tail. We see a generalization of Argyres-Seiberg duality, where the $SU(5)$ is weakly gauged within the $SU(6)$ flavour symmetry subgroup of some interacting SCFT with full flavour symmetry $SU(6)^2\times SU(2)$. Unlike the original Argyres-Seiberg duality where the $SU(6)\times SU(2)$ subgroup enhances to a full E_6 flavour symmetry, this flavour symmetry is not expected to enhance. The dual quiver is depicted in Figure 4.5.6.

We want to finish by mentioning the already famous T_N theory or fixture. Consider a linear quiver of $N-2$ $SU(N)$ gauge groups, with at both ends N fundamentals. The associated Riemann surface has $N-1$ simple punctures and two full punctures. The collision of $N-1$ punctures is a strong coupling limit of the original quiver, while in a dual description a weakly coupled superconformal $SU(N-1)$ tail appears. As usual, we interpret this as the statement that there was an $SU(N-1)\subset SU(N)$ weakly gauged which splits of as a tail $[N-1]\times SU(N-2)\times\ldots\times SU(2)$ coupled to a single fundamental, while leaving a triangular $SU(N)^3$ theory. The triangular theory is called T_N and represents an interacting superconformal field theory with an flavour symmetry subgroup $SU(N)^3$, which does not enhance generically.

There are simple rules to obtain the correct degrees of the poles of the degree i differentials at a particular puncture, depending on whether one considers the massless or massive Seiberg-Witten curve. For these rules, we refer to [17]. In the same paper, the interested reader may also find an extensive discussion on possible degenerations of the above described quivers, punctures that appear after a full degeneration and the allowed punctures on the set of thrice punctured spheres, i.e. the allowed building blocks for these quiver theories.

Chapter 5

Summary and Outlook

In this thesis the geometrical nature of $\mathcal{N} = 2$ super Yang-Mills theories stood central. In chapter one, the essential ingredients were provided which underlie the celebrated exact solution of the low energy effective pure $SU(2)$ action by Seiberg and Witten. The geometrical nature of this theory emerged from the smoothness of its quantum moduli space, which was shown to be equivalent to the moduli space of a certain elliptic curve. In particular, the monodromies of the gauge coupling and the physical requirement that $\text{Im}\tau(u) > 0$ suggested its identification with the modular parameter of the elliptic curve. This insight turned out to be sufficient to solve for the full low energy theory. A fundamental aspect of the theory is that although the theory is smoothly connected all over the moduli space, appropriate physical descriptions require different local coordinates on the moduli space or in more physical terms: electric-magnetic dual descriptions. Whereas physics described in a Lagrangian is inescapably a local construct, the Seiberg-Witten curves give a first indication of how to think about a global description of gauge (a family of) theories.

The solution begged for the natural question: does this construction generalize to more general $\mathcal{N} = 2$ gauge theories? The answer turned out to be positive. The inclusion of hypermultiplets showed a much richer singularity structure on moduli space leading to unexpected physical conclusions, primarily attributed to an intricate web of walls of marginal stability. Perhaps the most accessible, yet interesting theory is the theory with four flavours of hypermultiplets. Due to its vanishing beta function a more direct relation between the space of abelian IR and non-abelian UV theories could be obtained. In fact, the spectrum of the IR theory shows signs of $SL(2, \mathbb{Z})$ duality as long as flavour symmetry representations are appropriately permuted. The absence of lines of marginal stability guarantee the semiclassical BPS spectrum remains intact as one moves towards strong coupling, in contrast to the theories with less flavours. The duality properties of the IR theory and its spectrum can be regarded as rudimentary evidence of duality in the non-abelian UV theory. More evidence, or eventually proof, would be interesting in the sense it would enable us to better understand the microscopic dynamics of non-abelian theories.

Interestingly, the class of theories described by elliptic curves greatly enhanced when Witten discovered the Seiberg-Witten curves could be realized in extra space dimensions in the context of M-theory. Moreover, M-theory naturally incorporates particular details of the four dimensional theories, which from the four-dimensional perspective appeared somewhat ad hoc.

Gaiotto showed an even more thrilling aspect of the M-theory construction. The Seiberg-Witten curves which describe the physics on the Coulomb branch can be understood as covers of a punctured Riemann surface. There is even more structure to it, as the Seiberg-Witten curves are realized as a collection of k-differentials living in the k^{th} symmetric product of the cotangent bundle of the Riemann surface where $k = 1, \ldots, N$ for a general quiver gauge theory with largest rank group $SU(N)$. An intuitive idea about the boundaries of the moduli space was given by conjecturing the degenerations of the base Riemann surface correspond to all weakly coupled descriptions of the theory. This led to the discovery of unexpected dual descriptions of a single theory, in which for $N > 2$ also non-Lagrangian theories appeared. At the very least, it is remarkable that one starts with a seemingly innocent gauge theory and ends up with completely new SCFTs which would have been hard to guess as a dual description. It would be interesting to better understand these non-Lagrangian theories, as this might teach us about other paradigms than weakly coupled Lagrangian analysis in which to describe physics.

On the one hand, it is satisfying that M-theory knows so much about complicated structures appearing gauge theories and allows relative ease in obtaining results through the brane constructions. On the other hand, of course, extra dimensions still reside in the metaphysical corner of theoretical physics. Especially in the light of emergent geometry and gravity from CFTs in recent years, one might wonder about the implication arrows in the case of the M-theory description and the four dimensional gauge dynamics. In particular, does the geometry emerge from the gauge dynamics or are gauge dynamics fundamentally geometric?

Gaiotto's perspective on $\mathcal{N} = 2$ gauge theories has opened many doors. We try to present a list of references which originate from and have been inspired by [29]. Since this is a very active area of research the list should not be understood as being up-to-date, due to the simple fact the author has not been able during the thesis to reach cutting-edge science. Rather, he presents a list of references, some perhaps a bit dated, which are connected to [29] and deserve further study. In no particular order:

1. One of the first papers following $\mathcal{N} = 2$ dualities was the proposal for a rather interesting correspondence, the already celebrated AGT correspondence [36]. The correspondence states that integrals over partition functions of the 4d A_1 quiver gauge theories obtained by reducing the 6d $\mathcal{N} = (2, 0)$ on a Riemann surface $C_{n,g}$ are equivalent to correlators in 2d Liouville field theories defined on the same Riemann surface. This correspondence has spurred much research into possible generalizations and also explanations of the origin of the correspondence. For instance, a similar relation between the superconformal A_{N-1} theories and Toda conformal field theories was found shortly after the original discovery[75], while a recent claim for a proof of the correspondence can be found in [38].

