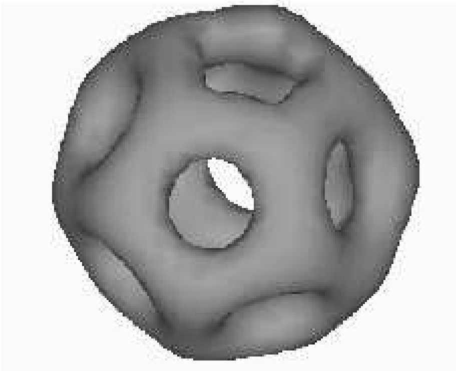
Modelling the Dynamics of BPS Monopoles

J. K. Slingerland

September 30, 1998

Supervisor: Erik Verlinde



Universiteit Utrecht, Report number:ITF-UU-98/08

The figure on the front cover was taken from [1]. It represents an energy density surface for a dodecahedral 7-monopole

Abstract

This thesis presents several aspects of the theory of BPS-monopoles. In particular, it deals with Nahm's construction of static monopoles and with the description of monopole dynamics through two different methods. The first of these is Manton's geodesic approximation, the second is an approach in which the monopoles are treated as point particles. It is shown how these methods are related and how they complement each other. To make this thesis more self-contained, I have included an introduction to the soliton techniques which are on the basis of the geodesic approximation and an elementary treatment of the ADHM-construction of instantons, which is on the basis of Nahm's construction for monopoles.

Contents

Preface

Motivation and History

This thesis is about magnetic monopoles, i.e. particles with magnetic charge. Despite the fact that such particles have never been observed, their properties have been studied theoretically over a period of more than a century and searches for monopoles have been conducted in accelerators each time a new energy regime was opened up to experiment. In this section, I will give a very brief overview of the history of the subject and try to indicate the main reasons for the continuing interest in monopoles.

The history of the study of monopoles can be divided into two parts, separated by the year 1931. Before 1931, there were no physical arguments for the existence of monopoles in nature. Still, there was some mathematical interest in magnetic monopoles, because the introduction of magnetic charges can make Maxwell's equations more symmetrical. It is easy to check that Maxwell's equations for the vacuum are invariant when we apply the following transformation to the fields.

$$\mathbf{E} \to -\mathbf{B}, \qquad \mathbf{B} \to \mathbf{E}$$

This symmetry is called electric/magnetic duality. It is broken when one introduces electric charges. However, it can be restored by the introduction of magnetic charges. When an electric current density j_μ and a magnetic current density j^m_μ are present, Maxwells equations take the form

$$\begin{array}{rcl} \nabla \cdot \mathbf{B} &=& j^m_0 \\ \nabla \cdot \mathbf{E} &=& j_0 \end{array} \qquad \begin{array}{rcl} (\nabla \times \mathbf{E})_i + \partial_t \mathbf{B}_i &=& -j^m_i \\ (\nabla \times \mathbf{B})_i - \partial_t \mathbf{E}_i &=& j_i \end{array} \tag{0.1}$$

and it is easy to check that this is invariant under the generalized electric/magnetic duality transformation given by

$$\begin{array}{rcl} \mathbf{E} &\to& -\mathbf{B} \\ j_\mu &\to& -j^m_\mu \end{array} \qquad \begin{array}{rcl} \mathbf{B} &\to& \mathbf{E} \\ j^m_\mu &\to& j_\mu \end{array} \tag{0.2}$$

In 1931, Dirac published an article in which he argued that the existence of a monopole somewhere in the universe would explain the quantization of electric charge [2]. This article excited a lot of interest in monopoles among physicists, since no other explanation for the quantization of electric charge was (and is) available. However, because monopoles were not found in experiments, and because there were mathematical difficultites with Dirac's monopoles, interest gradually died down again until, in 1974, 't Hooft and Polyakov independently found that monopoles appear naturally as regular, stable classical solutions in gauge theories whose gauge group has a compact covering [3, 4]. Although this class does not include the standard model, it is certainly conceivable that grand unified theories will include these 't Hooft-Polyakov monopoles.

Shortly after 't Hooft and Polyakov's discovery, it was conjectured that there might be a generalized electric/magnetic duality in certain theories, which would make it possible to study the properties of the gauge bosons of the theory at large coupling by looking at the properties of monopoles at small coupling [5, 6]. One could then translate the result back to the gauge particle picture by applying a duality transformation. This new way of studying the properties of gauge particles would be especially useful in the study of the strong force, because it would make it possible to study important non perturbative processes like quark confinement and hadron formation by doing perturbation theory in the dual picture.

An important step in this direction was made in 1994, when Seiberg and Witten found an exact low energy effective action for $N = 2$ supersymmetric Yang-Mills theory, a distant relative of QCD, assuming electric/magnetic duality [7]. Since then, the study of monopoles has once more come to the center of attention and monopoles and their dynamics are now a very active field of research.

Content of this thesis

In this thesis, we will not be very concerned about duality, but concentrate mostly on the dynamics of magnetic monopoles and especially on how to model these. The most important chapter of this thesis is therefore the last chapter, chapter 4, in which we treat two different models which can be used to describe monopole dynamics.

The first of these is the so called geodesic approximation. In this approach, monopole motion is described in terms of geodesic motion on a *moduli space,* a manifold of gauge equivalence classes of static monopole solutions. The geodesic approximation describes the low energy dynamics of monopoles of like magnetic charges very well, but it cannot describe situations that involve monopoles of opposite magnetic charges. Also, it does not incorporate radiation effects. The second model is an approximation which treats the monopoles as point particles. This model can incorporate monopoles of opposite magnetic charges and radiation effects, but it works only if the monopoles are far apart, which is a limitation that does not apply to the geodesic approximation.

The earlier chapters of this thesis introduce the necessary mathematical tools and treat some properties of static monopoles.

- Chapter 1 treats the basic techniques of soliton physics which we need later on. In particular, it explains the idea behind the geodesic approximation in an elementary setting.

- Chapter 2 is a brief description of the ADHM-construction of all self-dual instantons in Yang-Mills theory. This forms the necessary background for the description of Nahm's construction for monopoles in chapter 3. This chapter also serves the purpose of introducing Yang-Mills theory.

- Chapter 3 starts with an introduction of the Georgi-Glashow model, the model in which we want to study monopoles. It is shown how 't Hooft-Polyakov monopoles can appear in this model and how exact classical solutions which represent static monopole configurations can be obtained in the so called BPS-limit of this model. In particular, the Nahm construction which (in principle) yields all these solutions is treated in some detail and it is shown that, for magnetic charge k, there is a $4k-1$ dimensional family of static monopole solutions, modulo gauge transformations. This is an important result, because it allows us to conclude (in chapter 4) that the dimensions of the moduli spaces used in the geodesic approximation have to be $4k$.

Acknowledgements

Before continuing, I would like to thank my supervisor, Erik Verlinde. I would also like to thank my fellow students Luuk Hoevenaars, Hessel Posthuma, Alexander Heimel, Ivo Savonije, Jeroen Wiemer and Fieke Dekkers for creating a nice atmosphere and for many useful and interesting discussions. Finally, I would like to thank my parents and my girlfriend Theresa for their support.

Chapter 1

Solitons

Throughout this thesis, solitons will play a very important role. This chapter will start with a review of some of the basic concepts and techniques of soliton physics and culminate in the description of a method to describe the low energy dynamics of certain solitons, the so called geodesic approximation. Much of the material in the earlier part of this chapter is taken from [8], which is a good introduction to the subject.

1.1 Solitons: What are they ?

The word soliton has been used several times now and you may be wondering what we mean with it. This is actually not an easy question to answer. There are two reasons for this. First, the word soliton is used both for a certain kind of mathematical object and for a certain physical particles whose description involves these mathematical objects. Second, even the definition of a soliton as a mathematical object varies according to the context in which the word is used, so that the mathematical definitions in the literature (see e.g. [8]) are usually too restrictive to be of general use.

The only really general mathematical statement one can make about a soliton, is that it is a regular solution to the classical equations of motion for some field theory which keeps the energy for this theory finite.

If we were doing classical physics, it would be easy to see how such a solution could correspond to a physical particle. The finiteness of the energy implies that the energy density of a soliton will go to zero at spatial infinity. We can thus hope that, for any fixed time, we can find some area of space where (most of) the energy of the soliton is localised. A logical interpretation would then be that this area is occupied by a particle. Thus, a sufficiently "nice" soliton could be interpreted as a finite-sized particle moving through space.

Of course, we really want to do quantum physics and this makes the interpretation of the mathematical soliton as a particle much more difficult. Still, even in quantum physics, the term soliton refers to a finite-sized particle related to a classical finite energy solution. Rather than go into details about the nature of this relationship now, let us just start to do some calculations that involve solitons with this "particle picture" in mind. A more thorough understanding should then develop as we go along.

1.2 Solitons in 1+1 dimensions; the kink model

To get a feeling for soliton physics, let us consider a toy model with one space and one time dimension, governed by the following Lagrangian:

$$L(t) = \int dx \left\{ \frac{1}{2}(\partial_t \phi)^2 - \frac{1}{2}(\partial_x \phi)^2 - U(\phi) \right\} \tag{1.1}$$

In this equation, we take U to be a potential with degenerate, discretely located minima. The value of U at these minima is taken to be zero and we will call the location of the i^{th} minimum $g^{(i)}$. The equation of motion which follows from (1.1) is

$$\partial_t^2 \phi - \partial_x^2 \phi = -\frac{\partial U}{\partial \phi} \tag{1.2}$$

and the energy is given by

$$E = \int dx \left\{ \frac{1}{2}(\partial_t \phi)^2 + \frac{1}{2}(\partial_x \phi)^2 + U \right\}. \tag{1.3}$$

That is, we want to find solutions to the equation of motion (1.2) which keep the energy $\mathcal{E}$ finite. Now we see that the energy density is the sum of three positive terms and hence each of these terms has to be integrable to make the energy converge. Imposing this condition on the first two terms in (1.3) gives us the boundary condition that $\partial_t \phi$ and $\partial_x \phi$ have to go to zero at spatial infinity. The third term in the energy density is the potential U. Since we have taken this to be larger than or equal to zero, it is clear that, in order for the energy to converge, the field configuration $\phi(x)$ has to be such that $U(\phi(x))$ goes to zero as x goes to plus or minus infinity. This gives us boundary conditions for the field ϕ itself. To keep the energy finite, $\phi(x)$ has to approach one of the values $g^{(i)}$ as x goes to infinity and another of the $g^{(i)}$ (possibly the same) as x goes to minus infinity.

Thinking from a "particle perspective," it makes sense to restrict ourselves to time independent solitons at the moment, since these would correspond to particles at rest. Also, once we have a time independent soliton, we can turn it into a freely moving soliton by applying a Lorentz boost. Of course the opposite is also true, so when we find all static solitons, we will effectively find all freely moving solitons in the process.

When we restrict to time independent configurations, the equation of motion is also much simplified; it reduces to

$$\partial_x^2 \phi = \frac{\partial U}{\partial \phi} \tag{1.4}$$

If we temporarily reinterpret the space coordinate x by choosing to think of it as a time coordinate, then we see that this equation is just the equation of motion for a particle moving in one space dimension (with coordinate ϕ), moving in a potential $-U$. The boundary condition that ϕ has to approach one of the $g^{(i)}$ as x goes to plus or minus infinity can also be fit into this picture. It just says that the particle moves from a certain asymptotic location $g^{(i)}$, which it used to occupy infinitely long ago (as x goes to minus infinity) to an asymptotic location $g^{(j)}$ which it will occupy in the far, far future (as x goes to infinity). When this condition is satisfied, then, in our t-independent setting, the boundary condition on $\partial_x\phi$ is also satisfied. The condition on $\partial_t\phi$ is of course trivially satisfied in the t-independent case.

