Gauge theories in 3 & 1 dimensions. A string inspired approach to supersymmetricGauge theories, D0-branes in a curved background

Wijnhout, J.S.

Publication date
2005

Document Version
Final published version

Gauge Theories in 3 & 1 Dimensions

A String Inspired Approach to Supersymmetric Gauge Theories, D0-branes in a Curved Background

Gauge Theories in 3 & 1 Dimensions

A String Inspired Approach to Supersymmetric Gauge Theories, D0-branes in a Curved Background

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 woensdag 20 april 2005, te 12:00 uur

door

Jeroen Sebastian Wijnhout

geboren te Raalte.

PROMOTIECOMMISSIE

PROMOTOR

prof. dr Jan de Boer

OVERIGE LEDEN

prof. dr Robbert Dijkgraaf

prof. dr Rob Myers

prof. dr Alexander Sevrin

dr Kostas Skenderis

prof. dr Jan Smit

prof. dr Erik Verlinde

FACULTEIT DER NATUURWETENSCHAPPEN, WISKUNDE EN INFORMATICA

This thesis was produced using Kile, a LaTeX editor for UNIX.

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)".

PUBLICATIONS

This thesis is based on the following publications:

① R. Boels, J. de Boer, R. Duivenvoorden and J. Wijnhout,
Nonperturbative superpotentials and compactification to three dimensions,
JHEP **03** (2004) 009 [hep-th/0304061].

② R. Boels, J. de Boer, R. Duivenvoorden and J. Wijnhout,
Factorization of seiberg-witten curves and compactification to three dimensions,
JHEP **03** (2004) 010 [hep-th/0305189].

③ M. Alishahiha, J. de Boer, A. E. Mosaffa and J. Wijnhout,
N = 1 g(2) sym theory and compactification to three dimensions,
JHEP **09** (2003) 066 [hep-th/0308120].

④ J. de Boer, E. Gimon, K. Schalm and J. Wijnhout,
Evidence for a gravitational myers effect,
Annals of Physics **313/2** (2004) 402–424 [hep-th/0212250].

⑤ J. de Boer, K. Schalm and J. Wijnhout,
General covariance of the non-abelian dbi-action: Checks and balances,
Annals of Physics **313/2** (2004) 425–445 [hep-th/0310150].

CONTENTS

PREFACE

One of the remarkable properties of string theory is the extra dimensions it predicts. In contrast to this, this thesis will primarily deal with string theory inspired models that live in *less* than four dimensions. Two such theories will be discussed, one is a supersymmetric three-dimensional gauge theory, the other a one-dimensional bosonic gauge theory coupled to gravity. Each theory will be treated in a separate part.

Part one deals with the supersymmetric gauge theories. These three-dimensional gauge theories have the remarkable property that certain exactly calculable quantities predict exact quantum properties of a related four-dimensional gauge theory. An important role is played by a classical integrable system, this role will be elucidated by relating the gauge theories to a D-brane setup.

Part two is about the interaction of D0-branes with gravity. The action for multiple D0-branes, a one-dimensional gauge theory, is difficult to determine due to ordering ambiguities. The application of a new symmetry principle is used to partially solve this problem. Also, evidence is found for interesting collective behavior of D0-branes in a curved background: a gravitational Myers effect.

Part I

Non-perturbative results in supersymmetric gauge theories

Chapter 1

Introduction and motivation

The first part of this thesis deals with gauge theories, supersymmetric gauge theories in particular. The goal of this chapter is to explain why these supersymmetric gauge theories are worth studying. One reason is one of the big open questions in high-energy physics: Finding an analytic description of the vacuum structure of QCD. Allthough the string theory inspired supersymmetric gauge theories, that will be presented in this thesis, are far from giving satisfactory explanations, they do provide a valuable laboratory where much can be learned about gauge theories in general. This chapter will start by briefly reviewing the main issues in QCD, gradually building up to recent advances in supersymmetric gauge theories.

Quantum Chromo Dynamics

The gauge theory of the strong interaction, Quantum Chromo Dynamics or simply QCD, has the special property of asymptotic freedom: at very high energies the theory resembles a free field theory, at low energies the theory is strongly coupled. Precisely the strongly coupled nature of QCD in the infrared (IR) makes it difficult to give a satisfactory theoretical description of the low energy quantum dynamics. Afterall perturbation theory, the favorite tool of most high-energy physicists, becomes cumbersome at strong coupling.

At this point in time it is not even possible to derive the properties of the vacuum of QCD from first principles. However, from experimental data and general arguments, one knows some of the properties it should have. The first property that comes to mind is confinement: although the elementary fields in QCD are the quarks and gluons, physical states are always colorless (i.e. baryons, mesons or glueballs).

A different property, although not completely unrelated, is the existence of a mass gap, which basically means that there are no states in the spectrum at zero energy. In particular this means that the strong force is a short-distance force (in reverse, in the absence of a mass gap, massless gluons would mediate a long-range force which does not speak in favor of confinement). A general consequence of a mass gap is the introduction of a mass scale in the theory. Note that classically, Yang-Mills theory in four-dimensions is scale invariant (and therefore cannot have an intrinsic mass scale classically), the introduction of a mass-scale is therefore a pure quantum effect.

The third property of the vacuum that is worth mentioning is that of chiral symmetry breaking. Take for example QCD for just two quarks (u and d), this has a global $SU(2)$ symmetry that interchanges the two flavors. However, in the limit of massless quarks, this symmetry is enhanced to $SU(2) \times SU(2)$, since the left- and right-handed parts of the massless spinors can be rotated independently. This chiral symmetry is spontaneously broken in QCD, the Goldstone bosons being the pions. At high energy the chiral symmetry should be restored. Since the breaking of the chiral symmetry is intimately connected with the existence of a mass gap (and therefore also confinement), it will be interesting to understand the mechanism of chiral symmetry breaking and restoration more thoroughly.

The above three properties of the vacuum of QCD are still ill understood from the theoretical point of view. It will be a great challenge to solve these issues.

SUPERSYMMETRIC GAUGE THEORIES

This thesis will not be focused on QCD, but rather on supersymmetric gauge theories. Supersymmetric gauge theories are interesting for many reasons. Many Grand Unifying Theories (GUTs), for example, need supersymmetry to achieve gauge coupling unification. Extensions to the Standard Model often involve supersymmetry, not unimportant in this respect is the fact that supersymmetry can, via loop cancellations, control radiative corrections to particle masses and can therefore be a solution to the hierarchy problem.

One can also view supersymmetric gauge theories as a laboratory to study QCD, viewing QCD as a perturbation from a supersymmetric gauge theory. Knowledge of supersymmetric gauge theories can then possibly be used to learn something about QCD. String theory in the form of D-brane setups, provides many excellent laboratories. The D-brane setups, for example, provide alternative perspectives one can use to describe gauge theories.

Besides being physically interesting, supersymmetric gauge theories are also easier to deal with. Powerful symmetry arguments sometimes lead to exact non-perturbative results. Take for example the superpotential for a theory involving only a chiral superfield; using supergraph techniques one can show that the superpotential is not modified perturbatively. This result can also be derived, following Seiberg, using the fact that the superpotential is holomorphic in the fields and coupling constants [1]. Sometimes holomorphy in combination with global symmetries of the superpotential allows one to derive the exact quantum superpotential (as is the case for the Wess-Zumino model). Note that the holomorphy argument by Seiberg only applies when the theory can be regularized in such a way that global symmetries are preserved.

Supersymmetry thus constrains the quantum corrections to the superpotential. The kinetic terms however are not protected by holomorphy and can receive many perturbative and non-perturbative corrections. Thus the power of holomorphy is the most useful in situations where the kinetic terms are not important, such as in the study of the vacua of supersymmetric gauge theories.

More symmetry usually means more analytic control, this is the case with supersymmetric gauge theories as well. The gauge theory with extended supersymmetry ($\mathcal{N} = 2$) can be shown to have an electric-magnetic duality symmetry. This symmetry interchanges the electric and magnetic charges in the theory and at the same time switches between strong and weak coupling. To understand electric confinement (such as the confinement of quarks and gluons, which are the "electric" degrees of freedom) it is necessary to understand the low energy, strongly coupled regime of the theory. Using the electric-magnetic duality one can turn this into a problem in the weakly coupled dual theory, where one can use perturbation theory to address the problem of confinement [2].

Lagrangians for gauge theories with $\mathcal{N} = 2$ supersymmetry are determined by the prepotential, a holomorphic function from which the Kähler potential, gauge couplings and scalar potential can be determined. Seiberg and Witten [2] showed that the problem of finding the non-perturbative prepotential can be turned into the problem of computing certain quantities on an elliptic curve. So out of this quantum gauge theory an effective geometry emerges that captures the strong coupling phenomena. For the gauge group $SU(2)$ the elliptic curve is a torus, with modulus τ. This modulus can be identified with the gauge coupling, both are controlled by the vacuum expectation value of the scalar field of the $\mathcal{N} = 2$ chiral multiplet.

Recent progress

Phenomenologically $\mathcal{N} = 2$ theories are not that interesting because $\mathcal{N} = 2$ multiplets can not contain chiral fermions. $\mathcal{N} = 1$ theories do allow chiral fermions, however since they have less supersymmetry it is also harder to deal with those theories. Still, having $\mathcal{N}=1$ supersymmetry does constrain the possible corrections to the superpotential significantly. Using the holomorphy argument of Seiberg [1] one can show that the superpotential has a one-loop perturbative contribution and that all other contributions are non-perturbative. Determining those non-perturbative contributions can be quite difficult.

Surprisingly recent progress, by Dijkgraaf and Vafa [3], shows that it is possible to determine the effective superpotential, in terms of the glueball superfield, by doing a perturbative computation in a zero-dimensional theory (a so-called Matrix Model). The end-result is an exact expression in the sense that the expectation values for the glueball fields are written as a non-perturbative summation over instanton contributions. Whether or not this effective action is a genuine Wilsonian effective action is not clear.

Roughly the argument, leading to this result, is as follows. If one is interested in calculating the exact superpotential in a background of gauge fields, the path-integral can be reduced to such a form that only planar diagrams in the zero-momentum limit contribute: the path-integral reduces to a matrix integral in the 't Hooft limit (the large N limit is not performed on the rank of the original gauge symmetry, therefore the results are valid for the gauge symmetry one started with). The superpotential can then be expressed, term by term, in the gauge field background. The end-result is a series of *perturbative* corrections to the superpotential. Note that this is not in conflict with the non-renormalization theorem of Seiberg, since the $\mathcal{N}=1$ theory has an anomalous $U(1)_R$ symmetry. After minimizing this superpotential with respect to the gauge field background the superpotential can be expressed as a power series in the characteristic energy scale (the scale reminiscent of dimensional transmutation):

$$W \sim \Lambda^3 + \ldots \sim \Lambda_0^3 \mathrm{e}^{-\frac{8\pi^2}{g^2}} + \ldots, \tag{1.1}$$

showing the non-perturbative nature of the calculations. Allthough the original calculation was done using a chain of dualities in string theory, the argument can be completely given in field theory terms [4].

Actually, it turns out this is only a special example of a more general scheme: $d-4$-dimensional theories can be used to calculate superpotentials of d-dimensional theories [5]. A three-dimensional gauge theory would correspond to a -1-dimensional theory (or algebraic theory, no integrals). Three-dimensional theories, however, are not as interesting as four-dimensional theories. Fortunately, some three-dimensional theories do predict the correct results for holomorphic quantities in related four-dimensional theories. The main focus of the first part of this thesis will be a specific example of such a three-dimensional theory: a four-dimensional $\mathcal{N}=1$ theory compactified on a circle. As will be shown, the algebraic theory that underlies the three-dimensional theory (and hence aspects of the four-dimensional theory) is a classical mechanical integrable system.

The following series of chapters will start off by giving some necessary background material about (supersymmetric) gauge theories and integrable systems. All is necessary to understand the main result of this part of the thesis: Calculating four-dimensional superpotentials using an underlying integrable system. Chapter 2 will be used to review some relevant aspects of gauge theories, nothing new, but hopefully helpful to starters in the field. Things will become a bit more technical in the following chapters 3 and 4 dealing with necessary background material about supersymmetric gauge theories and integrable systems. These two chapters are necessary to understand the new material presented in chapter 5, which is about calculating quantum effective superpotentials using integrable systems.

CHAPTER 2

GAUGE THEORY

In this chapter a few well-known results, obtained in the theory of gauge fields, will be reviewed. The material presented in this chapter is not new, but it is relevant background material for the following chapters.

(2.1) LIMITATIONS OF PERTURBATION THEORY

The reason why it is difficult to determine the vacuum properties of quantum gauge theories, QCD in particular, is that at low energy the theory is strongly coupled. This means that perturbation theory breaks down and non-perturbative effects become essential. A good example is the instanton solution, a single instanton contributes a factor

$$e^{-\frac{8\pi^2}{g^2}} \tag{2.1}$$

to the path-integral. This expression does not have a Taylor expansion around $g = 0$, hence the instanton effect can never be seen in perturbation theory: perturbation theory is incomplete and exact quantities can never be obtained using this method.

One might hope that, at least, the summation over all Feynman diagrams yields results that capture all the non-perturbative physics. Even this is not the case since a perturbation series like

$$\sum_{n=0}^{\infty} a_n g^n \tag{2.2}$$

will only converge asymptotically since the number of Feynman diagrams at n-th order grows like $n!$, therefore

$$a_n \sim n!. \tag{2.3}$$

The fact that the number of diagrams at n-th order behaves like this can of course be checked using combinatorics. One can also make use of the fact that the path-integral is a generating functional for all Feynman diagrams. Since only the number of Feynman diagrams is relevant, it is sufficient to simply evaluate the path-integral in the zero-momentum limit, reducing it to an ordinary integral. Take for example Φ^4 theory, the path-integral in this case reduces to (see [6])

$$Z(g) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2 - \frac{1}{4}gx^4} dx = \sum Z_k g^k. \tag{2.4}$$

The numbers Z_k are an indication of the number of diagrams at k-th order. Remember that for this theory, the number of loops l is given by $l = k - 1$. The integral in equation 2.4 can be estimated for large k. Write

$$Z_k = \frac{1}{\sqrt{2\pi}k!} \int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2} \left(-\frac{x^4}{4}\right)^k dx \tag{2.5}$$

and note that this is an integral representation of the Γ function. In the limit of large k this can be written as

$$Z_k \sim \frac{(-1)^k}{\pi\sqrt{2}} 4^k (k-1)!, \tag{2.6}$$

confirming the large order behavior.

It was realized by ’t Hooft [7] that the perturbation series improve for $SU(N)$ at large N, if one uses the coupling $\lambda = g^2 N$. Then only planar diagrams contribute and the number of diagrams grows as a power law

$$a_n \sim c^n. \tag{2.7}$$

Perturbation series in λ can be analytic and can have a finite radius of convergence, if there is no essential singularity at $g = 0$ (that is, any singularity must behave like a pole). One can use this technique to learn about QCD or other quantum gauge theories. To show that the number of diagrams indeed grows as a power law, the same strategy as above is used. Take the path-integral for Φ^4 theory, with Φ an $N \times N$ matrix. In the large N limit the path-integral is a generating functional for planar diagrams only [8]

$$\mathrm{e}^{-N^2 F(g)} = \int d\Phi \mathrm{e}^{-\frac{1}{2}\mathrm{Tr}\,\Phi^2 - \frac{g}{N}\mathrm{Tr}\,\Phi^4}, \tag{2.8}$$

where F is the generating function for connected planar diagrams. Note that this integral can be evaluated and that at large order it results in

$$F(g)_k \sim (48)^k \tag{2.9}$$

reflecting that the number of connected planar diagrams grows as a power law.

(2.2) CONFINEMENT

CONFINEMENT OR MAXIMAL SCREENING?

One of the most prominent features of QCD is confinement; the effective degrees of freedom are color singlets, whereas the fundamental degrees of freedom carry color charge. Though ill understood from a purely theoretical point of view, this phenomenon is often attributed to the existence of flux tubes and a mass gap. However, the issue of confinement (as observed in experiments) is a bit more subtle. The subtlety lies in the fact that confinement can also occur through the mechanism of "maximal screening". Usually, however, by confinement is meant the confinement of flux into flux tubes. This mechanism is still not understood completely. The two mechanisms will now be described briefly, one by one.

Maximal screening is a completely straightforward and well-understood mechanism (see for example [9] for a review). Consider two quarks and start to pull them apart slowly. At some point the distance between the quarks reaches the QCD scale, i.e. the scale at which the coupling becomes strong. The energy density of the gluon field between the quarks can then be estimated as follows

$$\mathcal{E} \sim g^2/r^4 \sim g^2\Lambda^4_{QCD} \sim \Lambda^4_{QCD}. \tag{2.10}$$

It turns out that this energy density is enough to create a quark–anti-quark pair, since

$$m_q^4 \sim \Lambda^4_{QCD}. \tag{2.11}$$

This effect is called maximal screening, the original quark pair is screened completely because pulling apart the quarks creates a gluon field energetic enough to pair produce quarks. This game can even be played with QED, provided the coupling is high enough (which it isn't in reality).

The question arises what is meant exactly with the confinement due to flux tubes. When flux tubes form, the potential between the quarks becomes linear. This can be easily seen using a Gaussian surface around the flux tube (see figure 2.1):

$$Q_{enc} = \oint_S \vec{E} \cdot \mathrm{d}\vec{\mathcal{A}} \Rightarrow E \propto q/A \Rightarrow V \propto d. \tag{2.12}$$

Here A is the surface of the Gaussion box. The linear potential due to the flux *tube* between the quarks means that it requires a constant force to keep them apart. As the coupling constant increases with increasing distance (asymptotic freedom), it is impossible to separate the quarks completely. Now it is possible to state what is usually meant by "proving confinement". Confinement is proven if the existence of flux tubes can be shown. So this is really confinement of flux and not solely

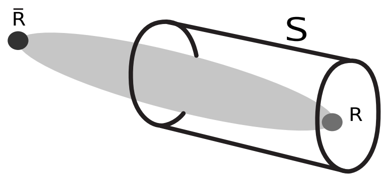


Figure 2.1: *A flux tube between two color charges.*

the confinement of charges (which can also occur through the maximal screening mechanism discussed above).

FLUX TUBES

Flux tubes can be characterized by their end-points (i.e. the representation R and $\bar{R}$ of the quarks at both end-points, see figure 2.1). To characterize the flux tube it is not sufficient to give the representation R. This is because the representation is not a conserved quantity in the following sense. Gluon number is not a conserved quantity, therefore one can always take a gluon from the flux tube and combine it with the charge R to make it look like a different representation. This new configuration possibly has a lower tension (flux tube energy density), causing the system to decay to this energetically more favorable flux tube. One quantity is conserved however: the charge(s) under the center of the gauge group. Those charge(s) can be used to classify the flux tubes. As an example consider $SU(N)$, the center consists of elements

$$U = \mathrm{e}^{2\pi \mathrm{i} k/N}\mathbb{I}, \quad k = 0, \dots, N-1, \tag{2.13}$$

which together form the Abelian group $\mathbb{Z}_N$. For $SU(N)$ any representation can be built out of direct products of the fundamental N and anti-fundamental $\bar{N}$ representations. Because the anti-fundamental representation is the complex conjugate of the fundamental representation, it has opposite charge. The adjoint representation is built out of $N \otimes \bar{N}$, therefore it has charge zero. Combining any representation $\mathcal{R}$ with the adjoint representation in a direct product will not affect the total charge under the center. It will remain equal to the charge that corresponds to $\mathcal{R}$.

To each $\mathcal{R}$ one can associate a flux tube and to each flux tube one associates a tension $T_{\mathcal{R}}$. As explained above, labeling the flux tube using the representation $\mathcal{R}$ is awkward since one representation can turn into another simply by combining with one of the gluons in the flux tube. If this combination $\mathcal{R} \otimes adj$ contains a representation belonging to a flux tube with lower tension, then it is energetically favorable for the system to decay to this less energetic flux tube. In the case the direct product contains a singlet, the flux tube disappears alltogether. A special case is the adjoint representation, since $adj \otimes adj$ always contains a singlet, there cannot be any stable fluxtubes between adjoint matter. Groups that have a trivial center cannot have stable flux tubes at all.

THE MEISSNER EFFECT

Popular explanations of the existence of flux tubes in QCD use the so-called dual Meissner effect. The Meissner effect is a phenomenon of superconducting materials. In the superconducting phase pairs of electrons (Cooper pairs), described by a complex scalar field ϕ, condense. This superconducting condensate has some special properties: If an external magnetic field is applied a current is induced, creating an opposite magnetic field. The external magnetic field is expelled, although flux tubes form through which the magnetic field can penetrate the material. Within the flux tube the material is in the non-superconducting phase, allowing a magnetic field to be present. One can say that magnetic flux is confined in this system.

The flux tubes in this condensed matter system, Abrikosov vortices, can be described using a spontaneously broken $U(1)$ scalar field theory. In the context of relativistic field theories a $2{+}1$-dimensional Abelian Higgs model should be used. The existence of vortices in this theory was shown by Nielsen and Olesen [10]. The Lagrangian of the Abelian Higgs model reads

$$\mathcal{L} = -\tfrac{1}{4g^2}F^2 + (D\phi)^2 - V(\phi^\dagger\phi). \tag{2.14}$$

The potential is such that for field configurations satisfying

$$\phi^\dagger\phi = v^2 \Rightarrow \phi = v\mathrm{e}^{\mathrm{i}\theta n} \tag{2.15}$$

the minimum is reached. From this equation one deduces that n is an integer. In this minimum there are no massless degrees of freedom anymore. This means a Yukawa type potential for the electromagnetic interaction, hence electric fields are screened. There is a problem with this vacuum configuration: it is not a finite energy configuration, as the kinetic energy of the scalar field is not finite

$$E \approx \int \mathrm{d}V(\partial_\theta\phi)^2 \sim \pi n^2 v^2 \int \frac{\mathrm{d}r}{r} \to \infty. \tag{2.16}$$

A finite energy configuration can be created by modifying the value of the vector potential appropriately, such that the covariant derivative vanishes at infinity

$$D_\theta\phi = \tfrac{1}{r}\partial_\theta\phi + \mathrm{i}A_\theta\phi = 0 \Rightarrow \tfrac{1}{r}\partial_\theta\phi = -\mathrm{i}A_\theta\phi. \tag{2.17}$$

Combining this with equation 2.15 one gets

$$A_\theta = -n/r \Rightarrow \text{flux: } \Phi = \int r \mathrm{d}\theta A_\theta = -2\pi n, \tag{2.18}$$

showing that the magnetic flux is quantized.

The spontaneous breaking from $U(1) \to \mathbb{I}$, allows the appearance of vortex solutions. A more fancy way of stating this result is that this configuration is described by a map from the boundary ∂S, a circle, to a $U(1)$ phase. The tubes are topologically stable, with the topological charge being

$$\pi_1[U(1)] = \mathbb{Z}. \tag{2.19}$$

In this case the topological charge is just the integer number n. The general mechanism is as follows. When the symmetry group G is spontaneously broken to a subgroup H

$$G \to H, \tag{2.20}$$

then the topological charge is given by

$$\pi_1[G/H]. \tag{2.21}$$

So, flux tubes exist if $\pi_1[G/H]$ is non-trivial. If G is simply connected this is equal to $\pi_0[H]$.

If the Higgs mechanism is applied to an adjoint scalar field the symmetry is broken to the center of the gauge group. For example if $SU(N)$ is broken to its center $\mathbb{Z}_N$, the topological charge is

$$\pi_1[SU(N)/\mathbb{Z}_N] = \mathbb{Z}_N, \tag{2.22}$$

which corresponds to the charge classifying flux tubes, discussed before.

The central idea is now to copy this mechanism over to gauge theories with the electric and magnetic charges interchanged: a dual Meissner effect. The fundamental magnetic degrees of freedom will be magnetic monopoles, which will therefore be discussed before the mechanism of confinement in gauge theories is discussed.

MONOPOLES

For a theory to have monopoles á là 't Hooft-Polyakov, a breaking pattern like $SU(2)$ to $U(1)$ should occur. The appropriate model is still the Abelian Higgs model, see equation 2.14, but now in $3+1$ dimensions. For stable point-like topological charges to exist the second fundamental group of the vacuum manifold M should be non-trivial

$$\pi_2(M). \tag{2.23}$$

At infinity fields should be pure gauge, so the field configuration there is characterized by a map from the sphere at infinity to the moduli space (vacuum manifold). If the gauge symmetry is broken from G to H then the vacuum manifold is given by $M = G/H$.

The 't Hooft-Polyakov monopole solution with charge one is given by the following "hedgehog" configuration

$$\phi(\vec{r}) = \eta \frac{\vec{r}\cdot\vec{\sigma}}{r} f(r), \qquad f(0) = 0, \quad f(\infty) = 1 \tag{2.24}$$

$$\vec{A}(\vec{r}) = \frac{\vec{r}\times\vec{\sigma}}{r^2} a(r), \qquad a(0) = 0, \quad a(\infty) = 1. \tag{2.25}$$

The solutions for the smooth functions f and a are only known in certain limiting cases (such as the Prasad-Sommerfeld monopole, obtained by saturating the Bogomolnyi bound).

The hedgehog configuration can be gauge rotated to the form $\phi \propto \sigma_3$ if one uses the gauge transformation

$$U(\theta,\varphi) = \exp(-\mathrm{i}\sigma_2\theta/2)\exp(-\mathrm{i}\sigma_3\varphi/2). \tag{2.26}$$

Note that the gauge transformation becomes singular at the point where the gauge symmetry is restored ($\vec{r} = \vec{0}$), this means that at $\vec{r} = \vec{0}$ one cannot rotate ϕ to σ_3. In fact, it is undefined there. Applying this transformation to the gauge field yields

$$U\vec{A}U^\dagger - \tfrac{\mathrm{i}}{g}U^\dagger\vec{\nabla}U. \tag{2.27}$$

One can then project onto the gauge field that corresponds to this direction in isospace (there is only a contribution from the pure gauge term):

$$\vec{A}_3 \propto \mathrm{Re}\left(-\tfrac{\mathrm{i}}{g}\mathrm{Tr}\,\phi U^\dagger\vec{\nabla}U\right) \propto \frac{\cos\theta}{gr\sin\theta}\hat{\varphi}. \tag{2.28}$$

In terms of this $U(1)$ field, the monopole indeed carries one unit of magnetic charge

$$\vec{\nabla}\times\vec{A}_3 = \frac{\vec{r}}{gr^3}. \tag{2.29}$$

The dual Meissner effect: Abelian projection

The Meissner effect explains the confinement of magnetic flux due to the condensation of electric charges. In QCD one expects confinement of electric flux to take place. Perhaps a dual Meissner effect, in which electric and magnetic charges are interchanged, is responsible for confinement in QCD? Such a mechanism is indeed proposed by 't Hooft (see [11] and references therein). The mechanism roughly works as follows (for $SU(N)$)

- The non-Abelian gauge symmetry is fixed, leaving the Cartan subgroup $U(1)^{N-1}$ unfixed.
- Point-like singularities in the gauge fixing condition show the existence of field configurations with magnetic charge. After the electric charges are integrated out, an Abelian theory with magnetic charges only is left.
- A electric-magnetic duality transformation is performed, describing the system in terms of electric charges.
- If one can show that the Meissner effect occurs in this theory, one has proven confinement of electric charge in the original gauge theory.

This mechanism will be now discussed in a bit more detail for $SU(2)$. To fix the non-Abelian gauge symmetry completely one uses an adjoint scalar field. This scalar field can be a fundamental field of the theory, but it can also be a composite field (as has to be the case for pure Yang-Mills). This adjoint scalar field, called X, is subjected to the gauge fixing condition

$$X = \mathrm{diag}(\lambda, -\lambda), \tag{2.30}$$

which obviously fixes the gauge symmetry to $U(1)$. Then the following question arises: How can one spot a monopole? Suppose there is a point x_0 in space where $\lambda = 0$, resulting in an enhancement of the gauge symmetry to $SU(2)$ at that point. For a generic adjoint operator $X = \vec{x} \cdot \vec{\sigma}$, this means that there are three constraints. The solution space to these three constraints is, generically, a co-dimension three object. Therefore in three spatial dimensions these constraints describe isolated point-like objects. In other words, the enhanced gauge symmetry is expected only at points in space (or time-like curves in four-dimensional space-time). In a small neighborhood around the singularity (after gauge fixing) the configuration looks like

$$X = \mathrm{diag}(\epsilon, -\epsilon) = \sigma_3 \epsilon. \tag{2.31}$$

After performing the inverse transformation of the one used for the monopole (equation 2.26) this configuration is turned into a hedgehog configuration. At x_0 the full gauge symmetry is restored, just as in the monopole case. This is enough information to identify the appearance of monopoles in this partially fixed gauge. Around the singularity the field configuration is that of a monopole. The theory is now effectively described by

- one massless photon
- two massive charged vector fields
- one scalar field λ

- topological point-like objects with magnetic charge: monopoles

For $SU(N)$ the argument is completely analogous. Here one has to focus on a $SU(2)$ subgroup of $SU(N)$. The solution space will, again, be a co-dimension three object.

As mentioned before, if one can show that the monopoles condense, one can argue that an electricmagnetic version of the Meissner effect (the dual Meissner effect) is responsible for the confinement of electric flux. Condensation of magnetic monopoles can be shown in $\mathcal{N} = 2$ supersymmetric gauge theories [2], providing strong evidence for confinement of electric flux in that theory.

CHAPTER 3

SUPERSYMMETRY

This chapter will give a brief review of, and introduction to, $\mathcal{N} = 2$ and $\mathcal{N} = 1$ supersymmetric gauge theories. Since the theories central in this thesis are $\mathcal{N} = 2$ theories deformed to $\mathcal{N} = 1$, $\mathcal{N} = 2$ supersymmetry will be discussed first. After discussing the work of Seiberg and Witten on $\mathcal{N} = 2$ theories, the more recent progress of Dijkgraaf and Vafa on $\mathcal{N} = 1$ theories is presented. For a more complete and thorough treatment of supersymmetry and supersymmetric gauge theories the reader is referred to the following references: [12, 13].

(3.1) GAUGE THEORIES WITH $\mathcal{N} = 2$ SUPERSYMMETRY

UNCONSTRAINED SUPERSPACE

Unconstrained $\mathcal{N} = 2$ superspace has four commuting and eight anti-commuting coordinates, grouped in a single superspace variable z_M

$$z_M = \{x^\mu, \theta_{\alpha i}, \bar{\theta}^{\dot{\alpha} i}, \quad i = 1, 2\}. \tag{3.1}$$

The $\mathcal{N} = 2$ supersymmetry algebra has an $SU(2)_R$ inner-automorphism that acts in the fundamental representation on the index i of the anti-commuting coordinates θ^i. The $\mathcal{N} = 2$ chiral superfield is obtained by imposing the following conditions on a generic $\mathcal{N} = 2$ superfield:

$$\begin{aligned} \bar{D}_{\dot{\alpha} i} \Psi(z) &= 0 \\ D^{\alpha i} D^j_\alpha \Psi &= \bar{D}^i_{\dot{\alpha}} \bar{D}^{\dot{\alpha} j} \Psi \quad i, j = 1, 2 \end{aligned} \tag{3.2}$$

leading to the following expansion in $\mathcal{N}=1$ superfields

$$\Psi(x_+,\theta) = \Phi(x_+,\theta^1) + \sqrt{2}\theta^{\alpha 2} W_\alpha(x_+,\theta^1) + \theta^2\theta^2 G(x_+,\theta^1). \tag{3.3}$$

The coordinate x_+^μ is defined as $x_+^\mu = x^\mu + \mathrm{i}\theta_i\sigma^\mu\bar{\theta}^i$, Φ is an $\mathcal{N}=1$ chiral superfield and W_α the $\mathcal{N}=1$ gauge superfield (containing the $\mathcal{N}=1$ vector superfield V). The auxiliary superfield G can be expressed in terms of Φ and W as follows

$$G(x_+,\theta^1) = -\frac{1}{2}\int \mathrm{d}^2\bar{\theta}^1 \Phi(x_+ - \mathrm{i}\theta_1\sigma\bar{\theta}^1)^\dagger \mathrm{e}^{-2V(x_+ - \mathrm{i}\theta_1\sigma\bar{\theta}^1)}. \tag{3.4}$$

The $\mathcal{N}=2$ gauge superfield is defined by replacing the superderivatives D in equation 3.2 with their covariantized forms

$$\mathcal{D}_{\alpha i} = D_{\alpha i} + \mathrm{i}A_{\alpha i}, \quad \bar{\mathcal{D}}_{\alpha i} = \bar{D}_{\alpha i} + \mathrm{i}\bar{A}_{\alpha i}. \tag{3.5}$$

The A and $\bar{A}$ fields are supergauge fields that take values in the Lie algebra of the gauge group. The gauge superfield is similar in structure to the chiral superfield, therefore the gauge superfield is often referred to as the chiral superfield.

In terms of components, the $\mathcal{N}=2$ chiral supermultiplet has a gauge field, two Weyl spinors and a complex scalar field, all in the adjoint representation of the gauge group:

$$\begin{array}{ccc} & A_\mu & \\ \lambda_+ & & \lambda_- \\ & \phi & \end{array} \tag{3.6}$$

The Weyl spinors $\lambda_\pm$ form a Dirac spinor and transform under the fundamental of the $SU(2)_R$ R-charge, therefore theories with $\mathcal{N}=2$ chiral multiplets are necessarily non-chiral. The gauge field and scalar transform trivially under the $SU(2)_R$. There is also an Abelian R-symmetry, acting only on the $\mathcal{N}=1$ chiral multiplet part of the full multiplet:

$$\lambda_- \to \mathrm{e}^{\mathrm{i}\alpha}\lambda_-, \qquad \phi \to \phi. \tag{3.7}$$

This chiral $U(1)_R$ symmetry is anomalous at the 1-loop level.

