Integrable Structure in Low Energy Effective Supersymmetric Gauge Theory

Integrable Structure in Low Energy Effective Supersymmetric Gauge Theory

Academisch Proefschrift

ter verkrijging van de graad van doctor

aan de Universiteit van Amsterdam

op gezag van de Rector Magnificus

prof. mr. P. F. van der Heijden

ten overstaan van een door het college voor promoties ingestelde

commissie, in het openbaar te verdedigen in de

Aula der Universiteit

op vrijdag 14 Oktober, te 12:00 uur

door

Rutger Herman Boels

geboren te Leimuiden.

Promotiecommissie

Promotor

prof. dr. J. de Boer

Overige leden

prof. dr. P. van Baal

prof. dr. R.H. Dijkgraaf

prof. dr. T.J. Hollowood

prof. dr. E.P. Verlinde

dr. J.W. van de Leur

dr. K. Skenderis

dr. M.M. Taylor

Faculteit der Natuurwetenschappen, Wiskunde en Informatica

This work is part of the research program of the "Stichting voor Fundamenteel Onderzoek der Materie (FOM)", which is financially supported by the "Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO)".

This thesis is based on the following publications:

- R.H. Boels, J. de Boer, R.J. Duivenvoorden and J.S. Wijnhout,
  Non-perturbative superpotentials and compactification to three dimensions,
  Journal of High Energy Physics **03** (2004) 009 [hep-th/0304061].

- R.H. Boels, J. de Boer, R.J. Duivenvoorden and J. Wijnhout,
  Factorization of Seiberg-Witten curves and compactification to three dimensions,
  Journal of High Energy Physics **03** (2004) 010 [hep-th/0305189].

- R.H. Boels and J. de Boer,
  Classical spin chains and exact three-dimensional superpotentials,
  Nuclear Physics B **715** (2005) 234-274 [hep-th/0411110].

Contents

Chapter 1

Introduction

Over the last decades it has become increasingly clear that the standard model of particle physics does a fantastically good job of describing the known basic forces in nature, apart of course from gravity which is described by the separate theory of general relativity. However, much of this understanding is based on a perturbation theory in the various coupling constants and a natural question is what happens when some or even all of these couplings become large. This is a problem in particular in understanding the strong nuclear force, which among other things should bind quarks into protons. The quantum field theory which describes this force, quantum chromodynamics (QCD) is weakly coupled on small distance scales and becomes strongly coupled when viewed on larger distance scales. The 'large' and 'small' in the previous sentence refer to a distance scale of about the radius of a proton. This phenomenon of weak coupling at short distance scales is called asymptotic freedom and is of fundamental importance in describing high energy phenomena. However, when one is interested in, for instance, how exactly quarks are bound into protons or why never lone particles with color charge are seen, we need to know how the theory behaves at larger distance scales where perturbation techniques are not applicable.

A possible resolution to this problem is to put the theory on a lattice which discretizes space-time and try to solve the problem by doing computer calculations. This is however technically difficult to do in four space-time dimensions, especially if one is interested in dynamical chiral quarks. Also, ideally one would like to have an *analytic* understanding of the low energy structure of QCD. An obvious other avenue of attack is then the study of simplified models which are as close as possible to the model we are interested in while being completely solvable. A spectacularly successful class of models can be obtained by introducing supersymmetry, a symmetry which interchanges bosons and fermions. A surprise in the study of these models at the low energy scales we are interested in has been that they are related in a natural way to particular classical completely integrable systems known from classical mechanics and much-studied in mathematics. Integrable systems are multi-particle systems in classical mechanics which in a precise sense contain so much symmetry that they are actually in principle completely solvable, which is a very rare property for n-body systems. This thesis is a part of the study of the interconnection of these integrable systems and supersymmetric gauge theories at low energies.

The above paragraph gives a direct physical motivation for the study of supersymmetric gauge theories in four dimensions: supersymmetric gauge theory as a laboratory or toy-model for understanding strong coupling effects in gauge theories. There are however more reasons to study (supersymmetric) gauge theories and their relation to integrable systems, which we will list in order of distance to physics (as perceived by the author):

- Supersymmetry may have a more direct connection to the standard model of particle physics than as a toy model: it is widely believed that the standard model contains some form of broken supersymmetry which is restored at sufficiently high energy scales, which might even be as low as a (few) TeV. If the low energy supersymmetry scenario is realized in nature, it may be probed by upcoming experiments.

- Supersymmetry is a one of the usual main components of string theory, both as a symmetry of the world-sheet theory and as a symmetry of the effective space-time theory. Historically, advances in string theory triggered advances in the understanding of (supersymmetric) field theory and vice-versa since four-dimensional supersymmetric field theory arises in several different contexts within string theory. A prime recent example of this is for instance the Dijkgraaf-Vafa construction which was motivated from geometric engineering in string theory. It makes use of a Hermitian matrix model to construct a potential which describes the vacuum structure of $\mathcal{N} = 1$ supersymmetric gauge theory and will be discussed in the next chapter of this thesis.

- Closely related to the above point is the fact that integrable systems were encountered in string theory in the early days of the theory where they arose in the study of minimal models in $c \leq 1$ non-critical bosonic string theory. The integrable systems arising here are mainly integrable hierarchies of the KP and Toda type. Interestingly, these systems are found in the double scaling limit of certain matrix models which are closely related to the ones used in the Dijkgraaf-Vafa construction. A natural question is then what the precise integrable structure of that construction is.

- There is a large mathematical interest in integrable systems which has led to a large literature which can be mined for techniques and results to be used in the above sketched physical contexts.

- Finally, one might be able to turn the last item around and use physical intuition from the large literature on supersymmetric field theory to study the mathematical structure of integrable systems.

Although the above motivations are also interesting, in this introduction the focus will be on the physical motivation that was given first in this chapter: supersymmetric field theory as a model. First we will review some of the expected properties of QCD in the low energy limit and introduce one of the main tools in studying this limit: the low energy effective action.

1.1 QCD in the low energy limit

Quantum chromodynamics is part of the standard model of particle physics. In this section we give a very brief account of its most salient properties, see e.g. [1] for further details. It is a gauge theory with non-Abelian gauge group $SU(3)$ and contains 6 flavors of quarks all of which have non-zero masses. Its classical action can be written as

$$S_{\text{class}} = -\frac{1}{g^2}\int d^4x \frac{1}{4} F^a_{\mu\nu} F^{\mu\nu}_a - \sum_f \bar{\psi}_f \left(i\not{D} - m_f\right)\psi_f, \tag{1.1}$$

with

$$\begin{aligned} F^a_{\mu\nu} &= \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + f^a_{bc} A^b_\mu A^c_\nu && (1.2)\\ \not{D}\psi_f &= \gamma^\mu(\partial_\mu - iA^a_\mu t^a)\psi_f && (1.3)\end{aligned}$$

and t^a matrices which form a unitary irreducible representation of the algebra of the gauge group. For physical quarks this representation is the fundamental representation, while the gluons transform in the adjoint. Apart from the local $SU(3)$ gauge symmetry and the global Lorentz symmetry, this theory enjoys a global $SU(6) \times SU(6)$ non-anomalous flavor symmetry in the limit of vanishing quark masses. Equal mass terms for different quark flavors break this symmetry to a diagonal $SU(6)$ subgroup and unequal mass terms break the symmetry further. The charge of the gauge group is often referred to as color, which leads to the name chromodynamics.

The above action can be used for instance in calculations of physical quantities in renormalized perturbation theory. Higher order effects in this perturbation famously generate a non-trivial *negative* β function which indicates asymptotic freedom. The β function describes exactly how the coupling constant depends on the scale at which the theory is probed. The exact value of the coupling constant must be determined by experiment and this introduces a scale into the theory. The fact that the β function is negative indicates that the theory is weakly coupled at high energies. The theory becomes strongly coupled and perturbation theory breaks down at small energy scales. The scale at which the coupling constant is of order 1 in QCD is about 250 MeV.

At these scales (and lower) however quarks are bound into protons and mesons and therefore contain very interesting physics, which is inaccessible to direct analytic analysis through perturbation theory. A time-tested approach to physical problems inaccessible by existing mathematical technology as these is the formulation of effective models based on physical intuition and consistency conditions in which various expected properties of the theory and their interrelation can be studied. In the context of high energy physics physical intuition usually takes the form of an analysis of symmetry properties. The idea is to construct a Lagrangian which is invariant under all the expected symmetry properties of the (low energy) vacuum and use this to derive vacuum expectation values and scattering amplitudes which should be valid in the low energy approximation.

1.1.1 Effective actions

Before the expected general properties of vacua in non-Abelian gauge theories will be discussed, let us consider a particular class of low energy effective actions for QCD as an introductory example (see for instance [1], chapter 19, or [2]).

Since the masses of the up and down quark are much lighter than the QCD scale, these flavors are usually singled out and taken to be massless as a first approximation. In the limit of vanishing mass for the up and down quarks the theory has a $SU(2) \times SU(2) \times U(1)$ non-anomalous flavor symmetry. From experiment we know that the non-diagonal (chiral) part of this symmetry is realized in the vacuum in the Goldstone mode: the pions are the Goldstone bosons which eventually get a small (relative to the QCD scale) mass by the explicit symmetry breaking terms in the Lagrangian. We therefore like to write down a theory which is Lorentz-invariant, is invariant under $SU(2) \times SU(2) \times U(1)$, contains scalar degrees of freedom (since the pions are scalars) and which produces spontaneous symmetry breaking. Since $SO(4) = SU(2) \times SU(2)$, we can start off with the $SO(4)$ sigma-model,

$$\mathcal{L} = -\frac{1}{2}\partial_\mu \phi_n \partial^\mu \phi^n + \frac{m}{2}\left(\phi_n \phi^n\right) - \frac{\lambda}{4}\left(\phi_n \phi^n\right)^2 , \tag{1.4}$$

and take m and λ to be positive. Note that higher derivative or higher order terms added to this Lagrangian would be irrelevant in the renormalization group sense, so are suppressed by the UV cutoff scale which, in this case, is expected to be close to the QCD scale. The field ϕ have four components which we can think of as being analogous to a Euclidean four-vector. The vacuum ϕ_n^0 of this theory has non-zero length. If the vacuum is chosen to be given by $\phi_n^0 = \{0, 0, 0, |\phi^0|\}$, it is easy to see that this breaks $SO(4)$ down to $SO(3) \sim SU(2)$ and identifies the rotations of the first three components as the left-over isospin symmetry group. Any isospin transformation of the vacuum leaves it invariant, so we have a space of vacua called a moduli space. The massless excitations of the vacuum are the Goldstone bosons, so we would like to split off the dynamics of these degrees of freedom. To do this, note that at every point the field can be written as an $SO(4)$ symmetry transformation of a fixed vector, say $\bar{\phi}_n(x) = \{0, 0, 0, \sigma(x)\}$, so that

$$\phi_n = R_{nm}\bar{\phi}_m = R_{n4}\sigma. \tag{1.5}$$

Plugging this back into (1.4) and using the relation $RR^T = 1$ we get

$$\mathcal{L} = -\frac{1}{2}\partial_\mu\sigma\partial^\mu\sigma - \frac{1}{2}\sigma^2\partial_\mu R_{n4}\partial^\mu R_{n4} + \frac{m}{2}\sigma^2 - \frac{\lambda}{4}\sigma^4, \tag{1.6}$$

Note that the kinetic term for the R_{n4} fields should contain the degrees of freedom describing the pion which is understood as fluctuations of the non-trivial vacuum. The kinetic term is proportional to the energy of the pions. Now we want to integrate out the dynamics of the vacuum field σ. Since we are only interested in describing the effective degrees of freedom at low energy we can expand in powers of R, and obtain in first order an effective Lagrangian

$$\mathcal{L}_{eff} \sim -\frac{1}{2}\frac{m}{\lambda}\partial_\mu R_{n4}\partial^\mu R_{n4} + \text{h.o.} \tag{1.7}$$

If one works out the explicit connection to the pion field by the symmetry properties one gets an effective description of the pion field which reproduces low energy pion scattering reasonably well. Actually, to get a good description also the small explicit chiral symmetry breaking mass terms must be incorporated. This changes the description as shown above in the sense that an effective potential is added to system: instead of a moduli space of vacua a preferred direction for the vacuum is chosen. This is called vacuum alignment, an idea which will come back in the next chapter.

It is important to note that in the above example the input was formed by an analysis of the symmetries of the problem, an analysis of the modes in which these symmetries were realized and a focus on the low-energy approach. Actually, in this example one does not need to know anything about the high energy theory (QCD): the symmetry and input on symmetry breaking pattern constrains the problem so much it is solvable in terms of just a few unknowns. However, this type of analysis does for instance not explain how exactly the chiral symmetry of the problem is dynamically spontaneously broken. Ideally, one would like to derive the low energy effective action directly from the QCD path integral by integrating out all the physics above a certain cutoff scale. It is not known for pure QCD how to do this because one would need to understand the infrared behavior of the theory.

If supersymmetry is introduced into the problem however, this constrains the theory much more as we will see in more detail in chapter 2. Seiberg and Witten [3, 4] were able to show that for a four-dimensional $SU(2)$ gauge theory with the ingredients (1) extended $\mathcal{N} = 2$ supersymmetry, (2) a knowledge of the symmetries involved, (3) a perturbative calculation and (4) some consistency conditions basically fixes the full low energy effective action. Just as in the pion Lagrangian example, the vacuum in this case is actually a moduli space of vacua, but the construction in the supersymmetric case is much more involved. Since the physical interest is in a theory with no supersymmetry, one might wonder what happens if supersymmetry is broken. As a first step in this direction the extended supersymmetry may be broken to $\mathcal{N} = 1$ supersymmetry. A nice way of doing this is giving a mass to part of the particles in the theory such that the remaining low energy theory only contains $\mathcal{N} = 1$ fields. It turns out that, just as for vacuum alignment in the pion effective Lagrangian, the moduli space of vacua is lifted and a specific vacuum is chosen. There is however in this case much more control over the vacuum itself. If the partial supersymmetry breaking analysis is done in Seiberg-Witten theory one can show that certain magnetic monopoles must condense, which is in agreement with the expected properties of the theory as described in the next subsection.

1.1.2 Expected properties

QCD in the low energy limit is expected to have several distinctive qualitative properties. In the example above we already encountered one: the theory must display spontaneous chiral symmetry breaking. Another very basic property is that it must have a mass-gap. This is the statement that the exact spectrum of the theory must not include any massless states. The most mediagenic property however is that of color confinement, which states that the asymptotic states of the theory should be singlets (colorless) under the gauge group.

Note that the above property could in principle be a consequence of the lightness of the lightest quarks: when pulling two quarks apart it might be energetically more favorable to pull a new quark anti-quark pair from the vacuum, creating two quark pairs instead of ’free’ quarks. This mechanism is referred to as dynamical screening. However, based on lattice studies and other arguments (like the aforementioned Seiberg-Witten argument in supersymmetric gauge theories) it is expected that even in the absence of light quark matter non-Abelian gauge theories also confine in the sense explained below.

One way of putting the problem is that it must be explained why the low energy vacuum of a non-Abelian gauge theory does not want to admit flux-lines running through it. It so happens that there is a nice analogy between this problem and the Meissner effect in superconductors: when a superconducting material is placed in a magnetic field, the resulting constant current is such that the induced magnetic field exactly cancels the applied magnetic field. There is therefore no net magnetic field inside the conductor. If the magnetic field is increased, at some point the magnetic field will be forced into the superconductor. It turns out to be energetically favorable for the magnetic field to enter the superconductor in flux tubes: only such flux configurations are allowed that a electron transported around a flux tube will pick up an extra unobservable phase factor of 2π. Theoretically, superconductivity is the result of electric degrees of freedom condensing which is well described by the Abelian Higgs model. The endpoints of a finite flux tube in a superconductor would be describing a pair of Dirac magnetic monopoles, but since the flux is confined to a tube this would involve a linear potential between these monopoles. 't Hooft took this condensed matter analogy as a model for the confinement of electric (color) flux: if magnetic degrees of freedom (magnetic poles) would condense this would cause the electric degrees of freedom (the gluonic field) to confine. This conjectured picture of color confinement is referred to as the dual Meissner effect. Non-Abelian gauge theory is expected to have confinement of color flux as a dual Meissner effect. As noted before this picture is reinforced through lattice studies and model calculations in, for instance, supersymmetric field theory.

1.2 Integrable systems

The formulation of an effective action for mesons in QCD is a nice illustration of yet another general principle: this effective action predicts results by calculating its classical minima. In short, given an effective action we are doing *classical* field theory to get quantum answers. The meson effective action also describes the pions as the excitations of certain coordinates on the moduli space of the vacuum. If we would consider these coordinates to be the coordinate of a classical particle, we would obtain an effective action for classical (relativistic) particles coupled through some non-linear (higher order) terms.

Classical non-relativistic multi-particle systems are usually very hard to solve. An illustrative example is the three body problem in Newtonian gravity. The solution of the two-body problem was a triumph for classical mechanics and goes back to Newton. In contrast, the three body problem has proven to be unsolvable in the general case. It can be approached through numerical analysis for instance and displays some surprising possible periodic solutions like three bodies chasing each other around a figure eight. One of the main differences between the two and three body problem is the existence of a symmetry in the two body problem, which allows it to be reduced to an effective one body problem. By Noethers theorem symmetry translates in the existence of conserved charges, functions of the phase space variables which Poisson commute with the Hamiltonian. Two special classes of multi-body problems which are solvable by the existence of enough symmetry can now be formulated. Suppose the phase space of a certain classical system is $2n$ dimensional.

- the classical system is *integrable* if it admits n independent conserved charges
- the classical system is *completely integrable* if it admits n independent conserved charges which Poisson commute amongst themselves

Note that the need for the extra condition in complete integrability has a nice quantum mechanical analogue in the theory of angular momentum: although the entire angular momentum vector is conserved, the maximal commuting subset of this vector is two dimensional: one can for instance only measure the overall size and one component of angular momentum simultaneously. In this thesis integrable will always mean completely integrable.

In principle completely integrable systems can be solved through quadrature (algebraic manipulations and integrations), although in practice this may be very difficult. Note that, since there is a commuting set of conserved charges, time evolution can be generated by *any* of these charges which are therefore also known as Hamiltonians. If the system is periodic, a theorem of Liouville states that for every (real) completely integrable system there is a coordinate system in which the time evolution generated by any Hamiltonian is *linear*. Since the time evolution is also periodic according to the theorem these coordinates are referred to as action-angle variables. Solving an integrable system therefore amounts to finding the action-angle variables.

There is a neat trick due to Lax [5] to write an integrable system such that its conserved charges at least are obvious. The trick is to find finite dimensional matrices $M(q_i, p_i)$ and $K_j(q_i, p_i)$ such that the time evolution equations can be written as a matrix commutator

$$\frac{\partial o_i}{\partial t_j} = \{o_i, H_j\} \quad \longleftrightarrow \quad \frac{\partial M}{\partial t_j} = [M, K_j]\,, \tag{1.8}$$

since in that case traces of powers of M are automatically conserved. In the above equation o_i stands for all phase space variables. Traces of powers of M are completely symmetric functions of the eigenvalues of the finite matrix M, so these provide a basis for the Hamiltonians. However, the existence of a Lax representation is in general not guaranteed and there may be several Lax representations (with different matrix sizes for instance!) for the same system.

A particularly nice example of a completely integrable system which will play a major role in this thesis is the periodic Toda chain. This system can be viewed as a system of N particles on a circle interacting through an exponential nearest neighbor interaction. Its Hamiltonian u_2 is given as

$$u_2 = \frac{1}{2}\sum_{i=1}^{N} p_i^2 + \Lambda^2 e^{q_1-q_2} + \ldots + \Lambda^2 e^{q_{N-1}-q_N} + \Lambda^2 e^{q_N-q_1}. \tag{1.9}$$

This system is completely integrable [6]. The Lax matrix of the periodic Toda chain[1] is given as

$$M = \begin{pmatrix} p_1 & \Lambda^2 e^{q_1-q_2} & 0 & \ldots & z \\ 1 & p_2 & \Lambda^2 e^{q_2-q_3} & \ldots & 0 \\ 0 & 1 & p_3 & \ldots & 0 \\ . & . & . & \ldots & \Lambda^2 e^{q_{N-1}-q_N} \\ e^{q_N-q_1} z^{-1} & . & . & 1 & p_N \end{pmatrix},$$

z is a so-called spectral parameter which is complex. As will be further elaborated upon in chapter 3 the chain admits N independent conserved quantities $u_k = \frac{1}{k}\operatorname{Tr} M^k$ which generate time Hamiltonian flows by a matrix commutator just as in equation (1.8),

$$\frac{\partial M}{\partial t_k} = [M, \mathcal{L}(M^{k-1})], \tag{1.10}$$

where $\mathcal{L}$ is a certain map. Viewed in this way the Hamiltonian (1.9), which is equal to $\frac{1}{2}\operatorname{Tr} M^2$, is just one of N independent Hamiltonians which all generate independent time flows. The properties of the Lax matrix can be studied by calculating its eigenvalues. The characteristic equation of the Lax matrix is also called the spectral curve for a Lax matrix associated to a periodic system. The spectral curve associated to the Lax matrix M of the periodic Toda chain is

$$\det(x - M) \equiv P_N(x) + (-1)^N (z + \Lambda^{2N} z^{-1}) = 0, \tag{1.11}$$

which describes, when viewed as a manifold in $\mathbb{C}^2 \cup \{\infty\}$, a compact Riemann surface and will be identified with the so-called Seiberg-Witten curve of a four-dimensional $\mathcal{N} = 2$ gauge theory. Note that the coefficients in the polynomial P are completely symmetric functions of the eigenvalues. The trace of powers the Lax matrix are also symmetric functions of the eigenvalues, but in a different basis. This shows that the spectral curve of the Lax matrix is itself a conserved quantity.

[1] Listed here is the Lax matrix for the periodic Toda chain associated to (the fundamental representation of) the affine Lie algebra A_N^1. There is a natural generalization to other affine Lie algebras and a natural reduction to the semi-simple Lie algebras, see for instance [6, 7]

Besides systems in classical mechanics with a finite dimensional phase space there are systems with an infinite dimensional phase space. This corresponds to replacing the integer index i in equation (1.8) by a continuous one

$$o_i \to o(x). \tag{1.12}$$

As is shown in [8], if one takes for instance the following Poisson bracket for a phase space of functions $o(x)$ (with summation replaced by integration)

$$\{o(x), o(y)\} = \frac{\partial}{\partial x}\delta(x-y). \tag{1.13}$$

Then studying a Hamiltonian of the form

$$H[o] = \int_{-\infty}^{\infty} dx \frac{o(x)^3}{3!} - \frac{1}{2}\left(\frac{\partial o(x)}{\partial x}\right)^2 \tag{1.14}$$

will generate the ubiquitous KdV equation when combined with the usual rules of functional derivation. It can be shown [8] that this system admits an infinite number of Poisson commuting Hamiltonians and can therefore at least naively be classified as completely integrable in the above described sense. Infinite dimensional phase space integrable systems are usually referred to as integrable hierarchies and can be formulated in a bewildering variety of ways (see for instance the introduction of [9] for a quick overview of some approaches).

1.3 Integrability in low energy effective supersymmetric gauge theory

As noted before, the low energy effective action for any quantum field theory will be some classical field theory which is in general not integrable (solvable). It is therefore a remarkable fact that the solutions for low energy effective theories as constructed by Seiberg and Witten are directly related to known periodic integrable multi-particle systems. In this relation, the coordinates which parametrize the moduli space of vacua in the Seiberg-Witten analysis are identified with the conserved quantities in the integrable system. This connection relates these integrable systems to gauge theories with extended supersymmetry. A natural question is then how this relationship develops if that extended supersymmetry is broken down to simple supersymmetry since in that case, as was touched upon in this introduction, there is interesting physics on the gauge theory side such as confinement and chiral symmetry breaking.

Before this question can be answered, we will find that to make the link between gauge theory and integrable system as strong as possible we will have to compactify one spatial direction in the theory on a circle. The moduli space of vacua of $\mathcal{N} = 2$ supersymmetric field theories in four dimensions get modified once the theory is compactified on a circle. Surprisingly, it has been proven that there are coordinates on the new moduli space which are actually independent of the radius of that circle for theories.

In this thesis the breaking of extended $\mathcal{N} = 2$ supersymmetry to $\mathcal{N} = 1$ in a four dimensional gauge theory compactified on $\mathbb{R}^3 \times S^1$ is studied from the perspective of the underlying integrable system. The breaking of supersymmetry will be induced by turning on a classical polynomial superpotential. The goal is to construct the exact quantum potential for these particular gauge theories and it is conjectured that this exact potential can be constructed naturally from the integrable system.

When expressed in the coordinates on the moduli space which are independent of the radius of the circle, the sought exact quantum potential also does not depend on the circle radius. This implies that the minima of this conjectured exact potential can be interpreted as both three and four dimensional vacua. For gauge theories with $U(N)$ gauge symmetry and various amounts of matter in the fundamental representation the conjectured potentials are analyzed in this thesis and it is shown that they consistently reproduce known results in four dimensional physics and also yield some new technical results. In the final chapter we will comment on how these results might be embedded in a deeper mathematical structure. The results presented in this thesis show that the link between integrable systems and supersymmetric gauge theory is even stronger than was known before and might provide the right structure to understand supersymmetric low energy effective gauge theory.

Chapter 2

Field Theory Background

This chapter describes the main field theory background for the research presented in this thesis. It starts off by introducing supersymmetric gauge theory in four dimensions, followed by a brief discussion on the construction of low energy effective actions (Seiberg-Witten theory) and low energy effective superpotentials (Dijkgraaf-Vafa theory). This chapter begins with material which can be found in the literature (see e.g. the reviews [10], [11]) and leads to the formulation of the general conjecture first posed in this form in [12]. We will follow the standard supersymmetry conventions of [13], to which the reader is also referred for general background material on supersymmetry.

2.1 Supersymmetric gauge theories in four dimensions

The starting point of the analysis is the definition of the field theories we will be dealing with: a supersymmetric version of the QCD Lagrangian (1.1) is needed. Supersymmetry is an essentially unique extension of the usual symmetries of Minkowski space and since we will not be considering gravity in this thesis all space time symmetries are taken to be global. The most striking non-trivial part of the symmetry algebra is given by anti-commutation relations for the supersymmetry generators, which are Weyl-spinors

$$\{Q^I_\alpha, \bar{Q}_{\dot{\alpha}J}\} = \sigma^\mu_{\alpha\dot{\alpha}} P_\mu \delta^I_J \tag{2.1}$$

$$\{Q^I_\alpha, Q^J_\beta\} = 2\epsilon_{\alpha\beta} Z^{IJ}. \tag{2.2}$$

Here P is the momentum vector, the antisymmetric matrix Z is called the central charge matrix (which commutes with everything), $\sigma^\mu = \{1, \sigma^i\}$ with σ^i the Pauli matrices, and finally

$$\epsilon = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

We have allowed for extended supersymmetry indexed by the integers I, J. The minimal amount of supersymmetry is obtained by considering just one Weyl spinor, which contains four real components known as supercharges. For extended supersymmetry, the total number of supercharges in a given theory must be an integer multiple of four. This integer number is denoted by $\mathcal{N}$. More generally, the total number of supercharges in a supersymmetric theory in d dimensions is $\mathcal{N}$ times the minimal dimension of a spinor in d dimensions. For three dimensions for instance, the minimal dimension of a spinor is 2. The supersymmetry algebra has an $SU(\mathcal{N})$ internal symmetry which acts on the I, J indices and is called R-symmetry. In addition it has a chiral $U(1)_R$ global phase rotation symmetry if the central charges are zero.

Because the supersymmetry generators are fermionic they transform a bosonic state in a fermionic one. A careful analysis of what states are allowed in a supersymmetric theory gives that there is a class of states which transform non-trivially only under *half* the supersymmetry algebra. These states are called chiral or anti-chiral, depending on which half (in the chiral sense, hence the name) of the supersymmetry algebra these states transform non-trivially. This phenomenon is in essence a direct extension of the usual difference between massless and massive representations of the Poincare algebra. The main difference is that in supersymmetric theories there is an extra parameter, the central charge Z^{ij}. When that central charge matrix is 'diagonalized' by the internal R symmetry rotations to

$$Z = \text{Diag}\{Z_1\epsilon, Z_2\epsilon, \ldots\}, \tag{2.3}$$

the chiral states are those that saturate the so-called BPS-bound,

$$\exists i \,|\, M = Z_i. \tag{2.4}$$

States which satisfy this bound form special, smaller than usual representations of the supersymmetry group, just as massless fields do in a usual Lorentz invariant theory. In supersymmetric field theory the central charges are usually topological charges, so at first we will set them to zero.

An important building block for a supersymmetric field theory is of course fields which carry a representation of the symmetry algebra (2.2). First we will construct theories with $\mathcal{N} = 1$ and later $\mathcal{N} = 2$ supersymmetry. In standard field theory, fields are functions defined on space and since supersymmetry is an extension of space-time symmetry it is useful to look for a "supersymmetric space"or superspace first. Since the supersymmetry generators are themselves fermionic it is natural to expect that the extra coordinates we need are fermionic as well. In order to make this more precise the connection between regular coordinates and the underlying symmetry algebra is needed. Regular space-time can naturally be described by the a group coset construction: a point in space can be represented as an element of the Poincare group identified under the action of the Lorentz group. The Lorentz group independent part of a general element of the supersymmetric space-time symmetry group is given by

$$P = e^{i\left(-x\cdot P+\theta Q+\bar{\theta}\bar{Q}\right)}. \tag{2.5}$$

In this expression one recognizes the θ and $\bar{\theta}$ as extra *fermionic* coordinates and we have adopted the Wess and Bagger [13] conventions for the definition of products of Weyl fermions,

$$\theta Q = \theta^{\alpha} Q_{\alpha} \qquad\qquad \bar{\theta}\bar{Q} = \bar{\theta}_{\dot{\alpha}}\bar{Q}^{\dot{\alpha}}. \tag{2.6}$$

In this expression indices are raised and lowered using the matrix ϵ (2.3). If we act on (2.5) with a general would-be supersymmetry transformation from the right

$$e^{i\left(\xi Q+\bar{\xi}\bar{Q}\right)}e^{i\left(-x\cdot P+\theta Q+\bar{\theta}\bar{Q}\right)}, \tag{2.7}$$

an explicit expression for the supersymmetry generator is obtained by using the Baker-Campbell-Hausdorff formula and the supersymmetry algebra. This is completely analogous to how one usually derives the representation of the momentum operator on fields as $P = -i\frac{\partial}{\partial x^{\mu}}$. Their explicit expression will not be needed here and can be found in the literature. Now one introduces a general superfield on this extended space as $\kappa(x,\theta,\bar{\theta})$, which has a natural expansion in terms of complex space-time fields, which will be restricted to fields with spin ≤ 1,

$$\kappa(x,\theta,\bar{\theta}) \quad = \quad \phi(x) + \theta\xi(x) + \bar{\theta}\bar{\xi(x)} + \theta\theta m(x) + \bar{\theta}\bar{\theta}n(x) + \theta\sigma^{\mu}\bar{\theta}v_{\mu} \tag{2.8}$$

$$+\theta\theta\bar{\theta}\bar{\lambda}(x) + \bar{\theta}\bar{\theta}\theta\psi(x) + \theta\theta\bar{\theta}\bar{\theta}d(x). \tag{2.9}$$

In this expression σ^{μ} equals $\{\mathbb{I}_{2\times 2}, \sigma^{i}\}$ with σ^{i} the Dirac matrices. The supersymmetry transformation of this field can be worked out applying the constructed supersymmetry generators to it and comparing to the expression for a general superfield. The component of this field which has the most θ's transforms as a total derivative under supersymmetry transformations. A supersymmetric action can therefore easily be constructed by integrating this component over all of space. This component is called the top component. Because integration over fermionic coordinates is defined such that $\int d\theta^2\theta\theta = 1$ and $\int d\bar{\theta}^2\bar{\theta}\bar{\theta} = 1$, the top component can be isolated by integrating the superfield as

$$\kappa(x,\theta,\bar{\theta})\lfloor_{\text{D-term}} = \int d\theta^2 d\bar{\theta}^2 \kappa(x,\theta,\bar{\theta}). \tag{2.10}$$

Since the product and sum of general superfields is, again, a general superfield a large class of supersymmetric Lagrangians can be written down as the integral over the top component of a function of the general superfield. It is a useful fact that in such a construction the top component of the superfield itself, called the D-term, will never appear with a derivative acting on it. Therefore the D-term is not a dynamical, but an auxiliary field and may be integrated out by imposing it's field equation as a constraint on the other component fields in the superfield.

In the general expression displayed above the superfield κ is complex and has in total 16 (complex) degrees of freedom which is overly unrestrictive for the construction of gauge field theories. More useful building blocks are obtained by imposing various constraints on this field. For this purpose introduce the superspace derivatives $D, \bar{D}$

$$D_\alpha = \frac{\partial}{\partial \theta^\alpha} + i\sigma^\mu_{\alpha\dot{\alpha}}\bar{\theta}^{\dot{\alpha}}\partial_\mu \tag{2.11}$$

$$\bar{D}_{\dot{\alpha}} = -\frac{\partial}{\partial \bar{\theta}^{\dot{\alpha}}} - i\theta^\alpha\sigma^\mu_{\alpha\dot{\alpha}}\partial_\mu, \tag{2.12}$$

which obey schematically

$$\{\bar{D}, Q\} = 0, \qquad \{\bar{D}, \bar{Q}\} = 0. \tag{2.13}$$

Because of these anti-commutation relations, a condition like $\bar{D}_{\dot{\alpha}}\kappa_{\text{chiral}} = 0$ yields a consistent supersymmetric reduction of the general superfield. When written in coordinates $y^\mu = x^\mu + i\theta\sigma^\mu\bar{\theta}$ the superfield κ_{chiral} has a superfield expansion,

$$\kappa_{\text{chiral}}(y, \theta) = \kappa_{\text{chiral}}(y) + \theta\psi(y) + \theta^2 F(y), \tag{2.14}$$

where we have introduced the notational convention of denoting chiral superfields and their scalar components by the same variable. Which of the two interpretations is meant is clear from the context. The expansion shows that the field content of the chiral superfield consists of a complex scalar field κ_{chiral}, a Weyl fermion ψ and another complex scalar field F. Note that κ_{chiral} does not depend on half of the fermionic superspace coordinates. Given chiral fields it is easy to construct new chiral fields by multiplication and addition and they therefore form a ring. In particular any holomorphic function of a chiral superfield is again a chiral superfield. Working out the supersymmetry transformations shows that the top component (known as F-term) of chiral or anti-chiral fields transform as a total derivative under supersymmetry transformations. Supersymmetric actions can therefore easily be constructed by integrating the F-term of a holomorphic function of a chiral field over space. The F-term of any superfield can be isolated by fermionic integration as

$$\kappa_{\text{chiral}}(x, \theta)\lfloor_{\text{F-term}} = \int d\theta^2 \kappa_{\text{chiral}}(x, \theta). \tag{2.15}$$

Analogous to the case of the general superfield, it turns out that a general Lagrangian constructed as an analytic function of a chiral superfield will not contain kinetic terms for the top component of this superfield. The top component may therefore be integrated out. Completely analogous to chiral fields one can also define anti-chiral fields by imposing $D_\alpha \bar{\kappa}_{\text{anti-chiral}} = 0$

Another constraint one can impose on the general superfield (2.8) is a reality constraint $\kappa^\dagger = \kappa$. This field contains a (real) vector, leading to the name vector or gauge multiplet, denoted below by V. A particularly simple example of a real superfield is the sum of a chiral and an anti-chiral superfield, which in particular contains the derivative of a scalar function as a vector field. This is the motivation to define gauge transformations of a vectorfield V as

$$V \to V + \kappa_{\text{chiral}} + \bar{\kappa}_{\text{anti-chiral}}. \tag{2.16}$$

A particular gauge (known as Wess-Zumino gauge) can be chosen in which the lowest components of the vector superfield are set to zero, which is computationally useful since then only a finite number of powers of V are non-zero. This gauge will be used implicitly in the following. One can now use the gauge transformations in superspace and study the formulation of a supersymmetric gauge theory. For this purpose the following multiplets are constructed

$$W_\alpha = -\frac{1}{4}\bar{D}\bar{D}e^{-V}D_\alpha e^V \tag{2.17}$$

$$\bar{W}_{\dot{\alpha}} = -\frac{1}{4}DDe^{-V}\bar{D}_{\dot{\alpha}}e^V. \tag{2.18}$$

If we now define

$$\mathcal{L} = \frac{1}{2}Re\left(\tau \int d\theta^2 \text{Tr}\,(W^\alpha W_\alpha)\right), \tag{2.19}$$

with the complexified gauge coupling constant $\tau = \frac{2}{g^2} + i\frac{\theta}{2\pi^2}$ and with the trace over the gauge indices, this can be shown to yield

$$\mathcal{L} = -\frac{1}{4g^2}F_{\mu\nu}F^{\mu\nu} - \frac{\theta}{16\pi^2}F_{\mu\nu}\tilde{F}^{\mu\nu} + \text{fermions}. \tag{2.20}$$

This is the (bosonic part of the) $\mathcal{N} = 1$ supersymmetric gauge theory Lagrangian. Matter can also be coupled in a manifestly supersymmetric way to the gauge field as described in the literature. This gives extra terms in the Lagrangian of the form

$$\mathcal{L}_{matter} \sim \int d\theta^2 d\bar{\theta}^2 \tilde{Q}e^{-2V}Q. \tag{2.21}$$

2.1.1 Quantum effects

Up to this point we have basically described the structure of classical supersymmetric gauge theories. However, supersymmetry restricts the form of any quantum effective action. The simple observation is that in chiral integrals of the type appearing in, for instance, equation (2.19) the integrand should be an analytic (holomorphic) function: it should only depend on the chiral fields, not the anti-chiral fields. This fact, together with symmetry and decoupling arguments, constrains the theory so much it becomes exactly solvable in some cases. Arguments of this type go by the general name of 'holomorphy' or 'Seibergology' in honor of its inventor [14].

An example of how these arguments work is the β function in $\mathcal{N}=1$ pure $SU(N)$ super-Yang-Mills[1]: the gauge coupling appears in equation (2.19) is a parameter in a chiral part of the Lagrangian. In particular, it can be viewed as the vacuum expectation value of an infinitely heavy chiral field, since all constants r are chiral (and anti-chiral) fields because $D_\alpha r = 0$ ($\bar{D}_{\dot\alpha} r = 0$). If the gauge coupling is defined in a gauge theory with a UV cutoff at M, then it can be evaluated in a theory with a different UV cutoff M' keeping low energy physics fixed, by using the gauge theory beta function. The couplings must be related through

$$\tau'(M') = \tau(M) + f\left(\tau, \log\left(\frac{M}{M'}\right)\right), \tag{2.22}$$

where f is an analytic function of τ. Since shifting τ by $2\pi i$ should not change physics (since this corresponds to shifts in the θ angle), f should be a $periodic$ function of τ. Now we can expand the beta function, as a derivative with respect to $t = \log\left(\frac{M}{M'}\right)$, as a Fourier series

$$\frac{d}{dt}\tau = \beta(\tau) = \sum_{n\in\mathbb{Z}} c_n e^{\tau n}, \tag{2.23}$$

where the last equality simply follows from periodicity and analyticity. Note that this particular expansion does not have a weak-coupling limit ($\tau \to \infty$) if n is larger than zero, so we restrict the sum over n to $n \leq 0$. It follows that the form of the contributions to this beta function is a constant c_0 and a sum over non-perturbative (instanton) effects, since $\tau \sim \frac{1}{g^2}$. The constant c_0 is determined at one loop. The $\mathcal{N}=1$ theory has furthermore an anomalous $U(1)_R$-symmetry. Since the gauge coupling is considered to be a background field and can therefore be assigned transformation properties, this symmetry can be restored by assigning an appropriate transformation law to τ which shifts it by an arbitrary imaginary constant. Since this should be a symmetry of the theory, the beta function should be invariant under

[1]the argument as presented here can be found in [15].

a shift of its argument by an imaginary number. Since the right-hand side of (2.23) should be an analytic function, this is impossible unless c_n is zero for all $n > 0$. Therefore the only non-zero coefficient is c_0. This coefficient can be calculated by a one-loop computation. Especially the latter argument about the absence of non-perturbative corrections is prone to false conclusions however. If a beta function is defined as the (logarithmic) change in coupling constant if the scale at which this coupling constant is defined is varied, instead of the scale of the UV-cutoff as above, this latter scale is not the scale which brakes $U(1)_R$, so this symmetry can not be restored by assigning transformation properties.

There are a number of comments to make about an analysis along the above lines. First of all, in computations of the above sort the implicit assumption is made there is a regularization of the theory which preserves supersymmetry and any symmetries used in deriving results. Furthermore, arguments of the above type require a specific choice of regularization scheme: if one redefines the real vector fields in order to obtain canonical kinetic terms for the gauge fields then an anomaly in the measure contributes to the β function of the redefined gauge theory couplings[15]. This shows that in practice, even disregarding minor details like the exact mechanism of complete supersymmetry breaking, it might be difficult to relate results based on holomorphy arguments to, for instance, the QCD scale quantitatively. However, this is the same type of problem as relating the effective scale introduced in, for instance, dimensional regularization to measured quantities like the coupling constant at a certain scale. Qualitatively, the above argument illustrates the fact that the structure of the full quantum theory is much more under control using supersymmetry.

2.1.2 Supersymmetric low energy effective actions

In this thesis the focus is on the low energy effective dynamics of a supersymmetric gauge theory. For that purpose we would like to calculate a low energy effective action of the same type displayed in the introduction. In principle this effective action should be calculable from an exact description of the full quantum theory. Lacking that description, the main tool is the use of holomorphy arguments to constrain the low energy effective quantum theory. First of all we assume that supersymmetry is unbroken. Furthermore, terms in the Lagrangian with more derivatives than two are assumed to be suppressed by an effective scale, which is reasonable by dimensional analysis. It turns out [13] that for chiral fields the most general $\mathcal{N} = 1$ supersymmetric gauge invariant Lagrangian with at most two derivatives is described by a non-linear sigma model: the chiral fields κ_i can be interpreted as coordinates on a manifold. This Lagrangian takes the form

$$\mathcal{L} \sim \int d\theta^2 d\bar{\theta}^2 K(\kappa_i, \kappa_j^\dagger) + Re\left(\int d\theta^2 w(\kappa_i) + \tau(\kappa_i)\mathrm{Tr}\, WW\right), \tag{2.24}$$

where w and τ in general are holomorphic functions of the indicated arguments and K is a so-called Kähler potential. w is called the superpotential. Only the holomorphic data w, τ is expected to have special properties in the quantum theory by holomorphy arguments, the Kähler potential can not be determined by these arguments for a general $\mathcal{N} = 1$ theory since it is not holomorphic. Below a method is introduced to be able to apply holomorphy arguments to the Kähler potential as well.

Suppose we start with a $U(N)$ gauge theory coupled to a scalar field Φ which transforms in the adjoint. For a generic large vev of Φ the gauge theory is completely realized in the Higgs-phase and the gauge group is broken as $U(N) \to U(1)^N$. This is the regime in which perturbation theory is reliable since the effective coupling constant stops running at a very high scale, where the theory is weakly coupled. The question is then what the theory looks like for small values of $< \Phi >$. For this type of argument to work nicely, the scalar field and the gauge field must be part of the same supersymmetry multiplet. On-shell, this can only be done if the theory enjoys $\mathcal{N} = 2$ supersymmetry, so we want to construct the most general $\mathcal{N} = 2$ supersymmetric gauge theory Lagrangian with a $U(1)^N$ gauge symmetry.

The Lagrangian (2.24) is $\mathcal{N} = 1$ supersymmetric. Since $\mathcal{N} = 2$ supersymmetry is a special case of this class of theories, the $\mathcal{N} = 2$ low energy effective Lagrangian can be obtained by studying (2.24) for a $U(1)^N$ gauge group coupled to a scalar field Φ which transforms in the adjoint of the gauge group. For special values of the coupling constants the resulting Lagrangian is $\mathcal{N} = 2$ supersymmetric. It can be shown (see e.g. [11]) that the Lagrangian of the full theory thus obtained is specified by one holomorphic function of N variables $F(z_i)$, called the prepotential[2].

$$\mathcal{L} = Re\,\tau_{ij}(z^k)F^i_{\mu\nu}F^{j\mu\nu} + Im\,\tau_{ij}(z^k)F^i_{\mu\nu}\tilde{F}^{j\mu\nu} + Im\partial^\mu z_{Dj}\partial_\mu z^j + \text{fermions}, \tag{2.25}$$

with

$$\tau_{ij} = \frac{\partial^2}{\partial z^i \partial z^j}F(z) \tag{2.26}$$

$$z_{Dj} = \frac{\partial}{\partial z^j}F(z). \tag{2.27}$$

As part of the analysis, the superpotential w is required to vanish (for theories with no matter in the fundamental) in order to obtain $\mathcal{N} = 2$ supersymmetry. In other

[2]This is more clear in the unconstrained $\mathcal{N} = 2$ superspace formulation [11], which also directly shows the extended supersymmetry.

words, adding it will break the extended supersymmetry to $\mathcal{N} = 1$. A second thing to note is that the coupling constants of the effective $U(1)^N$ gauge theory are given by the second derivative of the prepotential. A final remark concerns the Kähler potential: here it is related by $\mathcal{N} = 2$ supersymmetry to the holomorphic sector of the theory and therefore holomorphy arguments now also determine this part of the theory.

2.2 Seiberg-Witten theory

In complex analysis holomorphic functions are highly constrained by their behavior on the boundary of a connected coordinate patch, so one might hope that in the supersymmetric field theory holomorphy might work out the same: as indicated before, in the limit of large vevs for Φ, some quantities can be calculated which is knowledge of a boundary region of the moduli space. This hope is justified, as was shown by Seiberg and Witten [3, 4]. Although below we will continue to consider for definiteness a pure $\mathcal{N} = 2$ $U(N)$ gauge theory, the reasoning employed is much more widely applicable. Useful general reviews on Seiberg-Witten theory include, but are not limited to [11, 2].

Classically, the microscopic theory is governed by a vacuum where the scalar potential (which stems from integrating out auxiliary (D-term) fields) vanishes,

$$[\Phi, \Phi^\dagger] = 0, \tag{2.28}$$

which implies that the complex scalar can be diagonalized by gauge rotations. This gives

$$< \Phi >_{cl} = \sum a_i H^i \tag{2.29}$$

with H^i the generators of the Cartan subalgebra, extended with $H^N = \mathbb{I}_{N\times N}$ for the overall $U(1)$ gauge field. The N complex parameters a_i parametrize a manifold of vacua. However, they are not really good coordinates since they are not invariant under Weyl reflections (which interchange the different Cartan generators H^i amongst each other). Better coordinates are symmetric polynomials of the a_i, for which there are two natural choices defined through the following generating functionals:

$$\det(x - \phi^{cl}) = \sum_{i=0}^{N} s_i(a^{cl}) x^{N-i} \tag{2.30}$$

$$-\mathrm{Tr}\,\log(1 - \phi^{cl} x) = \sum_{i=1}^{N} u_i(a^{cl}) x^i + h.o. \tag{2.31}$$

Here we have neglected the higher order terms for the u_i polynomials since they do not generate new independent polynomials. These different sets of polynomials are related through the Newton-Girard formulas, which can be derived by considering the identity

$$e^{\mathrm{Tr}\,A} = \det e^A, \tag{2.32}$$

for matrices $A = \log(x - \Phi)$, which yields

$$\mathrm{Tr}\,\log(x - \Phi) = \log(\det(x - \Phi)). \tag{2.33}$$

These expressions can both be expanded as formal power series in $1/x$, from which formulas relating u_i and s_i can be read of. The derivative with respect to x of the left hand side of this equation, $\mathrm{Tr}\left(\frac{1}{x-\Phi}\right)$, is called the resolvent of Φ because it is a generating functional for traces of powers of Φ. Either the set u_i or s_i provide a set of Weyl-invariant coordinates for the classical moduli of the gauge theory, although in practice one works with the coordinates a_i since they are most directly linked to the physics of the theory; they provide for instance the most direct picture of the gauge symmetry breaking.