If the potential U has a unique minimum $g^{(0)}$, then it is easy to see that the only solution to this system is given by $\phi(x) = g^{(0)}$. For a potential with several (degenerate) minima however, more interesting solutions are also possible. At very negative x, ϕ can start off at the location of one minimum of U, only to approach at the location of another minimum for very large x; our 'particle' can now move between two minima.

For our toy model we can in fact study this phenomenon by solving the equation of motion analytically. Multiplying both sides of (1.4) by $\partial_x\phi$ and integrating, we see that

$$\partial_x\phi = \pm\sqrt{2U(\phi)} \tag{1.5}$$

If ϕ is an invertible function of x, then this in turn implies

$$\frac{\partial x}{\partial\phi} = \pm\frac{1}{\sqrt{2U(\phi)}} \tag{1.6}$$

Note that the condition that $\phi(x)$ is invertible will of course never be satisfied for the trivial solutions to the equations of motion which are given by $\phi(x) = g^{(i)}$ for some i. Therefore we can not expect to find these by the method we are describing here. This method is obviously only suited for the study of nontrivial soliton solutions. Integrating (1.6), we obtain

$$x - x_0 = \pm\int_{\phi(x_0)}^{\phi(x)} d\psi \left\{\frac{1}{\sqrt{2U(\psi)}}\right\} \tag{1.7}$$

where, in this equation, we may still choose x_0 and $\phi(x_0)$ to suit our needs. We can now choose a potential, substitute it into the equation above and try to solve for ϕ.

As we are interested in potentials with multiple minima, we take

$$U(\phi) = \frac{1}{4}\lambda\left(\phi^2 - \frac{m^2}{\lambda}\right)^2$$

This has potential has minima at $\phi = \pm\frac{m}{\sqrt{\lambda}}$. Substituting U in (1.7) and choosing $\phi(x_0)$ equal to zero gives

$$x - x_0 = \pm\frac{1}{\sqrt{2}m}\log\left|\frac{m/\sqrt{\lambda} - \phi}{m/\sqrt{\lambda} - \phi}\right|$$

which can be inverted to give

$$\phi(x) = \pm \frac{m}{\sqrt{\lambda}} \tanh\left(\frac{m}{\sqrt{2}}(x - x_0)\right) \tag{1.8}$$

These solutions of (1.4) do indeed approach one of the minima of the potential when x goes to infinity and the other one when x goes to minus infinity. In figure 1.1(a), we have plot the graph of a solution with $x_0 = 0$. The overall shape of this graph has provided the solutions given by (1.8) with the name "kinks". Figure 1.1(b) shows a plot of the energy density of the kink with $x_0 = 0$. We can see clearly that the mass/energy of this kink is concentrated around $x = 0$, as we would hope.

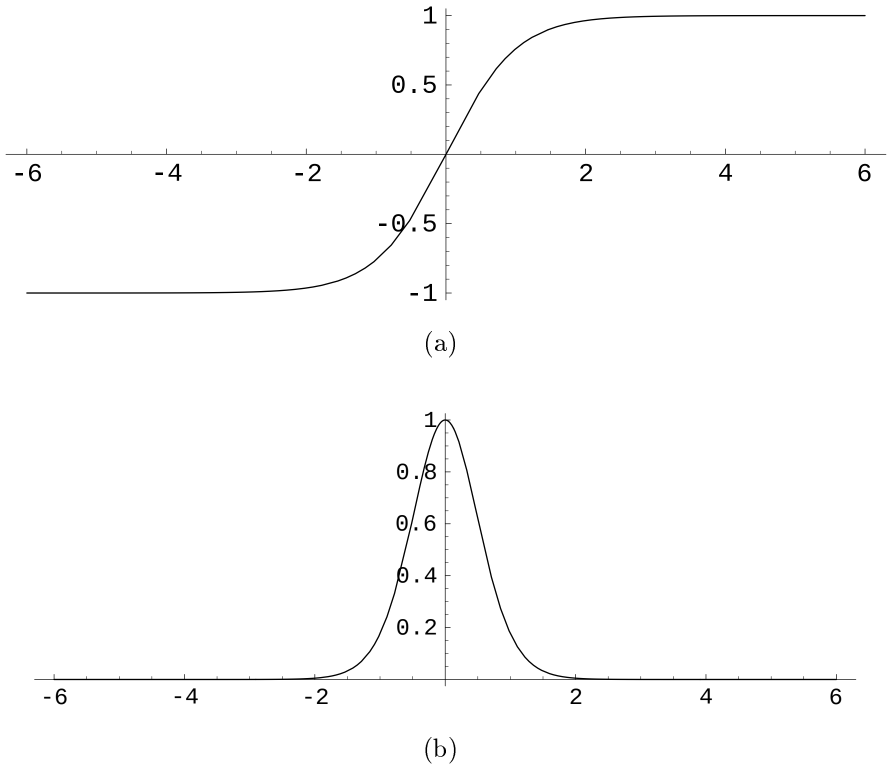


Figure 1.1: (a) graph of the kink solution
(b) graph of the kink's energy density

The kink model is very simple, but it has some interesting features which have analogues in the more complicated models we will study later. One of these is the way we derived boundary conditions on solutions to the static equations of motion through the requirement of minimal energy. In the following sections, we will often come back to the kink model when we need an explicit example of a general feature of soliton physics.

Before we go on, note as an aside that if we generalise the models studied in this section by adding some more scalar fields, but keep working in one space dimension, we can still use the analogy of a particle moving in a potential $-U$ (of course, the particle will now be moving in a higher dimensional space). These generalised models are already far more difficult and no general solution like (1.7) is known for these. A technique which can often still provide some solutions is to use trial functions with a number of free parameters to determine the trajectory of the 'particle' and solve for the time dependence afterwards. See [8] for details.

1.3 Topological Sectors

In the previous section, we noted that a finite energy solution to the equations of motion has to approach one of the minima of the potential at the edge of space. More explicitly, for every solitonic solution $\phi(x)$ to the equations of motion (1.2) we had to have

$$\begin{array}{ll} \phi(x) \to g^{(i)} & (x \to +\infty) \\ \phi(x) \to g^{(j)} & (x \to -\infty) \end{array}$$

for some i, j. If we now have another solution ψ which does not have the same boundary conditions at infinity, then it is clear that it will be impossible to deform this solution continuously into ϕ without making the energy become infinite somewhere in the process. Thus, if we allow only deformations which leave the energy finite, then solutions with different boundary conditions are topologically different. We can now classify the solutions into *topological sectors* by the possible pairs of boundary values and we call these values the *topological indices* of a solution. Because time evolution of a solution is just another form of continuous deformation (and because it obviously leaves the energy finite), a solution always remains in the same topological sector, so we can say that topological indices are preserved quantities.

In higher dimensional theories, one often still has topological sectors into which field configurations fall according to their boundary conditions. However, there is usually no clear equivalent of the topological indices, because the edge of space no longer consist of a finite number of points. In stead, one often has conserved quantities called *topological charges,* which label field configurations which are topologically distinct. These charges usually have corresponding conserved currents. For our $1+1$-dimensional system, we can also define such a topological charge Q, with a corresponding conserved current k. We can take

$$\begin{aligned} Q &= \phi(\infty) - \phi(-\infty) \\ k^{\mu} &= \epsilon^{\mu\nu}\partial_{\nu}\phi \end{aligned}$$

Using the antisymmetry of the epsilon tensor, one easily verifies that we have

$$\partial_{\mu}k^{\mu} = 0,$$

so k is indeed a conserved current. Also, we have

$$Q = \int_{-\infty}^{\infty} k_0 \, dx,$$

from which we see that Q is indeed the charge corresponding to this current.

Of course, for one dimensional models, the topological charge is not very useful, as the topological indices already contain all the information that could be gained from it (and more). Still, I thought it would be useful to introduce the charge Q here, because it is the simplest example of a topological charge I know. We will see more interesting topological charges later.

1.4 Quantisation of Static Solitons; Particle Interpretation

In this section, we will go into the quantisation of static solitons and in particular into that of the kink. We will also try to give the kink solution an interpretation in terms of particles.

To quantise static solitons, we use a semi-classical approximation to the full quantum field theory. I will first give a short review of this method for theories with a finite number of coordinates and then generalise to field theory, where the number of coordinates is infinite.

Let us start by looking at the theory with Lagrangian given by

$$L = \sum_{i=1}^{n} (\partial_t \phi^i)^2 - V(\phi)$$

Static classical solutions for this theory are given by the coordinate vectors ϕ for which the potential energy V has a stationary point. In order to get a stable solution, this stationary point has to be a (local) minimum. Around such a minimum, attained, say, for $\phi = \tilde{\phi}$ the potential may be expanded as follows:

$$V(\phi) = V(\tilde{\phi}) + (\phi^i - \tilde{\phi}^i)\partial_i\partial_j V(\tilde{\phi})(\phi^j - \tilde{\phi}^j) + O((\phi - \tilde{\phi})^3) \tag{1.9}$$

If we now change coordinates to a system in which the matrix $\partial_i\partial_j V(\tilde{\phi})$ is diagonal and the origin is in $\tilde{\phi}$, then we can rewrite the former expansion as

$$V(\phi) = V(\tilde{\phi}) + \omega_i^2 \eta_i^2 + O(\eta^3) \tag{1.10}$$

where the η^i are the new coordinates and the ω_i^2 are the eigenvalues of $\partial_i\partial_j V(\tilde{\phi})$, which we all know to be positive, because U has a minimum in $\phi = \tilde{\phi}$. We see that, to second order in η, this is just the potential for a system of non coupled harmonic oscillators with frequencies ω_i. This can be solved exactly, giving the well known energy levels

$$E_{\vec{n}} = V(\tilde{\phi}) + \hbar \sum_{i=1}^{n} \left(n_i + \frac{1}{2}\right) \omega_i$$

We may now hope that the third and higher order terms in the expansion are small as compared to the second order term and if they are, we can treat them as perturbations.

For field theory, the situation is exactly analogous, but for the fact that the number of "coordinates" goes to infinity. For the Lagrangian (1.1), the potential energy is given by

$$V[\phi] = \int dx \left\{ \frac{1}{2}(\partial_x \phi)^2 + U(\phi) \right\}$$

again, static classical solutions for this system are given by "coordinate vectors" (i.e. functions) $\phi(x)$ for which the potential energy is stationary. If $\tilde{\phi}(x)$ is such a function then, using partial integration, we can expand the potential energy as

$$V[\phi] = V[\tilde{\phi}] + \int dx \left\{ (\phi - \tilde{\phi}) \left[-(\partial_x)^2 + \left(\frac{\partial^2 U}{\partial \phi^2} \right)_{\tilde{\phi}} \right] (\phi - \tilde{\phi}) \right\} + O((\phi - \tilde{\phi})^3) \qquad (1.11)$$

This expansion is the field theoretical analogue of the expansion (1.9). The second term in (1.11) is just the second order change in $V[\phi]$ as the field is changed from $\tilde{\phi}$ to ϕ. The first order term in $\phi - \tilde{\phi}$ is zero, because the configuration ϕ we are expanding around is a stationary point of the potential energy functional V.