2. Holographic duals of quiver gauge theories have been described in [30]. It would be interesting to see how non-Lagrangian SCFTs, appearing in the large N-limit, are described in terms of their gravity duals.

3. The phenomenon of wall-crossing in $\mathcal{N} = 2$ gauge theories has been studied in this new language. A series of papers describes the development. For instance, see [20].

4. Other interesting directions include the breaking of supersymmetry to $\mathcal{N} = 1$ by giving a mass to the chiral multiplet, or coupling only a $\mathcal{N} = 1$ vector multiplet to a theory. This has been investigated in [22] in the context of the T_N theory. Understanding the theories with less flavours in the context of the Riemann surfaces was described in [28].

5. Lastly, a recent classification of UV complete $\mathcal{N} = 2$ gauge theories was given in [14]. Also, it is mentioned of each of theories whether they have a Seiberg-Witten solution.

Appendix A

Electric-magnetic Duality

A peculiar symmetry of the source free Maxwell equations is the symmetry in the exchange of electric and magnetic fields. Although the physical relevance of electric-magnetic duality in the Maxwell action is as of yet unclear, the spontaneously broken non-abelian gauge theories studied in this thesis also exhibit a form of electric-magnetic duality on the Coulomb branch. In the gauge theories we study this duality is more natural because magnetic monopoles appear in the spectrum. Since the dualization of the gauge theories is very similar to the ordinary Maxwell duality, we briefly review Maxwell duality in this appendix.

Consider the classical Maxwell action given by:

$$S = \int \mathrm{d}x^4 \frac{1}{g^2} F^{\mu\nu} F_{\mu\nu} + A_\mu J_e^\mu \tag{A.0.1}$$

with $F_{\mu\nu}$ the antisymmetric field strength tensor and where we have added an electric source. Determining the equations of motion we find the inhomogeneous Maxwell equations, while we write down the homogeneous equation as well:

$$\begin{aligned} \partial_\mu F^{\mu\nu} &= J_e^\nu \\ \partial_\mu \tilde{F}^{\mu\nu} &= 0 \end{aligned} \tag{A.0.2}$$

Here, $\tilde{F}_{\mu\nu} = \frac{1}{2}\epsilon_{\mu\nu\rho\sigma} F^{\rho\sigma}$ is the Hodge dual of the field strength. The homogeneous equation is non-dynamical and is automatically satisfied by the definition of the field strength in terms of the vector potential A_μ:

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{A.0.3}$$

We may treat the field strength as unconstrained if we instead impose the Bianchi identity via a Lagrange multiplier A_μ^D in the action:

$$S = \int \mathrm{d}x^4 \frac{1}{g^2} F^{\mu\nu} F_{\mu\nu} + A_\nu^D \partial_\mu \tilde{F}^{\mu\nu} \tag{A.0.4}$$

On the level of the partition function the Lagrange multiplier implies a change of the functional integration measure: $\int \mathcal{D}F_{\mu\nu} \mathcal{D}\, A_\mu^D$ instead of $\int \mathcal{D}A_\mu$. In the presence of a magnetic source, the Maxwell equations would look like:

$$\begin{aligned} \partial_\mu F^{\mu\nu} &= J_e^\nu \\ \partial_\mu \tilde{F}^{\mu\nu} &= J_m^\nu, \end{aligned}$$

We now recognize the second term in (A.0.4) as corresponding to a (magnetic) photon coupling to a magnetic current. Indeed, interpreting A^D_μ as an actual dynamical field we may integrate the source term by parts. A Gaussian functional integral over $F_{\mu\nu}$ remains. Performing it, we find:

$$S = \int \mathrm{d}x^4 \frac{1}{g_D^2} F_D^{\mu\nu} F^D_{\mu\nu} + A^D_\nu \partial_\mu F_D^{\mu\nu} \tag{A.0.5}$$

where now $F^D_{\mu\nu} = \partial_\mu A^D_\nu - \partial_\nu A^D_\mu$ and $\frac{1}{g_D} = g$.

Comparing the action (A.0.1) with (A.0.5), we see few, yet important differences. First of all the similarities between the two actions suggest an equivalent physical theory. The variables, however, are fundamentally different in the sense that A_D couples to magnetic charges and A couples to electric charges. Furthermore, the magnetic theory has a coupling constant proportional to the inverse of the electric theory, exhibiting a strong-weak duality.

Appendix B

Supersymmetry Multiplets and Superfields Expansions

In this section we will discuss the $\mathcal{N} = 2$ BPS representations: the vector multiplet and the hypermultiplet.

B.1 Vector and Hypermultiplet

The component fields of the vector multiplet are given by:

$$\begin{array}{ccc} & A_\mu & \\ \lambda & & \psi \\ & \phi & \end{array}$$

The multiplet consists of a massless vector, two Weyl fermions and a complex scalar. It can be understood in $\mathcal{N} = 1$ language as the union of a chiral and vector multiplet.

The component fields of the hypermultiplet are given by:

$$\begin{array}{ccc} & \psi_q & \\ q & & \tilde{q}^\dagger \\ & \psi_{\tilde{q}}^\dagger & \end{array}$$

It consists of two complex scalars and two Weyl fermions.

The rows of the tables show the action of $SU(2)_R$ R-symmetry. Both multiplets are BPS saturated, meaning they are annihilated by half of the $\mathcal{N} = 2$ supercharges.

B.2 Expansion of Superfields

We will give the component expansions of the superfields. Conventions concerning factors of i, of 2 and minus signs differ in the literature. We adopt conventions from [5]. Other references are [27] and [37]. Irrespective of the conventions, the structure of the expansions is the same and this is all we are interested in.