For generic values of the a_i the gauge symmetry is classically completely broken $U(N) \to U(1)^N$. If the symmetry breaking scale is large enough[3], the quantum low energy effective action is calculable in this case in perturbation theory (just as it is for the weak sector in the standard model), where it can be shown to only receive 1 loop contributions. In addition to the perturbative contributions, the prepotential will receive contributions from instantons. Since the generic form of these contributions is known and the prepotential itself must be a holomorphic (possibly multi-valued) function of the a_i and the effective $\mathcal{N} = 2$ scale Λ, the prepotential as a function of the low energy effective field a can be shown [16] to take the form

$$F(a_i) = \tau \sum_{i=1}^{N} (a_i)^2 - \frac{1}{8\pi i} \sum_{i,j=1}^{N} (a_i - a_j)^2 \log\left(\frac{(a_i - a_j)^2}{\Lambda^2}\right) + \sum_{i=0}^{\infty} F_k(a_i)\Lambda^{2Nk}. \tag{2.34}$$

Note that the coordinate $\sum_i a_i$ is completely decoupled from the slightly complicated renormalization effects. This is because this is the coordinate of the overall $U(1)$ gauge theory. Since instanton contributions are accompanied by powers of Λ^{2N}, the expansion in the last term is referred to as the instanton expansion. There are two known ways of obtaining the precise form of the above prepotential (the unknown coefficients in the instanton expansion). One is by simply calculating the instanton contributions directly. This is a hard calculation in general and was only done recently by Nekrasov by employing high-powered mathematics [17]. The original answer of Seiberg and Witten was derived by clever consistency arguments,

[3]More precisely if $|a_i - a_j| \gg \Lambda \quad \forall i \neq j$.

symmetry considerations and holomorphy. Their answer is that the prepotential can be obtained by calculating period integrals on an auxiliary Riemann curve which is fibered over the vacuum moduli space. This takes the following form for $U(N)$

$$\begin{aligned} a_i(\{s_j\}) &= \oint_{\alpha_i} \lambda_{sw} \\ \frac{\partial F(\{s_j\})}{\partial a_j} &= \oint_{\beta_j} \lambda_{sw} \\ \lambda_{sw}(\{s_j\}) &= x d \log\left(P(x) + \sqrt{P(x)^2 - 4\Lambda^{2N}}\right). \end{aligned} \tag{2.35}$$

Where λ_{sw} is a meromorphic differential on the surface which is such that the derivative with respect to any of the s_j is a *holomorphic* differential and α_i, β_j are canonically intersecting cycles on the surface. The polynomial P is given in terms of the s_j coordinates on the moduli space by

$$P(x) = \det(x - \Phi^{cl}). \tag{2.36}$$

Note that the polynomial P is of order N. The Riemann surface is given as a vector $(x, z) \in \mathbb{C}^2$ constrained as

$$y^2 = P(x)^2 - 4\Lambda^{2N}. \tag{2.37}$$

The solutions to this equation can be represented pictorially in the following way: y is either plus or minus the square root of the polynomial on the right-hand side. These can be represented by two sheets of paper. The polynomial has in general $2N$ single, paired roots. The branch cut of the logarithm will be chosen such that the roots will be connected by a branch cut: if x moves across the branch cut, y jumps $y(x) \rightarrow -y(x)$. Note that the length of the branch cut is set by the scale Λ. Hence we can imagine the two sheets of paper being connected by pieces of plumbing, which connect the corresponding cuts on both sheets. Adding points at $\pm\infty$ now gives a complete, compact, complex Riemann surface. Something special happens if two roots of the polynomial $P(x)^2 - 4\Lambda^{2N}$ collide, since then there is one less piece of plumbing needed to built the full space of solutions. This means that the effective genus of the surface decreases.

The equations (2.35) for the prepotential can be used to express the first derivative of the prepotential with respect to the various a_i in the a_j itself. This can then be integrated to give the full prepotential. The coupling constant matrix ($\tau_{ij} = \frac{\partial^2 F}{\partial a_i \partial a_j}$) of the effective $U(1)^N$ gauge theory obeys

$$\frac{\partial^2 F(\{s_j\})}{\partial a_i \partial a_j} = \sum_k \left(\frac{\partial a_j}{\partial s_k}\right)^{-1} \frac{\partial}{\partial s_k} \frac{\partial F(\{s_j\})}{\partial a_j} = \int_{\beta_j} \omega_i. \tag{2.38}$$

The last quantity is then identified as the period matrix of the Riemann surface since it is the β period integral over the normalized one-forms. The coupling constant

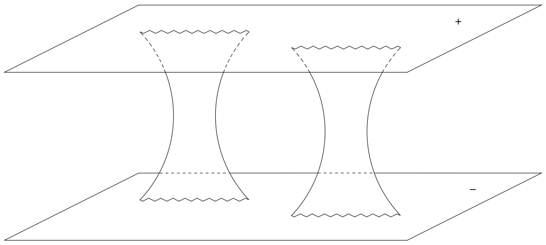

Figure 2.1: *Pieces of "plumbing" which connect the cuts on two sheets of solutions to equation (2.37).*

matrix also naturally appears in a geometrical construct associated to any Riemann surface. These surfaces have a so-called Jacobian J which is a complex torus $N-1$ dimensional torus with period matrix $\tau_{ij} = \int_{\beta_i} \omega_j$, where ω_j is a normalized [4] basis of holomorphic differentials on the surface. More explicitly for coordinates $\vec{\psi}$ on the torus,

$$\psi_j \sim \psi_j + n_j + \tau_{ji} m_i \qquad \forall m_i, n_i \in \mathbb{Z}. \tag{2.39}$$

In addition to the $N-1$ dimensional torus there is a separate complex circle related to the overall $U(1)$ gauge group. This extra circle is understood to be included if we refer to the Jacobian of a $U(N)$ gauge theory below.

An important aspect of the Seiberg-Witten analysis is the realization that there are special points (submanifolds) on the moduli space of vacua for which the description in equations (2.35) becomes singular: one or more of the α cycles collapse. The interpretation of this is that there are extra degrees of freedom in the theory which have become massless. For each degree of freedom becoming massless the moduli space is restricted to a submanifold of co-dimension 1. In order to identify exactly what the degrees of freedom are which become massless at the degeneration locus, one can compute how the combination of $\begin{pmatrix} \vec{a} \\ \vec{a}_D \end{pmatrix}$ transforms when circling close around the locus where the effective theory breaks down. This gives an idea how a and a_D behave near the singularity. Comparing these calculations to the known exact mass formula for (regular) BPS 't Hooft-Polyakov monopoles dyons one can see that these are precisely the degrees of freedom becoming massless. In order to obtain the full low-energy effective description of the gauge theory at these points in the moduli space, the extra massless degrees of freedom have to be added to the Lagrangian at these points. Based on the gauge theory lore described in the introduction one would very much like these degrees of freedom to condense in some clear-cut way, providing a nice and clean picture of confinement. This can be shown

[4]the integrals over the α cycles is $\int_{\alpha_i} \omega_j = \delta_{ij}$

to happen by breaking the supersymmetry to $\mathcal{N} = 1$ by turning on a superpotential.

Before we study supersymmetry breaking there are some general remarks to make. In general, if one would like to analyze the low energy effective description of a certain $\mathcal{N} = 2$ supersymmetric gauge theory, all that has to be done is the identification of a Riemann surface and a certain differential on it. This identification can be tricky in general but has been done for large classes of gauge theories, even for higher dimensional ones by interpreting them as four dimensional field theories with an infinite number of fields. This provides a large "phenomenology" of $\mathcal{N} = 2$ theories to work with.

2.2.1 Breaking $\mathcal{N} = 2$ to $\mathcal{N} = 1$

As mentioned before, turning on a microscopic polynomial superpotential for the adjoint scalar (actually for the $\mathcal{N} = 1$ chiral multiplet of which this is the scalar component),

$$W = \sum_{k=1}^{n+1} \frac{g_k}{k} \operatorname{Tr} \Phi^k, \tag{2.40}$$

will classically break $\mathcal{N} = 2$ to $\mathcal{N} = 1$. Theories with a high order superpotential are not renormalizable in 4 dimensions, but this is not a problem if we are only considering effective theories in the sense that this superpotential can be interpreted as being derived from a theory which lives at even higher scales for which this is the effective theory at some intermediate scale. This is exactly the point of view taken in this thesis. This form of supersymmetry breaking is called soft supersymmetry breaking.

Classically, the equation

$$W'(\Phi) \equiv g_{n+1}(\Phi - w_1)(\Phi - w_2)\ldots(\Phi - w_n) = 0, \tag{2.41}$$

must hold and therefore each of the eigenvalues a_i of Φ in equation (2.29) needs to be equal to one of the stationary point of the superpotential, w_j. The moduli space is reduced to a finite set of points, where at least a $U(1)^N$ gauge symmetry remains unbroken. More precisely, if w_i is occupied N_i times, in other words N_i of the a_j are equal to w_i, then the gauge symmetry is broken classically to

$$U(N) \longrightarrow U(N_1) \times \ldots \times U(N_n). \tag{2.42}$$

The different $U(N_i)$ gauge theories are expected to confine in the infrared in a full quantum analysis.

Generically the superpotential will also lift the entire quantum moduli space, but care has to be taken on those subspaces where dyons become massless. On these points the relevant $\mathcal{N}=1$ chiral superfields are

$$u_k, a_i, m, \tilde{m}. \tag{2.43}$$

Here u_k are the chiral superfields for which $<u_k>=\frac{1}{k}\mathrm{Tr}\ <\Phi^k>$ and a_i is the chiral superfield part of the different $U(1)$ $\mathcal{N}=2$ gauge multiplets. These operators were already there in the $\mathcal{N}=2$ case. New are m and $\tilde{m}$ which are the chiral multiplets for the dyons. Combining symmetry arguments, non-renormalization theorems, the existence of decoupling limits and taking into account the new degrees of freedom in the superpotential one can show that [18]

$$W=\sqrt{2}\sum_{i=1}^{l-r}\tilde{m}_i a_i m_i+\sum_{k=1}^{n+1} g_k u_k. \tag{2.44}$$

is actually an exact superpotential[5]. The bosonic part of this superpotential must now be minimized to find the correct $\mathcal{N}=1$ vacua. Note that both a_i (as the bosonic part of a_i) and u_k (as the bosonic part of u_k) should be functions of the vacuum moduli space variables s_i in equation (2.30). For a_i this relation is given by (2.35), but for u_k care has to be taken in how to define the relation properly. This involves a choice in the gauge theory. In this thesis the natural definition by the gauge theory resolvent is used,

$$\frac{1}{k}\mathrm{Tr}\ <\Phi^k>=\frac{1}{k}\frac{1}{2\pi i}\oint_{\infty} dx x^{k-1}\lambda_{sw}, \tag{2.45}$$

where the loop runs counterclockwise. This has the correct classical limit (relationship between s_i and u_i as in 2.33) for operators with $k\leq 2N$, since for these operators the right-hand side of the above equation is exactly equal to (2.33). Note that the Seiberg-Witten form also contains one (canonical) power of x. This form appears naturally in the matrix model analysis of Dijkgraaf and Vafa, as explained in [19].

Using the expressions for a_i and the gauge theory resolvent (2.45) one can minimize the exact superpotential [18] in order to find the supersymmetric vacua of the theory. The upshot of this analysis is that the Seiberg-Witten curve factorizes in the following way: there exist polynomials H, T, G and f of order indicated by their subscripts for which

$$\begin{aligned} y^2=P_N^2(x)-4\Lambda^{2N} &= H_{N-k}^2(x)T_{2k}(x)\\ G_{n-k}^2(x)T_{2k}(x) &= W'(x)^2+f_{n-1}(x). \end{aligned} \tag{2.46}$$

[5]In writing this potential the assumption is made that no mutually non-local monopoles become massless,which would have prevented the duality transformations necessary to arrive at the above form. However, in this thesis we will not consider these monopoles: we will not allow for Argyres-Douglas points.

The precise calculation of this factorization from the exact superpotential (2.44) is performed for gauge theories with matter that transforms in the fundamental representation in the second appendix to chapter 4. In other words, $N-k$ one-cycles have collapsed, which is equal to the number of monopoles becoming massless. The exact location of the double roots in the polynomial H are related to the non-zero vevs of the magnetic monopoles appearing in (2.44), which shows confinement being realized in the 't Hooft-Mandelstam scenario. However, since we are confined to studying holomorphic quantities like the superpotential this analysis only applies to the structure of the vacuum. It would be extremely interesting to obtain the Kähler potential of this $\mathcal{N}=1$ theory which contains the kinetic terms. In order to derive the result about the vacuum structure all that was needed was input on the symmetries of the theory and the implicit assumption of a mass gap in the theory: we assumed that generically the only degrees of freedom were those of a $U(1)^N$ gauge theory and the only degrees of freedom becoming light were the BPS monopoles (which subsequently condense).

The field theory analysis of the partial supersymmetry breaking reduces by the above analysis to finding those values of the coefficients of the polynomial P for which equation (2.46) holds. This can be formulated as a field theory recipe: first the values of the polynomial p must be found for which $P^2-4\Lambda^{2N}$ has $N-k$ double roots. After this the first equation of (2.46) is satisfied and the moduli space is k dimensional. In order to satisfy the second equation, the superpotential must be expressed in the remaining moduli space coordinates and then minimized. In the literature these two steps are sometimes done in one step by constructing a superpotential which incorporates the double root constraint as Lagrange multipliers.

To identify what happens, for instance, to the gauge symmetry of the theory one needs to take classical limits: in the classical limit the field Φ is distributed over the minima of the superpotential as in equation (2.41), which gives a direct interpretation in the form of the breaking pattern of the gauge symmetry as in (2.42). The classical limit is defined as any limit where the $\mathcal{N}=2$ scales to zero which will become clear below. In principle one solution to the factorization problem in equation (2.46) may admit multiple classical (scaling) limits. The genus $\hat{g}$ of the effective Riemann surface indicates to how many subgroups $U(N)$ is broken classically as

$$\hat{g}+1=\#\,\text{subgroups}. \tag{2.47}$$

Since one solution to the factorization problem gives a fixed genus $\hat{g}$, the phenomenon of multiple classical limits may only connect breaking patterns which lead to the same number of subgroups, e.g there might be one classical limit with $U(4)\to U(2)\times U(2)$ and another limit with $U(4)\to U(3)\times U(1)$, so $\hat{g}$ is a good order parameter.

2.2.2 $\mathcal{N} = 1$ perspective

With the above remarks in mind, the symmetry breaking pattern as displayed in equation (2.46) is

$$U(N) \stackrel{\text{classical}}{\longrightarrow} U(N_1) \times \ldots \times U(N_n) \stackrel{\text{quantum}}{\longrightarrow} U(1)^k, \tag{2.48}$$

where k is the number of N_i that are not equal to zero. The above analysis is from the $\mathcal{N} = 2$ perspective. One can also take a more $\mathcal{N} = 1$ perspective, since in the limit of an infinite superpotential the low energy theory is a strictly $\mathcal{N} = 1$ $U(N_1) \times \ldots \times U(N_n)$ gauge theory. In order to connect to known results in $\mathcal{N} = 1$ supersymmetric theories it is useful to first study just one of these subgroups. Since locally every critical point of a superpotential can be well described by a quadratic potential, the analysis will be restricted first to $\mathcal{N} = 2$ theory broken by a quadratic superpotential $\mu \mathrm{Tr}\,\Phi^2$. In the limit of large μ we are left with a pure $U(N)$ gauge theory. The effective scale of this theory must be related to the scale of the theory upstairs. In fact, if the scale Q at which the theory is probed is larger than μ the theory is a $\mathcal{N} = 2$ theory for which the coupling constant at the scale Q is

$$\frac{1}{g^2(Q)} = \frac{c_0}{2\pi} \log\left(\frac{Q}{\Lambda}\right), \tag{2.49}$$

where c_0 is the one-loop β function coefficient which is $2N$ for pure glue $\mathcal{N} = 2$ $U(N)$ theories. At scales Q lower than μ the theory is an effective $\mathcal{N} = 1$ theory for which

$$\frac{1}{g^2(Q)} = \frac{c_0'}{2\pi} \log\left(\frac{Q}{\Lambda_{eff}}\right). \tag{2.50}$$

Here c_0' is the one-loop β function coefficient which is $3N$ for pure glue $\mathcal{N} = 1$ $U(N)$ theories. The coupling constant must be a continuous function of the scale and therefore must match at the threshold value of the scale, the scale μ. This gives

$$\Lambda_{eff}^3 = \mu\Lambda^2. \tag{2.51}$$

The $\mathcal{N} = 1$ $U(N)$ gauge theory is expected to confine in the infrared, leaving a $U(1)$ gauge theory. In this theory some instanton contributions to the vev of the chiral superfield can be calculated. The contribution to the field $S = \mathrm{Tr}\, W^\alpha W_\alpha$ with the trace over the $SU(N)$ gauge theory indices, can be calculated to be

$$<S^N>= \Lambda_{eff}^{3N}. \tag{2.52}$$

The vacuum expectation value of a product of chiral operators in a supersymmetric field theory has the special property

$$<S^N>=<S>^N \tag{2.53}$$

by independence of the location of the operators due to supersymmetry, the fact that chiral field do not transform under half the supersymmetry algebra and cluster decomposition arguments. Therefore

$$<S>= \left(\Lambda_{eff}^{3N}\right)^{\frac{1}{N}} = \epsilon_i \Lambda_{eff}^3 \tag{2.54}$$

holds, where ϵ_i is a N-th root of unity, $\epsilon_i = e^{2\pi \frac{t}{N}}, t \in \{1, \ldots, N\}$. Since a field which transforms non-trivially under chiral symmetry transformations gets a vev this shows chiral symmetry breaking. The value on the right hand side is a fractional power of the usual instanton contribution, it is referred to as a fractional instanton contribution. Regular instanton contributions have an interpretation as certain gauge field configurations which are saddle points in the Euclidean path integral. It would be interesting to find out if there is a similar statement to be made about fractional instantons.

The vev for S written above is reproduced by the Veneziano-Yankielowicz [20] effective superpotential

$$W_{VY}(S) = S\left[\log\left(\frac{\Lambda_{eff}^{3N}}{S^N}\right) + N\right]. \tag{2.55}$$

Actually this superpotential can be derived from another argument. The needed vev can be calculated by taking a derivative with respect to the effective coupling in equation (2.24). Since the effective coupling can simply be replaced by the logarithm of an effective scale, can always take a derivative of the effective superpotential with respect to $\log(\Lambda_{eff}^{3N})$ to obtain $< S >$. According to the Intriligator-Leigh-Seiberg [21] linearity principle, this is also true non-perturbatively, and the coupling $\log(\Lambda_{eff}^{3N})$ will only appear linearly in the full effective superpotential. Given a superpotential as a function of the effective scale the superfield S can then be 'integrated in' by a Legendre transform [22, 21]. The idea is that a given superpotential is the low-scale effective superpotential of a superpotential which lives at a higher scale. Integrating out massive fields from that higher scale superpotential then should give back the given low-scale effective superpotential. If it can be proven that the extra massive fields can be integrated out exactly, the higher scale superpotential can then also be calculated exactly by the Legendre transformation. The application of 'integrating in' in the example below should further clarify the procedure.

The simplest possible superpotential which involves only the scale which can be written down is a superpotential proportional to the scale to the appropriate power,

$$W = N\Lambda_{eff}^3 = N(\Lambda_{eff}^N)^{\frac{1}{N}}. \tag{2.56}$$

In the second line we effectively used the presumption that this can be rewritten in terms of the instanton factor. This superpotential can also be obtained by Legendre transforming with respect to the source for S as

$$W = B + S\log\left(\frac{\Lambda_{eff}^{3N}}{B^N}\right). \tag{2.57}$$

Integrating out the field S in the above superpotential gives back 2.56. If instead the field B is integrated out, equation (2.55) is obtained.

The above observations from the $\mathcal{N}=1$ perspective were made for gauge symmetry breaking to one subgroup in the parent $\mathcal{N}=2$ gauge theory. If the gauge symmetry is instead broken to several subgroups, then the superpotential will in general look like a sum of Veneziano-Yankielowicz terms for the different gauge subgroups plus corrections suppressed by the overall scale of the superpotential. An exact expression for the expansion coefficients in terms of the classical superpotential coupling constants and the relationship of above calculations to the $\mathcal{N}=2$ perspective is the subject of Dijkgraaf-Vafa theory, to be discussed below. Note that in the case the gauge symmetry is broken to several subgroups the 'integrating in' procedure will give a superpotential for the field $\sum_i S_i$, since this is in (2.24) the field which multiplies the coupling constant, which is the source used above for Legendre transforming.

2.2.3 Dijkgraaf-Vafa

Based on the previous work [23], in [24, 25, 26] Dijkgraaf and Vafa proposed that the sought after superpotential can be calculated by studying the 't Hooft large M dynamics of an auxiliary matrix model. This proposal was derived within string theory, but can be understood purely within field theory. The matrix model in question is specified by the partition function

$$Z = \int dK_{M\times M}\, e^{-\frac{1}{g}\mathrm{Tr}\, W_{cl}(K)}. \tag{2.58}$$

Here the integral is over Hermitian $M \times M$ matrices and the potential is taken to be exactly equivalent to the classical superpotential (2.40). Models of this type have been widely studied in a variety of applications and a lot about them is known, especially about their large M limits (see for instance [27]). The crucial observation which renders these models manageable is the fact that Hermitian matrices can be diagonalized by a unitary transformation. This allows one to perform the integral over the off-diagonal entries of the matrix model exactly and reduces the model to a model only depending on eigenvalues. The price to be payed for this integration is the introduction of a measure factor in the model,

$$Z = \int \prod_k d\lambda_k \prod_{i<j} (\lambda_i - \lambda_j)^2\, e^{-\frac{1}{g}\sum_l W_{cl}(\lambda_l)}. \tag{2.59}$$

The determinant which appears is called the Vandermonde determinant[6] and can be exponentiated to be included in the potential,

$$Z = \int \prod d\lambda_i e^{-\frac{1}{g}\left(\sum_l W_{cl}(\lambda_l) + g\sum_{i\neq j}\log(\lambda_i - \lambda_j)\right)}. \tag{2.60}$$

The logarithmic contribution causes a repulsive force between the eigenvalues, so we expect that the vacuum state of this theory is a distribution of eigenvalues around the critical points of the classical potential W, since there is a competition between the forces generated by the classical potential and by the logarithm. The standard avenue of attack for these kind of path integrals is the method of steepest descent, so we study the region where g becomes very small and calculate the saddle point. In principle this forces one to choose a particular contour for the integrals over the eigenvalues, but the issue of choosing contours in the context of these particular matrix models has not been addressed yet. In a sense this does not matter since the primary interest will be in the matrix model as a generating function for an asymptotic (large M) expansion. The equation of motion for an eigenvalue is

$$\frac{1}{g} W'_{cl}(\lambda_i) = \sum_{j\neq i} \frac{1}{\lambda_j - \lambda_i}. \tag{2.61}$$

This equation is still difficult to solve, but if we multiply both side of the equation by $\frac{1}{M}$, the right-hand side starts to look like the discretization of an integral. For this to be literally true, the number of eigenvalues M has to be taken to infinity. In order to keep the structure of the equation fixed in this limit, it must be taken keeping $gM \equiv \mu$ fixed. This limit can be shown [27] to be precisely the limit for which in a diagrammatic expansion only planar diagrams survive (i.e. diagrams which can be drawn on a sphere). The continuum limit of equation (2.61) is

$$\frac{1}{\mu} W'_{cl}(\lambda(y)) = \int \frac{1}{\lambda(x) - \lambda(y)} dx. \tag{2.62}$$

By defining a density of eigenvalues as $\left(\frac{d\lambda}{dx}\right)^{-1} = \rho(\lambda)$ this equation can be written as the so-called loop equation

$$\frac{1}{\mu} W'_{cl}(\lambda') = \int \frac{\rho(\lambda)}{\lambda - \lambda'} d\lambda. \tag{2.63}$$

The function on the right-hand side is called the resolvent of the matrix model field Φ. As shown in [27] to which the reader is also referred for some examples, the loop equation determines the density of eigenvalues, and therefore determines the saddle-point configuration which dominates the eigenvalue integrals (2.60). The

[6] This determinant is named after the french mathematician Vandermonde, although according to [28] he did not actually write it down.

key observation is that the density of eigenvalues ρ must have compact, possibly disconnected support in the complex plane. This gives the function of the right-hand side very specific analytic properties which allow it to be determined.

Our primary interest however will be slightly more geometric and first we will derive, following [24], an equation for the resolvent which defines a Riemann surface. A standard trick is to multiply (2.61) by $\frac{1}{M}\frac{1}{x-\lambda_i}$ and sum over i. Now define a function $\omega(x) = \frac{1}{M}\sum_i \frac{1}{x-\lambda_i}$ which is the discrete version of the right-hand side of (2.62) and can be used to rewrite the above equation as

$$\frac{1}{\mu}W'_{cl}(x)\omega(x) + \frac{1}{\mu}\sum_i \frac{W'_{cl}(\lambda_i) - W'_{cl}(x)}{x-\lambda_i} + \frac{1}{M}\omega'(x) + \omega(x)^2 = 0 \tag{2.64}$$

with $\mu \equiv gN$. In the large M limit, the function on the right-hand side of (2.62) had the interpretation as a resolvent function for the density of eigenvalues. If M is taken to be large, while μ fixed the equation written above becomes algebraic. The second term in that equation is defined to be a function $f_n(x)$,

$$f_n(x) \equiv \mu \sum_i \frac{W'_{cl}(\lambda_i) - W'_{cl}(x)}{x-\lambda_i}. \tag{2.65}$$

This function f is a polynomial since it cannot have singularities by inspection of its definition and is of order n if the superpotential is of order $n+1$ (see 2.40). With the definition $y = 2\mu\omega(x) + W'(x)$, equation (2.64) transforms into

$$y^2 = W'_{cl}(x)^2 - 4f_n, \tag{2.66}$$

and a Riemann surface appears which bears a strong resemblance to the curve appearing in (2.46). However, from the matrix model point of view the coefficients of the polynomial f are free parameters, while in (2.46) they are fixed by the factorization conditions. This discrepancy will be fixed below. Since the function ω has the interpretation a resolvent function for the density of eigenvalues, it can be seen from (2.66) that the eigenvalues are in general grouped into a maximum of n different cuts. This also follows from considering the classical limit where it corresponds to breaking the classical symmetry group of the matrix model $U(M)$ to n (or less) subgroups:

$$U(M) \to U(M_1) \times \cdots \times U(M_n). \tag{2.67}$$

There are now several ways to obtain the free energy of the matrix model in the large M limit, which all lead to the same equations. In this thesis the one immediately relevant is the calculation in appendix B of [29] (see also [24] for a nice physical interpretation of the equations which will be arrived at in (2.70)). There it is shown that the free energy can be calculated from the density of eigenvalues in the large N limit by considering

$$Z \sim e^{\frac{1}{g^2}\mathcal{F}^{(0)}+\cdots}, \tag{2.68}$$

with the so-called sphere contribution given as

$$F^{(0)} = \int d\lambda\, \rho(\lambda) W(\lambda) + \int d\lambda d\lambda'\, \rho(\lambda)\rho(\lambda') \log\left(\frac{\lambda - \lambda'}{\Lambda}\right) + \sum_i \mu_i \int d\lambda\, \rho_i(\lambda). \quad (2.69)$$

Here the dots indicate terms which are sub-leading in the large M limit, a scale has been introduced by dimensional arguments and the ρ_i ($\rho = \sum \rho_i$) are have been defined to be the densities of eigenvalues in the various different cuts. The μ_i are chemical potentials for the variables $S_i = \int \rho_i(\lambda) d\lambda$. Calculation of the saddle-point of this system and then Legendre transforming gives an expression for the value of the free energy as a function of the S_i. The upshot of the analysis is that $F^{(0)}$ can be calculated by the calculation of certain period integrals on the effective Riemann surface as

$$\begin{aligned} S_i(\{f_k\}) &= \oint_{\alpha_i} d\lambda_{DV} \\ \frac{\partial \mathcal{F}}{\partial S_j}(\{f_k\}) &= \oint_{\beta_j} d\lambda_{DV} \\ d\lambda_{DV}(\{f_k\}) &= \sqrt{W'_{cl}(x)^2 + f_n(x)}dx. \end{aligned} \quad (2.70)$$

The differential $d\lambda_{DV}$ is a holomorphic differential when varied with respect to the coefficients of the polynomial f. The S_i can be identified as the 't Hooft couplings for the different $U(M_i)$ subgroups in equation (2.67).

As was noted above, the coefficients in f are unfixed, while if this would be the non-singular part of the Seiberg-Witten curve from (2.46) they would have to be fixed. Since the variables S_i in (2.70) are essentially a rewriting of these variables fixing these will fix the curve. The question is then what conditions they have to satisfy in order for the connection to Seiberg-Witten curves to work. The answer given by [23] is that the variables S_i can be determined by calculating the minima of a superpotential which is constructed as

$$W_{\text{eff}} = \sum_i N_i \frac{\partial \mathcal{F}}{\partial S_i} + \tau_{class} \sum_i S_i. \quad (2.71)$$

In this expression the N_i are the different classical subgroups of the gauge group $U(N)$ in (2.42). They should not be confused with the M_i which are the sizes of the different $U(M_i)$ symmetry groups in the classical breaking pattern

$$U(M) \to U(M_1) \times \cdots \times U(M_n) \quad (2.72)$$

of the matrix model.

It can be proven that the minimum of equation (2.71) is indeed such that on-shell the matrix model curve in (2.66) is the non-singular part of the Seiberg-Witten curve [23]. A motivation that this is indeed the correct potential is that the equations of motion following from it basically state that there is one eigenvector of the period matrix which describes the decoupled overall $U(1)$ theory. The notation of the variables S_i is no coincidence as the reader might have guessed: these variables are identified as the expectation values of the glueball variables $S_i = \mathrm{Tr}_{\,i} W^\alpha W_\alpha$ for the different classical subgroups of the microscopic gauge group $U(N)$. As explained above, we expect for every of these glueball variables a Veneziano-Yankielowicz to arise, which in the Dijkgraaf-Vafa framework can be shown to arise from the measure of the matrix model, since they are the leading terms in the 't Hooft $\frac{1}{M}$ expansion of the volume of the $U(M_i)$ subgroups and the S_i are identified with gM_i. Minimizing W with respect to S yields the vevs of the S_i.

Although matrix model computations like above present a major advance in the study of supersymmetric gauge theory vacua, the question can be asked whether one actually computes a low energy effective potential: this can only be true if the glueball is the only low energy degree of freedom in the system. At the very least however, the superpotential constructed above encodes the values of the glueball vevs, which is important information. It is unclear however if the coefficients of terms like $S_1^{12} S_3^{16}$ have an independent physical meaning.

2.2.4 Connection to integrable systems

Above the details of $U(N)$ $\mathcal{N} = 2$ pure glue gauge theory are discussed, but the techniques mentioned however are far more general. In particular, for large classes of gauge theories Seiberg-Witten curves and differentials are known and a handy classification of them can be formulated in terms of completely integrable periodic integrable systems. From a physics point of view this is very surprising as there is a priori no reason why these classical systems should have anything to do with supersymmetric quantum field theory in the infrared. However, one of the central ingredients of the Seiberg-Witten solution is the appearance of Riemann surfaces. From a mathematics point of view it has been known since the seventies, beginning of the eighties that Riemann surfaces play a big role in the theory of integrable systems. In particular, finite dimensional completely integrable periodic integrable systems always give rise to Riemann surfaces as spectral curves, just as was shown in the introduction for the periodic Toda system. For that system, the spectral curve was written down in equation (1.11), reproduced here for convenience,

$$P_N(x) + (-1)^N (z + \Lambda^{2N} z^{-1}) = 0.$$

This equation can easily be solved for z. If we then define $y \equiv z + (-1)^N \frac{P(x)}{2}$ we obtain

$$y^2 = P_N(x)^2 - 4\Lambda^{2N}, \tag{2.73}$$

which can be recognized as the same curve as the Seiberg-Witten curve in equation (2.37) for $U(N)$ $\mathcal{N} = 2$ pure glue gauge theory. In fact, the gauge theory resolvent (2.45) is naturally interpreted as $d \log z$. The essential information in both theories is therefore related as

$$\text{Seiberg-Witten curve} \longleftrightarrow \text{spectral curve.} \tag{2.74}$$

More specifically, this provides for a precise one-to-one map of 4D gauge theory moduli to the conserved quantities (action variables) of the integrable system. The connection of Seiberg-Witten curves to integrable systems was noted in [30] and further elaborated upon in [31] for pure glue gauge theory based on all classical gauge groups. This connection has since been studied in many other papers. See for instance the first page of [32] for an overview over correspondences. A natural question is then what the angle variables correspond to on the gauge theory side. A better way of putting this question is: which gauge theory has the angle variables as extra field theory moduli? It turns out that this gauge theory can be obtained by compactifying one spatial direction on a circle.

2.3 Compactifications on $\mathbb{R}^3 \times S^1$

Four dimensional gauge theories with one spatial direction compactified on a circle have been studied from various perspectives and for various reasons in the literature (relevant here are for instance [33, 34, 35]). One reason to study them is because they interpolate between a simpler three dimensional gauge theory and the physical case of four dimensions. Also, theories of this sort appear when one studies gauge theory at finite temperature, although there the compactified direction is Euclidean time and any supersymmetry will be broken. The study of $U(N)$ supersymmetric gauge theory with one spatial direction compactified on a circle starts in three dimensions.

2.3.1 $\mathcal{N} = 4$ theories in $d = 3$

A four dimensional $\mathcal{N} = 2$ gauge theory can be dimensionally reduced to three dimensions, yielding a three-dimensional $\mathcal{N} = 4$ gauge theory (8 supercharges). Under the dimensional reduction, the four-dimensional vector field A_μ decomposes in one scalar field r and a three dimensional vector field A_α. The three-dimensional vector can in turn be dualized to a second scalar via $\partial_\alpha \sigma = \epsilon_{\alpha\beta\gamma} F^{\beta\gamma}$. The magnetic charge acts as a shift on the scalar σ and therefore by charge quantization σ is a periodic variable. The two real scalars combine into a complex scalar z, which is part of a 3d hypermultiplet. Therefore, in three dimensions, the vector multiplet is dual to a hypermultiplet. This duality can be performed directly in the Lagrangian for Abelian gauge groups (see e.g. [34]), but not for non-Abelian groups, similar to what happens with electric-magnetic duality in four dimensions. If we dimensionally reduce from four to three dimensions, σ will be a periodic variable, but r is unconstrained. The natural variable in which to express z is

$$e^{z_i} \sim e^{r_i + i\sigma_i}. \tag{2.75}$$

If we instead compactify from four to three dimensions the variable r_i becomes periodic as well. The periodicity of r is not apparent in the variable written above, but this is simply due to the coordinate choice.

Classically, the complex variable z is $r + i\sigma$, and the action only depends on $Z + \bar{Z}$, where Z is the superfield with lowest component z. In perturbation theory the action remains a function of $Z + \bar{Z}$ only, but the relation between the vev of z, r and σ can become quite complicated. Non-perturbatively, the action no longer needs to be a function of $Z + \bar{Z}$, as instantons generate a non-trivial dependence on the zero mode of σ.

The compactification of a pure glue four dimensional $U(N)$ $\mathcal{N} = 2$ gauge theory on $\mathbb{R}^3 \times S^1$ yields a three-dimensional theory with a moduli space which is parametrized by the vevs of ϕ and z. The gauge symmetry is broken generically to $U(1)^N$, and both ϕ and z are diagonal. The moduli space is a hyperKähler (by $\mathcal{N} = 4$ supersymmetry) manifold of real dimension $4N$, and contains the moduli space of the four-dimensional theory that was parametrized only by ϕ as a subspace. What we gain by going to three dimensions are the vevs of z, and these parametrize a $2N$-torus, which can be identified with the Jacobian[7] of the Seiberg-Witten curve [33]. According to [33], one of the complex structures of the moduli space is actually independent of the radius R. Furthermore they showed that this complex structure is the one that will be relevant once we break to $\mathcal{N} = 2$ (in $d = 3$) by turning on a classical superpotential. In other words, there are holomorphic coordinates on the moduli space in which the quantum superpotential can be expressed independent of R. Therefore, the vacuum structure we find in three plus one dimensions should be directly related to the vacuum structure in four dimensions. One should be careful however with limiting values of the radius since one of the assumptions in deriving the Seiberg-Witten result on three dimensional gauge theories is for instance the

[7]With the understanding, as stressed above, that one of the directions in the Jacobian is actually a lone circle direction of the overall $U(1)$ of the $U(N)$ gauge group we are considering here.

compactness of the scalar z which is invalid in the decoupling limit $R \to 0$. In that limit everything should first be expressed in the 3 dimensional gauge coupling and the radius and only then the limit should be taken.

The above paragraph gives a direct answer to the question which motivated the study of gauge theories compactified on a circle: these are precisely the gauge theories which contain the extra moduli needed to find a gauge theory counterpart to the angle variables in the integrable system. The above analysis was mainly for $U(N)$ gauge theories, but can be generalized in a number of ways to include, for instance, matter in the fundamental.

2.3.2 Central problem

Given the above analysis for gauge theories compactified on a circle a natural question is what the analog of Dijkgraaf-Vafa theory is for these theories and what the interrelation of this question is with the underlying integrable system. Just like in the four dimensional case, soft supersymmetry breaking can be studied by turning on a classical superpotential in one compact and three flat dimensions,

$$\int d^3x d\phi d^2\theta \, \mathrm{Tr} \, W(\Phi) \tag{2.76}$$

with, just as in (2.40)

$$W(\Phi) = \sum_{i=1}^{n+1} \frac{g_i}{i} \Phi^i .$$

In four dimensions the gauge symmetry was broken according to (2.46), and since the complex structure of the moduli space did not depend on R, it is expected that for every finite value of R this remains true. Therefore, we expect that the moduli space collapses to a finite collection of tori of dimension $2k$. Note that this is different from the four dimensional case in the sense that some of the gauge theory moduli (the ones associated to the Jacobian of the curve) remain unfixed.

Just as in the four dimensional case where we encountered fractional instanton effects after soft supersymmetry breaking, it is important to understand known non-perturbative effects in three-dimensional $\mathcal{N} = 2$ gauge theories first before the general case is considered.

2.3.3 $\mathcal{N} = 2$ perspective

Non-perturbative physics in three dimensions is due to three-dimensional instantons, which from the four dimensional point of view are monopoles. They are classified by $\pi_2(U(N)/U(1)^N) = \mathbf{Z}^{N-1}$. Indeed, for each simple root there is a corresponding embedding $SU(2) \subset U(N)$, and for each such embedding there is a corresponding elementary monopole. A general monopole configuration is therefore labeled by a set of integers $\{n_1, \ldots, n_{N-1}\}$, counting the number of elementary monopole constituents. In a pure $\mathcal{N} = 2$ theory in three dimensions, one can count the number of gaugino zero modes in a general monopole background using the Callias index theorem [36], and one finds that there are two zero modes only if one $n_i = 1$ and all other n_j vanish. We need two fermionic zero modes in order to get a non-trivial contribution to the superpotential, and therefore only the single elementary monopoles contribute to superpotential. Their contribution can be explicitly evaluated and the final result for the superpotential reads [37]

$$W_{\text{quantum}} = e^{(Z_1-Z_2)/g_3^2} + e^{(Z_2-Z_3)/g_3^2} + \ldots + e^{(Z_{N-1}-Z_N)/g_3^2}. \tag{2.77}$$

Here, Z_i represents the diagonal entries of Z, and in the exponents one recognizes the simple roots of $U(N)$; g_3 is the three-dimensional Yang-Mills coupling. The result (2.77) is exact, and only has runaway solutions. In a sense, (2.77) is less subtle than the Veneziano-Yankielowicz effective superpotential (2.55), because (2.77) involves a sum of ordinary instantons, and fractional instantons play no role.

2.3.4 $\mathcal{N} = 2$ deformations on $\mathbb{R}^3 \times S^1$

Now we examine what happens if the $4d$ theory is put on $\mathbb{R}^3 \times S^1$. In addition to the monopoles that contributed to (2.77), there is one more non-trivial gauge field configuration that contributes, which is the Kaluza-Klein monopole. This is present due to the existence of large gauge transformations along the S^1 [38].

The KK monopole adds one extra contribution to (2.77), and it becomes [33]

$$W_{\text{quantum}} = e^{(Z_1-Z_2)/g_3^2} + \ldots + e^{(Z_{N-1}-Z_N)/g_3^2} + e^{-1/Rg_3^2} e^{(Z_N-Z_1)/g_3^2}. \tag{2.78}$$

The three-dimensional answer is recovered by taking $R \to 0$, whereas the decompactification limit is $R \to \infty$ while keeping the dynamical (4d) scale

$$\Lambda_{eff}^{3N} \equiv e^{-\frac{1}{g_4^2}} \equiv e^{-\frac{1}{Rg_3^2}} \tag{2.79}$$

fixed. In order to study the minima of (2.78) first a different set of variables is introduced.

$$\begin{aligned} y_1 &= e^{(Z_1-Z_2)/g_3^2} \\ &\vdots \\ y_{N-1} &= e^{(Z_{N-1}-Z_N)/g_3^2} \\ y_0 &= e^{-1/Rg_3^2} e^{(Z_N-Z_1)/g_3^2}. \end{aligned} \tag{2.80}$$

The variables y_i are not unconstrained but obey

$$\prod_{i=0}^{N-1} y_i = \Lambda_{eff}^{3N}. \tag{2.81}$$

Note that this change of variables is only well-defined if R is *finite*. To impose this constraint, we introduce a Lagrange multiplier field L, and with this field the superpotential (2.78) can be rewritten as

$$W_{\text{quantum}} = y_1 + \ldots + y_{N-1} + y_0 + L \log\left(\frac{\Lambda_{eff}^{3N}}{\prod_{i=0}^{N-1} y_i}\right). \tag{2.82}$$

The minima of (2.82) are easily found, they are in terms of the N-roots of unity ϵ_i we encountered before

$$y_{i,\text{min}} = L_{\text{min}} = \Lambda_{eff}^3 \epsilon_i, \qquad W_{\text{min}} = N\Lambda_{eff}^3 \epsilon_i. \tag{2.83}$$

Indeed, the dependence on R has dropped out of (2.82) and (2.83), and the results are identical to the result in (2.56). This result is therefore equivalent what was obtained through the Veneziano-Yankielowicz superpotential. Actually, the field S can be integrated in directly in equation (2.82) by the same procedure as before. First write

$$W_{\text{quantum}} = y_1 + \ldots + y_{N-1} + y_0 + L \log\left(\frac{B^N}{\prod_{i=0}^{N-1} y_i}\right) + S \log\left(\frac{\Lambda_{eff}^{3N}}{B^N}\right). \tag{2.84}$$

Integrating out S gives the same superpotential as before. Integrating out B gives equation (2.82) with the variable L identified as the glueball field S. The deeper reason this works is simply the fact that L couples in the same way to the source $\log \Lambda$ as S or in the more general case, $\sum S_i$.

Before discussing the general case, we discuss the simplest deformation of an $\mathcal{N} = 4$ theory to a $\mathcal{N} = 2$ theory, namely by a mass term $W(\Phi) = \frac{1}{2} m\Phi^2$. In the presence of such a mass term, there is a pure $\mathcal{N} = 2$ gauge theory at low energies, whose scale Λ_{eff} is related to the high energy $\mathcal{N} = 4$ scale Λ via scale matching as

$$\Lambda_{eff}^3 = m\Lambda^2. \tag{2.85}$$

If we substitute this in (2.82), and redefine $y_i \to m y_i$, the superpotential becomes

$$W = m \sum_{i=0}^{N-1} y_i + S \log\left(\frac{\Lambda^{2N}}{\prod_{i=0}^{N-1} y_i}\right). \tag{2.86}$$

The extrema of this superpotential are as before, we merely made a change of variables. The superpotential (2.86) only contains gauge degrees of freedom, but could also have been chosen to include the diagonal entries of Φ in the superpotential. They would simply appear through an extra mass term,

$$W = m \sum_{i=0}^{N-1} y_i + S \log \left(\frac{\Lambda^{2N}}{\prod_{i=0}^{N-1} y_i} \right) + \frac{1}{2} m \sum_{i=0}^{N-1} \phi_i^2. \tag{2.87}$$

The minima of (2.87) are the same as before, as $\phi_i = 0$ at the extremum. As can be seen by comparing this expressions to equation (1.9), (2.87) is equal to $\frac{1}{2} m \mathrm{Tr} M^2$, with M the Lax matrix of the periodic Toda chain. The simplicity of this final answer is one of the motivations for a conjecture on how the exact quantum superpotential looks like in general.

2.4 Central conjecture

We are now in a position to formulate and motivate the central conjecture studied in this thesis[8]. The key observations are that the only known non-perturbative effects contributing to the superpotential are regular (not fractional) instantons and there are no candidate degrees of freedom which can become light anywhere on the moduli space. Therefore, given a classical superpotential which breaks $\mathcal{N} = 4$ to $\mathcal{N} = 2$, all that needs to be done to find the quantum superpotential is finding the right holomorphic coordinates on the (hyperKähler) moduli space. Considering the fact that for a quadratic superpotential in pure $U(N)$ theories these coordinates were provided by the underlying integrable system and the fact that this system naturally provides for a unified description of the moduli space, these right coordinates are conjectured to be the holomorphic phase space variables of the underlying integrable system.

Let us first make this precise in the case of pure $U(N)$ gauge theories. For every term in the classical superpotential a corresponding term must be found on the integrable systems side. It is natural to look for these as an expression in the action variables of the integrable system. Since the Seiberg-Witten curve and spectral curve of the integrable system are the same, equation (2.45) can be used in the integrable system to calculate natural counterparts to the terms in the classical superpotential. The resulting expression is conjectured to be the exact quantum superpotential for pure $U(N)$ gauge theories compactified on a circle; a conjecture which will be investigated in detail in the next chapter. Here the q and p are naturally interpreted as chiral superfields (and are therefore naturally complex). Since for low order $< N$

[8]related observations and, indeed, some of the motivation for this conjecture can be found in [39, 40, 41]

superpotentials the calculation through the gauge theory resolvent simply enforces the Newton-Girard formula, in these cases we simply have

$$\int d^4xd^2\theta\,\mathrm{Tr}\,W(\Phi) \longrightarrow \int d^4xd^2\theta\,\mathrm{Tr}\,W(M), \tag{2.88}$$

where M is the Lax matrix. For these cases, an exact quantum superpotential is obtained by inserting the Lax matrix directly into the classical superpotential. In order to properly identify the field theory degrees of freedom one can compare to the quadratic superpotential case. Field theory results give equation (2.87), while the above conjecture gives the same, identifying q and p with Z and Φ. As mentioned before, this superpotential should reproduce both three and four dimensional vacuum structure. A very basic check is therefore that on the minima of the conjectured superpotential the Seiberg-Witten curve (hence the spectral curve) should factorize for pure glue gauge theory just as it did in four dimensions (2.46). This is a stringent check on the validity of the constructed superpotential, as it must reproduce four dimensional gauge theory results.

Above we mainly discussed $U(N)$ gauge theory but using the same reasoning as above one can construct a candidate exact potential W_{int} for *any* four-dimensional $\mathcal{N}=2$ gauge theory, compactified on $\mathbb{R}^3 \times S^1$, broken to $\mathcal{N}=1$ by turning on a classical superpotential for which a Seiberg-Witten description and an underlying integrable system is known. From the $\mathbb{R}^3 \times S^1$ perspective this procedure identifies the phase space coordinates of the integrable system with an open coordinate patch in the quantum moduli space of the gauge theory. The structure of the four-dimensional gauge theory vacua we are interested in can be derived from a $\mathcal{N}=2$ perspective. In this derivation a precise definition must be given of the operators appearing in the classical superpotential, such as the relationship between $\mathrm{Tr}\,\Phi^k$ and the moduli space parameters appearing in the Seiberg-Witten curve. For any definition used in the gauge theory derivation, a candidate exact superpotential W_{int} can be constructed on the integrable systems side by using the same definition. In this thesis the definition stated in equation (2.45) will be used, but based on the evidence we will find for the conjecture made in this section it is natural to suspect that the minima of W_{int} reproduce the gauge theory results for any definition as long as one uses the same definition on both sides.

2.4.1 Some subtleties

An issue now can be whether or not the integrable system provides for a global coordinate system: it is easy to imagine that several coordinate patches are needed to describe the full moduli space. Indeed, there can be more than one integrable system corresponding to the same four dimensional gauge theory, which all provide coordinate patches for part, but perhaps not all of moduli space. If a certain integrable system covers only part of the moduli space this implies that the minima of the constructed W_{int} for this system only describe some, but not all possible gauge theory vacua of the four dimensional gauge theory. The analysis and structure of all possible gauge theory vacua can in general be very complicated as will be illustrated in 4 and it would a priori be surprising if the full structure was reproduced on the integrable system side.

Another issue is the edge of the phase space: from the integrable system point of view solutions to the equations of motion which require one or several of the coordinates to be infinite are usually discarded from the solution set. Since these points are on the edge of the coordinate patch for the moduli space it is not clear whether these should correspond to good gauge theory vacua. This can either be a coordinate singularity (a canonical coordinate transformation will map this point to a finite solution) or real singular solutions which should be discarded. A third issue is the relationship to the microscopic variables: for the pure $U(N)$ case this relationship was provided for by comparing a semi-classical limit of the holomorphic coordinates to an instanton computation. The latter can be difficult in general. A final issue is the fact that there can be more than one integrable system describing the same Seiberg-Witten curve, which all provide coordinate patches for the gauge theory which may or may not overlap.

With these considerations in mind we then have two related conjectures:

Conjecture 1 *The minima of W_{int} correspond to vacua of the four-dimensional gauge theory. In particular, the Seiberg-Witten curve factorizes precisely as expected for these vacua.*

and

Conjecture 2 *W_{int} is the exact quantum superpotential for (4d) $\mathcal{N} = 2$ gauge theories compactified on a circle softly broken to $\mathcal{N} = 1$ by turning on a classical superpotential*

where of course proof of the (more mathematical) first one provides strong evidence for the second (more physical) one. If true, these conjectures provide for instance a way of studying partial supersymmetry breaking in four dimensions which is technically much simpler than the matrix model. For instance, no matrix integrals have to be performed. Mathematically, they provide a motivation to study equilibria of a certain class of potentials in finite, periodic integrable systems.

In this thesis the above conjectures will be investigated for pure $U(N)$ gauge theories (chapter 3) and for $U(N)$ gauge theories with matter in the fundamental representation (chapter 4). The first conjecture will be proven for pure $U(N)$ gauge theories. It is worthwhile to note that in the literature several related investigations of the above conjecture have appeared. In [42] both $U(N)$ gauge theory and $U(N)^*$ gauge theory were treated, and a different proof on conjecture 1 provided. In [43] the case of pure glue SO/SP gauge theories was analyzed in several examples. New results for gauge theory based on the exceptional group G_2 were obtained in [44]. The conformal case of $SU(2)$ gauge theories with 4 flavors was studied in [45]. There have even been some results for a class of 5 dimensional theories in [46].

Based on the evidence which will be found for these conjectures in the coming two chapters, the question may be asked if this conjecture can be verified by studying, for instance, the field theory relationship between Seiberg-Witten and Dijkgraaf-Vafa in more detail. From a purely field theory perspective this seems difficult to do. Although the expressions for the matrix model free energy and the $\mathcal{N}=2$ prepotential look similar, the field theory origin of the terms is completely different. The prepotential consists of a logarithmic expression which is derived in perturbation theory and a non-perturbative expansion, while the matrix model free energy consists of a similar logarithmic expression which now has an non-perturbative origin, and a perturbative expansion. Seiberg-Witten and Dijkgraaf-Vafa seem much more related from an integrable systems perspective. The matrix model is usually considered to be related to the Toda lattice hierarchy while in Seiberg-Witten theory the periodic Toda system arises. A natural suspicion is therefore that the two theories may be related directly by studying their integrable properties in detail. This suspicion will be the focus of the final chapter.