To reach the analogue of (1.10), we now need to solve the following eigenvalue equation:

$$\left[-(\partial_x)^2 + \left(\frac{\partial^2 U}{\partial \phi^2} \right)_{\tilde{\phi}} \right] \eta = \omega^2 \eta \qquad (1.12)$$

This will result in eigenvalues ω_i and orthonormal eigenfunctions or "normal modes" η_i, where i is some index variable. If we now write $\phi - \tilde{\phi} = \sum_i c_i(t)\eta_i(x)$ then we can rewrite the expansion (1.11) as

$$V[\phi] = V[\tilde{\phi}] + \frac{1}{2} \sum_i \omega_i^2 c_i^2 + O(c^3)$$

using orthonormality of the η_i. This formula corresponds to (1.10). The Lagrangian (1.1), too, can now be given in terms of the c_i. We have

$$L = -V[\tilde{\phi}] + \frac{1}{2} \sum_i \left\{ (\dot{c}_i(t))^2 - \omega_i^2 c_i^2 \right\} + O(c^3) \qquad (1.13)$$

Again, we see that, to second order, this is just a sum of non coupled harmonic oscillators, be it an infinite sum this time. Solving this system, we get exactly the same expression as before for the energy levels:

$$E_{\vec{n}} = V[\tilde{\phi}] + \hbar \sum_i (n_i + \frac{1}{2})\omega_i$$

Note that, because of the infinite summation over i, we have infinite ground state energy. This is a normal situation in field theory and no cause for worry as we can just "shift the energy scale by an infinite constant" to make the energy of physical states become finite again.

As an explicit example, we will now apply the method we just sketched to the kink model. For this model, the potential energy is

$$V[\phi] = \int dx \left\{ \frac{1}{2}(\partial_x\phi)^2 + \frac{\lambda}{4}\left(\phi^2 - \frac{m^2}{\lambda}\right)^2 \right\}$$

Before studying the actual kink solution, it will be useful to study the semi-classical quantisation of one of the two trivial field configurations which correspond to the absolute minima of the potential. Let us take $\phi_1(x,t) = m/\sqrt{\lambda}$ (quantisation of $-\phi_1$ is of course exactly analogous). Writing $\psi = \phi - \phi_1$, the expansion of the potential becomes:

$$V[\phi] = \int dx \left\{ \frac{1}{2}\psi\left(-(\partial_x)^2 + 2m^2\right)\psi + m\sqrt{\lambda}\psi^3 + \frac{1}{4}\lambda\psi^4 \right\}$$

We see that, for small λ, the cubic and quartic terms will be small compared to the quadratic terms and thus we can hope to treat them by perturbation. Here we will only solve the quadratic problem which remains if the cubic and quartic terms are ignored. This means that the next thing to do is solving the eigenvalue equation (1.12). For the present case, this reduces to

$$\left(-(\partial_x)^2 + 2m^2\right)\eta = \omega^2\eta$$

This results in a continuous spectrum of eigenvalues ω_k^2 with corresponding eigenfunctions η_k given by

$$\begin{aligned} \omega_k^2 &= k^2 + 2m^2 \\ \eta_k &= e^{ikx} \end{aligned} \tag{1.14}$$

In this approximation the energy is equal to

$$E_n = \int_0^\infty dk \left\{ \left(n(k) + \frac{1}{2}\right)\hbar\sqrt{k^2 + 2m^2} \right\}$$

We see that this corresponds to the energy of an ensemble of free particles with mass $\sqrt{2}m\hbar$, $n(k)$ of which have momentum between k and $k + dk$. These particles are the elementary quanta of this theory and we will call them mesons. The state that has $n(k) = 0$ for all k can be interpreted as a state with no particles present whatsoever and is therefore identified with the vacuum.

Our next step is to apply the semi-classical approximation to the kink solution and we will see that this yields a second kind of particle, which is not present in the perturbations of the vacuum and which interacts with the mesons.

For our kink, we take the solution with positive sign and $x_0 = 0$ in (1.8) and denote it ϕ_k, that is, we define

$$\phi_k(x) := \frac{m}{\sqrt{\lambda}} \tanh\left(\frac{m}{\sqrt{2}}(x)\right) \tag{1.15}$$

Expanding V around ϕ_k and writing $\psi = \phi - \phi_k$, we get

$$V(\phi) = V(\phi_k) + \int dx \frac{1}{2}\psi\left(-(\partial_x)^2 - m^2 + 3\lambda\phi_k^2\right)\psi + \lambda(\phi_k\psi^3 + \frac{1}{4}\psi^4)$$

Again, we retain only the quadratic terms and solve the corresponding problem. This means solving the eigenvalue equation

$$\left[-(\partial_x)^2 - m^2 + 3m^2\tanh^2(\frac{mx}{\sqrt{2}})\right]\eta(x) = \omega^2\eta(x)$$

or, in terms of $z = mx/\sqrt{2}$:

$$\left[-\frac{1}{2}(\partial_z)^2 + 3\tanh^2(z) - 1\right]\tilde{\eta}(z) = \frac{\omega^2}{m^2}\tilde{\eta}(z) \tag{1.16}$$

with $\tilde{\eta}(z) = \eta(x)$. The solution to this equation can be found in Morse and Feschbach [9]. In order of ascending magnitude, there are two discrete eigenvalues, followed by a continuum. We shall name the discrete eigenvalues $\omega_{d,0}^2$ and $\omega_{d,1}^2$ to distinguish them from the continuum eigenvalues, which will be called ω_q^2, with $q \in [0,\infty)$. The same notation is used for the eigenfunctions. We have

$$\begin{array}{rcll} \omega_{d,0}^2 &=& 0, & \text{with} \quad \tilde{\eta}_{d,0}(z) = 1/\cosh^2(z) \\ \omega_{d,1}^2 &=& \frac{3}{2}m^2, & \text{with} \quad \tilde{\eta}_{d,1}(z) = \sinh(z)/\cosh^2(z) \\ \omega_q^2 &=& m^2(2+\frac{1}{2}q^2), & \text{with} \quad \tilde{\eta}_q(z) = e^{iqz}(3\tanh^2(z) - 1 - q^2 - 3iq\tanh(z)) \end{array}$$

The first thing we notice in this spectrum is the presence of a mode with zero frequency. This is not surprising, it is just a consequence of the fact that the potential energy $V(\phi)$ is invariant under spatial translations, while the kink solution ϕ_k is not. We see that the zero mode $\eta_{d,0}$ is proportional to the derivative of ϕ_k and thus proportional to the variation of ϕ_k when it is shifted along the x-axis by an infinitesimal amount, confirming that this is the zero mode corresponding to translation invariance.

Though not surprising, the presence of a zero mode does present us with a problem; it is unlikely that, with the "spring constant" $\omega_{d,0}$ equal to zero, the wave function of the kink system will remain localised around the kink solution. Instead, it will probably tend to spread in the direction of $\eta_{d,0}$, making our approximation invalid. The zero mode will also cause computational difficulties; higher order terms in the perturbation series will have the zero frequency in their denominators. It is obvious that we should really develop a formalism for dealing with zero modes before we go on. We will do this in the next section. However, for now, let us assume that the wave function is localised around the kink at $t = 0$ and that the spreading in the $\eta_{d,0}$-direction takes place very slowly. At least on a short time scale we should then be able to ignore the effect of the zero mode on the energy of the system and carry on our analysis as if nothing had happened.

The equation for the energy that follows from (1.4) is

$$E_n = \frac{2\sqrt{2}m^3}{3\lambda} + \left(n_{d,1} + \frac{1}{2}\right)\hbar\sqrt{\frac{3}{2}}m + \sum_q \left\{ \left(n(q) + \frac{1}{2}\right)\hbar m\sqrt{\frac{1}{2} + 2q^2} \right\} \tag{1.17}$$

Where the first term is the classical kink energy and the second and third terms are the quantum corrections. I have kept the summation sign, because writing an integration would introduce an artificial difference in the way the discrete and continuum modes are treated. If one wants to make this (divergent) formula at least formally correct, a good option is to restrict the theory to a finite volume and introduce periodic boundary conditions, thus discretising the q-spectrum.

I will now give a particle interpretation of these results as promised. First of all, note that the lowest energy state in the kink spectrum can not be identified as a vacuum. There are at least two good reasons for this. First, the kink state is not translation invariant. Second, we already identified the lowest energy states located around the two absolute minima of the potential as vacua; those two states obviously have lower energy than all other states, including the lowest energy state in the kink spectrum. If the kink state of lowest energy is not a vacuum, there must be one or more particles present. If there were more than one particle, we would expect several quantum numbers characterising the particles' momenta to appear in the energy. Also, if the particles were far enough apart for their interactions to be negligible, we would expect to find a several zero modes, one for a shift in the position of each particle. As we don't see any of this naturally arise, we assume we are dealing with one particle at rest. Of course, in principle there could also be several particles at rest near each other. This would then most likely be some kind of bound state, which is the same as one particle for our purposes. Calculation of the mass of the kink state [10] shows that it is equal to $2\sqrt{2}m^3/3\lambda$ to order $\hbar\lambda^0$, which, for small λ, is much larger than the mass of a meson. This shows that we really have a new kind of particle here. We also see that the presence of this species of particle could never have been derived in perturbation theory; the mass tends to to infinity as the perturbation constant tends to zero.

Now we have established the kink as a kind of extended particle, we can interpret the excitations of the $\omega_{d,1}$-mode as excited states of this particle. All we have to do now is give an interpretation to the continuum modes. This can be done by looking at the asymptotic behaviour of the functions $\tilde{\eta}_q$. We see that

$$\tilde{\eta}_q(z) \to e^{i\left(qz \pm \frac{1}{2}\delta(q)\right)} \qquad (z \to \pm\infty)$$

with

$$\delta(q) = -2\arg(3q/(2-q^2))$$

Remembering that $z = mx/\sqrt{2}$ and $\tilde{\eta}_q(z) = \eta_q(x)$, we see that, for large $|x|$, $\eta_q(x)$ is just the wave function for a freely moving meson with momentum $\hbar mq/\sqrt{2}$. Somewhere on its "path" from $x = -\infty$ to $x = +\infty$, this meson has its phase shifted by an amount $\delta(q)$. This phase shift has to be due to some kind of interaction. In our calculation of the energy levels for the vacuum sector, the mesons did not interact with each other. Therefore, it is most likely that the phase shift is due to an interaction with the kink particle located around the origin. Thus, we can interpret the continuum modes as states of mesons that scatter off the kink.

1.5 Treatment of Zero Modes

In this section, I will deal with the treatment of zero modes in the canonical quantisation formalism. Most of this section is based on the treatment by Christ and Lee [11]. Some remarks on the validity of their method and on finding zero modes are added.

1.5.1 Finding Zero Modes

In section 1.4, we introduced zero modes. For clarity, I will start this paragraph with a definition of a zero mode.

Definition 1 *Suppose we have a solution $\tilde{\phi}$ to the equations of motion for some theory. A* zero mode *of $\tilde{\phi}$ is then a time independent deformation η of this solution such that for the potential energy V of the theory, one has $V[\tilde{\phi} + \lambda\eta] = V[\tilde{\phi}] + O(\lambda^3)$ as λ approaches zero.*

Notice that, because the zero mode is time independent, it will not only keep the potential energy fixed to second order, but also the total energy and the action.

Now when we have a classical solution $\tilde{\phi}$, we may find its zero modes straight from the above definition. To do that, we have to deform $\tilde{phi}$ to $\tilde{\phi} + \delta\phi$, write down the second order variation of the potential energy and somehow find out which deformations of $\tilde{\phi}$ make this second order variation equal to zero. For the Lagrangian (1.1), we have done exactly this in section (1.4) and this led to the result that the zero modes were just the solutions to the equation

$$\left[-(\partial_x)^2 + \left(\frac{\partial^2 U}{\partial \phi^2}\right)_{\tilde{\phi}}\right]\eta = 0$$

This equation is just equation (1.12) with the eigenvalue ω^2 set to zero. Now we note the following: The above equation is just the linearisation of the equation of motion (1.2) for this model. That is, the above equation is exactly the equation that the time independent deformation η has to satisfy to make sure that the deformed solution $\tilde{\phi} + \eta$ is still a solution to the equation of motion (1.2) to first order in η. I will now argue that this is not a coincidence, but that zero modes should in fact always be solutions to the linearised equations of motion. The argument runs as follows. Given a basis $\{\phi_i\}$ for the configuration space of our theory, we may write $\eta = \eta_i\phi_i$ and consequently

$$S[\tilde{\phi} + \eta] = S[\tilde{\phi}] + \partial_i S[\tilde{\phi}]\eta_i + \partial_i\partial_j S[\tilde{\phi}]\eta_i\eta_j + O(\eta^3) \tag{1.18}$$

If we define the vector $DS[\phi]$ by $DS[\phi]_i = \partial_i S[\phi]$ and the matrix $D^2S[\phi]$ by $D^2S[\phi]_{ij} = \partial_i\partial_j S[\phi]$, then we may rewrite the above expansion as

$$S[\tilde{\phi}+\eta] = S[\tilde{\phi}] + <\eta, DS[\tilde{\phi}]> + <\eta, D^2S[\tilde{\phi}]\eta> + O(\eta^3) \tag{1.19}$$

Similarly, we may then write

$$DS[\tilde{\phi}+\eta] = DS[\tilde{\phi}] + D^2S[\tilde{\phi}]\eta + O(\eta^2) \tag{1.20}$$

This equation is going to be important to us, because the derivative DS is related to the equations of motion for the theory. To be more precise: if $DS[\phi]$ is zero, then ϕ is a solution to the equations of motion.