The prepotential

The actions, invariant under $\mathcal{N}=2$ supersymmetry, that will appear in this thesis use a prepotential $\mathcal{F}$, a holomorphic function of a chiral superfield. The action is then obtained by integration the prepotential over $\mathcal{N}=2$ superspace

$$S = \int \mathrm{d}^4x\mathrm{d}^4\theta \ \mathrm{Tr}\,\mathcal{F}(\Psi) + c.c. \tag{3.8}$$

This can be written in $\mathcal{N}=1$ language by integrating over half of the anti-commuting variables, $\theta^{\alpha 2}$. The result is the general expression

$$L = -\tfrac{1}{2}\int \mathrm{d}^2\theta \frac{\partial \mathcal{F}}{\partial \Phi^a} G^a - \int \mathrm{d}^2\theta \frac{\partial^2 \mathcal{F}}{\partial \Phi^a \partial \Phi^b} W^{a\alpha} W^b_\alpha. \tag{3.9}$$

After plugging in the definition (equation 3.4) for the auxiliary superfield one obtains

$$L = -\tfrac{1}{2}\int \mathrm{d}^2\theta \mathrm{d}^2\bar{\theta} \frac{\partial \mathcal{F}}{\partial \Phi^a} \mathrm{e}^{-2V} \Phi^{a\dagger} - \int \mathrm{d}^2\theta \frac{\partial^2 \mathcal{F}}{\partial \Phi^a \partial \Phi^b} W^{a\alpha} W^b_\alpha. \tag{3.10}$$

So that the Kähler potential generated by the prepotential $\mathcal{F}$ reads

$$K(\Phi, \mathrm{e}^{-2V}\Phi^\dagger) = \frac{\partial \mathcal{F}}{\partial \Phi^a} (\mathrm{e}^{-2V}\Phi^\dagger)^a \tag{3.11}$$

and the coupling constant matrix is given by

$$\tau_{ab}(\Phi) = \frac{\partial^2 \mathcal{F}}{\partial \Phi^a \partial \Phi^b}. \tag{3.12}$$

For renormalizable $\mathcal{N}=2$ theories the classical prepotential is necessarily restricted to

$$\mathcal{F}_{class} = \tfrac{1}{2}\tau_{class}\Psi^2. \tag{3.13}$$

In that case the scalar potential $V(\phi)$, obtained by integrating out the auxiliary field from the chiral superfield Φ, is equal to

$$V(\phi) = \tfrac{1}{2g^2}\mathrm{Tr}\,[\phi, \phi^\dagger]^2. \tag{3.14}$$

Because of this relatively simple form, the structure of the exact pre-potential can be determined exactly. The coupling constant g comes from $\tau = \frac{4\pi \mathrm{i}}{g^2} + \frac{\theta}{2\pi}$. Quantum corrections to the prepotential are built up out of the 1-loop contribution and non-perturbative instanton terms [14]

$$\mathcal{F} = \frac{\mathrm{i}}{2\pi}\Psi^2 \ln\left(\frac{\Psi^2}{\Lambda^2}\right) + \sum_{n=1}^{\infty} \mathcal{F}_n \left(\frac{\Lambda}{\Psi}\right)^{4k} \Psi^2. \tag{3.15}$$

The instanton contributions determine the constants $\mathcal{F}_n$. The energy scale Λ is set by the coupling constant via dimensional transmutation:

$$\Lambda^{3N}\mathrm{e}^{2\pi \mathrm{i}\tau} = \Lambda_0^{3N}\mathrm{e}^{2\pi \mathrm{i}\tau_0}. \tag{3.16}$$

(3.1.1) The Seiberg-Witten solution

The classical supersymmetric vacua are obtained by solving [2]

$$F_{\mu\nu} = 0, \quad D_\mu \phi = 0, \quad \lambda_\pm = 0, \quad V(\phi) = 0. \tag{3.17}$$

From this it is clear that ϕ is gauge equivalent to a constant. Furthermore, it has to obey

$$[\phi, \phi^\dagger] = 0. \tag{3.18}$$

Since ϕ takes values in the Lie algebra of the gauge group, this equation is satisfied if ϕ is a sum over the elements of the Lie algebra that commute with each other: the Cartan subalgebra. For $SU(N)$, the classical vacua are then given by $r = N - 1$ complex parameters a_i:

$$\phi = \sum_{i=1}^{r} a_i H_i. \tag{3.19}$$

The parameters $a_i \in \mathbb{C}$ label different vacua. However not all different combinations of the a_i give inequivalent vacua. There are gauge transformations (that necessarily connect equivalent vacua) that only permute the elements of the Cartan subalgebra H_i, and therefore effectively permute that a_i's. Those transformations are called Weyl transformations. Weyl transformations form the symmetry group of the root system that corresponds to the Lie algebra for this group. Thus two vacua a_i and a'_i are equivalent if the a'_i are a permutation of the a_i. The space of inequivalent vacua can then be written as

$$\mathcal{M} = \mathbb{C}^r/\text{Weyl}. \tag{3.20}$$

This moduli space is then parameterized by the Weyl invariant parameters u_i, implicitly defined by

$$\det(x - \phi) = x^N + u_1 x^{N-1} + \cdots + u_r. \tag{3.21}$$

The kinetic terms of the action are determined by the metric on the moduli space, which in turn is derived from the prepotential. For $\mathcal{N} = 2$ theories it can be written as

$$\mathrm{ds}^2 = \mathrm{Im}\, \frac{\partial^2 \mathcal{F}}{\partial a^i \partial a^j} \mathrm{da}^i \mathrm{d\bar{a}}^j. \tag{3.22}$$

It is convenient to rewrite this using the dual parameters $a_{D,i} = \partial \mathcal{F}/\partial a^i$:

$$\mathrm{ds}^2 = \mathrm{Im}\ \mathrm{da}^i \mathrm{d\bar{a}}_{D,i}. \tag{3.23}$$

The variables a and a_D can be seen as a section of a bundle X of complex dimension $2r$. The Kähler moduli space $\mathcal{M}$ is the base of this bundle. As the variable a

determines the mass of the gauge bosons in the spontaneously broken vacuum, the dual variable a_D determines the mass of the monopoles

$$m \sim |a_D| . \tag{3.24}$$

In this language the coupling constants are expressed as

$$\tau_{ij} = \frac{\partial^2 \mathcal{F}}{\partial a^i \partial a^j} = \frac{\partial a_{D,i}}{\partial a_j}. \tag{3.25}$$

Such an $\mathcal{N} = 2$ theory has an interesting exact $Sp(2r, \mathbb{Z})$ duality. Classically the theory is invariant under $Sp(2r, \mathbb{R})$ rotations. To see this, consider the kinetic term, which is

$$\mathcal{L}_{kin} \sim \partial_\mu \begin{pmatrix} \phi & \phi_D \end{pmatrix}^\dagger \begin{pmatrix} 0 & \mathrm{i} \\ -\mathrm{i} & 0 \end{pmatrix} \partial^\mu \begin{pmatrix} \phi \\ \phi_D \end{pmatrix}. \tag{3.26}$$

This kinetic term is manifestly invariant under symplectic transformations. For the case $r = 1$ one then has

$$\begin{pmatrix} a \\ a_D \end{pmatrix} \to M \begin{pmatrix} a \\ a_D \end{pmatrix}, \quad M \in Sp(2r, \mathbb{R}) \tag{3.27}$$

generated by

$$\begin{aligned} S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad &: \quad \begin{array}{ll} \tau & \to \frac{-1}{\tau} \\ a & \to -a_D \\ a_D & \to a \end{array} \\ T_\beta = \begin{pmatrix} 1 & \beta \\ 0 & 1 \end{pmatrix} \quad &: \quad \begin{array}{ll} \tau & \to \tau + \beta \\ a & \to a \\ a_D & \to a_D + \beta a. \end{array} \end{aligned} \tag{3.28}$$

The shift in the θ-angle is only insignificant if they are shifted by amounts of 2π, therefore β must be an integer. This restricts the duality group to be

$$Sp(2r, \mathbb{Z}) \subset Sp(2r, \mathbb{R}). \tag{3.29}$$

A geometric approach

Since the a_i parameters do not label the inequivalent vacua properly, it is better to express the metric on the moduli space and the gauge coupling constants in terms of the good parameters on the moduli space u_i. This amounts to expressing the variables a and a_D in u. The goal is then eventually to express all physical quantities, such as the coupling constant and metric on the moduli space, in terms of the good variables u of the moduli space. One can then learn from these expressions how these physical quantities depend on the choice of vacuum.

The pair a, a_D has a monodromy in the u-plane, i.e. after going through a loop in the u-plane the a and a_D get shifted by an $Sp(2r, \mathbb{Z})$ transformation. The monodromies for the theory with gauge group $SU(2)$ are derived in great detail in [2], their results will be described here briefly. One of the main results of this paper of Seiberg and Witten is that there are three characteristic monodromies in the u-plane. One monodromy at infinity (looping through an infinitely large circle in the u-plane) and two monodromies associated with the points in the moduli space where monopoles become massless, i.e. when (see equation 3.24)

$$a_D = 0 \Rightarrow u = \pm 2\Lambda^2. \tag{3.30}$$

As an example, consider the monodromy at infinity. From the general formula 3.15 one can then see that for large values of u (and of a) the expression for the prepotential reduces to that of the one-loop contribution, therefore

$$a_D = \frac{\partial \mathcal{F}}{\partial a} = \frac{2\mathrm{i}a}{\pi} \ln \frac{a}{\Lambda} + \frac{\mathrm{i}a}{\pi}. \tag{3.31}$$

After traversing a full circle at infinity in the u-plane, $\ln u$ picks up an additional $2\pi\mathrm{i}$, which means for $a(u) \sim \sqrt{2u}$

$$\ln a \to \ln a + \pi \mathrm{i}. \tag{3.32}$$

Completing a full circle in the u-plane corresponds to half a circle in the a-plane: $a \to -a$. This leads to the following monodromy

$$\begin{aligned} a_D &\to -a_D + 2a \\ a &\to -a, \end{aligned} \tag{3.33}$$

which is indeed an $Sp(2, \mathbb{Z})$ transformation. Together with the other two monodromies these transformations form a subgroup of $Sp(2, \mathbb{Z})$, known as $\Gamma(2)$.

It turns out that the knowledge of the monodromies is enough to determine the complete solution to the problem. That is, it is possible to find $a(u)$ and $a_D(u)$ by using the monodromies and the fact that the metric on the moduli space has to be positive definite

$$\mathrm{Im}\, \tau(u) > 0. \tag{3.34}$$

The solution can be written in terms of a family of hyperelliptic curves, implicitly defined by

$$y^2 = (x^2 + u)^2 - 4\Lambda^4. \tag{3.35}$$

This surface can be constructed by glueing together two complex planes (sheets) along two "cuts"(see figure 3.1). One cut is running from $u - 2\Lambda^2$ to $u + 2\Lambda^2$, the other from $-u - 2\Lambda^2$ to $-u + 2\Lambda^2$. Then the planes are compactified in the same

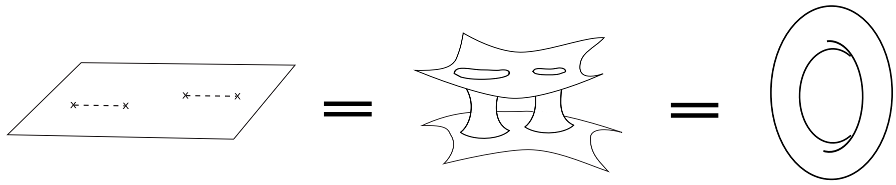


Figure 3.1: *A sheet (left) with two cuts and the corresponding two sheeted cover (middle) with the cuts blown up. The Riemann surface (a torus) is formed by compactifying the two sheets separately (right).*

way one compactifies a plane to a sphere by adding a point at infinity (one-point compactification). This curve becomes singular when $u = \pm 2\Lambda^2$, the points at which monopoles become massless. If the dynamically generated scale Λ goes to zero, one arrives at the classical limit where all cuts shrink to a point.

The modular parameter of this genus one Riemann surface (i.e. a torus) depends on u. Remarkably this modular parameter is exactly equal to the coupling constant of the $\mathcal{N} = 2$ gauge theory

$$\tau(u) = \tau_{torus}(u). \tag{3.36}$$

The modular parameter can be expressed in terms of the periods of this Riemann surface. A crucial element is the Seiberg-Witten differential defined on this curve. This differential $\mathrm{d}\omega$ is a meromorphic one-form, such that $\partial_u \mathrm{d}\omega$ is a holomorphic one-form. The dependence of a and a_D on u can then be calculated as follows

$$a(u) = \frac{1}{2\pi \mathrm{i}} \oint_A \mathrm{d}\omega = \frac{\sqrt{2}}{\pi} \int_{-1}^{1} \mathrm{d}x \frac{\sqrt{x-u}}{\sqrt{x^2 - 4\Lambda^4}} \tag{3.37}$$

$$a_D(u) = \frac{1}{2\pi \mathrm{i}} \oint_B \mathrm{d}\omega = \frac{\sqrt{2}}{\pi} \int_{1}^{u} \mathrm{d}x \frac{\sqrt{x-u}}{\sqrt{x^2 - 4\Lambda^4}}. \tag{3.38}$$

Here A and B form the canonical basis of one-cycles on the torus. From this the coupling constant can be derived

$$\tau(u) = \frac{\partial_u a_D}{\partial_u a}. \tag{3.39}$$

If a different basis of one-cycles is chosen the a and a_D are transformed by an $Sp(2, \mathbb{Z})$ element.

The power of the approach put forward by Seiberg and Witten is that the difficult questions in gauge theory (how do the coupling constants and metric on moduli space depend on the choice of vacuum) can be completely and exactly answered by studying the well-known Riemann surfaces. For the general case ($SU(r+1)$), the solution can be conveniently described in terms of a Riemann surface E of genus $g = r$. This Riemann surface can be represented as a hyperelliptic curve, or two complex planes joined along r cuts.

$$y^2 = P_N(x)^2 - 4\Lambda^{2N}, \tag{3.40}$$

with

$$P_N(x) = \det(x-\phi) = x^N + u_1 x^{N-1} + \cdots + u_{N-1}. \tag{3.41}$$

The formula used to obtain the coupling constants are easily generalized from equations 3.37 to 3.39:

$$\begin{aligned} a_i &= \tfrac{1}{2\pi \mathrm{i}} \oint_{A_i} \mathrm{d}\omega \\ a_{Di} &= \tfrac{1}{2\pi \mathrm{i}} \oint_{B_i} \mathrm{d}\omega \\ \tau_{ij} &= \partial_{u_k} a_{Di} (\partial_{u_j} a_k)^{-1}. \end{aligned} \tag{3.42}$$

(3.1.2) Confinement

In the introductory chapters the issue of confinement was put forward as one of the main reasons to study supersymmetric gauge theories. In [2], Seiberg and Witten used the results of the geometric approach to argue that the dual Meissner effect occurs in the deformed $\mathcal{N}=2$ theory. For completeness an extremely short version of their derivation will be shown here. Interested readers are referred to [2, section 5.6].

First, the $\mathcal{N}=2$ supersymmetry is broken to $\mathcal{N}=1$ by adding a superpotential

$$W = m \mathrm{Tr}\, \phi^2 = m\ u(A). \tag{3.43}$$

Then it is argued that, around the point in moduli space where monopoles become massless, the quantum effective superpotential should be modified to include the effect of the light monopoles. To describe the monopoles as fundamental particles, an electric-magnetic duality transformation is performed. All in all the exact quantum effective superpotential, from which the quantum vacua can be derived, can be written as

$$W = \sqrt{2} A_D M \tilde{M} + m u(A_D). \tag{3.44}$$

With M the chiral superfield representing the light monopoles and A_D the dual photon (magnetic charge). From this it is easy to derive the condition for a supersymmetric vacuum:

$$\mathrm{d}W = 0 \Rightarrow M = \tilde{M} = \sqrt{-\tfrac{1}{2}\sqrt{2} m u'(0)}. \tag{3.45}$$

This shows that monopoles condense and that the magnetic Higgs mechanism is in effect. This is an explicit realization of the dual Meissner effect proposed by 't Hooft. It is strong evidence that confinement in this supersymmetric gauge theory can occur through the dual Meissner effect.

(3.2) ADDING A SUPERPOTENTIAL

(3.2.1) FACTORIZATION OF SEIBERG-WITTEN CURVES

The $\mathcal{N}=2$ is broken to $\mathcal{N}=1$ if a superpotential for the $\mathcal{N}=1$ chiral multiplet Φ is added to the Lagrangian

$$\mathcal{L} \to \mathcal{L} + \int \mathrm{d}^2\theta \mathrm{Tr}\, W(\Phi). \tag{3.46}$$

In order to have a vacuum configuration, the scalar potential needs to be zero, therefore, in addition to equation 3.17 one should demand that the superpotential

$$W(\Phi) = \sum_{k=1}^{n+1} \frac{g_k}{k} \Phi^k \tag{3.47}$$

is extremized

$$W'(\phi) = g_{n+1} \prod_{k=1}^{n} (\phi - w_k \mathbb{I}) = 0. \tag{3.48}$$

Since the $\mathcal{N}=2$ vacuum conditions require ϕ to be diagonal, this basically means that the a_i can not be chosen arbitrarily anymore. Instead, they have to be picked from the set of w_k's. Depending on the degeneracy of this choice, the gauge symmetry is broken to one or more subgroups, i.e.

$$U(N) \to U(N_1) \times \cdots \times U(N_n), \text{ with } N_1 + \cdots N_n = N. \tag{3.49}$$

The numbers N_i are called filling numbers and correspond to the number of a's that are put equal to w_i.

Classically all expectation values involving the fermions only are zero. However, a one-loop effect in an instanton background gives a vacuum expectation value

$$\left\langle \lambda_-^{2N} \right\rangle \sim \mathrm{e}^{-S_{instanton}} = \Lambda_0^{3N} \mathrm{e}^{2\pi \mathrm{i}\tau} = \Lambda^{3N}, \tag{3.50}$$

breaking the $U(1)_R$ symmetry to $\mathbb{Z}_{2N}$. Special properties of the $\mathcal{N}=1$ supersymmetry algebra [15, 16] can be used to show that expectation values of chiral operators always factorize into 1-point functions of chiral operators:

$$\left\langle \lambda_-^{2N} \right\rangle = \left\langle \lambda_-^2 \right\rangle^N \Rightarrow \left\langle \lambda_-^2 \right\rangle \sim \mathrm{e}^{2\pi \mathrm{i}k/N} \Lambda^3, \quad k = 0, \ldots, N-1, \tag{3.51}$$

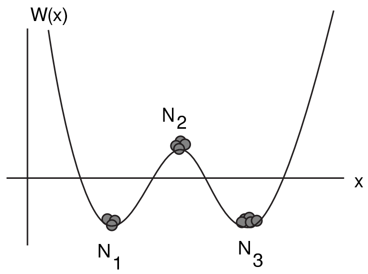


Figure 3.2: *A fourth order potential with eigenvalues distributed over the critical points. Note that the superpotential $W(\Phi)$ appears as $W'(\phi)^2$ in the scalar potential, so the eigenvalues are really localized at the minima of the scalar potential.*

with $\Lambda^3 = \Lambda_0^3 \exp(2\pi \mathrm{i}\tau_0/N)$. This means that the symmetry is broken to $\mathbb{Z}_2$ due to the vacuum expectation value of the fermion bilinear. Also the N different values this expectation value can take, means that there are N distinct quantum vacua. This result appears to be coming from a contribution from an instanton with charge $1/N$, since

$$\mathrm{e}^{\frac{2\pi\mathrm{i}\tau}{N}} = \mathrm{e}^{-\frac{8\pi^2}{gN}+\mathrm{i}\frac{\theta}{N}}. \tag{3.52}$$

Therefore the breaking of $U(1)_R$ to $\mathbb{Z}_2$ is often referred to as a fractional instanton effect. The fact that this glueball field

$$S = \lambda_-^2 = \tfrac{1}{2}\epsilon_{\alpha\beta}\lambda_-^\alpha\lambda_-^\beta \tag{3.53}$$

acquires a vacuum expectation value was used by Veneziano and Yankielowicz [17] to construct a superpotential for this chiral superfield S, known as the Veneziano-Yankielowicz term

$$W_{VY}(S) = NS\left[1 - \log\frac{S}{\Lambda^3}\right]. \tag{3.54}$$

It is easy to check that extremizing this superpotential with respect to S gives the correct vacuum expectation value for S. The Veneziano-Yankielowicz superpotential is constructed such that the variation of W_{VY}, with respect to a $U(1)_R$ rotation, reproduces the full $U(1)_R$ anomaly. This anomaly can be computed exactly by a one-loop calculation (or by using the Fujikawa path-integral method). Although the Veneziano-Yankielowicz superpotential reproduces an exact result, one cannot claim that it is an honest Wilsonian effective action. For this to be true one would have to confirm that the glueball superfield S is in fact the lightest superfield. Caution is in order since lattice results, though not conclusive (see [18]), do indicate that there are other fields with competing mass around.

Of the $U(N_i)$ factors in the gauge group, only the $SU(N_i)$ parts confine, therefore at the quantum level one expects to find a symmetry

$$U(N) \to U(N_1) \times \cdots \times U(N_n) \to U(1)^k, \tag{3.55}$$

with k the number of N_i that are non-zero.

Adding a superpotential breaks the $\mathcal{N} = 2$ supersymmetry to $\mathcal{N} = 1$, therefore the approach of Seiberg and Witten can not be followed directly. That said, the $\mathcal{N} = 1$ solution is contained in the $\mathcal{N} = 2$ solution. Just like the classical vacua for the $\mathcal{N} = 1$ theory are special cases of the $\mathcal{N} = 2$ solution (the moduli are localized by minimizing the scalar potential), it turns out that this remains to be true in the quantum case. The procedure [19] will be demonstrated here. First one looks for the submanifold in moduli space where $N - k$ monopoles (or dyons) become massless. As explained before the curve degenerates (one-cycles becomes singular) when a monopole becomes massless

$$y^2 = P_N(x)^2 - 4\Lambda^{4N} = H^2_{N-k} T_{2k}(x). \tag{3.56}$$

H_{N-k} is a polynomial of degree $N - k$, T_{2k} a polynomial of degree $2k$. The second step is to extremize the superpotential, by varying the gauge invariant parameters u_i, on this submanifold. These extrema are the quantum vacua. To illustrate this procedure a simple example $U(2)$ with $W(\Phi) = \frac{g_2}{2}\Phi^2$ will be treated. In this case the Seiberg-Witten curve reads

$$y^2 = P_2(x)^2 - 4\Lambda^4 = \det(x - \phi)^2 - 4\Lambda^4, \tag{3.57}$$

with

$$\phi = \begin{pmatrix} a_1 & 0 \\ 0 & a_2 \end{pmatrix} \Rightarrow P_2(x) = x^2 - u_1 x + u_2 = x^2 - (a_1 + a_2)x + a_1 a_2. \tag{3.58}$$

Since the superpotential is quadratic it has only one extremum, therefore classically the eigenvalues a_1 and a_2 all have to be put at the same extremum: $U(2)$ is unbroken, so in the quantum case one is left with a single $U(1)$. This means that the curve should have one double zero. The condition for a double zero can be written as

$$P'_N(x) = 0, \qquad P_N(x)^2 = 4\Lambda^4. \tag{3.59}$$

From the first condition one can read off: $x = u_1/2$, together with the second condition this finally gives

$$u_2 = \frac{u_1^2}{4} + 2\eta\Lambda^2, \qquad \eta = \pm 1. \tag{3.60}$$

This relation defines the one-dimensional submanifold in the moduli space where a monopole is massless. Therefore hyperelliptic curves satisfying equation 3.60 all have one massless monopole. After plugging this relation into the superpotential one gets

$$\mathrm{Tr}\, W(\phi) = \frac{g_2}{4} u_1^2 + 2 g_2 \Lambda^2. \tag{3.61}$$

Extremizing this with respect to u_1 one sees immediately that

$$u_1 = 0, \qquad u_2 = 2\eta\Lambda^2, \tag{3.62}$$

which are the exact quantum vacua for this gauge theory. As one can see there are two vacua, as expected on general grounds. In both vacua the superpotential takes the following value

$$W_{vac} = 2 g_2 \Lambda^2. \tag{3.63}$$

As a check, substitute these expressions for the u_i in the curve

$$y^2 = (x^2 + 2\eta\Lambda^2)^2 - 4\Lambda^4 = x^2(x^2 + 4\epsilon\Lambda^2), \tag{3.64}$$

which indeed factorizes correctly. The regular part of the curve $T_2(x) = x^2 + 4\epsilon\Lambda^2$ can be written as

$$T_2(x) = \tfrac{1}{g_2^2} W'(x)^2 + f_0(x), \tag{3.65}$$

with $f_0(x) = 4\epsilon\Lambda^2$. This is a special case of the general result

$$G_{n-k}(x)^2 T_{2k}^2(x) = \tfrac{1}{g_{n+1}^2} W'(x)^2 + f_{n-1}(x). \tag{3.66}$$

G_{n-k} and f_{n-1} are polynomials of degree $n-k$ and $n-1$ respectively.

As one can see, determining the exact quantum vacua in this case is a matter of simple algebra. In chapter 5 an even easier method for obtaining the quantum vacua will be described.

(3.2.2) Dijkgraaf-Vafa: a string theory inspired approach

Although the Seiberg-Witten theory is very powerful and useful, it does not give any information about the glueball expectation values. In [3] Dijkgraaf and Vafa explain how to use a chain of string dualities to calculate the exact effective superpotential for the chiral glueball fields S. The chain of dualities involves D-branes wrapping Calabi-Yau manifolds, open and closed topological strings and a matrix model. Fascinating as these dualities are, strictly one does not need them (although without a string theory model of $\mathcal{N} = 1$ gauge theories it would have been much harder to find their results). The main result of their paper can be derived using field theory techniques only. It can be summarized as follows: To calculate the effective superpotential in terms of the glueball superfield, it is sufficient to sum over

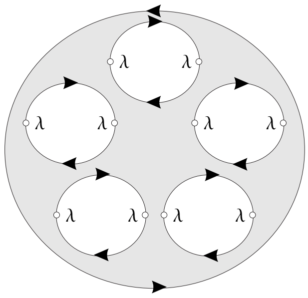


Figure 3.3: *A Feynman diagram with k loops that contributes to the S^k term of the effective superpotential. Every loop has one S insertion (made up out of two λ insertions).*

planar diagrams generated by W treated as a classical Lagrangian, involving only zero-momentum scalar fields running around in the loops, with gauge fields as external legs. The remarkable thing is that the summation over Feynman diagrams, i.e. a perturbative calculation, yields non-perturbative results for the glueball expectation values. These diagrams are conveniently described by a Matrix Model, that is a zero-dimensional field theory involving an adjoint scalar field. The appearance of this Matrix Model can be seen using the chain of dualities, however a purely field theoretic proof can also be given [4].

In formula the main result is stated as

$$W_{eff}(S_1,\ldots,S_k) = \sum_i \left(2\pi\mathrm{i}\tau_0 S_i + N_i \frac{\partial \mathcal{F}_0}{\partial S_i}\right). \tag{3.67}$$

Here the S_i are the glueball chiral superfields, one for every $U(1)$ factor that is left from each $U(N_i)$ factor. The perturbative sum is over n colored *planar* Feynman diagrams given by the classical action W_{tree} and can be written as

$$\mathcal{F}_0(S) = \sum_i -\tfrac{1}{2}S_i^2 \log\left(\frac{S_i}{\Lambda_0^3}\right) + \sum_k a_k S^k. \tag{3.68}$$

The coefficients a_k can be determined by summing all planar Feynman diagrams with k-loops evaluated at zero momentum. Every loop has two λ insertions, each contributing a factor of S to the final result. If the k holes are colored $i_1,\ldots,i_k \in \{0,1,\ldots,n-1\}$ then the graph gets multiplied by

$$S^k = S_{i_1}\cdots S_{i_k}. \tag{3.69}$$

The Feynman rules are derived from the superpotential $\mathrm{Tr}\, W(\Phi)$, now considered as the bosonic action of a zero-dimensional field theory, a matrix theory. Since only planar diagrams are included, the coefficients $a_k \sim c^k$ and that makes the (perturbative part) of $\mathcal{F}_0$ analytic in the S_i.

Note that extremizing W_{eff} enables one to express the superpotential completely in terms of the dynamically generated scale Λ and the coupling constants g_i. Also the values of the S_i in the minimum of the superpotential give the quantum corrected correlators of the fermion bilinears $S_i \sim \langle \mathrm{Tr}_i \lambda^2 \rangle$. The correlators take the form of fractional instanton expansions

$$\mathrm{d}W_{\mathrm{eff}} = 0 \Longrightarrow S_i = \sum b_k \Lambda^{3k}. \tag{3.70}$$

Hence non-perturbative results are obtained from perturbative calculations. For $k = 1$ the familiar fractional instanton result for the pure $\mathcal{N} = 1$ supersymmetric gauge theory is reproduced: $S_i \sim \Lambda^3$. Furthermore, if one calculates $\mathcal{F}$ up to k loops in graphs,the effects of up to k fractional instantons are taken into account.

Since the quantum theory has an Abelian $U(1)^k$ symmetry, there is also an effective gauge coupling matrix τ_{ij} given by

$$\tau_{ij} = \frac{\partial^2 \mathcal{F}_0}{\partial S_i \partial S_j}. \tag{3.71}$$

The whole procedure is best explained by working out an example. Consider the (renormalizable) superpotential

$$W(\phi) = \frac{m}{2}\phi^2 + \frac{g}{3}\phi^3. \tag{3.72}$$

This superpotential has two critical points, $\phi = 0$ or $\phi = -g/m$. In the vacuum where $\phi = 0$ the $SU(N)$ part of the gauge group $U(N)$ confines

$$U(N) \rightarrow^{S} \rightarrow U(1). \tag{3.73}$$

The Feynman rules are easily obtained (remember that for a zero-dimensional field theory, all momentum dependence can be dropped):

$$\text{propagator: } \frac{1}{m} \tag{3.74}$$

$$\text{cubic vertex: } \frac{g}{3}. \tag{3.75}$$

Since ϕ is an adjoint field, the Feynman diagrams are so-called fat or ribbon diagrams. For such diagrams it is often convenient to use the 't Hooft double line notation. Also, to the Feynman rules one should add the rule that every boundary loop gives a factor of S. The result up to two-loops is

$$\begin{aligned} \mathcal{F}_0 = & -\tfrac{1}{2} S^2 \log(S/\Lambda^3) && \text{[VY]} \\ & + \tfrac{1}{2} S^2 \log(m/\Lambda) && \text{[one-loop]} \\ & + \tfrac{2}{3} \frac{g^2}{m^3} S^3 && \text{[two-loop]} \\ & + \mathcal{O}(S^4). \end{aligned} \tag{3.76}$$

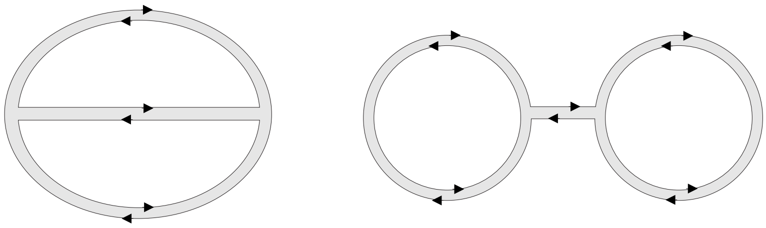


Figure 3.4: *The two planar two-loop diagrams, with combinatorial weight $\frac{1}{6}$ and $\frac{1}{2}$, that contribute to the order S^3 term in the free energy $\mathcal{F}_0(S)$.*

The first term in $\mathcal{F}$ gives rise to the VY term, in the matrix model it comes from the volume factor in the path-integral. The second term comes from the one-loop diagrams, of which there is only one. The symmetry factor of the diagram is $\frac{1}{2}$, the log comes from the usual momentum integration of a one-loop diagram (with cut-off Λ), but now in the zero-momentum limit. The third term is more interesting, it corresponds to the two-loop diagrams. For the case at hand, there are only two two-loop diagrams (see figure 3.4). The first diagram diagram can be created in three different ways (remember that the lines in the fat diagrams have an orientation) and the symmetry factor is $\frac{1}{2}$, further the diagram has two vertices and three propagators, so this diagram contributes to $\mathcal{F}$

$$(\frac{g}{3})^2 \frac{1}{m^3} \tfrac{1}{2} 3 S^3 = \frac{g^2}{6m^3} S^3, \tag{3.77}$$

the second diagram can be created in nine different ways, the symmetry factor is $\frac{1}{2}$ again, giving

$$(\frac{g}{3})^2 \frac{1}{m^3} \tfrac{1}{2} 9 S^3 = \frac{g^2}{2m^3} S^3. \tag{3.78}$$

These two contributions indeed add up to the third term. As mentioned before, the free energy $\mathcal{F}$ can be effectively obtained using an auxiliary matrix model. In this matrix model one takes ϕ to be an $\tilde{N} \times \tilde{N}$ matrix, think of it as an element of the Lie algebra of $U(\tilde{N})$. Note that $\tilde{N}$ is not to be confused with N, in fact $\tilde{N}$ is related to gaugino condensate $S \sim \tilde{N}$. The generating functional that will compute the quantum effective superpotential is a matrix integral

$$Z = \int \mathrm{d}\phi \, e^{\frac{1}{g_s} \mathrm{Tr}\, W(\phi)} = e^{\sum_{g=0} g_s^{2g-2} \mathcal{F}_g(S)}. \tag{3.79}$$

Here g_s is just an auxiliary parameter needed for picking out the planar diagrams (i.e. the genus zero diagrams, which are the diagrams contributing to $\mathcal{F}_0$). It is not to be confused with the gauge coupling (in fact it has dimension -3). The eigenvalues of ϕ can again be distributed over the critical points of W', giving a partition

$$\tilde{N} = \tilde{N}_1 + \ldots + \tilde{N}_n. \tag{3.80}$$

Now define the 't Hooft couplings for the matrix model as follows

$$S_i = g_s \tilde{N}_i. \tag{3.81}$$

Since one takes the large $\tilde{N}$ limit, function $\mathcal{F}_0$ will be analytic in the S_i. Thus it is possible to complexify the expression for $\mathcal{F}_0$. So, if one can compute the matrix integral in the large $\tilde{N}$ limit, effectively one computed the function $\mathcal{F}_0$ which, in turn, directly determines the exact quantum effective superpotential for the S_i. The appearance of this matrix model might seem a bit ad hoc. But remember the following. From the chain of dualities one learns that the superpotential is computed by zero-momentum planar diagram with λ-insertions. The matrix model can be seen as an effective *tool* to compute the amplitudes of those diagrams. To enforce that only planar diagrams contribute to the path-integral one can take the large $\tilde{N}$-limit in the matrix model. To avoid any confusion: $\tilde{N}$ is not N, that is, one gets finite rank results for the gauge theory using a large $\tilde{N}$ matrix model.

(3.2.3) Confinement

In the famous article by Veneziano and Yankielowicz [17] the effective superpotential, in terms of the glueball field S, for pure $\mathcal{N} = 1$ Yang-Mills theory was constructed. Using the fact that the quantum effective superpotential should reproduce the $U(1)_R$ anomaly, or rather a supermultiplet of anomalies, they arrived at

$$W_{VY} = hS \left[1 - \log \frac{S}{\Lambda^3}\right]. \tag{3.82}$$

Here h is the dual Coxeter number (for $SU(N), h = N$). The effective superpotential correctly reproduces the exact quantum vacua where the $U(1)_R$ is broken

$$S = \mathrm{e}^{2\pi i k/h} \Lambda^3, \qquad h \in \mathbb{Z}_h. \tag{3.83}$$

Now if one would expand the superpotential around this action, one would learn that the glueball field S picks up a quadratic term: the field S becomes massive. If the Veneziano-Yanckielowicz result would be a genuine quantum Wilsonian action, which can only be true if S is the lightest field around, then one can argue that there is a mass-gap. A necessary condition and strong argument for confinement.

However it is not all obvious whether or not the glueball field is the lightest field around. This is something that the Dijkgraaf-Vafa approach does not really address. What it does address is what happens to the expectation value and mass of the glueball field if a classical superpotential is added to the problem. In that case one can use this approach to calculate the exact superpotential, a superpotential that reproduces the exact expectation values for the glueball field. In this sense the Dijkgraaf-Vafa approach extends the range of applicability of the argument of Veneziano-Yankielowicz.

Chapter 4

Integrable systems

Some key concepts of the theory of integrable system will be introduced. The integrable systems that appear in later chapters of this thesis are introduced one by one.

(4.1) Symplectic geometry

A symplectic manifold M is a manifold that can be endowed with a non-degenerate closed two-form ω (a symplectic form). The non-degeneracy of the symplectic form means that ω^n defines a good volume form. From this it is clear that the dimension of M is always even, $\mathrm{dim} M = 2n$.

Tangent bundles are special cases of symplectic manifolds. The phase space of a classical mechanical system is a tangent bundle, therefore the phase space of a classical mechanical system is a symplectic manifold. In a classical mechanical system a Hamiltonian induces flow (i.e. the classical motion of the system) in the phase space. In the more general setting of symplectic geometry this corresponds to the notion of Hamiltonian vector fields. A scalar function h, the Hamiltonian, defined on the manifold defines a vector field, V_h, on the manifold:

$$\imath_{V_h}\omega = \mathrm{d}h. \tag{4.1}$$

Since the symplectic form ω is non-degenerate this relation can be inverted: given a Hamiltonian the corresponding flow can be calculated. The Hamiltonian dynamics of classical mechanical systems is often described using Poisson brackets. Since the symplectic form governs the flow in phase space, it comes as no surprise that the symplectic form determines the Poisson brackets too

$$\{f, g\} = \omega(V_f, V_g), \tag{4.2}$$

where f and g are two function on the manifold, V_f and V_g are their associated Hamiltonian vector fields.

Sometimes a classical mechanical system possesses a symmetry that effectively reduces the dimensionality of the problem. For example, a test particle moving in a spherically symmetric three-dimensional gravitational field is effectively described by a one-dimensional equation. This equation involves the radial coordinate only. In more technical terms such a reduction is called a symplectic quotient or Hamiltonian reduction. In the language of symplectic geometry it is described as follows. Suppose there is a flow ξ on the manifold that preserves the symplectic form $\mathcal{L}_\xi \omega = \mathrm{d}\imath_\xi \omega = 0$. This means that ξ is a Hamiltonian vector field: $\imath_\xi \omega = \mathrm{d}\mu$ for some function μ, the moment map, invariant under the flow of ξ. The phase space of the reduced system is then given by the following quotient $\mu^{-1}(0)/S_\xi$. Here S_ξ represents the action of the symmetry on μ. This quotient is again a symplectic manifold. To illustrate this procedure, consider a two-dimensional symplectic space $\mathbb{R}^2$ with symplectic form $\omega = \mathrm{d}q \wedge \mathrm{d}p$. This is invariant under the circular flow of $\xi = q\partial_p - p\partial_q$. This flow just rotates the coordinates p and q by an $SO(2)$ rotation. The moment map is given by

$$\imath_\xi \omega = q\mathrm{d}q + p\mathrm{d}p = \tfrac{1}{2}\mathrm{d}(p^2 + q^2 - R^2) \Rightarrow \mu(p, q) = \tfrac{1}{2}(p^2 + q^2 - R^2) \tag{4.3}$$

for some constant R. The space $\mu^{-1}(0)$ is then given by $p^2 + q^2 = R^2$, simply a circle in phase space. Identifying this circle under the $SO(2)$ action gives the symplectic quotient. In this case: a point.