The chiral multiplet is contained in the chiral superfield:

$$\Phi(y, \theta) = \phi(y) + \theta\chi(y) + \theta^2 F(y) \tag{B.2.1}$$

$$= \phi(x) + \theta\chi(x) + \theta^2 F(x) + i\bar{\theta}\bar{\sigma}^\mu\theta\partial_\mu\phi(x) - \frac{i}{2}\theta^2\bar{\theta}\bar{\sigma}^\mu\partial_\mu\chi(x) + \frac{1}{4}\theta^2\bar{\theta}^2\partial^\mu\partial_\mu\phi(x) \tag{B.2.2}$$

where $y^\mu = x^\mu - i\theta\sigma^\mu\bar{\theta}$ and in the second line we expanded around x^μ. $F(x)$ is an auxiliary field needed to obtain an off-shell realization of supersymmetry. This expansion is precisely the same for the chiral superfields $Q, \tilde{Q}$ making up the hypermultiplet.

The vector multiplet is contained in the real superfield:

$$V(x,\theta,\bar{\theta}) = -\bar{\theta}\bar{\sigma}^\mu\theta A_\mu(x) + i\theta^2\bar{\theta}\bar{\lambda}(x) - i\bar{\theta}^2\theta\lambda(x) + \frac{1}{2}\bar{\theta}^2\bar{\theta}^2 D(x) \tag{B.2.3}$$

where we have calculated the vector field in a particular gauge, the so-called Wess-Zumino gauge. Similarly as in the chiral multiplet, D is an auxiliary field. Due to the gauge, every term in the vector superfield contains some power of the anticommuting coordinates and we can consider:

$$e^V = 1 - \bar{\theta}\bar{\sigma}^\mu\theta A_\mu + i\theta^2\bar{\theta}\bar{\lambda} - i\bar{\theta}^2\theta\lambda + \frac{1}{2}\theta^2\bar{\theta}^2\left(D + \frac{1}{2}A_\mu A^\mu\right) \tag{B.2.4}$$

This particular form for the vector superfield is useful to describe interactions with chiral multiplets.

The gauge kinetic part is conventionally represented by the chiral superfield:

$$W_\alpha = -i\lambda_\alpha + \theta_\alpha D - \frac{i}{2}\left(\sigma^\mu\bar{\sigma}^\nu\theta\right)_\alpha F_{\mu\nu} + \theta^2\left(\sigma^\mu\partial_\mu\bar{\lambda}\right)_\alpha \tag{B.2.5}$$

which obeys a reality condition. For more details, and the relation between V and W we refer to [27].

Appendix C

Half Hypers, Bifundamentals and Trifundamentals

In this section we provide the Lagrangians for a half-hyper multiplet and subsequently use it to write down the Lagrangian for a trifundamental (half) hypermultiplet. Then, we take one decoupling limit at a time to arrive via the bifundamental and fundamental Lagrangian to the superpotential for the anfundamental. We follow the treatment as given in [34]. We use slightly different conventions: in [34] the Lie algebra is defined in mathematical conventions: the generators are antihermitian. We use physics conventions: hermitian generators.

As discussed in Section 2.3 and Section 3.8, a full $\mathcal{N}=2$ hypermultiplet for any gauge group is represented by two chiral multiplets:

$$Q_{fh} = \begin{pmatrix} Q \\ \tilde{Q}^* \end{pmatrix} \tag{C.0.1}$$

Our notation differs slightly from the one used in Section 2.4: we have redefined $\tilde{Q} \to \tilde{Q}^t$. These are the conventions of [34]. They are convenient especially to relate Q and $\tilde{Q}$ without the excessive use of daggers and other confusing notation.

An obvious way to reduce the degrees of freedom of such a hypermultiplet is to remove one of the chiral multiplets from which it is constructed. This is in a sense what one would call a half hypermultiplet. However, the price will be high: we will lose a manifest $\mathcal{N}=2$ supersymmetry. This loss of supersymmetry may be avoided by choosing a smarter constraint. We will relate Q and $\tilde{Q}$ in what resembles a Majorana constraint on a spinor to lose the superfluous degrees of freedom. To do just this, we define an antilinear involution τ. It is chosen to be antilinear since it has to map a representation R to its complex conjugate R^*. By demanding invariance under τ of the full hypermultiplet, we will obtain a half-hypermultiplet. Therefore, we want τ to be representation preserving. Since R is pseudoreal, τ, being antilinear, preserves the representation if it also contains a σ_2. Then τ is defined as:

$$\tau(Q_{fh}) = \sigma_2 \otimes \sigma_I (Q_{fh})^* \tag{C.0.2}$$

Here, σ_I acts on the components of Q_{fh} and reads:

$$\sigma_I = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \tag{C.0.3}$$

whereas σ_2 acts on the gauge doublets, the components within Q_{fh}. The complex conjugation has to be taken after the action of $\sigma_2 \otimes \sigma_I$. One can work out the action of τ explicitly to find:

$$\tau(Q_{fh}) = \begin{pmatrix} \sigma_2 \tilde{Q} & -\sigma_2 Q^* \end{pmatrix} \tag{C.0.4}$$

From this one may easily see $\tau^2 = 1$. Hence, the eigenvalues of τ are ± 1. From (C.0.4) it follows that eigenvectors of τ obey $\tilde{Q} = \pm\sigma_2 Q$. The half-hypermultiplet, written as a full hypermultiplet, is given by:

$$Q_{hh} = \begin{pmatrix} Q \\ \pm\sigma_2 Q^* \end{pmatrix} \tag{C.0.5}$$

with eigenvalues ± 1 under τ.