Chapter 3

Pure glue $U(N)$ gauge theory and the periodic Toda system

In the previous chapter a concrete proposal was made for the exact quantum superpotential for *any* $\mathcal{N}=2$ gauge theory on $\mathbb{R}^3 \times S^1$ softly broken to $\mathcal{N}=1$ by turning on a superpotential for the adjoint chiral multiplet. It was already shown that for a quadratic superpotential in a pure glue $U(N)$ theory this is remarkably similar to the superpotential obtained by studying non-perturbative effects. In this chapter this proposal is first further investigated by working out some examples and large classes of solutions. Then it will be proven in two different ways that the spectral curve of the Toda chain factorizes indeed precisely as in (2.46). This proves conjecture 1 in this specific case.

3.1 Periodic Toda chain

As mentioned in the introduction, the Lax matrix of the periodic Toda chain is given in variables $p_i \equiv \phi_i$ and $y_i = \Lambda^2 e^{q_i - q_{i+1}}$ as

$$M = \begin{pmatrix} \phi_1 & y_1 & 0 & \dots & z \\ 1 & \phi_2 & y_2 & \dots & 0 \\ 0 & 1 & \phi_3 & \dots & 0 \\ . & . & . & \dots & y_{N-1} \\ y_0 z^{-1} & . & . & 1 & \phi_N \end{pmatrix}. \tag{3.1}$$

Note that the variables y obey $\prod_{i=0}^{N-1} y_i^2 = \Lambda^{2N}$. One class of related Lax matrix representations can be obtained from the above one by conjugation with a diagonal matrix which can be used, for instance, to make the Lax matrix symmetric:

$$\hat{M} \equiv \begin{pmatrix} \phi_1 & \hat{y}_1 & 0 & \dots & \hat{y}_0 z \\ \hat{y}_1 & \phi_2 & \hat{y}_2 & \dots & 0 \\ 0 & \hat{y}_2 & \phi_3 & \dots & 0 \\ . & . & . & \dots & \hat{y}_{N-1} \\ \hat{y}_0 z^{-1} & . & . & \hat{y}_{N-1} & \phi_N \end{pmatrix}. \tag{3.2}$$

The variables $\hat{y}_i$ are defined as $\sqrt{y_i}$. Yet another representation, which will be used in the algebraic proof in the latter parts of this chapter, can be obtained by using the embedding $\widehat{gl}_N$ in gl_∞,

$$\tilde{M} = \begin{pmatrix} \dots & \dots & \dots & \dots & & \\ \dots & \phi_N & \hat{y}_0 & 0 & 0 & \\ \dots & \hat{y}_0 & \phi_1 & \hat{y}_1 & 0 & \dots \\ \dots & 0 & \hat{y}_1 & \phi_2 & \hat{y}_2 & \dots \\ & 0 & 0 & \hat{y}_2 & \phi_3 & \dots \\ & & \dots & \dots & \dots & \dots \end{pmatrix}. \tag{3.3}$$

The Lax matrix is then an infinite tridiagonal matrix $\tilde{M}$ periodic with period N. If we define a shift operator Z of order N by the equation $Z = D^N$, where D is the shift operator of order one with matrix elements $D_{ij} = \delta_{i,j-1}$, then clearly $\tilde{M}$ commutes with Z. By considering the action of $\tilde{M}$ on vectors that are eigenvectors of Z with eigenvalue z, so that they have only N independent components, we recover (1.10). By considering the action of Z on on vectors that are eigenvectors of $\tilde{M}$ with eigenvalue x one obtains a representation in terms of 2 by 2 matrices which will be discussed and used in chapter 4.

Another needed ingredient in the integrable system is the operator $\mathcal{L}$ which will act on integer powers of the Lax matrix M in order to generate time flows as in equation (1.10), reproduced here for convenience

$$\frac{\partial M}{\partial t_k} = [M, \mathcal{L}(M^{k-1})].$$

The Lax matrix M depends on a spectral parameter z and its powers can therefore be expanded as

$$M^n = \sum_r z^r M^n_{(r)} \tag{3.4}$$

Furthermore, define M^n_+ as the sum of the upper diagonal part of $M^n_{(0)}$ plus all negative powers of z $\sum_{r<0} z^r M^n_{(r)}$. M^n_- will be defined as the sum of the lower diagonal part of $M^n_{(0)}$ plus $\sum_{r>0} z^r M^n_{(r)}$, and finally M^n_0 as the diagonal part of $M^n_{(0)}$. In terms of these, $\mathcal{L}$ acts as,

$$\mathcal{L}(M^n) \equiv M^n_- + M^n_0 - M^n_+. \tag{3.5}$$

Note that in the infinite matrix representation $\tilde{M}^n_-$, $\tilde{M}^n_0$ and $\tilde{M}^n_+$ are simply the lower triangular, diagonal and upper triangular part of $\tilde{M}$ respectively. The linear operator (3.5) has a natural interpretation in terms of the affine root system of $\widehat{gl}_N$, it flips the signs of all positive roots but does nothing to the negative roots. For every $1 < n < N-1$ these $\mathcal{L}(M^n)$ generate independent isospectral flows on the phase space of the system as

$$\frac{\partial M}{\partial t_n} = [M, \mathcal{L}(M^{n-1})]. \tag{3.6}$$

These flows commute and clearly preserve the quantities $\mathrm{Tr}(M^l)$. The latter are the action variables of the periodic Toda chain, while the times t_n are essentially the angle variables, and altogether this demonstrates the integrability of the periodic Toda chain. The connection to the Hamiltonian formulation is provided for by the classical version of the R matrix formalism, as explained for untwisted affine Lie algebras in [6]. In this chapter the only ingredient needed from that formalism is the equivalence of the following equations of motion:

$$\begin{cases} \dfrac{\partial q}{\partial t_n} = \dfrac{1}{n}\{q, \mathrm{Tr}\, M^n\} \\ \dfrac{\partial p}{\partial t_n} = \dfrac{1}{n}\{p, \mathrm{Tr}\, M^n\} \end{cases} \qquad \longleftrightarrow \qquad \frac{\partial M}{\partial t_n} = [M, \mathcal{L}(M^{n-1})]. \tag{3.7}$$

Just as has been discussed in the previous chapter, the spectral curve associated to the Lax matrix M

$$\det(x - M) \equiv P_N(x) + (-1)^N(z + \Lambda^{2N} z^{-1}) = 0$$

is identified with the Seiberg-Witten curve of the four-dimensional $\mathcal{N} = 2$ $U(N)$ gauge theory.

The flows (1.10) preserve the product $\prod_{i=0}^{N-1} y_i^2 = \Lambda^{2N}$, and which can be enforced using a Lagrange multiplier S as indicated in the previous chapter. Since the flows do not affect Λ, we will not always write this Lagrange multiplier term explicitly in the remainder, but freely replace the product $\prod_{i=0}^{N-1} y_i^2$ by Λ^{2N} whenever appropriate.

The flows (1.10) also preserve the first and trivial Hamiltonian $\mathrm{Tr}(M)$. If we take M to be traceless, we are discussing the $SU(N)$ theory, while if we include the trace of M we are discussing the $U(N)$ theory. In the latter case, the moduli space also includes the expectation values of an extra complex scalar field which is obtained from the four dimensional diagonal $U(1) \subset U(N)$ gauge field after compactifying on a circle and dualizing the remaining three dimensional gauge field. This extra complex scalar is completely decoupled from the discussion since it cannot appear in the superpotential and we will ignore it in the remainder. In the remainder of the chapter we will assume the gauge group is $U(N)$ and briefly comment on the $SU(N)$ in subsection (3.8.3).

3.1.1 Conjecture details

According to the general conjecture we have to construct for a given superpotential a potential in the integrable system by using 2.45. It is easy to see that for potential terms $\mathrm{Tr}\,\Phi^n$, $n \leq N$ this corresponds to replacing the adjoint scalar in the classical superpotential directly by the Lax matrix. For sufficiently large m, the identification $\langle \Phi^m \rangle = \mathrm{Tr}(M^m)$ can no longer hold, because the right hand side will start to depend non-trivially on the spectral parameter. For higher order terms in the superpotential or for the calculation of higher order vevs there is a simple prescription which extends this easy replacement in accordance with the definition (2.45). This prescription will appear naturally in the course of the analysis and will be discussed in section 3.2.6.

3.1.2 Features of the proposal

Before showing some examples, we mention some general features of the superpotential (2.88).

First of all, it is interesting to observe that the quantum superpotential is simply obtained by substituting the Lax matrix in the classical superpotential for superpotentials of sufficiently low order. No matrix models need to be solved. Perhaps this is a manifestation of the observation that the superpotential of a d dimensional gauge theory appears to require solving a $d-4$ dimensional auxiliary theory, as suggested in [47]. This simply follows from the observation that a higher-than-four dimensional gauge theory can be reinterpreted as a four-dimensional gauge theory with an infinite number of fields. A -1-dimensional theory indeed requires no integrations whatsoever.

Another appealing feature of (2.88) is that it is polynomial in y_i, and therefore the expansion in powers of y corresponds to a monopole expansion. In the classical limit where $\Lambda \to 0$, all y_i should also be taken to zero, and all non-perturbative effects disappear. What is left is the classical superpotential expressed as a function of the ϕ_i.

In the periodic Toda chain, the center of mass coordinate $\sum q_i$ does not appear, and as a consequence the variables y_i are not independent but satisfy $\prod_i y_i = \Lambda^{2N}$. This center of mass coordinate corresponds to the $U(1)$ factor in $U(N) = U(1) \times SU(N)$, and this $U(1)$ is always left unbroken. It accordingly never appears in the superpotential. The remaining $U(1)^{k-1}$ gauge symmetry that is left unbroken after minimizing the exact superpotential (see (2.48)) will emerge as $k-1$ free parameters that are left after extremizing the superpotential. In other words, the superpotential will not have isolated vacua in general, but will have moduli spaces of vacua of real dimension $2(k-1)$. It is easy to see that the flows of the integrable system have to map extrema of the superpotential into extrema. Therefore what should happen is that $N-k$ of the flows leave the extremum invariant, while the remaining $k-1$ flows generate the $k-1$ free complex parameters dual to the unbroken $U(1)^{k-1}$. The gauge couplings of this unbroken $U(1)^{k-1}$ can be extracted from the spectral curve, by studying the period matrix of the spectral curve at the extrema.

The above picture is based on a comparison to the known results in four dimensions as obtained in the previous chapter, but we still need to show that this is indeed what happens. Intuitively, what happens is the following. The superpotential is a linear combination of action variables, and correspondingly generates a flow on the moduli space. In order for the superpotential to have an extremum, this flow needs to have a stationary point. Since all flows are linear motions on the Jacobian of the associated curve, this can only happen if the Jacobian and Seiberg-Witten curve degenerate. What is not clear is why the existence of a stationary point of a single flow should automatically imply that in fact $N-k$ flows are stationary at this point. Correspondingly, $N-k$ one-cycles of the Seiberg-Witten curve should collapse, and the Jacobian should degenerate to a $(k-1)$-dimensional complex torus.

A last feature we would like to point out is that the Lax matrix is not invariant under general permutations, i.e. under the Weyl group of $U(N)$. It is only form invariant under a $\mathbf{Z}_N$ subgroup of cyclic permutations of the ϕ_i and the y_i. A choice of classical vacuum configuration corresponds to choosing each ϕ_i equal to some a_j (the solution of $W'(x)=0$, see (2.41)), and all $y_i=0$. It is not guaranteed that all choices of ϕ_i have a corresponding minimum of the quantum superpotential. Because permutation symmetry is broken, it could be that one needs to order the ϕ_i in a suitable way in order to find a minimum of the quantum theory. We will see that this is indeed what happens. Precisely how one should order the eigenvalues in general is an interesting open problem that we have not yet been able to solve.

3.2 Examples

In this section we discuss in detail how our proposal works in various examples. For each of the classical vacua that gives rise to a symmetry breaking pattern as in (2.48), we expect to find $\prod_{N_i>0} N_i$ vacua, as each $U(N_i)$ low energy gauge group has N_i distinct quantum vacua. The various choices of vacua show up as choices of phases in the solutions below. We will compare the results to the four dimensional results as described in the previous chapter.

3.2.1 Gauge group $U(2)$

The first example that we will discuss is $U(2)$ with tree level superpotential

$$W = g_1 \mathrm{Tr}\,(\Phi) + \frac{g_2}{2}\mathrm{Tr}\,(\Phi^2) + \frac{g_3}{3}\mathrm{Tr}\,(\Phi^3). \tag{3.8}$$

The Lax operator in this case reads

$$M = \begin{pmatrix} \phi_1 & y_1 + z \\ 1 + y_0 z^{-1} & \phi_2 \end{pmatrix}. \tag{3.9}$$

The polynomial $P(x)$ appearing in the spectral curve equals

$$P(x) = (x-\phi_1)(x-\phi_2) - y_0 - y_1. \tag{3.10}$$

The resulting effective superpotential, using (2.88), reads

$$W_{\mathrm{eff}} = W(\phi_1) + W(\phi_2) + (y_0+y_1)(g_2 + g_3(\phi_1+\phi_2)) + S\log\left(\frac{\Lambda^4}{y_0 y_1}\right). \tag{3.11}$$

The equations for the extrema of W_{eff} are

$$\begin{array}{rclcrcl} W'(\phi_1) + g_3(y_0+y_1) & = & 0 & \qquad & g_2 + g_3(\phi_1+\phi_2) & = & \frac{S}{y_0} \\ W'(\phi_2) + g_3(y_0+y_1) & = & 0 & \qquad & g_2 + g_3(\phi_1+\phi_2) & = & \frac{S}{y_1} \end{array}$$
$$y_0 y_1 = \Lambda^4. \tag{3.12}$$

We now consider the various possible solutions of these equations.

<u>case 1: $S \neq 0$</u>. This necessarily implies that $y_0 = y_1 = \epsilon\Lambda^2$ with $\epsilon^2 = 1$. Next, we observe that the difference between the first two equations can be rewritten as

$$(\phi_1 - \phi_2)(g_2 + g_3(\phi_1+\phi_2)) = 0. \tag{3.13}$$

Since $S \neq 0$ and y_0 and y_1 can never be zero, the second factor in (3.13) cannot vanish, and therefore $\phi_1 = \phi_2$. The full solution is therefore the following. Define ϕ_0 to be any root of the equation

$$W'(\phi_0) + 2g_3\epsilon\Lambda^2 = 0, \tag{3.14}$$

then

$$\phi_1 = \phi_2 = \phi_0, \qquad y_0 = y_1 = \epsilon\Lambda^2, \qquad S = 2\epsilon\Lambda^2(g_2 + 2g_3\phi_0). \tag{3.15}$$

By taking the classical limit $\Lambda \to 0$, we see that ϕ_1 and ϕ_2 are both the same solution of the equation $W'(\phi) = 0$. Thus, this solution describes the maximally confining case with one solution of $W'(\phi) = 0$ doubly occupied, and the other not occupied. Classically the gauge symmetry is still $U(2)$, but it is broken quantum mechanically to $U(1)$. There are two different solutions depending on the choice of ϵ.

Let us compare this to the expected 4d field theory answer. According to the recipe in chapter (2), we first need to parametrize the locus where $P(x)^2 - 4\Lambda^4$ has a double zero, at say $x = x_0$. With $P(x) = x^2 + s_1 x + s_2$ we need that $P'(x_0) = 0$ and $P(x_0) = -2\eta\Lambda^2$ where $\eta^2 = 1$. This implies that $s_1 = -2x_0$ and $s_2 = x_0^2 - 2\eta\Lambda^2$. The effective superpotential, expressed in terms of s_1 and s_2, is equal to

$$W_{\text{eff}} = -g_1 s_1 + \frac{g_2}{2}(s_1^2 - 2s_2) + \frac{g_3}{3}(-s_1^3 + 3s_1 s_2), \tag{3.16}$$

and when evaluated on the locus where $P(x)^2 - 4\Lambda^4$ has a double zero at $x = x_0$ it becomes

$$W_{\text{eff}} = 2g_1 x_0 + g_2(x_0^2 + 2\eta\Lambda^2) + \frac{2g_3}{3}(x_0^3 + 6\eta\Lambda^2 x_0). \tag{3.17}$$

Extremizing this with respect to x_0 yields

$$2(W'(x_0) + 2g_3\eta\Lambda^2) = 0, \tag{3.18}$$

and for each solution x_0 the polynomial $P(x) = x^2 + s_1 x + s_2$ reduces to

$$P(x) = (x - x_0)^2 - 2\eta\Lambda^2. \tag{3.19}$$

This is exactly the same polynomial as the one that we obtain from (3.10) by substituting (3.15), with $\eta = \epsilon$. Thus, in the maximally confining case we reproduce the field theory answer.

<u>case 2: $S = 0$</u>. It is straightforward to find the solution in this case. We can take for example ϕ_1 to be completely arbitrary, so that $\phi_2 = -\phi_1 - g_2/g_3$. In addition, y_0 and y_1 are the two roots of the equation

$$y^2 + \frac{1}{g_3}W'(\phi_1)y + \Lambda^4 = 0. \tag{3.20}$$

To understand the classical limit we send $\Lambda \to 0$ and also take $y_i \to 0$. In that limit, ϕ_1 and ϕ_2 become the two distinct roots of the equation $W'(\phi) = 0$. Thus, this is the case where the gauge symmetry is classically broken to $U(1) \times U(1)$. In contrast to the previous situation, the quantum theory has a flat direction, parametrized by ϕ_1. As discussed in section (3.1.2), this flat direction represents the fact that at the quantum level there remains a $U(1) \times U(1)$ gauge symmetry. The first, diagonal $U(1)$ is completely decoupled and never present in the superpotential, the second

$U(1)$ is parametrized by ϕ_1. Equivalently, we can parametrize it by y_0/y_1, which is the combination that also represented this $U(1)$ classically. The polynomial (3.10) is equal to

$$P(x) = x^2 + \frac{g_2}{g_3}x + \frac{g_1}{g_3} = \frac{W'(x)}{g_3}. \tag{3.21}$$

As expected, the $P(x)$ does not depend on the free parameter ϕ_1 at all. According to (2.46) we get the same result in four dimensions. Therefore, also in this case we find complete agreement with 4d field theory.

In summary, the structure of the vacua in our proposal for three dimensions matches exactly the results in four dimensions, at least for $U(2)$ with a cubic superpotential.

3.2.2 Gauge group $U(3)$

Our next example is $U(3)$, again with tree level superpotential

$$W = g_1 \mathrm{Tr}\,(\Phi) + \frac{g_2}{2}\mathrm{Tr}\,(\Phi^2) + \frac{g_3}{3}\mathrm{Tr}\,(\Phi^3). \tag{3.22}$$

The Lax operator is

$$M = \begin{pmatrix} \phi_1 & y_1 & z \\ 1 & \phi_2 & y_2 \\ y_0 z^{-1} & 1 & \phi_3 \end{pmatrix}. \tag{3.23}$$

The polynomial $P(x)$ appearing in the spectral curve equals

$$P(x) = (x-\phi_1)(x-\phi_2)(x-\phi_3) - (y_0+y_1+y_2)x + (y_2\phi_1 + y_0\phi_2 + y_1\phi_3). \tag{3.24}$$

The resulting effective superpotential, again using (2.88), becomes

$$\begin{aligned} W_{\mathrm{eff}} &= W(\phi_1) + W(\phi_2) + W(\phi_3) + g_2(y_0+y_1+y_2) \\ &\quad + g_3(y_0(\phi_1+\phi_3) + y_1(\phi_1+\phi_2) + y_2(\phi_2+\phi_3)) + S\log\left(\frac{\Lambda^6}{y_0y_1y_2}\right). \end{aligned} \tag{3.25}$$

The equations for the extrema of W_{eff} are

$$\begin{aligned} W'(\phi_1) + g_3(y_0+y_1) &= 0 \qquad & g_2 + g_3(\phi_1+\phi_2) &= \tfrac{S}{y_1} \\ W'(\phi_2) + g_3(y_1+y_2) &= 0 \qquad & g_2 + g_3(\phi_2+\phi_3) &= \tfrac{S}{y_2} \\ W'(\phi_3) + g_3(y_2+y_0) &= 0 \qquad & g_2 + g_3(\phi_3+\phi_1) &= \tfrac{S}{y_0} \\ & & y_0y_1y_2 &= \Lambda^6. \end{aligned} \tag{3.26}$$

The first three equations can be used to solve for y_0, y_1 and y_2. When we substitute this in the next three equations, they become

$$\begin{aligned}(-W'(\phi_1) + W'(\phi_2) - W'(\phi_3))(g_2 + g_3(\phi_3 + \phi_1)) &= 2g_3 S \\ (-W'(\phi_1) - W'(\phi_2) + W'(\phi_3))(g_2 + g_3(\phi_1 + \phi_2)) &= 2g_3 S \\ (W'(\phi_1) - W'(\phi_2) - W'(\phi_3))(g_2 + g_3(\phi_2 + \phi_3)) &= 2g_3 S. \qquad (3.27)\end{aligned}$$

An interesting simplification appears if we consider the differences of pairs of these equations, which take the form

$$(\phi_i - \phi_j)(\phi_1 + \phi_2 + \phi_3 - 2a_1 - a_2)(\phi_1 + \phi_2 + \phi_3 - a_1 - 2a_2) = 0 \qquad (3.28)$$

for some i, j, and where a_1, a_2 are the two extrema of W',

$$W'(\phi) = g_3(\phi - a_1)(\phi - a_2). \qquad (3.29)$$

Again, we distinguish two cases. Either all ϕ_i are equal, or at least two of them are different.

case 1: All ϕ_i equal. This case is straightforward to analyze. All y_i are identical and equal to $\omega\Lambda^2$, with ω some third root of unity. All ϕ_i are also equal and a solution of the equation

$$W'(\phi_i) + 2\omega g_3 \Lambda^3 = 0. \qquad (3.30)$$

The polynomial in (3.24) becomes

$$P(x) = (x - \phi_1)^3 - 3\omega\Lambda^2(x - \phi_1) \qquad (3.31)$$

with ϕ_1 a solution of (3.30). For comparison with the field theory answer, we also compute $P(x)^2 - 4\Lambda^6$, which equals

$$P(x)^2 - 4\Lambda^6 = (x - \phi_1 - \omega^{\frac{1}{2}}\Lambda)^2(x - \phi_1 + \omega^{\frac{1}{2}}\Lambda)^2(x - \phi_1 - 2\omega^{\frac{1}{2}}\Lambda)(x - \phi_1 + 2\omega^{\frac{1}{2}}\Lambda). \qquad (3.32)$$

This has the expected form for the maximally confining case. Indeed, in the classical limit all ϕ_i reduce to either a_1 or a_2, so this situation is the one where classically the gauge group remains unbroken. In the quantum theory the gauge symmetry is broken to $U(1)$, and there are three different vacua depending on the choice of ω.

Let us briefly check that we get the same answer in field theory. First, we need to parametrize s_1, s_2 and s_3 in such a way that $P(x)^2 - 4\Lambda^6$ has two double zeroes. This parametrization is easily found to be

$$s_1 = -3t, \qquad s_2 = 3t^2 - 3\omega\Lambda^2, \qquad s_3 = -t^3 + 3\omega\Lambda^2 t. \qquad (3.33)$$

The effective superpotential, restricted to the locus (3.33) parametrized by t is

$$W_{\text{eff}}(t) = 3W(t) + 3g_2\omega\Lambda^2 + 6g_3\omega\Lambda^2 t \tag{3.34}$$

which is extremal for $W'(t) + 2g_3\omega\Lambda^2 = 0$. With this value for t, we can insert (3.33) in (3.24) to compute $P(x)$, and we immediately see it is the same as (3.31).

Both for $U(2)$ as well as $U(3)$ the maximally confining case had all ϕ_i equal, and all y_i equal as well. This continues to be true for $U(N)$, as we will elaborate on in section (3.2.3).

case 2: some ϕ_i different: in this case (3.28) implies that

$$\phi_1 + \phi_2 + \phi_3 = 2a_1 + a_2. \tag{3.35}$$

We could also have chosen the left hand side to be equal to $a_1 + 2a_2$, but we can then simply exchange a_1 and a_2, so there is no loss in generality in taking (3.35). The single linear relation (3.35) collapses the three equations in (3.27) into a single equation, since all differences vanish. We can choose to eliminate ϕ_3 from the equation using (3.35). We are then left with a single equation from (3.27), plus the additional equation $y_0y_1y_2 = \Lambda^6$. Those are two equations for the three variables S, ϕ_1, ϕ_2, and therefore there will be a free parameter in the solution. This is similar to what we saw for $U(2)$ when it is classically broken to $U(1) \times U(1)$. Here, as is clear from (3.35), we are considering the situation where $U(3)$ is classically broken to $U(1) \times U(2)$. In the quantum theory it is further broken to $U(1) \times U(1)$. One of these $U(1)$'s is the trivial diagonal $U(1)$, the other is the free parameter that we are find here.

The full solution can be parametrized as follows

$$\begin{array}{lllllll} \phi_1 & = & a_1 + \tilde{\phi}_1 & \quad & y_0 & = & \frac{\epsilon\Lambda^3}{\tilde{\phi}_2} \\ \phi_2 & = & a_1 + \tilde{\phi}_2 & & y_1 & = & \tilde{\phi}_1\tilde{\phi}_2 \\ \phi_3 & = & a_2 - \tilde{\phi}_1 - \tilde{\phi}_2 & & y_2 & = & \frac{\epsilon\Lambda^3}{\tilde{\phi}_1} \\ & & & S & = & -\epsilon g_3\Lambda^3. & \end{array} \tag{3.36}$$

with $\epsilon^2 = 1$ and $\tilde{\phi}_1, \tilde{\phi}_2$ subject to one constraint

$$\epsilon\Lambda^3 + \tilde{\phi}_1\tilde{\phi}_2(a_1 - a_2 + \tilde{\phi}_1 + \tilde{\phi}_2) = 0. \tag{3.37}$$

To compare with field theory we note that (3.36) substituted in (3.24) yields

$$P(x) = (x - a_1)^2(x - a_2) + 2\epsilon\Lambda^3, \tag{3.38}$$

from which we deduce

$$P(x)^2 - 4\Lambda^6 = (x - a_1)^2 (\frac{1}{g_3^2} W'(x)^2 + 4\epsilon\Lambda^3 (x - a_2)). \tag{3.39}$$

This is exactly the form that one would get from a field theory analysis and therefore the result agrees with the four-dimensional field theory expectation. Notice that the two solutions of $\epsilon^2 = 1$ correspond to two inequivalent quantum vacua, as expected for $U(3) \to U(1) \times U(2)$.

Thus, for $U(3)$ we also reproduce the complete vacuum structure that we expect from four dimensions. Before a final concrete example will be discussed, let us first check some broad classes of examples.

3.2.3 $U(N)$ with a quadratic superpotential

The analysis for a $U(N)$ theory with a quadratic superpotential,

$$W = g_1 \mathrm{Tr}\,(\Phi) + \frac{g_2}{2} \mathrm{Tr}\,(\Phi^2) + S \log\left(\frac{\Lambda^{2N}}{y_0 \dots y_{N-1}}\right) \tag{3.40}$$

can be done for general N. Substituting the Lax matrix for the adjoint scalar Φ into the superpotential gives the following effective superpotential

$$W_{\text{eff}} = \sum_{i=0}^{N-1} (g_2 y_i + g_1 \phi_{i+1} + \tfrac{1}{2} g_2 \phi_{i+1}^2) + S \log\left(\frac{\Lambda^{2N}}{y_0 \dots y_{N-1}}\right),$$

with the equations for the extrema

$$\begin{aligned} y_0 \dots y_{N-1} &= \Lambda^{2N} \\ g_2 - \frac{S}{y_i} = 0 &\Rightarrow y_i = \frac{S}{g_2} \\ g_2 \phi_i + g_1 &= 0. \end{aligned}$$

From the second equation we can learn that the y's are all equal, this enables us to solve for L using the first equation

$$S = \epsilon g_2 \Lambda^2, \qquad \epsilon^N = 1. \tag{3.41}$$

Substituting L into the equation for y_i yields

$$y_i = y = \Lambda^2 \epsilon.$$

Further, we also see that all the ϕ's occupy the root $x = -g_1/g_2$ of $W'(x)$

$$W'(x) = g_2 (x + \frac{g_1}{g_2}).$$

The ϕ's should always occupy the roots of $W'(x)$ classically ($\Lambda \to 0$), but here it is even true in the quantum case.

The superpotential in the extrema is then

$$W_{\text{extr}} = N(y - \frac{g_1^2}{2g_2}) = N(\Lambda^2 e^{\frac{2\pi i l}{N}} - \frac{g_1^2}{2g_2}).$$

To compare with field theory results we compute the characteristic polynomial of the Lax matrix $P_N(x) = \det(x1_N - \Phi)$. According to appendix A this is either a Chebyshev polynomial of the first or of the second kind. By evaluating $P_1(x)$ and $P_2(x)$ one can see that one has to pick the polynomials of the first kind

$$P_N(x) = 2y^{N/2} T_N\left(\frac{x+\frac{g_1}{g_2}}{2\sqrt{y}}\right).$$

These $P_N(x)$ are in perfect agreement with field theory results

$$\begin{aligned} P_N(x)^2 - 4\Lambda^{2N} &= 4\Lambda^{2N}\left(1 - T_N^2\left(\frac{x+\frac{g_1}{g_2}}{2\sqrt{y}}\right)\right) \\ &= 4\Lambda^{2N}\left(\left(\frac{x+\frac{g_1}{g_2}}{2\sqrt{y}}\right)^2 - 1\right) U_{N-1}^2\left(\frac{x+\frac{g_1}{g_2}}{2\sqrt{y}}\right) \\ &= \Lambda^{2(N-1)}\epsilon^{N-1}\left(\frac{1}{g_2^2}(W')^2 - 4\Lambda^2\epsilon\right) U_{N-1}^2\left(\frac{x+\frac{g_1}{g_2}}{2\sqrt{y}}\right). \end{aligned}$$

3.2.4 Maximally confining vacua

In this subsection the results of the previous one will be extended to a more general class of vacua. The particular class of vacua of interest will be the one for which the symmetry breaking pattern is

$$U(N) \overset{\text{classical}}{\longrightarrow} U(N) \overset{\text{quantum}}{\longrightarrow} U(1). \tag{3.42}$$

In field theory these vacua are those solutions to the factorization problem (2.46) for which the Seiberg-Witten curve fully degenerates. Since this factorization problem can in general be solved by Chebyshev polynomials, just as the polynomials appearing above, it is natural to suspect that the integrable system solution as presented in the case of a quadratic superpotential can be extended to a maximally confining solution to a general potential. These solutions will be constructed below for arbitrary N and for polynomial potentials of degree $n \leq N + 1$[1]. In the Chebyshev field theory solution, the factorization equations reduce to an equation on a free parameter

[1]This constraint will be relaxed in the following chapter in this thesis.

in the solution. Below it is shown that this remaining equation is also obtained by considering the simple vacua where all ϕ's are equal and all y's are a root of unity times the scale.

The maximally confining vacua of a $U(N)$ theory with superpotential

$$W(x) = \sum_{k=0}^{n} \frac{g_{k+1}}{k+1} x^{k+1} \tag{3.43}$$

are obtained by taking all $y_i = (\eta\Lambda)^2$ with $\eta^{2N} = 1$, and taking all $\phi_i = \phi$, where ϕ is such that $\mathrm{Tr}\ W'(M) = 0$. Because of the tridiagonal form of M this requirement on ϕ is equivalent to the condition $c_0 = 0$ for the coefficient in the Laurent expansion

$$W(\phi + \xi + y/\xi) = \sum_{k=-n}^{n} c_k \xi^k. \tag{3.44}$$

Expressed in terms of the coefficients of the superpotential, the condition $c_0 = 0$ reads

$$\sum_{k=0}^{n} g_{k+1} \left(\sum_{l=0}^{\lfloor k/2 \rfloor} \frac{k!}{(l!)^2 (k-2l)!} \phi^{k-2l} y^l \right) = 0. \tag{3.45}$$

with $\lfloor k \rfloor := \max\{m \in \mathbb{Z} | m \leq k\}$. This can be expressed in terms of a hypergeometric function

$$\sum_{k=0}^{n} g_{k+1} \phi^k \ {}_2F_1(-\frac{k}{2}, -\frac{k}{2} + \frac{1}{2}; 1; 4y/\phi^2) = 0. \tag{3.46}$$

On the other hand, according to (2.46) the factorization of the polynomial $P_N(x)$ appearing in the spectral curve imposes an apparently different condition on ϕ. With all ϕ_i equal, $P_N^2(x) - 4\Lambda^{2N}$ must have $N-1$ double zeroes:

$$P_N^2(x) - 4\Lambda^{2N} = T(x) H_{N-1}^2(x). \tag{3.47}$$

In this case P_N is a Chebyshev polynomial: $P_N(x) = 2(\eta\Lambda)^N T_N\left(\frac{x-\phi}{2\eta\Lambda}\right)$, with $\eta^{2N} = 1$. Together with the convenient definition of the Chebyshev polynomials of first and second kind

$$\begin{aligned} T_n(\cos x) &= \cos nx \\ U_n(\cos x) &= \frac{\sin\,(n+1)x}{\sin x}, \end{aligned} \tag{3.48}$$

this factorization is directly observed

$$\frac{P_N^2(x) - 4\Lambda^{2N}}{4\Lambda^{2N}} = \left(T_N^2(\frac{x-\phi}{2\eta\Lambda}) - 1 \right) = \left[\left(\frac{x-\phi}{2\eta\Lambda} \right)^2 - 1 \right] U_{N-1}^2 \left(\frac{x-\phi}{2\eta\Lambda} \right). \tag{3.49}$$

So after rescaling one obtains

$$T(x) = x^2 - 2\phi x + \phi^2 - 4y \tag{3.50}$$

for any maximally confining vacuum, irrespective of the rank of the gauge group.

From the point of view of four dimensional gauge theory, it is expected that $T(x)$ must be related to the superpotential as

$$T(x)\, G_{n-1}^2(x) = (W'(x))^2 + f_{n-1}(x) \tag{3.51}$$

for some polynomials $G_{n-1}(x)$, $f_{n-1}(x)$ of degree $n-1$. Equation (3.51) can be seen as a set of $2n+1$ conditions that the coefficients of each power of x match. Of these conditions $2n$ can be satisfied by appropriate choice of coefficients of $G_{n-1}(x)$ and $f_{n-1}(x)$. The final remaining condition turns out to be $c_0 = 0$, in (3.44), which relates ϕ and y to the superpotential at criticality.

To see this, examine the conditions that $G_{n-1}(x)$ and $f_{n-1}(x)$ can be found such that (3.51) is true. As a matter of convenience, we do a field redefinition

$$M \mapsto M - \phi I, \tag{3.52}$$

such that at criticality the redefined matrix M has $\phi = 0$ on the diagonal. Next a choice of scale is made such that $y = 1$. With these choices $T(x) = x^2 - 4$. By a substitution

$$x \mapsto \xi + \xi^{-1} \tag{3.53}$$

this can be written as a complete square

$$T = (\xi - \xi^{-1})^2. \tag{3.54}$$

So the entire left hand side of (3.51) is a complete square.

Note that on the r.h.s. of (3.51) the coefficients of terms proportional to x^n and higher powers of x are entirely determined by $W'(x)$. The polynomial $f_{n-1}(x)$ only serves to match the coefficients of lower powers of x. Taking the square root of (3.51) yields

$$\sqrt{T(x)G_{n-1}^2(x)} = W'(x) + \frac{1}{2}\frac{f_{n-1}(x)}{W'(x)} + \ldots = W'(x) + \mathcal{O}(x^{-1}). \tag{3.55}$$

With the substitution (3.53) , noting $x^{-1} \mapsto (\xi + \xi^{-1})^{-1} = \xi^{-1}/(1+\xi^{-2}) = \mathcal{O}(\xi^{-1})$, it is observed that (3.51) implies

$$(\xi - \xi^{-1})G_{n-1}(\xi + \xi^{-1}) = W'(\xi + \xi^{-1}) + \mathcal{O}(\xi^{-1}). \tag{3.56}$$

In order that a polynomial $G_{n-1}(x)$ of degree $n-1$ exists, the right hand side of (3.56) must be divisible by $(\xi - \xi^{-1})$. Choosing the $\mathcal{O}(\xi^{-1})$ polynomial in ξ^{-1} as follows almost guarantees this:

$$(\xi - \xi^{-1})G_{n-1}(\xi + \xi^{-1}) = W'(\xi + \xi^{-1}) - 2[W'(\xi + \xi^{-1})]_{-}. \tag{3.57}$$

Here $[g(\xi)]_{-}$ denotes the part of the Laurent series of g with strictly negative powers of ξ. The right hand side, written as a Laurent series in ξ is of the form

$$r.h.s. = c_0 + \sum_{k=1}^{n} c_k(\xi^k - \xi^{-k}). \tag{3.58}$$

This is divisible by $(\xi - \xi^{-1})$ iff $c_0 = 0$, with c_0 the coefficient of ξ^0 in the Laurent series of $W'(\xi + \xi^{-1})$. This is exactly the requirement for criticality, with the field redefinition and choice of scale that set $\phi = 0$ and $y = 1$.

3.2.5 Lifting solutions from $U(N)$ to $U(tN)$

It is known [23] that supersymmetric vacua of a $U(N)$ gauge theory with superpotential W can be lifted to supersymmetric vacua of a $U(tN)$ gauge theory with exactly the same superpotential. On the level of the Seiberg-Witten polynomial, this lifting involves Chebyshev polynomials and a few other ingredients. It turns out, as we will show in this section, that it is very simple to do this at the level of Lax operators. One simply takes t copies of the Lax operator of $U(N)$ to construct a Lax operator of $U(tN)$ that is periodic in steps of N. This new Lax matrix is an extremum for the same superpotential, and this replica trick therefore provides a very simple picture of how to lift vacua. To show technically how this works, we start with the Lax operator for $U(N)$

$$M_N = \begin{pmatrix} \phi_1 & y_1 & 0 & \dots & z \\ 1 & \phi_2 & y_2 & \dots & 0 \\ 0 & 1 & \phi_3 & \dots & 0 \\ . & . & . & \dots & y_{N-1} \\ y_0 z^{-1} & . & . & 1 & \phi_N \end{pmatrix} \tag{3.59}$$

and the corresponding polynomial $P_N(x)$ in the spectral curve

$$P_N(x) = \det(xI_N - M_N) + z + \Lambda^{2N} z^{-1}, \tag{3.60}$$

where the energy scale $\Lambda^2 \geq 0$ sets the condition

$$\Lambda^{2N} = \prod_{i=0}^{N-1} y_i. \tag{3.61}$$

As explained above, for a U(tN) theory with a superpotential W of degree $d \leq N+1$ a special form of M_{tN} can be considered so that the analysis can be reduced to that of the U(N) case, with the same superpotential. This is possible when the entries of M_{tN} are periodically identified like

$$\phi_i \equiv \phi_{i+N} \qquad\qquad y_i \equiv y_{i+N} \tag{3.62}$$

so that the only non zero entries of M_{tN} are

$$(M)_{i,i-1} = 1 \qquad\qquad (M)_{1,tN} = z \tag{3.63}$$

$$(M)_{i,i+1} = y_i \qquad\qquad (M)_{tN,1} = y_0 z^{-1} \tag{3.64}$$

$$(M)_{i,i} = \phi_i \tag{3.65}$$

The condition set by the energy scale is written as

$$\prod_{i=0}^{tN-1} y_i = \Lambda^{2tN} \geq 0. \tag{3.66}$$

Note in particular that the (physical) energy scale of the $U(tN)$ theory should be real. This condition can be satisfied by a family of t different complex valued Λ^N in the U(N) theory, $\{\Lambda^N, \eta^2\Lambda^N, \ldots, \eta^{2(t-1)}\Lambda^N\}$, with η a $2t$-th root of unity.

The polynomial $P_{tN}(x)$ in the spectral curve of M_{tN} is

$$P_{tN}(x) = \det(xI_{tN} - M_{tN}) + z + \Lambda^{2tN} z^{-1}. \tag{3.67}$$

The determinant can be calculated by considering a gauge equivalent matrix. By a gauge transformation M_{tN} can be brought into a form that is invariant under cyclic permutations of order N. We define

$$G = \mathrm{diag}(1, z^{1/tN}, z^{2/tN}, \ldots, z^{(tN-1)/tN}). \tag{3.68}$$

Then $\tilde{M}_{tN} = G M_{tN} G^{-1}$ has factors of $z^{1/t}$ and $z^{-1/t}$ democratically distributed over the $(\tilde{M})_{i,i-1}$ and $(\tilde{M})_{i,i+1}$ entries respectively. $\tilde{M}_{tN}$ satisfies

$$S \tilde{M}_{tN} S^{-1} = \tilde{M}_{tN}, \tag{3.69}$$

where

$$S = \begin{pmatrix} 0 & I_{(t-1)N} \\ I_N & 0 \end{pmatrix}.$$

is the matrix that generates a cyclic permutation of order N on tN elements, $S^{-1} = S^T$.

In case the degree of the superpotential is small enough, $\deg(W) \leq N+1$, the cyclic invariance, (3.69), ensures that the equations of motion collapse to those of the U(N) theory with the same superpotential. First, because in $\tilde{M}_{tN}$ the ϕ_i and y_i appear only linearly, derivatives with respect to ϕ_{i+kN} and y_{i+kN} can be replaced with derivatives with respect to ϕ_i and y_i

$$\begin{aligned}\partial_{\phi_{i+N}} \mathrm{tr}\left(\tilde{M}_{tN}^{n+1}\right) &= (n+1)\mathrm{tr}\left(\tilde{M}_{tN}^{n}\partial_{\phi_{i+N}}\tilde{M}_{tN}\right)\\ &= (n+1)\mathrm{tr}\left(S\tilde{M}_{tN}^{n}S^{-1}S(\partial_{\phi_{i+N}}\tilde{M}_{tN})S^{-1}\right)\\ &= (n+1)\mathrm{tr}\left(\tilde{M}_{tN}^{n}(\partial_{\phi_i}\tilde{M}_{tN})\right)\\ &= \partial_{\phi_i}\mathrm{tr}\left(\tilde{M}_{tN}^{n+1}\right).\end{aligned} \tag{3.70}$$

Second, in $\tilde{M}_{tN}$ the ϕ_i and y_i appear only on the diagonal and the super-diagonal respectively

$$(\tilde{M}_{tN})_{ij} = \phi_i\delta_{i,j} + y_i z^{-1/tN}\delta_{i,j-1} + z^{1/tN}\delta_{i,j+1}, \tag{3.71}$$

The diagonal elements of $\tilde{M}_{tN}^n$ each depend on at most n consecutive y_i's and ϕ_i's. Therefore, all equations of motion for ϕ_1 up to ϕ_{tN} can be both mapped to equations of motion for ϕ_1 to ϕ_N and these also consistently collapse onto the equations for the first N ϕ_i's. The same holds for the equations of motion for all y_i's. Also the equation of motion for the Lagrange multiplier, (3.66) maps to (3.61).

$\tilde{M}_{tN}$ can be explicitly written as a $t \times t$ matrix of which each entry is itself one of four $N \times N$ matrices

$$\tilde{M}_{tN} = \begin{pmatrix} A & D & 0 & \dots & \dots & 0 & E\\ E & A & D & 0 & \dots & \dots & 0\\ 0 & E & A & D & \dots & \dots & \dots\\ \dots & 0 & E & A & \dots & \dots & \dots\\ \dots & \dots & \dots & \dots & \dots & \dots & \dots\\ \dots & \dots & \dots & \dots & \dots & \dots & D\\ D & 0 & \dots & 0 & \dots & E & A \end{pmatrix}. \tag{3.72}$$

The $N \times N$ matrices A, D, and E are of the following form

$$A = \begin{pmatrix} \phi_1 & y_1 z^{-1/tN} & 0 & \dots & \dots & 0\\ z^{1/tN} & \phi_2 & y_2 z^{-1/tN} & 0 & \dots & 0\\ 0 & z^{1/tN} & \dots & \dots & \dots & \dots\\ \dots & 0 & \dots & \dots & \dots & \dots\\ \dots & \dots & \dots & \dots & \dots & \dots\\ \dots & \dots & \dots & \dots & \dots & y_{N-1}z^{-1/tN}\\ 0 & \dots & \dots & \dots & z^{1/tN} & \phi_N \end{pmatrix} \tag{3.73}$$

$$D = y_0 z^{-1/tN}\begin{pmatrix} 0_{1\times(N-1)} & 0_{(N-1)\times(N-1)}\\ 1 & 0_{(N-1)\times 1} \end{pmatrix} \tag{3.74}$$

$$E = z^{1/tN} \begin{pmatrix} 0_{1\times(N-1)} & 1 \\ 0_{(N-1)\times(N-1)} & 0_{(N-1)\times 1} \end{pmatrix}. \tag{3.75}$$

Because of (3.69), there exists a basis of simultaneous eigenvectors of $\tilde{M}_{tN}$ and S. The eigenvectors of S fall into t N-dimensional subspaces, each of which is labeled by a different t-th root of unity ω^r, $\omega^t = 1$. The eigenvalue equation for $\tilde{M}_{tN}$ is written in a basis of S eigenvectors $v_r^{(\alpha)} = (v^{(\alpha)}, \omega^r v^{(\alpha)}, \omega^{2r} v^{(\alpha)}, \ldots, \omega^{(t-1)r} v^{(\alpha)})$, with $\alpha = 1, 2, \ldots, N$.

For each r the eigenvalue equation $\tilde{M}_{tN} v_r^{(\alpha)} = \lambda_r^{(\alpha)} v_r^{(\alpha)}$ becomes

$$(A + \omega^r D + \omega^{-r} E) v_r^{(\alpha)} = \lambda_r^{(\alpha)} v_r^{(\alpha)}, \tag{3.76}$$

and therefore

$$\det(M_{tN}) = \det(\tilde{M}_{tN}) = \prod_{r=1}^{t} \det\left(A + \omega^r D + \omega^{-r} E\right). \tag{3.77}$$

The U(tN) polynomial reads

$$\begin{aligned} P_{tN}(x) &= \prod_{r=1}^{t} \det\left(x I_N - A - \omega^r D - \omega^{-r} E\right) + z + \Lambda^{2tN} z^{-1} \\ &= \prod_{r=1}^{t} \left(P_n(x) - \omega^{-r} z^{1/t} - \eta^2 \Lambda^{2N} \omega^r z^{-1/t}\right) + z + \Lambda^{2tN} z^{-1} \end{aligned} . \tag{3.78}$$

The polynomials $P(x)$ are by construction independent of z. A convenient choice to evaluate (3.78) is to take $z^{1/t} = e^{\frac{i\pi}{t}} |\Lambda^{\frac{1}{2tN}}|$. Recall that ω is a t-th root of unity such that ω^r hits all t different eigenvalues of the $\mathbb{Z}_t$ cyclic permutation matrix S. Hence $P_{tN}(x)$ can be written as

$$\begin{aligned} P_{tN}(x) &= \prod_{r=1}^{t} \left(P_N(x) - \eta \Lambda^N \left(e^{i\pi \frac{4r-1}{2t}} + e^{-i\pi \frac{4r-1}{2t}}\right)\right) \\ &= 2^t \eta^t \Lambda^{tN} \prod_{r=1}^{t} \left(\frac{1}{2\eta\Lambda^N} P_N(x) - \cos(\frac{4r-1}{2t}\pi)\right). \end{aligned} \tag{3.79}$$

The latter expression defines the n-th Chebyshev polynomial of the first kind, which is defined as

$$T_t(x) = 2^{n-1} \prod_{k=1}^{t} \left(x - \cos(\frac{2k-1}{2t}\pi)\right) \tag{3.80}$$

with

$$s := \max\{\sigma \in \mathbb{Z} : 2\sigma \leq t\}. \tag{3.81}$$

Equation (3.80) can also be written as

$$\begin{aligned} T_n(x) &= 2^{n-1}\prod_{p=1}^{s}\left(x-\cos(\frac{2(2p)-1}{2t}\pi)\right)\prod_{q=1}^{s}\left(x-\cos(\frac{2(2q-1)-1}{2t}\pi)\right) \\ &= 2^{n-1}\prod_{p=1}^{s}\left(x-\cos(\frac{4p-1}{2t}\pi)\right)\prod_{q=1}^{s}\left(x-\cos(\frac{4(t-q)+3}{2t}\pi)\right) \qquad , \\ &= 2^{n-1}\prod_{r=1}^{t}\left(x-\cos(\frac{4r-1}{2t})\right) \end{aligned} \tag{3.82}$$

so

$$P_{tN} = 2\Lambda^{tN}\eta^t T_t\left(\frac{P_N(x)}{2\eta\Lambda^N}\right). \tag{3.83}$$

Thus the periodic Ansatz for the $U(tN)$ theory yields t times the number of vacua found in the $U(N)$ theory. The polynomial P_{tN} in (3.83) agrees precisely with what had been found in field theory in [23].