(4.2) INTEGRABILITY

An n-dimensional classical mechanical system (the phase space being $2n$-dimensional) is called integrable if one can find n independent conserved quantities

$$I_i(q, p) \quad i = 1, \ldots, n. \tag{4.4}$$

If the Hamiltonian for this system is called H then this condition can be written as

$$\dot{I}_i = \{I_i, H\} = 0. \tag{4.5}$$

If all the conserved quantities mutually commute

$$\{I_i, I_j\} = 0, \tag{4.6}$$

the system is called completely integrable.

For integrable systems it is possible to describe the motion of the system in terms of action-angle variables. The action variables are the conserved quantities I_i, the motion of the angle variables θ_i is then completely linear

$$\ddot{\theta}_i = 0 \Rightarrow \theta_i(t) = \theta_i(0) + \tau_i I_i t. \tag{4.7}$$

The τ_i and $\theta_i(0)$ are integration constants. The action-angle variables are related to the canonically conjugate variables (p, q) via a canonical transformation.

There is a class of systems that allows one to efficiently write down the equations of motion using just two $n \times n$ matrices $L(p, q)$ and $M(p, q)$

$$\dot{L} = [L, M]. \tag{4.8}$$

The matrices L and M are called a Lax pair, after their creator [20]. Although it is generically hard to find a Lax pair, once it is found it is trivial to find the n conserved quantities. In fact they are given by

$$I_k = \mathrm{Tr}\, L^k, \qquad k = 1, \ldots, n. \tag{4.9}$$

It is easy to see that traces of powers of the Lax matrix L are conserved

$$\begin{aligned} \dot{I}_k &= \frac{\mathrm{d}}{\mathrm{d}t} \mathrm{Tr}\, L^k = k \mathrm{Tr}\,(\dot{L} L^{k-1}) \\ &= k \mathrm{Tr}\,([L, M] L^{k-1}) = k \mathrm{Tr}\,(L M L^{k-1} - M L^k) = 0, \end{aligned} \tag{4.10}$$

here the cyclic property of the trace and the equations of motion were used.

Constructing a Lax pair for a system is therefore a sufficient condition for proving the integrability of the system. Note that not all integrable systems admit a Lax pair formulation.

(4.3) THE PERIODIC TODA CHAIN

A very relevant integrable system for the description of the quantum vacua of four-dimensional $\mathcal{N} = 1$ gauge theories is the periodic Toda chain. This integrable system is a classical mechanical system of n points on a circle. The dynamics of the Toda chain is described by the following Hamiltonian

$$H = \tfrac{1}{2} \sum_{i=1}^{n} p_i^2 + \Lambda^2 \sum_{i=1}^{n} \mathrm{e}^{q_{i+1} - q_i}, \quad q_{n+1} = q_1. \tag{4.11}$$

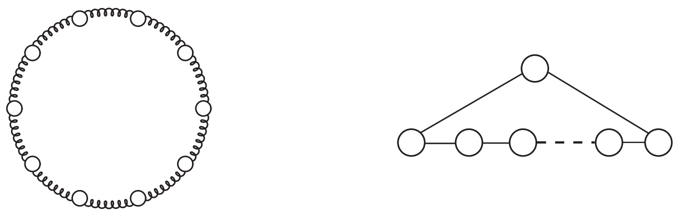


Figure 4.1: *The periodic Toda system describes beads connected by springs, providing the exponential interaction. The Dynkin diagram of the affine Lie algebra ($\hat{A}_7^{(1)}$) associated to this system is drawn on the right.*

The q_i are the coordinates (angles) of the points along the chain, p_i the conjugate momenta. The equations of motion are given by

$$\ddot{\rho}_i = -\Lambda^2(2\mathrm{e}^{\rho_i} - \mathrm{e}^{\rho_{i+1}} - \mathrm{e}^{\rho_{i-1}}), \text{ with } \rho_i = q_{i+1} - q_i. \tag{4.12}$$

After taking a closer look at these equations it becomes clear that they can be written more elegantly in terms of the Cartan matrix, A_{ij}, of the affine Lie algebra $\hat{A}_{n+1}^{(1)}$

$$\ddot{\rho}_i = -\Lambda^2 A_{ij}\mathrm{e}^{\rho_j}. \tag{4.13}$$

Readers unfamiliar with affine Lie algebras are referred to appendix A for a crash course. The special properties of the Cartan matrix are responsible for the integrability of the Toda chain. Also it is no coincidence that the Dynkin diagram of $\hat{A}_{n+1}^{(1)}$ looks exactly like the Toda chain (see figure 4.1). This system is now easy to generalize, just replace the Cartan matrix with that of an other affine Lie algebra. It turns out that you actually get away with that, the resulting system is always integrable.

The use of the affine Lie algebra $\hat{A}_{n+1}^{(1)}$ goes further than the appearance of the Cartan matrix. In fact the Toda chain admits a Lax formulation, and the Lax matrix L is given by the Chevalley generators of $\hat{A}_{n+1}^{(1)}$. In full generality (i.e. for arbitrary affine Lie algebra) the Lax matrix reads

$$L = \sum_{i=1}^{r} p_i H_i + \sum_{\alpha\in\Delta_s} (y_i E_i^+ + E_i^-) + \frac{y_0}{z} E_0^- + z E_0^+, \tag{4.14}$$

with z a complex parameter, r the rank of the algebra and Δ_s the system of simple roots α_i. The y_i are defined as

$$y_i = \Lambda^2 \mathrm{e}^{\alpha_i \cdot q}. \tag{4.15}$$

The inner product should be read as follows. The simple roots for any affine Lie algebra can be embedded in a r or $r+1$-dimensional vector space. For $\hat{A}_r^{(1)}$ one can write

$$\begin{aligned} \alpha_i &= e_i - e_{i+1}, \qquad i = 1, \ldots, r \\ \alpha_0 &= -(\alpha_1 + \alpha_2 + \cdots + \alpha_r) = -e_1 + e_{r+1}. \end{aligned} \tag{4.16}$$

With the e_i a basis of $\mathbb{R}^r$. Then the inner product becomes

$$\alpha_i \cdot q = q_i - q_{i+1}. \tag{4.17}$$

Note that the variables y are not all independent, in fact they obey the constraint

$$\Lambda^{2h} = \prod_{i=0}^{r} y_i^{a_i}, \tag{4.18}$$

here h is the dual Coxeter number of the affine Lie algebra and the a_i are is the mark of the i^{th} root in the Dynkin diagram. This constraint basically follows from the definition of the affine root

$$\alpha_0 = -\sum_{i=1}^{r} a_i \alpha_i. \tag{4.19}$$

For $\hat{A}_r^{(1)}$ one has $h = r+1$ and $a_i = 1$. Using appendix A it is now possible to write down an explicit form of the Lax matrix for $\hat{A}_r^{(1)}$

$$L = \begin{pmatrix} p_1 & y_1 & 0 & \ldots & z \\ 1 & p_2 & y_2 & \ddots & 0 \\ 0 & 1 & p_3 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & y_r \\ \frac{y_0}{z} & \ldots & 0 & 1 & p_{r+1} \end{pmatrix}. \tag{4.20}$$

(4.3.1) Relation to Seiberg-Witten theory

The Lax matrix defined in equation 4.14 contains the complex parameter z, this parameter is usually called a spectral parameter. Lax matrices with a spectral parameter have an associated hyperelliptic curve, defined as follows (in the case the Lax matrix is an $N \times N$ matrix)

$$0 = \det(x - L) = P_N(x)^2 + (-1)^N \left(z + \frac{\Lambda^{2N}}{z} \right). \tag{4.21}$$

With $P_N(x)$ a polynomial of degree N in x. After defining $y = 2z + (-1)^N P_N(x)$ this can be cast into the standard form for hyperelliptic curves

$$y^2 = P_N(x)^2 - 4\Lambda^{2N}. \tag{4.22}$$

gauge group	$SU(r+1)$	$SO(2r)$	$SO(2r+1)$	$Sp(2r)$	G_2
affine dual algebra	$\hat{A}_r^{(1)}$	$\hat{D}_r^{(1)}$	$\hat{A}_{2l-1}^{(2)}$	$\hat{D}_{r+1}^{(2)}$	$\hat{D}_4^{(3)}$

Table 4.1: *The mapping between gauge group and affine algebra.*

This curve is of precisely the same form as the Seiberg-Witten curve in equation 3.40. In [21] Martinec and Warner indeed argued that the Seiberg-Witten curve for gauge group G is given by the Toda system based on the dual affine Lie algebra of this gauge group. For $G = SU(r+1)$ this is simply $\hat{A}_r^{(1)}$, however for the other groups this is not so simple. In table 4.1 the affine algebras that reproduce the Seiberg-Witten curves are listed. The reason why one has to take the *dual* affine Lie algebra instead of the normal affine Lie algebra is rather easy to explain. Suppose one would take the affine extension of the Lie group belonging to the gauge symmetry. The spectral curve obtained that way would have a term Λ^{2h}, with h the Coxeter number. However the correct Seiberg-Witten curves have a term $\Lambda^{2h^\vee}$ instead. The dual Coxeter and Coxeter numbers are interchanged when dualizing the affine Lie algebra (see appendix A). Therefore one is forced to take the dual affine Lie algebra. The fact that the spectral curve of the Toda system is equivalent to the Seiberg-Witten curve, already shows the relevance of the Toda system for Seiberg-Witten theory. However the question *why* the Toda system appears is a different story (see section 5.3).

(4.4) Hitchin systems

There is a systematic way to associate an integrable system in combination with a gauge group to a Riemann surface. This construction, due to Hitchin, is based on the moduli space of principal bundles over a Riemann surface.

Take a Riemann surface Σ of genus g and consider the space $\mathcal{A}$ of $Lie(G)$-valued $(0,1)$ forms $A = A_{\bar{z}}(z,\bar{z})\mathrm{d}\bar{z}$. There is also a $(1,0)$ form Φ defined as $\Phi = \Phi_z(z,\bar{z})\mathrm{d}z$. The Hitchin construction starts with the co-tangent bundle $T^*\mathcal{A}$. Being a co-tangent bundle this space is naturally a symplectic manifold. The integrable system arises after constructing a symplectic quotient (or Hamiltonian construction). For this a Hamiltonian vector field is needed, i.e. a vector field that leaves the symplectic form invariant $\mathcal{L}_\xi\omega = 0$. The moment map (see section 4.1) is the Hamiltonian function μ that corresponds to this vector field

$$\mathrm{d}\mu = \imath_\xi\omega. \tag{4.23}$$

To be more specific, the moment map is given by:

$$\mu = \bar{\mathcal{D}}_A \Phi = \bar{\partial}\Phi + A \wedge \Phi + \Phi \wedge A. \tag{4.24}$$

The symplectic quotient is then constructed by identifying the zero locus of μ (i.e. $\mu^{-1}(0)$) under the stability group of the equation $\mu = 0$. In the case of a Hitchin system the stability group is the complete gauge group

$$\mathcal{P} = \frac{\mu^{-1}(0)}{\mathcal{G}}. \tag{4.25}$$

The quotient $\mathcal{P}$ is again a symplectic space. In fact it is $T^*\mathcal{N}$, with $\mathcal{N} = \mathcal{A}/\mathcal{G}$ the orbit space of $\mathcal{A}$. For genus 0 and 1 systems the Hitchin construction by itself does not produce interesting dynamical systems. However if the construction is based on a marked Riemann surface (at the point z) one should solve for

$$\mu = \bar{\mathcal{D}}_A \Phi = 2\pi\mathrm{i}\, \delta^2(z) u. \tag{4.26}$$

Hamiltonians on the cotangent bundle $\mathcal{P}$ can be constructed using elements $e_i \in H^1(\Sigma, T^{(1,0)}\Sigma)$ (see [22])

$$h_{k,i} = \int_\Sigma e_i \wedge \mathrm{Tr}\, \Phi_z^k. \tag{4.27}$$

This integrable system describes the moduli space of $\mathcal{N} = 2$ supersymmetric Yang-Mills with a massive adjoint hypermultiplet [22]. The field Φ corresponds to the complex adjoint scalar from the hypermultiplet. In fact this integrable system is equivalent to a spin-Calogero-Moser system, from which one may derive the Toda system in an appropriate limit to describe the pure $\mathcal{N} = 2$ theory.

(4.4.1) Elliptic spin Calogero-Moser from Hitchin

As a simple, but relevant, example of a Hitchin system, a torus T_τ (modulus τ) with one marked point is considered (the following derivation is an expanded version of [23, §7.12]):

$$z \sim z + 1, \quad z \sim z + \tau \qquad z \in \mathbb{C}. \tag{4.28}$$

At the marked point, $z = 0$, an arbitrary element u from the Lie algebra of G is chosen. The integrable system is then given by solutions of the equation

$$\bar{\mathcal{D}}_A \Phi = 2\pi\mathrm{i}\, \delta^2(z) u. \tag{4.29}$$

To solve this equation one needs to find the space $\mathcal{N}$, the orbit space of $\mathcal{A}$, the space of gauge fields A. This comes down to finding all principal G-bundles over the torus. Such a bundle is defined by the transition functions between the various patches. The torus can be covered with two such patches:

$$U_0 = D_\epsilon(0), \quad U_\infty = T_\tau - \{0\}, \tag{4.30}$$

U_0 is a small disk around the marked point $z = 0$ and U_∞ is the whole torus except the marked point. For simplicity the field strength is taken to be zero, i.e. in every patch the following relation holds

$$A = h^{-1}\bar{\partial}h, \tag{4.31}$$

for some $Lie(G)$-valued function h. To solve equation 4.29 in each patch separately one can use the fact that $\bar{\mathcal{D}}_A$ is a covariant derivative. Since A is pure gauge, one can write

$$\bar{\mathcal{D}}_A\Phi = h^{-1}\left(\bar{\partial}_{A=0}\hat{\Phi}\right)h. \tag{4.32}$$

With $\hat{\Phi}$ obeying

$$\begin{array}{lcl} U_0 & : & \partial_{\bar{z}}\hat{\Phi}_0 = 2\pi \mathrm{i}\delta^2(z)\hat{u}, \quad \text{with } \hat{u} = h_0 u h_0^{-1} \\ U_\infty & : & \partial_{\bar{z}}\hat{\Phi}_\infty = 0. \end{array} \tag{4.33}$$

Using a contour integral around $z = 0$ in the U_0 patch, it becomes clear that $\hat{\Phi}_0$ has a simple pole at $z = 0$ with residue $\hat{u}$. The solutions in the two patches are related by the transition function

$$g_{0\infty} = h_\infty h_0^{-1}. \tag{4.34}$$

This transition function is only defined in the overlap between the two patches, that is the annulus $U_0 - \{z = 0\}$. For a torus the transition function takes the special form $g_{0\infty} = \exp(q/z)$, with q an element in the Cartan subalgebra

$$q = \sum_i q_i H_i. \tag{4.35}$$

So in the U_∞ patch the solution reads

$$\hat{\Phi}_\infty = \mathrm{e}^{q/z}\left(\frac{\hat{u}}{z} + \mathcal{O}(1)\right)\mathrm{e}^{-q/z}. \tag{4.36}$$

It is convenient to decompose $\hat{\Phi}_\infty$ and $\hat{u}$:

$$\begin{aligned} \hat{u} \quad &= \textstyle\sum_i \hat{u}_i + \sum_\alpha \hat{u}_\alpha E_\alpha \\ \hat{\Phi}_\infty &= \textstyle\sum_i \hat{\Phi}_i H_i + \sum_\alpha \hat{\Phi}_\alpha E_\alpha \mathrm{d}z. \end{aligned} \tag{4.37}$$

Comparing this to equation 4.36 one sees that

$$\hat{\Phi}_i \sim \frac{\hat{u}_i}{z}, \quad \hat{\Phi}_\alpha \sim \exp(\alpha(q)/z)\hat{u}_\alpha/z. \tag{4.38}$$

Here $\alpha(q)$ should be read as the N-dimensional root vector α acting on the collection of position of the nodes on the chain q_i. For example, for $SU(N)$

$$\alpha_i = e_i - e_{i+1} \Rightarrow \alpha(q) = q_i - q_{i+1}, \tag{4.39}$$

where the e_i form an orthonormal basis of $\mathbb{R}^N$. From the first condition, $\hat{\Phi}_i \sim \frac{\hat{u}_i}{z}$, one infers that

$$\hat{u}_i = 0 \text{ and } \hat{\Phi}_i = p_i = \text{constant}, \tag{4.40}$$

since an elliptic function with a single pole of order one is a constant function. The second constraint in equation 4.38 has a unique solution for $\hat{\Phi}$, given in terms of the Lamé function[1]:

$$\hat{\Phi}_\alpha(q;z) = -\hat{u}_\alpha \frac{\sigma(z-\alpha(q))}{\sigma(z)\sigma(\alpha(q))} \mathrm{e}^{\alpha(q)\zeta(z)}. \tag{4.41}$$

Here σ and ζ are the Weierstrass sigma and zeta functions, related to the Weierstrass p-function $\wp$ as follows:

$$\begin{aligned}
\wp(x;\omega_1,\omega_2) &= \tfrac{1}{x^2} + \textstyle\sum_{(m,n)\neq(0,0)} \left(\tfrac{1}{(x+2m\omega_1+2n\omega_2)^2} - \tfrac{1}{(2m\omega_1+2n\omega_2)^2}\right) \\
\wp(x) &= \zeta'(x) \\
\zeta(x) &= \tfrac{\sigma'(x)}{\sigma(x)} \\
\left(\tfrac{\mathrm{d}^2}{\mathrm{d}z^2} - \wp(x)\right)\Phi(x;z) &= 2\wp(z)\Phi(x;z).
\end{aligned} \tag{4.42}$$

All in all the expectation value of the adjoint scalar from the hypermultiplet is found to be

$$\hat{\Phi} = \left(\sum_i p_i H_i + \sum_\alpha \hat{\Phi}_\alpha\right) \mathrm{d}z, \tag{4.43}$$

which is exactly the Lax matrix of the elliptic spin Calogero-Moser system with Hamiltonian:

$$H = \sum_i p_i^2 - \sum_\alpha \wp(\alpha(q))(\hat{u}_\alpha \hat{u}_{-\alpha}). \tag{4.44}$$

This is just like the ordinary elliptic Calogero-Moser system except here there is a different coupling strength $(\hat{u}_\alpha \hat{u}_{-\alpha})$ for each node of the periodic chain.

(4.4.2) Periodic Toda from elliptic spin Calogero-Moser

It is well known that the Toda system can be obtained as a special limit of the elliptic Calogero-Moser system. The Hamiltonian of the ordinary elliptic Calogero-Moser

[1] Lamé functions are the elliptic equivalents of the spherical harmonics.

system reads:

$$H = \tfrac{1}{2}p^2 - \tfrac{1}{2}m^2 \sum_{\alpha} \wp(\alpha(x); \omega_1, \omega_2). \tag{4.45}$$

The periods ω_1 and ω_2 of the torus are now explicitly written. The periodic Toda system is recovered by sending the real part of ω_2 to infinity while keeping ω_1 fixed:

$$\omega_1 = -\mathrm{i}\pi, \qquad \mathrm{Re}\,(\omega_2) \to \infty. \tag{4.46}$$

The coordinates q_i and the coupling parameter m are changed

$$q_j = Q_j + 2\omega_2 j/N, \qquad m = M\mathrm{e}^{\omega_2/N}, \tag{4.47}$$

while Q_j and M are kept fixed. After taking the limit one is left with the Hamiltonian for the Toda system. For the gauge group $SU(N)$ this is

$$H = \tfrac{1}{2}p^2 - \tfrac{1}{2}M^2 \sum_i \mathrm{e}^{\alpha_i \cdot Q} = \tfrac{1}{2}p^2 - \tfrac{1}{2}M^2 \left(\sum_{i=1}^{N-1} \mathrm{e}^{Q_{i+1}-Q_i} \right) - \tfrac{1}{2}M^2 \mathrm{e}^{Q_N - Q_1}. \tag{4.48}$$

Note that for the Toda system one only sums over the simple roots of the Lie algebra, for the Calogero-Moser system one needed to sum over all positive roots.

Taking the limit of the elliptic spin Calogero-Moser system is essentially the same, however, in that case there are multiple coupling parameters

$$m_\alpha = \hat{u}_\alpha \hat{u}_{-\alpha}. \tag{4.49}$$

The $\hat{u}$'s have to be chosen such that the m_α's are non-zero if α is a simple root (since those are the only terms that give a contribution to the Toda Hamiltonian). In the limit all m_α go to $M_\alpha \mathrm{exp}(\omega_2/N)$, leading to a Toda system with different couplings for each interaction between two nodes. The difference in the couplings can be canceled by shifting the coordinates Q appropriately.

CHAPTER 5

THREE-DIMENSIONAL SUPERSYMMETRIC GAUGE THEORIES

After having introduced supersymmetric gauge theories, Seiberg-Witten theory, the Dijkgraaf-Vafa approach and integrable systems, it is time to put everything together. Starting from section 5.2 till the end of the chapter, original work published in [24, 25, 26] will be presented. The work entails the calculation of the non-perturbative quantum superpotential using a classical mechanical system. This calculation involves a conjecture that will be tested in several ways.

(5.1) COMPACTIFYING ON A CIRCLE

Before embarking upon the study of three-dimensional gauge theories, it is in order to explain why one should be interested in *three-dimensional* theories in the first place. The upshot is that certain quantities calculated in three-dimensional gauge theories can be used to determine four-dimensional quantities. The central argument is due to Seiberg and Witten [27, section 3].

The starting point is the familiar four-dimensional $\mathcal{N}=2$ theory compactified on $\mathbb{R}^{1+2}\times S^1$. For convenience the discussion in this section will be limited to $SU(2)$. If the radius R of the circle is much greater than the distance scale set by Λ, the theory is effectively four-dimensional and the results obtained in chapter 3 apply. In that case the effective theory has a $U(1)$ symmetry; the moduli space is described by a complex scalar u. The effective action is then given by the familiar kinetic and topological terms one can write down for a $U(1)$ gauge field

$$S_{eff} = \int \mathrm{d}^4x \frac{1}{4g(u)^2} F_{\mu\nu}F^{\mu\nu} + \frac{\mathrm{i}\theta(u)}{64\pi^2}\epsilon_{\mu\nu\rho\kappa}F^{\mu\nu}F^{\rho\kappa}. \tag{5.1}$$

The coupling constant and theta-angle can be easily read-off once $\tau(u)$ is determined, using the Seiberg-Witten method introduced in chapter 3.

This action still contains the full four-dimensional fields. In order to write down the effective three-dimensional action one rewrites the gauge field in terms of two scalars b and σ. First, the fourth component of the gauge field is evidently a scalar from a three-dimensional perspective

$$A_4 = \frac{b}{\pi R}. \tag{5.2}$$

It so happens that the remaining three-dimensional gauge field is dual to a scalar

$$\partial_\alpha \sigma = \epsilon_{\alpha\beta\gamma}F^{\beta\gamma}, \quad \alpha, \beta, \gamma = 0, 1, 2. \tag{5.3}$$

The effective action in three-dimensions can then be written as [27]

$$L_{eff,3d} = \int \mathrm{d}^3x \frac{1}{\pi R g^2} \left|\partial b\right|^2 + \frac{g^2}{64\pi^3 R} \left|\partial\sigma - \tfrac{\theta}{\pi}\partial b\right|^2. \tag{5.4}$$

From this effective Lagrangian one can read-off that this three-dimensional gauge theory has two additional real moduli: b and σ. Since the two extra moduli are periodic, the complete moduli space can be seen as a torus fibered over the u-plane. For a more general gauge group the moduli space is a hyper-Kähler manifold of dimension $4k$, with k the number of left-over $U(1)$ factors. The general structure of this moduli space is that of a k-torus fibered over the u_i-plane. Although the moduli space, being hyper-Kähler, has three different complex structures, the one inherited from the complex u_i-plane plays a special role. In [27] it is argued that this complex structure is independent of the radius R.

(5.2) The conjecture

It turns out that this R independent complex structure, inherited from the u-plane, is exactly the complex structure that appears in the breaking of $\mathcal{N} = 2$ to $\mathcal{N} = 1$ by adding a superpotential. This means that in order to calculate the quantum vacuum structure in four-dimensions one can also calculate the three-dimensional quantum vacuum structure. The question now is: how does one calculate the quantum effective superpotential in three dimensions. First of all, there are no fractional instanton contributions, only regular ones (coming from the monopoles in the four-dimensional theory). Further in four-dimensions one has the difficulty of monopoles becoming massless, making it necessary to alter the quantum effective superpotential to include their effects. In three dimensions there are no monopoles to become massless, hence the classical superpotential is not corrected. The only thing left to do is to identify the variables in which to express the superpotential, such that when extremizing the superpotential with respect to these variables, one obtains the correct quantum vacua. The conjecture, put forward in [24], says that the “right”variables are given by the Lax matrix of the Toda system one can associate to this gauge symmetry (see table 4.1). The conjecture is now easily stated

$$\boxed{W_{eff} = W_{tree}(L).} \tag{5.5}$$

That is, substitute the Lax matrix for the adjoint scalar in the superpotential. Since the Lax matrix contains a spectral parameter z, the final result could contain terms depending on z. These terms are simply dropped. Then after minimizing with respect to the variables, p_i and y_i, of the Toda system one finds the exact quantum vacua.

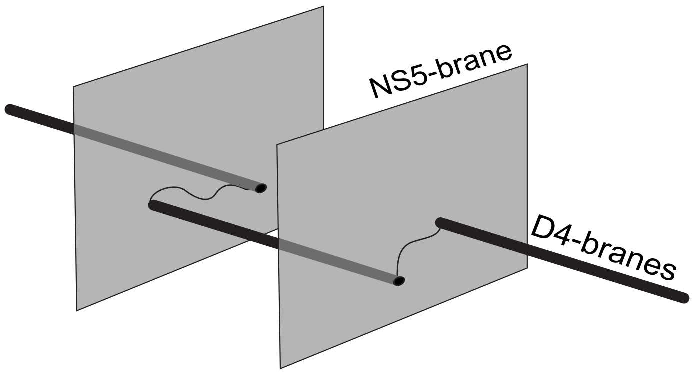
The idea of using an integrable system to calculate quantities in a supersymmetric gauge theory is not new in itself, see for example: [28, 29, 30, 31].

(5.3) Towards a proof of the conjecture

The way the Lax matrix of the Toda system pops-up out of the blue in this conjecture, is quite unsatisfactory. In this section a D-brane configuration, describing some supersymmetric gauge theory, will be used to show how the Toda system comes into play. The line of reasoning is based on a paper by Kapustin [32]. In this paper a D-brane setup of NS5-branes and stacks of D4-branes is used to show that the vacua of this system can be described by a Hitchin system, an integrable system. This is where the paper of Kapustin stops, however one can push his results a bit further. The D-brane system describes a gauge theory with adjoint hyper multiplets. To get rid off those one can take a limit in which these hypermultiplets become infinitely massive. It is the goal of this section to show that in this limit the Hitchin system turns into the Toda system and that the vacuum expectation value of the adjoint scalar Φ is equal to the Lax matrix of the Toda system, thereby giving an explanation for the appearance of the Lax matrix. Note that no superpotential is added here. The consequences of adding a superpotential will be commented on at the end of this section.

(5.3.1) THE D-BRANE SETUP

Take k NS5 branes extended in the $0, 1, 2, 3, 4, 5$ directions with stacks of D4-branes, suspended between those NS5-branes, extending in the $0, 1, 2, 3, 6$ directions (see figure 5.1). One can create such a configuration by taking N D4-branes and adding k NS5-branes. At the intersection points between the D4-branes and an NS5-brane, the D4-branes can split and move over the NS5-brane in the $4, 5$ directions. The gauge symmetry of this system is $SU(N) \times \cdots \times SU(N)$. The number of D4-branes suspended between the i-th and $i + 1$-th NS5-brane is the same everywhere. Since the stacks of D4-branes can move freely on the NS5-branes along the $4, 5$ directions it is convenient to define $v = x^4 + \mathrm{i}x^5$. The coordinate x^6 is taken to be periodic with period $2\pi L$. The bosonic fields living on the D4-branes are a gauge field and five



	0	1	2	3	4	5	6	7	8	9
NS5	–	–	–	–	–	–	·	·	·	·
D4	–	–	–	–	·	·	–	·	·	·

Figure 5.1: *Stacks of D4-branes suspended between NS5-branes.*

real adjoint scalars, corresponding to the fluctuations of the brane in the $x^{4,5,7,8,9}$ directions. Since the D4-branes are of finite size in the x^6 direction, the world-volume theory is basically $1 + 3$-dimensional. The five-dimensional theory then effectively becomes the four-dimensional $\mathcal{N} = 4$ super Yang-Mills theory.

The geometry of the brane-setup breaks supersymmetry to $\mathcal{N} = 2$. The $\mathcal{N} = 4$ vector multiplet decomposes into an $\mathcal{N} = 2$ vector multiplet and an $\mathcal{N} = 2$ hyper multiplet. However, since the D4-branes end on the NS5-branes, the boundaries of the D4-branes are not free to fluctuate in all directions. In this case the Dirichlet boundary conditions, disallowing fluctuations in the $x^{6,7,8,9}$ directions at the intersection between the D4 and NS5-branes, set the hyper multiplet to zero. The vector multiplet remains untouched [33].

In the case there are two stacks of D4-branes on one single NS5-brane, extending in opposite directions, there can be strings stretching between those stacks. Those strings would, in a way, restore the degrees of freedom in the hypermultiplet. Each end-point of the strings would transform in the fundamental representation belonging to the stack it ends on. The mass of the hypermultiplet is proportional to the distance between the stacks along the NS-brane. The massive hypermultiplet, like massive quarks, has a vacuum expectation value of zero. This basically means that the stacks are fixed in the $x^{6,7,8,9}$ direction. This is easily understood from the string theory perspective: a semi-infinite brane has to end on another brane (here the semi-infinite D4-branes end on the NS5-brane). The hypermultiplet can cause fluctuations in the $x^{6,7,8,9}$ directions. In the limit where the mass of the hypermultiplet goes to zero, the two semi-infinite D4-branes join up and a full D4-brane can move off the NS5-brane. This is in agreement with the fact that massless quarks can have a vacuum expectation value.

To summarize, the D-brane system describes a $d = 4, \mathcal{N} = 2$ gauge theory with gauge symmetry $SU(N) \times \ldots \times SU(N)$. The strings stretching between two adjacent D4 stacks give k bi-fundamental hypermultiplets Q_α. The Q_α hypermultiplet transforms in the $(N, \bar{N})$ representation. The complex adjoint scalar in the $\mathcal{N} = 2$ vector multiplet corresponds to the movement of the D4-branes on the NS5-branes (i.e. the $x^{4,5}$ directions). The two complex adjoint scalars in the hypermultiplet correspond to movement in the $x^{6,7,8,9}$ directions. For D4-branes suspended between NS5-branes, only the vector multiplet survives as the D4-branes cannot move off the NS5-branes in the $x^{7,8,9}$ directions. The bi-fundamental hyper multiplet arises because strings stretching between stacks of D4-branes ending on a single NS5-branes cause fluctuations in the $x^{6,7,8,9}$ directions at the intersection of the D4 and NS5-brane.

The next step is to compactify the x^3 direction. This allows the use of a kind of mirror symmetry that relates the Coulomb branch of this three-dimensional theory

	0	1	2	3	6	4	5	7	8	9
D4'	–	–	–	·	·	–	–	·	·	·
D4	–	–	–	–	–	·	·	·	·	·

Table 5.1: *The D4'-D4 brane setup (note the unusual order of the labels).*

(the "electric" theory) to the Higgs branch of another theory (the "magnetic" theory) [34]. The Higgs branch doesn't receive any quantum corrections, so calculations will be much easier there. The Higgs branch and Coulomb branch are interchanged under an S-duality. In this case the winning combination of dualities is a TST duality:

a) First apply a T-duality on x^3, the NS5-branes become IIB NS5-branes, the D4-branes become D3-branes.

b) Then apply the S-duality, turning the NS5-branes into D5-branes and leaving the D3-branes invariant.

c) Finally apply a T-duality on x^3 again. The D5-branes end up as D4'-branes extending in the $x^{0,1,2,4,5}$ directions. The D3-branes eventually are turned into the original D4-branes again (i.e. they are left invariant by the duality sequence).

So one is left with D4-branes wrapped around a torus T^2 (x^3, x^6), the world-volume theory of these D4-branes is essentially three-dimensional. The D4'-branes manifest themselves as impurities in this three-dimensional theory, localized at points on the torus. As said, this three-dimensional theory in the Higgs branch does not receive quantum corrections, hence it can be treated completely classically. One might wonder how this classical theory is able to encode the non-trivial coupling constant dependence of the original theory in the Coulomb branch. The answer is that results in the magnetic theory will depend non-trivially on the modulus of the torus, which in turn is related to the coupling constant of the electric theory due to the several duality transformations that are performed:

$$\tau = \frac{L}{\lambda}, \tag{5.6}$$

where L is the period of the x^6 directions and λ is the coupling constant in the electric theory.

According to [32] the supersymmetric vacua are given by the following equations on the torus

$$\begin{aligned} &F_{z\bar{z}} - [\Phi_z, \Phi_z^\dagger] = 0 \\ &\bar{D}_A \Phi_z = -\tfrac{\pi}{RL} \textstyle\sum_{\alpha=1}^{k} \delta^2(z - z_\alpha)\, \mathrm{diag}(m_\alpha, -\mu, \ldots, -\mu). \end{aligned} \tag{5.7}$$

This precisely defines a Hitchin system on a marked torus (see figure 5.2). The field A is a gauge field on the torus (more precisely it is a $U(N)$ connection). This gauge field comes from the D4-branes wrapping the torus. The adjoint scalar field Φ parameterizes the location of the D4-branes along the D4'-branes in the $x^{4,5}$ directions. These moduli correspond to the moduli of the adjoint scalar field in the vector multiplet of the original "electric" theory. Hence the scalar field appearing in 5.7 can be identified with the adjoint scalar from the vector multiplet. Since there are multiple stacks of D4-branes one would expect multiple adjoint scalar fields Φ, one for every stack. This is not the case since equation 5.7 should be interpreted as an impurity theory in which there is one stack of N D4-branes where the D4'-branes are treated as impurities (represented by the delta-functions). The parameters m_α correspond

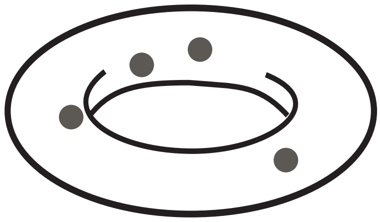


Figure 5.2: *The punctured torus of the Hitchin system, the black dots represent the location of the D4'-branes.*

to the masses of the hypermultiplets located at the impurities (i.e. the D4'-branes). The parameter μ corresponds to a deformation of the system. On the electric side it is caused by a non-trivial background geometry [33], whereas on the magnetic side it corresponds to a Fayet-Iliopoulos parameter. This deformation is necessary to allow for a massive hypermultiplet in the case of a single NS5-brane. To see this suppose for a minute that $\mu = 0$. Since the x^6 direction is compact, the D4-branes (intersecting and breaking at the NS5-branes) should eventually come back to itself, suggesting that the following should hold

$$\sum_\alpha m_\alpha = 0. \tag{5.8}$$

This means that in the case of just one NS5-brane, the hypermultiplet is massless. However after turning on the parameter μ the background geometry is twisted in such a way that the D4-brane does not have to come back on itself, it fails a distance μ. This in effect allows for a non-zero hypermultiplet mass, even in the case of just one NS5-brane, since the condition now reads

$$\sum_\alpha m_\alpha = (N-1)\mu. \tag{5.9}$$

The brane setup required to reproduce the pure $\mathcal{N} = 2$ theory is then seen to consist of one NS5-brane with one stack of D4-branes, with a non-zero parameter μ. Ultimately this is described by a Hitchin system satisfying

$$\bar{D}_A \Phi_z = -\frac{\pi}{RL}\delta^2(z)\, m\, \mathrm{diag}(1, -\tfrac{1}{N-1}, \ldots, -\tfrac{1}{N-1}). \tag{5.10}$$

In the limit of $m \to \infty$ the hypermultiplets decouple and the Hitchin system should, if the conjecture is correct, approach a Toda system.