We can then construct the Lagrangian for the half hypermultiplet by plugging in the constraint (C.0.5) into the hypermultiplet Lagrangian from Section 2.3. We rewrite the kinetic term for $\tilde{Q}$ slightly, namely we transpose $\tilde{Q}$ and scale $-2V \to V$ to obtain:

$$S = \int \mathrm{d}\theta^2 \mathrm{d}\bar{\theta}^2 \left(Q^\dagger e^V Q + \tilde{Q}^t e^{-V} \tilde{Q}^* \right) + 2\sqrt{2}\mathrm{Re} \int \mathrm{d}\theta^2 \left(\tilde{Q}^t \Phi Q \right) \tag{C.0.6}$$

Let us apply the constraint in the kinetic term for $\tilde{Q}$:

$$\begin{aligned} \tilde{Q}^t e^{-V} \tilde{Q}^* &= Q^t \sigma_2^t e^{-V} \sigma_2^* Q^* \\ &= Q^t \sigma_2^{-1} e^{-V} \sigma_2 Q^* \\ &= Q^t e^{V^t} Q^* \\ &= Q^\dagger e^V Q \end{aligned} \tag{C.0.7}$$

Here, we used the fact that the vector superfield is expanded in a basis of hermitian generators T_a. Since the representation is pseudoreal, $\sigma_2 T_a \sigma_2 = -T_a^* = -T_a^t$. The total Lagrangian for a half hypermultiplet then equals:

$$S = \int \mathrm{d}\theta^2 \mathrm{d}\bar{\theta}^2 \left(2Q^\dagger e^V Q \right) \pm 2\sqrt{2}\mathrm{Re} \int \mathrm{d}\theta^2 \left(Q^t \sigma_2^t \Phi Q \right) \tag{C.0.8}$$

$\mathcal{N} = 2$ supersymmetry is preserved; the supercharges now relate Q, Q^* instead of $Q, \tilde{Q}^*$. It means that the $SU(2)_R$ now acts on the doublet $(q, \pm\sigma_2 q^*)$; the quark and the antiquark are now in the same supersymmetry multiplet instead of two separates.

Now, it is easy to construct a trifundamental half hypermultiplet, the building stone of the Gaiotto construction of $SU(2)$ quiver gauge theories. We simply couple the half hyper to three $SU(2)$ gauge fields:

$$\begin{aligned} S = &\int \mathrm{d}\theta^2 \mathrm{d}\bar{\theta}^2 \left(Q^*_{abc} e^{(V_1)^a{}_{a'}} Q^{a'bc} + Q^*_{abc} e^{(V_2)^b{}_{b'}} Q^{ab'c} + Q^*_{abc} e^{(V_3)^c{}_{c'}} Q^{abc'} \right) \\ &\pm \sqrt{2}\mathrm{Re} \int \mathrm{d}\theta^2 \left(Q_{abc} (\Phi_1)^{aa'} Q_{a'}{}^{bc} + Q_{abc} (\Phi_2)^{bb'} Q^a{}_{b'}{}^c + Q_{abc} (\Phi_3)^{cc'} Q^{ab}{}_{c'} \right) \end{aligned} \tag{C.0.9}$$

Notice one can lower and raise $SU(2)$ indices with the ϵ tensor. We use this to rewrite the superpotential for later convenience:

$$W = \epsilon^{bb'} \epsilon^{cc'} Q_{abc} (\Phi_1)^{aa'} Q_{a'b'c'} + \epsilon^{aa'} \epsilon^{cc'} Q_{abc} (\Phi_2)^{bb'} Q_{a'b'c'} + \epsilon^{aa'} \epsilon^{bb'} Q_{abc} (\Phi_3)^{cc'} Q_{a'b'c'} \tag{C.0.10}$$

A half hypermultiplet in the trifundamental representation effectively provides four half hypers for each of the gauge groups.

The Coulomb branch parameter associated to the symmetry breaking play the role of a bare mass for the bifundamental hyper. Sending this parameter to zero can be interpreted in two ways. One can say the gauge symmetry is restored and the situation has returned to the previous. However, as explained in Section 3.8, we can also see it as causing an enhanced flavour symmetry. This interpretation obviously holds only when the original gauge group was very weakly coupled. Both enhanced symmetries are $SU(2)$.

Let us now see how we can reduce these trifundamentals to bifundamentals and fundamentals. To reach a bifundamental we want to get rid of one of the gauge symmetries. Let us spontaneously break the first gauge symmetry by setting $(\Phi_1)^a{}_{a'} = m_1(\sigma_3)^a{}_{a'}$. To rewrite the superpotential in this form, we make use of two identities:

$$\begin{aligned}\epsilon^{bb'}\epsilon^{cc'}Q_{abc}(\Phi_1)^{aa'}Q_{a'b'c'} &= -Q_{abc}(\Phi_1)^a{}_{a'}Q^{a'bc}\\ &= \epsilon^{bb'}\epsilon^{cc'}\epsilon^{a'd'}Q_{abc}(\Phi_1)^a{}_{d'}Q_{a'bc}\\ &= m_1 Q_2^{bc}Q_{1bc} + m_1 Q_1^{bc}Q_{2bc}\\ &= 2m_1\tilde{Q}^{bc}Q_{bc}\end{aligned} \tag{C.0.11}$$

where we defined:

$$\begin{aligned}Q_{bc} &= Q_{1bc} = -Q^2{}_{bc}\\ \tilde{Q}_{bc} &= Q_{2bc} = Q^1{}_{bc}\end{aligned} \tag{C.0.12}$$

The remaining Yukawa couplings become in terms of (C.0.12):

$$\begin{aligned}\epsilon^{aa'}\epsilon^{cc'}Q_{abc}(\Phi_2)^{bb'}Q_{a'b'c'} &= \epsilon^{cc'}\left(-Q_{1bc}(\Phi_2)^{bb'}Q_{2b'c'} + Q_{2bc}(\Phi_2)^{bb'}Q_{1b'c'}\right)\\ &= -2\epsilon^{cc'}\tilde{Q}_{bc}(\Phi_2)^{bb'}Q_{b'c'}\end{aligned} \tag{C.0.13}$$

Here we used the convention $\epsilon^{12} = -\epsilon_{12} = -1$. Thus, (C.0.11) and (C.0.13) give:

$$W = 2m_1 Q_{bc}\tilde{Q}^{bc} - 2\epsilon^{cc'}\tilde{Q}_{bc}(\Phi_2)^{bb'}Q_{b'c'} - 2\epsilon^{bb'}\tilde{Q}_{bc}(\Phi_2)^{cc'}Q_{b'c'} \tag{C.0.14}$$

Interestingly, we have recovered a superpotential for a bifundamental hypermultiplet with bare mass m_1. It is easy to see that the kinetic terms also reduce to those of a bifundamental. After symmetry breaking, we are left with an abelian vector multiplet. This means within the vector multiplet, there will be no interactions since they all live in the adjoint. By $\mathcal{N} = 2$ supersymmetry, in low energy, there cannot be a coupling between the abelian vector multiplet and the bifundamental. Effectively then, we reduced a full $SU(2)$ gauge symmetry to a $U(1)$ global symmetry. To this $U(1)$ we may associate the bare mas parameter. Setting it to zero corresponds from the bifundamental hyper point of view as to an enhanced flavour symmetry. In this case, since the bifundamental is in a real representation of the gauge groups, it has an enhanced flavour symmetry of $USp(1) \cong SU(2)$. The derivation of this is very analogous to the derivation of the enhanced flavour symmetry of an ordinary fundamental given in Section 3.8.