3.2.6 High order potentials

As noted in the beginning of this chapter, the prescription of replacing the adjoint scalar in the classical superpotential by the Lax matrix is ambiguous for high order superpotentials. The resolution is to use the fact that given a Lax matrix M_N of size $N\times N$, there is a corresponding Lax matrix M_{tN} of size $tN\times tN$ which is constructed using the replica procedure given above. For small m, $\mathrm{Tr}(M_N^m) = t^{-1}\mathrm{Tr}(M_{tN}^m)$, but for larger values of m the right hand side starts to depend on the spectral parameter, whereas the left hand side does not. We therefore propose

$$\langle\Phi^m\rangle = \lim_{t\to\infty}\frac{1}{t}\mathrm{Tr}(M_{tN}^m) \tag{3.84}$$

as the right definition for arbitrary m. Note that the replica procedure yields

$$P_{tN}(x) = 2\Lambda^{tN}T_t(\frac{P_N(x)}{2\Lambda^N}). \tag{3.85}$$

With this expression for $P_{tN}(x)$ and the identities for Chebyshev polynomials

$$\begin{aligned} T_n'(x) &= nU_{n-1}(x) \\ T_n^2(x)-1 &= (x^2-1)U_{n-1}^2(x), \end{aligned} \tag{3.86}$$

it is easy to see that this proposition yields a resolvent that agrees with the field theory result of [48],

$$\langle\mathrm{tr}\frac{1}{x-\Phi}\rangle = \frac{P'(x)}{\sqrt{P^2(x)-4\Lambda^{2N}}}. \tag{3.87}$$

Therefore, M_{tN} for large t provides a master field for Φ. For the simple case of a quadratic superpotential (see section 3.2.4) M takes a simple form, and M_{tN} for infinite t can be written in a simple way in terms of Cuntz variables. In this way it also provides a master field for the Gaussian matrix model, as observed in [49], but this appears to be a coincidence that happens only for quadratic superpotentials. Whether there is a direct way other than, as will be discussed in equation (3.118), to extract the matrix model resolvent remains to be seen.

3.2.7 Gauge group $U(4)$

We are now in a position to use the techniques of the preceding section to study our final example $U(4)$, again with a cubic superpotential

$$W = g_1 \mathrm{Tr}\,(\Phi) + \frac{g_2}{2}\mathrm{Tr}\,(\Phi^2) + \frac{g_3}{3}\mathrm{Tr}\,(\Phi^3). \tag{3.88}$$

The Lax operator, polynomial $P(x)$, superpotential W_{eff} and equations of motion are straightforward generalizations of (3.23), (3.24), (3.25) and (3.26), with Λ^6 replaced by Λ^8. Instead of giving a lengthy and tedious analysis of the possible solutions of the equations of motion, we will simply present a solution for each of the critical points that we expect, based on the knowledge of the possible solutions in four dimensions. These are given explicitly in section 3.3 of [50].

To describe the qualitative form of the solutions, we write $W'(x) = (x - a_1)(x - a_2)$. Classically, each of the ϕ_i is equal to either a_1 or a_2. We denote the number of ϕ_i for which $\phi_i = a_j$ by N_j, so that $N_1 + N_2 = 4$. However, this does not fully specify the different solutions. For $U(4)$ with a cubic superpotential, we know that at a minimum $P(x)^2 - 4\Lambda^8$ has two double zeroes. These double zeroes are distributed over $P(x) - 2\Lambda^4$ and $P(x) + 2\Lambda^4$, since these two factors cannot have a common zero. We denote by $r_\pm$ the number of double zeroes in $P(x) \pm 2\Lambda^4$, so that $r_+ + r_- = 2$.

When all ϕ_i are equal, with $(N_1, N_2) = (4, 0)$ or $(0, 4)$, we have a maximally confining vacuum, and these were already described in detail in (3.2.4).

When $(N_1, N_2) = (2, 2)$, we can use the results in (3.2.5) to find solutions by lifting solutions in $U(2)$ to $U(4)$. One easily finds that these solutions have $(r_+, r_-) = (2, 0)$ or $(0, 2)$. In addition, they necessarily have $(\phi_1, \phi_2, \phi_3, \phi_4) = (a_1, a_2, a_1, a_2)$ or (a_2, a_1, a_2, a_1).

However, this does not exhaust all solutions with $(N_1, N_2) = (2, 2)$. There are also solutions that have $(r_+, r_-) = (1, 1)$. In addition, we have not yet considered solutions with $(N_1, N_2) = (1, 3)$ or $(3, 1)$. It turns out that all missing solutions are part of one family, that can be described as follows. First, by an overall rescaling and by shifting x by a constant we can always choose $W(x)$ so that

$$W'(x) = x^2 + x\frac{\Lambda^4}{a^3} - a^2, \tag{3.89}$$

for some parameter a. The solution is then

$$\begin{array}{rclcrcl}
\phi_1 &=& \frac{a^2}{\phi_3} & & y_0 &=& a^2\frac{\phi_3+\phi_4}{\phi_3} \\
\phi_2 &=& \frac{a^2}{\phi_4} & & y_1 &=& a^2 + \frac{a^4}{\phi_3\phi_4} \\
\phi_3 &=& \phi_3 & & y_2 &=& a^2\frac{\phi_3+\phi_4}{\phi_4} \\
\phi_4 &=& \phi_4 & & y_3 &=& a^2 + \phi_3\phi_4
\end{array}$$

$$S = \frac{\Lambda^4}{a}. \tag{3.90}$$

with ϕ_3, ϕ_4 subject to the constraint

$$(\phi_3+\phi_4)(a^5 + a^3\phi_3\phi_4) + \phi_3\phi_4\Lambda^4 = 0. \tag{3.91}$$

As explained in [50], there are two classical limits, one is $\Lambda \to 0$ while keeping a fixed, the other is $\Lambda \to 0$ while keeping Λ^4/a^3 fixed. The first corresponds to $(N_1, N_2) = (2,2)$ and $(r_+, r_-) = (1,1)$, the second to $(N_1, N_2) = (1,3)$ or $(3,1)$.

We have accounted for all vacua that we expect in four dimensions. One interesting feature of the solutions is that in the case with $(N_1, N_2) = (2,2)$, all classical limits have $\phi_1 = \phi_3$ and $\phi_2 = \phi_4$, but there is no solution whose classical limit obeys $\phi_1 = \phi_2$ and $\phi_3 = \phi_4$. It therefore appears that one should be careful in choosing the right ordering of the eigenvalues, not all orderings will give rise to a solution of the quantum equations of motion. This is not a contradiction, since the choice of Lax matrix breaks the S_N symmetry to Z_N, and there is no symmetry that arbitrarily permutes the eigenvalues.

The solution in (3.90) has one free parameter, which corresponds to the extra $U(1)$ that appears when breaking $U(4) \to U(2) \times U(2)$.

3.3 The semi-classical expansion

In the previous section we presented several examples of superpotentials and their extrema. The solutions depend in a non-trivial way on Λ, and it is important to understand the nature of the semi-classical expansion, certainly if we want to compare our results to direct field theory calculations in three dimensions. A precise understanding of the semi-classical expansion is probably also important in order to understand the relation to the 4d description in terms of gluino bilinear superfields, as we will discuss in section 2.2.2.

What is the general structure of the semi-classical expansion that one would expect to find? In four dimensions in a situation where the the gauge group is classically broken to $\prod_i U(N_i)$ the effective superpotential explicitly takes the form of a semi-classical expansion. In addition to the Veneziano-Yankielowicz superpotentials (2.55) for each unbroken $U(N_i)$, there are many additional terms coming from the planar diagrams of the matrix model. These are polynomial and give ultimately rise to an expansion in positive but possibly fractional powers of Λ. The low-energy scales Λ_i of the unbroken $U(N_i)$ that appear in the Veneziano-Yankielowicz superpotential (2.55) are obtained by scale matching and given by

$$\Lambda_i^{3N_i} = g_n^{N_i} \Lambda^{2N} \prod_{j \neq i} (a_j - a_i)^{N_i - 2N_j}, \tag{3.92}$$

with a_i and g_n defined in (2.41).

On $\mathbb{R}^3 \times S^1$, the Veneziano-Yankielowicz superpotential is no longer appropriate, the relevant superpotential is instead given by (2.78) or equivalently (2.82). Therefore, we expect that the superpotential on $\mathbb{R}^3 \times S^1$ should admit an expansion of the form

$$W = \sum_i \sum_{j=0}^{N_i - 1} y_j^{(i)} + \text{higher order terms}, \tag{3.93}$$

where the $y_j^{(i)}$ are analogues of the variables y_i for each of the classically unbroken $U(N_i)$ groups, and they should therefore obey

$$\prod_{j=0}^{N_i - 1} y_j^{(i)} = \Lambda_i^{3N_i} \tag{3.94}$$

for each i. The superpotential in the form (3.93) only depends on variables y_i, not on the eigenvalues ϕ_i of Φ. In addition, it depends on a choice of classical vacuum, whereas the superpotential (2.88) with the ϕ_i included describes all vacua. Therefore, a description like (3.92) can only emerge after we integrate out the ϕ_i.

To illustrate how an expansion like (3.93) can emerge, we consider the case $U(4) \to U(2) \times U(2)$, with superpotential $W(x) = x^3/3 - a^2 x$, so that $W'(x) = (x-a)(x+a)$. As a first step, we integrate out the ϕ's in the quantum superpotential. This can be done explicitly, the solutions for the ϕ_i read

$$\phi_1 = \sqrt{a^2 - y_0 - y_1} \qquad \phi_3 = \sqrt{a^2 - y_2 - y_3}$$
$$\phi_2 = -\sqrt{a^2 - y_1 - y_2} \qquad \phi_4 = -\sqrt{a^2 - y_0 - y_3}.$$

Notice that we chose the signs of the square roots in such a way that in the classical limit $y_i \to 0$ we indeed end up in a vacuum where $U(4)$ is broken to $U(2) \times U(2)$.

Therefore integrating out the ϕ_i also involves the selection of a classical vacuum configuration. If we insert the values for ϕ_i in the superpotential, and expand the result to second order in y, we get

$$W_{\text{eff}} = S\log(\frac{\Lambda_4^8}{y_0y_1y_2y_3}) - \frac{1}{2a}(y_0-y_2)(y_1-y_3) + \dots$$

If we next integrate out S, this becomes

$$W_{\text{eff}} = -\frac{1}{2a}\left(y_0y_1 + \frac{\Lambda^8}{y_0y_1} - y_1y_2 - \frac{\Lambda^8}{y_1y_2}\right). \tag{3.95}$$

Interestingly, this depends on only two independent variables, namely y_0y_1 and y_1y_2, and therefore it is already of the form (3.93). Indeed, if we define

$$y_0^{(1)} = -\frac{y_0y_1}{2a}, \qquad y_1^{(1)} = -\frac{\Lambda^8}{2ay_0y_1}, \qquad y_0^{(2)} = \frac{y_1y_2}{2a}, \qquad y_1^{(2)} = \frac{\Lambda^8}{2ay_1y_2}, \tag{3.96}$$

then (3.95) is of the form (3.93), and $y_0^{(1)}y_1^{(1)} = \Lambda_1^6 = \Lambda^8/(2a)^2$, $y_0^{(2)}y_1^{(2)} = \Lambda_2^6 = \Lambda^8/(2a)^2$, completely in agreement with (3.94) and (3.92).

Of course, the above result is simply the semi-classical result at leading order. It would be very interesting to go beyond the leading order, and to understand in detail how the expansion is organized. We have not studied this in detail, but expect the following. In general the value of the superpotential is invariant under the flows (1.10) of the integrable system. Some of the flows are stationary at the extremum, but some are not, and that is why there is a $k-1$ complex parameter family of minima (see also the discussion in section 3.1.) If we could redefine our y variables in such a way that $N-k$ of them are independent of the $k-1$ flows that do not degenerate, then the superpotential should be a non-trivial function of these $N-k$ variables only. This is exactly the number of independent variables that appears in the semi-classical expansion (3.93). Hence we expect that the semi-classical expansion appears naturally by integrating out the variables ϕ_i and L, and by subsequently redefining the complex variables y_i in a suitable way, exactly as in our example above.

To conclude this section, we illustrate in the case of $U(4)$ with a cubic potential $W(x) = x^3/3 - a^2x$ how the solutions found in 3.2.7 can be expanded in (possibly fractional) powers of Λ.

For the maximally confining case, with $(N_1, N_2) = (4, 0)$, we have

$$\begin{aligned} \phi_i &= a - \omega\frac{\Lambda^2}{a} + \dots \\ y_i &= \omega\Lambda^2 + \dots \\ S &= 2a\omega\Lambda^2 + \dots \end{aligned} \tag{3.97}$$

with ω a fourth root of unity.

For $(N_1, N_2) = (2, 2)$ with $(r_+, r_-) = (2, 0)$ we find

$$\begin{aligned}
\phi_1 = \phi_3 = -\phi_2 = -\phi_4 &= a - (\xi + \frac{1}{\xi})\frac{\Lambda^2}{2a} + \dots \\
y_0 = y_2 &= \xi\Lambda^2 + \dots \\
y_1 = y_3 &= \frac{1}{\xi}\Lambda^2 + \dots \\
S &= 0.
\end{aligned} \tag{3.98}$$

Here, and in the solutions below, ξ indicates the free parameter that is related to the additional unbroken $U(1)$ that one gets in the corresponding vacuum solution.

For $(N_1, N_2) = (2, 2)$ with $(r_+, r_-) = (1, 1)$ the expansion reads

$$\begin{aligned}
\phi_1 &= a - (\xi - \tfrac{1}{\xi})\tfrac{\Lambda^2}{2a} + \dots & y_0 &= \xi\Lambda^2 + \dots \\
\phi_2 &= -a - (\xi + \tfrac{1}{\xi})\tfrac{\Lambda^2}{2a} + \dots & y_1 &= -\tfrac{1}{\xi}\Lambda^2 + \dots \\
\phi_3 &= a + (\xi - \tfrac{1}{\xi})\tfrac{\Lambda^2}{2a} + \dots & y_2 &= -\xi\Lambda^2 + \dots \\
\phi_4 &= -a + (\xi + \tfrac{1}{\xi})\tfrac{\Lambda^2}{2a} + \dots & y_3 &= \tfrac{1}{\xi}\Lambda^2 + \dots \\
& & S &= \frac{\Lambda^4}{a} + \dots .
\end{aligned} \tag{3.99}$$

This result differs considerably from (3.98), illustrating the fact that these are two different solutions.

Finally, for $(N_1, N_2) = (3, 1)$ the result is an expansion in

$$\epsilon = \left(\frac{\Lambda^4}{2a}\right)^{1/3}, \tag{3.100}$$

which reads

$$\begin{aligned}
\phi_1 &= a - \tfrac{\epsilon}{\xi} + \dots & y_0 &= \tfrac{2a\epsilon}{\xi} + \dots \\
\phi_2 &= a - \tfrac{\epsilon^2}{a} + \dots & y_1 &= \epsilon^2 + \dots \\
\phi_3 &= a - \epsilon\xi + \dots & y_2 &= \epsilon^2 + \dots \\
\phi_4 &= -a + \epsilon(\xi + \tfrac{1}{\xi}) + \dots & y_3 &= 2a\epsilon\xi + \dots \\
& & S &= 2a\epsilon^2 + \dots .
\end{aligned} \tag{3.101}$$

3.4 Interpretation of glueball fields

Using the identification of the Lagrange multiplier in the term

$$S \log \left(\frac{\Lambda^{2N}}{\prod_{i=0}^{N-1} y_i^2} \right) \tag{3.102}$$

as the sum of S_i, we can find the superpotential as a function of this field and the couplings by integrating out all p_i and y_i from a given superpotential. Below it is shown how this reproduces known results obtained through matrix models.

3.4.1 Gauge group $U(2)$

For convenience we specialize to the superpotential

$$W(x) = \frac{x^3}{3} - a^2 x, \tag{3.103}$$

which has extrema at $x = \pm a$. Since we are going to study the maximally confining case there is only one chiral superfield S, which can be recovered from the superpotential

$$W_{\text{eff}} = W(\phi_1) + W(\phi_2) + (y_0 + y_1)(\phi_1 + \phi_2) + S \log \left(\frac{\Lambda^4}{y_0 y_1} \right)$$

by integrating out the ϕ's and the y's degrees of freedom. This means we have to solve for the extrema of the superpotential in terms of the ϕ's and y's, so we can simply use the results of section 3.2.1[2]. The only thing we have to do is to choose the vacuum. In the classical limit ϕ is $\pm a$, we pick $\phi_{\text{clas}} = a$. To integrate out the ϕ's we pick the corresponding solution from section 3.2.1 (with $g_3 = 1, g_2 = -a^2, g_1 = 0$)

$$\begin{aligned} y &= y_0 = y_1 \\ \phi &= \phi_1 = \phi_2 = \sqrt{a^2 - 2y}. \end{aligned} \tag{3.104}$$

Plugging this into the superpotential leaves us with

$$W_{\text{eff}} = -\frac{4}{3}(a^2 - 2y)^{\frac{3}{2}} + S \log \left(\frac{\Lambda^4}{y^2} \right). \tag{3.105}$$

The next step is to integrate out the y's

$$\frac{\partial W}{\partial y} = 0 \Rightarrow S^2 = 4y^2(a^2 - 2y). \tag{3.106}$$

[2]Since we do not want to integrate out S we have to be careful *not* to use the equation of motion for it: $y_0 y_1 = \Lambda^4$.

To solve (3.106) we write

$$\xi = \frac{S}{4a^3} \text{ and } y(S) = A(\xi)\frac{S}{2a}.$$

Then (3.106) can be written as

$$A^2(\xi) - 4\xi A^3(\xi) = 1.$$

In principle there are three solutions for $A(\xi)$, however not all solutions have the right classical limit ($\Lambda \to 0$). From the solution presented in section 3.2.1 we learn that in the classical limit we need to have $y = \frac{S}{2a}$, therefore:

$$y = \frac{S}{2a} + \mathcal{O}(S^2), \text{ hence } A(0) = 1.$$

The solution for $A(\xi)$ (with $A(0) = 1$) is (see for example [51])

$$A(\xi) = \sum_{n=0}^{\infty} \frac{2^{2n}}{n+1} \binom{\frac{3n-1}{2}}{n} \xi^n = \frac{1}{12\xi} + \frac{1}{6\xi} \sin\left(\arcsin\left(\frac{216\xi^2 - 1}{3}\right)\right)$$

yielding the following expression for y

$$\begin{aligned} y(S) &= \frac{S}{2a} \sum_{n=0}^{\infty} \frac{2^{2n}}{n+1} \binom{\frac{3n-1}{2}}{n} \left(\frac{S}{4a^3}\right)^n = -\frac{a^2}{6}\left(-1 + 2\sin\left(\frac{1}{3}\arcsin\left(1 - \frac{27S^2}{2a^6}\right)\right)\right) \\ &= \frac{S}{2a} + \frac{S^2}{4a^4} + \frac{5S^3}{16a^7} + \mathcal{O}(S^4). \end{aligned}$$

Substituting this solution into the effective superpotential gives us, in principle, a closed expression valid to all orders in S. However the form of this expression is not particularly illuminating and therefore we expand the superpotential in S

$$W_{\text{eff}} = -\frac{4a^3}{3} + 2S\left(1 - \log\left(\frac{S}{\Lambda^2 m}\right)\right) - \left(\frac{S^2}{2a^3} + \frac{S^3}{3a^6} + \frac{35S^4}{96a^9} + \dots\right), \quad (3.107)$$

here $m = 2a$ is the mass of the fluctuations of ϕ around the classical extremum $\phi = a$.

Let us compare this with the four dimensional answer [19], in that case we would write (for gauge group $U(N)$)

$$W_{\text{eff}} = NS\left(1 - \log\left(\frac{S}{\Lambda^3}\right)\right) - NS\log\left(\frac{\Lambda}{m}\right) - N\frac{\partial \mathcal{F}}{\partial S}.$$

For a superpotential $W = \frac{g}{3}\Phi^3 + \frac{m}{2}\Phi^2$ (i.e. $m = 2a, g = 1$) the function $\mathcal{F}$ is given by

$$\mathcal{F} = \frac{2}{3}\frac{g^2}{m^3}S^3 + \frac{8}{3}\frac{g^3}{m^6}S^4 + \frac{56}{3}\frac{g^4}{m^9}S^5 + \dots \Rightarrow 2\frac{\partial \mathcal{F}}{\partial S} = \frac{S^2}{2a^3} + \frac{S^3}{3a^6} + \frac{35S^4}{96a^9} + \dots$$

So we see that equation (3.107) is in good agreement with the four-dimensional answer.

3.4.2 Going from $\mathbb{R}^4$ to $\mathbb{R}^3 \times S^1$

Once we identify the Lagrange multiplier with the glueball field $S = \sum_i S_i$, there is a concrete procedure to find the superpotential as a function of S from (2.88). One may wonder whether one can also go back and start with the result on $\mathbb{R}^4$ and construct the superpotential on $\mathbb{R}^3 \times S^1$. We don't know whether this can be done in general, but a step in this direction is to show how one can obtain (2.78) from (2.55). The procedure is very similar to the path integral derivation of 2d mirror symmetry given in [52]. Starting with the Veneziano-Yankielowicz superpotential

$$W_{VY}(S) = S\left[\log\left(\frac{\Lambda^{3N}}{S^N}\right) + N\right], \tag{3.108}$$

we first write it in a form as if it were built out of $U(1)$ pieces rather than $U(N)$,

$$W_{VY} \sim \sum_{t=1}^{N} S_t\left[\log\left(\frac{\Lambda^3}{S_t}\right) + 1\right] + \sum_{t=1}^{N-1} Z_t(S_t - S_{t+1}). \tag{3.109}$$

If the variables Z_t are integrated out, all S_t are identified, and we are back at the form (3.108) of the superpotential. However, we proceed by integrating out the S_t from (3.109) instead. This yields

$$W \to \Lambda^3(e^{Z_1} + e^{Z_2 - Z_1} + \ldots + e^{-Z_{N-1}}), \tag{3.110}$$

which is indeed precisely of the form (3.109). It would be very interesting to understand whether and how these transformations can be generalized, perhaps in the spirit of 2d mirror symmetry, to more complicated situations.

3.4.3 Interpretation of the individual S_i

A full reconstruction of the superpotential (2.71) from (2.88) requires us to not only find the right interpretation of $S = \sum S_i$, but also of the individual S_i. To do this the nature of the semi-classical expansion discussed in section 5 is probably crucial. If we could write the superpotential in the form (3.93) with constraints (3.94), we could try to impose the constraints (3.94) using various Lagrange multiplier fields L_i similar to those used in (2.82), and it would then be natural to identify those with the S_i. To lowest order, this would simply boil down to integrating in the S_i in each individual gauge group and therefore correctly reproduce the Veneziano-Yankielowicz superpotentials (2.55). However, this program does not seem to work for the simplest guesses one can make.

3.5 Factorization equations in terms of commutation relations

The Lax matrix of the periodic Toda chain played an essential role in the construction of the effective superpotential in three dimensions. One of the motivations of this work was to try to find a direct relation between the periodic Toda chain and the Dijkgraaf-Vafa matrix model. We have not yet succeeded in finding a direct relation, but nevertheless we have found in examples that the supersymmetric vacua in three dimensions have a very nice interpretation in terms of the Toda integrable system. In this section it will be proven that the equilibria of the proposed class of $3d$ potentials indeed reproduces the expected four dimensional gauge theory vacua.

The equations of motion derived from (2.88) imply that the flow generated by $W(M)$ degenerates. This translates to the equations

$$[M, W'(M)_+ - W'(M)_-] = 0, \qquad W'(M)_0 = 0. \tag{3.111}$$

The second equation $W'(M)_0 = 0$ appears because $\mathrm{Tr}(M)$ does not generate a flow, and has to be treated separately. It is absent for $SU(N)$ gauge theories. In (3.111) W' is the function obtained by differentiating $W(x)$ for some complex number x. A second observation, which follows from (1.11), is that

$$[M, P_N(M)_+ - P_N(M)_-] = 0. \tag{3.112}$$

The two equations (3.111), (3.112) define a reduction of the integrable Toda system. We have the following theorem

Theorem 1 *Equations 3.111 imply the existence of polynomials H_{N-k}, T_{2k}, G_{n-k}, f_{n-1}, U_k of degrees $N-k, 2k, n-k, n-1, k$, $1 \leq k \leq n$ respectively such that*

$$\begin{aligned}
(W'(M)_+ - W'(M)_-)^2 &= W'(M)^2 + f_{n-1}(M) && (3.113)\\
W'(M)_+ - W'(M)_- &= G_{n-k}(M)(U_k(M)_+ - U_k(M)_-) && (3.114)\\
(P_N(M)_+ - P_N(M)_-)^2 &= P_N(M)^2 - 4\Lambda^{2N} && (3.115)\\
P_N(M)_+ - P_N(M)_- &= H_{N-k}(M)(U_k(M)_+ - U_k(M)_-) && (3.116)\\
(U_k(M)_+ - U_k(M)_-)^2 &= T_{2k}(M). && (3.117)
\end{aligned}$$

This system of equations demonstrates not only that the extrema of the superpotential (2.88) are in exact one-to-one correspondence with the four-dimensional results (2.46). It also shows that we can construct those four-dimensional equations directly in terms of the Lax matrix. In particular, the matrix model resolvent [24] $2R(z) = -\sqrt{W'(z)^2 + f_{n-1}(z)} + W'(z)$ satisfies the very simple equation

$$R(M) = W'(M)_-. \tag{3.118}$$

Actually, equation (3.111) is tantalizingly close to a similar equation that can be derived for matrix integrals of the form

$$\int d\Phi e^{-W(\Phi)}. \tag{3.119}$$

If one defines a set of orthogonal polynomials with respect to the measure $e^{-W(x)}$, of the form $p_n = x^n + \ldots$, one can construct a semi-infinite matrix Q that acts on p_n as multiplication by x. This semi-infinite matrix is tridiagonal (i.e. $Q_{ij} = 0$ for $|i-j| > 1$), and obeys the "string equation" (see for example [53] and references therein)

$$[Q, -\frac{1}{2}(W'(Q)_+ - W'(Q)_-)] = 1. \tag{3.120}$$

The only difference between (3.111) and (3.120) is that one is replaced by zero. Equation (3.120) is also the equation that gives rise to the string equation in the matrix model description of minimal models coupled to gravity. Though (3.111) and (3.120) are very similar, we have not found a direct map between Q and M. Whereas Q is relevant for the orthogonal polynomials defined with respect to $e^{-W(x)}$, M seems to define orthogonal polynomials for a measure which coincides with the gauge theory resolvent instead. Various other relations between matrix integrals and the Toda lattice equations are discussed in e.g. [54, 55].

3.6 Algebraic proof

In this section we will provide an algebraic proof of the above theorem. The first step of the proof is to show that in addition to (3.111) we also have

$$[M, P_N(M)_+ - P_N(M)_-] = 0, \tag{3.121}$$

or equivalently

$$[\tilde{M}, P_N(\tilde{M})_+ - P_N(\tilde{M})_-] = 0. \tag{3.122}$$

This follows from (1.11). Indeed, (1.11) implies the matrix equation

$$P_N(M) + (-1)^N(z + \Lambda^{2N} z^{-1}) = 0, \tag{3.123}$$

which when expressed in terms of the infinite matrix $\tilde{M}$ implies

$$P_N(\tilde{M}) + (-1)^N(Z + \Lambda^{2N} Z^{-1}) = 0. \tag{3.124}$$

Therefore, $P_N(\tilde{M})_+ = (-1)^{N+1} Z$ and $P_N(\tilde{M})_- = (-1)^{N+1}\Lambda^{2N} Z^{-1}$, and (3.112) follows trivially from $[\tilde{M}, Z] = 0$.

The main idea of the proof is to show that one can compute a "greatest common divisor" of $P_N(M)_+ - P_N(M)_-$ and $W'(M)_+ - W'(M)_-$. This computation follows a version of the Euclid algorithm that one uses compute the greatest common divisor of two integers. The crucial step is to show that if $[M, A(M)_+ - A(M)_-] = 0$, then for all polynomials B there is a polynomial C such that $C(M)_+ - C(M)_- = B(M)(A(M)_+ - A(M)_-)$ and in addition $[M, C(M)_+ - C(M)_-] = 0$. We will show this below, working exclusively with the infinite matrix representation of M (and dropping the tilde for convenience).

3.6.1 Some definitions

Let $gl(\infty)_N$ denote the elements in $gl(\infty)$ which are periodic with period N. X is said to be a matrix operator of degree k if

$$X \in gl(\infty)_N, X_{ij} \neq 0 \quad \Longrightarrow \quad i + k = j. \tag{3.125}$$

Any matrix can be written as a sum of matrices of fixed degree, $X = \sum_r X^{(r)}$, with $X^{(r)}$ of degree r. If there is a smallest r for which $X^{(r)} \neq 0$, we say that X is of minimal degree r, and similarly if there is a largest r for which $X^{(r)} \neq 0$ we say X is of maximal degree r.

We will only study operators with a finite maximal degree, which form a ring. By the inverse of a matrix we will mean an inverse in this ring. If this exists, it is defined as follows: write a general matrix R of maximal degree r as $R = R^{(r)} + \tilde{R}$, where $\tilde{R} = \sum_{j<r} R^{(j)}$. R is invertible in the ring if and only if $R^{(r)}$ is invertible, in which case it is defined as

$$R^{-1} = (R^{(r)})^{-1}(1 + \tilde{R}(R^{(r)})^{-1})^{-1} \equiv (R^{(r)})^{-1} \sum_{k\geq 0} (-1)^k (\tilde{R}(R^{(r)})^{-1})^k, \tag{3.126}$$

so that R^{-1} is of maximal degree $-r$.

3.6.2 Some facts

First we will collect a number of intermediate results which will be used to prove the theorem, beginning with the following

Lemma 1 *Let X be a matrix operator of degree $p \geq 0$.*

$$[X, M^+] = 0 \quad \text{iff} \quad X = \lambda (M^+)^p \tag{3.127}$$

for some constant $\lambda \in \mathbb{C}$

The proof of the first statement follows trivially from the second. Clearly, $X = \lambda(M^+)^p$ constitutes a one parameter family of solutions of $[X, M^+] = 0$. We will now show that this equation only has a one parameter family of solutions thereby proving the lemma. Write out the equation in components,

$$[X, M^+]_{ij} = (X_{i,i+p}M^+_{i+p,i+p+1} - M^+_{i,i+1}X_{i+1,i+p+1})\delta_{i+1+p,j} = 0. \tag{3.128}$$

Recall that none of the off-diagonal elements of M can be zero, since $\prod y_i^2 = \Lambda^{2N} \neq 0$. Therefore the above equation recursively determines all entries of X in terms of one arbitrary given entry, and therefore there is a one complex parameter family of solutions.

The above lemma is used to construct a proof for

Lemma 2 *Let $X \in gl(\infty)_N$ have maximal degree k. Then $[X, M] = 0$ implies*

$$X = \sum_{-\infty \leq i \leq k} \lambda_i M^i. \tag{3.129}$$

As before, we denote by $X^{(p)}$ the component of X of degree p. X can be trivially expanded as

$$X = \sum_{i \leq k} X^{(i)}. \tag{3.130}$$

Since M is of maximal degree one, $[X, M]$ is of maximal degree $k + 1$, and its component of highest degree equals

$$[X, M]^{(k+1)} = [X^{(k)}, M^{(1)}] = 0. \tag{3.131}$$

Using the previous lemma we have $X^{(k)} = \lambda_k M_+^k$. Now consider the operator $X' = X - \lambda_k M^k$ of maximal degree $k - 1$. This operator also commutes with M. By applying the same procedure we find λ_{k-1} such that $X'' = X' - \lambda_{k-1}M^{k-1}$ is of maximal degree $k - 2$. By continuing the procedure the lemma follows.

By itself the lemma is not that useful yet: we would like to obtain a finite (polynomial) series. Note that the natural operator with a finite series is the upper triangular part of a matrix of finite maximal degree since this matrix can be obtained by the above procedure in a finite number of steps,

$$X_+ = \left(\sum_{i=1}^{k} \lambda_i M^i\right)_+. \tag{3.132}$$

This observation leads to the following easily proven

Lemma 3 *Let $X \in gl(\infty)_N$ have maximal degree k and $[X, M] = 0$.*

- *X is symmetric then $X = A(M)$*
- *X is antisymmetric then $X = A(M)_+ - A(M)_-$*

with $A(M)$ some polynomial function in M of order k

Finally we prove the lemma

Lemma 4 *Let A be some some polynomial function in M of order k and $[M, A(M)_+ - A(M)_-] = 0$. Then $A(M)_0 = \lambda\mathbb{I}$ and*

$$(A(M)_+ - A(M)_-)^2 = A(M)^2 + f_{k-1}(M) + 2\lambda A(M), \tag{3.133}$$

with f_{k-1} a polynomial function of order $k-1$, $\lambda \in \mathbb{C}$.

Adding $A(M)$ to $A(M)_+ - A(M)_-$ yields $[M_+, A(M)_0] = 0$. By lemma 1 $A(M)_0 = \lambda\mathbb{I}$. Now consider $(A(M)_+ - A(M)_-)^2$. Since this is a symmetric operator, it is by lemma 3 equal to a polynomial, say $U(M)$. Then

$$U(M) - A(M)^2 = -2\lambda A(M) - \lambda^2 - 4(A_+(M)A_-(M)). \tag{3.134}$$

This proves the lemma, since the last term is symmetric and of maximal degree $k-1$.

3.6.3 Proving the theorem

With these auxiliary results the stage is set for proving the assertions in theorem 1. The fifth assertion is a consequence of lemma 3. The first is a simple application of lemma 4 using the second condition in (3.111). This leaves (3.114), (3.116). We will find the greatest common divisor indicated in these equation through the following algorithm, which is a variant of the algorithm of Euclid:

1. There is a polynomial function $A(M)$ in M of maximum degree $N - n$ such that
$$(P_N(M)_+ - P_N(M)_-)^{(\pm N)} = A(M)^{(\pm(N-n))}(W'(M)_+ - W'(M)_-)^{(\pm n)}. \tag{3.135}$$
2. Define
$$\tilde{P}(M)_+ - \tilde{P}(M)_- \equiv (P_N(M)_+ - P_N(M)_-) - A(M)(W'(M)_+ - W'(M)_-). \tag{3.136}$$
Note that this operator commutes with M and is of maximum degree $\leq N-1$. The existence of the polynomial $\tilde{P}$ is guaranteed by lemma 3.
3. Repeat the above steps with the operators $(W'(M)_+ - W'(M)_-)$ and $\tilde{P}(M)_+ - \tilde{P}(M)_-$.

4. At some point the subtraction step will give an operator of degree zero which commutes with M and is thus proportional to the identity matrix. The other operator is now by construction the greatest common divisor.

Some remarks are in order. First of all, the algorithm stops after a finite number of steps. Furthermore, the maximum degree of the gcd is of course n which is obtained if the algorithm terminates after one step. Denoting the maximum degree of the gcd by k we see we have proven (3.114) and (3.116) thereby finishing the proof of theorem 1.

We have now shown explicitly that the superpotential constructed through (1) indeed reproduces correctly the factorization of the Seiberg-Witten curve for softly broken $\mathcal{N} = 2 \to \mathcal{N} = 1$ super Yang-Mills theory. This is strong evidence that they give a correct description of the F-terms of the low-energy effective field theory.

3.7 Algebraic-geometrical proof

There is a long literature relating solutions of the equations of motion of the periodic Toda chain to algebraic geometric quantities, see e.g. [56] [57] [58]. The basic result that allows one to do so is that the commuting flows of the Toda system correspond to linear flows on the Jacobian of the associated spectral curve (1.11). The action variables correspond to moduli of the spectral curve, and are left invariant by the commuting flows. The angle variables are literally coordinates on the Jacobian of each of the surfaces. Therefore, the main problem is to understand the relation between the coordinates on the Jacobian and the variables that appear in the Lax matrix. This relation involves a set of g points (more precisely, a divisor of degree g) on the spectral curve (of genus g). On the one hand, given g points P_i, one other point P_0, and a basis of the holomorphic one-forms ω_i, the map

$$\{P_i\} \to \sum_i \int_{P_0}^{P_i} \omega_j \tag{3.137}$$

maps the g points to a point on the Jacobian, and every point on the Jacobian can be written in this way via the Jacobi inversion theorem. Therefore, in order to understand the relation between the Lax matrix and the Jacobian, we need to extract g points on the spectral curve from the Lax matrix. One way to do this is to consider eigenvectors of the Lax matrix (1.10). Viewed as functions on the spectral curve, the entries of the eigenvectors generically have poles at g points, which we take to be the P_i. Equivalently, we can consider the spectrum of the matrix M' obtained from the Lax matrix by removing the last row and column. This is a matrix of rank $N-1$, and the spectral curve of a generic Lax matrix is indeed a curve of genus $N-1$.

It would be interesting to study in more detail the construction and physical meaning of the points P_i, and their behavior under the flows of the Toda lattice, but we will not do that here. To prove that the Seiberg-Witten curve factorizes in the appropriate way, it suffices to use theorem 4 in [58], which gives an explicit expression for the velocities of the Toda flows on the Jacobian. Adapted to our situation, this theorem states that the velocities are given by

$$a_j = \mathrm{Res}_{x=\infty}(\omega_j A(x)), \tag{3.138}$$

for the flows given by $\dot{M} = [M, A(M)_+ - A(M)_-]$. Here, ω_j is a basis for the holomorphic one-forms on the Riemann surface. For a surface of the form $y^2 = P_N(x)^2 - 4\Lambda^{2N}$, such a basis is

$$\omega_j = \frac{x^{j-1}}{\sqrt{P_N(x)^2 - 4\Lambda^{2N}}}, \qquad j = 1, \ldots, N-1. \tag{3.139}$$

In order to make the flow given by $A(x) = W'(x)$ stationary, which is according to (3.111) a necessary condition to find an extremum of the superpotential, we therefore need that

$$\mathrm{Res}_{x=\infty}\left(\frac{x^{j-1}}{\sqrt{P_N(x)^2 - 4\Lambda^{2N}}} W'(x)\right) = 0, \qquad j = 1, \ldots, N-1. \tag{3.140}$$

Though necessary, these equations are not yet sufficient, since we have not yet taken the additional constraint $W'(M)_0$ in (3.111) into account. This constraint is in fact equivalent to $\mathrm{Tr}(W'(M)) = 0$, as one can see from the proof of lemma 4. If W' has a large order, $W'(M)$ appears to depend explicitly on the spectral parameter, but as explained above it is then natural to consider the infinite Lax matrix (3.3) instead, and view the trace as the trace over any N consecutive diagonal entries of the infinite matrix (since the matrix is periodic, this is well-defined). With this interpretation of $W'(M)$ for any W', one then obtains that $\mathrm{Tr}(W'(M))$ is proportional to

$$\mathrm{Res}_{x=\infty}\left(\frac{P_N'(x)}{\sqrt{P_N(x)^2 - 4\Lambda^{2N}}} W'(x)\right). \tag{3.141}$$

Thus, the complete set of equations we need to solve are (3.140) and (3.141 = 0), which can be summarized as

$$\mathrm{Res}_{x=\infty}\left(\frac{x^{j-1}}{\sqrt{P_N(x)^2 - 4\Lambda^{2N}}} W'(x)\right) = 0, \qquad j = 1, \ldots, N. \tag{3.142}$$

These equations have no solution if the order of W' is less than N. The only way in which we can have a stationary flow is if the Riemann surface degenerates, so that the genus becomes smaller and there are fewer holomorphic one-forms. So we will assume that the surface degenerates with

$$P_N^2(x) - 4\Lambda^{2N} = H_{N-k}^2(x)T_{2k}(x), \tag{3.143}$$

in which case the constraints we need to solve read

$$\mathrm{Res}_{x=\infty}\Big(\frac{x^{j-1}}{\sqrt{T_{2k}(x)}}W'(x)\Big) = 0, \qquad j = 1, \ldots, k. \tag{3.144}$$

To obtain this, observe that $P_N'/\sqrt{P_N^2 - 4\Lambda^{2N}} \to V_{k-1}/\sqrt{T_{2k}}$ for some polynomial V_{k-1} if the curve degenerates according to (3.143). Equations (3.144) are equivalent to

$$\frac{W'(x)}{\sqrt{T_{2k}(x)}} = G_{n-k}(x) + \sum_{l>k}\frac{c_l}{x^l}, \tag{3.145}$$

with G_{n-k} a polynomial of order $n-k$ (recall that n is the order of W'). Multiplying the left and right hand side by $\sqrt{T_{2k}}$, and taking squares, shows that

$$W'(x)^2 = T_{2k}(x)G_{n-k}(x)^2 - f_{n-1}(x), \tag{3.146}$$

where $f_{n-1}(x)$ is a polynomial whose order is at most $n-1$. Thus, we have in (3.143) and (3.146) once again precisely reproduced the factorization of the Seiberg-Witten curve (2.46).

3.8 Comments

3.8.1 Stationary flows

Although the equations of motion that we obtained from the superpotential imply that one particular flow is stationary, namely the flow generated by $W(M)$ via equation (3.111), the results in equations (3.113)–(3.117) show that $N-k$ inequivalent flows have become stationary. One can take any polynomial A of degree $N-k-1$, and in view of lemma 3 we can always write

$$A(M)(U_k(M)_+ - U_k(M)_-) = V(M)_+ - V(M)_- \tag{3.147}$$

for some polynomial V, and M commutes with all these operators. Since A contains $N-k$ free parameters, there are $N-k$ linearly independent flows that are stationary.

3.8.2 Number of moduli

We can also easily verify that the Lax matrices that extremize the superpotential have the right number of moduli that one expects from field theory. At low energies, for a solution of the form (3.113)–(3.117), we expect an unbroken $U(1)^k$ gauge group, and therefore k complex moduli. One of these is related to the diagonal $U(1) \subset U(N)$ and is not present in the Lax matrix. Therefore, we expect that the Lax matrix has $k-1$ complex moduli. These moduli are easily understood: under the commuting flows, any extremum of the superpotential is mapped to another extremum, because an extremum is a stationary point of a particular flow and all flows commute. We already argued above that $N-k$ flows are stationary, leaving $k-1$ flows that act non-trivially on the Lax matrix. These are the flows generated by $\mathrm{Tr}(M^j)$ for $j = 2, \ldots, k$. These flows must act non-trivially, because if one of them were to act trivially, there would exist a polynomial V of degree less that k such that $[M, V(M)_+ - V(M)_-] = 0$, and in that case we would be able to reduce the solution (3.113)-(3.117) even further. (We tacitly assumed that T_{2k} has no further quadratic factors in our solution). The above results are also easily understood in algebraic geometric framework given above.

3.8.3 $SU(N)$ versus $U(N)$

So far we discussed the theory with gauge group $U(N)$, and it is easy to see what needs to be changed if we are interested in gauge group $SU(N)$ instead. The only difference is that we need to drop the additional equation $W'(M)_0 = 0$ in (3.111), and that we should work with a traceless Lax matrix. In the algebraic approach in section 3.6, we can therefore no longer put $\lambda = 0$ when we apply lemma 4 to $A = W'$, and therefore we end up with precisely the same results as for $U(N)$, the only difference being that the order of f has to be n instead of $n-1$. This is consistent with the fact that $\mathrm{Tr}(M) = 0$, so that the linear term in W and thus the constant term in W' should have no effect whatsoever on the factorization and drop out of all equations. A change of the constant term in W' can be compensated by a change of f_n and indeed does not affect the physics in any way.

From the algebraic geometric perspective, changing $U(N)$ to $SU(N)$ implies that in (3.142) and (3.144) the range of j should be $1 \leq j \leq N-1$ and $1 \leq j \leq k-1$ instead. This then implies that we can have a term c_k/x^k in equation (3.145), and therefore f in (3.143) can be of order n instead of $n-1$. Thus we reach the same conclusion as above.

3.9 Conclusions

In this chapter we have described the low-energy effective superpotential for supersymmetric gauge theory based on the group $U(N)$ on $\mathbb{R}^3 \times S^1$. We have shown that it yields precisely the same vacuum structure as one obtains on $\mathbb{R}^4$, but the relation between the two is highly non-trivial. The results are a first step towards a direct derivation of the integrable system that underlies the Dijkgraaf-Vafa matrix model from field theory.

We believe that the formulation of the theory on $\mathbb{R}^3 \times S^1$ offers some advantages over the formulation on $\mathbb{R}^4$. For instance, the generalization of the periodic Toda chain to arbitrary gauge groups is known [31], and therefore the present formalism should also be applicable to groups like G_2 and E_6, for which the Dijkgraaf-Vafa matrix model has not yet been worked out. In [43] the case of pure glue SO/SP gauge theories was analyzed in several examples and for G_2 this has been studied in [44]. As indicated in the previous chapter, we also know the relevant integrable system for various other gauge theories, as summarized in [32], such as $N = 2$ theories with matter [59], $N = 4$ super Yang-Mills theory [60], certain conformally invariant $N = 2$ theories with gauge groups of quiver type [60, 40], and even for some 5d theories [61, 46]. In [42] both $U(N)$ gauge theory and $U(N)^*$ gauge theory were treated, and a different proof on conjecture 1 provided. It would be interesting if the central conjectures of the previous chapter can be proven in all of these cases as well. Ultimately, we would like to understand in all these cases the nature of the reduction of the integrable system in a supersymmetric vacuum.

In string theory, there is no obvious way to describe 3d field theory results using topological string theory, since this would require some 7d topological string theory. The integrable system can however in some cases be extracted from string theory using dualities and suitable brane configurations [62], and it is worthwhile to explore this connection further, see for instance [63].

The integrable system itself plays a crucial role in this discussion. There are many physical properties that beg for a nice explanation in terms of the integrable system, such as for example the loop equations and generalized Konishi anomaly of [48]. At the same time, there are many features of the integrable system we have not yet used, such as the explicit description of its solutions in terms of algebraic-geometric data [58], and such as the existence of additional flows associated to Whitham times (see e.g. [64] and references therein). The latter may help in finding the correct interpretation of the gluino bilinear superfields S_i.

We found it particularly elegant that lifting vacua from $U(N)$ to $U(tN)$ had such a nice interpretation in terms of a simple replica trick for the Lax matrix. In this context it is amusing to observe that there is a close relation between replica tricks, random matrix theory and the Toda lattice hierarchy [65], but whether that is of any relevance to the present discussion remains to be seen.

It is worthwhile to note that the second proof of factorization given in this chapter is probably much more widely applicable than just $U(N)$ gauge theory. Since the analog of [58] has been worked out for Toda systems based on untwisted affine Lie algebras in [7], the proof seems to lift to at least the D-series of gauge theories as well. However, according to [31] the integrable system needed to describe gauge theory based on the non-simply-laced classical Lie algebras B_l and C_l is the periodic Toda system associated to the dual of the Lie algebra, i.e. the affine algebra with long and short roots interchanged. These dual algebras can be identified with twisted affine Lie algebras and for this particular class of groups we have not been able to find the analog to [58]. Most probably this is just a technicality and we are confident that [7] can be generalized to include those cases as well. It would also be interesting to study the algebro-geometrical proof for the exceptional groups, like G_2

Appendix A A recurrence relation for $P_N(x)$

In this appendix we derive a recurrence relation for characteristic polynomials of the following type

$$P_N(x) = \det(x1_N - \Phi) = \begin{vmatrix} x-\phi_1 & -y_1 & 0 & . & . & . & 0 & -z \\ -1 & x-\phi_2 & -y_2 & 0 & . & . & . & 0 \\ 0 & -1 & x-\phi_3 & -y_3 & 0 & . & . & 0 \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & -1 & x-\phi_{N-2} & -y_{N-1} \\ -\frac{y_0}{z} & 0 & . & . & . & . & -1 & x-\phi_{N-1} \end{vmatrix}. \quad (3.148)$$

This determinant can be expressed in terms of determinants of the following form

$$G_N(x) = \begin{vmatrix} x-\phi_1 & -y_1 & 0 & . & . & . & 0 & 0 \\ -1 & x-\phi_2 & -y_2 & 0 & . & . & . & 0 \\ 0 & -1 & x-\phi_3 & -y_3 & 0 & . & . & 0 \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & -1 & x-\phi_{N-2} & -y_{N-1} \\ 0 & 0 & . & . & . & . & -1 & x-\phi_{N-1} \end{vmatrix}. \quad (3.149)$$

Expanding $P_N(x)$ along the bottom line and keeping only z-independent terms we get

$$\begin{aligned} P_N(x) &= (x-\phi_{N-1})G_{N-1}(x) - (-1)(-y_{N-1})G_{N-2}(x) + \\ &\qquad (-1)^{N+1}(-\frac{y_0}{z})(-1)^N(-z)G^+_{N-2}(x) \qquad (3.150) \\ &= (x-\phi_{N-1})G_{N-1}(x) - y_{N-1}G_{N-2}(x) - y_0 G^+_{N-2}(x), \end{aligned}$$

here $G^+_N(x)$ is equal to $G_N(x)$ with shifted ϕ's and y's (i.e. $\phi_i \to \phi_{i+1}, y_i \to y_{i+1}$).

The $G_N(x)$ and $G^+_N(x)$ are tri-diagonal and therefore satisfy the recurrence relations

$$G_N(x) = (x-\phi_{N-1})G_{N-1}(x) - y_{N-1}G_{N-2}(x) \qquad (3.151)$$

$$G^+_N(x) = (x-\phi_N)G^+_{N-1}(x) - y_N G^+_{N-2}(x). \qquad (3.152)$$

As a special case, take all the y's and ϕ's equal, then $G^+_N(x) = G_N(x)$ and the recurrence relation for $P_N(x)$ is

$$P_N(x) = (x-\phi)G_{N-1} - 2yG_{N-2} = G_N(x) - yG_{N-2},$$

in this case it is easy to show that the P_N satisfies the same recurrence relation as the G_N

$$P_N(x) = (x-\phi)P_{N-1}(x) - yP_{N-2}(x), \qquad (3.153)$$

which is, up to some rescaling, the Chebyshev recurrence relation. Therefore the $P_N(x)$ (with all y's and ϕ's equal) are given by the Chebyshev polynomials of the first or of the second kind.

Chapter 4

Matter in the fundamental and the XXX spin chain

In the previous chapters the proposal for an exact superpotential was mainly worked out for pure glue $U(N)$ gauge theory. An interesting question is then if and how this discussion extends to systems with additional matter, especially with matter that transforms in the fundamental representation of the gauge group which models quarks. This is the subject of the present chapter. As explained before, our proposal may be checked by calculating the vacua of the constructed superpotential and comparing them to the gauge theory expectations.