(5.3.2) How the brane setup leads to the Toda system

To show that the Toda system in fact describes the pure $\mathcal{N} = 2$ theory, two things need to be proven:

1. One needs to show that the Hitchin system derived from the brane setup corresponds to the Hitchin system from which the spin Calogero-Moser system is derived (see section 4.4).
2. The limit in which the $\mathcal{N} = 2$ theory with bi-fundamental hyper-multiplets becomes a pure $\mathcal{N} = 2$ theory should be the limit in which the spin Calogero-Moser system becomes the Toda system.

The Hitchin system that will lead to the desired spin Calogero-Moser system is discussed in section 4.4.1. The Hitchin system that describes the D-brane system is characterized by a zero fieldstrength corresponding to the gauge field A and certain restrictions on the "spin" variables u. In order to be able to derive the Toda system from the Hitchin system one needs to make sure none of the couplings in the spin Calogero-Moser system, that connects the two systems, become zero. Specifically this means that:

$$\hat{u}_i = 0 \text{ and } \hat{u}_\alpha \hat{u}_{-\alpha} \neq 0 \quad \forall_{\alpha \in \Delta_s}, \tag{5.11}$$

where Δ_s are the simple roots. The Hitchin system under consideration with just one marked variable (see equation 5.7) becomes

$$\bar{D}_A \Phi_z = -\frac{\pi}{RL}\, \delta^2(z)\, \mathrm{diag}\left(m, -\tfrac{m}{N-1}, \ldots, -\tfrac{m}{N-1}\right) = 2\pi \mathrm{i} u \delta^2(z). \tag{5.12}$$

The right-hand side corresponds to the variable u, which after some (yet undetermined) gauge transformation can be brought into the form

$$\hat{u} = huh^{-1} = \sum_i \hat{u}_i + \sum_\alpha \hat{u}_\alpha E_\alpha. \tag{5.13}$$

If $\hat{u}$ still satisfies equation 5.11 then the Hitchin system derived from the D-brane setup is of the correct form. To show that it is always possible to find such a gauge transformation h, one can make use of the fact that there is a convenient matrix representation of the E_α. For $SU(4)$ for example the matrix E_α for simple roots is given by

$$\sum_{\alpha \in \Delta_s} u_\alpha E_\alpha + u_{-\alpha} E_{-\alpha} = \begin{pmatrix} 0 & u_{\alpha_1} & 0 & 0 \\ u_{-\alpha_1} & 0 & u_{\alpha_2} & 0 \\ 0 & u_{-\alpha_2} & 0 & u_{\alpha_3} \\ 0 & 0 & u_{-\alpha_3} & 0 \end{pmatrix}. \tag{5.14}$$

The other non-diagonal elements of the matrix u correspond to the higher roots. Given this it is sufficient to prove that there exists a matrix $\hat{u}$, with zeroes on the diagonal and non-zero elements on the upper and lower diagonal, that has the following eigenvalues

$$N-1, \qquad -1 \text{ (N-1 times)}. \tag{5.15}$$

This would mean that $\hat{u}$ can be diagonalized to u. This is equivalent to proving that there exist a gauge transformation that takes u into $\hat{u}$. As an educated guess one can try the following matrix

$$\hat{u} = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}. \tag{5.16}$$

By calculating the characteristic equation of this matrix one can show that the eigenvalues are indeed of the correct form. To illustrate this explicitly, the calculation will be performed for the 4×4 matrix, generalization to arbitrary N is straightforward. The characteristic equation is determined by

$$c_4(x) = \begin{vmatrix} -x & 1 & 1 & 1 \\ 1 & -x & 1 & 1 \\ 1 & 1 & -x & 1 \\ 1 & 1 & 1 & -x \end{vmatrix}. \tag{5.17}$$

Using the fact that the determinant does not change if one row is added to another, one can write this as

$$c_4(x) = \begin{vmatrix} -x-1 & 0 & 0 & x+1 \\ 0 & -x-1 & 0 & x+1 \\ 0 & 0 & -x-1 & x+1 \\ 1 & 1 & 1 & -x \end{vmatrix} = (x+1)^3 \begin{vmatrix} -1 & 0 & 0 & 1 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & -1 & 1 \\ 1 & 1 & 1 & -x \end{vmatrix} \tag{5.18}$$

by subtracting the last row from all the other rows. The matrix can be brought to a triangular form by adding the first three rows to the fourth row

$$c_4(x) = (x+1)^3 \begin{vmatrix} -1 & 0 & 0 & 1 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -x+3 \end{vmatrix}, \tag{5.19}$$

showing the the characteristic equation is given by

$$c_4(x) = (x+1)^3(x-3). \tag{5.20}$$

From this it is clear that the eigenvalues are indeed 3 and -1 (multiplicity 3). As said, the calculation is easily extended to arbitrary N, confirming that it is possible to gauge rotate u in such a way that $\hat{u}$ has the properties listed in equation 5.13.

Equation 5.12 is solved as explained in section 4.4. The solution will be a Lamé function, expressed in the parameters u, q and z (see equation 4.41). Here the role of u is replaced by a single parameter m, which controls the mass of the bi-fundamental hypermultiplets as well as the strength of the deformation.

To get rid off the effect of the bi-fundamental hypermultiplets, their masses are tuned to infinity, which in this case also means tuning the deformation parameter m to infinity. How exactly those masses are tuned to infinity shouldn't really matter, therefore one can perform the limit as is necessary to recover the Toda system. Using this freedom, the q's are shifted according to equation 4.47. In this limit the Toda system will be recovered, thereby proving that the solution of the "electric" gauge theory is obtained from the Lax matrix of the Toda system:

$$\boxed{\text{vector multiplet: } \Phi_{\text{quantum}} = \text{Lax matrix of corresponding Toda system}} \tag{5.21}$$

Until now the way the entries of the Lax matrix are related to the moduli of the gauge theory was not discussed at all. In principle there are $2N$ complex variables in the Lax matrix: p_i and Q_i. These correspond to the parameters of the moduli space of the three-dimensional gauge theory (which is indeed $4N$-dimensional). So the end result is that the moduli space of the three-dimensional theory is described by the Lax matrix of the Toda system. Combining this result with the fact that the three-dimensional superpotential only needs to be written down in the "correct" variables to describe the exact four-dimensional superpotential, one has a good argument in favor of the conjecture.

A superpotential can be added to this setup by inserting a second NS5-brane and bending it in the $w = x^7 + \mathrm{i}x^8$ direction according to [19]

$$w = W'(v). \tag{5.22}$$

To keep the configuration supersymmetric the D4-branes have to intersect the NS5-branes. Since the D4-branes are located at $w = 0$, and the D4-branes cannot be placed at an angle without breaking supersymmetry, this implies that

$$W'(v) = W'(\Phi) = 0. \tag{5.23}$$

The location of the D4-branes along the straight NS5-brane is not arbitrary anymore, instead they need to be located at the extrema of W. Combining this with the result that the vacuum value of Φ (in the case where no classical superpotential is present) is given by the Lax matrix of the Toda system, one has reason to believe that the condition $W'(L) = 0$ gives the supersymmetric quantum vacua. To proof the conjecture one would have to check whether or not the Hitchin system of equation 5.7 still appears after adding a superpotential.

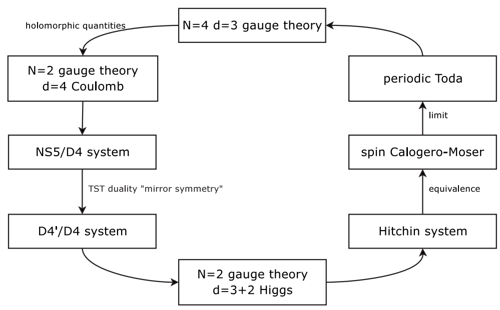


Figure 5.3: *A schematic overview of the relation between the periodic Toda system and the $\mathcal{N} = 2$ gauge theory.*

(5.4) Performing the calculation in three dimensions

There are several ways to check the conjecture. The first method is to compare the results with results obtained from Seiberg-Witten theory. The Seiberg-Witten curve is given by the Lax matrix (see section 4.3.1)

$$y^2 = (\det(x\mathbb{I} - L))^2 - 4\Lambda^{4N}. \tag{5.24}$$

Then, according to the conjecture, one has to extremize the superpotential with respect to the variables of the Lax matrix. The solutions that are a result of this should then be substituted into the Seiberg-Witten curve. If the curve factorizes as in equations 3.56 and 3.66 one can be certain that one has found the correct quantum vacua. In [25] a general proof for the correct factorization, in the case of an $U(N)$ gauge group, is given. Note that the y_i variables in the Lax matrix are not all independent, instead they obey the constraint (for $U(N)$, see 4.18 for the general result)

$$\Lambda^{2N} = \prod_{i=0}^{N-1} y_0 \cdots y_N. \tag{5.25}$$

It is technically easier to treat the y_i variables as unconstrained and use a Lagrange multiplier ℓ to enforce the constraint

$$W_{eff} = W_{tree}(L) + \ell \log \left(\frac{\Lambda^{2N}}{\prod_{i=0}^{N-1} y_0 \cdots y_N} \right). \tag{5.26}$$

This superpotential is then supposed to be extremized with respect to ℓ, p_i and all of the y_i.

The second method to check the conjecture is to try to reproduce the results of the Dijkgraaf-Vafa method. One could compare the value of the superpotential in the vacua (thus expressing the superpotential in terms of the energy-scale Λ and the coupling constants only). However it is more interesting to reproduce the superpotential expressed in terms of the glueball fields S. The problem here is to figure out how the glueball field S enters the superpotential. In the maximally confining case, i.e. when $U(N)$ breaks to a single $U(1)$ factor, there is only one S field. In this case one can use the "integrating in" procedure [35]: find the value of the superpotential in the vacuum

$$W_{min}(g_i, \Lambda) \tag{5.27}$$

and replace it by

$$W_{min}(g_i, \Omega) + S \log \left(\frac{\Lambda^{2N}}{\Omega^{2N}} \right). \tag{5.28}$$

The variable S can be integrated in by integrating out Ω.

It does not matter if the integrating in procedure is performed before or after minimizing the superpotential, therefore one can also add the $S \log \left(\frac{\Lambda^{2N}}{\Omega^{2N}} \right)$ term to equation 5.26. From equation 5.26 it is clear that Ω appears only in the log-terms. Having realized this one can combine the two $\log$-terms

$$S \log \left(\frac{\Lambda^{2N}}{\Omega^{2N}} \right) + \ell \log \left(\frac{\Omega^{2N}}{\Pi_i y_i} \right) \tag{5.29}$$

and integrate out Ω explicitly to obtain

$$W_{tree}(L) + S \log \left(\frac{\Lambda^{2N}}{\Pi_i y_i} \right), \tag{5.30}$$

after which the variables p, y and S can be integrated out.

Comparing this with equation 5.26 reveals that the Lagrange multiplier can be identified with S. This remarkable result will be used in some of the examples to reproduce the Dijkgraaf-Vafa results. If there is more than one glueball field, the Lagrange multiplier can be identified with the sum of the glueball fields

$$\ell = \sum_i S_i. \tag{5.31}$$

The rest of this chapter is devoted to giving some examples that support this conjecture. More examples can be found in [24, 25, 26, 36].

(5.4.1) Example: $U(N)$ with quadratic superpotential

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

$$W = g_1 \mathrm{Tr}\,(\phi) + \frac{g_2}{2}\mathrm{Tr}\,(\phi^2) + \ell \log\left(\frac{\Lambda^{2N}}{y_0 \cdots y_{N-1}}\right) \tag{5.32}$$

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

$$W_{\text{eff}} = \sum_{i=0}^{N-1}(g_2 y_i + g_1 p_{i+1} + \tfrac{1}{2}g_2 p_{i+1}^2) + \ell \log\left(\frac{\Lambda^{2N}}{y_0 \cdots y_{N-1}}\right) \tag{5.33}$$

with the equations for the extrema

$$y_0 \cdots y_{N-1} = \Lambda^{2N} \tag{5.34}$$

$$g_2 - \frac{\ell}{y_i} = 0 \Rightarrow y_i = \frac{\ell}{g_2} \tag{5.35}$$

$$g_2 p_i + g_1 = 0. \tag{5.36}$$

From the second equation one can learn that the y's are all equal, making it possible to solve for ℓ using the first equation

$$\ell = \epsilon g_2 \Lambda^2, \qquad \epsilon^N = 1. \tag{5.37}$$

Substituting ℓ into the equation for y_i yields

$$y_i = y = \Lambda^2 \epsilon. \tag{5.38}$$

As expected there are N different solutions

$$\epsilon = \mathrm{e}^{2\pi \mathrm{i} k/N}, \quad k = 0, \ldots, N-1. \tag{5.39}$$

Further, from equation 5.36 follows that all the p's occupy the root of the tree-level superpotential $x = -g_1/g_2$ of $W'(x)$

$$W'(x) = g_2(x + \frac{g_1}{g_2}). \tag{5.40}$$

The p'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

$$W_{\text{extr}} = N(y - \frac{g_1^2}{2g_2}) = N(\Lambda^2 \mathrm{e}^{\frac{2\pi \mathrm{i} k}{N}} - \frac{g_1^2}{2g_2}). \tag{5.41}$$

To compare with field theory results, the characteristic polynomial $P_N(x) = \det(x1_N - \Phi)$ needs to be computed. According to appendix B this is either a Chebyshev polynomial of the first or 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

$$P_N(x) = 2y^{N/2} T_N\left(\frac{x+\frac{g_1}{g_2}}{2\sqrt{y}}\right). \tag{5.42}$$

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

$$\begin{aligned} P_N(x)^2 - 4\Lambda^{2N} &= 4\Lambda^{2N}\left(T_N^2\left(\frac{x+\frac{g_1}{g_2}}{2\sqrt{y}}\right) - 1\right) \\ &= 4\Lambda^{2N}\left(\left(\frac{x+\frac{g_1}{g_2}}{2\sqrt{y}}\right)^2 - 1\right) U_{N-1}^2\left(\frac{x+\frac{g_1}{g_2}}{2\sqrt{y}}\right) \\ &= \Lambda^{2(N-1)}\epsilon^{N-1}\left(\frac{1}{g_2^2}(W')^2 - 4\Lambda^2\epsilon\right) U_{N-1}^2\left(\frac{x+\frac{g_1}{g_2}}{2\sqrt{y}}\right). \end{aligned} \tag{5.43}$$

Here T_N and U_N are the Chebyshev polynomials of the first and second kind respectively. The form of this curve is in correspondence with the general expectation equations 3.56 and 3.66. The appearance of the Chebyshev polynomials was first demonstrated, using completely different techniques, in [37].

(5.4.2) Example: Integrating in the glueball fields

$U(2)$ with cubic superpotential

In this example the conjecture will be tested by reproducing the Dijkgraaf-Vafa superpotential for a cubic tree-level superpotential

$$W(x) = \frac{x^3}{3} - a^2 x. \tag{5.44}$$

Since the gauge group is $U(2)$ the Lax matrix reads

$$L = \begin{pmatrix} p_1 & y_1 + z \\ 1 + y_0/z & p_2 \end{pmatrix}. \tag{5.45}$$

This superpotential has extrema at $x = \pm a$. Since the integrating in procedure can be used to integrate in a single glueball field S, the discussion is necessarily limited to the maximally confining case in which the gauge group is broken a single $U(1)$. As discussed before the chiral field S has to be equal to the Lagrange multiplier ℓ in that case:

$$\ell = S. \tag{5.46}$$

As explained above, the S-dependence can be recovered from the superpotential

$$W_{eff} = W(p_1) + W(p_2) + (y_0 + y_1)(p_1 + p_2) + S \log\left(\frac{\Lambda^4}{y_0 y_1}\right) \tag{5.47}$$

by integrating out the p's and the y's degrees of freedom. This means solving for the extrema of the superpotential by varying it with respect to the p's and y's. Note that, since S should not be integrated out, the equation of motion for S, $y_0 y_1 = \Lambda^4$, should not be used. In general the maximally confining vacua are obtained by using the following ansatz

$$y_j = y \quad \forall_j, \qquad p_i = p \quad \forall_i. \tag{5.48}$$

This ansatz is proven in [24, section 4.4], but in this case one can check it explicitly. Using this ansatz one has to solve only two equations

$$\frac{\partial W}{\partial p} = 0 \text{ and } \frac{\partial W}{\partial y} = 0. \tag{5.49}$$

Under the simplification equation 5.48 the superpotential becomes

$$W_{eff} = 2W(p) + 4yp + S \log\left(\frac{\Lambda^4}{y^2}\right). \tag{5.50}$$

The equation for p yields

$$p = \sqrt{a^2 - 2y}. \tag{5.51}$$

Plugging this into the superpotential yields

$$W_{eff} = -\frac{4}{3}(a^2 - 2y)^{\frac{3}{2}} + S \log\left(\frac{\Lambda^4}{y^2}\right). \tag{5.52}$$

The next step is to integrate out the y's

$$\frac{\partial W}{\partial y} = 0 \Rightarrow S^2 = 4y^2(a^2 - 2y). \tag{5.53}$$

To solve equation 5.53 it is useful to rewrite it slightly in such a way that only the combination $S/2a$ appears. To this end one defines

$$\xi = \frac{S}{4a^3} \text{ and } y(S) = A(\xi)\frac{S}{2a}. \tag{5.54}$$

Then 5.53 can be written as

$$A^2(\xi) - 4\xi A^3(\xi) = 1. \tag{5.55}$$

In principle there are three solutions for $A(\xi)$, however not all solutions have the right classical limit ($\Lambda \to 0$). To identify the unphysical solutions a scaling argument is used to determine the low energy form of the solution (i.e. Λ close to zero but not equal to zero). The glueball field scales as Λ^3, the variable y as Λ^2 and the coupling constant a as Λ. From equation 5.53 it is then clear that at low energy one has

$$S^2 = 4y^2a^2 \Rightarrow y = \pm\frac{S}{2a} \text{ (low energy)}. \tag{5.56}$$

As expected there are two choices since on general grounds one expects two vacua (see the discussion surrounding equation 3.51). For the rest of the discussion the solution $y = \frac{S}{2a}$ is selected. Thus:

$$y = \frac{S}{2a} + \mathcal{O}(S^2), \text{ hence } A(0) = 1. \tag{5.57}$$

The solution for $A(\xi)$ (with $A(0) = 1$) is (see for example [38])

$$A(\xi) = \sum_{n=0}^{\infty} \frac{2^{2n}}{n+1}\binom{\frac{3n-1}{2}}{n}\xi^n = \frac{1}{12\xi} + \frac{1}{6\xi}\sin\left(\tfrac{1}{3}\text{arcsin}\left(216\xi^2 - 1\right)\right) \tag{5.58}$$

yielding the following expression for y

$$\begin{aligned} y(S) &= \frac{S}{2a}\sum_{n=0}^{\infty}\frac{2^{2n}}{n+1}\binom{\frac{3n-1}{2}}{n}\left(\frac{S}{4a^3}\right)^n \\ &= -\frac{a^2}{6}\left(-1 + 2\sin\left(\tfrac{1}{3}\text{arcsin}\left(1 - \frac{27S^2}{2a^6}\right)\right)\right) \\ &= \frac{S}{2a} + \frac{S^2}{4a^4} + \frac{5S^3}{16a^7} + \mathcal{O}(S^4). \end{aligned} \tag{5.59}$$

Substituting this solution into the effective superpotential yields, in principle, a closed expression valid to all orders in S. However the form of this expression is rather cumbersome, therefore the superpotential is expanded in S

$$W_{\text{eff}} = -\frac{4a^3}{3} + 2S\left(1 - \log\left(\frac{S}{\Lambda^2 m}\right)\right) - \left(\frac{S^2}{2a^3} + \frac{S^3}{3a^6} + \frac{35S^4}{96a^9} + \dots\right), \tag{5.60}$$

here $m = 2a$ is the mass of the fluctuations of ϕ around the classical extremum $\phi = a$. It is interesting to note that the Veneziano-Yankielowicz term (the log term) appears automatically in this effective action. The reason that this term (which is sometimes difficult to reproduce in Matrix Models) appears is because of the integrating in procedure, which effectively assumes that the Veneziano-Yankielowicz term should be present [39]. This means that the Veneziano-Yankielowicz term is introduced via a backdoor: the integrating in procedure.

To test the conjecture this result should be compared with the four-dimensional answer [40, eq. 3.1]. According to the Dijkgraaf-Vafa method the superpotential has the following general from (for gauge group $U(N)$)

$$W_{\text{eff}} = NS\left(1 - \log\left(\frac{S}{\Lambda^3}\right)\right) - NS\log\left(\frac{\Lambda}{m}\right) - N\frac{\partial \mathcal{F}}{\partial S}. \tag{5.61}$$

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

$$\mathcal{F} = \frac{2}{3}\frac{g^2}{m^3}S^3 + \frac{8}{3}\frac{g^3}{m^6}S^4 + \frac{56}{3}\frac{g^4}{m^9}S^5 + \ldots \Rightarrow 2\frac{\partial \mathcal{F}}{\partial S} = \frac{S^2}{2a^3} + \frac{S^3}{3a^6} + \frac{35S^4}{96a^9} + \ldots \tag{5.62}$$

This shows that equation (5.60) is in excellent agreement with the four-dimensional answer.

$U(4) \to U(2) \times U(2)$ WITH CUBIC SUPERPOTENTIAL

In this section the gauge group is $U(4)$ and the superpotential as in equation (5.44). The goal is to calculate the quantum corrections around the following (classical) vacuum

$$p_{1,3} = a \qquad p_{2,4} = -a. \tag{5.63}$$

Classically this vacuum breaks the $U(4)$ to a $U(2) \times U(2)$ symmetry. So in this case there are two chiral superfields involved: S_1 and S_2. However, it is not known (yet) how to integrate in these glueball fields. Only for the sum $S = S_1 + S_2$ it is known how to integrate it in: identify it with the Lagrange multiplier. To integrate in S the same approach as in the previous section is used.

Using the $U(4)$ Lax matrix

$$L = \begin{pmatrix} p_1 & y_1 & 0 & z \\ 1 & p_2 & y_2 & 0 \\ 0 & 1 & p_3 & y_3 \\ y_0/z & 0 & 1 & p_4 \end{pmatrix} \tag{5.64}$$

and plugging it into the tree-level superpotential gives

$$W_{eff} = \sum_{i=1}^{4} W(p_i) + y_{i-1}(p_{i-1} + p_i), \text{ with } p_0 = p_4. \tag{5.65}$$

Extremizing the superpotential with respect to the p's gives the following four equations

$$p_1 = \sqrt{a^2 - y_0 - y_1} \qquad p_3 = \sqrt{a^2 - y_2 - y_3} \tag{5.66}$$

$$p_2 = -\sqrt{a^2 - y_1 - y_2} \qquad p_4 = -\sqrt{a^2 - y_0 - y_3}. \tag{5.67}$$

After substituting this into the superpotential, the superpotential is expressed in terms of the y's only. In principle one should proceed to integrate out the y's, however the algebra involved is rather messy, therefore it is easier to expand the superpotential as a power series in the y's. Up to second order the superpotential then reads

$$W_{\text{eff}} = S\log(\frac{\Lambda_4^8}{y_0y_1y_2y_3}) - \frac{1}{m}(y_0-y_2)(y_1-y_3)+\ldots \tag{5.68}$$

Integrating out the y's yields

$$y_0 = \frac{Sm}{2y_3}, \quad y_1 = -y_3, \quad y_2 = -\frac{Sm}{2y_3} \tag{5.69}$$

and leads to the following effective superpotential

$$W_{\text{eff}} = S\left(2+\log\left(\frac{4\Lambda_4^8}{m^2S^2}\right)\right). \tag{5.70}$$

The scale Λ_4, corresponding to the $U(4)$, can be related to the scales of the $U(2)$'s Λ_2 (using the notation $\Lambda_2^3 = m\Lambda^2$)

$$\Lambda_2^6 = \frac{\Lambda_4^8}{m^2} = m^2\Lambda^4 \Rightarrow \Lambda_4^2 = m\Lambda. \tag{5.71}$$

So finally one arrives at

$$W_{\text{eff}} = 2S\left(1-\log\left(\frac{S}{2m\Lambda^2}\right)\right). \tag{5.72}$$

In order to compare this with the four-dimensional result [40]

$$W_{\text{eff}} = 2\left(-S_1\log\left(\frac{S_1}{m\Lambda^2}\right) - S_2\log\left(\frac{S_2}{m\Lambda^2}\right)+S_1+S_2+\ldots\right) \tag{5.73}$$

this effective superpotential should be expressed in terms of $S = S_1 + S_2$. This means the difference $S_1 - S_2$ needs to be integrated out. Since action is expanded only to first order in the chiral superfields, the two chiral superfields don't mix and integrating out $S_1 - S_2$ is trivial. The result is that $S_1 = S_2 = S/2$, substituting this in the superpotential gives back equation (5.72).

(5.5) Extending the conjecture to G_2

The conjecture has been tested for $U(N)$ [24, 25] and the other classical gauge groups $SO(N)$, $Sp(2N)$ [36] pretty well. A further test for the conjecture would be to test it for the exceptional gauge groups, of which for example G_2 is the simplest.

There are field theory results known to check the results of the conjecture against. However in a way this conjecture can be used to generate new results. For example, for G_2 it is not known how to write the effective superpotential, expressed in the glueball field S, using a matrix integral. Using the integrating in procedure it is, in principle, possible to derive this without resorting to the calculation of Feynman diagrams.

(5.5.1) CLASSICAL FIELD THEORY RESULTS

In this section a brief overview of the field-theory results of the quantum vacua of G_2 gauge theory will be given. More details can be found in [26]. Central in the discussion is the characteristic polynomial of the adjoint G_2 scalar field

$$P_{G_2}(x) = \frac{1}{x}\det(x\mathbb{I} - \phi) = x^6 - 2ux^4 + u^2x^2 - v. \tag{5.74}$$

The adjoint scalar field ϕ is represented by the seven-dimensional representation. In a vacuum, the D-term equations allows this field to be chosen diagonally

$$\phi_{class} = \text{diag}(\phi_1 + \phi_2, 2\phi_1 - \phi_2, \phi_1 - 2\phi_2, 2\phi_2 - \phi_1, \phi_2 - 2\phi_1, -\phi_1 - \phi_2, 0). \tag{5.75}$$

However the two variables $\phi_{1,2}$ are not suitable to parameterize space of classical vacua. Instead the two quantities $u = \frac{1}{4}\text{Tr}\,\phi^2$, $v = \frac{1}{6}\text{Tr}\,\phi^6 - \frac{1}{96}(\text{Tr}\,\phi^2)^3$ parameterize the two-dimensional moduli space. They can also be used to write down a tree-level superpotential

$$W_{tree} = g_2 u + g_6 v. \tag{5.76}$$

The theory with this superpotential has two classical vacua. In the first vacuum the symmetry is not broken

$$\phi_1 = \phi_2 = 0 \Rightarrow P_{G_2}(x) = x^6. \tag{5.77}$$

In the second vacuum it is broken to $SU(2) \times U(1)$

$$\phi_1 = \phi_2 = e = \left(-\frac{g_2}{4g_6}\right)^{1/4} \Rightarrow P_{G_2} = (x^2 - e^2)^2(x^2 - 4e^2). \tag{5.78}$$

(5.5.2) FINDING THE QUANTUM VACUA

THE DUAL AFFINE LIE ALGEBRA OF G_2

The calculation of the quantum vacua can not be done without knowing the Lax matrix for the G_2 system. In order to find the Lax matrix, a suitable explicit representation for the affine dual of the G_2 Lie algebra should be constructed. Since this

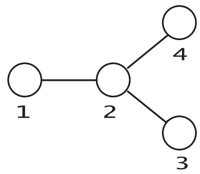


Figure 5.4: *The Dynkin diagram of the D_4 algebra. The S_3 permutation symmetry interchanges the outer most nodes.*

is not exactly standard material, a procedure to do this will be described here (also see appendix A). First of all the dual affine Lie algebra of G_2, $(G_2^{(1)})^\vee$, is equivalent to a twisted version of the affine $SO(8)$ Lie algebra: $D_4^{(3)}$. The D_4 algebra has a special property called triality. The origin of this property is the S_3 symmetry of the Dynkin diagram of D_4 (figure 5.4). Twisting the algebra means identifying the algebra under the action of the symmetry of the Dynkin diagram: permutations of the roots $\alpha_1, \alpha_3, \alpha_4$. The subalgebra invariant under these permutations is an G_2 algebra and will be the horizontal algebra of the affine algebra under construction. This G_2 horizontal algebra is easily constructed

$$h_1 = H_1 + H_3 + H_4 \quad e_1^+ = E_1^+ + E_3^+ + E_4^+ \tag{5.79}$$
$$h_2 = H_2 \quad e_2^+ = E_2^+. \tag{5.80}$$

The Cartan matrix of the horizontal G_2 subalgebra is

$$A_{G_2} = \begin{pmatrix} 2 & -3 \\ -1 & 2 \end{pmatrix}. \tag{5.81}$$

and the Cartan matrix of $D_4^{(3)}$ is

$$A = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -3 & 2 \end{pmatrix}. \tag{5.82}$$

To get the explicit Chevalley generators of the full $D_4^{(3)}$ algebra, the only thing that needs to be done is to find the Chevalley generators corresponding to the affine root. The element of the Cartan subalgebra corresponding to the affine root is given by

$$h_0 = -(a_1^\vee h_1 + a_2^\vee h_2) = -2h_1 - 3h_2, \tag{5.83}$$

where $a_i^\vee$ are the co-marks derived from 5.82 (remember that for an affine Lie algebra the co-marks is the eigenvector, belonging to eigenvalue zero, of the Cartan matrix). By demanding the canonical commutation relations for the Chevalley generators

$$[h_i, e_j^{\pm}] = \pm A_{ij} e_j^{\pm} \quad , \quad [e_i^+, e_j^-] = \delta_{ij} h_i, \tag{5.84}$$

it is easy to find $e_0^{\pm}$. Note that allthough the fundamental of G_2 is seven-dimensional the Lax will be an 8×8 matrix since the whole construction will be based on the fundamental of D_4. The explicit form of the Lax matrix that will be used reads

$$\begin{pmatrix} p_1+p_2 & y_2 & 0 & 0 & 0 & y_1 & -z & 0 \\ 1 & 2p_1-p_2 & 0 & ay_1 & by_1 & 0 & 0 & -z \\ 0 & 0 & p_1-2p_2 & -a & b & 0 & 0 & y_1 \\ 0 & a & -ay_2 & 0 & 0 & -a & ay_1 & 0 \\ 0 & b & by_2 & 0 & 0 & b & by_1 & 0 \\ 1 & 0 & 0 & -ay_2 & by_2 & 2p_2-p_1 & 0 & 0 \\ -\frac{y_0}{z} & 0 & 0 & a & b & 0 & p_2-2p_1 & y_2 \\ 0 & -\frac{y_0}{z} & 1 & 0 & 0 & 0 & 1 & -p_1-p_2 \end{pmatrix}, \tag{5.85}$$

with $a = \sqrt{1/2}$, $b = \sqrt{3/2}$. The constraint the y's need to obey is

$$36 y_0 y_1^2 y_2 = \Lambda^8. \tag{5.86}$$

The unusual normalization is kept here in order to comply with literature.

Using the Lax matrix

In order to compute the quantum vacua, according to the conjecture, all that needs to be done is to plug-in the Lax matrix into the tree-level superpotential (equation 5.76). Following [26], the classical superpotential that will be studied is

$$W = g_2 u - g_6 v + \frac{4 g_6}{27} u^3. \tag{5.87}$$

The Seiberg-Witten curve has a different form if compared to the $U(N)$ case. For G_2 the spectral curve is given by

$$P_{quan}(z,x) = 3\left(z - \frac{\Lambda^8}{36z}\right)^2 - x^2\left(z + \frac{\Lambda^8}{36z}\right)(6x^2 - 2u) - x^2 P_6(x,u,v) = 0, \tag{5.88}$$

where

$$P_6(x,u,v) = x^6 - 2ux^4 + u^2x^2 - v. \tag{5.89}$$

Note that this curve is not of the hyperelliptic type. A proposal for a hyperelliptic curve will be introduced later on.

The quantum vacua are now obtained and checked as usual. First the W_{dual} is extremized

$$\mathrm{d}W_{eff}(L) = 0, \tag{5.90}$$

then the results are plugged into the spectral curve

$$\det(x\mathbb{I} - L) = 0 \tag{5.91}$$

and the factorization of this curve is checked. This procedure is worked out in detail in [26, section 3], complete agreement with known field theory results is found.

To be more explicit the superpotential is written out (expanding u and v):

$$W = g_2 u + g_6(-v + \frac{4}{27}u^3) = \frac{1}{4}g_2 \mathrm{Tr}\,(\Phi^2) - \frac{1}{6}g_6 \mathrm{Tr}\,(\Phi^6) + \frac{11}{864}g_6 \mathrm{Tr}\,(\Phi^2)^3 \ . \tag{5.92}$$

When the Lax matrix L is substituted for the adjoint scalar Φ in this expression, some terms that explicitly depend on z appear. It is convenient to get rid off those, therefore the superpotential used for calculations is corrected such that the z dependent terms disappear

$$\begin{aligned} W = &\frac{1}{4}g_2 \mathrm{Tr}(L^2) - \frac{1}{6}g_6 \mathrm{Tr}\,(L^6) + \frac{11}{864}g_6 \mathrm{Tr}\,(L^2)^3 - \frac{5}{2}g_6 \mathrm{Tr}\,(L^2)(z + \frac{y_0 y_1^2 y_2}{z}) \\ &+ \ell \log\left(\frac{y_0 y_1^2 y_2}{\Lambda^8/36}\right) \ . \end{aligned} \tag{5.93}$$

Here we have also imposed the constraint on y_i using a Lagrange multiplier ℓ. The equations for the extrema of this superpotential are written down in appendix C.

As said before, some new results can be obtained as well. These results, discussed below, are the superpotentials in terms of the glueball field S. Both qualitatively different solutions, broken and unbroken gauge symmetry, will be discussed.

CASE 1: UNBROKEN GAUGE SYMMETRY

Using a, by now familiar, procedure one can obtain the equations for the extrema of the superpotential. For G_2 these equations are too difficult (see appendix C) to solve in generality. However, solutions can be found by trial and error. It is not too difficult to see that a solution is given by

$$p_1 = p_2 = 0, \quad y_1 = \frac{2}{3}y, \quad y_2 = \frac{1}{3}y, \quad y_0 = y, \tag{5.94}$$

where $y = 3^{1/4}\Lambda^2/2$. This solution corresponds to the situation in which the gauge group is classically unbroken. In this case gauge invariant parameters read $u = 3^{1/4}2\Lambda^2$ and $v = 4/3^{1/4}\Lambda^6$.

In order to integrate in the glueball field S one needs to forget, for a moment, that a solution for the y_i has been found. That is, one discards the fact that y is known and treats it as an undetermined variable. In that case, the superpotential reads

$$W = 4g_2 y - \frac{32}{27} g_6 y^3 + S \log\left(\frac{3\Lambda^8}{16y^4}\right) \ . \tag{5.95}$$

The next step is then to integrate out y

$$\frac{\partial W}{\partial y} = 4g_2 - \frac{32}{9}\, g_6 y^2 - 4\frac{S}{y} = 0 \ , \tag{5.96}$$

which is an equation that can be used to solve for y in terms of S. In fact the solution can be given in power series of S. Up to $\mathcal{O}(S^8)$ one finds

$$y = \frac{1}{g_2}\, S + \frac{8}{9}\, \frac{g_6}{g_2^4}\, S^3 + \frac{64}{27}\, \frac{g_6^2}{g_2^7}\, S^5 + \frac{2048}{243}\, \frac{g_6^3}{g_2^{10}}\, S^7 \ . \tag{5.97}$$

Plugging the above expression for y into the effective superpotential one gets

$$W = -4S\left(\log\left(\frac{2S}{3^{1/4} g_2 \Lambda^2}\right) - 1\right) - \frac{32}{27}\, \frac{g_6}{g_2^3}\, S^3 - \frac{128}{81}\, \frac{g_6^2}{g_2^6}\, S^5 - \frac{8192}{2187}\, \frac{g_6^3}{g_2^9}\, S^7 \ . \tag{5.98}$$

up to order eight in the glueball field S. If the conjecture is correct and valid for G_2 then this would be the superpotential in terms of the glueball superfield S.