Now suppose that $\tilde{\phi}$ is a solution to the equations of motion and η is a zero mode of $\tilde{\eta}$. In that case, we know that the second and third terms in the expansion (1.19) are zero. We would like to show that $\tilde{\phi}+\eta$ still solves the equations of motion to first order in η, as this implies our claim that eta is a solution to the linearised equations of motion. This means we have to show that the second term in the expansion (1.20) for DS is zero. To do this, we first change to a basis $\{\hat{\phi}_i\}$ in which the symmetric matrix $D^2S(\tilde{\phi})$ is diagonal. In terms of the new basis, we may write the third term in the expansion (1.19) as follows

$$<\eta, D^2S[\tilde{\phi}]\eta> = \lambda_i\hat{\eta}_i^2 = 0$$

where the λ_i are the eigenvalues of $D^2S[\tilde{\phi}]$ and the $\hat{\eta}_i$ are the components of η in the new basis. Now if $\tilde{\phi}$ is a (semi)stable classical solution, then all the eigenvalues λ_i of D^2S will be smaller than or equal to zero and hence it follows for all i that $\hat{\eta}_i$ can be non zero only if λ_i is zero. From this, one can then easily see that $D^2S[\tilde{\phi}]\eta = 0$ and hence, we can conclude from (1.20) that $\tilde{\phi}+\eta$ will solve the equations of motion to first order in η.

The result we just derived has given us a powerful tool to find zero modes; we can just solve the linearised equations of motion and check if the solutions we find are zero modes. Note that this does not have to be the case. The reason for this lies in the precise nature of the correspondence between the equations of motion and the equation $DS = 0$. When we know that $<\xi, DS[\tilde{\phi}]> = 0$ for all test functions ξ, then we can derive from this (by partial integration) that $\tilde{\phi}$ satisfies the equations of motion. Conversely, if we know that $\tilde{\phi}$ satisfies the equations of motion, we may derive from this that $<\xi, DS[\tilde{\phi}]> = 0$ for all test functions ξ. But now it becomes important how "test functions" are defined. If a solution η to the linearised equations of motion does not come into this category, then $<\eta, DS[\tilde{\phi}+\eta]>$ may be unequal to zero and may even diverge. Now from (1.19) and (1.20), we see that we can write

$$S[\tilde{\phi}+\eta] = S[\tilde{\phi}] + <\eta, DS[\tilde{\phi}+\eta]> + O(\eta^3)$$

and hence, the deformation η will certainly not leave S fixed if $<\eta, DS[\tilde{\phi}]>$ is not zero. In fact,η will make the action diverge if $<\eta, DS[\tilde{\phi}+\eta]>$ diverges.

A typical example of how things can go wrong is the theory given by

$$\mathcal{L}(\phi) = \int dx \left\{ (\partial_t \phi)^2 - (\partial_x \phi(x))^2 + \frac{1}{2}\phi(x)^2 \right\}$$

The potential in this Lagrangian is clearly minimised by the unique field configuration with $\phi(x,t) = 0$ for every x and t. The equation of motion for this Lagrangian is

$$\partial_t^2 \phi - \partial_x^2 \phi - \phi = 0$$

Note that this is a linear equation and thus equal to its own linearised version. This means that zero modes will have to satisfy the time independent version of this equation. Solving this, we find two linearly independent solutions:

$$\phi_1(x) = \cosh(x)$$
$$\phi_2(x) = \sinh(x)$$

However, these solutions are not zero modes in the sense of the definition above, because one can easily verify that they make the action diverge. In the following, we will often call zero modes in the sense of the definition above *normalisable zero modes,* whereas we will call solutions to the linearised equations of motion *non-normalisable zero modes.* The term "normalisable zero modes" stems from the fact that these zero modes have to leave some norm finite that follows from the requirement that they leave the action, or equivalently, the energy finite.

The above gives a general recipe for finding zero modes. However, physically interesting potential energy functionals often have an interesting feature, which makes it much easier to find their zero modes. These potentials are often invariant under certain groups of transformations of the fields. We say that the zero modes that arise from these are a consequence of spontaneous symmetry breaking. Zero modes that arise from such symmetry breaking can be found explicitly by differentiating with respect to the symmetry's parameters. Often, the symmetry involved is not just a symmetry of the potential energy, but also a symmetry of the total energy. If this is the case, the quantisation of time independent solutions simplifies considerably, as we will see. The kink model is a typical example of a theory with the above feature; the model has translation symmetry, which has one dimensional orbits, the kink breaks this symmetry and one zero mode arises as a result. Also, the translation symmetry is a symmetry of the kink's potential energy as well as a symmetry of its total energy.

1.5.2 Collective Coordinates

In this section, I will describe a formalism for the semi-classical quantisation of solitons that does take zero mode fluctuations into account (unlike the naive formalism we used in section 1.4). This formalism is due to Christ and Lee [11]. It can be used for the quantisation of both time dependent and time independent classical soliton solutions, but we shall be most interested in the static case. In the static case, the Christ-Lee formalism does not necessarily deal with all zero modes that are imaginable according to definition 1, but it does handle the ones due to variations along level surfaces of the potential energy, as we will see. Zero modes that are not of this type are not dealt with separately, but this need not cause very much concern, because, if the original classical solution was stable, we may hope that quantum fluctuations in the directions of these modes will be be restricted by quartic or higher order terms in the potential.

For our account of the Christ-Lee formalism, we will work with a generalisation of our toy Lagrangian (1.1) to arbitrary space dimensions. We use a slightly different notation, with an explicit coupling constant g, following [11]:

$$L = \int d\vec{x} \left\{ \frac{1}{2}\partial^\mu \phi \partial_\mu \phi - g^{-2} U(g\phi) \right\} \tag{1.21}$$

If we expand the potential U around one of its minima, the quadratic terms will have order g^0, the cubic terms order g^1 and so on. In particular, we see that the potential is purely quadratic in the limit of g going to zero.

Let us suppose that we have found all solutions to the Euler-Lagrange equations for (1.21) which have a certain energy $\mathcal{E}$. Moreover, let us suppose that we have parameters $z_1^0, \ldots, z_K^0$ which parametrise this family of solutions. We can then write a particular solution in this family as $\phi(\vec{x}, t, z_1^0, \ldots, z_K^0)$. We can remove the explicit time dependence of such a solution by replacing the time independent parameters z_0^k by time dependent parameters z_k, that is, we can write

$$\phi(\vec{x}, t, z_1^0, \ldots, z_K^0) = \phi(\vec{x}, 0, z_1(t, z_1^0, \ldots z_K^0), \ldots, z_K(t, z_1^0, \ldots z_K^0))$$

It is not difficult to understand why we can do this. The parameters z_k^0 effectively describe all possible initial conditions for the fields of a solution with energy $\mathcal{E}$, while t describes the time evolution of the solution. However, every point on the orbit of a solution with energy $\mathcal{E}$ through field space is also a good initial point for such an orbit (Note that we use the fact that time evolution keeps the energy fixed here). In short, we can say that time evolution is really a motion on the space parametrised by the z_k^0.

In fact, we can do even better than we did above; we may choose the z_k to be of the form

$$z_k(t) = z_k^0 + u_k t \tag{1.22}$$

where the u_k are constants. The easiest way to see that this is possible, is to take one of the parameters z_k to be the parameter that describes the flow of the equations of motions along the manifold parametrised by the z_0. This will then have the form above while the others will be time independent.

If the family σ consists of (multi)soliton solutions to the equations of motion, then there could be a more physical interpretation for the speeds u_k as well. In cases where the individual solitons are far removed from each other, the u_k (or at least some of them) could be identified with the velocity components of the solitons. The z_k^0 could then be identified with the solitons' initial positions.

Note that, in the important special case where the $\phi(\vec{z})$ are all static solutions that correspond to minima of the potential energy, the u_k will all be zero and the z_0^k will just be a set of coordinates for the manifold of minima of the potential energy.

All the solutions $\phi(\vec{x}, \vec{z})$ are of order g^{-1}. We can see this as follows. If we define

$$\sigma(\vec{x}, \vec{z}) := g\phi(\vec{x}, \vec{z})$$

Then the functions σ will satisfy the equation

$$\partial_x^2 \sigma = \frac{\partial U(\sigma)}{\partial \sigma}$$

This equation is independent of g and therefore, so is σ. Hence ϕ is of order g^{-1}. To get more explicit perturbation series, we will often write $g^{-1}\sigma$ in stead of ϕ.

To get an approximation to the quantum theory, the idea is to expand the field ϕ around the classical solution as follows:

$$\phi(\vec{x}, t) = g^{-1}\sigma(\vec{x}, z_1, \dots, z_K) + \sum_{n=K+1}^{\infty} q_n(t)\psi_n(\vec{x}, z_1, \dots, z_K) \tag{1.23}$$

where for any $\vec{z}$, the $\psi_n(\vec{z})$ are a complete orthonormal set of functions subject to the constraints

$$\int d\vec{x} \left\{ \psi_n \frac{\partial \sigma}{\partial z_k} \right\} = 0 \tag{1.24}$$

In the static case, the functions $\frac{\partial \sigma}{\partial z_k}(\vec{z})$ are just the zero modes of $\phi(\vec{z})$ that lie along the equipotential surface parametrised by the z_k^0. Thus, the constraints above just make sure that all the ψ_n are orthogonal to all zero modes that lie along equipotential surfaces. On the one hand, this means we will no longer have any problems with the quantisation of coordinates associated to quantum fluctuations in the directions of zero modes. On the other hand, we will somehow have to incorporate a description of these fluctuations in our formalism if we want it to make any physical sense. We do this by elevating the parameters z_k into dynamical coordinates. These are called *collective coordinates*, presumably because a change in one of the z_k affects all the modes associated with the q_n collectively.

Having made the expansion (1.23), the rest of the quantisation process is fairly straightforward: First, one writes the Hamiltonian in terms of the new coordinates q_n and z_k and their conjugate momenta. Then, after dealing with some ordering difficulties, one substitutes the appropriate operators for these. Finally, one tries to find the solution to the resulting quantum mechanical problem in terms of a power series in the coupling constant g. A brief sketch and results of this procedure will be given in section 1.5.3.

In the rest of this section, I will give a geometrical interpretation to the expansion (1.23) and comment briefly on its range of validity. To come to our geometrical interpretation of (1.23), let us first have a look at the more conventional expansion given by

$$\phi(\vec{x}, t) = g^{-1}\sigma(\vec{x}, z_1, \dots, z_K) + \sum_{n=1}^{\infty} q_n(t)\psi_n(\vec{x}), \tag{1.25}$$

where we do not impose restrictions on the complete orthonormal set of functions ψ_n. In this expansion, the ψ_n correspond to a fixed choice of basis in the infinite dimensional vector space M of allowed field configurations $\phi(\vec{x})$. The q_n are used as coordinates that describe the motion of a field configuration $\phi(\vec{x},t)$ through M. The expansion (1.25) is the expansion of the fields we would use in a situation without zero modes. However, we are now in a situation with zero modes. Therefore, we had to somehow change the expansion (1.25) in such a way that we could deal with zero modes separately. Let us now describe how one comes from the expansion (1.25) to the expansion (1.23).