These expectations will be reviewed first. This lightning review will show that for gauge theory with fundamental matter, the vacua which survive the deformation to $\mathcal{N} = 1$ after turning on the superpotential are much more involved. Then the integrable system needed will be reviewed, in this case certain degenerations of the classical XXX spin chain, which is followed by an analysis of the superpotentials constructed following the general conjecture in chapter 2. It will be shown in several examples and some general classes of potentials that also in this more complicated case the conjectured superpotentials on $\mathbb{R}^3 \times S^1$ constructed in the underlying integrable system reproduce the structure of the expected vacua. This is surprising since much of the extra structure on the gauge theory side resides in vevs for the quark fields and there is no natural quark field in the integrable system. There is however in this case some fine print concerning certain magnetic vacua. Along the way some interesting technical results will be encountered, such as a general solution to a factorization problem of the Seiberg-Witten curve in one class of vacua, analogous to the one described in equation (2.46) for pure $U(N)$ gauge theory. Also, it is shown how our superpotential is directly related to the Affleck-Dine-Seiberg superpotential by integrating in the meson fields.

This chapter consists basically of material from [66]. The main difference is that in this chapter some more results on obtaining the magnetic vacua are included.

4.1 Gauge theory expectations

In this section we will very briefly review the zoo of phases of (the low energy effective theory corresponding to) $\mathcal{N}=2$ $U(N_c)$ 4 dimensional gauge theories with N_f ($\leq 2N_c$) fundamental matter hypermultiplets, softly broken to $\mathcal{N}=1$ by turning on a superpotential (the literature on this subject is vast, see e.g. [67], [68] or the review in [69] and references therein). In the following we will denote the scalar components of the supersymmetric multiplets by the same symbol as the multiplets themselves.

4.1.1 Classical supersymmetric gauge theories with fundamental matter on $\mathbb{R}^4$

The perturbative field content of the $\mathcal{N}=2$ supersymmetric gauge theory we will be studying consists of the $U(N)$ gauge vector multiplet, which contains an $\mathcal{N}=1$ adjoint scalar superfield Φ, and N_f pairs of massive hypermultiplets $(\tilde{Q},Q)$ in the (anti) fundamental. The global classical symmetries are an $SU(2)\times U(1)$ R-symmetry and an $SU(N_f)\times SU(N_f)\times U(1)_B$ flavor symmetry in the limit of vanishing masses. In Wess and Bagger [13] conventions the superpotential for theories with $\mathcal{N}=2$ supersymmetry and fundamental matter is

$$\mathcal{W}_{\mathcal{N}=2}=\sqrt{2}\tilde{Q}_i^a\Phi_a^bQ_b^i+\sqrt{2}\tilde{Q}_i^a m_j^i Q_a^j, \tag{4.1}$$

$$m_i^j=\mathrm{Diag}(m_1,m_2\ldots,m_{N_f}). \tag{4.2}$$

It is well known that this theory is perturbatively asymptotically free in the UV for $N_f<2N_c$ and finite for $N_f=2N_c$. We will break $\mathcal{N}=2$ softly to $\mathcal{N}=1$ by turning on a polynomial superpotential $W_{n+1}(\phi)=\sum_{k=0}^{n+1}\frac{g_k}{k}\mathrm{Tr}\,(\Phi^k)$ for the $\mathcal{N}=1$ adjoint scalar Φ. These theories have a rich vacuum structure, which can be studied by a variety of methods. Below we briefly present the results of a field theoretical analysis. The first step is calculating the minima of the classical gauge theory, in order to identify the possible phases of the theory. Since some of them will not receive quantum corrections the results there are exact. The D-term equations of the theory read

$$\begin{aligned} [\Phi, \Phi^\dagger] &= 0 \\ Q^i_a (Q^\dagger)^b_i - \tilde{Q}^i_a (\tilde{Q}^\dagger)^b_i &= \nu \delta^b_a, \qquad \nu \in \mathbb{R}. \end{aligned} \tag{4.3}$$

with ν the Fayet-Illiopoulos parameter, which in almost all cases must be zero in order for the classical theory to have a supersymmetric vacuum. We will not consider non-zero ν in this chapter.[1] The F-terms give

$$\begin{aligned} W'(\Phi)^b_a + \sqrt{2}\tilde{Q}^a_i Q^i_b &= 0 \\ \Phi^b_a Q^i_b + m^i_j Q^j_a &= 0 \\ \tilde{Q}^b_i \Phi^a_b + \tilde{Q}^a_i m^i_j &= 0. \end{aligned} \tag{4.4}$$

As is well known, the first D-term equation states that gauge transformations can be used to diagonalize $\Phi = \text{Diag}(\phi_1, \phi_2, \phi_3, \ldots \phi_{N_c})$. The F-terms state that the quark fields are either zero or eigenvectors of Φ with $\phi_a = -m_i$ for some a and i. If the quark masses are all distinct, we can label the quark fields and the adjoint scalar eigenvalues for which this condition is satisfied by $a = i = 1, \ldots, r$. For every r there are $\binom{N_f}{r}$ ways of choosing masses for which this condition holds. For unequal masses there are then r distinct eigenvectors of Φ which we can group into 'diagonal' $N_c \times N_f$ matrices for Q and $\tilde{Q}^t$,

$$Q = \text{Diag}(Q_1, \ldots, Q_r, 0, \ldots) \tag{4.5}$$

$$\tilde{Q}^t = \text{Diag}(\tilde{Q}_1, \ldots, \tilde{Q}_r, 0, \ldots). \tag{4.6}$$

By diagonal we mean $a_{ij} = 0$ if $i \neq j$ and if $i, j > \min(N_c, N_f)$. If some masses happen to be equal, than instead of a diagonal Q matrix we get a block diagonal matrix Q with sub-blocks Q_i, but we will first assume unequal masses. The remaining D and F term equations now split into equations for the different diagonal elements in Q. We get

$$|Q_i|^2 = |\tilde{Q}_i|^2. \tag{4.7}$$

Next, we order the masses and the eigenvalues of the adjoint scalar such that the first r of each are equal. The value of the quark vevs and the remaining eigenvalues of the adjoint scalar can be determined from the remaining F term equations,

$$W'(-m_i) + \sqrt{2}\tilde{Q}_i Q^i = 0, \qquad i = 1, \ldots, r \quad \text{(no sum)} \tag{4.8}$$

$$W'(\phi_a) = 0, \qquad a = r+1, \ldots, N_c. \tag{4.9}$$

If $W'(-m_i)$ is zero for some i, the quark fields Q_i should be zero and do not give rise to a Higgs branch. If $W'(-m_i)$ is non-zero, the quark gets a vev which Higgs-es

[1] In [67] ν is not necessarily the FI parameter, since there $SU(N)$ gauge theory is considered.

the corresponding $U(1)$ gauge group. Clearly, the number of vacua depends on the detailed form of the superpotential. For example, a quadratic superpotential will force all the $\tilde{N} = N_c - r$ eigenvalues of the adjoint scalar ϕ_a which are not equal to some $-m_i$ to be equal to each other. The resulting $U(\tilde{N})$ theory is known to have $\tilde{N}$ supersymmetric vacua.

When some masses are equal (as they will be in most of the cases we consider) more elaborate vacua are possible. See the analysis in, for instance, [67, 68]. The moduli space splits into two different branches: a baryonic branch and a non-baryonic branch. The latter branch contains the Coulomb branch. In summary the following happens:

If N_i eigenvalues of the adjoint scalar become fixed at one mass we retain a $U(N_i)$ gauge theory with a number M_i of massless flavors. If $W'(-m_i)$ is not zero, the combination of equations (4.7) and (4.9) (suitably generalized for matrices Q and $\tilde{Q}$) forces Q and $\tilde{Q}$ to acquire N_i non-zero eigenvalues which completely Higgs the $U(N_i)$ theory. If $W'(-m_i)$ is zero, at least some Q and $\tilde{Q}$ must be zero (and, in fact, they may all be). The solutions are then labeled by an index $l \leq \min(\frac{M_i}{2}, N_i), l \neq N_i$ and generically the quark fields both have l degrees of freedom, constrained by the remaining D term equation. The theory that described the remaining degrees of freedom is then a $U(N_i - l)$ gauge theory with $M_i - 2l$ massless flavors and $l(M_i - l)$ massless neutral Goldstone hypermultiplets from the broken flavor symmetry. The vacua described in this paragraph constitute the non-baryonic branch.

In addition to the vacua described above, there is another class of vacua possible when $M_i \geq N_i$ and $W'(-m_i) = 0$. This class is called the baryonic branch. These are vacua where the gauge group is fully Higgsed and the baryonic operators

$$\begin{aligned} B^{i_1 \ldots i_{N_i}} &= Q^{i_1}_{a_1} \ldots Q^{i_{N_i}}_{a_{N_i}} \epsilon^{a_1 \ldots a_{N_i}} \\ \tilde{B}^{i_1 \ldots i_{N_i}} &= \tilde{Q}^{i_1}_{a_1} \ldots \tilde{Q}^{i_{N_i}}_{a_{N_i}} \epsilon^{a_1 \ldots a_{N_i}} \end{aligned} \tag{4.10}$$

have a non-zero vacuum expectation value. In general, this branch is described by gauge invariant operators B, $\tilde{B}$ and M, subject to constraints. Note that the condition $W'(-m_i) = 0$ plays the same role as above.

Summarizing, the theory has several phases, which we distinguish by the expectation values of the quark fields. If these are zero, then we are on the Coulomb branch. If they are nonzero then we are either on the baryonic or the non-baryonic Higgs branches. It can happen that the expectation values of the quark fields on these branch can be tuned to zero, in this case we can continuously interpolate between Coulomb and Higgs branch. The special point on which they connect is the root of the Higgs branch, and in order for it to touch the Coulomb branch classically, $W'(-m_i)$ must be zero. A question is whether or not this last point might be modified in the quantum theory: while the metric on the Higgs branch does not get quantum corrected, this does not mean that the (position of the) boundary of moduli space can not get quantum corrections. In the study of the integrable system we usually take $W'(-m)$ to be zero, however we have not found a special status of this extra condition on the superpotential: the Higgs branch continues to be there in the quantum analysis.

As is well known, the moduli space of supersymmetric gauge theory is governed by holomorphy and the non-renormalization theorems. First of all, the Coulomb and Higgs branch structure reviewed above persists in the quantum case: the Kähler metric cannot contain mixing terms. Furthermore, the metric on the Higgs branch does not get quantum corrected, see [67]. The quantum corrections to the Coulomb branch are quite well known since the work of Seiberg and Witten. Note that these arguments fail on the special points of the moduli space on which extra degrees of freedom become massless, which then have to be added to the description. Below, we will review the field theory analysis of the quantum corrections to the Coulomb branch found above.

4.1.2 Weak coupling analysis

First we want to study the theory in a regime where the theory is expanded around saddle points of the classical superpotential. We therefore want to fix the eigenvalues of the adjoint such that W' vanishes for those values. Let us consider $W'(x) = \prod_{i=1}^{n}(x - a_i)$. If $|a_i - a_j| >> \Lambda, i \neq j$, then the saddle point analysis should give a good description of the theory. The gauge group is broken to $U(N_i)$ subgroups centered at the a's, where each subgroup has an effective quadratic potential $\sim \mu\Phi^2$ for the adjoint scalar of this subgroup, and the masses of the hypermultiplets get shifted by a_i. We can therefore reduce the analysis to the analysis of a $\mathcal{N} = 1$ theory perturbed by a quadratic potential with a certain number of massless fundamentals which can then be combined into an analysis for the general superpotential. The vacuum structure now depends on the number of massless flavors charged under this subgroup, M_i, since this number determines the effective low energy degrees of freedom of the theory. These can be determined, for instance, by analysis of the anomaly consistency conditions (see e.g. [70, 71]). We will analyze the system below mostly in the limit of vanishing masses and a large but finite μ.

The case $M_i < N_i$

The mass of the effective potential is very large (but finite), so the adjoint scalars can be integrated out of the superpotential (4.1) to give an effective theory for the meson field $M_i^j = Q_i^a Q_a^j$. Now add to this Lagrangian the known $N = 1$ non-perturbative contribution of the Affleck-Dine-Seiberg [72] superpotential to arrive at the potential

$$\mathcal{W} = -\frac{1}{2\mu}\mathrm{Tr}\, M^2 + (N_i - M_i)\Lambda_i^{\frac{3N_i - M_i}{N_i - M_i}} \det(M)^{\frac{-1}{N_i - M_i}}. \tag{4.11}$$

This potential is exact by the usual combination of holomorphy and symmetry arguments, supplemented by knowledge of weak coupling limits and analyticity, see e.g. [73]. The $\mathcal{N} = 1$ scale Λ_i can be related to the $\mathcal{N} = 2$ scale by scale matching, $\Lambda_i^{3N_i - M_i} = \mu_i^N \Lambda^{2N_i - M_i}$. We can now analyze the minima of the above potential by using the technique outlined in [73]. We find that the meson matrix can be diagonalized by flavor rotations to a matrix with two different entries. Let the number of entries of one type, say α_1 be given by an integer $r \leq [\frac{M_i}{2}]$. In general we will have $2N_i - M_i$ different solutions for every α_1, leading to $2N_i - M_i$ different vacua at every r. If $r = \frac{M_i}{2}$ then we have $N_i - M_i/2$ vacua.

Bare quark masses can be added to (4.11), which generically break the full flavor symmetry group. For a given r branch with massless quarks there are $\binom{M_i}{r}$ inequivalent ways of deforming this by giving masses to the quarks. The vacua obtained as minima of potentials of the form (4.11) will be referred to as 'electric' vacua.

The case $M_i = N_i$

In addition to the meson field degrees of (gauge invariant) freedom this case also has one baryonic degree of freedom. The effective $\mathcal{N} = 1$ theory is described by a linear sigma model for the meson and baryons. Adding to this the potential inherited from the $\mathcal{N} = 2$ theory we get an effective potential

$$\mathcal{W} = -\frac{1}{2\mu}\mathrm{Tr}\, M^2 + K(\det(M) - B\tilde{B} - \Lambda_1^{2N_i}). \tag{4.12}$$

The vacua of this theory include the same type of vacua as for $N_i < M_i$ when $B = \tilde{B} = 0$, so we can analytically continue the above counting results to this case. In addition to these, there is one vacuum with non-zero expectation value for the baryons $B, \tilde{B}$.

The case $M_i > N_i$

The non-renormalization theorem guarantees that the baryonic branches will not receive any quantum corrections. In that case the gauge theory is fully Higgsed

Table 4.1: *Number of extra vacua for $\mathcal{N}=2$ softly broken to $\mathcal{N}=1$ by a quadratic superpotential. Here $\hat{N}_i \equiv N_c - N_f$.*

case	# vacua	labeled by
all M_i	$\binom{M_i}{r}(2N_i - M_i)$	$r = 0, \ldots, [M_i]/2]$
	$\binom{M_i}{r}(N_i - \frac{M_i}{2})$	$r = M_i/2$
$M_i \geq N_i$	1	$r = N_i$
$M_i > N_i$	$\binom{M_i}{\hat{r}}(\hat{N}_i - \hat{r})$	$\hat{r} = 0, \ldots, \hat{N}_i - 1$
	1	$\hat{r} = \hat{N}_i$

and there is no non-trivial gauge theory quantum dynamics. For the non-baryonic branches the same argument which gave (4.11) holds and the counting argument of that potential holds. The number of vacua for the baryonic branch can be most easily counted through the Seiberg dual [70] formulation. This is a theory which is indistinguishable from the original theory in the infrared limit. In the case at hand, the dual theory consists of magnetic quarks and gluons for a gauge theory with dual gauge group $U(\hat{N}_i \equiv M_i - N_i)$ and M_i flavors. The effective superpotential for these degrees of freedom is given by

$$\mathcal{W}_d = \tilde{q}_i M^i_j q^j - \frac{1}{2\mu} \mathrm{Tr}\, M^2 + \left((N_i - M_i) \Lambda^{\frac{2N_i - M_i}{N_i - M_i}} \cdot \det(M)^{\frac{-1}{N_i - M_i}} \right), \qquad (4.13)$$

where the part between parenthesis is only added when the rank of the meson matrix is M_i, since in this case the IR theory is a pure glue gauge theory. This superpotential has two branches of solutions, depending on whether the meson matrix is degenerate or not (compare the classical analysis for the Higgs branches). If it is not degenerate, then the dual quarks can be integrated out and the analysis is the same as in the case $M_i < N_i$. In particular, we get the same number and degeneration pattern of vacua as we obtained there. We will refer to this case as 'electric' vacua. If the meson matrix *is* degenerate, the solutions are classified by an integer $\hat{r}$, just as in the previous analysis, but which now appears for the dual quarks and mesons. We will refer to these vacua as 'magnetic'. The difference in the situation in the electric case is the coupling to the meson matrix. This coupling gives a mass to all quarks. In the IR we are therefore left with a classical gauge group $U(\hat{N}_c - \hat{r})$ with no massless flavors which gives $\hat{N}_c - \hat{r}$ vacua with a surviving $U(1)$. In addition to these two branches we also have a baryonic branch, just as in the electric case.

The weak coupling analysis now yields the table 4.1, where we have only listed phases which appear *in addition* to the ones already found in the cases listed above it. The binomial factors are included to facilitate comparison to the results further on in the chapter. Note that the binomial factors are included and we stress that this is the result for zero effective quark mass. In this table we have abused our notation slightly by labeling the baryonic branches by the integer r or $\hat{r}$.

4.1.3 Strong coupling analysis

The final step consists of analyzing the Coulomb branch dynamics encoded in the Seiberg-Witten curve of the gauge theory and especially on its singular points. We take the Seiberg-Witten differential λ_{sw} from [74],

$$\lambda_{sw}(z) = \frac{z}{y}\left(P'(z) - \frac{1}{2}(P(z) - y(z))\frac{Q'(z)}{Q(z)}\right) dz = z\,\mathrm{dln}\,(y + P(z))\,, \tag{4.14}$$

with

$$y^2 = P(z)^2 - 4\Lambda^{2N_c - N_f}\,(z + m_i)^{N_f} \equiv P^2 - 4Q \tag{4.15}$$

the Seiberg-Witten curve. As noted in [74], there are several proposals for the exact form of the Seiberg-Witten curve for this specific gauge theory. Specifically, there can be a term of form $\Lambda^{2N_c - N_f}$ times a polynomial of order $N_f - N_c$ if $N_f \geq N_c$ in the polynomial P. This ambiguity concerns the Seiberg-Witten relationship of the fields a with the classical moduli. Specifying a definition of these operators in the quantum theory fixes the ambiguous polynomial, see e.g. [75]. In addition, we have the same ambiguity of the relation of the quantum operators $\mathrm{Tr} < \Phi^k >$ with their classical counterparts as discussed in the previous chapter. There we fixed the definition to be

$$\frac{1}{k}\mathrm{Tr} < \Phi^k > = \frac{1}{k}\frac{1}{2\pi i}\oint_\infty dz z^{k-1}, \lambda_{sw},$$

where the loop runs counterclockwise. Note that the Seiberg-Witten form also contains one power of z.

Turning on a superpotential lifts most of the moduli space of the parent $\mathcal{N} = 2$ theory; the only points not lifted are those at which new degrees of freedom (dyons) become massless or which are fixed by non-zero expectation values of the quarks. In this chapter we do not consider theories in which mutually non-local dyons become massless and we take all quarks to have the same mass m. Theories of the sort we are studying in this chapter have been studied by various other methods in the literature, for instance by using geometric engineering [76], Konishi anomaly techniques [48], exact potentials [18], or by using a brane setup [18]. The upshot is that the gauge theory vacua are described by a factorized Seiberg-Witten curve of the form

$$\begin{aligned} P(x)^2 - 4Q(x) &= (x+m)^{2r} H^2_{N_c - r - n + l}(x) F_{2n-2l}(x) \\ (W_n'^{(c)}(x) + f_{n+1}(x)) &= B_l^2(x) F_{2n-2l}(x), \end{aligned} \tag{4.16}$$

where r labels the Higgs branches as before, H is a polynomial in x of order $N_c - r - n + l$, B_l is a polynomial of degree l and $W_n'^{(c)}(x)$ is the derivative of the classical superpotential. This is the direct generalization of (2.46).

In [69] an extra contribution for $n \geq 2N_c - N_f$ to the non-singular part of the Seiberg-Witten curve was reported by revisiting the field theory analysis in [18]. The quantum operators used in that derivation however are *not* the quantum operators defined through the resolvent. In this thesis, as emphasized in chapter 2, the quantum operators are defined by the resolvent. In appendix B it is shown that if that definition is used in the gauge theory calculation, the extra contribution in the calculation disappears.

The baryonic Higgs branch, as seen from the Coulomb branch, is much more easy, since the gauge theory is in this case fully Higgs-ed. The only way this can be reflected in the Seiberg-Witten curve is the case when the whole curve factorizes, $y^2 = H^2$. More concretely, this argument can be substantiated by calculations such as those in [67]. The starting point is the realization that the root of the baryonic branch consists of a single point, invariant under the unbroken discrete $\mathbb{Z}_{2N_c - N_f}$ residual R-symmetry group. However, this is only true if the masses of the quarks are equal. Giving one quark a mass will in general split the baryonic branch into 2 sub-branches: one branch remains baryonic. The other branch will turn into a non-baryonic Higgs branch.

A small technicality

In the following we will repeatedly use a partially integrated form of expression (2.45). Partial integration will give a boundary term in the loop integral because the logarithm has a branch cut which we choose to be on the negative real axis. However, this boundary term mostly cancels in calculations (as it should) since we will only calculate the pole at infinity. Here the function $\ln(y + P(z)) \sim \ln\left(z^N + l.o.\right) = \ln(z^N) + \ln(1 + \mathcal{O}(1/z))$. The last part is a nice holomorphic function in a neighborhood of $z = \infty$. Now the boundary term cancels against the term coming from $\oint z^{k-1} \ln z^N$. The only exception to this canceling mechanism are terms

$$\frac{1}{2\pi i} \oint d(\log(z^N)) = N. \tag{4.17}$$

These terms we will however not need in this chapter since they correspond to the trace of the identity operator. We can therefore write

$$\frac{1}{k} \mathrm{Tr} < \Phi^k >= \frac{-1}{2\pi i} \oint_\infty dz z^{k-1} \ln\left(\frac{1}{z^{N_c}}(y(z) + P(z))\right). \tag{4.18}$$

4.1.4 Compactifying on a circle

Above we reviewed the vacuum structure of four-dimensional $\mathcal{N} = 1$ gauge theories. Once we compactify on a circle, the vacuum structure remains the same, as explained in chapter 2, but the moduli are different. The same line of reasoning as employed in the previous chapter then gives that the natural moduli space is $4N$-dimensional and naturally consists of the Seiberg-Witten curve plus the Jacobian. By the same reasoning as employed before, for a generic point on moduli space we can still try to find the right variables in which to express a conjectured exact quantum superpotential. In the analysis using the spin chain below, we will encounter complex variables p_i and q_i. In the classical limit, they can be roughly thought of as the vacuum expectation values of the adjoint scalar Φ and of the two additional scalars respectively.

The precise relation of these variables to the quantum operators in the theory is something we will not discuss in detail, but will certainly suffer from an ambiguity similar to the one discussed in the previous subsection. In addition, from the integrable system point of view any set of variables p_i', q_i' that is related to the original variables p_i, q_i by a canonical transformation is perfectly acceptable. We will see this explicitly in the next sections. Again, which of the canonical variables have the simplest interpretation in terms of the microscopical definition of the gauge theory is something which is outside the scope of the present chapter.

4.2 The classical $SL(2, \mathbb{C})$ spin chain and its degenerations

In this section we review the classical inhomogeneous twisted periodic $SL(2)$ spin chain. In this thesis we will refer to this system as the XXX spin chain. We follow mainly [32]. The curve with fundamental hypermultiplets is a generalization of the curve without, so the integrable system lurking behind should reduce in some limit to the periodic Toda chain. Since the spin-chain will be formulated in terms of 2 by 2 matrices, the periodic Toda system will first be written in terms of such matrices.

4.2.1 Periodic Toda in terms of 2×2 matrices

As was shown in chapter (3), the periodic Toda system can be defined by an infinite-dimensional Lax matrix which commutes with the shift-by-N operator (is N-periodic). We can write this infinite dimensional matrix acting on an eigenvector as a recursion relation,

$$\Lambda^2 e^{q_n - q_{n-1}} f_{n-1} + \phi_n f_n + f_{n+1} = \lambda f_n, \tag{4.19}$$

which can be written as a matrix equation (in just 'site-local' variables for the matrix),

$$\begin{pmatrix} f_{n+1} \\ e^{q_n} f_n \end{pmatrix} = \begin{pmatrix} \lambda - \phi_n & -\Lambda^2 e^{-q_n} \\ e^{q_n} & 0 \end{pmatrix} \begin{pmatrix} f_n \\ e^{q_{n-1}} f_{n-1} \end{pmatrix}. \tag{4.20}$$

Applying N matrices of this type will lead to the shift-by-N Toda transfer matrix T_{Toda}:

$$T_{Toda} = \prod_{i=N}^{1} \begin{pmatrix} \lambda - \phi_i & -\Lambda^2 e^{-q_i} \\ e^{q_i} & 0 \end{pmatrix}. \tag{4.21}$$

Since our original infinite dimensional matrix was periodic with period N, we should focus on eigenvalues of the transfer matrix. These eigenvalues are given by the roots of the characteristic equation of this 2 by 2 matrix which is

$$\det(T_{Toda} - w) = w^2 - \operatorname{Tr} T_{Toda} w + \det(T_{Toda}) = 0. \tag{4.22}$$

By defining $y = 2w - \operatorname{Tr} T_{Toda}$ and using $\det(T_{Toda}) = \Lambda^{2N}$ we get the familiar form of the spectral curve,

$$y^2 = (\operatorname{Tr} T_{Toda})^2 - 4\Lambda^{2N}. \tag{4.23}$$

4.2.2 The spin chain

The above formulation of the periodic Toda chain can be used to search for generalizations. Let us consider the following matrices (slightly confusingly also called Lax matrices),

$$L_j(\lambda) = \lambda - \lambda_j + \sum_{k=1}^{3} S_k^{(j)} \sigma^k, \tag{4.24}$$

$$L_j = \begin{pmatrix} \lambda - \lambda_j + S_1^j & S_2^j - iS_3^j \\ S_2^j + iS_3^j & \lambda - \lambda_j - S_1^j \end{pmatrix}, \tag{4.25}$$

where λ_i are constants (trivial Casimirs) known as impurities and the σ^j are the Pauli matrices. We define σ^0 to be the identity matrix. For this system to be integrable the Lax matrices should satisfy quadratic r-matrix relations (see e.g. [77], chapter 2),

$$\{L_i(\lambda) \overset{\otimes}{,} L_j(\lambda')\} = \delta_{ij} \left[r(\lambda - \lambda'), L_i(\lambda) \otimes L_j(\lambda') \right], \tag{4.26}$$

where we choose a rational r-matrix $\frac{1}{2(\lambda' - \lambda)} \sum_{a=0}^{3} \sigma^a \otimes \sigma^a$. This r-matrix is proportional to the interchange operator $P : A \otimes B \to B \otimes A$ and the r-matrix relation is invariant under multiplication on the left or right of both Lax matrices with a constant matrix. With this r-matrix the S_j should satisfy a classical $SL(2,\mathbb{C})$ Poisson algebra,

$$\{S_i^a, S_j^b\} = i\delta_{ab}\epsilon_{ijk}S_k^a. \tag{4.27}$$

Now define the transfer matrix and the associated spectral curve in analogy to the Toda case,

$$T(\lambda) = \prod_{j=N}^{1} L_j(\lambda) \tag{4.28}$$

$$\det(T(\lambda) - w) = w^2 - \operatorname{Tr} T w + \det(T) \equiv w^2 - P_N(\lambda)w + Q_{2N}(\lambda) = 0, \tag{4.29}$$

where the subscripts denote the degree of the polynomials. Defining $y = 2w - \operatorname{Tr} T_{Toda}$ we obtain

$$y^2 = (P_N)^2 - 4Q_{2N}. \tag{4.30}$$

An important step in analyzing integrable systems is obtaining its conserved quantities. The spin chain variables have a natural quadratic Casimir, the spin length $(S^a)^2$, which commutes with all S_i^b. If we fix the values of the Casimirs, there remains a Poisson manifold of dimension $2N$. The remaining variables can be expressed in terms of N coordinates and momenta (see for instance appendix A) with the usual Poisson bracket. To completely solve the integrable system we therefore need determine N conserved quantities. Observe that products of Lax matrices at different sites obey the same r-matrix relations, e.g. $\{L_1(\lambda)L_2(\lambda) \overset{\otimes}{,} L_1(\lambda')L_2(\lambda')\} = [r(\lambda-\lambda'), L_1(\lambda)L_2(\lambda)\otimes L_1(\lambda')L_2(\lambda')]$. Taking the trace on both factors in the direct product now gives a conserved quantity of the spin chain. A convenient basis can be obtained by studying $\{T(\lambda), T(\lambda')\}$, which gives the coefficients of the trace of the transfer matrix as conserved quantities.

Comparing with the Seiberg-Witten curve for $U(N_c)$ with $N_f = 2N_c$ flavors in e.g. [78] shows that the Q term should be (up to the modular functions of the conformal gauge coupling factor $e^{i\tau}$, which we will discuss momentarily) $\prod_{i=1}^{2N_c}(\lambda + m_i)$, where m_i are the masses of the fundamental hypermultiplets[2]. In the XXX chain the polynomial Q_{2N} is,

$$Q_{2N}(\lambda) = \prod_{i=1}^{N}((\lambda-\lambda_i)^2 - K_i^2), \tag{4.31}$$

where K_i^2 is the obvious quadratic Casimir of the Poisson algebra (4.27) at site i. We therefore obtain 2 masses from each site which are furthermore equal if K_i is

[2]strictly speaking the m_i as used here are modular functions themselves. If we take a limit to the case of $N_f < 2N_c$, the corresponding limit on the scale of the problem will reduce the $m(\tau)$ to simple constant masses.

zero. Note that the full spectral curve of the spin chain only contains Casimirs of the algebra and is therefore itself a conserved quantity.

As a final step in identifying gauge theory quantities, the modular functions have to be inserted in front of the Q, or for the more general ($N_f < 2N_c$) case introduce a scale into the problem. It has already been noted that the integrability condition (4.26) is invariant under multiplication of the Lax matrices by a constant matrix. We can therefore take generalized boundary conditions without affecting integrability, or in other words take a modified transfer matrix $\tilde{T} = CT$, where we take C to be

$$C = \begin{pmatrix} 1 & 1 \\ \frac{-1}{4}(1-g^2) & 0 \end{pmatrix}. \tag{4.32}$$

In the case where $N_f = 2N_c$, g is the modular function $\frac{\theta_2^4+\theta_1^4}{\theta_2^4-\theta_1^4}$. To really fit the curve in [78] λ should be shifted by $g\mu$ with $\mu = \frac{1}{2N_c}\sum m_i$ and the masses should be redefined in the spin chain to be $\mu_i = m_i - \mu$. However, since we will be reducing to non-finite gauge theories in a moment we will not do that here. In any case, in most cases considered in this chapter the gauge theory scale will be introduced in another way than by the above C matrix, see subsection 4.2.3.

4.2.3 Degenerations

In gauge theory the degeneration to lesser amounts of flavors is performed by integrating out massive degrees of freedom. On the level of the Seiberg-Witten curve this translates into taking several masses to infinity while keeping the product of these masses with the gauge theory effective scale fixed to a new effective scale. If our integrable models describe the gauge theory correctly, there should be a limit in the integrable system which reproduces this procedure. We will discuss how to take limits on the spin chain to integrate out masses. For one site and in different variables this was discussed in [32].

The $SL(2,\mathbb{C})$ algebra at each site has a Casimir which we will want to keep fixed in this chapter since it relates to the masses. Also the connection to the Toda chain is not clear in the variables S_i, so we will re-express the $SL(2,\mathbb{C})$ algebra in new variables q, p, K for each site. The Heisenberg-like variables q, p obey the standard Poisson algebra $\{q_i, p_j\} = \delta_{ij}$. An additional reason to switch to Heisenberg variables is that the 3d superpotential we will consider has known non-perturbative contributions in the pure glue sector of the form $\sum e^{q_{j+1}-q_j}$ and these appear naturally in Heisenberg like variables. Hence, these new variables are much closer to the naive gauge theory degrees of freedom. After some analysis (see the appendix (A)) we arrive at the Lax matrix:

$$L_j = \begin{pmatrix} \lambda - p_j & (m_{2j} + p_j)e^{-q_j} \\ -(m_{2j-1} + p_j)e^{q_j} & \lambda + (m_{2j} + m_{2j-1}) + p_j \end{pmatrix}, \tag{4.33}$$

where one of three possible forms is chosen, namely the one which is the most convenient one for taking limits. The other two forms are connected to this one by a (possibly singular) canonical transformation. In this form the inhomogeneities λ_j and the Poisson-algebra Casimirs K_j have been expressed in terms of the masses. We want to study degenerations to the Lax matrix of the Toda chain(4.20) which can be done by taking certain scaling limits. Immediately a problem presents itself: if we introduce the modular functions through a C matrix as above, then not all masses are multiplied by the coupling function in general which prevents us from taking masses to infinity directly. The origin of this problem is the site-locality of the masses at sites and the site global form of the coupling matrix. Since the masses are site-local, we will study a local limiting procedure below. However, this global versus local problem is the source of the existence of the different coordinate patches we will encounter later.

Local limits

As noted in the previous subsection, the integrable structure of the chain is invariant under multiplication by a constant matrix. Multiply (4.33) by constant matrices $A_J = \begin{pmatrix} 1 & 0 \\ 0 & \alpha_j \end{pmatrix}$ and $B_J = \begin{pmatrix} 1 & 0 \\ 0 & \beta_j \end{pmatrix}$ on the right and left respectively to obtain

$$L_j = \begin{pmatrix} \lambda - p_j & \alpha_j(m_{2j} + p_j)e^{-q_j} \\ -\beta_j(m_{2j-1} + p_j)e^{q_j} & (\lambda + (m_{2j} + m_{2j-1}) + p_j)\alpha_j\beta_j \end{pmatrix}, \tag{4.34}$$

At every site there are four different limits one can take:

1. Take both α_j and β_j to zero, keeping $\beta_j m_{2j-1}$ and $\alpha_j m_{2j}$ fixed (at Λ). We will call this limit (0)
2. Take α_j to zero and β_j to one, keeping $\alpha_j m_{2j}$ fixed (at Λ). We will call this limit (1^u).
3. Take β_j to zero and α_j to one, keeping $\beta_j m_{2j-1}$ fixed (at Λ). We will call this limit (1^d).
4. The trivial limit $\alpha = 1$ and $\beta = 1$, which we will call (2)

Note that the first limit reproduces the Toda chain 2 by 2 matrices. The gauge theory scale is introduced in the above by dimensional analysis and by comparing to the Toda chain. In principle we could also take an entirely different set of limits where in the end all the top left components of the Lax matrix would be put to zero when taking the Toda limit. However, it seems inconsistent to take such a limit on one of the Lax matrices and then take the above limit (0) on the next one: this procedure would for instance not reproduce a Seiberg-Witten curve of $U(N_c)$ gauge theory without flavors in the case where all masses are integrated out. Since taking the above described limit or the alternative limit at *all* matrices will give the same spectral curve we will only consider the ones in the list above.

We will use the following notation for limits: the total quark content of the chain will be given by a vector $M := (e_{N_c}, e_{N_c-1}, \ldots, e_1)$, where e_i indicates the limit taken on a site.

4.3 The proposal

In this section we present our proposal for the exact superpotential of $\mathcal{N} = 2\ U(N_c)$ gauge theory with N_f hypermultiplets in the fundamental representation on the manifold $\mathbb{R}^3 \times S^1$, which is broken to $\mathcal{N} = 1$ by adding a (polynomial) superpotential W for the adjoint superfield Φ.

4.3.1 Motivation

As explained in the second chapter, in order to write down the quantum superpotential we merely have to find good holomorphic variables on the moduli space. The integrable system and its degeneration described in the previous subsection has all the required properties and therefore provides us with variables that are valid on at least an open subset of the moduli space. Therefore, our proposal simply boils down to using the conserved charges of the integrable system to build the quantum superpotential. There is a subtlety regarding exactly which conserved quantity we should associate to the operator $\mathrm{tr}(\Phi^k)$. As discussed in section 4.1.3 we will use to gauge theory resolvent to give a precise quantum meaning to $\mathrm{tr}(\Phi^k)$ and to define it unambiguously in terms of the conserved charges of the integrable system. This quantum definition is also the one for which the factorization of the Seiberg-Witten curve takes on the simple universal form of equation (18).

4.3.2 The superpotential

The form of the superpotential in 4d for the softly broken $\mathcal{N}=2$ theory is

$$W = W(\Phi) + \sqrt{2}\tilde{Q}^i_a m^j_i Q^a_j + \sqrt{2}\tilde{Q}^i_a \Phi^a_b Q^b_i. \tag{4.35}$$

In the pure glue case it was natural to replace the moduli space coordinates in the superpotential by (a specific basis of) the conserved quantities (Hamiltonian functions) in the integrable system, i.e. traces of powers of the Lax matrix. In the present case that would translate to replacing the moduli space coordinates by certain Hamiltonians of the XXX chain. In the direct application of this idea we then run into the problem of what to do with the Q's: since there are no dynamical variables in the chain which correspond naturally to the fermionic quark fields or the composite meson field variables. This is highly reminiscent of the lack of natural glueball field variables in the periodic Toda chain. Note that, just as the on-shell glueball fields can be calculated from the factorization of the curve, on-shell we can recover the meson vevs [48] as well. Imagine integrating out the quarks, irrespective of whether this is allowed. We are then left with an effective superpotential of the same degree as the original one. However, in order to integrate out the quarks we need to choose a particular vacuum around which to expand. One would therefore expect to find only one branch, with fixed r, of the full vacuum structure. For an attempt to calculate an integrable system potential in a similar way see appendix C.

In the main body of this chapter a more naive approach will be followed by taking $W(\Phi)$ and replacing the gauge theory Casimirs directly by the corresponding expression calculated through the gauge theory resolvent. Note that this (correctly) reduces to the pure glue gauge theory in that particular limit.

In the following several vacua of this proposed superpotential will be studied and, in contrast to the naive expectation, several r-branches are described by the minima of one superpotential. In fact, for a quadratic potential and $N_f < N_c$ all expected vacua will be found in one go. Note that that there can be different solutions to the equations of motion which yield the same Seiberg-Witten curve. These solutions therefore describe the same vacuum from the four dimensional point of view. However, on $\mathbb{R}^3 \times S^1$ they obviously correspond to inequivalent vacua.

4.4 Addition and multiplication maps

On field theory grounds some vacua of gauge theories with different matter content and/or different rank of the gauge group should be connected through the so-called addition and multiplication maps (see for instance [69]). In this section we show that these maps both have a natural counterpart in the integrable system. We will use that the e.o.m. derived from a polynomial superpotential can be written as,

$$\frac{\delta W}{\delta o_i} = \frac{-1}{2\pi i}\oint_{\infty} W'(x)\frac{\frac{\delta P}{\delta o_i}}{\sqrt{P^2-4Q}}. \tag{4.36}$$

Here o_i is either p_i or q_i. Hence if the derivative of the transfer matrix with respect to the coordinates vanishes the equations of motion are satisfied.

4.4.1 Addition map

Suppose we have a solution of the spin chain e.o.m derived for a fixed superpotential $W = \sum_{k=0}^{n+1}\frac{g_k}{k}\mathrm{Tr}\,(\Phi^k) = \sum u_k g_k$ in a N_c, N_f parent gauge theory. Using this it must then be shown that (some of the) solutions to the e.o.m of spin chain corresponding to a daughter gauge theory with the same superpotential and $\tilde{N}_c = N_c+1, \tilde{N}_f = N_f+2$ can be generated. Likewise, there must be a solution in the $\tilde{N}_c, \tilde{N}_f$ which can be projected down to the parent theory. In the field theory analysis [69] this map is known as the addition map and below we show how it has a natural interpretation in our proposal.

Start with the transfer matrix of the N_f, N_c theory and add one diagonal Lax matrix $(x+\tilde{m})\mathbb{I}_{2\times 2}$ to it, which can be obtained by adding a spin chain Lax matrix with 2 masses $\tilde{m}$ and setting $p \to -\tilde{m}$. Then it easily seen that the spectral curve of the resulting transfer matrix has the form $(x+\tilde{m})^2(P_N^2 - 4Q)$ appropriate for a $\tilde{N}_c = N_c+1, \tilde{N}_f = N_f+2$ theory, with 2 flavors of mass $\tilde{m}$ on the Higgs branch. It then must be shown that the e.o.m. of the two theories have a simultaneous solution which has $p \to -\tilde{m}$, so that a solution in the parent theory can be lifted to the daughter. The generating function for the e.o.m. in the $\tilde{N}_c$ theory is obtained by adding a matrix at site $N+1$ to a $U(N_c)$ transfer matrix,

$$\mathrm{Tr}\,T_{\tilde{N}_c}(x) = \mathrm{Tr}\left(\begin{pmatrix} x-p_{\tilde{N}_c} & (\tilde{m}+p_{\tilde{N}_c})e^{-q_{\tilde{N}_c}} \\ -(\tilde{m}+p_{\tilde{N}_c})e^{q_{\tilde{N}_c}} & x+2\tilde{m}+p_{\tilde{N}_c}\end{pmatrix}\begin{pmatrix} t_{11} & t_{12} \\ t_{21} & t_{22}\end{pmatrix}\right), \tag{4.37}$$

$$\begin{aligned}\mathrm{Tr}\,T_{\tilde{N}_c}(x) \;&=\; (x-p_{\tilde{N}_c})t_{11} + (\tilde{m}+p_{\tilde{N}_c})e^{-q_{\tilde{N}_c}}t_{21} - \\ &\qquad (\tilde{m}+p_{\tilde{N}_c})e^{q_{\tilde{N}_c}}t_{12} + (x+2\tilde{m}+p_{\tilde{N}_c})t_{22}.\end{aligned} \tag{4.38}$$

It is easy to see that the e.o.m. for $o_i \in p_i, q_i$ $i \neq N_c+1$ simply project down to the $U(N_c)$ theory when $p_{\tilde{N}_c} = -\tilde{m}$, since only the t's depend on these coordinates. In this case the e.o.m. can be derived from

$$\frac{\delta T_{\tilde{N}_c}}{\delta o_i} = (x+m)\frac{\delta T_{N_c}}{\delta o_i}. \tag{4.39}$$

The equation of motion for $q_{\tilde{N}_c}$ is also solved by this value for $p_{\tilde{N}_c}$. The equation obtained by varying with respect to $p_{\tilde{N}_c}$ is quadratic in $e^{q_{\tilde{N}_c}}$ on this solution since in that case,

$$\mathrm{Tr}\, T \sim t_{22} - t_{11} - e^{-q_{\tilde{N}_c}} t_{21} - e^{q_{\tilde{N}_c}} t_{12}, \tag{4.40}$$

and the Seiberg-Witten curve is independent of $e^{q_{\tilde{N}_c}}$ on the solution. The resulting equations can be solved for $e^{q_{\tilde{N}_c}}$ which yields two solutions. The solutions for $p_{\tilde{N}_c}$ and $q_{\tilde{N}_c}$ map two solutions to the e.o.m. of the daughter theory to the parent. These two solutions for $q_{\tilde{N}_c}$ correspond to the same Seiberg-Witten curve in the parent theory. Vice-versa, one solution to the parent theory e.o.m. can be lifted to two solutions in the daughter with equivalent Seiberg-Witten curve. It remains to be shown that the e.o.m. projected down to the parent theory can be derived from the same superpotential in both theories, but this follows trivially from (4.36). Note that on shell we have

$$\tilde{u_k} = \frac{1}{k}(\tilde{m})^k + u_k. \tag{4.41}$$

4.4.2 Multiplication map

The multiplication map can be derived in the 2×2 formalism slightly simpler than it was derived for the pure glue gauge theory in the $N_c \times N_c$ form in chapter 3. The derivation below is the extension of the integrable systems multiplication map to theories with fundamental matter. Consider a transfer matrix for a $U(tN_c)$ gauge theory with tN_f quarks which is the product of t transfer matrices T_{N_c} of the parent $U(N_c)$ theory.

It needs to be shown that solutions to the e.o.m. for the $U(tN_c)$ can be projected solutions to the e.o.m. of the $U(N_c)$ theory by periodically identifying coordinates and momenta. The difference between the e.o.m. for q_j and q_{at+j} can be studied in the $U(tN_c)$ theory, with a an integer,

$$\left(\frac{\delta}{\delta q_j} - \frac{\delta}{\delta q_{at+j}}\right) P_{tN_C}. \tag{4.42}$$

The transfer matrix can be written in a site-local form. Hence taking derivatives with respect to coordinates q_j and q_{at+j} (or momenta) and periodically identifying with period t will yield the same transfer matrix in both cases. Therefore the right hand side of equation (4.42) vanishes and the e.o.m. of the $U(tN_c)$ theory can be projected down to $U(N_c)$ theory.

It remains to be shown that the equations of motion can be derived from the same superpotential and how the Seiberg-Witten curves are related. Below we will first show that the curve in the $U(tN)$ theory is related to the curve in the $U(N)$ theory as

$$P_{tN_c}^2 - 4Q_{tN_c} = H_{2(t-1)N_c}^2 (P_{N_c}^2(x) - 4\det(T_{N_c})), \tag{4.43}$$

with H a polynomial of order $2(t-1)N_c$. In particular, if one curve factorizes as in equation (4.16) for a particular solution in that theory, so does the multiplication map of that curve when evaluated on the solutions which have been generated by the multiplication map of that solution.

The trace of the product of transfer matrices, $\mathrm{Tr}\, T_{tN_c}$, only depends on the two eigenvalues of T_{N_c}, so the trace of the product of t of them can be written as a function of the trace and the determinant of T_{N_c}. In fact, since for 2 by 2 matrices

$$\mathrm{Tr}\,(T_{tN_c}) = \mathrm{Tr}\,(T_{N_c}^t) = \mathrm{Tr}\,(T_{N_c})\mathrm{Tr}\,(T_{N_c}^{t-1}) - \det(T_{N_c})\mathrm{Tr}\,({T_{N_c}}^{t-2}), \tag{4.44}$$

we have

$$\frac{\mathrm{Tr}\,(T_{N_c}^t)}{(2\det(T_{N_c}))^{t/2}} = \frac{2\mathrm{Tr}\,(T_{N_c})}{2\sqrt{\det(T_{N_c})}} \frac{\mathrm{Tr}\,(T_{N_c}^{t-1})}{2\det(T_{N_c})^{(t-1)/2}} - \frac{\mathrm{Tr}\,(T_{N_c}^{t-2})}{2(\det(T_{N_c}))^{(t-2)/2}}, \tag{4.45}$$

or defining $C_t = \frac{\mathrm{Tr}\,(T_{N_c}^t)}{2(\det(T_{N_c}))^{t/2}}$ and $x \equiv C_1 = \frac{\mathrm{Tr}\,(T_{N_c})}{2\sqrt{\det(T_{N_c})}}$

$$C_t(x) = 2xC_{t-1}(x) - C_{t-2}(x), \tag{4.46}$$

which is the Chebyshev recursion relation (for Chebyshev polynomials of the *first* kind with the above initial conditions). Putting it all together we get,

$$P_{tN_c}(x) = 2(\det(T_{N_c}))^{t/2} C_t \left(\frac{P_N(x)}{2\sqrt{\det(T_{N_c})}} \right), \tag{4.47}$$

as is expected on purely field theory grounds (see e.g. [23]) and which combines nicely with the determinant of T_{tN_c} to give the correct factorization of the curve.

$$\begin{aligned} P_{tN_c}^2 - 4Q_{tN_c} &= 4\det(T_{N_c})^t \left(C_t^2 \left(\frac{P_{N_c}(x)}{2\sqrt{\det(T_{N_c})}} \right) - 1 \right) \\ &= \det(T_{N_c}))^{t-1} U_{t-1}^2 \left(\frac{P_{N_c}(x)}{2\sqrt{\det(T_{N_c})}} \right) \left(P_{N_c}^2(x) - 4\det(T_{N_c}) \right). \end{aligned}$$

Note that the first part of the last equation can be written as the square of a *polynomial*. This can be easily proven using the parity relations for the Chebyshev polynomials. Since the scale in the $U(tN_c)$ theory must be real, $\det(T_{N_c})$ must be a t-th root of unity and there are t inequivalent ways of lifting a solution of the $U(N_c)$ theory to a solution of the $U(tN_c)$ gauge theory.

Using the above analysis of the Seiberg-Witten curve, it can now be shown that the equations of motion of parent and daughter theory in the multiplication map are derived from the same superpotential. From equation (4.36) we get that

$$\frac{\delta W}{\delta o_i} = \frac{-1}{2\pi i}\oint_{\infty} W'(x) \frac{C_t'\left(\frac{P_N(x)}{2\sqrt{\det(T_{N_c})}}\right)\frac{\delta P_N}{\delta o_i}}{U_{t-1}\left(\frac{P_{N_c}(x)}{2\sqrt{\det(T_{N_c})}}\right)\sqrt{P_N^2-4Q_N}}. \tag{4.48}$$

The following relation can easily be proven using the generating functions for the Chebyshev polynomials,

$$C_t'(x) = tU_{t-1}(x), \tag{4.49}$$

and it is then easy to see that the equations of motion in both parent and daughter theory are derived from the same superpotential. Note that there is no freedom to choose limits in these versions of the multiplication and addition maps, although when the above maps are combined with finite non-singular canonical transformations there certainly is such freedom.

4.5 Examples

This section contains several examples in which the vacua of several of our proposed superpotentials are calculated. Unless stated otherwise these examples are for a quadratic superpotential with $W'(-m) = 0$ and with equal mass fundamental hypermultiplets.