CASE 2: G_2 BROKEN TO $SU(2) \times U(1)$

The equations of motion coming from the potential (5.93) also have another solution. In fact it can also be seen that the following ansatz solves the equations given in appendix C

$$p_1 = p, \quad p_2 = 2p, \quad y_0 = \pm\frac{4\Lambda^4}{9e^2}, \quad y_1 = -\frac{3e^2}{4}, \quad y_2 = \pm\frac{\Lambda^4}{9e^2}, \tag{5.99}$$

where $e = (-g_2/4g_6)^{1/4}$ and p is given by

$$p^2 = -\frac{1}{12}\left(e^2 \pm \frac{10\Lambda^4}{27e^2}\right) \ . \tag{5.100}$$

Looking at the limit $\Lambda \to 0$, one can see that this solution corresponds to the situation where the gauge group is classically broken to $SU(2) \times U(1)$.

The gauge invariant parameters are found to be

$$u = 3e^2 \mp \frac{\Lambda^4}{2e^2}, \quad v = \mp 4e^2\Lambda^4 + \frac{\Lambda^8}{3e^2} \mp \frac{\Lambda^{12}}{54e^6} \ , \tag{5.101}$$

and therefore the quantum superpotential reads

$$W = 3g_2e^2 + 4g_6e^6 \pm 2\sqrt{-g_2g_6}\,\Lambda^4\ , \tag{5.102}$$

which is the same as the one found in the field theory [26].

Since quantum mechanically the gauge symmetry is broken to $U(1)$, one would expect a free parameter in the solution. Note that the situation differs from the $U(N)$ case, there a center of mass $U(1)$ exists that did not manifest itself in the solution. Because G_2 is a simple Lie group it does not have this center of mass $U(1)$. Thus one would expect to see one free parameter in the solution. Obviously the solution 5.99 does not have a free parameter, which means that the ansatz $p_2 = 2p_1 = 2p$ somehow fixes this parameter. This leads to the conclusion that the solution is merely a special case in a one-parameter family of solutions.

Nevertheless one can still see the existence of the free parameter by considering flows in the integrable system. For G_2 there are two independent flows, generated by the Hamiltonians $\mathrm{Tr}\,L^2$ and $\mathrm{Tr}\,L^6$, or equivalently by u and v. The flows of the dynamical variables $\xi \in \{p_i, y_j\}$ are calculated by considering the Poisson brackets

$$F_k(\xi) = \frac{\partial}{\partial t_k}\xi = \{\xi, \mathrm{Tr}\,L^k\}. \tag{5.103}$$

The t_k is the time associated with the Hamiltonian $\mathrm{Tr}\,L^k$. To calculate the Poisson brackets one has to identify the coordinates that correspond to the conjugate momenta p_i; these are the x's appearing in

$$y_i = \exp(\alpha_i \cdot x), \tag{5.104}$$

with the α_i the simple roots of $D_4^{(3)}$

$$\alpha_0 = -(2\alpha_1 + \alpha_2), \quad \alpha_1 = (0, \sqrt{2}), \quad \alpha_2 = (\frac{1}{2}\sqrt{6}, -\frac{3}{2}\sqrt{2}). \tag{5.105}$$

The Poisson brackets then read

$$\{p_i, y_j\} = (\alpha_j)_i y_j, \tag{5.106}$$

where $(\alpha_j)_i$ is the i-th component of the j-th root in the basis (5.105). Using these brackets it is straightforward to calculate the flows $F_k(\xi)$. The result is that for all $\xi \in \{p_i, y_j\}$ the two flows F_2 and F_6 are related in the following way:

$$F_6(\xi) = -\frac{9}{4}\left(3\frac{g_2}{g_6} + 176\Lambda^4 + 2112\Lambda^8\frac{g_6}{g_2}\right)F_2(\xi). \tag{5.107}$$

So, indeed, there is exactly one independent flow and therefore precisely one free parameter in the solution. This establishes the fact that the symmetry is broken down to a single $U(1)$.

(5.5.3) Deriving the resolvent for G_2

In this section an alternative, but very useful, way to calculate the quantum effective action will be discussed. This method uses the so-called resolvent, a function that encodes the vacuum expectation values of operators of the form

$$\mathrm{Tr}\,\phi^k, \quad k = 0, \ldots, \infty. \tag{5.108}$$

The resolvent $R(x)$ is usually defined as

$$R(x) := \left\langle \mathrm{Tr}\,\frac{1}{x-\phi} \right\rangle = \sum_{k=0}^{\infty} \frac{\langle \mathrm{Tr}\,\phi^k \rangle}{kx^{k+1}}. \tag{5.109}$$

Since the superpotential is expressed in terms of traces of powers of ϕ, knowing the resolvent also means knowing the quantum effective superpotential. In the case of $U(N)$ the resolvent can be expressed in terms of the characteristic polynomial P_N that defines the Seiberg-Witten curve [16, 41, 42]

$$R(x) = \frac{P'_N(x)}{\sqrt{P_N(x)^2 - 4\Lambda^{4N}}}, \tag{5.110}$$

with

$$P_N(x) = \det(x\mathbb{I} - \phi) = x^N + u_1 x^{N-1} + \cdots + u_N. \tag{5.111}$$

In the classical limit the resolvent can be written as

$$R_{class}(x) = \frac{P'_N(x)}{P_N(x)} = \partial_x \log P_N(x) = \partial_x \log \det(x\mathbb{I} - \phi). \tag{5.112}$$

Inspired by this simple looking classical expression one can write the exact quantum result in a similar way

$$R_{quant}(x) = \partial_x \log\left(P_N(x) + \sqrt{P_N(x)^2 - 4\Lambda^{2N}} \right). \tag{5.113}$$

This equation can be generalized to G_2 and used to calculate properties of the quantum vacua. The generalization is based on the integrable system solution of section 4 in [25]. By know it should be familiar to the reader that the Seiberg-Witten curve has an underlying integrable system. The integrable system is characterized by the existence of a complete set of action-angle variables. In terms of these variables the evolution in the phase-space of the classical mechanical system becomes quite simple, half of the variables are conserved and the other half (the angle variables) evolve with constant velocity. Further, to the integrable system one can associate a Riemann surface, which is equivalent to the Seiberg-Witten curve. The conserved quantities then correspond to the moduli of this surface and the angle variables are coordinates on the Jacobian of this Riemann surface. In general, only a subset of the moduli correspond to action variables, and the number of flows need therefore not be equal to the dimension of the Jacobian. The equations of motion of the integrable system correspond to linear flows on the Jacobian.

The main idea of section 4 of [25] is that the superpotential is at an extremum if the velocities of the flows on the Jacobian are zero. The velocities of the flows are expressed in terms of the superpotential $W(x)$ and the one-forms ω_k (see [43]):

$$v_k(W) = \mathrm{res}_{x=\infty} \left(W'(x)\omega_k \right). \tag{5.114}$$

The one-forms ω_k are just the Seiberg-Witten forms. For $U(N)$ they are given explicitly by

$$\omega_k = \frac{x^{N-k}\mathrm{d}x}{\sqrt{P_N(x)^2 - 4\Lambda^{2N}}}, \quad k = 0, \ldots, N-1. \tag{5.115}$$

Residues can also be used to conveniently express the superpotential in terms of the resolvent:

$$W_{\text{quantum}} = \mathrm{res}_{x=\infty} \left(W(x)R(x) \right). \tag{5.116}$$

It turns out to be possible to express both the one-forms ω_k and the resolvent in terms of a single function:

$$\Omega(x, u_k) = (\log\det(x - L(z)))|_{z^0}, \tag{5.117}$$

by which the z-independent part of $\log\det(x - L(z))$ is meant. Further, $L(z)$ is the Lax matrix (with spectral parameter z) of the integrable system that underlies the Seiberg-Witten curve and the u_k are the moduli of this curve. As it stands, Ω is not well defined, because the z-independent part of some complicated function with branch cuts has to be extracted. One way to define Ω is as follows. Since the characteristic polynomial $\det(x - L(z))$ is symmetric under the interchange of z and $1/z$, one can write

$$\det(x - L(z)) = a_0^2 \prod_{t=1}^{r} (a_t - z)(a_t - 1/z) \tag{5.118}$$

which allows the following definition

$$\Omega \equiv 2\sum_{t=0}^{r} \log a_t. \tag{5.119}$$

Having defined Ω, the resolvent is given by

$$R(x) = \partial_x \Omega(x, u_k) \tag{5.120}$$

and the one-forms by

$$\omega_k = \frac{\partial \Omega}{\partial u_k} \mathrm{dx}. \tag{5.121}$$

Actually, the definition of Ω still suffers from minus sign ambiguities, which will be fixed by demanding that the resolvent that follows from this Ω has the expansion

$$R(x) = \sum_{i=1}^{\infty} \frac{\mathrm{tr}(L(z)^{i-1})|_{z^0}}{x^i}. \tag{5.122}$$

In order to show that this proposal makes sense, the resolvent and one-forms for $U(N)$ will be calculated. From the curve for $U(N)$

$$\det(x - L(z)) = P_N(x) + (-1)^N (z + \frac{1}{z}) \tag{5.123}$$

one easily derives that $a_0^2 = 1/a_1$ and

$$a_1 = (P + \sqrt{P^2 - 4})/2. \tag{5.124}$$

This yields for the function Ω

$$\Omega = \log((P + \sqrt{P^2 - 4})/2), \tag{5.125}$$

from which the usual resolvent follows

$$R(x) = \partial_x \Omega = \frac{P'(x)}{\sqrt{P(x)^2 - 4}} \tag{5.126}$$

and the one-forms

$$\omega_k = \partial_{u_k} \Omega \mathrm{dx} = \frac{x^{N-k}}{\sqrt{P(x)^2 - 4}} \mathrm{dx}. \tag{5.127}$$

In order to apply this procedure to G_2 the roots a_t for the G_2 curve must be computed first. The algebraic curve for G_2 is given by (see equation 5.88)

$$3\left(z - \frac{\Lambda^8}{36z}\right)^2 - x^2\left(z + \frac{\Lambda^8}{36z}\right)(6x^2 - 2u) - x^2 P(x) = 0 \tag{5.128}$$

written in terms of $y = z + \frac{\Lambda^8}{36z}$ this reads

$$3y^2 - x^2 y(6x^2 - 2u) - x^2 P(x) - \frac{\Lambda^8}{3} = 0. \tag{5.129}$$

This equation has two solutions:

$$y_\pm = x^2\left(x^2 - \frac{u}{3}\right) \pm \frac{1}{3}\sqrt{x^4(3x^2 - u)^2 + 3x^2 P(x) + \Lambda^8}, \tag{5.130}$$

yielding the four roots of the algebraic curve

$$z_{+\pm} = \frac{1}{2}y_+ \pm \frac{1}{6}\sqrt{9y_+^2 - \Lambda^8}, \qquad z_{-\pm} = \frac{1}{2}y_- \pm \frac{1}{6}\sqrt{9y_-^2 - \Lambda^8}. \tag{5.131}$$

To write down Ω one has to make a choice for the roots. One should pick one root from $\{z_{++}, z_{+-}\}$ and one from $\{z_{--}, z_{-+}\}$, so there are four possible choices:

$$\begin{aligned}\Omega &= \eta(\log z_{++} + \epsilon \log z_{--}), \text{ with } \epsilon^2 = \eta^2 = 1 \\ &= \eta(\log\left(\frac{1}{2}y_+ + \frac{1}{6}\sqrt{9y_+^2 - \Lambda^8}\right) + \epsilon \log\left(\frac{1}{2}y_- + \frac{1}{6}\sqrt{9y_-^2 - \Lambda^8}\right)) \end{aligned} \tag{5.132}$$

with the choice $\eta = 1, \epsilon = -1$ the resolvent reads

$$R(x) = \partial_x\Omega = \frac{3\partial_x y_+}{\sqrt{9y_+^2 - \Lambda^8}} - \frac{3\partial_x y_-}{\sqrt{9y_-^2 - \Lambda^8}}. \tag{5.133}$$

The expansion of the resolvent around $x = \infty$ should have the form of (5.122). Indeed the choice $\eta = 1, \epsilon = -1$ was correct, since after doing the expansion one ends up with

$$\begin{aligned} R(x) &= \frac{8}{x} + \frac{4u}{x^3} + \frac{4u^2}{x^5} + \frac{4u^3 + 6v}{x^7} + \frac{4u^4 + 16uv + \frac{20\Lambda^8}{3}}{x^9} \\ &+ \frac{4u^5 + 30u^2v + 30u\Lambda^8}{x^{11}} + \frac{4u^6 + 48u^3v + 6v^2 + \frac{250}{3}\Lambda^8}{x^{13}} \\ &+ \mathcal{O}\left(\frac{1}{x^{15}}\right). \end{aligned} \tag{5.134}$$

One can check that the coefficients in this expansion correspond to the traces of powers of the Lax matrix. Classically, one could write for the resolvent :$P_6'(x)/P_6(x)$, this would generate the correct expansion if one would set to zero the terms in the expansion that explicitly depend on Λ. Apparently this naive guess for the resolvent is correct up to order $1/x^7$.

Next the flow equations (5.114) are used to determine the minima of the superpotential. One therefore has to calculate the one-forms $\partial_{u,v}\Omega$

$$\omega_u = \partial_u\Omega = \partial_u a(x)R_1(x) + \partial_u b(x)R_2(x) \tag{5.135}$$

$$\omega_v = \partial_v\Omega = \partial_v a(x)R_1(x) + \partial_v b(x)R_2(x). \tag{5.136}$$

The conditions that the flows on the Jacobian vanish ($v_l = 0$) then imply:

$$\begin{aligned} x^2R_2(x)W'(x) &= r_v(x) + \sum_{l=1} \frac{c_l}{x^{2l+1}} \\ \left(-\frac{x^2}{3}R_1(x) + 4x^4(2u - 3x^2)R_2(x)\right)W'(x) &= r_u(x) + \sum_{l=1} \frac{d_l}{x^{2l+1}}. \end{aligned} \tag{5.137}$$

The polynomials r_u and r_v are the regular parts of the expansion. The flow equations for the $U(N)$ case allowed the factorization of the gauge theory and Matrix model curve (see [25] section 4) to be derived, in a similar spirit equations 5.137 should somehow define the analog of the Matrix model curve for G_2. Unfortunately, equation 5.137 is hardly a manageable equation, and it is therefore hard to draw general conclusions from these equations. In section 5.2 of [26] the factorization of the Matrix model curve for a superpotential with terms up to order six is worked out.

The remaining part of this section will be used to show that (5.137) is indeed equivalent to minimizing the superpotential. For definiteness the superpotential will be chosen to be $W'(x) = g_2 x + g_6 x^5$. The values of u and v in the minimum determine the Seiberg-Witten curve completely and therefore also the factorization properties of this curve. To study the factorization properties it is useful to consider the conditions for the curve to develop a double zero:

$$\begin{aligned} P_6(x_0) &= \pm 2\Lambda^4(x_0^2 - u/3) && (5.138)\\ P_6'(x_0) &= \pm 4\Lambda^4 x_0, && (5.139) \end{aligned}$$

these equations can be used to solve for u and v in terms of x_0. So there is only one free parameter, not two. Therefore the two equations (5.137) are replaced by the single equation

$$\mathrm{res}_{x=\infty}\left(\partial_{x_0}\Omega W'(x)\right) = 0. \qquad (5.140)$$

This equation can be used to solve for x_0, which allows the u and v to be expressed in terms of the coupling constants and the energy scale Λ. One can then substitute u and v into equation (5.138) and study its factorization properties. For the superpotential $W'(x) = g_2 x + g_6 x^5$, three classes of solutions are found:

First Solution

$$x_0 = 0 \Rightarrow P_6(x) - 2\Lambda^4(x^2 - u/3) = x^4(x^2 - 2\sqrt{2}\Lambda^2) \qquad (5.141)$$

Second Solution

$$x_0 = \eta\left(\frac{8}{3}\right)^{1/4}, \quad \eta^4 = -1 \Rightarrow P_6(x) - 2\Lambda^4(x^2 - u/3) = (x^2 \pm 2\mathrm{i}\sqrt{\frac{2}{3}}\Lambda^2)^3. \qquad (5.142)$$

This factorization has the same form as one would expect for a superconformal point.

Third Solution

$$x_0 = \epsilon \left(\Lambda^4 - 6e - \frac{5}{33}\sqrt{\frac{9}{e^2} - 66\frac{\Lambda^4}{e} + 22\Lambda^8} \right)^{1/4}, \quad \epsilon^4 = 1, \quad e = \frac{g_6}{g_2} \tag{5.143}$$
$$\Rightarrow P_6(x) - 2\Lambda^4(x^2 - u/3) = (x^2 - \alpha)(x^2 - \beta)^2$$

α and β can be expressed in terms of e and Λ.

In order to check the claim that equation (5.140) is equivalent to minimizing the superpotential, if suffices to minimize

$$W = \frac{g_2}{2}u + \frac{g_6}{6}v + A(P_6(x_0) \pm 2\Lambda^4(x_0^2 - u/3)) + B(P_6'(x_0) \pm 4\Lambda^4 x_0) \tag{5.144}$$

with respect to u, v, A, B and x_0. The Lagrange multipliers A and B are there to enforce the formation of a double zero. The calculations are pretty straightforward and complete agreement is found, suggesting that Ω indeed generates the one-forms as described.

The resolvent derived in this section also hints at the existence of a hyper-elliptic curve for G_2. This can be seen as follows. The resolvent can be written in the form

$$R(x) = \frac{r(x)}{x^2\sqrt{P_6(x)^2 - 4\Lambda^8(x^2 - u/3)^2}} \tag{5.145}$$

with $r(x)$ some function without poles. Comparing this resolvent to that of $U(N)$ suggests that

$$y^2 = P_6(x)^2 - 4\Lambda^8(x^2 - u/3)^2 \tag{5.146}$$

is in fact a hyper-elliptic curve for G_2. Indeed, in analyzing the factorization of the G_2 curve, expressions like $P_6(x) \pm 2\Lambda^4(x^2 - u/3)$ pop up everywhere. As it turns out, this hyperelliptic curve is just a small modification of an earlier (rejected) proposal for a hyperelliptic curve for G_2 (see [44]).

Note that the resolvent of the gauge theory contains arbitrarily high powers of the adjoint scalar field. The precise definition of such operators in the quantum theory depends on a choice of UV completion of the theory. The integrable system approach prefers one particular UV completion over the others. This choice is basically set by the following requirement that was imposed:

$$\mathrm{tr}(\Phi^i) \equiv \mathrm{tr}(L(z)^i)|_{z^0}. \tag{5.147}$$

In the case of $U(N)$, this was also the UV completion preferred by string theory. Here, the integrable system provides a natural UV completion for the exceptional gauge groups as well. It would be interesting to explore other UV completions. Other UV completions can be obtained by using a different representation for the Lax matrix. Previously the fundamental representation was used to describe the Lax matrix, choosing a different one will yield a different UV completion.

CHAPTER 6

CONCLUSIONS

In the first part of this thesis a method was proposed to calculate the exact quantum vacua of four-dimensional $\mathcal{N} = 2$ supersymmetric gauge theories deformed to $\mathcal{N} = 1$ by a superpotential. The effective superpotential could be calculated in the theory compactified on $\mathbb{R}^{1+2} \times S^1$. In the three-dimensional limit the superpotential was easy to calculate, no quantum correction had to be calculated. It was only a matter of finding the right variables such that the minima of the superpotential give the correct quantum vacua. Those variables were conjectured to be given by a classical mechanical system: the Toda system. This system is an integrable system with a Lax matrix description. The adjoint scalar field should be substituted with the Lax matrix of the Toda system. In formula

$$W_{eff} = W_{tree}(L). \tag{6.1}$$

Ample evidence for various gauge groups was found for this conjecture [24, 25, 36, 26]. Much of the evidence came from comparing the results with field theory calculations, Seiberg-Witten curve factorizations to be more specific.

Using a special technique, ”integrating in”, it was also possible to obtain the superpotential in terms of the glueball field S. This made it possible to compare the results with the recent progress known as the Dijkgraaf-Vafa method. Because the use of the integrable system involves no integrals, some results were very easy to obtain. Exact agreement was found with known results, sometimes new results were obtained. Among the new results is the exact expression equation 5.59.

The conjecture was also extended to the gauge group G_2. Although of academic interest only, being an exceptional gauge group it is particularly interesting because one cannot take a large N limit of an exceptional gauge group. Matrix Model methods usually rely on taking a large N limit (see for example [45]). If the conjecture is correct and extentable to G_2, the proposed method can be used to obtain completely new results. As far as known results are concerned, agreement was found, strengthening the confidence in the conjecture. A proposal for the resolvent, an effective way to summarize vacuum expectation values of chiral operators, was put forward and tested, finding a hyperelliptic Seiberg-Witten curve for G_2 along the way.

This thesis began by reviewing confinement and supersymmetric gauge theories. The following question naturally arises: "Did we make progress?". It is certainly true that the use of integrable systems is useful in studying non-perturbative aspects of supersymmetric gauge theories. Although non-perturbative knowledge is necessary in order to understand confinement, current methods are still not powerfull enough to proof it. The closest thing to an analytical proof is that of Seiberg and Witten for a mass deformed $\mathcal{N} = 2$ theory [2]. The calculation of the superpotential in terms of the fermion bilinear glueball field S is sometimes used as an argument for confinement. In those arguments the mass of the glueball field in the quantum vacuum is determined. Usually the glueball field is massive, signaling a mass gap (and hinting at confinement) if the glueball field is the field with lowest mass. The question of the existence of a mass gap is then turned into the question whether or not the glueball field is the lightest field. The results obtained in this thesis do not shed light on that question. However, the use of integrable systems does allow one to answer non-perturbative question using a completely algebraic approach. In fact the quantum effective superpotential is determined by the dynamics of a classical mechanical system. Since the algebraic methods are easier to tackle than the path-integrals one encounters in field theory, it is easier to find (exact) results. Those results will hopefully add to the understanding of the vacuum structure of gauge theories in general.

Part II

D0-branes in a curved background

CHAPTER 7

INTRODUCTION AND MOTIVATION

Often string theory is proposed to be a candidate for a theory of everything, by which is meant a quantum theory of the four fundamental forces. Before 1995 string theory was in the embarrassing situation of having at least five possible candidates for a theory of everything. However, all these string theories are related to one another through a web of dualities. In 1995 [46] it was realized that all string theories could be derived from one big elusive theory: M-theory. The five known string theories can be seen as special limits of this M-theory. Although many properties of M-theory are still ill understood, the M-theory picture has proven to be quite useful in understanding string theory.

In the literature M-theory is often defined in the following way:

- its low energy (compared to the Planck scale) description is $d = 11$ supergravity.
- compactified on a circle with radius R, it is equivalent to type IIA string theory with coupling $g_s = R/l_s$.

The second part of this thesis will deal mainly with the second defining property of M-theory. Type IIA string theory has, besides closed strings, many extended objects in its spectrum: the D-branes. The simplest of those, the D0-branes, are point-like BPS objects charged under the Ramond-Ramond one-form. The BPS property relates the mass and the charge of the D0-branes:

$$m_{D0} = q = \frac{1}{l_s g_s}. \tag{7.1}$$

One can wonder what the corresponding objects are in M-theory. It turns out that D0-branes correspond to M-waves (the quantum counterparts of the supergravity pp-waves [47, 48]). The momentum of the M-waves in the compactified dimension translates into RR one-form charge from the ten-dimensional string theory perspective. The easiest way to see this is to use the supersymmetry algebra of $d = 11$ supergravity. Rewriting this algebra in terms of ten-dimensional spinors reveals that the momentum in the compactified dimension becomes a central charge in the ten-dimensional supersymmetry algebra. From this one can derive the following formula for the mass (and charge) of a D0-brane:

$$m_{D0} = q = \frac{1}{R}. \tag{7.2}$$

Combining equations 7.1 and 7.2 yields the relation between the string coupling and the radius of the circle: $g_s = R/l_s$. One should be able to recover M-theory by taking the decompactification limit (i.e. taking $R \to \infty$). In that limit the D0-branes become massless and, since this limit takes the system to a strong coupling regime, the D0-branes dominate the dynamics (see for [49] for a review). This reveals that the physics of D0-branes should be able to capture the physics of M-theory, and therefore of all the string theories. It is well known that the low energy dynamics of D0-branes is described by supersymmetric Yang-Mills theory in $9 + 1$ dimensions, reduced to $0 + 1$ dimensions. The dimensionally reduced theory is a theory of supersymmetric matrix quantum mechanics, it is sometimes referred to as: Matrix theory.

There is something puzzling about this, there is a mismatch between the ten-dimensional space-time the D0-branes live in and the eleven-dimensional nature of M-theory. The puzzle is resolved by considering the nine-dimensional transverse space of the D0-branes as the transverse space of M-theory in the infinite momentum (or light-cone) frame. In the infinite momentum frame (IMF) only objects with momentum in the compactified dimension survive, which means that only objects with RR one-form charge survive (the D0-branes). This is the essence of the Matrix-theory conjecture: D0-branes capture the full physics of M-theory in the IMF [50].

Although not explicitly mentioned, the discussion about Matrix and M-theory was for a *flat background*. If one wants to address any cosmological questions in M-theory, then a description of M-theory in a curved background must be given. A natural starting point would be to try to describe D0-branes in a curved background. Since the arguments that lead to the Matrix theory conjecture did not mention the background explicitly, it is tempting to generalize the conjecture to hold for arbitrary backgrounds (see for example [51, 52], note that subtleties arise when the curved space has a compact part [53]). Hence D0-branes in a curved background will capture the physics of M-theory in that same background. This is the main motivation for studying D0-branes in a curved background. As it turns out, it is far from trivial to extend the action of D0-branes in a flat background to that of an action for a curved background. The larger part of the coming chapters will be devoted to introducing and applying a calculational tool that is helpful in constructing such an action.

Before discussing the interplay between D0-branes and gravity, the theory of D0-branes in a flat background will be reviewed in chapter 8. To set the stage the D-geometry axioms will then be discussed in chapter 9, followed by an extension of these axioms in the shape of a new symmetry principle, to be discussed in chapter 10. Finally the D-geometry axioms and symmetry principle will be put to work in chapter 11. Evidence for interesting collective behavior of D0-branes in a curved background will be presented in chapter 12.

CHAPTER 8

D0-BRANES IN A FLAT BACKGROUND

The low energy action for a collection of D0-branes in a flat background can be derived from the $d = 9 + 1$ supersymmetric Yang-Mills theory. To be more specific, to describe N D0-branes one takes the gauge group $U(N)$. The field content of this supersymmetric field theory is a gauge field A^a_μ and an adjoint Majorana-Weyl spinor λ^a_α. The field strength and covariant derivatives are defined as follows (in a basis where the gauge field is Hermitian):

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + \mathrm{i}[A_\mu, A_\nu] \tag{8.1}$$

$$D_\mu \lambda = \partial_\mu \lambda + \mathrm{i}[A_\mu, \lambda]. \tag{8.2}$$

The action is just that of the spinor minimally coupled to the gauge field:

$$S_{\mathrm{SYM}} = \frac{-1}{2g^2_{\mathrm{SYM}}} \int \mathrm{d}^{10}\mathrm{x}\, \mathrm{Tr}\, \left(\tfrac{1}{2} F_{\mu\nu} F^{\mu\nu} + \mathrm{i}\bar{\lambda}\Gamma^\mu \mathcal{D}_\mu \lambda\right). \tag{8.3}$$

To dimensionally reduce the theory to $0 + 1$ dimensions one has to compactify the theory on a nine-dimensional torus with a volume $V = (2\pi R)^9$ and drop all dependence on the compact dimensions. This then leads to the following simplifications (the index i runs over the compact dimensions):

$$F_{0i} \to \partial_0 A_i + \mathrm{i}[A_0, A_i] = \mathcal{D}_0 A_i, \quad F_{ij} \to \mathrm{i}[A_i, A_j], \tag{8.4}$$

$$\mathcal{D}_i \lambda \to \mathrm{i}[A_i, \lambda]. \tag{8.5}$$

Here, the A_i and λ are Lie algebra valued $N \times N$ matrices. Substituting these simplifications into the action and rewriting it in terms of $X_i = 2\pi l_s^2 A_i$ gives:

$$S_{\mathrm{D0}} = \frac{V}{g_{\mathrm{SYM}}^2} \int \mathrm{dtTr} \left(\frac{1}{2(2\pi l_s^2)^2} \dot{X}_i \dot{X}^i + \frac{1}{4(2\pi l_s^2)^4} [X_i, X_j]^2 \right. \\ \left. + \frac{\mathrm{i}}{2} \lambda \dot{\lambda} + \frac{1}{4\pi l_s^2} \bar{\lambda} \Gamma^i [X_i, \lambda] \right). \tag{8.6}$$

The $A_0 = 0$ gauge is chosen here, turning the covariant derivatives into ordinary ones. This action has the following form:

$$S_{\mathrm{D0}} = \int \mathrm{dtTr} \left(\tfrac{m}{2} \dot{X}_i^2 + \tfrac{\lambda^2}{4} [X_i, X_j]^2 + \tfrac{\mathrm{i}}{2} \lambda \dot{\lambda} + \tfrac{1}{4\pi l_s^2} \bar{\lambda} \Gamma^i [X_i, \lambda] \right), \tag{8.7}$$

from which one can infer (using $m = 1/(g_s l_s)$) that $\lambda^2 = 1/(4\pi^2 l_s^3)$ and $1/g_{\mathrm{SYM}}^2 = 4\pi^2 l_s^3/(V g_s)$. At a first glance it might seem odd that N point-particles are described by nine $N \times N$ matrices instead of N nine-dimensional vectors. This is easily explained if one remembers the string theory origin of the D0-branes. The diagonal components of the matrices form N nine-dimensional vectors, describing the position of the D0-branes. The off-diagonal elements correspond to the open strings stretching between the D0-branes. The effect of these strings is captured by the potential, the $[X, X]^2$ term.

The classical vacua are easily determined from equation 8.7, the potential reaches it minimum when the matrices X^i all commute:

$$\mathrm{d}V = 0 \Rightarrow [X^i, X^j] = 0 \quad \forall_{i,j}. \tag{8.8}$$

Configurations for which all the matrices X^i commute are configurations in which the D0-branes are all on top of each other, hence the matrices X^i are proportional to the unit matrix.

D-branes in general carry RR charge, naturally coupling to the RR $p+1$ form. A single D0-brane thus couples to the RR one-form

$$S_{CS} = \int C^{(1)} = \int \mathrm{d}\tau C_i^{(1)} \frac{\mathrm{d}X^i}{\mathrm{d}\tau}. \tag{8.9}$$

In the presence of an NSNS two-form B-field or an open string gauge field background, a single Dp-brane can even have couplings to the other RR forms as well

$$S_{CS} = \int P \left[\sum_n C^{(n)} \mathrm{e}^B \right] \mathrm{e}^{2\pi l_s^2 F}. \tag{8.10}$$

For the D0-brane this general expression reduces to 8.9. If this expression is generalized to the case of more D0-branes [54], additional terms proportional to commutator terms appear

$$S_{CS} = \int \left[\mathrm{e}^{\mathrm{i}\lambda i_X i_X} \sum_n C^{(n)} \mathrm{e}^{B} \right] \mathrm{e}^{2\pi l_s^2 F}. \tag{8.11}$$

Here i_X is the innerproduct operator as is common in the theory of differential forms, that is

$$i_X C^{(1)} = X^i C_i^{(1)}. \tag{8.12}$$

The effect of extra commutator terms, from a different origin, in the action for D0-branes will be studied in chapter 12.

CHAPTER 9

D-GEOMETRY

In the previous chapter it was shown that the action for D0-branes in a flat (gravitational) background can be derived from Yang-Mills theory in ten dimensions. To derive the action for D0-branes in a curved background one would perhaps be tempted to use Yang-Mills theory in a curved background and follow the same procedure as for the flat case. This procedure however uses T-duality, which requires isometries in the nine spatial directions. Generically a curved space does not have these isometries, rendering the T-duality method useless.

Also there is no obvious generalization of the action of a single D0-brane to that of multiple D0-branes since one is confronted with difficult ordering problems. How would one for example generalize the following expression

$$S = \int \mathrm{d}\tau \; x^2 \delta_{ij} \dot{x}^i \dot{x}^j \; ? \tag{9.1}$$

Would this be

$$S = \int \mathrm{d}\tau \; \delta_{ij} \mathrm{Tr}\, X^2 \dot{X}^i \dot{X}^j, \tag{9.2}$$

or would it be

$$S = \int \mathrm{d}\tau \; \delta_{ij} \delta_{kl} \mathrm{Tr}\, X^k \dot{X}^i X^l \dot{X}^j \; ? \tag{9.3}$$

Ordering problems like this make it extremely difficult to find an action for D0-branes in a curved background. In fact there could be completely new terms that should be added, terms that vanish in the point-particle case such as any expression involving commutators. It turns out that this is exactly the case [55][56][57]: finding the correct action for multiple D0-branes is more than an ordering problem, new terms have to be added.

It should be clear by now that one has to look for alternative methods. The most direct method would be to calculate the β-functions of the superstring with some D0-branes present. This method is troublesome because of the massive strings stretching between the D0-branes (see [58, 59] for an attempt). If direct methods do not work, or are technically too difficult, there are indirect methods that have proven to be very powerful. One method is to list all the requirements for a sensible action of D0-branes in a curved background. This list then serves as the axioms the action needs to satisfy.

(9.1) The D-geometry axioms

In [60] Douglas compiles a list of rules, axioms, the action for D0-branes in a curved background should satisfy. The idea behind this, in his own words, is to "state the problem in a self-contained way which we could give to a mathematician". The main idea behind his axioms is to view the degrees of freedom of the D0-brane, $d\ N \times N$ matrices, as ordinary coordinates in a larger dN^2-dimensional space endowed with a metric:

$$\text{matrix: } X^i \quad \rightarrow \quad \text{coordinate: } X^I \text{with } I = iab, \quad a, b = 1, \ldots, N. \tag{9.4}$$

The geometry of this space describes the D0-branes, explaining the term D-geometry. The action for D0-branes can then be written compactly as

$$S = \int \mathrm{d}\tau\, G_{IJ} \dot{X}^I \dot{X}^j - V(X). \tag{9.5}$$

Here G_{IJ} is the metric on the dN^2-dimensional space and $V(X)$ is the potential. Of course this is just another way of writing the action, no new information was added to it. The action is supposed to satisfy the following axioms [60]

- The action is assumed to have an $U(N)$ isometry.
- G_{IJ} is a metric on the dN^2-dimensional space.
- The classical moduli space of the action for N D0-branes is equal to the space of N unordered points on the d-dimensional space.
- If the D0-branes are partitioned into groups of D0-branes that are on top of each other in the following way: $N = N_1 + \cdots + N_n$, then the unbroken gauge group is equal to $U(N_1) \times \cdots \times U(N_n)$.
- Strings stretching between two D0-branes have a mass equal to the geodesic distance between the two D0-branes. For equation 9.5 this implies that the

mass terms in a quadratic fluctuation analysis should be proportional to the geodesic distance between the D0-branes.

- The action is single trace.

According to Douglas these axioms should be enough to determine the action completely. In practice however, one takes into account all known properties of the D0-brane action, axiomatic or not. The action is well known if all the D0-branes are separated far from each other (it reduces to the action of N regular point-particles). Also, there are partial results when the action is reduced to terms linear in the Riemann tensor [61].

(9.2) MATRIX VALUED DIFFEOMORPHISMS

Even after a quick look at the axioms it seems that one obvious ingredient appears to be missing. If one tries to write down the coupling of D0-branes to gravity, wouldn't it be logical to impose diffeomorphism invariance, the gauge symmetry that corresponds to gravitational couplings? For the point-particle action this procedure is very successful, imposing diffeomorphism invariance there completely determines the coupling of a point-particle to gravity (see [55] for a short review).

In [55] an attempt is made to impose diffeomorphism invariance on the action for D0-branes. This is far from straightforward since it is not clear what diffeomorphism invariance means in D-geometry. Of course, one has the diffeomorphism invariance in the dN^2-dimensional space: since G_{IJ} is a metric and $\dot{X}^I$ a vector the combination

$$G_{IJ}\dot{X}^I\dot{X}^J \tag{9.6}$$

is diffeomorphism invariant. However what is required is invariance under diffeomorphims of the d-dimensional space. It is not at all clear how *those* diffeomorphisms act on the X^I. Take for example the d-dimensional diffeomorphism (diff)

$$x^i = f^i(x) = x^i + f^i_{jk}x^jx^k + f^i_{jkl}x^jx^kx^l + \dots \tag{9.7}$$

and its generalization to dN^2-dimensional diffeomorphisms (DIFF or matrix valued diffeomorphisms)

$$X^I = F^I(X) = X^I + F^I_{JK}X^JX^K + F^I_{JKL}X^JX^KX^L + \dots \tag{9.8}$$

If one wants to impose diff invariance in the action for D0-branes one has to lift the transformation 9.7 to DIFF. That is, given f^i construct a unique F^I. In [55] it is shown that this is not possible. One can only find a lift from diff to DIFF if the diff are actually a subgroup of DIFF. It turns out that this is not the case [55], any attempt of assigning some F^I to a certain f^i will destroy the homomorphism property of the map, making it impossible to lift diff to DIFF. From this one can conclude that imposing general coordinate invariance in D-geometry is far from straightforward. In the next chapter a new symmetry principle will be put forward, a symmetry principle that implements general coordinate invariance.