Of the zero modes, the ones that have to be dealt with most urgently lie along the sub-manifold S of M which consists of field configurations that lie on classical paths with energy $\mathcal{E}$. This manifold is parametrised by the z_k^0. We would like to choose a new basis for M in such a way that the first K elements of this basis lie along S and the rest lie orthogonal to it. At least in the static case we would then be able to identify the zero modes with the first K vectors in this basis and this would enable us to treat them separately. However, it is easy to see that it is in general impossible to choose a basis so that this condition is satisfied at all points of S at the same time. What we can do is the following: for each point $(z_1^0,\ldots,z_K^0)$ of S, we can find a basis $\psi_n(\vec{x},z_1^0,\ldots,z_K^0)$ which fulfills our condition at this particular point. What we have done in the expansion (1.23), is to take a point $\vec{z}(t)$ on S and the coefficients q_n $(n>K)$ of the $\psi_n(\vec{z})$ that are orthogonal to S at this point as the coordinates that describe the motion of $\phi(x,t)$ through M.

It is now logical to ask whether it is possible to describe an arbitrary field configuration $\phi(\vec{x},t)$ in terms of these new coordinates. This will only be the case if, at all times t, we can choose a point $\vec{z}(t)$ on S so that $\phi(\vec{x})$ lies in the orthogonal complement to the tangent plane to S at $\vec{z}$. It is not immediately obvious that one can always choose such a point. Also, even if one can find a point $\vec{z}(t)$ for every t, then there could still be other difficulties which make it impossible to use the expansion (1.23) in practice. I will illustrate these problems with some simple graphical examples.

For simplicity, let us work with a two-dimensional space of field configurations M and let us assume for the moment that the sub-manifold S is just a circle. In this situation there will be only one relevant vector in each of the bases $\psi_n(\vec{z})$ and I will denote this $\psi(z)$. This situation is depicted in figure 1.2. In this figure we see the orbit of a field configuration $\phi(x,t)$ through M, as well as the orbit of the corresponding point $\vec{z}(t)$ through S. The contribution of $\psi(z(t))$ to $\phi(t)$ is indicated by a vector from $z(t)$ to $\phi(t)$ at several points on the orbit.

In this example, the expansion (1.23) works very well. The only problem that occurs is the fact that the coordinates for $\phi(x,t)$ are not uniquely determined: in stead of the points $z(t)$ chosen in the picture, we could also have chosen the diametrically opposed points. However, in practical (semi-classical) calculations, this is unlikely to cause any problem, because we would usually make a local approximation around points of S that would make it impossible to "see" from one side of the circle what happens on the other side

Let us now look at figure 1.3(a). Here, S is a "dented" closed curve in stead of a circle. We see that in this case, we can still find a point on S for each point of the orbit of ϕ. However, we also notice that the collective coordinates corresponding to these points may have to make a jump on S as ϕ varies. This would introduce an artificial discontinuity into our model. In our two-dimensional example, it is easy to see that this situation will

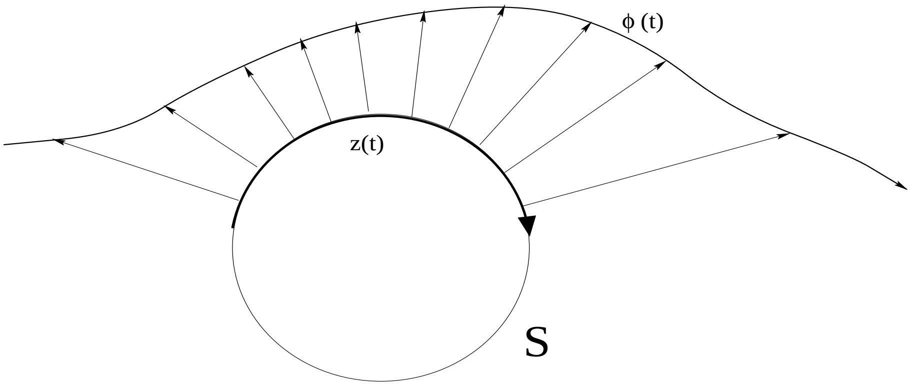


Figure 1.2: path of a field configuration ϕ and its orthogonal projection on the circle S

occur less and less as we require the orbit of ϕ to be closer to S. In particular, in the limit of weak coupling ($g \to 0$), we expect it to disappear, unless S has a pathological shape similar to that drawn in figure 1.3(b).

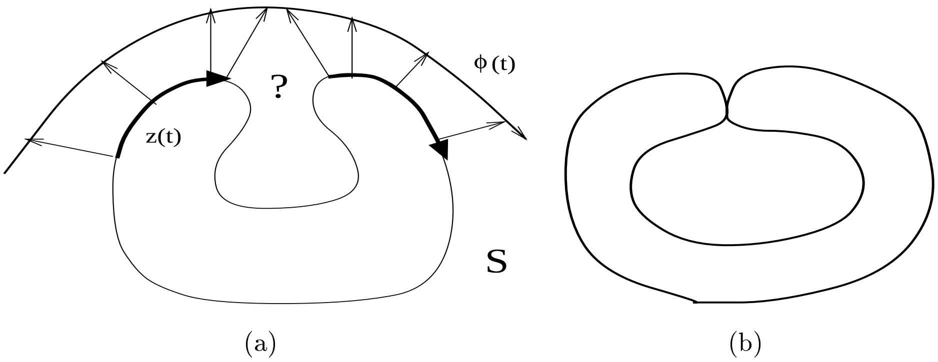


Figure 1.3: (a) path of a field configuration ϕ and its orthogonal projection on a dented curve S
(b) picture of a pathological curve

Now have a look at figure 1.4. Here we have taken the manifold S to be a line segment. In this example, we cannot always find collective coordinates for $\phi(x, t)$.

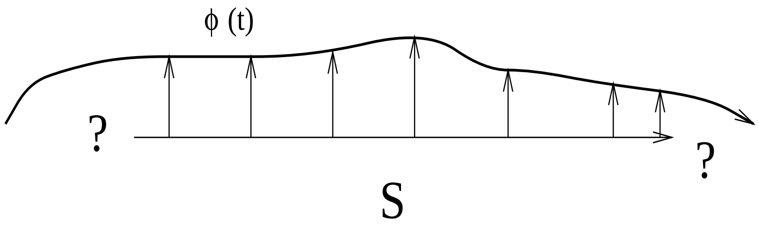


Figure 1.4: path of a field configuration ϕ and its orthogonal projection on the line segment S

Fortunately, there is some reason to believe that this problem will not occur in physically interesting situations, especially in finite-dimensional configuration spaces. I will now give a simple intuitive argument why.

First, we notice that the set S of equal energy solutions to the equations of motion will be a closed set if the energy and the action are continuous. This follows because S

is the intersection of the complete original of a point with respect to the action and the complete original with respect to a point of the energy. Let us now assume that the set S is also a manifold in the strict mathematical sense of the word. This will almost always be the case in physical applications. Given a point p in configuration space, we can now define a function f on S as follows: for every s in S we take $f(s)$ equal to the distance between s and p in configuration space. If S is now bounded in configuration space, then f will have to take a minimum value in a point z of S. The line through p and z will now have to intersect S at right angles, because if it did not, it would be possible to find a point on S near z with a smaller distance to p than z. Here it is crucial that S is a manifold, because this guarantees that one can go in a direction towards p without leaving S. If S is not a manifold, this does not have to be the case (cf. the line segment example with the endpoints of the line segment included in S). The argument can be extended to the case of closed, but unbounded manifolds S by restricting the distance function to a closed sphere around p that contains points of S. Of course this argument only works for finite-dimensional configuration spaces, but it does seem to give hope for the general situation.

Concluding, the main points we can make about the validity of the collective method are the following:

- In general, it may not be possible to find collective coordinates for all allowed field configurations, but the situation for physically interesting situations seems to be more hopeful.

- Discontinuous changes in the collective coordinates may sometimes be necessary to describe continuous fluctuations of the fields in time. Except in pathological cases, this problem will disappear at weak coupling. The speed at which this happens will of course depend on the geometry of the manifold of equal energy solutions to the equations of motion.

- There may be several sets of collective coordinates that describe the same field configuration. In the semi-classical approximation, this problem will probably disappear; only one of the possibilities will be seen.

1.5.3 Some Results of the Collective Coordinate Method

In this section, we will quantise the Lagrangian (1.21) in terms of the coordinates z_k and q_n introduced in the previous section. First we write the Lagrangian in terms of these:

$$\begin{aligned} L = &\int d\vec{x} \left\{ \frac{1}{2} \left[\sum_{k=1}^{K} \dot{z}_k \left(g^{-1} \frac{\partial \sigma}{\partial z_k} + \sum_{n=K+1}^{\infty} q_n \frac{\partial \psi_n}{\partial z_k} \right) + \sum_{n=K+1}^{\infty} \dot{q}_n \psi_n \right]^2 \right\} \\ &- \int d\vec{x} \left\{ \frac{1}{2} \left[g^{-1} (\nabla \sigma) + \sum_{n=K+1}^{\infty} q_n (\nabla \psi_n) \right]^2 + g^{-2} U \left(\sigma + g \sum_{n=K+1}^{\infty} q_n \psi_n \right) \right\} \end{aligned} \tag{1.26}$$

where the derivatives of σ and the ψ_n have been taken with the z_k and $\vec{x}$ regarded as independent variables. In the following we will sum over all repeated indices, where we take the convention that indices k, k' range from 1 to K and indices n, n' range from $K+1$ to infinity. Let us now introduce a mass matrix $M(z_k, q_n)$ with elements given by:

$$\begin{aligned} M_{k,k'} &= \int d\vec{x} \left\{ \left(g^{-1} \frac{\partial \sigma}{\partial z_k} + q_n \frac{\partial \psi_n}{\partial z_k} \right) \left(g^{-1} \frac{\partial \sigma}{\partial z_{k'}} + q_{n'} \frac{\partial \psi_{n'}}{\partial z_{k'}} \right) \right\} \\ M_{kn} &= M_{nk} = \int d\vec{x} \left\{ q_{n'} \frac{\partial \psi_{n'}}{\partial z_k} \psi_n \right\} \\ M_{nn'} &= \delta_{nn'} \end{aligned} \tag{1.27}$$

. Further, let us define

$$\bar{U}(\phi) := U(\phi) + \frac{1}{2}(\nabla\phi)^2.$$

It is not difficult to see that, with these definitions, the Lagrangian can be rewritten as

$$L = \frac{1}{2}\left(\dot{z}_k M_{kk'} \dot{z}_{k'} + 2\dot{z}_k M_{kn} \dot{q}_n + \dot{q}_n M_{nn'} \dot{q}_{n'}\right) - g^{-2} \int d\vec{x} \left\{ \bar{U}(\sigma + g q_n \psi_n) \right\}$$

(hence the name "mass matrix"). The momenta conjugate to z_k and q_n are now

$$\begin{aligned} p_k &= M_{kk'} \dot{z}_{k'} + M_{kn} \dot{q}_n \\ \pi_n &= M_{nk} \dot{z}_k + M_{nn'} \dot{q}_n' \end{aligned} \tag{1.28}$$

or, in matrix form

$$\begin{pmatrix} \vec{p} \\ \vec{\pi} \end{pmatrix} = M \begin{pmatrix} \dot{\vec{z}} \\ \dot{\vec{q}} \end{pmatrix}.$$

We can now write down the Hamiltonian:

$$H = \frac{1}{2} p_k (M^{-1})_{kk'} p_{k'} + p_k (M^{-1})_{kn} \pi_n + \pi_n (M^{-1})_{nn'} \pi_{n'} + g^{-2} \int d\vec{x} \left\{ \bar{U}(\sigma + g q_n \psi_n) \right\} \tag{1.29}$$

We would now like to convert this classical Hamiltonian into a differential operator by making the standard substitution $p_k \to -i\partial_{z_k}$, $\pi_n \to -i\partial_{\psi_n}$. However, the differential operator we obtain this way depends on the order of the factors in the first two terms of (1.29), as these contain both coordinates and their conjugate momenta. Therefore, to have a consistent quantisation scheme, we have to determine an ordering for the Hamiltonian before making the given substitution.