We continue to Higgs the second gauge group as well: $(\Phi_2)^b{}_{b'} = m_2(\sigma_3)^b{}_{b'}$. Instead of continuing with the expression (C.0.14), we return to equation (C.0.10). We will massage it a little bit further into a convenient form. Note that the first term may be rewritten as follows:

$$\begin{aligned}\epsilon^{bb'}\epsilon^{cc'}Q_{abc}(\Phi_1)^{aa'}Q_{a'b'c'} &= -Q_a{}^{bc}(\Phi_1)^a{}_{a'}Q^{a'}{}_{b'c'}\\ &= Q_a{}^{bc}(\sigma_2)^f_{a'}(\Phi_1)^a_{a'}(\sigma_2)_e{}^a Q^{a'}{}_{b'c'}\\ &= (\Phi_1)^{a'}_a Q^{abc}Q_{a'bc}\end{aligned} \tag{C.0.15}$$

We use this notation now for both terms:

$$W = m_1(\sigma_3)^{a'}_a\delta^{b'}_b Q^{abc}Q_{a'b'c} + \delta^{a'}_a m_2(\sigma_3)^{b'}_b Q^{abc}Q_{a'b'c} + \epsilon^{aa'}\epsilon^{bb'}Q_{abc}(\Phi_3)^{cc'}Q_{a'b'c'} \tag{C.0.16}$$

The a and b indices should be read as flavour indices whereas the c index remains a gauge index. We see that this Lagrangian corresponds to four half hypermultiplets or two full hypermultiplets coupled to an $SU(2)$ gauge field. The bare masses of the four half hypers are given by all possible combinations of $\pm m_1 \pm m_2$ as may be seen directly upon summing over the indices or recognizing a tensor (Kronecker) product between the Kronecker delta and σ_3. If the bare masses vanish, there is an enhanced $SO(4)$ flavour symmetry.

Finally then, let us also break the last $SU(2)$ gauge symmetry by choosing: $(\Phi_3)^c{}_{c'} = m_3(\sigma_3)^c{}_{c'}$. The superpotential becomes:

$$W = m_1(\sigma_3)^{a'}{}_a \delta^{b'}_b \delta^{c'}_c Q^{abc} Q_{a'b'c'} + \delta^{a'}_a m_2(\sigma_3)^{b'}{}_b \delta^{c'}_c Q^{abc} Q_{a'b'c'} + \delta^{a'}_a \delta^{b'}_b m_3(\sigma_3)^{c'}{}_c Q^{abc} Q_{a'b'c'} \quad \text{(C.0.17)}$$

All gauge indices have turned into flavour indices. For each chosen triple of $SU(2)$ indices we have a half hyper with mass $\pm m_1 \pm m_2 \pm m_3$ where for instance $m_1 + m_2 + m_3$ corresponds to Q_{111}.

The theory described by (C.0.17) has no kinetic terms. It is the $\mathcal{T}_{3,0}$ building block of Gaiotto represented by Figure C.0.1. Notice that we can associate the mass parameters of the of the flavour

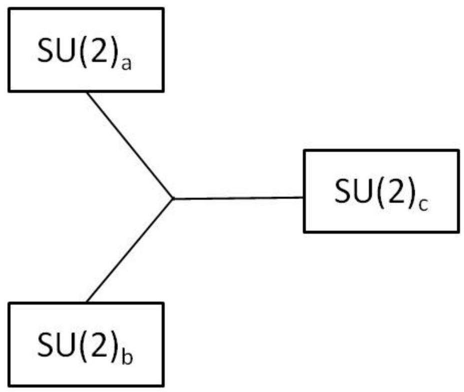

Figure C.0.1: Quiver representation of.

groups to the Coulomb branch parameters. The mass parameters break the flavour symmetry to its Cartan subgroup.

Appendix D

Explicit Expression Quadratic Differential

$$\lambda^2 = \frac{\frac{1}{4}\left((m_1+m_2)t^2 + \frac{1}{a}\frac{1}{\left(\frac{-1-\sqrt{1-4a}}{2a}\right)^2}(m_3+m_4)\right)^2}{(t-1)^2(t-q(a))^2t^2}$$

$$- \frac{(t-1)(t-q(a))\left(m_1m_2t^2 + \frac{u}{a}\frac{1}{\left(\frac{-1-\sqrt{1-4a}}{2a}\right)}t + \frac{1}{a}\frac{1}{\left(\frac{-1-\sqrt{1-4a}}{2a}\right)^2}m_3m_4\right)}{(t-1)^2(t-q(a))^2t^2}\mathrm{d}t^2$$

with

$$q(a) = \frac{-1+\sqrt{1-4a}}{-1-\sqrt{1-4a}}$$

Bibliography

[1] S. Theisen A. Brandhuber, J. Sonnenschein and S. Yankielowicz. M-theory and seiberg-witten curves: Orthogonal and symplectic groups. *Nuclear Physics B*, 504:175–188, 1997.

[2] S. Theisen A. Klemm, W. Lerche and S. Yankielowicz. Simple singularities and $\mathcal{N} = 2$ supersymmetric yang-mills theory. *Physics Letters*, B344:169–175, 1995.