As an aside, note that diff can be seen as a quotient of DIFF

$$\text{diff} = \frac{\text{DIFF}}{\mathcal{C}}. \tag{9.9}$$

The operation $\mathcal{C}$ equates all elements in DIFF that differ by terms that involves commutators (and hence go to zero in the limit $N \to 1$).

CHAPTER 10

THE PRINCIPLE OF BASE-POINT INDEPENDENCE

Given the fact that it is impossible to lift d-dimensional diffeomorphisms to diffeomorphisms in the larger dN^2-dimensional space, an alternative way to implement diffeomorphism invariance is discussed in this chapter: base-point independence.

The principle of base-point independence is usually based on the background field expansion. This tool, which is useful regardless of base-point independence, will be introduced first.

(10.1) THE BACKGROUND FIELD EXPANSION

As explained in chapter 9, the action for D0-branes is written as a non-linear sigma model (NLσM) in dN^2 dimensions

$$S[X] = \int \mathrm{d}\tau \; G_{IJ}(X)\dot{X}^I \dot{X}^J. \tag{10.1}$$

Here $X(\tau)$ is a map from the manifold M (of dimension one) to the target space Σ, a manifold of dimension dN^2:

$$X \quad : \quad M \to \Sigma. \tag{10.2}$$

It so happens that this NLσM is precisely of the same form as the bosonic action used in string theory. Now, it turns out that it is hard to do calculations with this action that yield manifestly covariant results. The reason is the following. Suppose one wants to calculate the effect of quantum fluctuations π around some classical background configuration $\bar{X}$, i.e. one writes

$$X = \bar{X} + \pi. \tag{10.3}$$

Since X is a map from the space-time M to a *point* in the target-space Σ, the quantum fluctuation $\pi = X - \bar{X}$ is a difference of two-points. Although X can be viewed as a coordinate on Σ and therefore has well defined transformation properties, the difference of two points has no nice geometrical meaning in curved space. The consequence of this is that the expansion of the action S in terms of π will not be manifestly covariant under diffeomorphisms of Σ. The background field expansion is a solution to this problem. Actually the solution is very simple: connect the background field $\bar{X}$ with the full quantum field X using a geodesic in the target-space.

The problem of finding a covariant expansion is now turned into the problem of finding a general expression for the geodesic. For the geodesic $\lambda(t)$ one can write the following boundary value problem

$$\begin{aligned} &\lambda(0) = \bar{X} \\ &\lambda(1) = X \\ &\frac{\mathrm{d}^2\lambda^I}{\mathrm{d}t^2} + \Gamma^I_{JK}\frac{\mathrm{d}\lambda^J}{\mathrm{d}t}\frac{\mathrm{d}\lambda^K}{\mathrm{d}t} = 0. \end{aligned} \tag{10.4}$$

Instead of specifying the end-points of the geodesic it is a lot simpler to specify the starting point and the initial velocity

$$\dot{\lambda}^I(0) = \xi^I, \tag{10.5}$$

with a judiciously chosen vector ξ. In fact the choice of ξ is such that the length of the geodesic is given by the length of ξ

$$s^2 = G_{IJ}\xi^I\xi^J. \tag{10.6}$$

The advantage of this, seemingly unimportant, change of boundary conditions is that the solution to the geodesics equation can be written as a power series in the *vector* ξ. This ultimately leads to a manifestly covariant expansion of the NLσM action in terms of ξ. In order to derive an explicit expression for the geodesic solution expanded in terms of ξ, one first writes the geodesic solution as a formal power series in the parameter t

$$\lambda^I(t) = \lambda^I(0) + \sum_{k=1}^{\infty}\frac{1}{k!}\left.\frac{\mathrm{d}^k\lambda^I(t)}{\mathrm{d}t^k}\right|_{t=0} t^k. \tag{10.7}$$

The $k = 1$ term of this expansion can be written as

$$\left.\frac{\mathrm{d}\lambda^I(t)}{\mathrm{d}t}\right|_{t=0} t = \xi^I t. \tag{10.8}$$

After plugging this expression into the geodesic equation the second term of the expansion can be expressed in terms of ξ

$$\left.\frac{\mathrm{d}^2\lambda^I}{\mathrm{d}t^2}\right|_{t=0} = -\Gamma^I_{JK}\left.\frac{\mathrm{d}\lambda^J}{\mathrm{d}t}\frac{\mathrm{d}\lambda^K}{\mathrm{d}t}\right|_{t=0} = -\Gamma^I_{JK}(\bar{X})\xi^J\xi^K. \tag{10.9}$$

One can continue this process by differentiating the geodesic equation with respect to the parameter t. This procedure results in the following expansion

$$\lambda^I(t) = \bar{X} + \xi^I t - \sum_{k=2}\frac{1}{k!}\Gamma^I_{J_1\ldots J_k}(\bar{X})\xi^{J_1}\ldots\xi^{J_k}t^k. \tag{10.10}$$

The generalized Christoffel symbols $\Gamma^I_{J_1\ldots J_k}$ are defined as follows

$$\Gamma^I_{J_1\ldots J_k} = \nabla_{J_k}\cdots\nabla_{J_3}\Gamma^I_{J_2 J_1}, \tag{10.11}$$

where the covariant derivatives act solely on the lower indices. The generalized Christoffel symbols are evaluated at the background point $\bar{X}$. With this power series solution it is possible to express π covariantly in terms of the tangent vector ξ

$$\pi^I = \lambda^I(1) - \bar{X}^I = \xi^I - \sum_{k=2}\frac{1}{k!}\Gamma^I_{J_1\ldots J_k}\xi^{J_1}\ldots\xi^{J_k}. \tag{10.12}$$

If this expression is used to expand the action, then the expansion will be a manifestly covariant expansion in terms of the vector ξ.

The vector ξ can also be used as a coordinate. Such a coordinate system will have a base-point, a point of origin, located at $\bar{X}$. Since the length of ξ is equal to the length of the geodesic, geodesics are straight lines in these coordinates (see figure 10.1). Straight lines obey the differential equation

$$\frac{\mathrm{d}^2\lambda^I}{\mathrm{d}t^2} = 0. \tag{10.13}$$

Comparing this with the geodesic equation one learns that, in this coordinate system the Christoffel symbol evaluated at the base-point vanishes

$$\Gamma^I_{JK}(\bar{X}) = 0. \tag{10.14}$$

As a consequence of this the generalized Christoffel symbols vanish too

$$\Gamma^I_{J_1\ldots J_k}(\bar{X}) = 0. \tag{10.15}$$

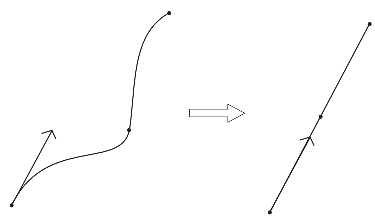


Figure 10.1: *Changing to Riemann Normal Coordinates straightens geodesics.*

A choice of coordinates satisfying equation 10.15 are called Riemann Normal Coordinates (RNC). Such a coordinate system has very convenient properties. For example, since the Christoffel symbol vanishes, partial derivatives are equal to the covariant derivatives. This leads to identities like

$$\xi^{K_1}\xi^{K_2}\partial_{K_1}\partial_{K_2}G_{IJ} = -\tfrac{1}{3}\xi^{K_1}\xi^{K_2}R_{IK_1JK_2}. \qquad (10.16)$$

The background field is also called the base-point of the Riemann Normal Coordinate system. When the term background field is used the fact that $\bar{X}$ is a field over M is emphasized. The term base-point stresses that $\bar{X}$ is also a point in Σ.

To covariantly expand the action in equation 10.1 one can now proceed as follows:

1. Write a Taylor series for the metric and convert "ordinary" derivatives acting on the metric into covariant derivatives acting on (combinations of) the Riemann tensor.
2. Substitute $\pi^I = \xi^I$.
3. Collect terms of the same order in ξ.

This is a rather lengthy and technical procedure, but due to the systematics in this procedure it boils down to remembering a few simple rules [62]. The rules are most easily described using an operator Δ, implicitly defined by

- $\Delta(\dot{X}^I) = \dot{\xi}^I$
- $\Delta(\dot{\xi}^I) = R^I_{JKL}\xi^J\xi^K\dot{X}^L$
- $\Delta(T_{I_1\ldots I_p}) = \xi^M\nabla_M T_{I_1\ldots I_p}$ for a tensor T
- $\Delta(\mathcal{O}_1\mathcal{O}_2) = \mathcal{O}_1\Delta(\mathcal{O}_2) + \Delta(\mathcal{O}_1)\mathcal{O}_2$.

The background field expansion is then obtained by systematically writing out the following expression

$$S = \int \mathrm{d}\tau \; \mathrm{e}^{\Delta} \left(G_{IJ}(X) \dot{X}^I \dot{X}^J \right). \tag{10.17}$$

The first few terms of this expansion are [63]

$$\begin{aligned} S[\bar{X} + \pi] = \int \mathrm{d}\tau \; (G_{IJ}(\bar{X}) &+ \tfrac{1}{3} R_{IKLJ}(\bar{X}) \xi^K \xi^L \\ &+ \tfrac{1}{6} \nabla_M R_{IKLJ}(\bar{X}) \xi^K \xi^L \xi^M) \dot{\xi}^I \dot{\xi}^J. \end{aligned} \tag{10.18}$$

(10.2) Shifting the base-point

From the previous section it is clear that RNC are a useful tool in constructing a manifestly covariant background expansion of the action for the non-linear sigma model. However the split of the full quantum field X in a background (base) and a fluctuation part is an arbitrary one. This means that the action should always be invariant under a change in the background compensated by a change in the fluctuations. In terms of RNC this means that the action should be invariant under a shift of the base-point.

On a formal level this is obvious, expressing the action in terms of a RNC system at $\bar{X}$ with tangent vector ξ

$$S = S[\bar{X}, \xi, G(\bar{X}), R(\bar{X}), \dots] \tag{10.19}$$

or at a different base-point $\bar{Y}$ and tangent ζ

$$S = S[\bar{Y}, \zeta, G(\bar{Y}), R(\bar{Y}), \dots] \tag{10.20}$$

does not matter. The functional form of both expressions is exactly the same (after all, the two expressions are related by simply renaming the variables). However to test the invariance explicitly one needs to relate the two coordinate systems. That is, the tangent vector ξ in the $\bar{X}$ system should be expressed in terms of the $\bar{Y}$ system. The way to relate the two systems is by realizing that both systems should describe the same point in M, that is

$$\bar{X} + \pi = \bar{Y} + \rho, \tag{10.21}$$

with ρ the fluctuation around $\bar{Y}$ and π the fluctuation around $\bar{X}$. Now the relation 10.12 can be used to relate ξ to ζ. After having chosen RNC in $\bar{X}$ one gets

$$\bar{X}^I + \pi^I = \bar{X}^I + \xi^I = \bar{Y}^I + \zeta^I - \sum_{k=2}^{\infty} \Gamma^I_{J_1 \dots J_k}(\bar{Y}) \zeta^{J_1} \dots \zeta^{J_k}. \tag{10.22}$$

If the two base-points are chosen (infitesimally) close to each other, $\bar{X} = \bar{Y} - \epsilon$, one can express ξ in terms of ζ and tensors in the $\bar{X}$ system:

$$\xi^I = \epsilon^I + \zeta^I - \epsilon^L \sum_{k=2}^{\infty} \partial_L \Gamma^I_{J_1 \ldots J_k}(\bar{X}) \zeta^{J_1} \ldots \zeta^{J_k}, \tag{10.23}$$

where

$$\begin{aligned} \Gamma^I_{J_1 \ldots J_k}(\bar{Y}) &= \Gamma^I_{J_1 \ldots J_k}(\bar{X}) + \epsilon^L \partial_L \Gamma^I_{J_1 \ldots J_k}(\bar{X}) + \mathcal{O}(\epsilon^2) \\ &= \epsilon^L \partial_L \Gamma^I_{J_1 \ldots J_k}(\bar{X}) + \mathcal{O}(\epsilon^2) \end{aligned} \tag{10.24}$$

was used. Since the $\partial\Gamma$ expressions are evaluated in RNC it is possible to convert them to manifestly covariant expressions

$$\xi^I = \epsilon^I + \zeta^I - \epsilon^L \sum_{k=2}^{\infty} \tfrac{1}{(k+1)!} R_{J_1(LI)J_2 \ldots J_k}(\bar{X}) \zeta^{J_1} \ldots \zeta^{J_k}, \tag{10.25}$$

with

$$R_{J_1(LI)J_2 \ldots J_k} = \tfrac{1}{k!} \nabla_{(J_k} \ldots \nabla_{J_3} R_{J_2|LI|J_1)} + (L \leftrightarrow I). \tag{10.26}$$

As an example the procedure will be applied to the action for the NLσM, expanded up to first order in the Riemann tensor (and no covariant derivatives)

$$S[\bar{X} + \pi] = \int \mathrm{d}\tau \left(G_{IJ}(\bar{X}) - \tfrac{1}{3} R_{IK_1JK_2}(\bar{X}) \xi^{K_1} \xi^{K_2} \right) \dot{\xi}^I \dot{\xi}^J, \tag{10.27}$$

which after plugging in equation 10.12 becomes

$$\begin{aligned} S = \int \mathrm{d}\tau & \left(G_{IJ}(\bar{Y}) - \tfrac{1}{3} R_{IK_1JK_2}(\bar{Y}) \zeta^{K_1} \zeta^{K_2} \right) \dot{\zeta}^I \dot{\zeta}^J \\ & - \tfrac{2}{3} R_{IK_1JK_2}(\bar{Y}) \epsilon^{K_1} \zeta^{K_2} \dot{\zeta}^I \dot{\zeta}^J - \tfrac{2}{3} R_{IK_1JK_2}(\bar{Y}) \epsilon^J \dot{\zeta}^I \tfrac{\mathrm{d}}{\mathrm{d}\tau} \left(\zeta^{K_1} \zeta^{K_2} \right). \end{aligned} \tag{10.28}$$

Here the fact that, in this approximation, $g(\bar{X}) = g(\bar{Y})$ and $R(\bar{X}) = R(\bar{Y})$ is used. The last two terms cancel, leaving the expression

$$S = \int \mathrm{d}\tau \left(G_{IJ}(\bar{Y}) - \tfrac{1}{3} R_{IK_1JK_2}(\bar{Y}) \zeta^{K_1} \zeta^{K_2} \right) \dot{\zeta}^I \dot{\zeta}^J \tag{10.29}$$

which is precisely the action expressed in RNC around $\bar{Y}$. Hence

$$S[\bar{X} + \pi] = S[\bar{Y} + \rho] \tag{10.30}$$

thereby proving base-point independence for the NLσM in this approximation (the general proof is similar but more technical).

(10.3) A NEW SYMMETRY PRINCIPLE

Since the base-point is essentially a point in d-dimensional space, the base-point shift imposes a kind of d-dimensional diff on the action for D0-branes. Although this diff cannot be lifted to the DIFF (see chapter 9) it is possible to lift base-point shifts in the d-dimensional space to base-point shifts in D-geometry. To see this, look at the base-point transformation equation 10.25. If one chooses the base-point such that all the D0-branes are on top of each other

$$\bar{X} = \bar{x}\mathbb{I} \tag{10.31}$$

and restricts the shifts to be diagonal as well

$$\epsilon^I = \epsilon^i \delta^{ab}, \tag{10.32}$$

then all tensors evaluated at the base-point decompose in a tensor over the d-dimensional space and a tensor in the matrix indices:

$$R_{IJKL}(\bar{x}\mathbb{I}) = R_{ijkl}(\bar{x})T_{a_ib_ia_jb_ja_kb_ka_lb_l}. \tag{10.33}$$

This leaves the tensor T undetermined. These matrix-tensors can be determined by imposing base-point invariance as a self-consistency condition, together with the D-geometry axioms. In [55] the following results were found

$$R_{IJKL} = R_{ia_ib_ija_jb_jka_kb_kla_lb_l} = R_{ijkl}\Delta_{Sym(a_ib_ia_jb_ja_kb_ka_lb_l)}, \tag{10.34}$$

with the tensor Δ defined by

$$\Delta_{a_ib_ia_jb_ja_kb_ka_lb_l} = \delta_{b_ia_j}\delta_{b_ja_k}\delta_{b_ka_l}\delta_{b_la_i}. \tag{10.35}$$

The tensor Δ basically takes for four matrices, multiplies them and takes the trace:

$$\Delta_{a_ib_ia_jb_ja_kb_ka_lb_l}X^{ia_ib_i}X^{ja_jb_j}X^{ka_kb_k}X^{la_lb_l} = \mathrm{Tr}\,X^iX^jX^kX^l. \tag{10.36}$$

From this it is now easy to see that equation 10.34 defines a completely symmetric tensor

$$R_{IJKL}X^IX^JX^KX^L = R_{ijkl}\mathrm{Str}\,X^iX^jX^kX^l. \tag{10.37}$$

The symbol Str is defined as the symmetrized trace, i.e.

$$\mathrm{Str}\,ABC = \tfrac{1}{6}\mathrm{Tr}\,(ABC + ACB + \dots). \tag{10.38}$$

The symmetrized trace prescription remains valid for the ∇R tensor

$$\begin{aligned}\nabla_M R_{IJKL} &= \nabla_{ma_mb_m}R_{ia_ib_ija_jb_jka_kb_kla_lb_l}\\ &= \nabla_m R_{ijkl}\Delta_{Sym(a_ib_ia_jb_ja_kb_ka_lb_laa_mb_m)}.\end{aligned} \tag{10.39}$$

Beyond this order (i.e. tensor of type $\nabla^n R, \quad n > 1$) the symmetrized trace prescription is no longer valid [55].

To summarize: Implementing base-point invariance as described above, lifts a symmetry from the d-dimensional space to the dN^2-dimensional space. Since the base-point symmetry in the d-dimensional space is equivalent to diff invariance [55] one can take the base-point invariance in D-geometry as a definition of diff invariance in D-geometry.

In contrast to the point-particle in a curved background, this symmetry principle does not determine the action of D0-branes in a curved background uniquely. This is reflected by the fact that the expressions for $\nabla_{M_1} \cdots \nabla_{M_n} R_{IJKL}$ for $n > 1$ are not fixed by the D-geometry axioms and base-point independence [55].

Chapter 11

Applications and checks

(11.1) Applying base-point independence

After having introduced the concept of base-point independence it is time to apply the technique to D-geometry. The basis of this D-geometry is the realization that the final action must be a constrained form of a dN^2-dimensional non-linear sigma model (NLσM).

$$S = \int d\tau \, G_{IJ}(X) \dot{X}^I \dot{X}^J \ . \tag{11.1}$$

Each index I describes a triplet $I = \{i; ab\}$ built from a d-dimensional space-time index i and two $U(N)$-indices a, b. The dN^2-dimensional metric $G_{IJ}(g_{ij}, R_{ijkl}, \ldots, X)$ is a *functional* of the d-dimensional metric g_{ij} and its derivatives. Using a dN^2-dimensional expansion in normal coordinates, one must impose the following conditions to obtain the d-dimensional base-point independent action for matrix normal coordinates [55]:

(a) When functionally expressed in terms of the d-dimensional constituents the dN^2 dimensional metric, Riemann tensor, and covariant derivatives thereof must obey *all* the usual identities of symmetry/antisymmetry, Bianchi identity, commutation relations of covariant derivatives, etc.

(b) The $U(N)$ indices must be such that the action is a single trace; i.e. no traces may occur within the functional expressions.

(c) It should have the right $U(1)$ limit for diagonal matrices.

(d) At linearized order the symmetrized ordering should emerge.

(e) Most importantly, *base-point independence* follows from the requirement that the "trace" of the dN^2-dimensional covariant derivative, acts as the d-dimensional covariant derivative:

$$\delta^{ab}\nabla_{i;ab}(\text{anything}) = \nabla_i(\text{anything}) \ . \tag{11.2}$$

For instance at order 4 and 5 in matrix normal coordinates $X^{i;ab}$ the two relevant dN^2-dimensional tensors are the Riemann tensor and its covariant derivative. Imposing the above matrix-geometry constraints one finds that in terms of d-dimensional curvature tensors, they are

$$\begin{aligned} R_{IJKL} &= R_{ijkl}\Sigma_{a_ib_ia_jb_ja_kb_ka_lb_l} \ , \\ \nabla_M R_{IJKL} &= \nabla_m R_{ijkl}\Sigma_{a_mb_ma_ib_ia_jb_ja_kb_ka_lb_l} \ , \end{aligned} \tag{11.3}$$

where $\Sigma_{a_1b_1\ldots a_nb_n}$ is the object that when contracted with n-matrices returns the symmetrized trace

$$\Sigma_{a_1b_1,\ldots,a_nb_n}O^{i_1a_1b_1}\cdots O^{i_na_nb_n} = \mathrm{Str}(O^{i_1}\cdots O^{i_n}) \ . \tag{11.4}$$

At order six, the fully symmetrized ordering is no longer consistent with the identity

$$[\nabla_N,\nabla_M]R_{IJKL} = R_{NMI}{}^{P}R_{PJKL} + \ldots \tag{11.5}$$

This illustrates why the symmetrized approximation corresponds to the linearized approximation.

(11.1.1) SECOND ORDER IN $\dot{X}$

To explicitly show how the matrix-geometry generates a base-point independent action, the application for the kinetic term — order two in derivatives — to order $\mathcal{O}(X^4)$ will be reviewed, i.e. it is shown that the action

$$L_2 = -\tfrac{1}{2}(\delta_{ij}\mathrm{tr}(\dot{X}^i\dot{X}^j) + \tfrac{1}{3}\mathrm{R}_{iklj}\mathrm{Str}(\dot{X}^i\dot{X}^jX^kX^l)) + \mathcal{O}(X^5) \tag{11.6}$$

is base-point independent. Writing the action in terms of a dN^2 NLσM

$$L = -\tfrac{1}{2}\eta_{AB}\Pi^A\Pi^B, \tag{11.7}$$

with

$$\Pi^A = E_I^A\dot{X}^I. \tag{11.8}$$

Capital letters refer to a multi-index notation in which a matrix Π^i is represented as $\Pi^I = \Pi^{i\alpha\beta}$. Sometimes it is easier to work in a local-Lorentz frame, in this case the matrix is written as: $\Pi^A = \Pi^{\imath ab}$. The vielbein relating the Π^I and the Π^A,

$$\Pi^A = E_I^A\Pi^I \Rightarrow \Pi^{\imath ab} = \sum_{i\alpha\beta} E_{i\alpha\beta}^{\imath ab}\Pi^{i\alpha\beta}, \tag{11.9}$$

has a convenient flat space representation:

$$\eta_{AB} = \eta_{\imath ab,\jmath cd} = \delta_{\imath\jmath}\delta_{ad}\delta_{bc}. \tag{11.10}$$

The flat metric is defined in such a way that:

$$\mathrm{tr}(X^i X^j)\delta_{ij} = \eta_{IJ}X^I X^J \Rightarrow \eta_{IJ} = \eta_{i\alpha\beta,j\gamma\delta} = \delta_{ij}\delta_{\alpha\delta}\delta_{\beta\gamma}. \tag{11.11}$$

The metric for curved space is then given by:

$$G_{IJ} = E_I^A E_J^B \eta_{AB} \tag{11.12}$$

Expanding in RNC the vielbein equals:

$$E_I^A = \delta_I^A + \frac{1}{12}R^A{}_{(PQ)I}X^P X^Q + \tfrac{1}{24}\nabla_P R^A{}_{(QR)I}X^P X^Q X^R + \dots \tag{11.13}$$

Substituting this into equation 11.7, and using 11.3, equation 11.6 is recovered. By virtue of the fact that eq. 11.3 is the solution to the matrix-geometry constraints, this action is basepoint independent. For this simple case, one can check it explicitly [55].

An instructive illustration of the power of the matrix-geometry method, is the following exercise. Although from the flat space limit it is clear that the one-form Π^A should be contracted with the tangent space metric η_{AB} of eq. 11.10, it is possible to start off with a more general form

$$L_2 = -\tfrac{1}{2}M_{AB}(X)\Pi^A\Pi^B. \tag{11.14}$$

Expanding in RNC as prescribed one obtains the action (with the base-point $\bar{X}$):

$$\begin{aligned} -2L_2 = \big\{ & M_{AB}|_{\bar{X}} + \nabla_C M_{AB}|_{\bar{X}}\, X^C \\ & + \tfrac{1}{2}\left(\tfrac{2}{3}M_{QB}\mathrm{R}^Q{}_{CDA} + \nabla_D\nabla_C M_{AB}\right)\Big|_{\bar{X}} X^C X^D + \dots\Big\}\, \dot{X}^A\dot{X}^B. \end{aligned} \tag{11.15}$$

Comparing with the flat space case one reads off: $M_{AB}|_{\bar{X}} = \eta_{AB}$. Assume that it is possible to set $\nabla\dots\nabla\ M|_{\bar{X}} = 0$, then one has $M(X)_{AB} = \eta_{AB}$ (remember: in RNC partial and covariant derivatives are the same). This results in the action:

$$\begin{aligned} -2L_2 &= (\eta_{AB} + \tfrac{1}{6}\mathrm{R}_{B(CD)A}X^C X^D)\dot{X}^A\dot{X}^B \\ &= \eta_{AB}\Pi^A\Pi^B. \end{aligned} \tag{11.16}$$

Given that the vielbein is constructed from tensors obeying the matrix-geometry constraints, the properties the action needs to have can be checked without explicit calculations. The only thing that needs to be verified is that M_{AB} also satisfies the matrix-geometry constraints. Single traceness of the action follows from the fact that $M_{AB} = \eta_{AB}$ has no internal $U(N)$ contractions. The correct $U(1)$ limit follows from the vanishing of all the covariant derivatives $\nabla \ldots \nabla M = \nabla \ldots \nabla \eta = 0$. The crucial property to check is base-point independence. Note that, since the expansion of M_{AB} is $SO(dN^2)$ covariant, it is manifest that the action is invariant under any matrix valued diffeomorphism. However, the tensor M_{AB} should also be a functional of the d-dimensional metric, and its derivatives. This functional will be consistent with base-point independence, if under a shift in base-point, it is parallel transported in the d-dimensional sense. As before, this is guaranteed if

$$\epsilon^k \delta^{ab} \nabla_{kab} = \epsilon^k \nabla_k, \tag{11.17}$$

on the tensor M_{AB}. For $M_{AB} = \eta_{AB}$ this is obviously so, so that the result is indeed base-point independent. Note that this is a truly non-trivial constraint that does not follow from the fact that the Lagrangian L_2 is a scalar quantity under matrix valued diffeomorphisms.

Focussing on M_{AB} alone, these results can be extended to arbitrary order in X; at all orders $\nabla \ldots \nabla M$ can be set to vanish, in a way consistent with the matrix-geometry constraints. In particular, the non-trivial identity

$$\begin{aligned} \nabla_{[C}\nabla_{D]} M_{AB}\big|_{\bar{X}} = \mathrm{R}_{CDA}{}^{Q}\, M_{QB}|_{\bar{X}} + \mathrm{R}_{CDB}{}^{Q}\, M_{QA}|_{\bar{X}} \\ = \mathrm{R}_{CDBA} + \mathrm{R}_{CDAB} = 0. \end{aligned} \tag{11.18}$$

and higher order analogues are satisfied. For these, it is crucial that M_{AB} evaluated at the base-point is equal to the tangent space metric. In section 11.1.2, where the consistency of the base-point independence approach for higher derivative terms is discussed, it will be shown that these non-trivial identities do impose constraints. Thus the two-derivative action, equation 11.7, is recovered here. The lesson is that base-point independence and the other constraints of matrix geometry are automatically satisfied when $\nabla \ldots \nabla M$ can be put to zero. The power of the above argument, by generalization, is that the expansion of the action in RNC allows to, almost, read off whether the action is base-point independent and therefore a candidate for the non-Abelian generalization of the DBI-action.

(11.1.2) Higher order corrections: Fourth order in $\dot{X}$

In section 11.1.1 the covariant expansion approach was used as a straightforward approach to determine the crucial properties of the action. Therefore this route will be followed for the action L_4, the part of the DBI-action of fourth order in $\dot{X}$, as well. Before doing that, it will be instructive to show explicitly why the symmetrized trace approximation starts to fail at this order. If the action would be the symmetrized trace, then it could be written as:

$$L_4 = -\tfrac{1}{8}\mathrm{Str}\,(g_{ij}(X)g_{kl}(X)\dot{X}^i\dot{X}^j\dot{X}^k\dot{X}^l)$$
$$= -\tfrac{1}{8}(\delta_{ij}\delta_{kl}\mathrm{Str}\,(\dot{X}^i\dot{X}^j\dot{X}^k\dot{X}^l) + \tfrac{2}{3}\mathrm{R}_{imnj}\delta_{kl}\mathrm{Str}\,(X^m X^n \dot{X}^i\dot{X}^j\dot{X}^k\dot{X}^l)) + \ldots . \tag{11.19}$$

Using the base-point transformation (a constant shift ϵ)

$$\Delta X_i = \epsilon_i + \tfrac{1}{6}\epsilon^k \mathrm{Sym}\,(Z^{p_1}Z^{p_2})\mathrm{R}_{p_1(ki)p_2}, \tag{11.20}$$

one gets for the variation (schematically and up to first order in the Riemann tensor)

$$\Delta L_4 \propto \epsilon \mathrm{RStr}\,(\dot{Z}^3\mathrm{Sym}\,(\dot{Z}Z)) + \epsilon\mathrm{RStr}\,(\dot{Z}^4 Z) \neq 0. \tag{11.21}$$

From this it is obvious that the two combinatorial structures cannot be combined, hence the symmetrized trace prescription does not yield a base-point independent action.

To determine what the correct ordering is, the same steps are followed as before. The starting point is an action of order four in one-forms Π^A contracted with an arbitrary symmetric four-tensor $M_{ABCD}(X)$:

$$L_4 = -\tfrac{1}{8}M_{ABCD}(X)\Pi^A\Pi^B\Pi^C\Pi^D\ . \tag{11.22}$$

Expanding in RNC up to second order in X, one finds

$$-8L_4 = \Big\{ M_{ABCD}|_{\bar{X}} + \nabla_F M_{ABCD}|_{\bar{X}}\, X^F$$
$$+ \tfrac{1}{2}\left(\tfrac{4}{3}M_{QBCD}\mathrm{R}^Q{}_{EFA} + \nabla_E\nabla_F M_{ABCD}\right)\Big|_{\bar{X}} X^E X^F \Big\}\dot{X}^A\dot{X}^B\dot{X}^C\dot{X}^D. \tag{11.23}$$

The lesson from section 11.1.1 is to check whether or not one can put $\nabla\ldots\nabla\ M|_{\bar{X}} = 0$. This would ensure the correct $U(1)$ limit and base-point independence. Suppose it is possible. In that case the action would be:

$$-8L_4 = \Big\{ M_{ABCD}|_{\bar{X}} + \tfrac{2}{3}M_{QBCD}\mathrm{R}^Q{}_{EFA}\Big|_{\bar{X}} X^E X^F\Big\}\dot{X}^A\dot{X}^B\dot{X}^C\dot{X}^D. \tag{11.24}$$

From the flat space limit, [1]

$$-8L_4^{flat} = M_{ABCD}|_{\bar{X}}\,\dot{X}^A\dot{X}^B\dot{X}^C\dot{X}^D = \mathrm{Str}\,(\dot{X}^i\dot{X}^j\dot{X}^k\dot{X}^l)\delta_{ij}\delta_{kl}, \tag{11.25}$$

one learns that:

$$M_{ABCD}|_{\bar{X}} \equiv \eta_{ABCD} = \delta_{ab}\delta_{cd}\Sigma_{\alpha_1\beta_1\ldots\alpha_4\beta_4} \neq \eta_{AB}\eta_{CD}, \tag{11.26}$$

[1] The flat space limit follows indirectly from explicit string computations which show that the non-abelian DBI-action at order F^4 is given by the symmetrized trace [64, 65, 66, 67]. This breaks down at higher orders; see [68] for the latest status.

Note that the single trace requirement means that the tensor M is not proportional to two tangent metrics. This will be important.

The crucial question is: Are all the requirements (a)-(e) of matrix geometry satisfied? Since the assumption was that all covariant derivatives on M can be put to zero, the $\mathrm{U}(1)$ limit is automatically right. The single trace condition is met, by construction. What about the base-point independence? This is also guaranteed if one can truly put all covariant derivatives on M to zero (evaluated in the base-point). It therefore remains to test this assumption. Because M is not the tensor product of the metric tensor, it is not clear that a covariantly constant M_{ABCD} exists. The litmus tests are the identities involving commutators of covariant derivatives on M_{ABCD}; the simplest being

$$\nabla_{[E}\nabla_{F]}\, M_{ABCD}|_{\bar{X}} = 2(\mathrm{R}_{EFA}{}^{Q} M_{QBCD} + \mathrm{R}_{EFB}{}^{Q} M_{QACD} + \mathrm{R}_{EFC}{}^{Q} M_{QABD} + \mathrm{R}_{EFD}{}^{Q} M_{QABC})\Big|_{\bar{X}} . \tag{11.27}$$

This directly shows that a covariantly constant M, not proportional to the metric, is not consistent since the right-hand side of the equation will not evaluate to zero. Recall that in the two derivative case, this was explicitly tested. There M_{AB} was equal to the metric, and this problem did not occur since

$$\nabla_{[C}\nabla_{D]} M_{AB}\big|_{\bar{X}} = \mathrm{R}_{CDA}{}^{Q}\; M_{QB}|_{\bar{X}} + \mathrm{R}_{CDB}{}^{Q}\; M_{QA}|_{\bar{X}} = \mathrm{R}_{CDBA} + \mathrm{R}_{CDAB} = 0. \tag{11.28}$$

The mathematical cause for the inconsistency of a covariantly constant M_{ABCD} is the single trace requirement. It was responsible for the fact that M_{ABCD} could not equal two tangent space metrics. One might object that the identity 11.28 appears to be irrelevant as in the action $\nabla_E \nabla_F M_{ABCD}$ is contracted with the symmetric combination $X^E X^F$. The remaining discussion will show why this is not so.