Now if we had expanded the field ϕ in terms of a complete set of orthonormal functions $\Psi_s(\vec{x})$ with coefficients $Q_s(t)$ in stead of using the collective coordinate expansion (1.23), then the Hamiltonian would have taken the form

$$H = -\frac{1}{2}\sum_s \frac{\partial^2}{\partial Q_s^2} + V(\vec{Q}) \tag{1.30}$$

for some potential energy function V. Here we have no ordering difficulties, as coordinates and momenta do not occur in the same term. We can now determine the correct ordering for (1.29) by requiring that, as a differential operator, this expression is equal to (1.30), expressed in terms of the variables q_n and z_n.

First, we give a formula which enables us to express (1.30) in terms of an arbitrary set of variables. Suppose we have to sets of variables $x_1 \ldots x_l$ and $y_1 \ldots y_l$ with the x_i given as functions of the y_i. Then we have

$$\sum_{i=1}^{l} \frac{\partial^2}{\partial x_i^2} = \sum_{i,j=1}^{l} \frac{1}{J}\frac{\partial}{\partial y_i}(m^{-1})_{ij} J \frac{\partial}{\partial y_j} \tag{1.31}$$

where the elements of the matrix m are given by

$$m_{ij} = \sum_{r=1}^{l} \frac{\partial x_r}{\partial y_i}\frac{\partial x_r}{\partial y_j}$$

and where J is the square root of the determinant of m. I will not prove this formula here, but just use it as a given. In our case, the matrix m is just the mass matrix M, as can be easily checked. Here we encounter a problem, namely the fact that M is an infinite dimensional matrix. Even if we assume that the above formulae are valid for infinite dimensional matrices, we still need the determinant of M to apply them. Fortunately, it turns out that we can reduce the determinant of M to that of a finite dimensional $(K \times K)$-matrix. To this end we write M as the product of a matrix Q and its transpose Q^t, with Q of the form

$$Q = \begin{pmatrix} A & B \\ 0 & 1 \end{pmatrix}$$

where A is a $(K \times K)$-matrix and B is a $(K \times \infty)$-matrix. If we take $B_{kn} = M_{kn}$ then it follows that we will indeed have $M = QQ^t$ if A satisfies

$$(A^2)_{kk'} = M_{kk'} - M_{kn}M_{nk'}.$$

Since the right hand side of this equation is a symmetric matrix, we will be able to find an A which satisfies this requirement. Looking back, we see that

$$J = \sqrt{\det(M)} = \det(Q) = \det(A) = \sqrt{\det(A^2)} = \sqrt{\det(M_{kk'} - M_{kn}M_{nk'})} \tag{1.32}$$

Now we can write down the correctly ordered form of the Hamiltonian:

$$
\begin{aligned}
H &= \frac{1}{2}J^{-1}\left[p_k(M^{-1})_{kk'}Jp_{k'} + p_k(M^{-1})_{kn}J\pi_n + \pi_n(M^{-1})_{nk}Jp_k + \pi_n(M^{-1})_{nn'}J\pi_{n'}\right] + \\
& g^{-2}\int d\vec{x}\left\{\bar{U}(\sigma + gq_n\psi_n)\right\} \qquad (1.33)
\end{aligned}
$$

If we consider the p_k and π_n as operators, then this Hamiltonian describes the quantum mechanics of our theory.

Now that we have the Hamiltonian, we would like to find some way of solving its eigenvalue equation

$$
H\Psi = E\Psi. \qquad (1.34)
$$

to some order in g, starting with order g^{-2}, which should be the leading contribution, as the classical solution was of this order. The first thing we need to do is to find out which terms in the equation (1.34) are actually of order g^{-2}. To do this, we need to have some information on the g-dependence of $p_n\Psi$ and $\pi_n\Psi$. We will have to make a guess, based on the behaviour of the classical momenta p_k and π_n. Before quantisation, these were given by the expression (1.28). On expanding the matrix elements in these expressions we can read off their g-dependence and we find that the p_k were of order g^{-2}, while the π_n were of order g^0. At least in the semi-classical approximation that we use, we want the expectation values of the operators p_k and π_n to exhibit the same behaviour. Therefore, we shall conjecture

$$
p_k\Psi = O(g^{-2}) \qquad (1.35)
$$

and

$$
\pi_n\Psi = O(g^0). \qquad (1.36)
$$

Armed with these conjectures we can extract the terms of order g^{-2} from the eigenvalue equation (1.34) as follows: First we note that the $(M^{-1})_{kk'}$ and the $(M^{-1})_{nk}$ are of order g^2, while the $(M^{-1})_{nn'}$ are of order g^0. If we now take into account that the derivatives $\frac{\partial}{\partial q_n}$ and $\frac{\partial}{\partial z_k}$ do not lower the order of terms when they act on J or on elements of M^{-1}, then we see that the terms in (1.34) involving $(M^{-1})_{nn'}$, $(M^{-1})_{kn}$ and $(M^{-1})_{nk}$ are of higher order then the remaining two terms and can therefore be ignored to lowest order. Next, it is easy to see that, to order g^2, we must have

$$
(M^{-1})_{kk'} = (M_0^{-1})_{kk'}
$$

where M_0 is the $K \times K$ matrix with elements

$$
(M_0)_{kk'} = \int d\vec{x}\left\{\frac{\partial\sigma}{\partial z_k}\frac{\partial\sigma}{\partial z_{k'}}\right\} \qquad (1.37)
$$

Further, we note that we do not have to expand the complicated factors J and J^{-1} in the term containing $(M^{-1})_{kk'}$. This is because all the terms in which J has to be differentiated are of higher order in g then the ones in which we have to differentiates Ψ. This leaves us with terms in which the factors J and J^{-1} cancel each other. In the same way, we get rid of all terms in which $(M^{-1})_{kk'}$ has to be differentiated.

Using all this, we find that the terms of order g^{-2} in (1.34) involve only the z_k and can be written

$$g^{-2}\left[\frac{1}{2}(M_0^{-1})_{kk'}\frac{\partial}{\partial z_k}\frac{\partial}{\partial z_{k'}}+\int d\vec{x}\left\{\bar{U}(\sigma(\vec{x},\vec{z}))\right\}\right]\Psi=E\Psi \tag{1.38}$$

Now note that the second term in this equation is just a constant, independent of $\vec{z}$, because it is exactly the potential energy of the classical solution corresponding to $\sigma(\vec{x},\vec{z})$. This means that, to this order in g, we can work with an effective Hamiltonian given by

$$H_{-2}=\frac{1}{2}(M_0^{-1})_{kk'}p_kp_{k'}. \tag{1.39}$$

This is just the Hamiltonian for geodesic motion on the manifold S parametrised by the z_k^0, when this is endowed with the metric $(M_0^{-1})_{kk'}$. Notice $(M_0^{-1})_{kk'}$ is just the Riemannian metric on S, which comes from the $\mathcal{L}^2$ metric on the space of field configurations. It turns out that in many interesting models a description of the low energy dynamics of solitons in terms of geodesic motion on a manifold like S is possible. We will return to this idea in section 1.6. Before proceeding, let us make two interesting remarks about the Hamiltonian (1.39).

First of all, we see from the definition (1.37) that the zero modes have to be normalisable to make the metric nonsingular. This is just an expression of the fact (noted in section 1.5.1) that zero modes have to be normalisable to make the energy converge.

Secondly, we see that each of the momenta p_k commutes with the Hamiltonian (1.39) to leading order (commuting one of the p_k with M_0^{-1} will in general produce changes in higher orders). This means that to leading order, we should be able to choose Ψ to be an eigenfunction of H_{-2} and all the p_k at the same time. It would seem likely that our approximation scheme will work best if we choose the expectation values of the p_k equal to the classical values of the p_k.

In ([11]), Christ and Lee make the above argument more explicit. They incorporate the conjectures (1.35) and (1.36) in an ansatz for the energy eigenfunctions Ψ_α:

$$\Psi_\alpha(\vec{z},\vec{q})=e^{i\xi(\vec{z})/g^2}\chi_\alpha(\vec{z},\vec{q}) \tag{1.40}$$

where $\xi(\vec{z})$ is independent of g and χ_α is independent of g at least to leading order.

Then, they find the equation (1.38) in terms of ξ.

$$\frac{1}{2}\frac{\partial\xi}{\partial z_k}(M_0^{-1})_{kk'}\frac{\partial\xi}{\partial z_{k'}}+\int\bar{V}(\sigma(\vec{x},\vec{z}))d\vec{x}-g^2E=0 \tag{1.41}$$

They prove that it is possible to choose a unique function ξ so that it satisfies (1.41) and the condition

$$\frac{\partial \xi}{\partial z_k} = (M_0)_{kk'} u_{k'} \tag{1.42}$$

and that with ξ chosen in this way, the terms of order g^{-1} in the eigenvalue equation (1.34) vanish. The constants u_k in (1.42) are the same as those appearing in (1.22) and thus we have the identity $\dot{z}_k = u_k$ before quantisation. With (1.28) we then see that our approximation is indeed improved if we choose Ψ so that it gives the momenta p_k their classical values (to leading order), just as we expected before.

From (1.42), we can also note another interesting fact: Substituting into equation (1.38), we see that the energy $\mathcal{E}$ in this equation is given by

$$E = g^{-2}\left[u_k (M_0)_{kk'} u_{k'} + \int d\vec{x} \left\{ \bar{V}(\sigma(\vec{x}, \vec{z})) \right\}\right]$$

which is just the energy of the classical solution.

Let us now go on with the choice of S that eliminates the terms of order g^{-1} and go on with a brief description of the terms of order g^0. These are obtained by expanding M^{-1} and $\bar{U}$ to second order in the quantum fluctuations $q_n \psi_n$. We will limit ourselves to the case where the classical solutions to the equations of motions we started from are static. This means that all the u_k are zero. The general case is treated in [11]

To obtain the terms of order g^0, we look back to the Hamiltonian (1.33). Using our previous observations on M^{-1}, it is not difficult to see that the order g^0 contributions from the terms that contain one of the p_k will always have to contain a factor $\frac{\partial \xi}{\partial z_k}$ for some k. Using equation (1.42) we see that all these terms will vanish, as in our case, we have $\vec{u} = 0$. This means we are left with the contributions that come from the last two terms in the Hamiltonian. From this point it follows easily that, in the static case, the terms of order g^0 in the eigenvalue equation (1.34) are given by:

$$\left[\frac{1}{2}\pi_n \pi_n + \frac{1}{2} q_n F_{nn'} q_{n'} - E_\alpha\right] \chi_\alpha = 0 \tag{1.43}$$

where $F_{nn'}$ is the n, n' matrix element of the operator F given by

$$F[\psi] = \left[-\nabla^2 + \frac{\partial^2 U(\sigma(\vec{z}))}{\partial \sigma^2}\right] \psi. \tag{1.44}$$

In the static case, the zero modes $\frac{\partial \sigma}{\partial z_k}$ are exactly all the eigenstates of this operator with eigenvalue zero. Therefore, we can choose the ψ_n so, that they are eigenfunctions of the operator F, as well as orthogonal to the zero modes. $\frac{\partial \sigma}{\partial z_k}$. With this choice of ψ_n, we can rewrite (1.43) as

$$\left[\frac{1}{2}\pi_n \pi_n + \frac{1}{2} q_n \omega_n(\vec{z}) q_n - E_\alpha\right] \chi_\alpha = 0 \tag{1.45}$$

where $\omega_n(\vec{z})$ is the eigenvalue corresponding to $\psi_n(\vec{z})$

Note that, even though we started from a static classical solution, this does not mean that the operator $\vec{z}$ which appears in this equation is independent of time. In fact, there is a good argument why this is usually not the case; suppose for a moment that z is time independent. In that case we are just dealing with a system of infinitely many harmonic oscillators and the usual energy spectrum will follow:

$$E_\alpha = \sum_n (N_n + \frac{1}{2})\omega_n(\vec{z}). \tag{1.46}$$

where the N_n are occupation numbers. Now if the ω_n depend on $\vec{z}$, then the system will in general be able to lower its potential energy by moving from states with high $\vec{\omega}(\vec{z})$ into states with lower $\vec{\omega}(\vec{z})$. Naturally, this process will make $\vec{z}$ time dependent.