4.5.1 Gauge group $U(2)$

2 flavors

We will do this case in some detail, to illustrate the (in)equivalence of the various limits that one can take on the full XXX chain. In the notation introduced in section 4, there are three different and inequivalent limits that one can take, namely $(2,0)$, $(1^u,1^u)$ and $(1^u,1^d)$. Below we list for each of these three cases the value of $P_2(x)$, $Q_2(x)$, $u_1 = \mathrm{Tr}(\Phi)$ and $u_2 = \frac{1}{2}\mathrm{Tr}(\Phi^2)$. The latter we compute using the gauge theory resolvent, so that it is slightly different from the naive quantity that one would extract from $P_2(x)$ (see equation (4.139)). For simplicity we also introduce the notation $y = e^{q_1 - q_2}$. We find for the case $(2,0)$

$$\begin{aligned}
P_2(x) &= (x-p_1)(x-p_2) - y\Lambda(p_1+m_1) - y^{-1}\Lambda(p_1+m_2) \\
Q_2(x) &= \Lambda^2(x+m_1)(x+m_2) \\
u_1 &= p_1+p_2 \\
u_2 &= \frac{1}{2}(p_1^2+p_2^2) + \Lambda^2 + \Lambda y(p_1+m_1) + \Lambda y^{-1}(p_1+m_2),
\end{aligned} \tag{4.50}$$

for the case $(1^u, 1^u)$

$$\begin{aligned}
P_2(x) &= (x-p_1)(x-p_2) - y\Lambda(p_1+m_1) - y^{-1}\Lambda(p_2+m_2) + \Lambda^2 \\
Q_2(x) &= \Lambda^2(x+m_1)(x+m_2) \\
u_1 &= p_1+p_2 \\
u_2 &= \frac{1}{2}(p_1^2+p_2^2) + \Lambda y(p_1+m_1) + \Lambda y^{-1}(p_2+m_2),
\end{aligned} \tag{4.51}$$

and finally for $(1^u, 1^d)$

$$\begin{aligned}
P_2(x) &= (x-p_1)(x-p_2) - y\Lambda^2 - y^{-1}(p_1+m_1)(p_2+m_2) + \Lambda^2 \\
Q_2(x) &= \Lambda^2(x+m_1)(x+m_2) \\
u_1 &= p_1+p_2 \\
u_2 &= \frac{1}{2}(p_1^2+p_2^2) + \Lambda^2 y + y^{-1}(p_1+m_1)(p_2+m_2).
\end{aligned} \tag{4.52}$$

As we expected, these three cases are related to each other by canonical transformations. The second case is obtained from the first one by the transformation

$$p_1 \to p_1 - \Lambda y^{-1}, \qquad p_2 \to p_2 + \Lambda y^{-1}, \tag{4.53}$$

while the third case is obtained from the second one by the transformation

$$y \to \Lambda y/(p_1+m_1). \tag{4.54}$$

Notice that this last transformation is not well-defined everywhere on the phase space. This illustrates the general phenomenon that different degenerations can in principle describe different open subsets of moduli space, though on overlaps they are always related by a canonical transformation.

The vacua of a purely quadratic superpotential (proportional to u_2) in the case of equal masses $m_1 = m_2 = m$ are easily found. In the third case $(1^u, 1^d)$, there are three types of solutions. The first are the $r = 0$ solutions

$$p_1 = p_2 = \pm\Lambda, \quad y = -\frac{\Lambda \pm m}{\Lambda}, \tag{4.55}$$

for which

$$P_2^2(x) - 4Q_2(x) = (x \mp 2\Lambda)^2(x^2 \pm 4m\Lambda + 4\Lambda^2). \tag{4.56}$$

This follows the general factorization pattern summarized in (4.16). Secondly, we have the $r = 1$ solutions

$$p_1 = \frac{-m \pm \sqrt{m^2 - 4\Lambda^2}}{2}, \quad p_2 = \frac{-m \mp \sqrt{m^2 - 4\Lambda^2}}{2}, \quad y = 1 \tag{4.57}$$

for which

$$P_2^2(x) - 4Q_2(x) = (x + m)^2(x^2 - 4\Lambda^2). \tag{4.58}$$

This shows that both solutions give rise to the same Seiberg-Witten curve. They are nevertheless inequivalent solutions for the gauge theory on $\mathbb{R}^3 \times S^1$.

There is one more solution which is not strictly speaking a solution. It corresponds to the baryonic branch. Consider the $\epsilon \to 0$ limit of

$$p_1 = p_2 = -m + \epsilon, \quad y = \epsilon/m. \tag{4.59}$$

One can easily check that the p_i and q_i equations of motion vanish in this limit, and that the Seiberg-Witten curve becomes

$$P_2^2(x) - 4Q_2(x) = (x + m + \Lambda)^2(x + m - \Lambda)^2. \tag{4.60}$$

However, this solution is at the boundary of the coordinate system.

The $r = 0$ and $r = 1$ solutions can be mapped directly to the first two limits $(2, 0)$ and $(1^u, 1^u)$. However, the baryonic branch is more complicated, as this is exactly where the map between the $(1^u, 1^u)$ and $(1^u, 1^d)$ degenerates. One can take limits of the coordinates in $(2, 0)$ and $(1^u, 1^u)$ so that the Seiberg-Witten curve takes the form (4.60), but in none of those limits do the equations of motion also vanish. In any case, the baryonic root appears to be never a bona fide finite solution of the equations.

An alternative limit for $N_f = 2$

There is in fact one more limit to do the calculation for. This limit is actually not contained in the form (4.34), but arises when one of the other two options for Heisenberg variables in appendix (A) is chosen. There is just one possibility which is not trivially related to either itself or the previous explored cases: take the Lax matrix on the first site to be the Toda one and the Lax matrix on the second site to be

$$L_2 = \begin{pmatrix} \lambda - p_2 & (m_1 + p_2)(m_2 + p_2)e^{-q_2} \\ -e^{q_2} & (\lambda + (m_2 + m_1) + p_2) \end{pmatrix}. \tag{4.61}$$

We find for this case

$$\begin{aligned}
P_2(x) &= (x-p_1)(x-p_2) - y^{-1}\Lambda^2 - y(p_1+m_2)(p_1+m_1) \\
Q_2(x) &= \Lambda^2(x+m_1)(x+m_2) \\
u_1 &= p_1+p_2 \\
u_2 &= \frac{1}{2}(p_1^2+p_2^2) + \Lambda^2 + \Lambda y(p_2+m_1)(p_2+m_2) + \Lambda^2 y^{-1}, \qquad (4.62)
\end{aligned}$$

This seems similar to what was obtained in the limit $(1^u, 1^d)$ by taking $y \to 1/y$, but note that the factor which multiplies Λy contains just p_2 and not p_1. By the application of

$$y \to \frac{\Lambda}{y(p_1+m_1)}. \qquad (4.63)$$

this case is mapped to the case $(2,0)$. However, the $r=1$ vacuum found there in the case of equal masses is located precisely at the point where the above map degenerates. This is reflected in the fact that the e.o.m. for a quadratic superpotential (and equal masses) for the alternative limit of this subsection only have two $r=0$ solutions. The $r=1$ solutions are missing, which indicates that the moduli space coordinates of this particular alternative limit are a poor choice.

Since the map displayed above will always connect the alternative limit to the limit (2), there will always be a possibility for degeneration in this case. Further on in this chapter it will be found that for $N_f < N_c$ all vacua can be obtained from a $(1^u, 1^u, \ldots, 0, 0)$ limit. This will make that class of limits (experimentally) much better behaved than the alternative. This is the reason that in this chapter generically the form of (4.33) is chosen for the Lax matrices instead of the option (4.61).

1 and 3 flavors

For the limits $(1^u, 0)$ and $(1^d, 0)$ we find the expected 3 genus zero SW curves. For $(2, 1^d)$ we find 3 different phases ($r = 0, 1$ and a baryonic root). The $r=1$ branch is triply degenerate, as expected. The baryonic root is located on the edge of the phase space as above and consists of one solution. The gauge theory analysis gives in this case 2 phases in addition to the three found. These are the phases obtained from the Seiberg dual quark potential in equation (4.13) when the meson matrix is degenerate. However, from the Coulomb branch point of view, these phases are contained in the baryonic root. Since the Higgs branch roots found for the case $N_f = 3$ are triply degenerate, correctly corresponding to all solutions, it is expected that the extra magnetic phases would show up as degenerate solutions as well.

To see if the needed extra vacua are not hidden by a bad choice of moduli space coordinates such as the choice in the alternative limit displayed in the previous section, we can study the case of two zero masses and one non-zero mass. The non-zero mass can either be in the 2 or in the 1^d of the limit $(2, 1^d)$. In the case 1^d we continue to find 5 solutions. Two $r = 1$ solutions and the baryonic branch have the same double root structure in the on-shell Seiberg-Witten curve. The same structure is found if the non-zero mass is the top mass in the 2.

More interesting things happen if the non-zero mass is the bottom mass in the 2. In addition to the vacua found above, an extra baryonic vacuum appears at a finite point in moduli space. The baryonic vacua found give Seiberg-Witten curves

$$\begin{aligned} y^2 &= x^2(x + m - \Lambda)^2 \\ y^2 &= (-x^2 + m\Lambda + \Lambda y)^2. \end{aligned} \tag{4.64}$$

The solution to the e.o.m. which give the first vacuum diverges in the limit $m- > \infty$ and also does not admit a holomorphic decoupling limit. The second Seiberg-Witten curve corresponds to the baryonic vacuum found before. These results seem to suggest that the 'magnetic' vacua of the theory can be found in some coordinate patches on gauge theory moduli space, but not in all coordinate patches. The fact that one mass seems to play a privileged role in the chain seems reminiscent of the ordering problem we encountered in chapter (3) for taking classical limits.

4 **flavors**

This is a simpler version of the case studied in [45] and for direct comparison we take the superpotential in this case to be purely quadratic $W(x) = x^2$. The difference is that in that article the flavor masses were all assumed to be unequal and the solution was calculated in the weak coupling limit. We will not present the calculation here, but we find the expected 2 electric branches: two $r = 0$ branches and a quadruply degenerate $r = 1$ branch.

2 **flavors with unequal masses**

We study this in the limit $(1^u, 1^d)$. Since this example is more involved we will elaborate slightly. We start by imposing a superpotential of the form $W(x) = \frac{1}{2}x^2$. The e.o.m. for the momenta are easily solved. We are then left with one equation for the coordinates which, using the transformation $e^{q_1} = b_1, e^{q_2} = b_2$ reads,

$$\frac{b_2\left(\Lambda^2\left({b_1}^2 - {b_2}^2\right)^2 + {b_1}^2\left(b_1\, b_2\, {m_2}^2 - \left({b_1}^2 + {b_2}^2\right) m_2\, m_3 + b_1\, b_2\, {m_3}^2\right)\right)}{{b_1}^3 - b_1\, {b_2}^2} = 0. \tag{4.65}$$

This yields 5 solutions, a trivial one ($b_2 = 0$) and the four solutions of the quartic equation. The Seiberg-Witten curve for the trivial solution is

$$y_{\text{triv}}^2 = \left(x^2 - \Lambda^2 + m_2\, m_3 + x\; (m_2 + m_3)\right)^2 . \tag{4.66}$$

The solution to the quartic equation is not very insightful by itself and we were unable to show the factorization of the Seiberg-Witten curve directly. However, factorization can be shown through a different route[3]. Note that, since we are first and foremost interested in roots of the Seiberg-Witten curve, we can try to find it's double roots when evaluated on the solution for the momenta. We shall use the quadratic part of the numerator of equation (4.65) repeatedly to simplify expressions. Using this, it can be shown that the spectral curve simplifies to

$$y^2 = -4\,\Lambda^2\; (x + m_2)\; (x + m_3) + \left(x^2 + \frac{b_1(m_2 + m_3)}{b_1 + b_2} x + \frac{2\Lambda^2 b_1 - b_2}{b_1}\right)^2 . \tag{4.67}$$

We first try to find its double roots by studying the cubic equation $\frac{\partial y^2}{\partial x} = 0$. Happily, this equation factorizes into a linear and a quadratic equation. The linear equation has a root at

$$-\frac{(m_2 + m_3) b_1}{b_1 b_2} . \tag{4.68}$$

This root is a candidate for a double root of the spectral curve (4.67). Plugging this root back into the curve we see that it is indeed a double root of that curve iff equation (4.65) holds. We have therefore found the double root of the spectral curve. Using polynomial division gives the non-singular part, which has the simple form

$$(x)^2 - \frac{4\,\Lambda^2\, b_2}{b_1} . \tag{4.69}$$

Numerical evaluation of the double roots shows that these consists of solutions which connect (at $m_2 = m_3$) to a doubly degenerate $r = 1$ branch and the expected two solutions for the $r = 0$ branch.

2 flavors with a cubic potential

We study a superpotential of the form $W(x) = \frac{1}{3}x^3 - a^2 x$, using the limit $(1^u, 1^d)$. Solving the equations for the coordinates yields three independent solutions. Two of these are related by $\Lambda \to -\Lambda$. The third one yields two genus one Seiberg-Witten curves of the same form,

$$y^2 = -4\,(m + x)^2\,\Lambda^2 + \left(-a^2 + x^2 + 2\Lambda^2\right)^2 . \tag{4.70}$$

[3]another route would be to compute the common divisors of the of the curve expressed in spin chain variables and the remaining equation of motion

The extra contribution in the P-polynomial exactly cancels the leading term of the Q polynomial to form $W'(x)^2 + f_{n-1}$ just as expected from the field theory analysis in appendix B. There are 2 main branches of solutions remaining with two remaining equations in each of them. Examination of these equations shows these have again 3 sub-branches of solutions. Since the equations are related by $\Lambda \to -\Lambda$ for the two main branches, we only have to explore one set of these 3 sub-branches. The first sub-branch gives again the above genus 1 curve (quadruply degenerate this time). The second sub-branch gives rise to 4 different genus zero curves and the last one consists of two doubly degenerate $r = 1$ curves.

1 **flavor with a cubic potential**

We consider the same potential as above and the e.o.m. give in this case 2 main branches of solutions. One main branch consists of two sub-branches. One of them numerically corresponds to 6 genus zero curves, whereas the other sub-branch contains (numerically) a root at approximately $x = -m$. However, we were unable to obtain an analytical expression for these last curves. The remaining main branch corresponds to two genus one curves, as expected.

A holomorphic coordinate patch for the baryonic root at $N_f = 3$

The alternative limit on the 2×2 Lax matrices displayed in (4.61) did not yield as much information in the case $N_c = 2, N_f = 2$ as the regular limits studied in all other cases above. In particular, it did contain the baryonic root in that case. Naively, there is no reason to suspect that this would be different for the case $N_f = 3$. However, with the limit 1^u on one Lax matrix and the alternative one on the second does yield different results. With the superpotential $W(x) = \frac{1}{2}x^2 + mx$ and equal masses m for the quarks we find that the theory has 3 (non-degenerate) different vacua with Seiberg-Witten curves

$$\begin{aligned} y^2 &= (m+x)^2(m+x-\Lambda)^2 \\ y^2 &= (m+x)^2((m+x)^2+4\Lambda^2) \\ y^2 &= (m+x-4\Lambda)^2((m+x)^2+4\Lambda^2). \end{aligned} \tag{4.71}$$

The most interesting thing about these solutions is that the baryonic first Seiberg-Witten curve is located at a *finite* point in moduli space

$$p_1 = -m, \qquad p_2 = -m, \qquad y = -1. \tag{4.72}$$

It would certainly be interesting to understand why in this case the baryonic branch and the $r = 1$ branch are covered by this particular limit and not in the case of

$N_f = 2$. Furthermore, it is slightly suspicious that we recover only three solutions and not the expected 5. This is because the $r = 1$ branch found in the alternative limit is not triply degenerate.

4.5.2 Gauge group $U(3)$

4 flavors

We study the limit $(1^u, 2, 1^d)$, since this case resembles the correct limit for $N_c = 2, N_f = 2$ and find $2+2+1$ vacua for $r = 0, 1, 2$ respectively. These vacua belong to the electric phase. Furthermore we find the direct (addition map) lift of the $N_c = 2$, $N_f = 2$ baryonic branch. It appears that we are missing the same (magnetic) vacua here as we did in the $N_c = 2$, $N_f = 3$ case. The $r = 2$ branch yields six times the same curve, the two $r = 1$ branches four each, and the $r = 0$ just has one solution for each curve, as expected from the field theory analysis.

2 flavors

We find 4 $r = 0$ and 2 branches of the solution using the $(2, 0, 0)$ limit. The $r = 1$ branch (which has two solutions for each curve) corresponds to the addition map applied to $N_c = 2, N_f = 0$, since it has the structure $(x + m)^2$ times the curve for $N_c = 2, N_f = 0$ with the same superpotential. The $(1^u, 1^d, 0)$ limit shows one extra phase for a solution of the e.o.m. which is on the edge of the phase space ($q_1 \to \infty$). This phase has Seiberg-Witten curve factorization

$$y^2 = (x+m)^2((x+m)^2 + 2\Lambda^2)((x+m)^2 - 2\Lambda^2). \tag{4.73}$$

which is a genus 1 curve. This is an example of a solution which is not expected from the gauge theory analysis which lies on the boundary of the phase space.

3 flavors

For this interesting case we checked several different limits. Just as in the $N_c = N_f = 2$ case we generically find the expected 3 $r = 0$ and 3 (with 3 solutions each) $r = 1$ electric branches. Only limits which contain a link like in the $N_f = 2 = N_c$ case, i.e. $(1^u, 1^d)$ also produce the baryonic root.

4.5.3 Gauge group $U(4)$

In order to check the multiplication and the addition map we study the four flavored theory. The multiplication map can be studied by taking the limit $(1^u, 1_d, 1^u, 1^d)$. This gives 4 phases at $r = 0$, 4 at $r = 1$, and 2 phases at $r = 2$. Actually, the phase at $r = 2$ has a $(x+m)^6$ piece in the Seiberg-Witten curve which suggests it is an (unphysical) $r = 3$ branch, but if we shift the mass in the superpotential from $\frac{1}{2}x^2 + mx$ to a slightly different mass $\tilde{m}$ we see that only 4 roots of the Seiberg-Witten curve remain at $x = -m$. In addition we find the multiplication map lift of the baryonic branch. However, in addition to these phases we also find a Seiberg-Witten curve of the form $((x+m)^8 + \Lambda^8)$ and one of the form $(m+y)^4((y+m)^2 - \Lambda^2)((y+m)^2 + \Lambda^2)$ which are genus 1 curves, both very similar to equation (4.73). Just as in that case, these solutions are on the edge of the phase space. In fact, the genus 1 curves seems to be the addition map lift of this 'wrong curve'. We checked also a different limit for this case, which yields the same phases with the exception of the baryonic root.

4.5.4 The case $N_f \leq N_c$

We take the limit $(0^{N_c - N_f}(1^u)^{N_f})$, N_c to be larger than to 2 and the superpotential to be $W(x) = \frac{1}{2}x^2 + mx$. We take the masses to be all unequal. We compute the superpotential,

$$W = \frac{1}{2}\sum_{i=1}^{N_c} p_i^2 + \sum_{i=1}^{N_f} \Lambda e^{q_{i-1}-q_i}(p_i + m_i) + \sum_{i=N_f+1}^{N_c} \Lambda^2 e^{q_{i-1}-q_i} + m\sum_{i=1}^{N_c} p_i. \quad (4.74)$$

To see this is indeed the correct expression for the quadratic potential note the following: for Lax matrices at a site which contains masses, the constant lower right matrix element only contributes to the trace of the transfer matrix at order x^{N_c-3}. Since a quadratic potential is needed that can be calculated from the coefficient of x^{N_c-2} and x^{N_c-1} that element can be safely ignored for calculational purposes. The general form of the quadratic superpotential can then be inferred by comparing to the Toda case. From equation (4.74) we easily obtain and solve the equations of motion for p,

$$p_i + m = -\Lambda e^{q_{i-1}-q_i} \quad i = 1, \ldots, N_f \quad (4.75)$$
$$p_i + m = 0 \quad i = N_f + 1, \ldots, N_c \quad (4.76)$$

Her the variable q_0 is defined as q_N. The equations of motion for q are a bit more involved,

$$\begin{array}{rcll} e^{q_{i-1}-q_i}(p_i+m_i) &=& e^{q_i-q_{i+1}}(p_{i+1}+m_{i+1}) & i=1,\ldots,N_f-1 \\ e^{q_{N_f-1}-q_{N_f}}(p_{N_f}+m_{N_f}) &=& \Lambda e^{q_{N_f}-q_{N_f+1}} & i=N_f \\ \Lambda e^{q_{i-1}-q_i} &=& \Lambda e^{q_i-q_{i+1}} & i=N_f+1\ldots N_c-1 \\ \Lambda e^{q_{N_c-1}-q_{N_c}} &=& e^{q_{N_c}-q_1}(p_1+m_1) & i=N_c, \end{array} \tag{4.77}$$

or using the solution for p and the definition $e^{q_{i-1}-q_i}=z_i$,

$$\begin{array}{rcll} z_i(m_i-\Lambda z_i-m) &=& z_{i+1}(m_{i+1}-\Lambda z_{i+1}-m) & i=1,\ldots,N_f-1 \\ z_{N_f}(m_{N_f}-\Lambda z_{N_f}-m) &=& \Lambda z_{N_f+1} & i=N_f \\ z_i &=& z_{i+1} & i=N_f+1\ldots N_c-1 \\ \Lambda z_{N_c} &=& z_1(m_1-\Lambda z_1-m) & i=N_c. \end{array} \tag{4.78}$$

The variables z_i obey the constraint $\prod z_i = 1$. At this point it is important to note that the equations of motion indeed have a solution, up to the constraint. Also, in this case all solutions give finite coordinates. Evaluated on the solutions for the momenta the transfer matrix becomes

$$T(x)=\Lambda^N\prod_{i=N_c}^{N_f+1}\begin{pmatrix} \frac{x}{\Lambda}+\frac{m}{\Lambda} & z_i \\ -1 & 0 \end{pmatrix}\prod_{j=N_f}^{1}\begin{pmatrix} \frac{x}{\Lambda}+z_j+\frac{m}{\Lambda} & z_j(\frac{m_j}{\Lambda}-z_j-\frac{m}{\Lambda}) \\ -1 & z_j \end{pmatrix}. \tag{4.79}$$

Here we have obtained the variables z_j by inserting '1' between Lax matrices L_iL_{i-1} as

$$\begin{pmatrix} 1 & 0 \\ 0 & e^{q_{i-1}} \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & e^{-q_{i-1}} \end{pmatrix}, \tag{4.80}$$

and realizing we are only interested in the eigenvalues of the transfer matrix.

Equal masses

First take all masses equal to m. Then N_f simple quadratic equations have to be solved, which are solved for N_f less than N_c as $z_i=\epsilon_i\sqrt{-z_{N_c}}$, with ϵ_i equal to either $+1$ or -1. The constraint now reads,

$$i^{N_f}\left(z_{N_c}\right)^{N_c-\frac{N_f}{2}}\prod_{i=1}^{N_f}\epsilon_i=1. \tag{4.81}$$

For even N_f this constraint can be solved straightforwardly for z_{N_c} for any choice of ϵ_i. We therefore have a total number of $2^{N_f}(N_c-N_f/2)$ solutions. When N_f is odd, we have to square the constraint which gives $z_{N_C}=(-1)^{\frac{1}{2N_c-N_f}}$. Inserting this back into the constraint then fixes one of the ϵ, arriving at the same number of solutions. On the basis of the field theory analysis we expect $\binom{N_f}{N_f/2}(N_c - N_f/2) + \sum_{r=0}^{N_f/2-1}\binom{N_f}{r}(2N_c - N_f)$ solutions for even N_f and $\sum_{r=0}^{N_f/2-1/2}\binom{N_f}{r}(2N_c - N_f)$ for odd N_f, which can easily be shown to be equal to the number of found solutions using $2^k = \sum_{i=0}^{k}\binom{k}{i}$. If N_f equals N_c the above analysis can be repeated by solving everything in terms of $z_i = \epsilon_i |z_{N_c}|$ (note that there are $N-1$ epsilons). Then we see that there are $2N_c$ solutions by squaring the constraint, and the constraint fixes one of the epsilons for the two possibilities $(z_{N_c})^{N_c} = \pm 1$. For the rest of this subsection we set the mass m for convenience to zero and consider $N_f < N_c$ since the case $N_f = N_c$ is a simple generalization of the argument below. The mass can be restored by shifting $x \to x + m$ and shifting $p \to p - m$ in the solution.

The above counting argument gives the total number of solutions for all r branches. Below we show how to obtain the factorization of the curve. Use the solution of the z_j in terms of the z_N and (cyclically) redistribute factors of $\sqrt{z_{N_c}}$ in (4.79) to get,

$$T(x) = \Lambda^N \prod_{i=N_c}^{N_f+1} \begin{pmatrix} \frac{x}{\Lambda} & \sqrt{z_{N_c}} \\ -\sqrt{z_{N_c}} & 0 \end{pmatrix} \prod_{j=N_f}^{1} \begin{pmatrix} \frac{x}{\Lambda} + i\epsilon_j\sqrt{z_{N_c}} & \sqrt{z_{N_c}} \\ -\sqrt{z_{N_c}} & i\epsilon_j\sqrt{z_{N_c}} \end{pmatrix}. \tag{4.82}$$

Define $\tilde{x} \equiv \frac{x}{\Lambda\sqrt{z_{N_c}}}$, so

$$T(\tilde{x}) = (\sqrt{z_{N_c}}\Lambda)^N \mathrm{Tr} \prod_{i=N_c}^{N_f+1} \begin{pmatrix} \tilde{x} & 1 \\ -1 & 0 \end{pmatrix} \prod_{j=N_f}^{1} \begin{pmatrix} \tilde{x} + i\epsilon_j & 1 \\ -1 & i\epsilon_j \end{pmatrix}. \tag{4.83}$$

This expression for the transfer matrix consists of building blocks

$$B \equiv \begin{pmatrix} \tilde{x} & 1 \\ -1 & 0 \end{pmatrix} \quad A^+ \equiv \begin{pmatrix} \tilde{x} + i & 1 \\ -1 & i \end{pmatrix} \quad A^- \equiv \begin{pmatrix} \tilde{x} - i & 1 \\ -1 & -i \end{pmatrix}. \tag{4.84}$$

These matrices share a basis of eigenvectors and commute. In fact,

$$A^+ \cdot A^- = \tilde{x} B \tag{4.85}$$

This immediately leads to the identification of the r-branches in this solution: we take a set of $N_f - r$ A^+'s and r A^-'s, $r \leq N_f$, z_N the appropriate solution of (4.81) and apply (4.85) to get the transfer matrix

$$T(\tilde{x}) = (\sqrt{z_{N_c}}\Lambda)^N (\tilde{x})^r B^{N_c - N_f + r} (A^{\pm})^{N_f + r}, \tag{4.86}$$

which will yield a spectral curve of the form $y^2 \sim \tilde{x}^{2r} G(\tilde{x})$. The $\pm$ depends on whether $r \leq [\frac{N_f}{2}]$ or not. From equation (4.81) we see that we have $N_c - \frac{N_f}{2}$ solutions of this type for even N_f and $r \leq \frac{N_f}{2}$. Therefore for $r < \frac{N_f}{2}$ there are

$2\binom{N_f}{r}(N_c - \frac{N_f}{2})$ solutions for each r. The factor of two comes from the fact that for fixed r both the solutions with r A^+'s as the solutions with r A^- contribute. For $r = \frac{N_f}{2}$ there are therefore only $N_c - \frac{N_f}{2}$ solutions. For odd N_f we have to take into account that there has to be either one A^+ or A^- more than the other. Then we have $\binom{N_f}{r}(2N_c - N_f)$ solutions at each r-level, so the counting of states works out level by level.

Now writing the building block matrices in an eigenvalue basis gives a transfer matrix with diagonal elements:

$$T_{11}(\tilde{x}) = (\sqrt{z_{N_c}}\Lambda)^N \left(\frac{\tilde{x}}{2} + \sqrt{(\frac{\tilde{x}}{2})^2 - 1}\right)^{N_c - N_f} \left(\frac{\tilde{x}}{2} + i + \sqrt{(\frac{\tilde{x}}{2})^2 - 1}\right)^{N_f - r} \left(\frac{\tilde{x}}{2} - i + \sqrt{(\frac{\tilde{x}}{2})^2 - 1}\right)^{r} \quad (4.87)$$

$$T_{22}(\tilde{x}) = (\sqrt{z_{N_c}}\Lambda)^N \left(\frac{\tilde{x}}{2} - \sqrt{(\frac{\tilde{x}}{2})^2 - 1}\right)^{N_c - N_f} \left(\frac{\tilde{x}}{2} + i - \sqrt{(\frac{\tilde{x}}{2})^2 - 1}\right)^{N_f - r} \left(\frac{\tilde{x}}{2} - i - \sqrt{(\frac{\tilde{x}}{2})^2 - 1}\right)^{r}. \quad (4.88)$$

The square of the difference between the diagonal elements is the discriminant of the characteristic equation of the 2 by 2 matrix, so is the Seiberg-Witten curve,

$$y^2 = (T_{11} - T_{22})^2. \quad (4.89)$$

By first reducing by using (4.85) it is easily seen that the remaining Seiberg-Witten curve has the form polynomial times $\sqrt{(\frac{\tilde{x}}{2})^2 - 1}$ and thus factorizes.

Motivated by [79] we show how to write this in term of linear combinations of Chebyshev polynomials. Chebyshev polynomials of first and second kind obey the identities

$$T_n(x) = \frac{1}{2}\left(\left(x + \sqrt{(x^2-1)}\right)^n + \left(x - \sqrt{(x^2-1)}\right)^n\right) \quad (4.90)$$

$$U_{n-1}(x)\sqrt{x^2-1} = \frac{1}{2}\left(\left(x + \sqrt{(x^2-1)}\right)^n - \left(x - \sqrt{(x^2-1)}\right)^n\right). \quad (4.91)$$

It is easily seen by expanding

$$\left(\frac{\tilde{x}}{2} + a \pm \sqrt{(\frac{\tilde{x}}{2})^2 - 1}\right)^r = \sum_{k=0}^{r} \binom{r}{k} a^{r-k} \left(\frac{\tilde{x}}{2} \pm \sqrt{(\frac{\tilde{x}}{2})^2 - 1}\right)^k, \quad (4.92)$$

that the trace of the transfer matrix is a linear combination of Chebyshev polynomials of the first kind, while the Seiberg-Witten curve is a linear combination of Chebyshev polynomials of the second kind times the expected non-singular piece.

Unequal masses

In this case far less specific information than above can be obtained since to obtain more concrete results in this case the constraint equation $\prod z_i = 1$ has to be solved, which is hard to do in general. However, note that inserting the solution for the z_i in term of z_{N_c} into equation (4.79) leads to a transfer matrix of the form

$$T(x) = \Lambda^N \prod_{i=N_c}^{N_f+1} \begin{pmatrix} \frac{x}{\Lambda} + \frac{m}{\Lambda} & z_{N_c} \\ -1 & 0 \end{pmatrix} \prod_{j=N_f}^{1} \begin{pmatrix} \frac{x}{\Lambda} + \tilde{z}_j + \frac{m}{\Lambda} & z_{N_c} \\ -1 & \tilde{z}_j \end{pmatrix}, \tag{4.93}$$

where $\tilde{z}_j$ are the roots of

$$z_j(\frac{m_j}{\Lambda} - z_j - \frac{m}{\Lambda}) = z_{N_c}. \tag{4.94}$$

Set $\hat{x} \equiv \frac{x+m}{\Lambda\sqrt{z_{N_c}}}$ and $\hat{z}_j = \frac{\tilde{z}_j}{\sqrt{z_{N_c}}}$ to obtain

$$T(x) = \left(\sqrt{z_{N_c}}\Lambda\right)^N \prod_{i=N_c}^{N_f+1} \begin{pmatrix} \hat{x} & 1 \\ -1 & 0 \end{pmatrix} \prod_{j=N_f}^{1} \begin{pmatrix} \hat{x} + \hat{z}_j & 1 \\ -1 & \hat{z}_j \end{pmatrix}, \tag{4.95}$$

The building block matrices in this transfer matrix all commute so can be diagonalized simultaneously, as before. The eigenvalues now differ from previous cases only by the value of a in equation (4.92) since the differences in the Lax matrices between the cases with and without equal masses are proportional to the identity matrix. Hence we can (again) write the spectral curves of these transfer matrices in terms of Chebyshev polynomials and the spectral curve factorizes in the appropriate manner,

$$y^2_{\text{diff mass}} = H^2 \sqrt{\left(\frac{\hat{x}}{2}\right)^2 - 1}. \tag{4.96}$$

From the point of view of gauge theory it is interesting that we find a concrete realization of the polynomials studied in [79] generating the factorization for both equal and unequal masses for the full range of $N_f \leq N_c$.

4.5.5 All massive vacua for $N_f \leq N_c$

In line with the calculation in chapter 3 we show how to generalize the above computation to the massive vacua of an arbitrary (single trace) polynomial superpotential for $N_f \leq N_c$. Since the factorization into a genus zero curve shown above is very special, the solution of the integrable system equations of motion shown is a natural starting point to look for a solution for all polynomial superpotentials

$W = \sum_{k=1}^{n+1} g_k \frac{x^k}{k}$. From the gauge theory point of view for massive vacua the Seiberg-Witten curve must factorize into a genus zero curve which we take to be

$$P^2 - 4Q = H^2 \left(\left(\frac{x+c}{\sqrt{z_{N_c}}\Lambda} \right)^2 - 4 \right). \tag{4.97}$$

According to equation (4.16) there must then be polynomials G and f such that

$$\left(\left(\frac{x+c}{\sqrt{z_{N_c}}\Lambda} \right)^2 - 4 \right) G_{n-1}^2(x) = W_n'(x)^2 + f_{n-1}(x) \tag{4.98}$$

holds. Since this factorization problem is exactly the same as in the pure glue case which was analyzed in chapter 3 we only state here that for such polynomials to exist, the constant c must be such that the coefficient c_0 in

$$W'(\sqrt{z_{N_c}}\Lambda(\xi + \frac{1}{\xi}) - c) = \sum_{i=-n}^{n} c_i \xi^i \tag{4.99}$$

is zero. We will first show below that a slight generalization of that solution indeed solves the equations of motion up to one equation which we then show to be equal to the equation expected from the gauge theory.

We take our solution of the equations of motion in (4.36) to be equations (4.76) and (4.78) together with the constraint $\prod z_i = 1$, where in equation (4.76) we replace m by the variable c. We have already shown that solutions of this type yield factorized Seiberg-Witten curves with non-singular part as in equation (4.98), but we still have to show that these solutions actually solve the equations of motion.

First change variables as $x = \sqrt{z_{N_c}}\Lambda\tilde{x} - c$ in the integral of equations (4.36) above. The equations of motion for the q_i are satisfied because the appropriate derivative of the trace of the transfer matrix evaluated on the equation of motion gives (after some manipulations)

$$\frac{\delta T_N}{\delta q_i} = \sqrt{z_{N_c}}\Lambda \begin{pmatrix} 0 & -\sqrt{z_{N_c}} \\ -\sqrt{z_{N_c}} & 0 \end{pmatrix} T_{N-1}, \tag{4.100}$$

where T_{N-1} is the transfer matrix obtained by eliminating Lax matrix L_i, evaluated on the solution to the equation of motion. In particular, T_{N-1} is diagonalized a known basis of eigenvectors. In this basis the other matrix becomes

$$\begin{pmatrix} 0 & -\sqrt{z_{N_c}}\Lambda \\ -\sqrt{z_{N_c}}\Lambda & 0 \end{pmatrix} \to \begin{pmatrix} 0 & \frac{1}{2}(x + \sqrt{x^2-4}) \\ \frac{1}{2}(x - \sqrt{x^2-4}) & 0 \end{pmatrix}. \tag{4.101}$$

Therefore the trace of equation (4.100) vanishes and the e.o.m. for the q's are satisfied.

The e.o.m. for the momenta can be derived from the transfer matrix

$$\frac{\delta T_N}{\delta p_i} = \begin{pmatrix} -1 & z_i \\ 0 & 0 \end{pmatrix} T_{N-1}. \tag{4.102}$$

Changing basis to the basis of eigenvectors of the other Lax matrices in the transfer matrix (and some manipulations) yields

$$\begin{pmatrix} -1 & \hat{z}_i \\ 0 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} \frac{\hat{z}_i+\frac{1}{2}(\tilde{x}-\sqrt{\tilde{x}^2-4})}{\sqrt{\tilde{x}^2-4}} & \frac{\hat{z}_i+\frac{1}{2}(\tilde{x}+\sqrt{\tilde{x}^2-4})}{\sqrt{\tilde{x}^2-4}} \\ \frac{-\hat{z}_i+\frac{1}{2}(-\tilde{x}+\sqrt{\tilde{x}^2-4})}{\sqrt{\tilde{x}^2-4}} & \frac{-\hat{z}_i+\frac{1}{2}(-\tilde{x}-\sqrt{\tilde{x}^2-4})}{\sqrt{\tilde{x}^2-4}} \end{pmatrix}. \tag{4.103}$$

Therefore the trace of equation (4.102) becomes

$$\operatorname{Tr}\frac{\delta T_N}{\delta p_i} = \frac{(2\hat{z}_i+\tilde{x})(T^{11}_{N-1}-T^{22}_{N-1}) - \sqrt{\tilde{x}^2-4}(T^{11}_{N-1}+T^{22}_{N-1})}{\sqrt{\tilde{x}^2-4}}, \tag{4.104}$$

with

$$T_{N-1} = \begin{pmatrix} T^{11}_{N-1} & 0 \\ 0 & T^{22}_{N-1} \end{pmatrix}. \tag{4.105}$$

Now we can also express the Seiberg-Witten curve on the solution in terms of the transfer matrix T_{N-1} using the fact that on that solution

$$T_N = \sqrt{z_{N_c}}\Lambda \begin{pmatrix} \frac{\hat{z}_i+\frac{1}{2}(\tilde{x}-\sqrt{\tilde{x}^2-4})}{\sqrt{\tilde{x}^2-4}} & 0 \\ 0 & \frac{\hat{z}_i+\frac{1}{2}(\tilde{x}+\sqrt{\tilde{x}^2-4})}{\sqrt{\tilde{x}^2-4}} \end{pmatrix} \begin{pmatrix} T^{11}_{N-1} & 0 \\ 0 & T^{22}_{N-1} \end{pmatrix}. \tag{4.106}$$

As explained above, the Seiberg-Witten curve is the difference between the diagonal values, which gives something proportional to equation (4.104). We therefore arrive at

$$\left.\frac{\frac{\delta P}{\delta p_i}}{\sqrt{P^2-4Q}}\right|_{\text{solution}} = \frac{1}{\sqrt{z_{N_c}}\Lambda\sqrt{\tilde{x}^2-4}}. \tag{4.107}$$

Hence the equations of motion for all the momenta collapse into 1 equation. Combining this computation with the shift in x and changing variables $x \rightarrow \xi + \frac{1}{\xi}$ yields

$$\oint W'(\sqrt{z_{N_c}}\Lambda(\xi+\frac{1}{\xi})-c)\frac{1}{\xi}d\xi = 0, \tag{4.108}$$

which is equivalent to equation (4.99) expected from the gauge theory point of view. We have therefore proven that this particular solution solves the equations of motion for the integrable system and is such that equation (4.16) holds. Note that this holds for all $N_f \leq N_c$ and for arbitrary masses.

As a technical aside, note that in the Toda case the derivatives of the transfer matrix with respect to the momenta can be summed such that on the solution the following equation holds:

$$\sum_i \frac{\partial P}{\partial p_i} = -\frac{\partial P}{\partial x}. \tag{4.109}$$

This reproduces, when evaluated in the resolvent, exactly the condition in chapter 3 that $\mathrm{Tr}\, W'(M)$ should vanish, where M is the periodic Toda chain Lax operator.

It is instructive to realize at this point that the factorized Seiberg-Witten curves obtained in this subsection should be the full field theory answer to the factorization problem for massive vacua of gauge theories with fundamental matter with $N_f \leq N_c$. In particular, they account for all possible vacua in a unified way. However, the above construction is constructive in the sense that we do not prove that these are all possible factorizations, i.e. that there are no others. On the other hand, they are all possibilities that are expected from the weak coupling point of view by a simple counting argument and we expect this to cover all field theory possibilities. To compare to other approaches to these particular gauge theories we calculate in the final subsection the vevs for the adjoint scalar in these vacua.

4.5.6 Adjoint scalar vevs for massive vacua

The vacuum expectation values for the adjoint scalar in the massive vacua discussed above are given by

$$< u_k >= -\frac{1}{2\pi i} \oint x^{k-1} \log\left(\frac{(P(x) + \sqrt{P^2 - 4Q})}{x^{N_c}} \right) dx. \tag{4.110}$$

For the massive vacua

$$P(x) + \sqrt{P^2 - 4Q} = 2T_{11} = 2(\sqrt{z_{N_c}}\Lambda)^{N_c} \prod_{i=1}^{N_c} \left(\frac{\tilde{x}}{2} + z_i + \sqrt{\left(\frac{\tilde{x}}{2}\right)^2 - 1} \right), \tag{4.111}$$

with $\tilde{x} = \frac{x+c}{\sqrt{z_{N_c}}\Lambda}$. The expectation values can be split into a part which also appears in the pure glue case and extra terms depending on the flavors,

$$\begin{aligned} < u_k > &= < u_k >_{\text{pure glue}} + < u_k >_{\text{matter}} \\ < u_k >_{\text{pure glue}} &= -\frac{1}{2\pi i} \oint x^{k-1} N_c \log\left(\frac{\sqrt{z_{N_c}}\Lambda \left(\frac{\tilde{x}}{2} + \sqrt{\left(\frac{\tilde{x}}{2}\right)^2 - 1} \right)}{x} \right) dx \\ < u_k >_{\text{matter}} &= -\frac{1}{2\pi i} \oint x^{k-1} \log\left(1 + \frac{z_i}{\frac{\tilde{x}}{2} + \sqrt{\left(\frac{\tilde{x}}{2}\right)^2 - 1}} \right) dx. \end{aligned} \tag{4.112}$$

Calculating the glue term first we get, by splitting the logarithm into

$$\log\left(1 + \frac{c}{x}\right) + \log\left(\frac{1}{2} + \sqrt{\frac{1}{4} - \frac{1}{\tilde{x}^2}} \right), \tag{4.113}$$

two finite integrals which can be performed. The first vev is easily calculated to be

$$< u_1 >_{\text{pure glue}} = -c\, N_c. \tag{4.114}$$

The needed contour integral of both logarithms in equation (4.113) can be evaluated, with the help of the expansion[4]

$$\log\left(\frac{1}{2} + \sqrt{\frac{1}{4} - \frac{1}{\tilde{x}^2}}\right) = -\sum_{i=1}^{\infty} \frac{1}{2i} \binom{2i}{i} \frac{1}{\tilde{x}^{2i}}, \tag{4.115}$$

to give

$$< u_k >_{\text{pure glue}} = \frac{N_c}{k} \sum_{l=0}^{\left[\frac{k}{2}\right]} \binom{k}{2l} \binom{2l}{l} (z_{N_c}\Lambda^2)^l (-c)^{k-2l}. \tag{4.116}$$

The matter contribution can be calculated by using the change of variables $\tilde{x} \to \xi + \frac{1}{\xi}$. As an intermediate result

$$-\frac{1}{2\pi i} \oint_{\xi=\infty} (\xi + \frac{1}{\xi})^j (1 - \frac{1}{\xi^2}) \log\left(1 + \frac{z_i}{\xi}\right) = \sum_{m=0}^{\left[\frac{j}{2}\right]} \frac{1}{j-m+1} \binom{j}{m} (-z_i)^{j-2m+1} \tag{4.117}$$

is obtained by straightforward computation. Plugging this result into the needed integral we arrive at

$$< u_k >_{\text{matter}} = (\sqrt{z_{N_c}}\Lambda)^k \sum_{i=1}^{N_f} \sum_{j=0}^{k-1} \sum_{m=0}^{\left[\frac{j}{2}\right]} \binom{k-1}{j} \binom{j}{m} \frac{(-z_i)^{j-2m+1}}{j-m+1} \left(\frac{-c}{\sqrt{z_{N_c}}\Lambda}\right)^{k-1-j}. \tag{4.118}$$

The full answer for the adjoint scalar vev in the massive vacua is now the sum of equations (4.116) and (4.118). Comparing above expressions to the ones obtained by a matrix model calculation in [80] we immediately identify

$$c \to -T \qquad z_{N_c}\Lambda^2 \to R. \tag{4.119}$$

Their constraint is exactly the constraint $\prod z_i = 1$. This can be seen by multiplying the two different solutions to equation (4.94) for each site with a mass. For the solutions of that equation a consistent sign has to be chosen to connect to the matrix model result. With these identifications the relation to equations (67), (68), (69) and (70) in [80] is obvious. The relation with the remaining equality (71) is more obscure, but for $k = 2$ there is perfect agreement. It would certainly be interesting to check the matrix model calculation further against the above field theory result.

[4]Take a derivative on the left hand side w.r.t. $y \equiv \frac{1}{x}$ and observe that this is $-\frac{C(y)}{\sqrt{1-4y}}$ with $C(y)$ the generating function for the Catalan numbers.

4.5.7 Toward a general proof of factorization

In this final subsection we indicate how a proof of factorization of the Seiberg-Witten curve can be obtained up to a technical detail. In chapter 3 we relied on a theorem in [58] which relates straight line flows on the Jacobian of the spectral curve to the equations of motion of the periodic Toda chain. This connection is lacking in the case of the classical spin chain, although the fact that it is integrable and periodic strongly suggests that the reasoning should hold in exactly the same way. In this chapter however, we are mainly interested in the critical points of the potential as can be obtained by using equation (4.36), reproduced here for convenience,

$$\frac{\delta W}{\delta o_i} = -\frac{1}{2\pi i}\oint W'(x)\frac{\frac{\delta P}{\delta o_i}}{\sqrt{P^2-4Q}}.$$

These equations can be written as

$$\frac{\delta W}{\delta o_i} = \frac{1}{2\pi i}\sum_j \frac{\delta s_j}{\delta o_i}\oint W'(x)\frac{x^{N_c-j}}{\sqrt{P^2-4Q}}. \tag{4.120}$$

Assuming that the transformation Jacobian has rank $N-1$, the equations of motion amount to the vanishing of certain one forms on the Riemann-surface, a condition actually identical to the conditions derived in chapter 3. Therefore, modulo the technical assumption, this amounts to a proof of factorization as in equation (4.16) of the spectral curve of the integrable system in exactly the same way as derived from the gauge theory. Since in the quadratic superpotential case we found all gauge theory vacua, it is expected for the general superpotential case which can be though of as being built out of a number of quadratic potentials that all expected gauge theory vacua will be found.

4.6 Integrating in the mesons

A natural question in the integrable framework is whether or not one can 'integrate in' the mesons, just as the glueball field can be integrated in. A natural reason to suspect that this might be possible is the fact that the solutions we found for $N_f < N_c$ are calculationally very similar to vacua of the potential (4.11). The gauge theory interpretation of the solutions of both superpotentials is also similar. In addition, pure field theory arguments indicate that it might be possible to integrate in the mesons and even the baryons [22]. Finally, in [81] it was shown that the minima of the ADS potential and the minimum of the matrix model potential agree for a quadratic potential. In this section we show how to directly integrate in the meson fields for a quadratic potential in the case $N_f \leq N_c$ and how this reproduces both four and three dimensional results simultaneously.

The quadratic XXX superpotential (obtained from equation (4.74), reproduced here for convenience) is for the limit 1^u on the first N_f sites

$$W = \mu \left(\sum_{i=1}^{N_c} p_i^2 + 2 \sum_{i=1}^{N_f} \Lambda e^{q_{i+1}-q_i}(p_i + m_i) + 2 \sum_{i=N_f+1}^{N_c} \Lambda^2 e^{q_{i+1}-q_i} \right).$$

Here q_{N+1} is identified with q_1. A convenient way to rewrite this is in terms of new variables $y_i = \Lambda e^{q_{i+1}-q_i}(p_i + m_i)$. These satisfy one constraint which we impose using a Lagrange multiplier S, and we get:

$$W = \mu \sum_{i=1}^{N_c} (p_i^2 + 2y_i) + S \log \left(\frac{\Lambda^{2N_c - N_f} \prod_{j=1}^{N_f} (p_j + m_j)}{\prod y_j} \right). \tag{4.121}$$

In this equation Λ only appears in the logarithm and it couples to S. Therefore, S is to be identified with the familiar glueball field of pure $\mathcal{N} = 1$ theories. We now introduce new Lagrange multipliers M_i for the masses and new fields $\tilde{m}_i$ and we rewrite (4.121) as

$$W = \mu \sum_{i=1}^{N_c} (p_i^2 + 2y_i) + S \log \left(\frac{\Lambda^{2N_c - N_f} \prod_{j=1}^{N_f} (p_j + \tilde{m}_j)}{\prod y_j} \right) + \sum_{i=1}^{N_f} (m_i - \tilde{m}_i) M_i. \tag{4.122}$$

If we integrate out M we regain (4.121). On the other hand, we can also integrate out p_i, $\tilde{m}_i$ and all y except y_{N_c} and we find

$$W = \sum_{i=1}^{N_f} \left(-\frac{M_i^2}{2\mu} + m_i M_i \right) + 2\mu y_{N_c} + (N_c - N_f - 1)S + S \log \left(\frac{\Lambda^{2N_c - N_f} 2^{N_c} \mu^{N_c}}{y_{N_c} S^{N_c - N_f - 1} \prod M_j} \right). \tag{4.123}$$

Integrating out S gives

$$W = \sum_{i=1}^{N_f} \left(-\frac{M_i^2}{2\mu} + m_i M_i \right) + 2\mu y_{N_c} + (N_c - N_f - 1) \left(\frac{\Lambda^{2N_c - N_f} 2^{N_c} \mu^{N_c}}{y_{N_n} \prod M_j} \right)^{\frac{1}{N_c - N_f - 1}}. \tag{4.124}$$

From now on we identify in the above expression $\prod M_j$ with the determinant of the meson matrix. In order to compare to the three-dimensional $\mathcal{N} = 2$ theories studied in [35], we want to take the limit $\mu \to \infty$, keeping $\Lambda^{2N_c - N_f} \mu^{N_c}$ fixed by scale matching. Furthermore we set the quark masses to zero and arrive at the superpotential for the four dimensional gauge theory we study, compactified on a circle [35],

$$W = Y + (N_c - N_f - 1) (Y \det M)^{\frac{1}{N_c - N_f - 1}}, \tag{4.125}$$

with a convenient normalization. If $N_f = N_c - 1$, then integrating out S in the final step yields the right quantum constrained moduli space.