Lacking a deeper insight in the requirements on M_{ABCD}, one is forced to check base-point independence by hand. Recall that the base-point transformation is given by:

$$\Delta X_i = \epsilon_i + \tfrac{1}{6}\epsilon^k \mathrm{Sym}\,(Z^{p_1} Z^{p_2}) \mathrm{R}_{p_1(ki)p_2}. \tag{11.29}$$

Using this transformation calculate the variation of L_4 with $M_{ABCD} \equiv \eta_{ABCD}$:

$$-8L_4 = \eta_{ABCD} E^A_I \dot{X}^I E^B_J \dot{X}^J E^C_K \dot{X}^K E^D_L \dot{X}^L = \delta_{ij}\delta_{kl} \mathrm{Str}\,(\dot{X}^i \dot{X}^j \dot{X}^k \dot{X}^l) + \tfrac{1}{3}\mathrm{R}_{i(kl)j}\delta_{mn}\mathrm{Str}\,(\dot{X}^i \dot{X}^m \dot{X}^n \mathrm{Sym}\,(\dot{X}^j X^k X^l)), \tag{11.30}$$

which reads

$$\begin{aligned} -\Delta L_4 &= \tfrac{1}{8}\tfrac{4\cdot 2}{3\cdot 2}\epsilon^k \mathrm{R}_{\beta(k\alpha)p_2} \mathrm{Str}\,(\dot{Z}^\alpha \dot{Z}^2 \mathrm{Sym}\,(\dot{Z}^\beta Z^{p_2})) \\ &\quad + \tfrac{2}{3\cdot 8}\mathrm{R}_{k(\alpha\beta)p_2}\epsilon^k \mathrm{Str}\,(\dot{Z}^\alpha \dot{Z}^2 \mathrm{Sym}\,(\dot{Z}^\beta Z^{p_2})) \\ &= \tfrac{1}{6}\epsilon^k \mathrm{Str}\,(\dot{Z}^\alpha \dot{Z}^2 \mathrm{Sym}\,(\dot{Z}^\beta Z^{p_2}))\left\{\mathrm{R}_{\beta(k\alpha)p_2} + \tfrac{1}{2}\mathrm{R}_{k(\alpha\beta)p_2}\right\}. \end{aligned} \tag{11.31}$$

Note that $\mathrm{Sym}\,(\ldots)$ expressions are treated as one block within the symmetrized trace. Using (note $A_{(ab)} = A_{ab} + A_{ba}$),

$$\mathrm{R}_{\alpha(p_1\beta)p_2} = -\tfrac{1}{2}\mathrm{R}_{p_1(\alpha\beta)p_2} + \tfrac{3}{2}\mathrm{R}_{\alpha\beta p_1 p_2}, \tag{11.32}$$

this simplifies to (with the notation $A_{p_1\ldots p_n} = A_{1\ldots n}$):

$$\Delta L_4 = -\tfrac{1}{4}\epsilon^6 \mathrm{R}_{3456}\delta_{12}\mathrm{Str}\,(\dot{Z}^1\dot{Z}^2\dot{Z}^3\mathrm{Sym}\,(\dot{Z}^4 Z^5)) \neq 0. \tag{11.33}$$

So the proposed action is indeed not base-point independent. To see that the single trace requirement is responsible — and hence the correlated inconsistency of choosing $M_{ABCD}(X) = \eta_{ABCD}$ — look at the corresponding calculation for L_2. The steps are analogous and the L_2 result comes down to removing the $\delta_{12}\dot{Z}^1\dot{Z}^2$ part in the L_4 result and adjusting some factors:

$$\Delta L_2 \propto \epsilon^6 \mathrm{R}_{3456}\mathrm{Str}\,(\dot{Z}^3\mathrm{Sym}\,(\dot{Z}^4 Z^5)). \tag{11.34}$$

Using the identity

$$\mathrm{Str}\,(ABCD\ldots) = \mathrm{Tr}\,(A\mathrm{Sym}\,(BCD\ldots)). \tag{11.35}$$

makes it possible to write ΔL_2 as

$$\Delta L_2 = \epsilon^6 \mathrm{R}_{3456}\mathrm{Str}\,(\dot{Z}^3\dot{Z}^4 Z^5), \tag{11.36}$$

which is obviously zero. In the case of ΔL_4 the identity in equation 11.35 is of no use, because of the extra $\delta_{12}\dot{Z}^1\dot{Z}^2$ factor. If one had not insisted one a single trace result and used $M_{ABCD} = \eta_{AB}\eta_{CD}$, then the variation of ΔL_4 would have had the structure:

$$\Delta L_4 = \Delta L_2 \delta_{12}\mathrm{Str}\,(\dot{Z}^1\dot{Z}^2) = 0. \tag{11.37}$$

So, as claimed, the single trace property spoils base-point independence. As a result of this, the importance of the identity (11.28) is confirmed. $\nabla_E\nabla_F\, M_{ABCD}|_{\bar{X}}$ should not be zero if one insists on base-point independence.

Fortunately is it not terribly difficult to find an correction term to L_4 that renders the action base-point independent while keeping the correct $\mathrm{U}(1)$ limit. One possible answer is:

$$L_4^C = \alpha \mathrm{R}_{1356}\delta_{24}\mathrm{Str}\,(\dot{X}^1\dot{X}^2\mathrm{Sym}\,(\dot{X}^4 X^5)\mathrm{Sym}\,(\dot{X}^3 X^6)). \tag{11.38}$$

Since R_{1356} is antisymmetric in 1 and 3, this result vanishes in the $U(1)$ limit. The variation of L_4^C equals

$$\begin{aligned}\Delta L_4^C &= \alpha \mathrm{R}_{1356}\delta_{24}\mathrm{Str}\,(\dot{Z}^1\dot{Z}^2\dot{Z}^4\epsilon^5\mathrm{Sym}\,(\dot{Z}^3 Z^6))\\ &+ \alpha \mathrm{R}_{1356}\delta_{24}\mathrm{Str}\,(\dot{Z}^1\dot{Z}^2\mathrm{Sym}\,(\dot{Z}^4 Z^5)\dot{Z}^3\epsilon^6)\\ &= \alpha\epsilon^6\mathrm{R}_{1465}\delta_{24}\mathrm{Str}\,(\dot{Z}^1\dot{Z}^2\dot{Z}^3\mathrm{Sym}\,(\dot{Z}^4 Z^5))\\ &= 4\alpha\Delta L_4.\end{aligned} \tag{11.39}$$

Requiring base-point independence determines the constant α to be $-\frac{1}{4}$.

So it is possible to find a correction term L_4^C that renders the total action base-point independent. However this correction term was not written in the form of a tensor M_{ABCD}. Therefore the obvious question is: What is the value for $\nabla \ldots \nabla\ M_{ABCD}|_{\bar{X}}$ that corresponds to L_4^C? Since there is no $\mathcal{O}(X)$ term in the U(1) limit, it seems logical to try to maintain this property for the non-abelian case: require that there is no $\nabla_F\ M_{ABCD}|_{\bar{X}}\, X^F$ term in the action. For the $\nabla_E\nabla_F\ M_{ABCD}|_{\bar{X}}\, X^E X^F$ part, one can only determine the part symmetric in EF and $ABCD$ from L_4^C. It is determined by:

$$\begin{aligned}&\nabla_E\nabla_F\ M_{ABCD}|_{\bar{X}}\, X^E X^F \dot{X}^A\dot{X}^B\dot{X}^C\dot{X}^D\\ &\qquad = 4\mathrm{R}_{1456}\delta_{23}\mathrm{Str}\,(\dot{X}^1\dot{X}^2\mathrm{Sym}\,(\dot{X}^4 X^5)\mathrm{Sym}\,(\dot{X}^3 X^6))\end{aligned} \tag{11.40}$$

Write this schematically as

$$\begin{aligned}&\nabla_E\nabla_F\ M_{ABCD}|_{\bar{X}}\, X^E X^F \dot{X}^A\dot{X}^B\dot{X}^C\dot{X}^D\\ &\qquad = \tfrac{1}{48}s_{(ABCD)(EF)}X^E X^F \dot{X}^A\dot{X}^B\dot{X}^C\dot{X}^D\end{aligned} \tag{11.41}$$

with

$$\begin{aligned}s_{A_1A_2A_3A_4A_5A_6} &= 4\mathrm{R}_{i_1i_4i_5i_6}\delta_{i_2i_3}\Sigma_{a_1b_1a_2b_2a_7b_7a_8b_8}S^{a_7b_7}_{a_4b_4a_5b_5}S^{a_8b_8}_{a_3b_3a_6b_6}\\ A_j &= i_j a_j b_j.\end{aligned} \tag{11.42}$$

The object S takes a set of matrices and combines them (symmetrically) into one matrix:

$$S^{a_1b_1}_{a_2b_2\ldots a_nb_n}O^{i_2a_2b_2}\cdots O^{i_na_nb_n} = \mathrm{Sym}(O^{i_2}\cdots O^{i_n})^{a_1b_1}. \tag{11.43}$$

As already shown explicitly, the action thus obtained is base-point independent. However, consistency of the more general approach demands that base-point independence can also be shown by proving the following:

$$\epsilon^i\delta^{ab}(\nabla_{iab}\nabla_F\ M_{ABCD}|_{\bar{X}}) = \epsilon^i(\nabla_i\nabla_F\ M_{ABCD}|_{\bar{X}}). \tag{11.44}$$

In order to check this, one has to know what the right-hand part of the previous equation is. It corresponds to the variation of the $\nabla_F\, M_{ABCD}|_{\bar{X}}\, X^F$ term. Since this term vanishes, one concludes that the right-hand side of equation 11.44 should be zero in the action. Recall that $\nabla_E \nabla_F\, M_{ABCD}|_{\bar{X}}$ can receive several contributions, of which only the S part contributes to the action. Of course, it would be the more satisfactory if it vanishes as a single formal tensor. Using the explicit representation for s_{ABCDEF} (equation 11.42), we find for the left-hand side of eq. (11.44):

$$\begin{aligned}
&\epsilon^i \delta^{ab} s_{1234\, iab\, 6} = \epsilon^i \delta^{ab} (\mathrm{R}_{14i6} \delta_{23} \Sigma_{1278} S^7_{4\, ab} S^8_{36} + \ldots) \\
&= \epsilon^i (\mathrm{R}_{14i6} \delta_{23} \Sigma_{1248} S^8_{36} + \ldots) \\
&\propto \epsilon^i (\mathrm{R}_{14i6} \delta_{23} \Sigma_{1248} S^8_{36} + \epsilon^i \mathrm{R}_{41i6} \delta_{23} \Sigma_{1248} S^8_{36} + \ldots) = 0.
\end{aligned} \tag{11.45}$$

This confirms the base-point independence. The other contributions to $\nabla\nabla M$, such as the parts anti-symmetric in E, F, are of no concern since these vanish in the action. It is, however, possible to add other corrections terms to M such that equation 11.44 is zero, even as a formal tensor.

Presumably these results can be generalized to arbitrary order in the velocity $\dot{X}$. However, the extension to higher order in X is far from trivial due to the complications from the inability to put $\nabla \ldots \nabla M$ to zero in matrix geometry. For every order in X it is necessary to find the appropriate tensor $\nabla\nabla \ldots \nabla\, M|_{\bar{X}}$, which is beyond present capabilities. The fact that it is possible to do so for fourth order in derivatives does give confidence that this is possible in general. This computation does explicitly show that, at the first corrective order in derivatives, the symmetrized trace approximation is no longer consistent with the single trace requirement in the presence of gravity.

(11.2) Checking T-duality

One of the consequences of the non-abelian nature of the coordinates X is that new terms can be present in the action, proportional to commutators, which have no $U(1)$ equivalent. Indeed for non-abelian D0-branes in flat space, string theory requires that at lowest order in derivatives there is a potential equal to

$$V = -\frac{T\lambda^2}{2} \mathrm{Tr}([X^i, X^j][X_i, X^j]) \; . \tag{11.46}$$

The form of the potential is dictated by consistency with T-duality. Under this stringy symmetry the potential and kinetic terms are exchanged. T-duality holds for any spacetime with isometries, and the curved space analogues of both the potential and the kinetic term should be consistent with the duality. In addition to constructing a base-point independent kinetic term, in [55] a potential for D0-branes in a curved background was also put forward. This conjectured potential passed a strong consistency test. It satisfied the non-trivial D-geometry constraint that fluctuations around a diagonal background have masses proportional to geodesic lengths. In this section it will be shown that the conjectured form of the potential reproduces the kinetic term after a T-duality transformation. This is strong confirmation that the proposed form of the potential is correct.

To generalize the expression 11.46 to curved space, an analogue of the vector $\dot{X}^I$, which can be contracted with the vielbein E_I^A, is needed. Define the "commutation" operator

$$\mathcal{D}(X)^{icb;ad} \equiv \delta^{ab} X^{icd} - \delta^{cd} X^{iab} \ . \tag{11.47}$$

Acting on a matrix M_{da} it returns the commutator

$$\mathcal{D}(X)^{icb;ad} M_{da} = [X^i, M]^{cb} \ . \tag{11.48}$$

Commutators obey the Leibniz rule and act as a derivation on the space of matrices. Analogous to the standard time derivative $\dot{X}$, one can expect any X^{iab} appearing inside a commutator to transform as a vector under matrix coordinate (i.e. base-point) transformations. The matrix-valued "commutation operator" $\mathcal{D}(X)^{I,ad}$ can therefore be pushed forward to the tangent space with the $SO(dN^2)$ vielbein. The commutator term is the basic building block of the flat space potential. A small calculation shows that it is equivalent to four building blocks, $\mathcal{D}(X)^{I,ad}$, contracted with exactly twice the $SO(dN^2)$ metric [55].

$$\begin{aligned} V_{flat} &= -\frac{T\lambda^2}{4}\eta_{IK}\eta_{JL}\mathcal{D}(X)^{I,ab}\mathcal{D}(X)^{J,bc}\mathcal{D}(X)^{K,cd}\mathcal{D}(X)^{L,da} \\ &= -\frac{T\lambda^2}{4}\eta_{IK}\eta_{JL}\mathrm{Tr}\mathcal{D}(X)^I\mathcal{D}(X)^J\mathcal{D}(X)^K\mathcal{D}(X)^L \ . \end{aligned} \tag{11.49}$$

In the last line the trace is only over the explicit $U(N)$ indices of the matrix valued $SO(dN^2)$ vector $\mathcal{D}(X)^{I,ab}$.

The generalization to curved space is now straightforward. Simply insert the appropriate number of vielbeins into the flat space potential:

$$\begin{aligned} V_{curved} &= -\frac{T\lambda^2}{4}\eta_{AC}\eta_{BD}E_I^A E_J^B E_K^C E_L^D \mathrm{Tr}\,\mathcal{D}(X)^I\mathcal{D}(X)^J\mathcal{D}(X)^K\mathcal{D}(X)^L \\ &= -\frac{T\lambda^2}{4}(\eta_{AC}\eta_{BD} + \frac{4}{12}\mathrm{R}_{C(PQ)A}X^P X^Q \eta_{BD})\mathrm{Tr}\,\mathcal{D}(X)^A\mathcal{D}(X)^B\mathcal{D}(X)^C\mathcal{D}(X)^D \\ &+ \mathcal{O}(\nabla^2) \\ &= -\frac{T\lambda^2}{4}(\mathrm{Tr}\,([X^i, X^k][X^j, X^l])\delta_{ij}\delta_{kl} \\ &+ \frac{1}{3}\mathrm{R}_{i(kl)j}\delta_{mn}\mathrm{Str}\,(X^k X^l[X^i, X^m][X^j, X^n])) + \ldots, \end{aligned} \tag{11.50}$$

The last two steps shows that in the linearized approximation it correctly reproduces the symmetrized result from [69], as is expected.

Under T-duality the parallel part of the curved space potential 11.50 must transform into the kinetic term 11.7. To check this, assume that the d-dimensional geometry is a product of a $(d-1)$-dimensional piece times a circle along the direction $i = 9$. The following expression for the dN^2 metric can then be written down

$$G_{i\alpha\beta,j\gamma\delta}(X^i) = \begin{pmatrix} G_{\mu\alpha\beta,\nu\gamma\delta}(X^\rho) & 0 \\ 0 & \delta_{\alpha\delta}\delta_{\beta\gamma} \end{pmatrix} \qquad \mu,\nu,\rho = 1,\ldots,d-1. \tag{11.51}$$

The expression in the lower right corner is a simply a consequence of the non-trivial form of the flat-space metric:

$$G^{flat}_{i\alpha\beta,j\gamma\delta} = \eta_{i\alpha\beta,j\gamma\delta} = \delta_{ij}\delta_{\alpha\delta}\delta_{\beta\gamma}\ . \tag{11.52}$$

Because the space-time is chosen to be a direct product form, the tangent space decomposes trivially:

$$\eta^{d-dim}_{a\alpha\beta,b\gamma\delta} = \begin{pmatrix} \eta^{(d-1)-dim}_{m\alpha\beta,n\gamma\delta} & 0 \\ 0 & \delta_{\alpha\delta}\delta_{\beta\gamma} \end{pmatrix}. \tag{11.53}$$

In particular, the component of the vielbein E_I^A along the circle is

$$E^{a;\epsilon\zeta}_{9\alpha\beta} = \delta^a_9\delta^\epsilon_\alpha\delta^\zeta_\beta\ . \tag{11.54}$$

In this background the potential splits into three parts. One has no tangent-vectors lying along the isometry direction; this will become the potential in the T-dual case. A second term has all components along the circle: since the flat limit corresponds to the commutator squared, this term will vanish. The crossterm with half the components along the circle is the interesting part. Using that the vielbein is trivial in the ninth direction, the crossterm equals

$$V^{cross}_{curv} = -\frac{T\lambda^2}{2}\eta_{9ab,9cd}\eta_{BD}E_J^B E_L^D \mathrm{Tr}\mathcal{D}(X)^{9ab}\mathcal{D}(X)^J\mathcal{D}(X)^{9cd}\mathcal{D}(X)^L\ . \tag{11.55}$$

The triviality of the vielbein allows the use of the following contraction identity,

$$\eta_{IJ}\mathcal{D}(X)^{I,ad}\mathcal{D}^{J,ef} = 2(X^k)^{af}(X_k)^{ed} - \delta^{af}(X^2)^{ed} - \delta^{ed}(X^2)^{af}, \qquad (11.56)$$

in the direction of the circle. Thus

$$V^{cross}_{curv} = -T\lambda^2\eta_{BD}E^B_J E^D_L \left[\mathrm{Tr}\left(X^9\mathcal{D}(X)^J\right)\mathrm{Tr}\left(X^9\mathcal{D}(X)^L\right) - \mathrm{Tr}\left(\mathcal{D}(X)^J\right)\mathrm{Tr}\left(X^9X^9\mathcal{D}(X)^L\right)\right]. \qquad (11.57)$$

Due to the defining property 11.48 of the commutation operator, the last term vanishes as $\mathrm{Tr}(D(X)^I) = [X^i, \mathbb{I}] = 0$. The remaining term yields

$$V^{cross}_{curv} = -T\lambda^2\eta_{BD}E^B_{j\alpha\beta}E^D_{\ell\gamma\delta}\left([X^j, X^9]^{\alpha\beta}[X^\ell, X^9]^{\gamma\delta}\right). \qquad (11.58)$$

Upon using the standard T-duality rule which replaces commutators with derivatives,

$$i\lambda[X^9, \mathcal{F}(X)] \rightarrow \partial_9\mathcal{F}, \qquad (11.59)$$

the kinetic term equation 11.7 is recovered exactly. Notice that the vielbeins have basically just gone along for the ride. The proof of T-duality in a flat background would be identical. This clearly shows the power of the matrix-geometry approach.

CHAPTER 12

EVIDENCE FOR A GRAVITATIONAL MYERS EFFECT

(12.1) CORRECTION TERMS AND THE MYERS EFFECT

The flat-space action (equation 8.7) derived from the supersymmetric Yang-Mills theory is only valid for low energies (in fact $\dot{X}^2 \ll 1$) and receives different types of corrections if one tries to extend the validity of this action.

Higher derivative corrections come in the form of the Dirac-Born-Infeld (DBI) action

$$S_{\mathrm{DBI}} = m \int \mathrm{d}\tau \mathrm{Tr}\, \sqrt{-\eta_{\mu\nu}\partial_\tau X^\mu \partial_\tau X^\nu}. \tag{12.1}$$

It is easy to see that in the limit $\dot{X} \ll 1$ the kinetic term as in equation 8.7 is recovered. R.C. Myers [54] showed that, in order for this action to be consistent with T-duality in a background of constant RR 4-form fieldstrength $F^{(4)}$, the potential must receive further contributions of the form (see also equation 8.11):

$$V = -\tfrac{\lambda^2}{4}\mathrm{Tr}\,\left([X^i, X^j]^2\right) - \tfrac{\mathrm{i}}{3}\lambda^2 \mathrm{Tr}\,\left(X^i X^j X^k\right) F^{(4)}_{0ijk}. \tag{12.2}$$

These type of corrections have very interesting consequences. To see this, take the following constant background field:

$$F^{(4)}_{0ijk} = \begin{cases} f\epsilon_{ijk} & i,j,k \in 1,2,3 \\ 0 & \text{otherwise} \end{cases}. \tag{12.3}$$

The minima of the potential are now given by configurations that satisfy

$$[X^i, X^j] = \mathrm{i} f \epsilon_{ijk} X^k. \tag{12.4}$$

If the background flux was turned off somehow, the matrices X^i would have to be proportional to the unit matrix, hence all D0-branes would be on top of each other. This configuration would have zero energy. However, with f not equal to zero this configuration does not minimize the potential anymore. Configurations that do minimize the potential are given by:

$$X^i = \tfrac{f}{2}\alpha^i, \tag{12.5}$$

with α^i a $N \times N$ matrix representation of the $SU(2)$ algebra. Instead of sitting on top of each other, the D0-branes are now separated. To get an idea for how far the D0-branes are separated one can calculate the following quantity:

$$\mathcal{R} = \sqrt{\tfrac{\sum_{i=1}^{3} \mathrm{Tr}\,(X^i)^2}{N}} = \tfrac{f}{2}\sqrt{N^2-1}. \tag{12.6}$$

The surprising thing is that this configuration actually has a lower energy than the diagonal configuration:

$$V = -\frac{\lambda^2 f^2}{6} N(N^2-1). \tag{12.7}$$

So, due to the background flux, the D0-branes find it favorable to blow up in three directions to form a higher dimensional extended object that couples to the background flux. This effect is called the Myers effect. As an analogy one can think of this as a dielectric effect. A neutral dielectric material only shows its bound electric charges, in the form of dipoles, after an external electric field is applied. Here the RR 4-form charge of the D0-branes only appears after turning on a RR 4-form background field.

(12.1.1) TACHYONIC INSTABILITIES

The fact that the configuration of having N D0-branes on top of each other is not the lowest energy configuration in a RR four form background means that there must be tachyonic modes signaling the instability. To find those tachyonic modes one calculates the spectrum of quadratic fluctuations around this configuration. As an illustration the case of two D0-branes will be worked out (see [56]).

As the background configuration the D0-branes are not taken exactly on top of each other, instead the following is used

$$\bar{X}^i = \begin{pmatrix} \bar{x}^i & 0 \\ 0 & \bar{x}^i + \Delta^i \end{pmatrix}. \tag{12.8}$$

In the end Δ can always taken to be zero. The fluctuation δX is written as

$$\delta X^i = \begin{pmatrix} a^i & b^i \\ \bar{b}^i & d^i \end{pmatrix}. \tag{12.9}$$

Pluggin $X = \bar{X} + \delta X$ into equation 12.2 gives

$$\delta^2 V\big|_{X=\bar{X}} = \int \mathrm{d}\tau b^i (\delta_{ij}\lambda^2\Delta^2 + \mathrm{i}F_{0ikj}\Delta^k)\bar{b}^j. \tag{12.10}$$

In this calculation the Gauss law (the equation of motion for A_0 that still needs to be imposed)

$$[\dot{X}^i, X^i] = 0, \tag{12.11}$$

states that only fluctuations transverse to the static diagonal background are dynamical. Physical fluctuations therefore satisfy $[\delta X^i, \bar{X}^i] = 0$. This means that the physical fluctuations are orthogonal to the D0-brane separation $\Delta^i \equiv (x_1^i - x_2^i)$. Therefore the physical fluctuations are the b^i.

Note that there are only quadratic fluctuations in the off-diagonal elements, representing the strings stretching between the branes. To determine to existence of tachyonic modes it is necessary to calculate the eigenvalues of the mass matrix that follow from this expression. Taking the flux background to be (in accordance with equation 12.3)

$$F_{0ijk} = \lambda^2 \rho \epsilon_{ijk}, \tag{12.12}$$

the eigenvalues become (listed as $(eigenvalue)_{degeneracy}$):

$$(\Delta^2)_2, (\Delta^2 - \rho\Delta)_2, (\Delta^2 + \rho\Delta)_2. \tag{12.13}$$

By judiciously choosing Δ (small) and ρ (close to unity) one can make two of the eigenvalues negative, signaling instabilities in the off-diagonal components of X: a hint for the Myers effect. An interesting aspect of this instability is that it does not occur if Δ is exactly zero. One could interpret this as the instability being marginally stable, having the D0-branes on top of each other is a stable configuration, but even the slightest perturbation will cause the system to expand into the Myers configuration. However one should be careful with such line of reasoning since technically the action that is used to derive these results is not valid if the D0-branes are taken exactly on top of each other [56]. Therefore the end-conclusion remains, there is an instability that causes a blow-up of the D0-brane configuration if the D0-branes are separated slightly.

(12.2) INSTABILITIES IN CURVED BACKGROUNDS

One can ask the question if there exist gravitational backgrounds that cause a Myers effect, i.e. causing an instability in D0-brane configurations. Since the action of D0-branes in a curved background is still under discussion, it is difficult to find exact solutions one can trust. However one can try, within the limits of applicability, to find the tachyonic instabilities. Those instabilities are a necessary condition for the Myers effect to occur.

The action for D0-branes is known in a certain approximation, that of low velocity ($\dot{X} \ll 1$) and weak curvature (such that terms quadratic in the Riemann tensor can be ignored), also known as the linearized approximation. In this approximation the ordering of all terms is completely symmetric. The computation is essentially the same as in section 12.1.1, however in the case of a gravitational background it is also necessary to take into account the fluctuations of the kinetic term. The result of this calculation is (see [56] for a complete derivation)

$$\begin{aligned}\delta^2 S = &-\int d\tau\, [b^\nu(\ 2\lambda^2 g_{\mu\nu}(\bar{x}) g_{\alpha\beta}(\bar{x})\Delta^\alpha\Delta^\beta \\ &\qquad -m(R_{\alpha\mu\nu\beta}(\bar{x}) + R_{\alpha\nu\mu\beta}(\bar{x}))(\dot{\bar{x}}^\alpha\dot{\bar{x}}^\beta + \tfrac{1}{12}\dot{\Delta}^\alpha\dot{\Delta}^\beta))\bar{b}^\mu\Big] \\ \equiv &-\int d\tau b^\nu m_{\nu\mu}(p)\bar{b}^\mu.\end{aligned} \tag{12.14}$$

Because the metric is positive definite by definition, the static configurations of D0-branes clearly have no tachyonic modes. Therefore one has to look at configurations where D0-branes move along the geodesics of the curved background.

To illustrate the method the computation will be done for D0-branes moving on a positive curvature background, a sphere, and a negative curvature background, a hyperboloid. An interesting distinction between the two will be found.

To facilitate the computation, the metric of the sphere is written in RNC. Since the sphere is a homogeneous space the transformation from the standard metric to RNC is easily found and one obtains the metric:

$$\mathrm{d}s^2_{S^2,RNC} = \mathrm{d}z_x^2 + \mathrm{d}z_y^2 - \frac{1}{3D^2}(z_x\mathrm{d}z_y - z_y\mathrm{d}z_x)^2 + \mathcal{O}(z^3). \tag{12.15}$$

One easily checks that geodesics through the origin $z^i = 0$ can be written as $z^i = v^i\tau$ as required. Furthermore one immediately reads off the Riemann tensor at the origin $z^i = 0$:

$$\begin{aligned} ds^2_{RNC} &= \eta_{\mu\nu}dx^\mu dx^\nu + \tfrac{1}{3}R_{\mu\alpha\beta\nu}(0)x^\alpha x^\beta dx^\mu dx^\nu + \ldots \\ \Rightarrow \quad & R_{xyyx}(0) = -\frac{1}{D^2}. \end{aligned} \tag{12.16}$$

Consider a system of two D0-branes both moving along the same great arc through the origin, but slightly separated along this arc. If one chooses the z^x-axis as this arc, this implies that $\dot{z}_i^y = 0$, $\Delta^y = 0$ and b^x is the unphysical fluctuation. Substituting this data into equation 12.14 leads to:

$$\delta^2 S = -\int d\tau \left[b_x \left(2\lambda^2\Delta^2\right) \bar{b}_x + b_y \left(2\lambda^2\Delta^2 + \frac{2m}{D^2}\left((\dot{\bar{z}}^x)^2 + \tfrac{1}{12}(\dot{\Delta}^x)^2\right)\right) \bar{b}_y \right] . \tag{12.17}$$

Again the mass-matrix for the unphysical fluctuation b^x is that given by D-geometry and strictly positive definite. For the physical b^y fluctuation the mass matrix is explicitly positive semidefinite, and there is no tachyon in this case. The physical picture explains this in simple terms. There are two forces acting: the force due to the geodesic deviation and the force of open string between the branes. In this positive curvature background the geodesic deviation works in the same direction as the force by the stretched open strings, leading to a stable configuration.

As promised the same procedure will be applied to a negatively curved space. For two nearly coincident D0-branes on a two-dimensional hyperboloid, the mass-matrix of fluctuations is simply equation 12.17 with the substitution $D^2 \to -D^2$. The gravitational tidal force has changed sign and can now counterbalance the attracting strings. Specifically there is a tachyonic instability in the spectrum if and only if the D-brane separation is small compared to the velocity times the curvature:

$$2\lambda^2\Delta^2 \ll \frac{2m}{D^2}\left((\dot{\bar{z}}^x)^2 + \tfrac{1}{12}(\dot{\Delta}^x)^2\right) . \tag{12.18}$$

Note, however, that for large separations the attractive force of the open strings dominates the repelling force of the background geometry, and the tachyon disappears.

To truly test whether a tachyonic instability is found, one question remains. The tachyon appears to be evident for large speeds or small separations. Both extremes, however are outside the range of validity of the action [56]. Large speeds violate the truncation to second order in derivatives, while separations Δ must be larger than the eleven-dimensional Planck length. One needs to check whether the inequality 12.18, signaling the instability, can be satisfied within these bounds. Without loss of generality, one may simplify the inequality by approximating $\dot{\Delta}^x \simeq 0$. Substituting $\dot{z}^x = v$ with $|v| \ll 1$, expressing m and λ in string units g_s, ℓ_s and multiplying both sides of equation 12.18 by ℓ_s^2, one obtains

$$\frac{\Delta^2}{\ell_s^2} \ll v^2 \frac{\ell_s^2}{D^2} \ll 1 . \tag{12.19}$$

In words, the condition for a tachyon to be present is that the separation in string units must be less then the velocity times the curvature in string units. Also used here is the implicit weak gravity and Born-Infeld condition: both sides must be less than unity. Most importantly, however, the separation Δ must also be larger than the eleven-dimensional Planck length

$$g_s^{1/3}\ell_s \ll \Delta^i \ll \ell_s \ . \tag{12.20}$$

The window where equation 12.19 can be satisfied within the range of equation 12.20 is when the velocity times the curvature is much larger than the eleven-dimensional Planck length:

$$g_s^{2/3} \ll v^2 \frac{\ell_s^2}{D^2} \ . \tag{12.21}$$

This is easily satisfied for very weak string coupling. The region where the "individual" geodesic solution is unstable thus falls easily within the region of validity of the action. This is evidence for the existence of a lesser energetic stable configuration. It suggests that a purely gravitational Myers effect exists.

CHAPTER 13

CONCLUSIONS

Finding the action for multiple D-branes proves to be a challenging problem. The approach taken in this thesis is that of D-geometry. The D-geometry axioms put forward by Douglas do not seem to be sufficient to derive the action for D0-branes in a gravitational background. Perhaps in principle they are, but in practice certainly not. Therefore a method that has proven to be successful in theoretical physics is used: imposing a symmetry constraint. The symmetry closest to gravity is that of general coordinate or diffeomorphism invariance. Since the diffeomorphism invariance, of the space the D0-branes move around in, can not be lifted to the group of coordinate transformations in the larger matrix space, essential in the D-geometry formulation, imposing this symmetry is not straightforward.

Finding inspiration from the manifestly covariant background field expansion, a new symmetry principle is used to implement a kind of diffeomorphism invariance in the D-geometry formulation. In short this symmetry principle entails the expansion of the action around some base-point and the independence of the expansion of the choice of base-point. This base-point independence can be taken as a definition of diffeomorphism invariance in D-geometry.

The base-point independence can be successfully applied to the action of D0-branes in a curved background. Results obtained using other methods are found to be consistent with base-point independence. A common claim, that the symmetrized trace prescription is a general solution to the ordering problem in the action for D0-branes, is not consistent with base-point independence. The inconsistencies are found when trying to extend the action to be valid for higher velocities (expansion in $\dot{X}$) or higher curvatures (expansion in R). Non-symmetrized trace corrections, rendering the action base-point independent, can be found. However those corrections terms are not uniquely determined by this new symmetry principle.

D0-branes can exhibit interesting collective behavior. Since there are strings stretching between the branes one would think that it is always energetically favorable to put the D0-branes on top of each other. However in a special background (a Ramond-Ramond four form flux) the D0-branes tend to blow up into a spherical configuration. This is usually referred to as the Myers effect [54]. This is very peculiar since a single D0-brane is insensitive to this flux. It turns out that a system of multiple D0-branes *can* couple to this flux. One can wonder if gravitational couplings can also show such collective behavior. Evidence for this is found [56] in the sense that D0-branes moving along geodesics on a negatively curved space experience an instability (tachyon) in the modes that represent the strings stretching between the D0-branes, just like in the original Myers effect. Of course it may happen that the instability does not lead to a nearby stable vacuum. To decide about the fate of the instability one would have to extend the validity of the action for D0-branes in a curved background. Hopefully it is clear by now that this is not an easy thing to do.

Appendix A

Affine Lie Algebras

This appendix provides the necessary background in affine Lie algebras, needed to understand part one of this thesis. Familiarity with Lie algebras is assumed, only the more advanced topics will be explained. Excellent introductions into the theory of (affine) Lie algebras can be found in [70][71][72].

(A.1) Simple Lie algebras

Central in the discussion of Lie algebras is the Cartan matrix. The Cartan matrix defines literally all the properties of the Lie algebra. Because of the importance of the Cartan matrix, its properties will be discussed here. A simple Lie algebra of rank r has associated to it an $r \times r$ matrix A, the Cartan matrix, satisfying the following properties

1. $A^{ii} = 2$
2. $A^{ij} = 0 \Leftrightarrow A^{ji} = 0$
3. $A^{ij} \in \mathbb{Z}$ and $A^{ij} < 0$ for $i \neq j$
4. $\det A > 0$
5. It cannot be brought in block diagonal form by rearranging columns and rows.

There is also a graphical representation of the Cartan matrix: Dynkin diagrams. A Dynkin diagram is a diagram with r nodes. Two nodes i and j are connected by one or more lines if A^{ij} is not equal to zero. The number of lines connecting the nodes i

Figure A.1: *The Dynkin diagram of G_2.*

and j is given by the formula

$$A^{ij}A^{ji} \text{ (no summation).} \tag{A.1}$$

In the case that $|A^{ij}| \geq |A^{ji}|$ one draws an arrow pointing from node i to j. As a simple example consider the following matrix

$$A = \begin{pmatrix} 2 & -3 \\ -1 & 2 \end{pmatrix}. \tag{A.2}$$

It is easy to check that this matrix satisfies all the above listed properties a Cartan matrix should satisfy. Therefore this matrix is the Cartan matrix of some simple rank two Lie algebra. The Lie algebra is called G_2. The Dynkin diagram of this Lie algebra is shown in figure A.1. The Cartan matrix for the Lie algebra of $SO(8)$, D_4 is

$$A_{D_4} = \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & -1 \\ 0 & -1 & 2 & -1 \\ 0 & -1 & -1 & 2 \end{pmatrix}. \tag{A.3}$$

The Cartan matrix for the Lie algebra of $SU(r+1)$, A_r is given by a tri-diagonal matrix

$$A_{A_r} = \begin{pmatrix} 2 & -1 & & \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & 2 \end{pmatrix}. \tag{A.4}$$

There is a particularly useful basis of the Lie algebra in which the role of the Cartan matrix becomes very clear. In this basis, the Chevalley basis, the whole algebra is generated by $3r$ elements h^i, e^i and f^i satisfying the following algebra

$$\begin{aligned} \left[h^i, h^j\right] &= 0 \\ \left[h^i, e^j\right] &= -A^{ij}e^j \\ \left[h^i, f^j\right] &= A^{ij}f^j \\ \left[e^i, f^j\right] &= \delta^{ij} \end{aligned} \tag{A.5}$$

All the other elements of the algebra are then generated by computing all Lie brackets possible for the elements h, e and f. The h-elements form an Abelian subalgebra, called the Cartan subalgebra. The e^i and f^i form pairs similar to the raising and lowering operators of the $SU(2)$ algebra (A_1).

It is more common to define the Cartan matrix in terms of the root system like this

$$A^{ij} = 2\,\frac{\alpha^{(i)} \cdot \alpha^{(j)}}{\alpha^{(j)} \cdot \alpha^{(j)}}. \tag{A.6}$$

Where the $\alpha^{(i)}$ are the simple roots. Knowing the root system is therefore equivalent to knowing (in principle) everything about the Lie algebra. The highest root of the root system is given by the following expression

$$\theta = \sum_i a_i \alpha^{(i)}. \tag{A.7}$$

The natural numbers a_i are called the marks (or Coxeter numbers) of the Lie algebra. For G_2 the marks are

$$\theta = \alpha^{(1)} + 2\alpha^{(2)}. \tag{A.8}$$

Of course, one can calculate the marks by constructing the highest weight. However it is much easier to use the affine extension of the Lie algebra, to be introduced in the next section.

The Coxeter number is easily calculated once the marks are known

$$g = 1 + \sum_i a_i. \tag{A.9}$$

Likewise, for the dual Coxeter number

$$g^\vee = 1 + \sum_i a_i^\vee. \tag{A.10}$$

The co-marks $a^\vee$ are defined below.