In fact, it is possible to get a feeling for this phenomenon at the classical level. Let us take as an example a two dimensional system with a Hamiltonian given by

$$H(x, p_x, y, p_y) = \frac{1}{2}(p_x^2 + p_y^2) + V(x, y) \tag{1.47}$$

Let us take the potential energy $V(x, y)$ equal to $e^{x^2}y^2$ for the moment. This potential energy is shown in figure 1.5.3

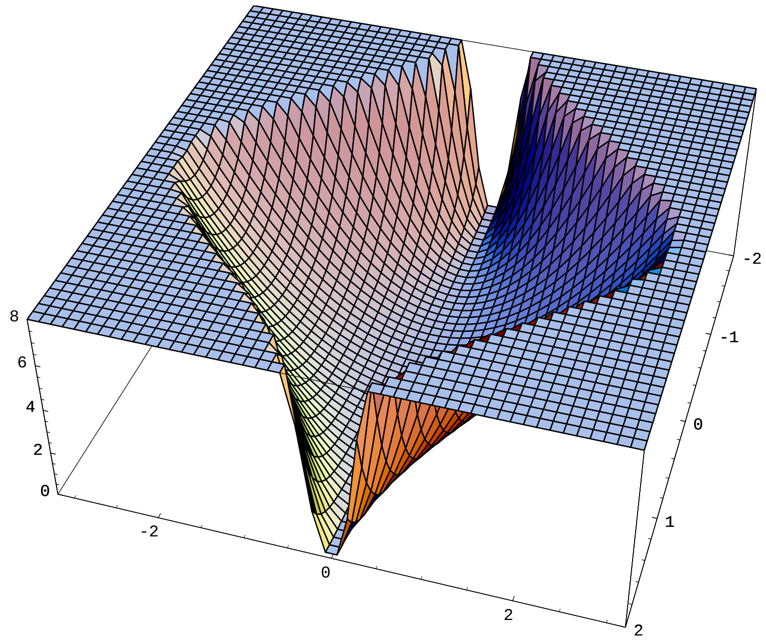


Figure 1.5: Plot of the the potential energy $V(x, y)$ in the Hamiltonian (1.47) as a function of x and y

It has a manifold of degenerate minima: the x-axis. Any point $(x, 0)$ will give a static classical solution and we see that the coordinate x can be taken as a collective coordinate, thus corresponding to $\vec{z}$ above. The Hamiltonian (1.47) can be taken to describe the motion of a point particle through a canyon landscape. The x-axis is the bottom of the canyon and to the sides of this axis, the walls of the canyon rise ever steeper as one moves away from the space origin. To get a feeling for the effects that quantum fluctuations will cause in this system, it makes sense to study what happens in the classical system as we slightly perturb the static solutions. Two numerical approximations for the temporal behaviour of the coordinates x and y due to different perturbations are plotted in figure (1.6).

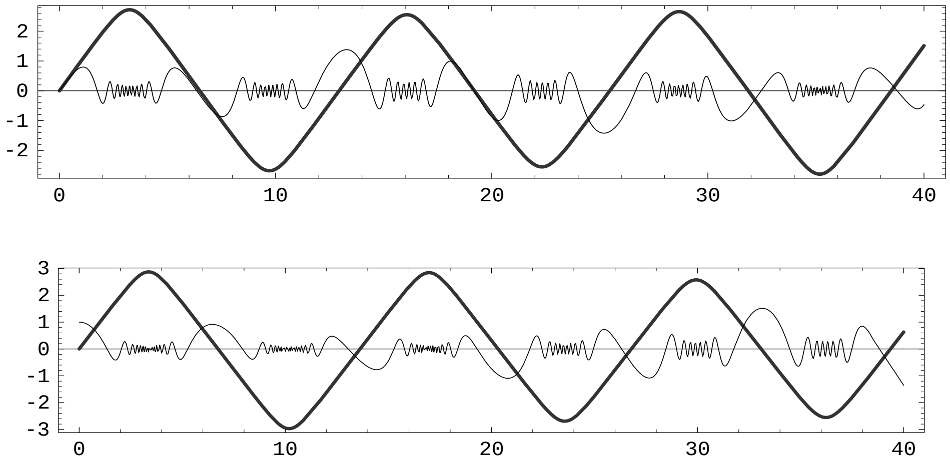


Figure 1.6: Numerical solutions to the equations of motion for (1.47) with initial conditions close to the static solution $x = y = 0$. The thick line gives $x(t)$, the thin line gives ten times $y(t)$.

From the graphs in this figure, we can clearly see that the "collective coordinate" x does not stay constant in time. In stead, it behaves roughly as if there was a potential on the x-axis which increases for increasing $|x|$. In terms of our canyon analogy, this means, that if we look only at the particle's movement along the bottom of the canyon, then it seems as if this bottom is not flat, but in stead rises as the walls of the canyon become steeper. Study of different potentials shows a similar behaviour. From the figure, we can also see how this behaviour of the motion along the canyon is linked to that of the motion orthogonal to it. As the walls of the canyon get steeper, more and more of the kinetic energy of the particle is transferred from the motion along the canyon into the motion perpendicular to it, until the motion of the particle along the canyon finally changes direction and then the reverse process takes place.

Now let us return to the system described by (1.45). If the frequencies ω_n are independent of $\vec{z}$, then the equation is independent of time. This still does not necessarily mean that the operator $\vec{z}$ will be independent of time (though we can expect that at least it's expectation value will be), but it does mean that (1.46) will give the right (time independent) energy spectrum. It also ensures that if we take the χ_α to be the usual eigenstates for a system of independent harmonic oscillators, they will provide eigenfunctions of the original Hamiltonian accurate through order g^0 on insertion into the ansatz (1.40).

Now when will the frequencies ω_n be independent of $\vec{z}$? A typical situation in which this occurs is the situation in which all the zero modes arise as a result of spontaneous breaking of a K parameter symmetry. In this case, one can identify the z_k with the parameters of this symmetry. The momenta $p_k \equiv \partial_{z^k}$ then become generators of the symmetry and hence they have to commute with the Hamiltonian. It follows that the matrix elements $F_{nn'}$ (and correspondingly the frequencies ω_n) have to be independent of $\vec{z}$, as promised. It also follows that the momenta p_k are conserved. Restricting to states with a definite value for each of these momenta, we have a non-degenerate ground state for the quantum theory to order g^0.

1.5.4 Implications for the Kink System

I will now briefly describe the implications of the previous section's results for the kink system. In this case we have only one collective coordinate; the one related to spatial translation. We will denote this Z. The momentum p_Z conjugate to Z is obviously just the total momentum operator of the system and therefore we will write $p_Z = P$. In terms of Z and P, the lowest order effective Hamiltonian (1.39) reduces to

$$\begin{aligned} H_{-2} &= \frac{P^2}{2\int dx(\partial_Z\phi_k)^2(x,Z)} && (1.48) \\ &= \frac{P^2}{2\int dx(\partial_Z\phi_k)^2(x-Z)} \\ &= \frac{P^2}{2\int dx(\partial_x\phi_k)^2(x)} \quad = \frac{P^2}{2M_k} && (1.49) \end{aligned}$$

where the M_k in the last line denotes the classical mass of the kink. That this is equal to $\int dx(\partial_x\phi_k)^2(x)$ can be easily seen using equations (1.5) and (1.3). Thus we find that, to lowest order in the coupling constant g, the quantum mechanics of the kink system can be described by the Hamiltonian for a non-relativistic free particle with mass equal to the classical mass of the kink. In fact, it is not difficult to see that the analogue to this result holds for any soliton described by the Lagrangian (1.1) which has just one zero mode.

As in the previous section, we can improve this result by one order of g if we require that the momentum corresponding to the collective coordinate be equal to its classical value. In the case of the kink, this means that we set the total momentum P equal to zero, thus "pinning the kink to the origin". Having done this, we could continue with the calculation of the order g^0 terms, as given in equation (1.43). Finally, we would have to solve the eigenvalue equation (1.45). Generally, this would be very difficult, because the frequencies ω_n in this equation would depend on the collective coordinates. However, for the kink system, this is is not the case. This follows either by the general argument given at the end of the previous section, or more directly by looking at the definition of the operator F given in (1.44). If one notes that the potential energy V in this definition is translation independent, then one sees that F itself must be independent of the collective coordinate Z. The ω_n will then also be independent of Z, as they are the eigenvalues of F. It follows that the energy spectrum of a static kink is given by the harmonic oscillator spectrum that we found in section (1.4) by just ignoring the zero mode. We conclude that the results obtained in this section are good for a kink at rest.

Now what if we want to describe a moving kink ? In that case it makes more sense to start from a time dependent classical solution; a Lorentz boosted version of the kink. Again, we can choose P to be equal to its classical value and doing this makes sure that the g^{-2} term in the energy is just equal to the classical energy of the moving kink. The quantum corrections are of order g^0 and it is shown in [11] that these are equal to the corrections for the static case.

1.6 Soliton Dynamics: The Geodesic Approximation

In principle, the formalism developed in the previous sections enables us to find a quantum description of soliton phenomena, starting from an appropriate classical solution. If we want to describe soliton dynamics, it would thus make sense to start with a classical solution that has two or more solitons and a non trivial time dependence. At this point we encounter a problem: it is usually very difficult to find time dependent classical solutions with multiple solitons.

In many cases, this difficulty makes application of the collective coordinate method as we have described it impossible. However, if we can find all static classical solutions with the required number of solitons, then we can make an approximation that circumvents this problem, at least in the limit where all the solitons' properties (including location) change very slowly. This approximation is the so called *geodesic* approximation. It was first proposed by Manton [12] in the context of classical BPS-monopoles (we will come back to these in chapter 4). I will now first describe how the approximation works and afterwards, I will give arguments as to why it should work.

First, one finds all the minimal-energy static classical solutions in the soliton sector one wishes to study. Though this is easier than finding time-dependent solutions, it is usually still quite difficult, especially as the number of solitons involved increases. In fact, it is not at all clear that static classical solutions will exist for arbitrary numbers of solitons. However, if there are no long-range inter-soliton forces, then we can hope to construct (approximate) static solutions by placing solitons in space at very large distances from each other. The energy of such solutions should then be equal to the sum of the masses of the individual solitons.

Second, one chooses a suitable set of collective coordinates for the space S of static solutions and calculates the lowest order effective Hamiltonian for the static system as given in (1.39). In particular, one has to find the metric M_0^{-1} in this equation. Remember that this was just the Riemannian metric induced by the $\mathcal{L}^2$ metric on the space of field configurations M. The resulting Hamiltonian is then used to describe the quantum mechanics of slowly moving solitons. Thus, one describes a configuration of slowly moving solitons as a point particle in geodesic motion on the moduli space of static solutions, endowed with the metric M_0^{-1}.