Another direction from equation (4.123) is integrating out y_{N_c} and S, which yields equation (4.11) upon putting the masses to zero and trivially redefining μ as 2μ. In this case we can also treat $N_f = N_c$ and arrive at the correct quantum constrained moduli space. Baryons can be integrated in by using the procedure described in one of the examples in [22], but this does not seem to yield illuminating results.

$N_c < N_f < 2N_c$

In principle equation (4.122) can be applied in this case as well by redefining the y variables appropriately. However, it is difficult to obtain a concrete general expression for the superpotential in this case, although this can still be done for $N_f = N_c + 1$. For the low order cases checked (up to $N_c = 3, N_f = 5$) the answer is exactly the potential expected for the electric branches of the theory.

4.7 Discussion

We have shown in this chapter that the XXX spin chain provides good holomorphic coordinates for at least part of the moduli space of $U(N)$ gauge theory with N_f fundamentals compactified on $\mathbb{R}^3 \times S^1$. Some coordinate patches touch the baryonic root at the edge of phase space, but there are sometimes other solutions to the equations of motion on this edge as well. These other solutions do not have a gauge theory interpretation. Being on the edge neither class is a bona fide vacuum of our superpotential, but it is certainly suggestive and it would be interesting to understand this further. We have shown in one specific case that there is a coordinate patch which includes the baryonic root holomorphically. Furthermore, we have given some evidence that the magnetic vacua may be covered by XXX coordinates in some cases. It remains an open question exactly what coordinate patch in the XXX chain covers what part of moduli space. Since the coordinates provide a coordinate patch for part of the moduli space it is natural to wonder whether there is a more general integrable system which provides coordinate patches for the other parts of moduli space, or even coordinates that are valid everywhere. It might be worthwhile to see if there is a more general formulation of the spin chain starting from the formulation in $SL(2, \mathbb{C})$ variables.

We have also shown that it is natural to integrate in the mesons into the superpotential by Legendre transforming with respect to the masses. The baryonic operators cannot be naturally integrated in and it remains an open question whether or not there is an extension of the integrable system in which there is a parameter multiplying the baryon. This parameter would then correspond to the variable b in [22] and could be used to integrate in the baryon. More generally it remains a question whether the meson and dyon vevs have an interpretation in the integrable system. An avenue of approach to this problem starts with noticing that the meson vevs are simply the value of the matrix model resolvent at $x = -m_i$. This gives at least on shell a method of backtracking to the integrable system what these vevs are.

Yet another open problem remains the microscopic interpretation of the phase space variables p_i and q_i in the three dimensional gauge theory. A way to examine this issue could be to study perturbative instanton contributions in the gauge theory to find the coordinate transformation of our results to the right microscopic variables, just as was used in chapter 3.

As indicated before, we expect that the factorization of the Seiberg-Witten curve on the minima of our superpotentials which is expected on field theory grounds can be proven along the lines of the second proof in chapter 3. We have indicated above how such a proof works in principle, but it would be interesting to close the small technical gap in that proof. Interesting generalizations of results in this chapter include but are not limited to studying other gauge groups, for which the integrable system is, to our knowledge, unknown and/or to higher dimensional gauge theories.

Appendix A Heisenberg variables

In this appendix we will rewrite (4.27) in terms of the Heisenberg variables for which $\{q_i, p_i\} = 1$. The aim will be to obtain something which resembles equation (4.20). We will drop the indices on the q's, p's and S's to declutter notation. First rewrite the algebra as

$$\{S_0, S_\pm\} = \pm S_\pm \qquad \{S_+, S_-\} = 2S_0. \tag{4.126}$$

We will begin with the Ansatz $S_0 \equiv -p - \lambda_i$ and read (4.126) as a system of differential (derivational) equations. The first can be easily solved by taking,

$$S_\pm = f_\pm(p) e^{\pm q}, \tag{4.127}$$

where the functions $f_\pm$ contain a constant of integration. Putting this into the second equation we obtain,

$$f_+(p)e^{-q}\{e^q, f_-(p)\} + e^q f_-(p)\{f_+(p), e^{-q}\} = 2(-p - \lambda_i). \tag{4.128}$$

Take the functions $f_\pm$ to be polynomials of degree 2, $f_\pm = a_\pm + b_\pm p + c_\pm p^2$. This yields three different solutions,

$$\left\{ \begin{array}{lcl} f_- &=& a_- \frac{2\lambda_i}{a_+} p - \frac{1}{a_+} p^2 \\ f_+ &=& a_+ \end{array} \right. \tag{4.129}$$

$$\left\{ \begin{array}{lcl} f_- &=& a_- + b_- p \\ f_+ &=& (-2\frac{\lambda_i}{b_-} + \frac{a_-}{b_-^2}) - \frac{1}{b_-} p \end{array} \right. \tag{4.130}$$

$$\left\{ \begin{array}{lcl} f_- &=& a_- \\ f_+ &=& a_+ - \frac{2\lambda_i}{a_-} p - \frac{1}{a_-} p^2 \end{array} \right. \tag{4.131}$$

The constants can be connected to λ_i and K_i through $S_0^2 + S_+ S_- = K_i^2$, which yields,

$$\left\{ \begin{array}{lcl} f_- &=& -\frac{\lambda_i^2 - K_i^2}{a_+} - \frac{2\lambda_i}{a_+} p - \frac{1}{a_+} p^2 \\ f_+ &=& a_+ \end{array} \right. \tag{4.132}$$

$$\left\{ \begin{array}{lcl} f_- &=& b_-(\pm K_i + \lambda_i) + b_- p \\ f_+ &=& -(\frac{\mp K_i + \lambda_i}{b_-}) - \frac{1}{b_-} p \end{array} \right. \tag{4.133}$$

$$\left\{ \begin{array}{lcl} f_- &=& a_- \\ f_+ &=& \frac{K_i^2 - \lambda_i^2}{a_-} - \frac{2\lambda_i}{a_-} p - \frac{1}{a_-} p^2 \end{array} \right. \tag{4.134}$$

where in the second line 2 solutions are possible (which differ only in which mass is in which off-diagonal entry).

Appendix B Seiberg-Witten curve factorization in the strong coupling approach

In this appendix it is shown how to calculate factorization loci of the Seiberg-Witten curve when turning on a classical superpotential in the gauge theory approach, particularly when the superpotential is of high degree ($W \sim \mathrm{Tr}\,\Phi^{n+1}$, $n \geq 2N_c - N_f$). In this appendix single trace superpotentials are studied for $N_f < 2N_c$.

The theory at the root of the r-th Higgs branch [67] is a $U(r) \times U(1)^{N_c - r}$ gauge theory, with N_f massless hypermultiplets charged under the $U(1)$ subgroups of $U(r)$. On special points along the root $l - r$ extra massless hypermultiplet degrees of freedom appear which are charged under the $U(1)$ subgroups. At these points the Seiberg-Witten curve factorizes as

$$P(x)^2 - 4Q(x) = (x-m)^{2r} \prod_{i=1}^{l-r} (x-p_i)^2 \prod_{j=1}^{2N-2l-2r} (x-q_j) \equiv H_l(x)^2 F_{2N-2l}(x). \tag{4.135}$$

Here $P(x) = \sum_{i=0}^{N_c} s_i x^{N_c-i}$ and $Q \equiv \Lambda^{2N_c-N_f}\left(\sum_{i=0}^{N_f} \tilde{m}_i x^{N_f-i}\right)$.

We will denote the extra massless dyons (and their scalar components) by d_i and $\tilde{d}_i$. Note that since r eigenvalues of Φ are fixed we are (classically) left with a $U(N_c-r)$ gauge group. For this theory we can write the superpotential

$$W = \sqrt{2}\sum_{i=1}^{l-r} \tilde{d}_i a_i d_i + \sum_{k=1}^{n+1} g_k u_k. \tag{4.136}$$

Since the theory is classically (for a generic point on the Higgs branch part of the vacuum) broken to a $U(N_c-r)$ gauge theory with N_f-2r flavors, the u_k correspond to multi-trace operators for $k > N_c - r$. Therefore, for $n < N_c - r$ the analysis below works[5].

There is an ambiguity in the definition of the operators u_k, since for higher values of k there can be several quantum operators with the same classical limit. We will denote the 'classical' u_k by $u_k^{(c)}$ in the following. In other words, $u_k^{(c)} = \frac{1}{k}\sum_i a_i^k$ for $P(x) = \prod_i (x-a_i)$. However, for the u_k appearing in the superpotential (4.136) we will use the quantum operators whose vacuum expectation values are obtained from the gauge theory resolvent. The quantum vevs $u_k^{(q)}$ can be expressed as

$$u_k^{(q)} = \frac{1}{2\pi i}\oint_{v=0} dv v^{-k-1}\left(\ln\left(P(1/v)\right) + \ln\left(1+\sqrt{1-4\frac{Q(1/v)}{P(1/v)^2}}\right)\right). \tag{4.137}$$

The first logarithm gives for $k \leq N_c$ exactly the Newton-Girard relations between power sums and the s_i, which can be written as symmetric polynomials of the eigenvalues. The second term produces the quantum corrections. Clearly, in the limit $\Lambda \to 0$, $Q \to 0$ and $u_k^{(q)} \to u_k^{(c)}$ as expected. We define $u_k^{(q)} = u_k^{(c)} + \Delta u_k$.

Before proceeding to the general case, we first calculate the first few corrections. Since only residues at $v = 0$ are needed, the quantum contributions Δu_k from the second logarithm can be expanded as

$$\Delta u_k = \frac{1}{2\pi i}\oint_0 dv v^{-k-1}\left(v^{2N_c-N_f}\frac{v^{N_f}Q(1/v)}{(v^{N_c}P(1/v))^2} + \frac{3}{2}v^{4N_c-2N_f}\left(\frac{v^{N_f}Q(1/v)}{(v^{N_c}P(1/v))^2}\right)^2 + \mathcal{O}(v^{6N_c-3N_f})\right). \tag{4.138}$$

The fraction in the integrand can now be Taylor-expanded around $v = 0$ since it does not contain poles any more. The quantum $u_k^{(q)}$ and the classical $u_k^{(c)}$ differ only

[5] For $n = N_c - r$ we do not have to add dyons in the case the Seiberg-Witten curve does not degenerate further. In this case the superpotential can be minimized by writing it first in terms of the $N_c - r$ eigenvalues. Then they can be taken to be equal to the $N - r$ different roots of W' such that the non-singular part of the curve is $W'_{N_c-r}(x)^2 + f_{N_f-2r}$. For $n \leq 2N_c - N_f$ this has the expected form of $W'(x) + f_{n-1}$. A more refined argument based on the gauge theory resolvent is also possible.

when $k \geq 2N_c - N_f$ and the first two corrections are

$$k = 2N_c - N_f \qquad \Delta u_k = \Lambda^{2N_c - N_f} \tag{4.139}$$

$$k = 2N_c - N_f + 1 \quad \Delta u_k = \Lambda^{2N_c - N_f}\left(\tilde{m}_1 + 2s_1\right) - \frac{3}{2}\Lambda^{4N_c - 2N_f}\delta_{2N_c - N_f, 1}. \tag{4.140}$$

The point of this calculation is that it matters which superpotential is used in (4.136) in a calculation along the lines of [18] to calculate Seiberg-Witten curve factorization loci. We will do this below for a (single trace) superpotential derived from the gauge theory resolvent. Equation (4.136) may be varied with respect to $u_k^{(q)}$ to obtain these. However, in that calculation we would need $\frac{\delta s_j}{\delta u_k^{(q)}}$. This is hard to calculate since the s_j are highly non-trivial functions of the $u_k^{(q)}$. It is easier to vary equation (4.136) with respect to the classical moduli, since part of the calculation is then standard. This yields

$$-\sum_{l=0}^{n+1} g_l \frac{\delta(u_l^{(c)} + u_l^{(q)})}{\delta u_k^{(c)}} = \sqrt{2}\sum_{i=1}^{N_c - r} \tilde{d}_i d_i \frac{\delta a_i}{\delta u_k^{(c)}}. \tag{4.141}$$

Repeating the steps in [18] gives,

$$\tilde{W}_n'(x) = B_q \frac{P(x)}{H_l(x)} + \mathcal{O}(\frac{1}{x}), \tag{4.142}$$

with

$$\tilde{W}'(x) \equiv \sum_{r=1}^{n+1} g_r x^{r-1} + \sum_{l=0}^{n+1}\sum_{r=1}^{n+1} g_l \frac{\delta u_l^{(q)}}{\delta u_r^{(c)}} x^{r-1}. \tag{4.143}$$

B_q is a polynomial[6] of degree less than or equal to $l-1$ and to get the neat first term in $\tilde{W}'(x)$ we have used the assumption that the degree of the potential is less than or equal to N_c. By comparing leading powers in x in equation (4.142), it follows that $N_c - n \leq l \leq N_c - 1$. Squaring equation (4.142) now gives an equation akin to equation (3.44) in [69],

$$\tilde{W'}_n^2(x) = B_q^2 F_{2n-2q} + f_{n-1}(x) + \frac{4Q(x)B_{l-1}^2}{H_{2l}^2(x)} + \mathcal{O}(\frac{1}{x}). \tag{4.144}$$

This equation indicates that $B_q^2 F_{2n-2q} = W_n'^{(c)}(x)^2 + g_{2n-2N_c+N_f}$ with g a polynomial of the indicated order. From other considerations we expect to find in the end $W'^{(c)}(x)^2 + \tilde{f}_{n-1}$ and below we show that this is actually the case if the extra contributions on both left and right hand side of this equation are taken into account.

[6]In [18] this polynomial was denoted by $H(v)$.

The left hand side of the equation can be rewritten in terms of the classical superpotential and its quantum contributions using equation (4.143). Using the obvious definition $\tilde{W}' \equiv W_n'^{(c)} + W'^{(q)}$ the left hand side of the equation reads

$$(W'^{(c)})^2 + 2W'^{(c)}W'^{(q)}(x) + (W'^{(q)}(x))^2. \tag{4.145}$$

Note that the extra terms on the right hand side appear at the same order as those on the left hand side. The extra terms on the left can be calculated by starting with

$$\frac{\delta u_l^{(q)}}{\delta u_r^{(c)}} = \frac{-1}{2\pi i}\oint_{x'=\infty} x'^{l-1} \frac{1}{\sqrt{1-\frac{4Q}{P^2}}+1-\frac{4Q}{P^2}}\left(\frac{4Q}{P^2}\right)\frac{\tilde{P}}{P}dx', \tag{4.146}$$

where $\tilde{P}$ denotes $\frac{\delta P_r}{\delta u_r^{(c)}}$. The last term in the integrand can be expressed as,

$$\frac{\tilde{P}}{P} = -\frac{\sum_{j=0}^{N_c} s_{j-r}x^{N_c-j}}{\sum_{k=0}^{N_c} s_k x^{N-k}}, \tag{4.147}$$

which gives after some massaging,

$$\frac{\tilde{P}}{P} = -x^{-r}(1-\frac{\sum_{j=N_c-r+1}^{N_c} s_j x^{-j}}{\sum_{k=0}^{N_c} s_k x^{-k}}). \tag{4.148}$$

The second term above is suppressed in the integrand by a factor of x^{-N_c-1} and will therefore not contribute to the residue integral as long as the order of the superpotential n does not exceed N_c. Plugging this result back into $W'^{(q)}(x)$ yields

$$W'^{(q)}(x) = \sum_{l=0}^{n+1}\sum_{r=1}^{n+1} g_l x^{r-1}\frac{1}{2\pi i}\oint_{x'=\infty} x'^{l-r-1}\frac{1}{\sqrt{1-\frac{4Q}{P^2}}+1-\frac{4Q}{P^2}}\left(\frac{4Q}{P^2}\right)dx'. \tag{4.149}$$

For functions of the form $f(x) = \sum_{i=0}^{\infty} a_i x^{p-i}$

$$f(x)|_+ = \sum_{i=0}^{p}\frac{x^{p-i}}{2\pi i}\oint_{x'=\infty} x'^{i-p-1}f(x')dx' \tag{4.150}$$

holds where $f(x)|_+ = \sum_{i=0}^{p} a_i x^{p-i}$ and the 'residua at infinity integral' is taken to be the same as in equation (2.45). Defining $r-1 = n-j$,

$$W'^{(q)}(x) = \sum_{l=0}^{n+1} g_l \sum_{j=0}^{n} x^{n-j}\frac{1}{2\pi i}\oint_{x'=\infty} x'^{j-n-1}x'^{l-1}\frac{\frac{4Q}{P^2}}{\sqrt{1-\frac{4Q}{P^2}}+1-\frac{4Q}{P^2}}dx' \tag{4.151}$$

gives

$$W'^{(q)}(x) = \left(W'^{(c)}(x)\frac{1}{\sqrt{1-\frac{4Q}{P^2}}+1-\frac{4Q}{P^2}}\left(\frac{4Q}{P^2}\right)\right)\Bigg|_+. \tag{4.152}$$

Therefore we arrive at an expression for the left hand side of equation (4.144)

$$W'^{(c)}(x)^2 + 2W_n'^{(c)}W'^{(q)}(x) + (W'^{(q)}(x))^2 = \\ W'^{(c)}(x)^2 + \left(W'^{(c)}(x)^2 \left(\frac{\frac{4Q}{P^2}}{1 - \left(\frac{4Q}{P^2}\right)} \right) \right)\Bigg|_+ . \quad (4.153)$$

The field theory moduli coordinates on the left hand side of equation (4.144) are the double roots, while on the right hand side these are the s_i. These are related by the factorization of the Seiberg-Witten curve and using this the extra term on the right hand side of equation (4.144) is

$$\frac{4Q(x)B_q^2}{H_{2l}^2(x)} = \frac{\frac{4Q}{P^2}}{1 - \frac{4Q}{P^2}}(W'^{(c)}(x)^2 + g_m). \quad (4.154)$$

In the above equation m denotes the order of the extra polynomial in the factorization formula. To start of it is of order $2n - 2N_c + N_f$, which is smaller than $2n$. This polynomial order was derived from equation (4.144). However, we see that the leading term which does not depend on the polynomial cancels on both sides, so the polynomial is of one order lower. Then the argument can be repeated until g is of order $n-1$. This completes the argument that even for high order superpotentials the factorization of the Seiberg-Witten curve is captured in equation (4.16) if one consistently uses the gauge theory resolvent to calculate the quantum operators u_k.

Appendix C An observation

A standard technique in matrix model calculations for theories with fundamental matter is integrating out the matter content: the resulting matrix model can then be tackled by the usual techniques. The modified potential takes a form like

$$\sim \log(\det(\Phi + m)) \quad (4.155)$$

for every integrated flavor. Considering that the translation for the pure glue case proceeded by directly identifying the matrix integral variable with the Lax matrix of the periodic Toda chain, we might try to postulate that one has to take the modified matrix model potential and insert the Toda chain Lax matrix into it to obtain the effective action on $\mathbb{R}^3 \times S^1$. This would be in line with the observation in [82] that matrix model calculated superpotentials for systems with matter in the fundamental can be constructed from data of systems without. Note that in this way we can introduce at most $N_f \leq N_c$ flavors. Also, this kind of potential leads to equations of motion of high order rapidly. We calculated (some of) the minima of a potential $W = mx + \frac{1}{2}x^2 + h\log(\det(x+m))$ for N_c equal to 2 and 3 to check this idea. Note that we leave a term h of dimension 2 in front of the potential undetermined for reasons which will become clear below. In the matrix model this is N_f.

Gauge group $U(2)$

We plug the periodic Toda chain Lax matrix into the potential. In a full calculation we would have to use the field theory resolvent to calculate the potential correctly, but here we will follow a more direct route: we simply take the z-independent part of the determinant before taking the logarithm. Solving the equations of motion we get 6 curves. Four of these have a double root. In addition, there is a doubly degenerate curve of the form

$$y^2 = ((x+m)^2 + 2\Lambda^2 + h\Lambda^2)((x+m)^2 - 2\Lambda^2 + h\Lambda^2). \tag{4.156}$$

Now if this curve is to have double points, h has to be ± 2, where it has a double root at $-m$. The rest curve has the form of $(x+m)^2 \pm 4\Lambda^2$. Note that the calculation with the XXX chain gives a minus sign in this curve. The remaining 4 curves can now be evaluated for the same two values of h, and these give 3 different curves since 2 solutions collapse to the same curve solution:

$$y^2 = ((x+m)^2 \mp 4\Lambda^2)(x+m)^2 \tag{4.157}$$

$$y^2 = (x+m)(m+x+4\sqrt{\mp 1}\Lambda)(m+x+2\sqrt{\mp 1}\Lambda)^2 \tag{4.158}$$

$$y^2 = (x+m)(m+x-4\sqrt{\mp 1}\Lambda)(m+x-2\sqrt{\mp 1}\Lambda)^2. \tag{4.159}$$

Note that we do not get the correct form of the curve for the two solution which probably should correspond to the $r = 0$ branch. Also, we get one more curve at $r = 1$ than is expected from the field theory analysis. To further check this type of construction we turned to $N_c = 3$.

Gauge group $U(3)$

We first have to fix the coefficient h. For this purpose we use the solution with equal momenta $p_1 = p_2 = p_3$. This gives twelve solutions to the equations of motion, all of which have 2 double roots in the Seiberg-Witten curve (for unfixed h even). This branch has a curve with double roots at $-m$ if and only if $h = 0$. We therefore try to find a h such that it reproduces at least one of the double roots of the $r = 1$ branch of the 3 flavored theory. We find that we can fix h such that 2 of the curves have the correct double roots (but not the single roots), but then we can numerically simply verify that the other curves do not have double roots at a recognizable location. This indicates that this approach to finding the factorization locus of Seiberg-Witten curves with matter in the fundamental in this particular setup does not work. It might be that the solutions of the equations of motion we have not considered do have the correct double roots. However, then one would still have to explain that there is a huge excess of solutions, for which there is no (intrinsic) reason to discard them. We have therefore checked that the constructed potential for three colors cannot (fully) describe a system with 2 flavors, as one might suspect.

Discussion

Although an intriguing observation at the level of two colors, the constructed potential in this particular form obviously does not reproduce four dimensional field theory answers for theories with more than 2 colors and is therefore not the correct integrable systems-inspired exact superpotential for systems compactified on a circle. However, based this observation we conjecture that there is a (generic for all N) non-polynomial superpotential for the periodic Toda chain which reproduces Seiberg-Witten curve factorizations for supersymmetric gauge theories with fundamental matter ($N_f = N_c$) which for 2 colors resembles the superpotential displayed above. Using the field theory resolvent to calculate the non-polynomial potential might be a good starting point to improve the above calculation. In any case, a further study of equilibria of non-polynomial potentials for the Toda chain would certainly be very interesting.

Chapter 5

Integrability in gauge theories from random partitions

In the previous two chapters we saw that the proposal for an exact superpotential for gauge theories on $\mathbb{R}^3 \times S^1$ from chapter 2 yields results in remarkable agreement with results found in field theory. A very natural question is then how this approach through integrable systems is related to other approaches to four dimensional gauge theories, softly broken from $\mathcal{N} = 2$ to $\mathcal{N} = 1$ by turning on a superpotential, especially the matrix models in Dijkgraaf-Vafa theory. In this chapter we will study a framework in which to answer this question for pure $SU(N_c)$ $\mathcal{N} = 2$ four-dimensional gauge theories from the point of view of integrability.[1].

5.1 Gauge theory expectations

The effective potential calculated in Dijkgraaf-Vafa matrix models gives, as reviewed in chapter 2, the on-shell values of the glueball fields. On-shell, the matrix model resolvent gives through equation (2.66) the non-singular part of the Seiberg-Witten curve, just as in the factorization problem for the Seiberg-Witten curve (2.46) [23]. Since the integrable systems approach in this thesis starts off in $\mathcal{N} = 2$ with the Lax matrix, it is a natural first question how the full Seiberg-Witten curve or even the Lax matrix can be obtained from the matrix model. The Hermitian matrix model calculates an exact potential for theories in which the extended supersymmetry is broken to $\mathcal{N} = 1$, so one might expect that if the symmetry breaking potential in the matrix model is tuned to zero one would get $\mathcal{N} = 2$ directly. However, this does

[1]This chapter is based on unpublished work with Xerxes Arsiwalla

not seem to make sense directly in the matrix model, since an integral of Hermitian matrices over 1 is not well defined and does not seem to incorporate, for instance, the rank of the microscopic $U(N)$ gauge group. Also, the glueball superpotential is calculated by an expansion around a stationary point of the superpotential and in the limit the potential vanishes this expansion diverges off-shell. A simple gauge theory argument also shows that the Seiberg-Witten prepotential and the matrix model free energy cannot be directly related, despite their apparent similarity: the $a \log a$ term in the prepotential is the result of a perturbative one-loop gauge theory calculation, while the expansion in the prepotential is an expansion in non-perturbative instanton effects. The matrix model free energy consists of an expansion which is related to a perturbative gauge theory computation while the $S \log S$ part of the free energy is the result of a non-perturbative instanton contribution.

From a purely gauge theory perspective the (indirect) relationship is clear from [83]: The recipe for obtaining the Seiberg-Witten curve directly from the matrix model is (1) take a potential of suitably high order ($N_c + 1$), (2) distribute the matrix model eigenvalues over *all* the N_c extrema and (3) calculate the glueball condensates from the Dijkgraaf-Vafa potential. Since on-shell the factorization conditions (2.46) have been proven to hold in [83], we must have

$$W'_{N_c}(x)^2 + f_{N_c-1}(x) = c(P_{N_c}(x)^2 - 4\Lambda^{2N_c}), \tag{5.1}$$

with c some constant. The fact that this equation must hold on-shell fixes the coefficients of f and c

$$\begin{aligned} c &= g_{N_c+1}^2 \\ f_i &= 0 \quad \text{for} \quad i = 1, \ldots, N_c - 1 \\ f_0 &= -4\Lambda^{2N_c}, \end{aligned} \tag{5.2}$$

and we obtain

$$W'(x) = g_{N_c+1} P(x). \tag{5.3}$$

$\mathcal{N} = 2$ supersymmetry is regained in the gauge theory if the potential is tuned to zero. By redefining the couplings in the potential as $g_i \to g_{N_c+1}\hat{g}_i$ this limit is equivalent to $g_{N_c+1} \to 0$, keeping the coupling constants $\hat{g}_i$ fixed. Note that in this limit pure $\mathcal{N} = 1$ quantities like the glueball vevs S_i vanish. There are quantities which turn out to be actually independent of g_{N_c+1} like the period matrix of the Riemann surface which contain non-trivial $\mathcal{N} = 2$ field theory data.

Calculation of the Seiberg-Witten curve in Dijkgraaf-Vafa theory through the above procedure seems a rather roundabout approach; Ideally one would like to find a formulation in which both the matrix model and Seiberg-Witten theory are incorporated. One interesting obstacle to such an approach is for instance the fact that the rank of the gauge group N_c enters both approaches differently. In the matrix model this rank only occurs in the definition of the superpotential (2.71) as the sum of the N_i. In the integrable systems approach, which is naturally connected to Seiberg-Witten theory, the rank of the gauge group is hard-wired into the rank of the Lax-matrix and the superpotential itself is independent of the rank.

More generally, a formulation as described above would be the natural framework to study the relationship of integrability as discussed in this thesis and the Dijkgraaf-Vafa approach, for the following reason: Generically, one expects the matrix model partition function to be a τ function of some integrable hierarchy, with the coupling constants of the potential assuming the role of times in the hierarchy. A τ function specifies in a very precise sense a solution to the system of differential equations with respect to the times which make up the hierarchy. The connection between matrix model partition functions and τ functions can be made precise for the Hermitian matrix model for both finite N (see e.g. [84]), in the double scaling limit (see e.g. [53]) and in a special case of the 't Hooft limit [85]. One way of specifying a τ function is by expressing it as a vev for a theory of free non-relativistic fermions (see e.g. [86]). For instance, let

$$H_n = \sum_i : \psi_i \tilde{\psi}_{n-i} : \tag{5.4}$$

denote the generators of a $U(1)$ current algebra. Then for any operator A which can be expressed as a proper fermion bilinear, the vev

$$\langle 0| e^{\sum_{n>0} t_n H_n} e^{A} e^{\sum_{n<0} t_n H_n} |0\rangle \tag{5.5}$$

is a τ function for the Toda lattice hierarchy [87], which is equivalent to the 2 component KP hierarchy [87, 9]. In particular τ functions obey an infinite set of partial differential equations which can be derived through Hirota's bilinear identity. This set must be complemented by an initial condition, specified by the operator A, to make a well-defined problem. A question one can ask of any multi-cut solutions to a Hermitian matrix model in the 't Hooft limit is then whether the partition function can be written as a τ function and if so, what (the analog of) the operator A is for this function. From a gauge theory perspective, if the coupling constants are the times of flows in an integrable hierarchy, the initial condition for the matrix model is $\mathcal{N} = 2$ supersymmetric gauge theory. One approach to the study of integrability in matrix models is therefore to start in $\mathcal{N} = 2$ gauge theory and try to find deformations which correspond to turning on a superpotential. This is the approach taken in this chapter.

The first order of business is then to find a formulation of the Seiberg-Witten solution to $\mathcal{N} = 2$ $SU(N_c)$ gauge theory which is most closely related to its integrability properties. The formulation used in this chapter is the one due to Nekrasov, which takes the form of a random partition model. The free energy of the partition function of that model reduces to the Seiberg-Witten prepotential in a certain limit. Interestingly, this partition function has a dual fermionic formulation as a τ function of the Toda lattice hierarchy. We will study what exactly the deformation of this τ function should look like to relate it to Dijkgraaf-Vafa matrix models.

Since any τ function can be expanded in terms of Schur-functions [86], an alternative approach to the study of integrability in matrix models is for instance to search for an expansion of the matrix model partition function in terms of these functions, such as [85]. Since Schur-functions are labeled by partitions, this might then also be related to the approach using Nekrasov's formula. In the appendix we show, using results from [85], how this alternative avenue of attack is implemented for single-cut solutions in Hermitian matrix models.

5.2 Nekrasov's formula for the prepotential

The general expression for the $\mathcal{N} = 2$ prepotential in equation (2.34) shows the perturbative sum can, in principle, be calculated as regular instanton contributions in field theory. In practice this is technically very difficult to do. Using high-powered mathematics Nekrasov [17, 88] was able to calculate all (regular) instanton contributions to a topologically twisted version of the $\mathcal{N} = 2$ $SU(N_c)$ gauge theory partition function, from which the free energy (prepotential) can be calculated. In the limit in which the topological twist disappears the usual $\mathcal{N} = 2$ theory is regained. This limit is denoted $\hbar \to 0$ below. In leading order in $\hbar$ the partition function behaves as

$$Z \sim e^{\frac{F^0(a)}{\hbar^2} + \sum_{i=0}^{\infty} F^i \hbar^{2i}}. \tag{5.6}$$

In a saddle point approximation which will be reviewed below, Nekrasov's formula was shown to reproduce precisely the Seiberg-Witten solution in [88]. In particular it was proven that, as the notation suggests, the prepotential can be calculated using the equations (2.35) derived by Seiberg and Witten by holomorphy arguments. The $\hbar \to 0$ limit is what we will focus on most in this chapter.

More concretely, Nekrasov calculates the (holomorphic part of the) $\mathcal{N} = 2$ $SU(N_c)$ gauge theory partition function to be [88]

$$Z_{\mathcal{N}=2}(\vec{a}, \hbar, \Lambda) = Z_{\text{pert}}(\vec{a}, \hbar, \Lambda) \sum_{\vec{k}} \Lambda^{2N_c|k|} \prod_{(l,i)\neq(n,j)} \left(\frac{a_l - a_n + \hbar(k_{l,i} - k_{n,j} + j - i)}{a_l - a_n + \hbar(j - i)} \right)$$

$$Z_{\text{pert}}(\vec{a}, \hbar, \Lambda) = e^{\sum_{n,l} \gamma_\hbar(a_n - a_l, \Lambda)}, \tag{5.7}$$

where $\vec{k}$ is a N_c colored partition, $|k|$ is the total number of boxes in the colored partition, $\hbar$ denotes the parameter which has to be taken to zero to reproduce regular

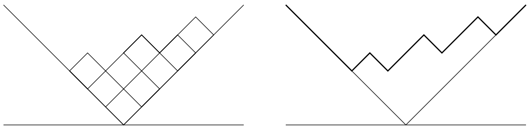

Figure 5.1: *The partition* $(5, 3, 1, 0, , 0, \ldots)$ *and its profile function*

four-dimensional gauge theory and $\gamma_\hbar$ is a function which admits an expansion in $\hbar$ of the form

$$\gamma_\hbar(x, \lambda) = \hbar^{-2}\left(\frac{1}{2}x^2 \log\left(\frac{x}{\Lambda}\right) - \frac{3}{4}x^2\right) - \hbar^0\left(\frac{1}{12}\log\left(\frac{x}{\Lambda}\right)\right) + \text{h.o.} \tag{5.8}$$

A colored partition $\vec{k}$ is an ordered N_c-tuple of partitions. Each element k_l is a partition which can be identified by a single Young diagram. Counted from below, the i-th row of the Young diagram k_l has length $k_{l,i}$. The function $\gamma_\hbar$ reproduces the perturbative contributions to the prepotential and arises through regularization of an ill-defined infinite product.

The prepotential F^0 in (5.6) was shown to be equivalent to the Seiberg-Witten prepotential in [88]. In order to compare Nekrasov's partition function to the Dijkgraaf-Vafa matrix model a short review of this proof will be given. The main idea is to rewrite the sum over partitions in the partition function (5.7) as an integral over continuous functions. In order to do that it is useful to define a so-called profile function for each partition. A single partition λ can be drawn in the Russian style as in the figure (5.2). The profile function $f_\lambda(x|\hbar)$ is defined as the function which describes its boundary. The functional form of the second derivative of the profile function can be worked out to be

$$\begin{aligned}\frac{1}{2}f''_\lambda(x|\hbar) = \delta(x) + \sum_{l=1}^{\infty}\big(&\delta(x - \hbar(\lambda_i - i + 1)) - \delta(x - \hbar(\lambda_i - i)) \\ &- \delta(x + \hbar(i-1)) + \delta(x + \hbar(i))\big)\,.\end{aligned} \tag{5.9}$$

Here the sides of the boxes are taken to be of size $\hbar$. The profile function of a colored partition is defined as a sum of partitions centered at different positions on the real axis,

$$f_{\vec{k}}(x|\hbar) = \sum_{j=1}^{N_c} f_{k_j}(x - a_j|\hbar). \tag{5.10}$$

Here it is assumed that the a_j are real. In [88] this is justified by the statement that non-real a can be defined by analytic continuation. Intuitively, when $\hbar \to 0$, the profile function becomes a continuous function because the bumpiness is scaled away. The shape of the profile function which dominates the path integral in this limit will indicate how the boxes of the colored partition are distributed over the real axes. Physical intuition suggests that the boxes will cluster around the centers a_l of the various sub-partitions. This picture is remarkably close to calculations in the large M limit of the Hermitian matrix model in chapter 2, but the difference is that here a potential is missing and in a precise sense only the measure of the matrix model is involved.

With the definition of the colored profile function and some properties of the γ function which can be found in ([88]), the partition function (5.7) can be rewritten as

$$Z_{\mathcal{N}=2}\left(\vec{a}, \hbar, \Lambda\right) = \sum_{\vec{k}} \exp\left(-\frac{1}{4}\int f''_{\vec{k}}(x|\hbar) f''_{\vec{k}}(y|\hbar)\gamma_\hbar(x-y,\Lambda)\right). \tag{5.11}$$

As suggested above, it is reasonable to to try to calculate the saddle-point of the above action in the $\hbar \to 0$ limit, where the sum over colored partitions is replaced by an integral over the colored profile function. Using the expansion (5.8) of the $\gamma_\hbar$ function the leading contribution to the partition function is

$$Z \sim \int df \exp\left(\frac{1}{\hbar^2}\int_{x<y} dxdy f''(x) f''(y)(x-y)^2 \log\left(\frac{x-y}{\Lambda} - \frac{3}{2}\right)\right). \tag{5.12}$$

The integral over the function f needs a precise definition: from the definition of f before taking the limit in equation (5.10) we obtain that

$$f(x) = \sum_{j=1}^{N_c} f_l(x - a_l). \tag{5.13}$$

Each f_l obeys some constraints which stem from the fact that they are continuous versions of discrete profile functions,

$$\begin{aligned} f_l(x) = |x|, &\qquad |x| >> 0, \\ |f_l(x) - f_l(y)| \le |x-y|. & \end{aligned} \tag{5.14}$$

The integral over functions of the form (5.13) which obey complicated constraints is very difficult to do in general because of the different positions a_l of the component profile functions. A resolution to this problem is to Legendre transform with respect to these positions. In the small $\hbar$ limit the colored profile looks like several disconnected regions of non-trivial ($\neq cx$ for some integer $|c| < N_c$) function connected by trivial, linear pieces. Denote the support of the non-trivial regions by U_i. By integrating over one such region we obtain

$$a_l = \frac{1}{2}\int_{U_i} x f''(x) dx. \tag{5.15}$$

In order for the Legendre transformation trick to work, a term must be added to (5.12) of the form $\frac{\xi_l a_l}{2\hbar^2}$. Since a_l is a function of f, the easiest way to add this term is by first defining a function σ which is piecewise linear with

$$x \in [-N_c + 2(l-1), -N_c + 2l] \qquad \sigma(x) = \xi_j x \tag{5.16}$$

for all $l = 1, \dots, N_c$ and then writing

$$\sum_{j=1}^{N} \xi_j a_j = \frac{1}{2} \int \sigma(f'(x)) dx. \tag{5.17}$$

Now the saddle-point problem is well defined and the equation of motion for the function $f(x)$ can be derived to be

$$\int_{y \neq x} dx dy (y-x) \left(\log \left(\frac{y-x}{\Lambda} \right) - 1 \right) f''(y) = \sigma'(f'(x)). \tag{5.18}$$

The right-hand side of the equation is either ξ_l or in the interval $[\xi_l, \xi_{l+1}]$ for some l, depending on the value of $f'(x)$: the function σ is not differentiable for $f'(x) = -N_c + 2l$ for $l = 0, \dots, N_c$. The value of it's derivative can be calculated as either a left or a right derivative,

$$\sigma'(-N_c + 2l) = \lim_{\epsilon \to 0^{\pm}} \frac{\sigma(x+\epsilon) - \sigma(x)}{\epsilon}. \tag{5.19}$$

Comparing the two definition then gives that σ' is in the interval $[\xi_l, \xi_{l+1}]$.

Nekrasov and Okounkov then go on to show that the first derivative of the profile function f which is the solution of (5.18) can be constructed as the (real part of the) value on the real axis of a certain conformal map $\Phi(x)$. This map is essentially $-i \log z$, with z a solution of $z + \frac{\Lambda^{2N_c}}{z} = P_N(x)$. ξ_l is given as the β-period integral of the Seiberg-Witten form and the free energy is simply given by the Seiberg-Witten formulas (2.35). This ends the lightning review of the work of Nekrasov and Okounkov.

5.2.1 Comparison a Hermitian matrix model

Note that formula (5.18) is remarkably similar to (2.69), if one only considers the equation on the support of the eigenvalues of the matrix-model. In fact, one is tempted to consider adding a potential term $e^{\frac{1}{\hbar^2}\int W(x) f'(x)}$ to the microscopic theory in equation (5.11), then to follow the derivation to the point above by varying with respect to $f'(x)$. This identifies for the leading terms in the $\hbar \to 0$ limit (the 't Hooft limit on the matrix model side)

$$f'(x) \sim \rho(x). \tag{5.20}$$

The tilde arises from undetermined constants which follow from partial integration. On the cuts we then arrive at the exact same equation as (2.69). Note that while this addition is motivated from Dijkgraaf-Vafa theory, it would be nice to understand what this means in the underlying topologically twisted theory used by Nekrasov.

The identification proposed above is in accordance with the naive field theory expectation that for the case in which the superpotential is switched of Seiberg-Witten theory should emerge from the matrix model. From the Hermitian matrix model point of view, this indicates that the measure of the matrix model is connected to Seiberg-Witten theory. However, the precise connection of the variables in Dijkgraaf-Vafa and Seiberg-Witten ('a' vs. 'S') theory is not clear from the above argument and it is known that this direct relation is complicated in general [89]. Furthermore, the Hermitian matrix model in Dijkgraaf-Vafa theory computes a superpotential which is only on-shell indirectly related to the prepotential of Seiberg-Witten theory. In other words, the above approach does not show yet how the equations of motion which arise in Dijkgraaf-Vafa theory should be interpreted in the framework of Nekrasov's partition function.

Note that the variable problem at least is not really a problem, since these variables concern different asymptotic expansions. The matrix model free energy in terms of S is the diagrammatic expansion of the matrix model partition function around a classical minimum of the matrix model potential. In this expansion the measure is considered to be 'smaller' than the potential. The variables 'a' are more naturally connected to a regime where the measure is considered to be larger than the potential.

Related observations to the ones in this subsection were made in [90] and especially in [91] (see the subsection on 2D Yang-Mills theory below), although there the regular (not the dual) partition function was studied. In general it would be very interesting to further understand (5.20) in more detail. Below we will provide a further motivation for the validity of this identification from the point of view of 2D Yang-Mills theory. Finally, it would be interesting to understand a possible field theory origin of the terms which have to be added to the partition function of the random partition model.

5.3 Free fermion formulation

From the viewpoint of integrability, the most interesting aspect of Nekrasov's work is the existence [88] of a free fermion formulation of the partition function which leads to the saddle point equation (5.18). This dual partition function can be represented as (we will specify the exact relation shortly)

$$Z_D(\xi_l, p, \Lambda, \hbar) = \langle p| e^{\frac{-H_1}{\hbar}} e^{\frac{1}{\hbar}\sum_{i=1}^{N_c} \xi_l C^{(l)}} e^{\frac{\Lambda^2 H_{-1}}{\hbar}} |p\rangle. \quad (5.21)$$

Here $|p>$ is defined to be the vacuum annihilated by ψ_k for all $k > -p$, $C^{(l)}$ is the charge defined as

$$C^{(l)} = \sum_{i=l \mod N+\frac{1}{2}} :\psi_i \tilde{\psi}_{-i}:, \quad (5.22)$$

and the operators H_n are defined as

$$H_n = \sum_i \psi_i \tilde{\psi}_{n-i}, \quad (5.23)$$

where the sum runs over the half-integers. The variables ξ_l are chosen to sum to zero. The relationship with equation (5.7) can be worked out by first taking into account the identity

$$e^{\frac{H_{-1}}{\hbar}} |p\rangle = \sum_\lambda \frac{R_\lambda}{\hbar^{|\lambda|} |\lambda|!}, |p; \lambda\rangle \quad (5.24)$$

where λ is a single partition, R_λ is the dimension of irreducible representation R_λ of the symmetric group which can be computed through the Hook length formula and $|p;\lambda\rangle$ is the state

$$|p;\lambda\rangle = \prod_{i=1}^{l(\lambda)} \psi_{-p-\frac{1}{2}+i-\lambda_i} \prod_{i=l(\lambda)}^{1} \tilde{\psi}_{p+\frac{1}{2}-i} |p\rangle. \quad (5.25)$$

The identity is a simple consequence of Taylor-expanding the exponential and identifying the states which can be built. The operator H_{-1} shifts one of the fermions which make up the state by one position. In the expression above for instance this means that the first fermion of the ground state is shifted by λ_1 positions. It is easy to see that therefore any state labeled by a partition λ can be built by application of $|\lambda|$ operators H_{-1}. This can be done in R_λ different ways. The formula can also be written as

$$e^{\frac{H_{-1}}{\hbar}} |p\rangle = \sum_\lambda \prod_{i<j} \frac{\hbar(\lambda_i - \lambda_j + j - i)}{\hbar(j-i)} |p;\lambda\rangle. \quad (5.26)$$

Equivalence to the form first given is established by first properly defining these products[2] and then using an identity in [92] (page 9, equation 4) to convert the result to the hook length formula. Some more details can be found in [93] (especially the last formula on page 26 is helpful). The double product on the right hand side

[2]Suppose the partition is of length n. Split the product in a part with $\{i, j < n\}$, a part with $\{i, j > n\}$ and one $\{i < n, j > n\}$. The first two parts are well-defined while the last one needs careful consideration.

of the equation is (for $\hbar = 1$) the dimension formula of irreducible representations of $U(N')$ for $N' > l(\lambda)$.

The operator in the middle of the fermion vev (5.21) suggests that the best way to calculate with it is to split up the fermion into a so-called N_c-component fermion. This is done by relabeling the fermions as

$$\psi_{N_c(r-\frac{1}{2})-l-\frac{1}{2}} = \psi_r^{(l)}. \tag{5.27}$$

A more graphical representation of the above formula can be constructed in the following way: write the single fermion modes as a sequence. Writing the corresponding N_c-component fermion below it yields a picture like

$$\begin{array}{ccccccc} \cdots & \psi_{-3/2} & \psi_{-1/2} & \psi_{1/2} & \psi_{3/2} & \psi_{5/2} & \cdots \\ & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \\ \cdots & \psi_{-1/2}^{(N_c-1)} & \psi_{-1/2}^{(N_c)} & \psi_{1/2}^{(1)} & \psi_{1/2}^{(2)} & \psi_{1/2}^{(3)} & \cdots \end{array} . \tag{5.28}$$

The relabeling of the conjugate fermions is determined by the demand that canonical anti-commutation relations are obtained, e.g.

$$\{\psi_k^{(n)}, \tilde{\psi}_l^{(m)}\} = \delta_{m,n}\delta_{k,-l}. \tag{5.29}$$

In the identity (5.26) the sum is over those partitions which correspond to the usual labeling of the fermions. In order to compare to sums over N_c-tuples of partitions as, for instance, those in Nekrasov's partition function (5.7), we would like to use the N_c component formalism. To get a flavor of what is required let us study the first non-trivial partition which consists of just one box, set the total charge $p = 0$ and take two colors ($N_c = 2$). The box partition gives a state

$$|\Box\rangle = \psi_{-1/2}\psi_{3/2}\psi_{5/2}\dots . \tag{5.30}$$

In terms of a 2-tuple of partitions as in (5.28) this is an empty first partition of charge -1 and second empty partition of charge $+1$. Note that it is important to keep the order of the charges in mind: a 2-tuple of two empty partitions with charge $(1, -1)$ corresponds to a different state (labeled by the uncolored partition $\{2, 1\}$). In the same way as above the single box partition can be decomposed into any N_c-tuple of partitions, where it will give an empty N_c-tuple of partitions of N_c-charge $(1, 0, \dots, 0, -1)$.

The simple example of the single box partition already shows that even if we start of with an partition with total charge zero ($p = 0$), in the N_c-tuple we will get all kinds of sub-charges for the partitions (which always sum up to zero). In general any single partition $\lambda = \{\lambda_1, \lambda_2, \dots\}$ may be decomposed uniquely into an N_c-tuple $\vec{k}$ of partitions and a set of N_c charges p_l (which sum up to p) according to

$$\{N_c(p_l + k_{l,i} - i) + l - 1 | l \in 1, \dots, N_c, i \in \mathbb{N}\} \leftrightarrow \{p + \lambda_j - j | j \in \mathbb{N}\} . \tag{5.31}$$

The uniqueness of the decomposition is ensured by the demand that every sub-partition k_i is such that

$$\exists m | k_{i,j} = 0 \qquad \forall j > m. \tag{5.32}$$

The above decomposition process (or rather its inverse) is called blending of partitions in [88].

With the above relabeling in hand we now turn to (5.26). We know from (5.24) that for any fixed number of boxes in the uncolored partition this is well-defined, a fact which relabeling will not change. To convert to (5.26) we simply rewrite the numerator. We obtain

$$\begin{aligned} e^{\frac{H_{-1}}{\hbar}} |p\rangle = & \sum_{\vec{k}, \{\vec{p} | \sum p_l = p\}} \prod_{s<r} \frac{1}{\hbar(r-s)} \\ & \prod_{i,j,l,m} \hbar(N(p_l + k_{l,i} - i) + l - N(p_m + k_{m,j} - j) - m) |\vec{p}; \vec{\lambda}\rangle. \end{aligned} \tag{5.33}$$

There is a constraint on the ranges of i, j, l, m, in accordance with the constraint as written above. This however will not matter in the final result. The products must be carefully defined just as in [17] (see also the appendix of [91]). The trick is to multiply by

$$1 = \frac{\prod_{i,j,l,m} \hbar(N_c(p_l - i) + l - N_c(p_m - j) - m)}{\prod_{i,j,l,m} \hbar(N_c(p_l - i) + l - N_c(p_m - j) - m)}, \tag{5.34}$$

and absorb the denominator into the definition of the product which contains the $k_{l,i}$. The numerator gives, through zeta function regularization, the perturbative contribution Z_{pert} to the partition function and we discard the infinite constant.

The above identity almost allows the fermion vev in (5.21) to be evaluated explicitly. The final ingredient needed is the value of $|k|$ in terms of the blended partitions. This can be calculated using power sums (see [88], a quadratic contribution to the classical prepotential can always be reabsorbed into a redefinition of the classical coupling constant) and comparing to (5.7) gives

$$Z_D = \sum_{\vec{k}, \{\vec{p} | \sum p_l = p\}} Z_{\mathcal{N}=2} \left(\hbar \{ p_l + \frac{l}{N} \}, \hbar, \Lambda \right) e^{\frac{1}{\hbar} \sum_l \xi_l p_l}. \tag{5.35}$$

In the limit $\hbar \to 0$ keeping $\hbar p_l$ fixed, the calculation of the saddle-point of this dual partition function reduces exactly to the calculation of the Seiberg-Witten solution from Nekrasov's partition function as reviewed above.