[3] S.L. Adler and W.A. Bardeen. Absence of Higher Order Corrections in the Anomalous Axial Vector Divergence Equation. *Phys.Rev.*, 182:1517–1536, 1969.

[4] L. Alvarez-Gaume and S.F. Hassan. Introduction to s-duality in $\mathcal{N} = 2$ supersymmetric gauge theories. Lecture Notes, 1997.

[5] L. Alvarez-Gaume and F. Zamora. Duality in quantum field theory and string theory. 1997.

[6] P.C. Argyres. S-duality and Global Symmetries in $\mathcal{N} = 2$ Supersymmetric field theory. *Adv.Theor.Math.Phys.*, 2:293–302, 1998.

[7] P.C. Argyres and A. Buchel. New S dualities in $\mathcal{N} = 2$ Supersymmetric $SU(2) \times SU(2)$ Gauge Theory. *JHEP*, 9911:014, 1999.

[8] P.C. Argyres and M.R. Douglas. New Phenomena in $SU(3)$ Supersymmetric Gauge Theory. *Nucl.Phys.*, B448:93–126, 1995.

[9] P.C. Argyres and A.E. Faraggi. The vacuum structure and spectrum of $\mathcal{N} = 2$ supersymmetric $su(n)$ gauge theory. *Physics Review Letters*, 74:3931–3934, 1995.

[10] P.C. Argyres and N. Seiberg. S-duality in $\mathcal{N} = 2$ Supersymmetric Gauge Theories. *JHEP*, 0712:088, 2007.

[11] P.C. Argyres and A.D. Shapere. The Vacuum Structure of $\mathcal{N} = 2$ superQCD with Classical Gauge Groups. *Nucl.Phys.*, B461:437–459, 1996.

[12] N. Arkani-Hamed and H. Murayama. Holomorphy, Rescaling Anomalies and Exact Beta Functions in Supersymmetric Gauge Theories. *JHEP*, 0006:030, 2000.

[13] P. Arvidsson. Superconformal Theories in Six Dimensions. 2006.

[14] L. Bhardwaj and Y. Tachikawa. Classification of 4d $\mathcal{N} = 2$ Gauge Theories. *JHEP*, 1312:100, 2013.

[15] A. Bilal. Duality in $\mathcal{N} = 2$ susy su(2) yang-mills theory: A pedagogical introduction to the work of seiberg and witten. pages 21–43, 1995.

[16] G. Bonelli and A. Tanzini. Hitchin systems, $\mathcal{N} = 2$ Gauge Theories and W-gravity. *Phys.Lett.*, B691:111–115, 2010.

[17] O. Chacaltana and J. Distler. Tinkertoys for gaiotto duality. *JHEP*, 1011:099, 2010.

[18] K. Chandrasekharan. *Elliptic Functions*. Springer-Verlag, 1st edition, 1985.

[19] S. Coleman. *Aspects of Symmetry*. Cambridge University Press, 6th edition, 1985.

[20] G.W. Moore D. Gaiotto and A. Neitzke. Wall-crossing, Hitchin Systems, and the WKB Approximation. 2009.

[21] E. D'Hoker and D.H. Phong. Seiberg-Witten Theory and Integrable Systems. 1999.

[22] Y. Tachikawa F. Benini and B. Wecht. Sicilian gauge theories and $\mathcal{N} = 1$ dualities. *JHEP*, 1001:088, 2010.

[23] B. Farb and D. Margalit. *A Primer on Mapping Class Groups*. Princeton University Press, 5.0 edition, 2011.

[24] H.M. Farkas and I. Kra. *Riemann Surfaces*. Springer-Verlag New York, 2nd edition, 1992.

[25] A. Fayyazuddin. Some Comments on $\mathcal{N} = 2$ Supersymmetric Yang-Mills. *Mod.Phys.Lett.*, A10:2703–2708, 1995.

[26] F. Ferrari and A. Bilal. The strong-coupling spectrum of the seiberg-witten theory. *Nuclear Physics B*, 469:387–402, 1996.

[27] J.M. Figueroa-O'Farrill. Busstepp lectures on supersymmetry. Summer School Lectures, 2000.

[28] K. Maruyoshi G. Bonelli and A. Tanzini. Wild Quiver Gauge Theories. *JHEP*, 1202:031, 2012.

[29] D. Gaiotto. $\mathcal{N} = 2$ dualities. *JHEP*, 1208:034, 2012.

[30] D. Gaiotto and J. Maldacena. The Gravity Duals of $\mathcal{N} = 2$ Superconformal Field Theories. *JHEP*, 1210:189, 2012.

[31] M. Henningson and P. Yi. Four-dimensional bps spectra via m-theory. *Phys.Rev.*, D57:1291–1298, 1998.

[32] A. Kashani-Poor and J. Troost. Transformations of Spherical Blocks. *JHEP*, 1310:009, 2013.

[33] S. Ketov. Solitons, monopoles and duality: From sine-gordon to seiberg-witten. Lecture Notes, 1996.

[34] C.A. Keller L. Hollands and J. Song. Towards a 4d/2d correspondence for sicilian quivers. *JHEP*, 1110:100, 2011.

[35] W. Lerche. Introduction to Seiberg-Witten Theory and Its Stringy Origin. *Nucl.Phys.Proc.Suppl.*, 55B:83–117, 1997.

[36] D. Gaiotto L.F. Alday and Y. Tachikawa. Liouville Correlation Functions from Four-dimensional Gauge Theories. *Lett.Math.Phys.*, 91:167–197, 2010.

[37] J.D. Lykken. Introduction to supersymmetry. Lectures, 1996.

[38] N. Haouzi M. Aganagic and S. Shakirov. A_n-Triality. 2014.

[39] A.I. Vainshtein M.A. Shifman and V.I. Zakharov. Exact Gell-Mann-Low Function in Supersymmetric Electrodynamics. *Phys.Lett.*, B166:334, 1986.

[40] A. Mikhailov. Bps states and minimal surfaces. *Nucl.Phys.*, B533:243–274, 1998.

[41] R. Miranda. *Algebraic Curves and Riemann Surfaces.* American Mathematical Society, 1st edition, 1995.