(A.2) Affine extensions

Affine Lie algebras are obtained by allowing more general Cartan matrices. For affine Lie algebras the condition

$$\mathrm{det} A > 0 \tag{A.11}$$

is replaced by

$$\mathrm{det} A_{\{i\}} > 0 \text{ for all } i = 0, \ldots, r. \tag{A.12}$$

The matrices $A_{\{i\}}$ are the principal minors of A, obtained by removing the i-th row and column. One can associate a Dynkin diagram to these Cartan matrices using

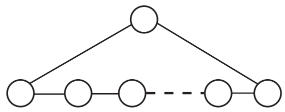


Figure A.2: $A_r^{(1)}$*, the affine extension of* A_r.

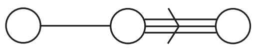


Figure A.3: $G_2^{(1)}$*, the affine extension of* G_2.

exactly the same rules as for the simple Lie algebras. The condition in equation A.12 then corresponds to the following statement: removing a single node from the affine Dynkin diagram should result in the Dynkin diagram of a semi-simple Lie algebra.

The Chevalley basis for the affine Lie algebra is still given by equation A.5, however now the index is running from 0 to r (r is the dimension of the Cartan matrix, it is not the rank anymore). So the description of the affine Lie algebra in terms of the Chevalley basis is just as simple as for the simple Lie algebras. However in the affine case the algebra does not close: it is infinite dimensional.

The marks and co-marks for the affine Lie algebra are particularly simple to calculate. The co-marks are given by the eigenvector, belonging to eigenvalue zero, of the affine Cartan matrix

$$\sum_j A^{ij} a_j^\vee = 0. \tag{A.13}$$

The $a_j^\vee$ are fixed by requiring that the co-mark for the affine node is zero: $a_0^\vee = 1$. To find the marks one uses the transposed affine Cartan matrix

$$\sum_j (A^T)^{ij} a_j = 0. \tag{A.14}$$

Again, $a_0 = 1$ is required. The Coxeter numbers are now given by

$$g = \sum_{i=0}^{r} a_i, \qquad g^\vee = \sum_{i=0}^{r} a_i^\vee. \tag{A.15}$$

To every simple Lie algebra one can associate an affine extension, defined by the fact that the marks of the affine extension reproduce the marks for the simple Lie algebra. For the A_r series the affine extension is obtained by adding a node and connecting it to the outermost two nodes. See figure A.2 for the affine extensions of A_r and G_2. The Cartan matrix of the affine extension of G_2 is

$$\begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -3 \\ 0 & -1 & 2 \end{pmatrix}. \tag{A.16}$$

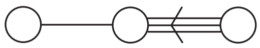


Figure A.4: *The Dynkin diagram of the dual affine Lie algebra of G_2, or $G_2^{(1)} = D_4^{(3)}$. Compared to the Dynkin diagram of $G_2^{(1)}$ only the direction of the arrow is reversed.*

The Dual Algebra

In this thesis the notion of a dual algebra appears. In terms of the Cartan matrix the dual algebra is very easy to define. The only thing one needs to do is to take the transpose of the Cartan matrix. Thus, if A is the Cartan matrix of Lie algebra $\mathfrak{g}$, then A^T is the Cartan matrix of the dual Lie algebra $\mathfrak{g}^\vee$. The A_r and D_r series are self-dual, the B_r and C_r series are each others dual.

Twisted Algebras

Sometimes the dual algebra can also be obtained by twisting another algebra. The twisting procedure will be described briefly here. Consider a Dynkin diagram of a simple Lie algebra. Generically this diagram has some discrete symmetry group Γ. This symmetry group interchanges nodes of the Dynkin diagram. On the level of the Lie algebra this means that this symmetry interchanges generators in the Chevalley basis. The twisted algebra is obtained by constructing a subalgebra, invariant under the symmetry group Γ. The relevant example in this thesis is D_4. Is has an S_3 symmetry, the invariant subalgebra being a G_2 algebra (see the discussion in section 5.5.2). After this twisting one can add an affine root to the algebra to obtain the $D_4^{(3)}$ algebra, which happens to be equivalent to the $G_2^{(1)\vee}$ algebra (see figure A.2). The Cartan matrix of this algebra is

$$\begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -3 & 2 \end{pmatrix}, \tag{A.17}$$

which is indeed the transpose of equation A.16.

APPENDIX B

A RECURRENCE RELATION FOR $P_N(x)$

In this appendix a recurrence relation for characteristic polynomial of the Lax matrix for $U(N)$ is derived. This characteristic polynomial is defined as

$$P_N(x) = \det(x\mathbb{I}_N - L), \tag{B.1}$$

written out more explicitly

$$P_N(x) = \begin{vmatrix} x-p_1 & -y_1 & 0 & . & . & . & 0 & -z \\ -1 & x-p_2 & -y_2 & 0 & . & . & . & 0 \\ 0 & -1 & x-p_3 & -y_3 & 0 & . & . & 0 \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & -1 & x-p_{N-2} & -y_{N-1} \\ -\frac{y_0}{z} & 0 & . & . & . & . & -1 & x-p_{N-1} \end{vmatrix}. \tag{B.2}$$

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

$$G_N(x) = \begin{vmatrix} x-p_1 & -y_1 & 0 & . & . & . & 0 & 0 \\ -1 & x-p_2 & -y_2 & 0 & . & . & . & 0 \\ 0 & -1 & x-p_3 & -y_3 & 0 & . & . & 0 \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & -1 & x-p_{N-2} & -y_{N-1} \\ 0 & 0 & . & . & . & . & -1 & x-p_{N-1} \end{vmatrix}. \tag{B.3}$$

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

$$\begin{aligned}P_N(x) &= (x - p_{N-1})G_{N-1}(x) - (-1)(-y_{N-1})G_{N-2}(x) \\ &+ (-1)^{N+1}(-\frac{y_0}{z})(-1)^N(-z)G^+_{N-2}(x) \\ &= (x - p_{N-1})G_{N-1}(x) - y_{N-1}G_{N-2}(x) - y_0G^+_{N-2}(x),\end{aligned} \tag{B.4}$$

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

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

$$G_N(x) = (x - p_{N-1})G_{N-1}(x) - y_{N-1}G_{N-2}(x) \tag{B.5}$$

$$G^+_N(x) = (x - p_N)G^+_{N-1}(x) - y_NG^+_{N-2}(x). \tag{B.6}$$

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

$$P_N(x) = (x - p)G_{N-1} - 2yG_{N-2} = G_N(x) - yG_{N-2}, \tag{B.7}$$

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

$$P_N(x) = (x - p)P_{N-1}(x) - yP_{N-2}(x), \tag{B.8}$$

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

Appendix C

Equations of motion for G_2

Here the equations of motion for G_2 with the potential given in (5.93).

For p_1 one gets

$$\begin{aligned}0 = 3g_2(2\phi_1 - \phi_2) + 2g_6\Big(&- 27y_2y_1\phi_2 + 9y_1y_0\phi_2 - 12y_2y_0\phi_2 + \frac{4}{3}y_0^2\phi_2 - 12y_0\phi_2^3 \\ &+ 9y_0\phi_1^2\phi_2 - 81\phi_2^3\phi_1^2 + 54\phi_2^2\phi_1^3 + \frac{1}{3}y_0^2\phi_1 + 6y_0\phi_1\phi_2^2 + 27y_2^2\phi_1 + 54y_1\phi_2^2\phi_1 \\ &- 81y_2\phi_1^2\phi_2 + 54y_2\phi_1\phi_2^2 - 6y_2y_0\phi_1 - 27y_1\phi_2^3 + 27\phi_2^4\phi_1\Big) \,.\end{aligned} \tag{C.1}$$

For p_2 one has

$$\begin{aligned}0 = 3g_2(2\phi_2 - \phi_1) + 2g_6\Big(&- \frac{8}{3}y_0^2\phi_2 + 27y_1^2\phi_2 - 27y_2\phi_1y_1 + 9y_0\phi_1y_1 - 12y_2y_0\phi_1 \\ &+ \frac{4}{3}y_0^2\phi_1 - 36y_0\phi_1\phi_2^2 + 3y_0\phi_1^3 - 81\phi_2^2\phi_1^3 + 27\phi_2\phi_1^4 + 6y_0\phi_1^2\phi_2 + 24y_0\phi_2^3 + 54y_1\phi_2\phi_1^2 \\ &- 27y_2\phi_1^3 + 54y_2\phi_1^2\phi_2 + 24y_2y_0\phi_2 - 81y_1\phi_2^2\phi_1 - 12y_1y_0\phi_2 + 54\phi_2^3\phi_1^2\Big) \,.\end{aligned} \tag{C.2}$$

For y_0 one finds

$$\begin{aligned}\frac{\ell}{y_0} = g_2 + g_6\Big(&- \frac{16}{3}y_0\phi_2^2 + \frac{4}{9}y_0^2 + \frac{8}{3}y_1y_0 + 18y_1\phi_1\phi_2 - 24\phi_1\phi_2y_2 - 12y_1y_2 \\ &+ \frac{16}{3}\phi_1y_0\phi_2 - 24\phi_2^3\phi_1 + 6\phi_1^3\phi_2 + \frac{2}{3}y_0\phi_1^2 + 6\phi_2^2\phi_1^2 + 12\phi_2^4 + 12y_2^2 - 6\phi_1^2y_2 \\ &+ 24\phi_2^2y_2 - 12y_1\phi_2^2 - \frac{16}{3}y_2y_0\Big) \,.\end{aligned} \tag{C.3}$$

For y_1 one gets

$$\frac{2\ell}{y_1} = 3g_2 + g_6\bigg(54y_1\phi_2^2 + \frac{4}{3}y_0^2 - 54\phi_1\phi_2y_2 + 18\phi_1y_0\phi_2 - 12y_2y_0 + 54\phi_2^2\phi_1^2 - 54\phi_2^3\phi_1 - 12y_0\phi_2^2\bigg) \; . \tag{C.4}$$

For y_2 one gets

$$\frac{\ell}{y_2} = 3g_2 + g_6\bigg(- 54y_1\phi_1\phi_2 - 24\phi_1y_0\phi_2 - 12y_1y_0 + 54\phi_1^2y_2 + 24y_2y_0 - 54\phi_1^3\phi_2 + 54\phi_2^2\phi_1^2 - 6y_0\phi_1^2 + 24y_0\phi_2^2 - \frac{8}{3}y_0^2\bigg) \; , \tag{C.5}$$

and finally for ℓ one has $y_0y_1^2y_2 = \Lambda^8/36$.

BIBLIOGRAPHY

[1] N. Seiberg, *Naturalness versus supersymmetric nonrenormalization theorems*, *Phys. Lett.* **B318** (1993) 469–475 [hep-ph/9309335].

[2] N. Seiberg and E. Witten, *Electric - magnetic duality, monopole condensation, and confinement in n=2 supersymmetric yang-mills theory*, *Nucl. Phys.* **B426** (1994) 19–52 [hep-th/9407087].

[3] R. Dijkgraaf and C. Vafa, *A perturbative window into non-perturbative physics*, hep-th/0208048.

[4] R. Dijkgraaf, M. T. Grisaru, C. S. Lam, C. Vafa and D. Zanon, *Perturbative computation of glueball superpotentials*, *Phys. Lett.* **B573** (2003) 138–146 [hep-th/0211017].

[5] R. Dijkgraaf and C. Vafa, *N = 1 supersymmetry, deconstruction, and bosonic gauge theories*, hep-th/0302011.

[6] J. Zinn-Justin, *Large order estimates in perturbation theory*, *Phys. Rept.* **49** (1979) 205–213.

[7] G. 't Hooft, *A planar diagram theory for strong interactions*, *Nucl. Phys.* **B72** (1974) 461.

[8] E. Brezin, *Planar diagrams*, *Phys. Rept.* **49** (1979) 221–227.

[9] M. J. Strassler, *On confinement and duality*, . Prepared for ICTP Spring School on Superstrings and Related Matters, Trieste, Italy, 2-10 Apr 2001.

[10] H. B. Nielsen and P. Olesen, *Vortex-line models for dual strings*, *Nucl. Phys.* **B61** (1973) 45–61.

[11] G. 't Hooft, *Under the spell of the gauge principle*, *Adv. Ser. Math. Phys.* **19** (1994) 1–683.

[12] J. Wess and J. Bagger, *Supersymmetry and supergravity*. Princeton University Press, 1992.

[13] E. D'Hoker and D. H. Phong, *Lectures on supersymmetric yang-mills theory and integrable systems*, hep-th/9912271.

[14] N. Seiberg, *Supersymmetry and nonperturbative beta functions*, *Phys. Lett.* **B206** (1988) 75.

[15] V. A. Novikov, M. A. Shifman, A. I. Vainshtein and V. I. Zakharov, *Instanton effects in supersymmetric theories*, *Nucl. Phys.* **B229** (1983) 407.

[16] 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].

[17] G. Veneziano and S. Yankielowicz, *An effective lagrangian for the pure n=1 supersymmetric yang-mills theory*, *Phys. Lett.* **B113** (1982) 231.

[18] I. Montvay, *Supersymmetric yang-mills theory on the lattice*, *Int. J. Mod. Phys.* **A17** (2002) 2377–2412 [hep-lat/0112007].

[19] J. de Boer and Y. Oz, *Monopole condensation and confining phase of n = 1 gauge theories via m-theory fivebrane*, *Nucl. Phys.* **B511** (1998) 155–196 [hep-th/9708044].

[20] P. D. Lax, *Integrals of nonlinear equations of evolution and solitary waves*, *Comm. Pure Appl. Math.* **21** (1968) 467–490.

[21] E. J. Martinec and N. P. Warner, *Integrable systems and supersymmetric gauge theory*, *Nucl. Phys.* **B459** (1996) 97–112 [hep-th/9509161].

[22] R. Donagi and E. Witten, *Supersymmetric yang-mills theory and integrable systems*, *Nucl. Phys.* **B460** (1996) 299–334 [hep-th/9510101].

[23] O. Babelon, D. Bernard and M. Talon, *Introduction to Classical Integrable Systems*. Cambridge University Press, 2003.

[24] R. Boels, J. de Boer, R. Duivenvoorden and J. Wijnhout, *Nonperturbative superpotentials and compactification to three dimensions*, *JHEP* **03** (2004) 009 [hep-th/0304061].

[25] R. Boels, J. de Boer, R. Duivenvoorden and J. Wijnhout, *Factorization of seiberg-witten curves and compactification to three dimensions*, *JHEP* **03** (2004) 010 [hep-th/0305189].

[26] M. Alishahiha, J. de Boer, A. E. Mosaffa and J. Wijnhout, *N = 1 g(2) sym theory and compactification to three dimensions*, *JHEP* **09** (2003) 066 [hep-th/0308120].

[27] N. Seiberg and E. Witten, *Gauge dynamics and compactification to three dimensions*, hep-th/9607163.

[28] N. Dorey, *An elliptic superpotential for softly broken n = 4 supersymmetric yang-mills theory*, *JHEPwef* **07** (1999) 021 [hep-th/9906011].

[29] N. Dorey, T. J. Hollowood and S. Prem Kumar, *An exact elliptic superpotential for n = 1* deformations of finite n = 2 gauge theories*, *Nucl. Phys.* **B624** (2002) 95–145 [hep-th/0108221].

[30] N. Dorey, T. J. Hollowood, S. Prem Kumar and A. Sinkovics, *Exact superpotentials from matrix models*, *JHEP* **11** (2002) 039 [hep-th/0209089].

[31] 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].

[32] A. Kapustin, *Solution of n = 2 gauge theories via compactification to three dimensions*, *Nucl. Phys.* **B534** (1998) 531–545 [hep-th/9804069].

[33] E. Witten, *Solutions of four-dimensional field theories via m-theory*, *Nucl. Phys.* **B500** (1997) 3–42 [hep-th/9703166].

[34] J. de Boer, K. Hori and Y. Oz, *Dynamics of n = 2 supersymmetric gauge theories in three dimensions*, *Nucl. Phys.* **B500** (1997) 163–191 [hep-th/9703100].

[35] K. A. Intriligator, R. G. Leigh and N. Seiberg, *Exact superpotentials in four-dimensions*, *Phys. Rev.* **D50** (1994) 1092–1104 [hep-th/9403198].

[36] M. Alishahiha and A. E. Mosaffa, *On effective superpotentials and compactification to three dimensions*, hep-th/0304247.

[37] M. R. Douglas and S. H. Shenker, *Dynamics of su(n) supersymmetric gauge theory*, *Nucl. Phys.* **B447** (1995) 271–296 [hep-th/9503163].

[38] *On-line encyclopedia of integer sequences, http://www.research.att.com/˜njas/sequences/* **sequence number A078531**.

[39] K. A. Intriligator, *'integrating in' and exact superpotentials in 4-d*, *Phys. Lett.* **B336** (1994) 409–414 [hep-th/9407106].

[40] R. Dijkgraaf, S. Gukov, V. A. Kazakov and C. Vafa, *Perturbative analysis of gauged matrix models*, *Phys. Rev.* **D68** (2003) 045007 [hep-th/0210238].

[41] F. Cachazo, N. Seiberg and E. Witten, *Phases of $n = 1$ supersymmetric gauge theories and matrices*, *JHEP* **02** (2003) 042 [hep-th/0301006].

[42] F. Cachazo, N. Seiberg and E. Witten, *Chiral rings and phases of supersymmetric gauge theories*, *JHEP* **04** (2003) 018 [hep-th/0303207].

[43] 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.

[44] M. Alishahiha, F. Ardalan and F. Mansouri, *The moduli space of the supersymmetric g_2 yang-mills theory*, *Phys. Lett.* **B381** (1996) 446–450 [hep-th/9512005].

[45] M. Aganagic, K. Intriligator, C. Vafa and N. P. Warner, *The glueball superpotential*, *Adv. Theor. Math. Phys.* **7** (2004) 1045–1101 [hep-th/0304271].

[46] E. Witten, *String theory dynamics in various dimensions*, *Nucl. Phys.* **B443** (1995) 85–126 [hep-th/9503124].

[47] C. M. Hull, *Exact pp wave solutions of 11-dimensional supergravity*, *Phys. Lett.* **B139** (1984) 39.

[48] P. K. Townsend, *M-theory from its superalgebra*, hep-th/9712004.

[49] A. Bilal, *M(atrix) theory: A pedagogical introduction*, *Fortsch. Phys.* **47** (1999) 5–28 [hep-th/9710136].

[50] T. Banks, W. Fischler, S. H. Shenker and L. Susskind, *M theory as a matrix model: A conjecture*, *Phys. Rev.* **D55** (1997) 5112–5128 [hep-th/9610043].

[51] N. Seiberg, *Why is the matrix model correct?*, *Phys. Rev. Lett.* **79** (1997) 3577–3580 [hep-th/9710009].

[52] A. Sen, *D0 branes on t(n) and matrix theory*, *Adv. Theor. Math. Phys.* **2** (1998) 51–59 [hep-th/9709220].

[53] M. R. Douglas, H. Ooguri and S. H. Shenker, *Issues in (m)atrix model compactification*, *Phys. Lett.* **B402** (1997) 36–42 [hep-th/9702203].

[54] R. C. Myers, *Dielectric-branes*, *JHEP* **12** (1999) 022 [hep-th/9910053].

[55] J. De Boer and K. Schalm, *General covariance of the non-abelian dbi-action*, *JHEP* **02** (2003) 041 [hep-th/0108161].

[56] J. de Boer, E. Gimon, K. Schalm and J. Wijnhout, *Evidence for a gravitational myers effect*, *Annals of Physics* **313/2** (2004) 402–424 [hep-th/0212250].

[57] J. de Boer, K. Schalm and J. Wijnhout, *General covariance of the non-abelian dbi-action: Checks and balances*, *Annals of Physics* **313/2** (2004) 425–445 [hep-th/0310150].

[58] H. Dorn, *Non-abelian gauge field dynamics on matrix d-branes in curved space and two-dimensional sigma-models*, *Fortsch. Phys.* **47** (1999) 151–157 [hep-th/9712057].

[59] H. Dorn and H. J. Otto, *Remarks on t-duality for open strings*, *Nucl. Phys. Proc. Suppl.* **56B** (1997) 30–35 [hep-th/9702018].

[60] M. R. Douglas, *D-branes and matrix theory in curved space*, *Nucl. Phys. Proc. Suppl.* **68** (1998) 381–393 [hep-th/9707228].

[61] I. Taylor, Washington and M. Van Raamsdonk, *Multiple d0-branes in weakly curved backgrounds*, *Nucl. Phys.* **B558** (1999) 63–95 [hep-th/9904095].

[62] S. Mukhi, *The geometric background field method, renormalization and the wess-zumino term in nonlinear sigma models*, *Nucl. Phys.* **B264** (1986) 640.

[63] L. Alvarez-Gaume, D. Z. Freedman and S. Mukhi, *The background field method and the ultraviolet structure of the supersymmetric nonlinear sigma model*, *Ann. Phys.* **134** (1981) 85.

[64] E. Bergshoeff, M. Rakowski and E. Sezgin, *Higher derivative superyang-mills theories*, *Phys. Lett.* **B185** (1987) 371.

[65] M. Cederwall, U. Gran, B. E. W. Nilsson and D. Tsimpis, *Supersymmetric corrections to eleven-dimensional supergravity*, hep-th/0409107.

[66] M. Cederwall, B. E. W. Nilsson and D. Tsimpis, *$D = 10$ super-yang-mills at o(alpha**2)*, *JHEP* **07** (2001) 042 [hep-th/0104236].

[67] E. A. Bergshoeff, A. Bilal, M. de Roo and A. Sevrin, *Supersymmetric non-abelian born-infeld revisited*, *JHEP* **07** (2001) 029 [hep-th/0105274].

[68] A. Sevrin and A. Wijns, *Higher order terms in the non-abelian d-brane effective action and magnetic background fields*, *JHEP* **08** (2003) 059 [hep-th/0306260].

[69] I. Taylor, Washington and M. Van Raamsdonk, *Multiple dp-branes in weak background fields*, *Nucl. Phys.* **B573** (2000) 703–734 [hep-th/9910052].

[70] J. Fuchs, *Affine Lie Algebras and Quantum Groups*. Cambridge Monographs On Mathematical Physics. Cambridge University Press, 1992.

[71] J. Fuchs and C. Schweigert, *Symmetries, Lie algebras and representations : a graduate course for physicists*. Cambridge Monographs On Mathematical Physics. Cambridge University Press, 1997.

[72] G. Bäuerle and E. de Kerf, *Lie algebras: finite and infinite dimensional Lie algebras and applications in physics Part One*. Studies in mathematical physics, vol. 1. Elsevier Science Publishing Company, 1990.

SAMENVATTING

DEEL ÉÉN: SUPERSYMMETRISCHE IJKTHEORIE

Eén van de meest geheimzinnige eigenschappen van de sterke wisselwerking is wel het feit dat de quarks nooit vrij voorkomen, maar altijd in zogenaamde kleurloze combinaties. Quarks zijn als het ware opgesloten. Als de Quantum Chromo Dynamica (QCD) ook daadwerkelijk een theorie van de sterke wisselwerking is, dan zou de opsluiting van de quarks ook met deze theorie beschreven moeten kunnen worden. Tot op de dag van vandaag is dat nog niet geheel gelukt. De reden daarvoor is dat opsluiting plaatsvindt bij lage energieën (typisch lager dan enkele honderden MeV). Juist bij deze lage energieën is de koppelingsconstante van QCD groter dan één. Dit maakt het toepassen van perturbatieve technieken vrij hopeloos. Zelfs al zou het lukken om de volledige reeks van Feynman diagrammen te sommeren, dan nog ontkomt men niet aan het feit dat deze reeks slecht asymptotisch convergeert. De beschrijving van de opsluiting van quarks met behulp van QCD vraagt dan ook om nieuwe niet-perturbatieve methoden.

Er is een theorie, de snaartheorie, die een frisse kijk op dit probleem kan geven. De effectieve veldentheorie voor een D-braan opstelling wordt gegeven door één of andere ijktheorie. Het is mogelijk om op deze manier ijktheorieën te construeren die veel gelijkenis vertonen met een supersymmetrische extensie van QCD. De dynamica van ijktheorieën bij lage energie kan op deze manier worden begrepen door naar de dynamica van de, beter begrepen, D-branen te kijken. Zo kunnen nieuwe inzichten worden verkregen in het mechanisme van de opsluiting van quarks.

Het is zelfs zo dat de snaartheorie kan worden gebruikt als een nieuwe invalshoek om veldentheoretische technieken te genereren. Een goed voorbeeld hiervan is het werk van Dijkgraaf en Vafa. Dit werk laat zien hoe men een exacte niet-perturbatieve superpotentiaal kan afleiden met behulp van een convergerende storingsreeks. Deze methode kan men uiteindelijk ook helemaal in de veldentheorie afleiden, maar hij vindt zijn oorsprong in een reeks dualiteiten toegepast op D-branen op Calabi-Yau variëteiten.

De methode van Dijkgraaf en Vafa komt erop neer dat men, in plaats van een volledige pad-integraal, slechts een gereduceerde matrixintegraal hoeft te berekenen. Er hoeft geen impuls afhankelijkheid te worden meegenomen en alleen de planaire diagrammen moeten worden berekend. Alhoewel dit al een enorme simplificatie is, kan het nog iets gemakkelijker door naar ijktheorieën in drie dimensies te kijken. Seiberg en Witten hebben laten zien dat de lage energie effectieve superpotentiaal in vier dimensies kan worden berekend door naar de drie-dimensionale theorie te kijken die men krijgt door compactificatie op een cirkel. Het blijkt dat in de zo verkregen drie-dimensionale theorie het relatief eenvoudig is om de superpotentiaal te berekenen. Er zijn geen fractionele instantonen of monopolen die bijdragen aan deze superpotentiaal, het is “slechts” een kwestie van het vinden van de juiste variabelen om deze superpotentiaal in uit te drukken.

Voor theorieën met $\mathcal{N} = 2$ supersymmetrie gebroken naar $\mathcal{N} = 1$, door middel van een superpotentiaal, bestaat het sterke vermoeden dat de juiste variabelen worden gegeven door de Lax matrix van een onderliggend integreerbaar systeem. De vacuüm structuur van ijktheorieën met uitgebreide supersymmetrie kan worden gevonden middels een hyperelliptische kromme. Deze kromme kan ook worden gezien als een spectrale kromme van een integreerbaar systeem met een Lax-paar formulering. Precies dit integreerbaar systeem, in dit geval een periodiek Toda systeem, geeft de juiste variabelen om de superpotentiaal in uit te drukken. Na het invullen van de Lax matrix in de superpotentiaal moeten alleen nog de extrema worden berekend. Aangezien de superpotentiaal een polynoom is, is dit een eenvoudige opdracht in vergelijking met het berekenen van een matrix integraal. De verkregen resultaten kunnen worden vergeleken met resultaten verkregen met behulp van de methodes van Seiberg-Witten en Dijkgraaf-Vafa. De drie-dimensionale berekeningen komen uitstekend overeen met de al bekende resultaten, iets wat het vertrouwen in het vermoeden versterkt.

In een poging om het vermoeden te bewijzen is een braan opstelling beschreven. Deze opstelling van NS5-branen met D4-branen daartussen opgehangen, beschrijft een ijktheorie met hypermultipletten, maar zonder superpotentiaal. Deze ijktheorie wordt effectief beschreven door een Hitchin systeem, een integreerbaar systeem gebaseerd op een Riemann oppervlak. In de limiet waarin de hypermultipletten oneindig zwaar worden gaat dit Hitchin systeem over in een periodiek Toda systeem. De adjoint scalar in dit systeem is precies gelijk aan de Lax matrix, een essentieel ingrediënt van het vermoeden. De verwachting is dat het aandraaien van een superpotentiaal deze redenering niet zal ontkrachten.

Het toepassen van een integreerbaar systeem bleek ook nuttig bij het bepalen van de vacua van theorieën met G_2 ijksymmetrie. Hier is het gelukt, naast het verifiëren van bekende resultaten, om de superpotentiaal uit te drukken in het chirale superveld S, iets wat met matrix technieken moeilijk gaat. Bij het uitwerken van een voorstel om de resolvent voor G_2 te berekenen, is een hyperelliptische Seiberg-Witten curve gevonden. Voorheen waren de curves die de vacuüm structuur voor G_2 beschreven van een ingewikkelder structuur dan de hyperelliptische.

Al met al kan worden geconcludeerd dat het toepassen van de Lax matrix, van het onderliggende integreerbaar systeem, resultaten geeft die precies in overeenstemming zijn met de resultaten van andere, vaak complexere, technieken. Vooral het uitdrukken van de superpotentiaal in het chirale superveld S is erg interessant. Het vermoeden bestaat namelijk dat dit veld het lichtste veld is. Berekeningen, waaronder die in dit proefschrift, tonen aan dat dit veld massief is. Dit is dan weer sterk bewijs voor het bestaan van een "mass gap", een van de onbewezen eigenschappen van Yang-Mills theorie. Zo'n "mass gap" wordt ook wel gezien als een noodzakelijke voorwaarde voor het bestaan van de opsluiting van quarks en veldlijnen. Helaas moet ook worden geconstateerd er geen sluitende argumenten zijn die aantonen of het superveld S ook daadwerkelijk het lichtste veld is.

DEEL TWEE: D0-BRANEN

De meest prominente kandidaat voor een allesomvattende theorie die uit de snaartheorie voortkomt is de zogenaamde M-theorie. Aangezien de verschillende snaartheorieën speciale gevallen zijn van M-theorie, hoopt men dat M-theorie een consistente beschrijving van de kwantumgravitatie kan geven. Het is dus de moeite waard om M-theorie te bestuderen.

Een van de manieren om iets over M-theorie te weten te komen is door type IIA snaartheorie te nemen in de limiet waar de snaarkoppeling zeer sterk wordt. In deze limiet wordt de dynamica in de type IIA snaartheorie volledig gedomineerd door de puntvormige D0-branen, soliton-achtige geladen objecten waarop snaren kunnen eindigen. De theorie van deze D0-branen beschrijft M-theorie in het Infinite Momentum Frame (IMF).

Om M-theorie op een gekromde ruimte te beschrijven, kan het dus nuttig zijn om D0-branen in een gekromde achtergrond te beschrijven. Nu is het echter zo dat het erg moeilijk is om een actie voor D0-branen in een gekromde achtergrond te vinden. Waar het eisen van diffeomorfisme invariantie de actie volledig vastlegt in het geval van slechts één D0-braan, is de situatie volledig anders in het geval van meerdere D0-branen. In dat laatste geval krijgt de dynamica namelijk een niet-commutatief karakter waardoor het zelfs niet meer duidelijk is hoe diffeomorfisme invariantie gedefinieerd moet worden.

Om een oplossing te vinden voor dit probleem kan men basispunt-onafhankelijkheid van de actie eisen. Het principe hierachter is eenvoudig. Men splits de velden die de D0-branen beschrijven in een achtergrond deel en een fluctuatie deel. Vervolgens expandeert men de actie in de fluctuatie velden, in een manifest covariante wijze. Het mag natuurlijk geen verschil maken hoe de opdeling in achtergrond en fluctuatie is gekozen, aangezien deze keuze volledig artificieel is. Vervolgens kan men transformaties van de velden definiëren die van de ene keuze naar de andere transformeren. De invariantie onder deze transformaties wordt ook wel basispunt-onafhankelijkheid genoemd.

Deze basispunt-onafhankelijkheid kan worden gezien als een definitie van diffeomorfisme invariantie voor de actie van meerdere D0-branen. In tegenstelling tot de actie voor één D0-braan, legt de basispunt-onafhankelijkheid de actie voor meerdere D0-branen niet volledig vast. Men kan proberen om de actie van één braan te gebruiken en deze generaliseren naar de actie voor meerdere branen. Hier stuit men echter, als direct gevolg van het niet-commutatieve karakter van de dynamica, op problemen met betrekking tot de ordening van de termen. De symmetrische ordening die vaak wordt aangenomen voor D0-branen blijkt alleen consistent te zijn met basispunt-onafhankelijkheid voor lage snelheden en kromming. Wel is het mogelijk om correcties te vinden op de symmetrisch geordende actie, zodanig dat de actie basispunt-onafhankelijk is. Aangezien de correcties niet uniek zijn, is de ordeningsproblematiek nog niet opgelost.

Een belangrijke eigenschap die elke actie voor branen, in een gekromde ruimte met een isometrie, moet hebben is T-dualiteit. Een voorstel voor de potentiaal in een gekromde achtergrond, die zodanig is geconstrueerd dat deze basispunt-onafhankelijk is, blijkt inderdaad consistent te zijn met T-dualiteit.

D0-branen op een negatief gekromde ruimte blijken eigenschappen te vertonen die doen denken aan een gravitationeel Myers effect. Het Myers effect is een bijzondere eigenschap van branen, waarbij een collectie van branen gevoelig blijkt te zijn voor een veld waaraan een enkel braan niet koppelt. De D0-branen, die normaal gesproken alleen koppelen aan de RR één-vorm, kunnen zo onder de invloed van een RR vier-vorm worden opgeblazen in een bol-achtige configuratie. Deze bol-achtige configuratie heeft een lagere energie dan wanneer men alle D0-branen op elkaar zet. Het bestaan van deze configuratie met lagere energie wordt gekenmerkt door het bestaan van tachyonen in het spectrum van kwadratische fluctuaties. Precies dit kenmerk kan ook worden gevonden in de theorie van bewegende D0-branen op een negatief gekromde achtergrond. Dit is geen bewijs voor het bestaan van zo'n lagere energie configuratie (de tachyonen wijzen op een instabiliteit, maar zeggen niets over het bestaan van een *stabiele* configuratie met lagere energie), maar het geeft zeker een reden om D0-branen in een gekromde achtergrond verder te bestuderen.

Dankwoord

Een proefschrift schrijven doe je niet alleen. Dat wil zeggen, het schrijven zelf wel, maar het tot stand komen van de inhoud ervan zeker niet. Zonder de ideeën en kennis van mijn promotor zou ik nooit in staat zijn geweest een proefschrift als dit te schrijven. Jan, dank voor al je hulp en de vrijheid die je mij hebt gegeven om mezelf te ontwikkelen. Voor jouw scherpe geest en uitgebreide kennis over dit gevarieerde vakgebied heb ik een grote bewondering.

Natuurlijk wil ik ook mijn kamergenoten en collega's bedanken. Robert, bedankt voor de vele nuttige discussies over snaren en vele andere dingen (en ook bedankt voor het gebruiken en testen van Kile!). Je bent een goede vriend en ik hoop dat je je draai kan vinden in North Carolina. Rutger, het was plezierig om op het laatst nog een kamer met je te delen. Ik denk dat we in die tijd veel aan elkaar hebben gehad.

De artikelen zijn tot stand gekomen door een vruchtbare samenwerking met onder andere Koenraad Schalm, Mohsen Alishahiha, Amir Mosaffa en natuurlijk mijn Amsterdamse collega's Jan, Robert en Rutger. Koenraad, jouw oog voor detail en nauwgezette manier van werken waren een voorbeeld. Mohsen and Amir, thank you very much for a pleasant long-distance collaboration.

Verder wil ik ook alle andere collega's op het instituut bedanken voor de prettige werksfeer.

Tenslotte nog enkele persoonlijke woorden. Allereerst speciale dank aan mijn paranimfen Mattijn en Stef. Aan mijn ouders ben ik veel dank verschuldigd vanwege hun onvoorwaardelijke steun en stimulerende "can do" instelling. Marloes, mijn lieve vriendin, bedankt voor alle steun en liefde die je mij de afgelopen jaren hebt gegeven. Je was, bent en blijft een bron van inspiratie.

Curriculum Vitae

Jeroen Sebastian Wijnhout was born in Raalte, the Netherlands on 30 August 1976. After graduating from the "Florens Radewijns College" in Raalte in the year 1995 he started his applied physics studies at the Universiteit Twente in the same year. The last eighteen months of these studies were spent at the Spinoza Instituut for theoretical physics at the Universiteit Utrecht, where he wrote his master thesis on "D0-branes, a consistency check of the definition of M-theory" under the supervision of prof. dr B. de Wit. Graduation, cum laude, followed in 2001.

In 2001 he accepted a Ph.D. position, at the Institute for Theoretical Physics in Amsterdam, to conduct research in string theory under the supervision of prof. dr J. de Boer. This research resulted in five publications on supersymmetric gauge theories and on D0-branes in a gravitational background. This thesis is a detailed description of the authors contribution to those publications. In addition to conducting research, the author also assisted in teaching several courses and attended a number of workshops and schools in the Netherlands and abroad.