5.4 Periodic Toda chain Lax matrix

The fermionic description of Nekrasov's formulation of the Seiberg-Witten theory is useful because it illuminates the connection between Seiberg-Witten theory and integrability. As an illustration of this we will derive the periodic Toda chain Lax matrix directly from the fermionic formulation. Then we will invert the logic and show in the partition language what kind of operators should generate the flows in the Toda system which played a large role in chapter 3.

The main tool is the observation that the $\hbar \to 0$ limit is remarkably similar to the dispersionless ('classical') limit known in integrable system theory [94] and explored in the context of the Seiberg-Witten solution in e.g. [95]. Actually, the last paper gives a definite prescription how to obtain the periodic Toda chain Lax matrix from the general description of the Toda lattice hierarchy described in [96]. For this calculation to work we need two elements. The first element is that the Toda lattice hierarchy should be N_c-periodic in the sense that the tau function τ' should obey

$$\begin{aligned} \tau'(p+N) &= \tau'(p) \\ \partial_{x_{N_c}}\tau' &= \partial_{y_{N_c}}\tau'. \end{aligned} \tag{5.36}$$

The second element is a certain condition on the wave function which is associated to the τ' function. which is written below in equation (5.46). We will first show both elements are met in the particular form of the Toda lattice tau function we are interested in.

The first constraint in (5.36) can be verified straightforwardly. If we define

$$\tau'(p) = e^{\sum_n n\frac{x_n y_n}{\hbar^2}} \langle p| e^{\frac{x_n H_n}{\hbar}} e^{\sum_{i=1}^{N_c} \xi_l C^{(l)}} e^{\frac{y_n H_{-n}}{\hbar}} |p\rangle, \tag{5.37}$$

it is easy to see that the second constraint in (5.36) holds exactly. The fact that the charge operators and the Hamiltonian operators H_{tN}, which were defined for any subscript in (5.23), commute for integer $t \in \mathbb{Z}$ follows from the observation that, in N_c-component fermion language, these operators are sums of the Hamiltonians $H^{(l)}_{\pm t}$. Note that the other Hamiltonians in general do not commute with the charges C_l, so the vev is far from trivial.

Tau functions are useful because they generate a certain hierarchy of differential equations. These can be derived by using the identity

$$[z^{-1}]\psi(z) \otimes \tilde{\psi}(z) G = [z^{-1}] G\psi(z) \otimes \tilde{\psi}(z), \tag{5.38}$$

which holds for all operators G which are the exponent of a sum over fermion bilinears. Here $[z^{-1}]f(z) = f_{-1}$ for any bi-infinite Laurent series $f(z) = \sum_i f_i z^i$. By wedging this identity between operators $\langle p+1| e^{\sum_{n>0} x_n H_n} \otimes \langle p'-1| e^{\sum_{n>0} x'_n H_n}$ and

$e^{\sum_{n<0} y_n H_n}|p\rangle \otimes e^{\sum_{n<0} y'_n H_n}|p\rangle$ one can derive Hirota's bilinear identity (see [9] for the general case or, alternatively, appendix C in [97] for the Toda hierarchy). This is most easily written in terms of wave functions. Define wave functions

$$\phi^{\infty}(z) = e^{\sum_n n\frac{x_n y_n}{\hbar^2}} \langle p+1|e^{\frac{x_n H_n}{\hbar}} \psi(z) e^{\sum_{i=1}^{N_c} \xi_l C^{(l)}} e^{y_n \frac{H_{-n}}{\hbar}} |p\rangle \tag{5.39}$$

$$\phi^{0}(z) = e^{\sum_n n\frac{x_n y_n}{\hbar^2}} \langle p+1|e^{\frac{x_n H_n}{\hbar}} e^{\sum_{i=1}^{N_c} \xi_l C^{(l)}} \psi(z) e^{y_n \frac{H_{-n}}{\hbar}} |p\rangle, \tag{5.40}$$

where

$$\psi(z) = \sum_{k\in\mathbb{Z}} \psi_{k+1/2} z^k. \tag{5.41}$$

The superscript indicates the fact that the wavefunction have either structure of an asymptotic expansion at $z=0$ or at $z=\infty$. In these wave functions a normalization factor $e^{\sum_n n\frac{x_n y_n}{\hbar^2}}$ has been added. The conjugate wave-functions can be defined analogously. For these wave functions Hirota's bilinear identity holds,

$$[z^{-1}]\phi^{\infty}(x,y,z)\tilde{\phi}^{\infty}(x',y',z) = [z^{-1}]\phi^{0}(x,y,z)\tilde{\phi}^{0}(x',y',z). \tag{5.42}$$

In the case at hand they also obey the interesting equations

$$\begin{aligned} \hbar\partial_{x_{tN_c}} \phi^{\infty,0}(z) &= z^{tN_c}\phi^{\infty,0}(z) \\ \hbar\partial_{y_{tN_c}} \phi^{\infty,0}(z) &= z^{-tN_c}\phi^{\infty,0}(z) \end{aligned} \tag{5.43}$$

for integer t. The wave functions $\phi^{0,\infty}(z) = \sum_i \phi_i^{0,\infty} z^i$ can be glued together into one wavefunction if track is kept of the powers of z involved,

$$\phi(z) = \begin{cases} \phi_i^0 z^i & i<0 \\ \phi_i^{\infty} z^i & i>0 \end{cases} \tag{5.44}$$

Using proposition 1.6 in [96] we see that if we include the overall charge p in the wave function in the obvious notation $\phi(z,p)$, this wave function is related to the wave matrices of that paper. If we expand

$$\phi(z,p) = \sum_j c(j,p) z^j, \tag{5.45}$$

the coordinates (j,p) are identified with (j,s) in that paper.

The calculation of the vev in (5.39) is in general difficult. It does however obey the constraint

$$z\partial_z \phi(z) = \left(\sum_n n x_n \partial_{x_n} - \sum_n n y_n \partial_{y_n}\right)\phi(z). \tag{5.46}$$

This equation can be shown to be true in the following way: assign to all fermion field an auxiliary charge T according to

$$\begin{aligned} T\left[\psi_{k+1/2}\right] &= k+\frac{1}{2} \\ T\left[\tilde{\psi}_{k+1/2}\right] &= -k-\frac{1}{2}. \end{aligned} \tag{5.47}$$

Here $T[A]$ denotes the charge of the operator A. Then the operators H_n and C_l have a definite charge of $T = n$ and $T = 0$. Now also assign charges to the coupling constants and z according to

$$\begin{aligned} T[x_n] &= -n \\ T[y_n] &= n \\ T[z] &= -1. \end{aligned} \tag{5.48}$$

The bra on the left of (5.39) has charge 1 compared to the ket and the fermionic wavefunction $\psi(z)$ has a fixed T-charge for each power of z. These sources of T-charge are discounted in the following since they are fixed. In order to give a non-zero vev, the operator between the brackets should then have total T-charge zero, since non-zero T-charge of the operator indicates that there is one power of H_k more than of H_{-k}. In general we can expand $\phi(z)$ in a formal power series in $\{x_n, y_n, z\}$

$$\phi(\{x_n, y_n, z\}) = \sum_{n_i, n_j, k} c_{n_i,n_j,k} \prod_{i,j} \left(x_i^{n_i} y_j^{n_j} z^k\right). \tag{5.49}$$

Since the fermion vev in the coefficient c has definite T charge zero, the following relation between the exponents should hold

$$\sum_i n_i i - \sum_j n_j j + k = 0. \tag{5.50}$$

Due to this relation, equation (5.46) holds. With the same type of argument one can show that the τ' function 5.37) obeys

$$\left(\sum_n n x_n \partial_{x_n} - \sum_n n y_n \partial_{y_n}\right) \tau = 0. \tag{5.51}$$

In the above derivation the assignment of T-charge was used in order to keep track of the operators involved in the vev, and the introduction of more general operators into the τ function would spoil (5.50) as the only coefficients for which we have non-zero coefficients c. As we will see below, it is very interesting to study the natural class of operators which can be inserted into (5.39) which leave the constraint (5.46) untouched. These operators must have T charge zero.

Equation (5.46) is the second needed condition in [95]. Note that in that paper it had to be imposed as an external equation, while here it appears naturally. It was already noted above that there is a natural map between the wave function as constructed as the fermion vev and the wave matrices studied in [96]. Actually, since the total charge is the column counting index in that paper, it is more natural to study a wave-vector: a wave-matrix at finite p which corresponds exactly to the wavefunction above, with the entries determined by the coefficients in the Laurent expansion. Furthermore, because of the periodicity conditions (5.36), we are dealing with the N_c-periodic Toda lattice hierarchy. This can also be seen from (5.43), which in addition shows that the periodicity parameter in this theory (ζ in [96], z in chapter 3) is z^{N_c}. To avoid unnecessary confusion in the following the definition $h \equiv z^{N_c}$ will be used.

In integrable system theory, the wave-vector is important because it is a key ingredient for writing down the integrable hierarchy in the formulation of [96]. From Hirota's bilinear equations it can be shown that the periodic wavevector $\vec{\phi}$ is the solution of an auxiliary linear problem. Essentially, this is an auxiliary problem of the Lax type, appearing here for an integrable hierarchy, not a classical integrable system. We have, including the parameter $\hbar$,

$$\hbar\partial_{x_n}\vec{\phi} = B_n\vec{\phi} \tag{5.52}$$
$$\hbar\partial_{y_n}\vec{\phi} = D_n\vec{\phi}, \tag{5.53}$$

for certain infinite matrices B and C. Here we will only need B_1, an matrix with periodic coefficients b_l on the diagonal and D_1, a strictly lower triangular matrix with periodic coefficients d_l

$$B_1 = \begin{pmatrix} \dots & \dots & \dots & \dots & & \\ \dots & b_N & 1 & 0 & 0 & \\ \dots & 0 & b_1 & 1 & 0 & \dots \\ \dots & 0 & 0 & b_2 & 1 & \dots \\ & 0 & 0 & 0 & b_3 & \dots \\ & & \dots & \dots & \dots & \dots \end{pmatrix}. \tag{5.54}$$

$$D_1 = \begin{pmatrix} \dots & \dots & \dots & \dots & & \\ \dots & 0 & 0 & 0 & 0 & \\ \dots & d_0 & 0 & 0 & 0 & \dots \\ \dots & 0 & d_1 & 0 & 0 & \dots \\ & 0 & 0 & d_2 & 0 & \dots \\ & & \dots & \dots & \dots & \dots \end{pmatrix}. \tag{5.55}$$

The variables b_l and d_l are in general functions of x_n, y_n. These matrices themselves obey flow equations which can be derived as the compatibility equations of (5.52). The flow equations for the matrices B_n and D_n form the Zakharov-Shabat representation of the Toda lattice hierarchy.

The parameter $\hbar$ in the fermionic vev bears a resemblance to Planck's constant, so a natural approach to the wave function (5.39) is to do a WKB approximation. This is precisely what is done in [95]. From equation (5.46), evaluated at the times in (5.21) we get

$$\frac{\hbar}{N_c}\frac{\partial}{\partial h}\phi(z) = -\left(\hbar^2\partial_{x_1} + \Lambda^2\hbar^2\partial_{y_1}\right)\phi(z). \tag{5.56}$$

The $\hbar \to 0$ limit now determines that the vector $\phi(z)$ is an eigenvector of the periodic Toda chain Lax matrix in coordinates related to the canonical ones, while the eigenvalue of this Lax matrix is a momentum coordinate, conjugate to the periodicity parameter $\frac{h}{N_c}$. The quantum representation of the momentum is simply $\hbar\partial_{\log h}$. This shows, for instance, that the Seiberg-Witten differential is proportional to 'pdq'.

The above analysis implies that in $\hbar \to 0$ limit the Seiberg-Witten curve emerges from Nekrasov's (dual) partition function as the spectral curve of the Lax matrix. The full Seiberg-Witten solution (2.35) then seems to follow by using the calculation in [95]. Note that the variables ξ_l parametrize in the limit the moduli of the curve, although the relationship between them is non-linear.

Actually, with hindsight it is surprising that the formulation of the Seiberg-Witten solution in terms of a fermionic vacuum expectation value was not found sooner. It was known since [30] that there was an integrable structure (the periodic Toda chain) behind the solution. Since the periodic Toda chain can be obtained straightforwardly as a known limit of the Toda lattice hierarchy, it is natural to suspect that there is a fermionic vev which gives the tau function (prepotential) and the closely related wave function. Then the only thing left to be fixed is the exact form of the initial conditions (the operator A in (5.5)) and the general form of these is fixed by [95]. The simplest operator which is non-trivially N-periodic is then the operator in (5.21). On the other hand, the fermionic vev is not the prepotential directly, but instead gives a dual of the prepotential, which might have thrown up some obstacles in deriving the explicit connection.

5.5 Relationship to 2D Yang-Mills

The above construction is very similar to (and in fact partly inspired by) the 't Hooft large $\tilde{N}$ limit of generalized 2D $U(\tilde{N})$ Yang-Mills theory on a sphere. This theory is defined as (see e.g. [98])

$$Z = \int DAD\phi \, \exp\left(\frac{i}{4\pi^2}\int \mathrm{Tr}\,\phi F + \frac{g_{ym}}{8\pi^2}\mathrm{Tr}\,\tilde{W}(\phi)\right), \tag{5.57}$$

with ϕ a scalar transforming in the adjoint and $\tilde{W}$ a general polynomial superpotential. Note that for a quadratic potential $\tilde{W} = \phi^2$ the adjoint field can be integrated out and a pure two-dimensional Yang-Mills theory is recovered. The above partition function can be calculated exactly by using, for instance, either standard path-integral techniques [98] or the existence of a lattice formulation which is exact [99]. The result for a quadratic potential is

$$Z = \sum_{\lambda} (\dim R_\lambda)^2 \exp\left(\frac{g_{ym}}{N} C_2(R)\right). \tag{5.58}$$

where λ is a partition which labels an *irreducible* representation R_λ of $U(\tilde{N})$ and C_2 is the quadratic Casimir of the gauge theory.

The reason this class of theories is interesting from the perspective of the topic of this chapter is the fact that in the large $\tilde{N}$ limit the partition function (5.58) is related to a Hermitian one matrix model with a polynomial potential [100, 101]. For a general polynomial potential $\tilde{W}(\phi)$ there is a corresponding matrix model potential of the same maximal order (but in slightly different variables). Furthermore, for each such potential the exact partition function can be calculated

$$Z = \sum_\lambda (\dim R_\lambda)^2 \exp\left(\frac{g_{ym}}{\tilde{N}} \sum_{j=1}^{k} \alpha_j C_j(R)\right). \tag{5.59}$$

The dimension formula is of course almost identical to the one used above in (5.26). This expression can be written in terms of free fermions, although now there is a cut-off in the number of fermions: this corresponds to the fact that in the above formula's one sums over irreducible representations and the Young diagrams corresponding to reducible representation should be excluded. In terms of Young diagrams this implies a cutoff on the height of the diagram: it may not be higher than $\tilde{N}$.

The kinship of 2D YM to Nekrasov's formulas for the prepotential was already noted in [88]. In [90] this relationship was explored as a suggestive rewriting of the dimension formula. The authors of that paper studied in addition the Gaussian Hermitian matrix model in terms of profile functions and the run-up to the Douglas-Kazakov phase transition. This calculation shows that terms in the polynomial potential of the Hermitian matrix model get identified with operators K_{k+1} according to

$$g_k K_{k+1} \sim g_k \mathrm{Tr}\, M^k \tag{5.60}$$

where M is the matrix model variable. Actually, in the calculations in [90] one should take, in their notation, $\epsilon = g_{ym}$ in contrast to the claims in that paper. The reason is that in formula (2.8) of [90] a factor of $\hbar^{|k|}$ is missing (compare equation (5.24)), which follows from the precise definition of the generally ill-defined infinite product. Tracing this factor through the calculation then gives that the large $\tilde{N}$ limit studied in [90], which is called a double scaling limit in that paper, can just be identified with the usual Douglas-Kazakov 't Hooft large $\tilde{N}$ limit. It is an interesting question what the partition sum with the extra parameter ϵ corresponds to, but it might simply be the 2D Yang-Mills theory again, written in slightly different variables.

The connection with profile functions comes because the Casimirs of the 2D gauge theory can be rewritten in terms of the profile functions by simply calculating the expansion coefficients. It is easy to see that, with the Casimirs taken to be (proportional to)

$$\tilde{N}^{k+1}C_k(\lambda) = \sum_i (\lambda_i + \tilde{N} - i)^k - (\tilde{N} - i)^k, \tag{5.61}$$

that there are coefficients c_j such that

$$C_k = \sum_{j=1}^{k+1} c_j \int dx x^j f''_{\vec{k},p}(x,\hbar). \tag{5.62}$$

The upper limit of this summation follows from the observation that $C_k \propto \lambda_i^k \propto K_{k+1}$. In 2D Yang-Mills theory, the large $\tilde{N}$ limit can be studied [100, 101] directly by taking a continuum limit in the partition function (5.59). This yields a Hermitian matrix model with a polynomial potential of the same order as the potential in (5.59).

The calculations in 2D Yang-Mills theory are, of course, related to (5.63). In fact, one may (almost) interpret that vev as a way of writing the 2D Yang-Mills partition function, where for finite $\tilde{N}$ one must impose a cut-off on the number of fermions which may get excited from the vacuum. However, for finite $\tilde{N}$ it is unnatural to study the colored fermions we studied above. Another, related difference is the occurrence in (5.63) of the term $\xi_l \int f'$ and it would be nice if this term also had a natural interpretation in the 2D Yang-Mills theory, perhaps as non-perturbative ($e^{-\tilde{N}}$) effects in the large $\tilde{N}$ limit.

In [91] the authors used the matrix model point of view to calculate a saddle-point of the 2D Yang-Mills partition function in the limit of large potential. They then study fluctuations around that saddle-point. Without further arguments however, this saddle-point approach does not allow one to simply take a limit to small potential directly as is done in that paper. Still, the observations from [91] are interesting in their own right and it would be nice to know the connection to for instance the fermionic formulation explored in this chapter. It appears that the calculations there focus on a different regime in the theory space: the fluctuations of the eigenvalues which are located in a cut are considered for both left and right fluctuations separately. In a different context, this type of fluctuations is studied in [102], where non-perturbative effects in the large $\tilde{N}$ are studied in a q-deformed 2D Yang-Mills theory on a torus. It is an interesting question how all this might be interrelated.

As a final point, it is interesting to note that the matrix model obtained in another lattice formulation of 2D Yang-Mills by Gross and Witten [103] has a double scaling limit which gives $\mathcal{N} = 2$ $SU(2)$ Yang-Mills theory directly [25]. Furthermore, in this double scaling limit the 2D Yang-Mills potential vanishes, in accordance with the other calculations in this section.

5.6 Deformations of the fermionic formulation and the matrix model

From the results in chapter 3 it is explicitly known what the equations are that have to be imposed (in the $\hbar \to 0$ limit) on the periodic Toda chain Lax matrix in order to obtain the factorization of the Seiberg-Witten curve expected from field theory,

$$[L, \mathcal{L}(W'(L))] = 0.$$

As a first step in the direction of obtaining this equation in the framework painted above, in this subsection the goal is to identify what the natural operators are which can be inserted into the fermion vev to model these time flows. In the next subsection we will argue that these operators should be the ones which correspond to turning on a superpotential in the gauge theory.

The property which we would like to keep fixed in the above derivation is (5.46) and not necessarily the periodicity condition. Remember that in the matrix model the periodicity of the Lax matrix, which is the rank of the microscopic gauge group, is not seen directly. Furthermore, the $\mathcal{N} = 2$ theory should be the initial conditions of the flows which are related to the coupling constants of the superpotential. With these considerations in mind, it is natural to try to model the addition of a superpotential as a change in the initial conditions of the $\mathcal{N} = 2$ theory itself. The natural operators which leave (5.46) invariant consist of operators which are sums of $\psi_k \tilde{\psi}_{-k}$. We therefore propose to study the (dual) partition function

$$Z_D(\xi_l, g_k, p, \Lambda, \hbar) = \langle p | e^{\frac{x_n H_n}{\hbar}} e^{\frac{1}{\hbar}\sum_{i=1}^{N_c} \xi_l C^{(l)}} e^{\frac{1}{\hbar^2}\sum_{k=2}^{\infty} \frac{g_k}{k+1} K_{k+1}} e^{\frac{y_n H_{-n}}{\hbar}} | p \rangle. \tag{5.63}$$

The operators K_{k+1} are taken to be

$$K_{k+1} \equiv \hbar^{k+1} \sum_r \left((r + \frac{1}{2})^{k+1} - (r - \frac{1}{2})^{k+1} \right) : \psi_r \psi_{-r} : . \tag{5.64}$$

This particular form is inspired by the analysis of the saddle-point equation above, considerations in 2D Yang-Mills theory (see below), and calculations in [93]. The authors of that last paper arrived at this particular form because of the fact that for uncolored partitions the following equation can be straightforwardly shown to hold,

$$K_{k+1} | p; \lambda \rangle = \left(\int dx x^{k+1} f_k''(x, \hbar) \right) | p; \lambda \rangle. \tag{5.65}$$

Repeating the derivation for colored partitions which lead to (5.35) shows that turning on the operators K^{k+1} corresponds to adding a term

$$\frac{1}{\hbar^2} \frac{g_k}{k+1} \sum_l \left(\int dx x^{k+1} f_l''(x - a_l, \hbar) \right) \tag{5.66}$$

to the free energy of the $\mathcal{N} = 2$ theory.

The missing information, compared to the previous two chapters, is exactly how the coupling constants g_k are related to flows in the Toda hierarchy. According to [93], [104] shows that g_k are the times in a Toda lattice hierarchy, but we have not yet been able to verify this concretely. A related question is how all the operators in this theory are related to each other. In particular, the resolvent of the matrix model should be related to the operator in (5.66). In the $\hbar \to 0$ limit, this will reduce to the usual matrix model hyperelliptic Riemann surface. If periodicity is imposed by hand, then there is also a natural Lax operator in the theory, since the derivation of the periodic Toda chain above then also holds. It would be natural to suspect that these two operators have a shared spectrum, i.e. *commute*. This would lead directly to an equation of the form (3.111). This would then fix the parameters ξ_l as functions of the coupling constants. However, it is not guaranteed that the wave functions studied above are also eigenfunctions of the matrix model resolvent. Alternatively, the needed equations of motion can be imposed by hand.

5.7 Discussion

In this chapter it is shown how the integrable structure explored in the previous chapters might be embedded into an integrable hierarchy. We have shown that the Seiberg-Witten curve can be derived directly as the spectrum of a Lax matrix of the periodic Toda chain in the dispersionless limit of Nekrasov's expression for the $\mathcal{N} = 2$ prepotential[3]. This was done in a certain fermionic formulation which is a τ function of the Toda lattice hierarchy. The Dijkgraaf-Vafa matrix model saddle point equation is seen to arise naturally from the addition of certain operators to the random partition model. The precise dictionary is determined by 2 observations: one is the comparison of the saddle-point equation derived from the (dual) random partition expression and the one derived from the (dual) Hermitian matrix model, (2.69). The other observation is that, when expressed in terms of profile functions, the $\hbar \to 0$ limit is equivalent to the limit $\epsilon \to 0$ studied in [90]. From the details of that calculation one can see that the identification of terms in the Hermitian matrix model with operators in (5.63) should also be in line with $\frac{1}{k+1}x^{k+1}f''(x) \leftrightarrow \mathrm{Tr}\,\phi^k$. It is shown how these operators may be then be incorporated into the fermionic formulation as deformations of an initial condition by specific operators. In order to compare to the calculations in the previous chapters of this thesis what remains to be shown is how flows in the coupling constants of these operators are related to flows in the Toda lattice hierarchy and how these reduce in the dispersionless

[3]After this chapter was written, we became aware that related calculations can be found for a 5D gauge theory in [105]

limit to the flows studied there. More generally, it remains an open question how the variables in all the different approaches are connected. This is currently under study.

It is interesting to see that the τ function (5.63) is related to a 2D Yang-Mills theory, a 4D gauge theory, and a 0D matrix integral in one limit. By the results in this thesis, the integrable structure also has a connection to gauge theory in 3D. A very naive counting would then lead one to suspect that 1-dimensional gauge theory will show up somewhere. Actually, our fermion vev is remarkably similar to one obtained in the $c = 1$ string [106], which has a representation as a 1D matrix model.

An obvious direction from the results in this chapter is to further understand the connection between Dijkgraaf-Vafa and Seiberg-Witten by extending the results of this chapter. The calculations in this chapter seem to indicate there is a much larger (string theory) structure underlying all of this. Studying sub-leading effects (in $\hbar$) will reveal how they are related and the possibility exists that our dictionary may need sub-leading corrections. From a $\mathcal{N} = 2$ perspective, a natural question is the exact relationship of the connection explored in this chapter to the field theory origin of the calculations of Nekrasov. In particular, there a certain topologically twisted version of $\mathcal{N} = 2$ four dimensional gauge theory is used to arrive at his results, and it would be nice to know exactly how the operators studied above compare to operators in the underlying $\mathcal{N} = 2$ theory. There is a slight puzzle here due to calculations in [93] who use an expression similar to (5.63) to calculate the generating functional of vacuum expectation values for a $U(1)$ gauge theory. There is however a small discrepancy. In general, the partition function written above is for $SU(N_c)$ gauge theory and it would be nice to also understand the $U(1)$ factor studied in [93], which might be slightly more difficult to do in the dual formulation. This is expected to account for the discrepancy.

A feature of the calculations written above which also deserves further study is the fact that for most of these calculations the fermion charges are real, a complex extension would be welcome. From a 2D Yang-Mills perspective there are also a number of open questions. In particular, although the matrix model which arises there is the Hermitian matrix model, in the random partition model in this chapter there is a term proportional to $\xi_l \int f'$ which does not have a finite $\tilde{N}$ interpretation. A natural infinite $\tilde{N}$ interpretation of these may be that, from a 2D Yang-Mills perspective, they are connected to non-perturbative effects in the 't Hooft limit. This is also under study. From a $\mathcal{N} = 1$ or Dijkgraaf-Vafa perspective, there are a large number of results which beg for an interpretation in the framework of this chapter. In particular many results obtained in matrix model language might be re-derived. For instance the puzzle on the distribution of eigenvalues in the semi-classical limit from chapter (3) might have a natural resolution in the framework above. As indicated above, we have some indications that even the equations of motion imposed in Dijkgraaf-Vafa theory might be derived in some form from the fermion vev. We leave the study of this interesting possibility for future work.

Appendix A Hermitian matrix models and sums over partitions

In this appendix we discuss one possible application of the above analysis we have in mind: it may provide for a much faster way of calculating Dijkgraaf-Vafa superpotentials including contributions which are sub-leading in the 't Hooft limit. However, at present it is unclear how the formalism used in this appendix connects to the general formalism in the rest of this chapter. In general the question can be asked if Hermitian matrix models admit an expansion as a sum over partitions directly. An expansion of this sort can be found in the work [85]. They show that the two-matrix model

$$Z(N,t,t^*) = \frac{1}{Z(N,0,0)} \int dM_1^{N\times N} dM_2^{N\times N} e^{\mathrm{Tr}\,(V_1(M_1)+V_2(M_2)-M_1M_2)} \tag{5.67}$$

with an Hermitian matrix M_1 and anti-Hermitian matrix M_2 and potentials $V_1(x) = \sum_{i=1}^{\infty} t_i x^i$ $V_2(x) = \sum_{j=1}^{\infty} t_j^* x^j$ has a natural expansion in terms of Schur-functions as

$$Z(N,t,t^*) = \sum_{\lambda} (N)_\lambda s_\lambda(t) s_\lambda(t^*). \tag{5.68}$$

Here $\lambda = \{\lambda_1, \lambda_2, \ldots\}$ is a (uncolored) partition, s_λ is a Schur function and

$$(N)_\lambda = \prod_{i=1}^{l(\lambda)} \prod_{j=1}^{\lambda_i} (N - i + j), \tag{5.69}$$

with $l(\lambda)$ is the length of the partition. $(N)_\lambda$ can be interpreted as the measure. This peculiar relation is shown by reducing the two-matrix model to an eigenvalue model. The integrals which arise there can also be written as a certain inner product on the space of Schur-functions and comparing the expressions give the desired result. Let us stress that this is a finite N equivalence statement, not an asymptotic equivalence. Furthermore it is shown in [85] that this is a tau function of the Toda lattice hierarchy by showing that (5.68) has a fermionic formulation of the form (5.5).

With a quadratic potential for one side it is easy to reduce this to a one matrix model and the expansion in equation (5.68) can in this case be related to directly to the 't Hooft large N expansion [85]. Note that the expansion in (5.68) must then give all terms in the $1/N$ expansion, including the sub-leading terms. In order to compare to results in the literature, especially in [107], we specialize to a potential $W = \frac{m}{2}x^2 + \frac{g}{3}x^3$ and substitute $N = S$. Calculating the free energy then gives an expansion

$$\begin{aligned} F(N, g_2, g_3) &= \log\left(\sum_\lambda (N)_\lambda s_\lambda(\{0, \frac{1}{2m}, \ldots\}) s_\lambda(\{0, 0, -\frac{g}{3}, \ldots\})\right) \\ &= \frac{g^2}{m^3}\left(\frac{2}{3}S^3 + \frac{1}{6}S\right) + \frac{g^4}{m^6}\left(\frac{8}{3}S^4 + \frac{7}{3}S^2\right) + \\ &\quad + \frac{g^6}{m^9}\left(\frac{56}{3}S^5 + \frac{332}{9}S^3 + \frac{35}{6}S\right) + \\ &\quad + \frac{g^8}{m^{12}}\left(\frac{512}{3}S^6 + \frac{1864}{3}S^4 + 338S^2\right) + \text{h.o.} \end{aligned} \tag{5.70}$$

in perfect agreement with the coefficients of the single cut pieces in equation 4.9 in [107]. Note that the above expansion contains sub-leading terms in the $1/N$ expansion since $N = S$.

The above expression only gives the single-cut solution and it is not obvious how this can be generalized to multi-cut solutions. In fact, since this expression is derived at finite N and contains the full partition function, it is not obvious that it even should be possible to derive the asymptotic expansion of a multi-cut solution in this framework, although it clearly would be interesting if it were. The above method does provide for a very quick way of calculating sub-leading corrections for the single-cut answers and seems to be unknown in the physics community. It would be interesting to relate the above formulae to (5.63), which should reduce to the case studied in this appendix if all ξ_l are set equal.

Bibliography

[1] S. Weinberg, *The quantum theory of fields*, vol. II. Cambridge University Press, 1996.

[2] M. E. Peskin, *Duality in supersymmetric Yang-Mills theory*, [hep-th/9702094].

[3] N. Seiberg and E. Witten, *Electric - magnetic duality, monopole condensation, and confinement in $\mathcal{N}=2$ supersymmetric Yang-Mills theory*, *Nucl. Phys.* **B426** (1994) 19–52 [hep-th/9407087].

[4] N. Seiberg and E. Witten, *Monopoles, duality and chiral symmetry breaking in $\mathcal{N}=2$ supersymmetric QCD*, *Nucl. Phys.* **B431** (1994) 484–550 [hep-th/9408099].

[5] P. D. Lax, *Integrals of nonlinear equations of evolution and solitary waves*, *Comm. Pure Appl. Math.* **21** (1968) 467–490.

[6] D. I. Olive and N. Turok, *Algebraic structure of Toda systems*, *Nucl. Phys.* **B220** (1983) 491.

[7] M. Adler and P. van Moerbeke, *Linearization of Hamiltonian systems, Jacobi varieties and representation theory*, *Adv. in Math.* **38** (1980), no. 3 318–379.

[8] A. Das, *Integrable models*. World Scientific, 1989.

[9] V. G. Kac and J. W. van de Leur, *The n-component KP hierarchy and representation theory*, *J. Math. Phys.* **44** (2003), no. 8 3245–3293.

[10] J. D. Lykken, *Introduction to supersymmetry*, [hep-th/9612114].

[11] E. D'Hoker and D. H. Phong, *Lectures on supersymmetric yang-mills theory and integrable systems*, [hep-th/9912271].

[12] R. Boels, J. de Boer, R. Duivenvoorden and J. Wijnhout, *Nonperturbative superpotentials and compactification to three dimensions*, *JHEP* **03** (2004) 009 [hep-th/0304061].

[13] J. Wess and J. Bagger, *Supersymmetry and supergravity*. Princeton University Press, 1992.

[14] N. Seiberg, *Naturalness versus supersymmetric nonrenormalization theorems*, *Phys. Lett.* **B318** (1993) 469–475 [hep-ph/9309335].

[15] N. Arkani-Hamed and H. Murayama, *Holomorphy, rescaling anomalies and exact beta functions in supersymmetric gauge theories*, *JHEP* **06** (2000) 030 [hep-th/9707133].

[16] N. Seiberg, *Supersymmetry and nonperturbative beta functions*, *Phys. Lett.* **B206** (1988) 75.

[17] N. A. Nekrasov, *Seiberg-Witten prepotential from instanton counting*, *Adv. Theor. Math. Phys.* **7** (2004) 831–864 [hep-th/0206161].

[18] J. de Boer and Y. Oz, *Monopole condensation and confining phase of $\mathcal{N} = 1$ gauge theories via m-theory fivebrane*, *Nucl. Phys.* **B511** (1998) 155–196 [hep-th/9708044].

[19] R. Dijkgraaf, S. Gukov, V. A. Kazakov and C. Vafa, *Perturbative analysis of gauged matrix models*, *Phys. Rev.* **D68** (2003) 045007 [hep-th/0210238].

[20] G. Veneziano and S. Yankielowicz, *An effective lagrangian for the pure $\mathcal{N} = 1$ supersymmetric Yang-Mills theory*, *Phys. Lett.* **B113** (1982) 231.

[21] K. A. Intriligator, R. G. Leigh and N. Seiberg, *Exact superpotentials in four-dimensions*, *Phys. Rev.* **D50** (1994) 1092–1104 [hep-th/9403198].

[22] K. A. Intriligator, *'integrating in' and exact superpotentials in $4 - d$*, *Phys. Lett.* **B336** (1994) 409–414 [hep-th/9407106].

[23] F. Cachazo, K. A. Intriligator and C. Vafa, *A large N duality via a geometric transition*, *Nucl. Phys.* **B603** (2001) 3–41 [hep-th/0103067].

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

[25] R. Dijkgraaf and C. Vafa, *On geometry and matrix models*, *Nucl. Phys.* **B644** (2002) 21–39 [hep-th/0207106].

[26] R. Dijkgraaf and C. Vafa, *Matrix models, topological strings, and supersymmetric gauge theories*, *Nucl. Phys.* **B644** (2002) 3–20 [hep-th/0206255].

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

[28] J. O'Connor and E. Robertson, "Alexandre-Théophile Vandermonde." http://www-history.mcs.st-andrews.ac.uk/Mathematicians/Vandermonde.html.

[29] F. Cachazo, N. Seiberg and E. Witten, *Chiral rings and phases of supersymmetric gauge theories*, *JHEP* **04** (2003) 018 [hep-th/0303207].

[30] A. Gorsky, I. Krichever, A. Marshakov, A. Mironov and A. Morozov, *Integrability and Seiberg-Witten exact solution*, *Phys. Lett.* **B355** (1995) 466–474 [hep-th/9505035].

[31] E. J. Martinec and N. P. Warner, *Integrable systems and supersymmetric gauge theory*, *Nucl. Phys.* **B459** (1996) 97–112 [hep-th/9509161].

[32] A. Gorsky, S. Gukov and A. Mironov, *Multiscale $\mathcal{N} = 2$ susy field theories, integrable systems and their stringy/brane origin. i*, *Nucl. Phys.* **B517** (1998) 409–461 [hep-th/9707120].

[33] N. Seiberg and E. Witten, *Gauge dynamics and compactification to three dimensions*, [hep-th/9607163].

[34] J. de Boer, K. Hori and Y. Oz, *Dynamics of $\mathcal{N} = 2$ supersymmetric gauge theories in three dimensions*, *Nucl. Phys.* **B500** (1997) 163–191 [hep-th/9703100].

[35] O. Aharony, A. Hanany, K. A. Intriligator, N. Seiberg and M. J. Strassler, *Aspects of $\mathcal{N} = 2$ supersymmetric gauge theories in three dimensions*, *Nucl. Phys.* **B499** (1997) 67–99 [hep-th/9703110].

[36] C. Callias, *Axial anomalies and index theorems on open spaces*, *Comm. Math. Phys.* **62** (1978), no. 3 213–234.

[37] I. Affleck, J. A. Harvey and E. Witten, *Instantons and (super)symmetry breaking in $(2+1)$ dimensions*, *Nucl. Phys.* **B206** (1982) 413.

[38] K.-M. Lee and P. Yi, *Monopoles and instantons on partially compactified d-branes*, *Phys. Rev.* **D56** (1997) 3711–3717 [hep-th/9702107].

[39] N. Dorey, *An elliptic superpotential for softly broken $\mathcal{N} = 4$ supersymmetric Yang-Mills theory*, *JHEP* **07** (1999) 021 [hep-th/9906011].

[40] N. Dorey, T. J. Hollowood and S. Prem Kumar, *An exact elliptic superpotential for $\mathcal{N} = 1*$ deformations of finite $\mathcal{N} = 2$ gauge theories*, *Nucl. Phys.* **B624** (2002) 95–145 [hep-th/0108221].

[41] N. Dorey, T. J. Hollowood, S. Prem Kumar and A. Sinkovics, *Exact superpotentials from matrix models*, *JHEP* **11** (2002) 039 [hep-th/0209089].

[42] T. J. Hollowood, *Critical points of glueball superpotentials and equilibria of integrable systems*, *JHEP* **10** (2003) 051 [hep-th/0305023].

[43] M. Alishahiha and A. E. Mosaffa, *On effective superpotentials and compactification to three dimensions*, *JHEP* **05** (2003) 064 [hep-th/0304247].

[44] M. Alishahiha, J. de Boer, A. E. Mosaffa and J. Wijnhout, *$\mathcal{N}=1$ $G(2)$ SYM theory and compactification to three dimensions*, *JHEP* **09** (2003) 066 [hep-th/0308120].

[45] T. J. Hollowood and K. Ohta, *Vacua of $\mathcal{N}=1$ supersymmetric QCD from spin chains and matrix models*, *JHEP* **02** (2005) 042 [hep-th/0405051].

[46] T. J. Hollowood, *Five-dimensional gauge theories and quantum mechanical matrix models*, *JHEP* **03** (2003) 039 [hep-th/0302165].

[47] R. Dijkgraaf and C. Vafa, *$\mathcal{N}=1$ supersymmetry, deconstruction, and bosonic gauge theories*, [hep-th/0302011].

[48] F. Cachazo, M. R. Douglas, N. Seiberg and E. Witten, *Chiral rings and anomalies in supersymmetric gauge theory*, *JHEP* **12** (2002) 071 [hep-th/0211170].

[49] R. Gopakumar, *N = 1 theories and a geometric master field*, *JHEP* **05** (2003) 033 [hep-th/0211100].

[50] F. Cachazo, N. Seiberg and E. Witten, *Phases of $\mathcal{N}=1$ supersymmetric gauge theories and matrices*, *JHEP* **02** (2003) 042 [hep-th/0301006].

[51] N. J. A. Sloane, *The on-line encyclopedia of integer sequences*, *published electronically at http://www.research.att.com/ njas/sequences/* (2005) sequence number A078531.

[52] M. Aganagic and C. Vafa, *Perturbative derivation of mirror symmetry*, [hep-th/0209138].

[53] P. H. Ginsparg, *Matrix models of 2-d gravity*, [hep-th/9112013].

[54] P. van Moerbeke, *Random matrices and permutations, matrix integrals and integrable systems*, *Astérisque* (2002), no. 276 411–433. Séminaire Bourbaki, Vol. 1999/2000.

[55] P. Zinn-Justin and J.-B. Zuber, *On some integrals over the $U(N)$ unitary group and their large N limit*, *J.Phys A* **36** (2003) 3173–3194.

[56] M. Kac and P. van Moerbeke, *A complete solution of the periodic Toda problem*, *Proc. Nat. Acad. Sci. U. S. A.* **72** (1975), no. 8 2879–2880.

[57] E. Date and S. Tanaka, *Analogue of inverse scattering theory for the discrete Hill's equation and exact solutions for the periodic Toda lattice*, *Progr. Theoret. Phys.* **55** (1976), no. 2 457–465.

[58] P. van Moerbeke and D. Mumford, *The spectrum of difference operators and algebraic curves*, *Acta Math.* **143** (1979), no. 1-2 93–154.

[59] A. Gorsky, A. Marshakov, A. Mironov and A. Morozov, *$\mathcal{N} = 2$ supersymmetric QCD and integrable spin chains: Rational case $N_f < 2N_c$*, *Phys. Lett.* **B380** (1996) 75–80 [hep-th/9603140].

[60] R. Donagi and E. Witten, *Supersymmetric Yang-Mills theory and integrable systems*, *Nucl. Phys.* **B460** (1996) 299–334 [hep-th/9510101].

[61] N. Nekrasov, *Five dimensional gauge theories and relativistic integrable systems*, *Nucl. Phys.* **B531** (1998) 323–344 [hep-th/9609219].

[62] A. Kapustin, *Solution of $\mathcal{N} = 2$ gauge theories via compactification to three dimensions*, *Nucl. Phys.* **B534** (1998) 531–545 [hep-th/9804069].

[63] J. S. Wijnhout, *Gauge theories in $3\&1$ dimensions*. PhD thesis, University of Amsterdam, 2005.

[64] H. Itoyama and A. Morozov, *Gluino-condensate (CIV-DV) prepotential from its Whitham- time derivatives*, *Int. J. Mod. Phys.* **A18** (2003) 5889–5906 [hep-th/0301136].

[65] K. Splittorff and J. J. M. Verbaarschot, *Replica limit of the Toda lattice equation*, *Phys. Rev. Lett.* **90** (2003) 041601 [cond-mat/0209594].

[66] R. Boels and J. de Boer, *Classical spin chains and exact three-dimensional superpotentials*, *Nucl. Phys.* **B715** (2005) 234–274 [hep-th/0411110].

[67] P. C. Argyres, M. R. Plesser and N. Seiberg, *The moduli space of $\mathcal{N} = 2$ SUSY QCD and duality in $\mathcal{N} = 1$ SUSY QCD*, *Nucl. Phys.* **B471** (1996) 159–194 [hep-th/9603042].

[68] G. Carlino, K. Konishi and H. Murayama, *Dynamical symmetry breaking in supersymmetric $SU(n(c))$ and $USp(2n(c))$ gauge theories*, *Nucl. Phys.* **B590** (2000) 37–122 [hep-th/0005076].

[69] V. Balasubramanian, B. Feng, M.-x. Huang and A. Naqvi, *Phases of $\mathcal{N} = 1$ supersymmetric gauge theories with flavors*, *Annals Phys.* **310** (2004) 375–427 [hep-th/0303065].

[70] N. Seiberg, *Electric - magnetic duality in supersymmetric non-Abelian gauge theories*, *Nucl. Phys.* **B435** (1995) 129–146 [hep-th/9411149].

[71] M. A. Shifman, *Nonperturbative dynamics in supersymmetric gauge theories*, *Prog. Part. Nucl. Phys.* **39** (1997) 1–116 [hep-th/9704114].

[72] I. Affleck, M. Dine and N. Seiberg, *Dynamical supersymmetry breaking in supersymmetric QCD*, *Nucl. Phys.* **B241** (1984) 493–534.

[73] K. Hori, H. Ooguri and Y. Oz, *Strong coupling dynamics of four-dimensional $\mathcal{N} = 1$ gauge theories from m theory fivebrane*, *Adv. Theor. Math. Phys.* **1** (1998) 1–52 [hep-th/9706082].

[74] E. D'Hoker, I. M. Krichever and D. H. Phong, *The effective prepotential of $\mathcal{N} = 2$ supersymmetric $SU(N(c))$ gauge theories*, *Nucl. Phys.* **B489** (1997) 179–210 [hep-th/9609041].

[75] S. G. Naculich, H. J. Schnitzer and N. Wyllard, *The $\mathcal{N} = 2$ U(N) gauge theory prepotential and periods from a perturbative matrix model calculation*, *Nucl. Phys.* **B651** (2003) 106–124 [hep-th/0211123].

[76] Y. Ookouchi, *$\mathcal{N} = 1$ gauge theory with flavor from fluxes*, *JHEP* **01** (2004) 014 [hep-th/0211287].

[77] O. Babelon, D. Bernard and M. Talon, *Introduction to classical integrable systems*. Cambridge Monographs on Mathematical Physics. Cambridge University Press, Cambridge, 2003.

[78] P. C. Argyres, M. R. Plesser and A. D. Shapere, *The Coulomb phase of $\mathcal{N} = 2$ supersymmetric QCD*, *Phys. Rev. Lett.* **75** (1995) 1699–1702 [hep-th/9505100].

[79] K. D. Kennaway and N. P. Warner, *Effective superpotentials, geometry and integrable systems*, [hep-th/0312077].

[80] Y. Demasure and R. A. Janik, *Explicit factorization of Seiberg-Witten curves with matter from random matrix models, Nucl. Phys.* **B661** (2003) 153–173 [hep-th/0212212].

[81] Y. Demasure, *Affleck-Dine-Seiberg from Seiberg-Witten*, [hep-th/0307082].

[82] M. Gomez-Reino, *Exact superpotentials, theories with flavor and confining vacua*, *JHEP* **06** (2004) 051 [hep-th/0405242].

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

[84] I. K. Kostov, *Conformal field theory techniques in random matrix models*, [hep-th/9907060].

[85] J. Harnad and A. Y. Orlov, *Scalar products of symmetric functions and matrix integrals*, *Theor. Math. Phys.* **137** (2003) 1676–1690 [nlin.si/0211051].

[86] T. Miwa, M. Jimbo and E. Date, *Solitons*, vol. 135 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 2000.

[87] K. Ueno and K. Takasaki, *Toda lattice hierarchy*, in *Group representations and systems of differential equations (Tokyo, 1982)*, vol. 4 of *Adv. Stud. Pure Math.*, pp. 1–95. North-Holland, Amsterdam, 1984.

[88] N. Nekrasov and A. Okounkov, *Seiberg-Witten theory and random partitions*, [hep-th/0306238].

[89] M. Matone, *Seiberg-Witten duality in Dijkgraaf-Vafa theory*, *Nucl. Phys.* **B656** (2003) 78–92 [hep-th/0212253].

[90] T. Matsuo, S. Matsuura and K. Ohta, *Large N limit of 2D Yang-Mills theory and instanton counting*, *JHEP* **03** (2005) 027 [hep-th/0406191].

[91] S. Matsuura and K. Ohta, *Localization on D-brane and gauge theory / matrix model*, [hep-th/0504176].

[92] I. G. Macdonald, *Symmetric functions and Hall polynomials*. The Clarendon Press Oxford University Press, New York, 1979. Oxford Mathematical Monographs.

[93] A. S. Losev, A. Marshakov and N. A. Nekrasov, *Small instantons, little strings and free fermions*, [hep-th/0302191].

[94] K. Takasaki and T. Takebe, *Integrable hierarchies and dispersionless limit*, *Rev. Math. Phys.* **7** (1995) 743–808 [hep-th/9405096].

[95] K. Takasaki and T. Nakatsu, *Isomonodromic deformations and supersymmetric gauge theories*, *Int. J. Mod. Phys.* **A11** (1996) 5505–5518 [hep-th/9603069].

[96] K. Ueno and K. Takasaki, *Toda lattice hierarchy*, in *Group representations and systems of differential equations (Tokyo, 1982)*, vol. 4 of *Adv. Stud. Pure Math.*, pp. 1–95. North-Holland, Amsterdam, 1984.

[97] V. Kazakov, I. K. Kostov and D. Kutasov, *A matrix model for the two-dimensional black hole*, *Nucl. Phys.* **B622** (2002) 141–188 [hep-th/0101011].

[98] M. Blau and G. Thompson, *Lectures on 2-d gauge theories: Topological aspects and path integral techniques*, [hep-th/9310144].

[99] B. E. Rusakov, *Loop averages and partition functions in $U(N)$ gauge theory on two-dimensional manifolds*, *Mod. Phys. Lett.* **A5** (1990) 693–703.

[100] M. R. Douglas and V. A. Kazakov, *Large N phase transition in continuum QCD in two dimensions*, *Phys. Lett.* **B319** (1993) 219–230 [hep-th/9305047].

[101] B. Rusakov and S. Yankielowicz, *Large N phase transitions and multicritical behavior in generalized 2-D QCD*, *Phys. Lett.* **B339** (1994) 258–262 [hep-th/9408046].

[102] R. Dijkgraaf, R. Gopakumar, H. Ooguri and C. Vafa, *Baby universes in string theory*, [hep-th/0504221].

[103] D. J. Gross and E. Witten, *Possible third order phase transition in the large N lattice gauge theory*, *Phys. Rev.* **D21** (1980) 446–453.

[104] A. Okounkov and R. Pandharipande, *The equivariant Gromov-Witten theory of P^1*, [math.ag/0207233].

[105] T. Maeda, T. Nakatsu, K. Takasaki and T. Tamakoshi, *Free fermion and Seiberg-Witten differential in random plane partitions*, *Nucl. Phys.* **B715** (2005) 275–303 [hep-th/0412329].

[106] R. Dijkgraaf, G. W. Moore and R. Plesser, *The partition function of 2-d string theory*, *Nucl. Phys.* **B394** (1993) 356–382 [hep-th/9208031].

[107] A. Klemm, M. Marino and S. Theisen, *Gravitational corrections in supersymmetric gauge theory and matrix models*, *JHEP* **03** (2003) 051 [hep-th/0211216].