[42] V. Khoze N. Dorey and M.P. Mattis. Multi-instanton Check of the relation between the Prepotential $\mathcal{F}$ and the modulus u in $\mathcal{N} = 2$ SUSY Yang-Mills theory. *Phys.Lett.*, B390:205–209, 1997.

[43] V.V. Khoze N. Dorey and M.P. Mattis. Multi-instanton calculus in $\mathcal{N} = 2$ supersymmetric gauge theory. 2. coupling to matter. *Physical Review D*, 54:7832–7848, 1996.

[44] V.V. Khoze N. Dorey and M.P. Mattis. On $\mathcal{N} = 2$ supersymmetric qcd with 4 flavors. *Nuclear Physics B*, 492:607–622, 1997.

[45] N. A. Nekrasov. Seiberg-Witten Prepotential from Instanton Counting. *Adv.Theor.Math.Phys.*, 7:831–864, 2004.

[46] D.I. Olive and P.C. West. Duality and Supersymmetric Theories. Proceedings, Easter School, Newton Institute, Euroconference, Cambridge, UK, April 7-18, 1997. 1999.

[47] H. Osborn. Topological Charges for $\mathcal{N} = 4$ Supersymmetric Gauge Theories and Monopoles of Spin 1. *Phys.Lett.*, B83:321, 1979.

[48] A.D. Shapere P.C. Argyres, M. Ronen Plesser. The coulomb phase of $\mathcal{N} = 2$ supersymmetric qcd. *Physics Review Letters*, 75:1699–1702, 1995.

[49] A.E. Faraggi P.C. Argyres and A.D.Shapere. Curves of Marginal Stability in $\mathcal{N} = 2$ superQCD. 1995.

[50] M. E. Peskin. Duality in supersymmetric yang-mills theory. pages 729–809, 1997.

[51] J. Polchinski. *String Theory Vol. 1: An Introduction to the Bosonic String.* Cambridge University Press, 1st edition, 1998.

[52] J. Polchinski. *String Theory Vol. 2: Superstring Theory and Beyond.* Cambridge University Press, 1st edition, 1998.

[53] L. O'Raifeartaigh I. Sachs R. Flume, M. Magro and O. Schnetz. Uniqueness of the Seiberg-Witten Effective Lagrangian. *Nucl.Phys.*, B494:331–345, 1997.

[54] R. Rajaraman. *Solitons and Instantons.* North-Holland Publishing Company, 1st edition, 1982.

[55] R.J. Rodger. A pedagogical introduction to the agt conjecture. Master Thesis, 2013.

[56] M. Stern S. Sethi and E. Zaslow. Monopole and Dyon bound states in $\mathcal{N} = 2$ supersymmetric Yang-Mills theories. *Nucl.Phys.*, B457:484–512, 1995.

[57] N. Seiberg. Supersymmetry and non-perturbative beta functions. *Physics Letters B*, 206(1), 1988.

[58] N. Seiberg. Naturalness versus supersymmetric non-renormalization theorems. *Physical Letters*, B318:469–475, 1993.

[59] N. Seiberg and E. Witten. Electric-magnetic duality, monopole condensation and confinement in $\mathcal{N} = 2$ supersymmetric yang-mills theory. *Nuclear Physics*, B426:19–52, 1994.

[60] N. Seiberg and E. Witten. Monopoles, duality and chiral symmetry breaking in $\mathcal{N} = 2$ supersymmetric qcd. *Nuclear Physics*, B431:484–550, 1994.

[61] M.A. Shifman and A.I. Vainshtein. Solution of the Anomaly Puzzle in SUSY Gauge Theories and the Wilson Operator Expansion. *Nucl.Phys.*, B277:456, 1986.

[62] J.H. Silverman. *The Arithmetic of Elliptic Curves.* Springer, 2nd edition, 2009.

[63] M.J. Strassler. An unorthodox introduction to supersymmetry. TASI Lectures, 2001.

[64] G. 't Hooft. Topology of the gauge condition and new confinement phases in non-abelian gauge theories. *Nuclear Physics*, B190:455–478, 1981.

[65] Y. Tachikawa. $\mathcal{N} = 2$ supersymmetric dynamics for dummies. Lecture Notes, 2013.

[66] Yuji Tachikawa. Some illustrative examples of argyres-seiberg-gaiotto duality. *Journal of Physics: Conference Series*, 462(1):012052, 2013.

[67] J. Terning. *Modern Supersymmetry: Dynamics and Duality.* Oxford University Press, 1st edition, 2006.

[68] J. Wess and J. Bagger. *Supersymmetry and Supergravity.* Princeton University Press, 2nd edition, 1992.

[69] P.C. West. *Introduction to Supersymmetry and Supergravity.* World Scientific Publishing Co, 1st edition, 1986.

[70] E.T. Whittaker and G.N. Watson. *A Course of Modern Analysis.* Cambridge University Press, 4th edition, 1950.

[71] E. Witten. Dyons of Charge $\frac{e\theta}{2\pi}$. *Phys.Lett.*, B86:283–287, 1979.

[72] E. Witten. Solutions of Four-Dimensional Field Theories via M-theory. *Nucl.Phys.*, B500:3–42, 1997.

[73] E. Witten and D. Olive. Supersymmetry algebras that include topological charges. *Physics Letters*, 78B(1), 1978.

[74] Scott A. Wolpert. Asymptotics of the spectrum and the selberg zeta function on the space of riemann surfaces. *Communications in Mathematical Physics*, 112(2):283–315, 1987.

[75] N. Wyllard. A_{N-1} conformal Toda field theory correlation functions from conformal $\mathcal{N} = 2$ $SU(N)$ quiver gauge theories. *JHEP*, 0911:002, 2009.

Acknowledgements

Over the course of the last year, there is a number of people I want to thank for their help and support.

First and foremost, I want to thank my supervisor Erik Verlinde for introducing me to the subject of supersymmetric gauge theories and giving me the freedom and confidence to explore this rich topic in theoretical physics. Especially giving me the chance to explain to him the recent developments was very motivating and caused me to have dinner quite frequently at the science park restaurant.

Furthermore, I would like to thank Paul de Lange for suggesting the subject of Seiberg-Witten theory in the first place and for being available whenever I had questions. For the same reason, I would like to thank Satoshi Nawata: for suggesting particularly useful literature and for always making time to answer numerous questions, whether during office hours, very at night or on an early Sunday morning. Also, I want to thank him for letting me into his own research. Although I have yet to understand his ideas, I do hope sometime in the future we could collaborate on one thing or the other. Also my gratitude goes out to Miranda Cheng and Irfan Ilgin, who helped me one way or the other at later stages of the project.

The master students at the fourth floor are sincerely acknowledged for providing on the one hand a motivating environment to study yet also good company outside whenever we are outside of the room. I will miss our group of students.

Without the love and support of my friends, brothers and parents last year would have been impossible. Last but not least, my deepest gratitude is reserved for Giselle.

Populaire Samenvatting

In de afgelopen eeuw is onze kennis van de natuur dramatisch toegenomen. En van de grote ontdekkingen is de quantum theorie geweest, de taal waarin de natuur zich op zeer kleine schalen laat beschrijven. Men kan stellen dat het belangrijkste product van het quantum-denken het standaard model is, dat beschrijft hoe de op dit moment aan ons bekende bouwstenen van de natuur met elkaar wisselwerken. Theoretische voorspellingen die op een groot aantal cijfers achter de komma overeenkomen met gemeten waardes zijn het standaardmodel niet vreemd. De grootste experimenten op aarde meten aan het standaard model, meest recent het LHC experiment in Genve. Maar niet alleen de significantie is opmerkelijk, ook de interne logica lijkt natuurgetrouw. Recent is dit opnieuw gebleken bij de vondst van het voorspelde Higgs-boson.

Ondanks al deze triomfen zijn er toch nog onbevredigende aspecten aan het standaard model. Men hoort vaak in de populaire media over de clash tussen de quantum theorie en de theorie van de zwaartekracht, maar zelfs binnen het standaard model, een consistente quantum theorie, zijn er onbegrepen zaken. Men kan het standaard model abstract en sluitend formuleren, doch zodra men het zijn gang laat gaan gebeuren er zaken waar wij wiskundig gezien maar moeilijk vat op kunnen krijgen. En van de zaken die we eigenlijk niet begrijpen is de zogenaamde opsluiting van quarks. Ook al durven we te zeggen dat we de natuur op haar fundamentele schaal begrijpen als bestaande uit onder andere quarks, blijkt op atomaire schalen dat quarks niet meer als elementaire bouwstenen maar slechts in pakketjes van twee (bijv. pionen) of drie (bijv. het proton en neutron) voorkomen. Men zegt ook wel dat de theorie die interacties tussen de quarks beschrijft op de schaal van het proton sterk gekoppeld is. Onze wiskundige beschrijving, visueel begrepen in de vorm van Feynman diagrammen, is alleen toepasbaar in het regime van de zwak gekoppelde systemen. Om de dynamica van sterk gekoppelde systemen te begrijpen zijn andere wiskundige methodes nodig. En methode heet (S)-dualiteit. Dualiteit wordt hier begrepen als de equivalentie van twee fysische beschrijvingen van een enkele, onderliggende theorie. Een bekend voorbeeld hiervan is de elektrisch-magnetische dualiteit, het duidelijkst te zien in de bronvrije Maxwell vergelijkingen die geheel symmetrisch zijn in het elektrisch en magnetisch veld. In de aanwezigheid van magnetische monopolen kan sterker worden gesteld dat onder uitwisseling van de elektrische en magnetische geladen vrijheidsgraden en de elektrisch en magnetisch velden een equivalente beschrijving van elektromagnetisme wordt verkregen. Het enige verschil schuilt in het feit dat de constante die de sterkte van de koppeling bepaalt genverteerd wordt. Dit betekent dat een sterk gekoppelde beschrijving ingewisseld wordt voor een zwak gekoppelde beschrijving en vice versa. Zo een S-dualiteit zou ons dus kunnen helpen de sterk gekoppelde systemen te begrijpen in termen van nieuwe variabelen. Merk op dat zo een dualiteit volledig discontinu is; wiskundig gezien kunnen we met S-dualiteit theorien slechts begrijpen wanneer deze extreem zwak of extreem sterk gekoppeld zijn en is er in principe geen (fysische) interpolatie tussen de extremen.

De grote ontdekking van Seiberg en Witten is geweest dat een enkele (supersymmetrische) quantumvelden theorie op drie gesoleerde punten als zwak gekoppeld kan worden begrepen in termen van telkens andere variabelen. Bovendien, de pracht van de ontdekking komt naar voren in het feit dat de interpolatie tussen deze drie punten geometrisch van aard is. Meer precies kan de fysische theorie begrepen worden als een vervormbare donut. Slechts wanneer n van de twee cirkels waaruit de donut bestaat een oneindige lengte krijgt, bestaat er een mathematische beschrijving van de fysica. Echter, de interpolatie tussen de extremen wordt nu begrepen als de vervorming van een donut. Ook al is het niet direct duidelijk hoe dit zich in fysica laat vertalen, wel levert dit inzicht een duidelijker beeld van het gebied tussen de extremen.

Het is niet de eerste keer dat geometrie opduikt als een natuurlijke beschrijving van fysische systemen. Eerder werd door Einstein gerealiseerd dat zwaartekracht een puur geometrische verklaring heeft. Het is echter wel opmerkelijker dat geometrie in quantumvelden theorien opduikt dan in de verklaring voor de zwaartekracht, aangezien zwaartekracht als eigenschap van ruimtetijd zelf wordt gezien, van nature geometrische begrippen, terwijl de dynamica van quantumvelden theorien niet direct lijkt gerelateerd aan geometrie. Toch lijken er nauwe verbanden te zijn tussen de geometrische vormen van Seiberg en Witten, en Gaiottos werk dat daar op heeft voortgeborduurd, en ruimtetijd zelf. Het onderzoeken van deze relatie is niet onderdeel van de scriptie. Wel wordt er een uitgebreide beschrijving gegeven van het werk van Seiberg, Witten en Gaiotto waarin geometrie de hoofdrol speelt en de verschillende extremen met elkaar verbindt.