Now why should this approximation work ? To answer this question, we have to look back to our derivation of the lowest order effective Hamiltonian H_{-2}. Note that the only input we used to derive this result was the assumption that we could expand an arbitrary field configuration ϕ in the form

$$\phi(\vec{x},t) = g^{-1}\sigma(\vec{x},z_1,\dots,z_K) + \sum_{n=K+1}^{\infty} q_n(t)\psi_n(\vec{x},z_1,\dots,z_K), \tag{1.50}$$

where $g^{-1}\sigma$ is the embedding into configuration space of some manifold S parametrised by the z_k and the ψ_n are subject to the constraints (1.24). The geometrical interpretation of (1.50) which I gave in section 1.5.2, would seem to imply that this expansion is valid if two conditions are satisfied. First: the manifold S must be sufficiently "nice". Second, and more importantly, the field configurations ϕ which we want to describe, must stay near S at all times.

Up to now, if we wanted to make an expansion around a classical solution with energy $\mathcal{E}$, we took S to be the manifold given by all field configurations that lay on a classical path with this energy. The second condition above was then automatically satisfied if the quantum corrections to the energy were small. Suppose now that we had started with a classical solution with energy $\mathcal{E}'$ just a little bit higher than $\mathcal{E}$. This solution would not be confined to S. In stead, it would be forced to remain on the manifold S' that consists of all field configurations that lie on classical paths with energy $\mathcal{E}'$. However, if the difference between the energies $\mathcal{E}$ and $\mathcal{E}'$ was small enough, then every point on S' would be very near a point of S. When this is the case, we can hope to describe an arbitrary time dependent field configuration that stays near S' in terms of the collective coordinates for S! We can then go on to derive the order g^{-2} effective Hamiltonian (1.39) and this gives us the geodesic approximation. Note that in this approximation, we will no longer be able to eliminate the terms of order g^{-1}, as we did in section 1.5.3. The effective Hamiltonian should therefore be valid only up to corrections of order g^{-1}, not up to corrections of order g^0 as before.

Of course the above argument is not a proof of the validity of the geodesic approximation. For example, I have not said anything about the definition of distance in configuration space, when it is clear that this would be absolutely crucial in a real proof. However, I hope that the argument does at least give some intuition.

Another thing that should be noted, is that the above argument differs from the argument given by Manton in his paper [12]. In this paper, Manton argued that the *classical* scattering of the solitons involved could be described as geodesic motion on the moduli space S. That is, if one works in the low speed limit and uses the following initial conditions: The initial field configuration should be on S and the initial speed (in field space) should be along S.

I will now give a nice argument which makes this approximation plausible. Let us again look at the mechanical system given by the Lagrangian

$$L = (\partial_t \phi, \partial_t \phi) - V(\phi). \tag{1.51}$$

This describes the motion of a non-relativistic particle through a "landscape" in which "height" is given by V. The notation with the inner product is chosen in order to cover both the finite dimensional and the infinite dimensional case (which includes field theory). The equations of motion for this system are given by

$$\partial_t^2 \phi_n = -\frac{\partial V}{\partial \phi_n} \tag{1.52}$$

Now suppose the potential V has a manifold S of degenerate (local) minima. This corresponds to a flat "valley" in our landscape. Each point of this valley corresponds to a static solution to the equations of motion. As usual, we will set the value of the potential energy on S to zero. As in section (1.5.2), we expect that, in a neighbourhood of S, we can find collective coordinates (z, ψ). Here, z would give local coordinates for S and ψ would describe the directions normal to S. This inspires us to write down the following ansatz for a solution $\phi(t)$ of (1.52):

$$\phi(t) = \Phi(z(t)) + \epsilon^2 \psi(t). \tag{1.53}$$

Here, Φ is an embedding of S into configuration space and epsilon is a small parameter. If we can describe solutions in this way, then we expect that $\psi(t)$ will describe rapid oscillations perpendicular to S, while $z(t)$ will describe a slow drift along S. Suppose we have a solution to the equations of motion which is of the above form. We can then write the energy of this solution as

$$\mathcal{E}(\Phi(z),\psi) = (\partial_t\phi, \partial_t\phi) + \epsilon^4 \frac{\partial^2 V(z)}{\partial\phi_n\partial\phi_m}\psi_n\psi_m \tag{1.54}$$

If we define the symmetric matrix L by $L_{mn} = \frac{\partial^2 V(z)}{\partial\phi_n\partial\phi_m}$, then we can rewrite this to give

$$\mathcal{E} = (\partial_t\phi(z), \partial_t\phi(z)) + L_{mn}(z)\psi_m\psi_n \tag{1.55}$$

Let us us denote the eigenvalues of $L(z)$ by $E_n(z)$. For all z, we know that all the $E_n(z)$ are greater than or equal to zero. This is because we assumed that the manifold S is completely composed of minima of the potential. Now suppose there is a constant C, greater than zero, so, that for all z, we have:

$$\inf_n E_n(z) > C. \tag{1.56}$$

in that case, we can make the following estimate for the energy:

$$\mathcal{E} > (\partial_t\phi(z), \partial_t\phi(z)) + C\epsilon^4(\psi,\psi). \tag{1.57}$$

If the energy of our solution is of order ϵ^4, then the above estimate tells us that $\partial_t\phi$ is of order ϵ^2 and that ψ is of order 1. Moreover, we know that this is true for the entire time interval for which the solution is defined, because the energy is a conserved quantity. Now suppose we can prove local existence of solutions of the form (1.53) for initial locations near S and low initial speeds. It then follows from the estimate above that we will have solutions of that form that are defined all times. Next to the estimates for ψ and $\partial_t\phi$, we also want to have estimates for the time derivatives of the z_k and ψ_n. I will only give an intuitive argument for these here. Since we expect the z_k to describe a slow drift along S, we do not expect the order of $\partial_t z$ to exceed that of $\partial_t\phi$. Let us therefore suppose that $\partial_t z$ is of order ϵ^2. With this, it follows from our estimate for $\partial_t\phi$ that $\partial_t\psi$ is of order 1. Another estimate we will make is that $\partial_t^2 z$ is of order ϵ^2. This basically means that we do not allow $\partial_t z$ to have fluctuations of its own order (ϵ^2) on a time scale of order 1. To me, this seems to be a reasonable assumption. We certainly do not exclude the possibility that $\partial_t z$ changes significantly over a longer time (say a time of order ϵ).

Let us now substitute the ansatz (1.53) into the equations of motion. This yields the equations

$$\partial_t^2\psi_n = \frac{1}{\epsilon^2}\partial_t^2\Phi_n(z(t)) - L_{nm}(z)\psi_m + O(\epsilon^2). \tag{1.58}$$

Using our estimates for $\partial_t z, \partial_t^2 z$ and ψ, we can conclude from these equations that $\partial_t^2\psi$ is of order 1. Next to the equations of motion above, we have the requirement that the ψ_n

should be normal to S. Since Φ is an embedding of S into configuration space, we can write this requirement in terms of the partial derivatives $\partial_{z_k}\Phi$. We must have (cf. (1.24))

$$(\psi, \partial_{z_k}\Phi) = 0. \tag{1.59}$$

Differentiating this twice with respect to time, we get

$$(\partial_t^2\psi, \partial_{z_k}\Phi) + 2(\partial_t\psi, \partial_t\partial_{z_k}\Phi) + (\psi, \partial_t^2\partial_{z_k}\Phi) = 0 \tag{1.60}$$

and when we substitute for $\partial_t^2\psi$ from (1.58), this leads to the following equation:

$$\frac{1}{\epsilon^2}(\partial_t^2\Phi, \partial_{z_k}\Phi) = (L(z)\psi, \partial_{z_k}\Phi) - 2(\partial_t\psi, \partial_t\partial_{z_k}\Phi) - (\psi, \partial_t^2\partial_{z_k}\Phi) + O(\epsilon^2). \tag{1.61}$$

The inner product on the left hand side of this equation is just the acceleration of the point $\Phi(z(t))$ in the directions along the manifold S. This means that, if the right hand side is zero, $\Psi(z(t))$ is in geodesic motion.

The first term on the right hand side is identically zero: The matrix L is symmetric and $\partial_{z_k}\Phi$ is an eigenvector of L with eigenvalue zero. Using our estimates for $\psi, \partial_t\psi, \partial_t^2\psi, \partial_t z$ and $\partial_t^2 z$, we can estimate the other terms to be of order ϵ^2. Thus we find

$$(\partial_t^2\Phi, \partial_{z_k}\Phi) = O(\epsilon^4). \tag{1.62}$$

This allows us to conclude that, on a time scale of order $\frac{1}{\epsilon}$, the velocity of Φ along S can only be modified by amounts of order ϵ^3. Since we estimated the speed of Φ along S to be of order ϵ^2 to start with, this means that, as we take ϵ to zero, the motion of Φ approaches geodesic motion on S.

Chapter 2

Yang-Mills Instantons

This chapter is about instanton solutions of the Euclidean Yang-Mills equations, or in short: Yang-Mills instantons. The Yang-Mills equations are the field equations of Yang-Mills theory, a field theory which I will introduce shortly. Instantons are solutions to these equations which have finite action. The requirement of finite action implies that the Lagrangian density of an instanton has to be localised in both space and time (hence the name "instanton"). Instantons have many uses in physics (see for example [13]), but we are mainly interested in them because of their connection with the monopoles we will be dealing with in chapter 3 and onward. These in turn are solitons of the so called Yang-Mills-Higgs equations.

The aim of this chapter then, is twofold. Firstly, we want to introduce some concepts from Yang-Mills theory which will be important in the rest of this thesis and secondly, we want to describe a method to find instanton solutions which we will later turn out to be of use in the search for monopoles.

2.1 Yang-Mills Theory

Yang-Mills theory is an example of a gauge theory. Mathematically, this is a principal bundle, that is, a fibre bundle which has a Lie group G as its fibre. In physical terms, Yang-Mills theory can be seen as an extension of electromagnetism. The most important ingredients for a gauge theory are a connection A on the principal bundle and a curvature F associated with this connection. In analogy with electromagnetism, A is called the vector potential (or gauge potential) and F is called the (gauge) field. The base space for the bundle is identified with physical space time and we shall take it to be $\mathbb{R}^4$. A and F can thus be seen as fields in space time.

For physical purposes, it is not very practical to describe gauge theories in terms bundles with Lie groups for fibres. In practice, one usually works with vector bundles in which the fibre is a vector space $\mathbb{C}^n$ on which the group G acts through a representation. One can then describe A, F and any other objects one wants to have in the theory in terms of coordinates on the vector spaces $\mathbb{C}^n$. Of course, before one can use coordinates, one has to choose a basis for each fibre in the bundle. Such a choice of bases is called a gauge. Gauges are not invariant under the action of the group G. One can go from one gauge to another by means of coordinate transformations with matrices in the chosen representation of G. This is called performing a gauge transformation. It should be clear that gauge transformations will only change the coordinate description of physical quantities and not the actual quantities themselves. Therefore, connections and fields which can be gauge transformed onto each other will be identified for physical purposes.

Let us make all this a bit more tangible. First, we define a gauge potential. This is a set of functions A_μ ($\mu = 1 \ldots 4$) on $\mathbb{R}^4$, which take values in a representation of the algebra of the group G. Here, we see G as a group of matrices and thus the A_μ will just be matrix valued functions. Under a gauge transformation, A_μ must transform as follows:

$$A_\mu \to g^{-1} A_\mu g + \frac{1}{e} g^{-1} \partial_\mu g,$$

where $g(x)$ is a differentiable G-valued function on $\mathbb{R}^4$ and e is the fundamental electric charge Once we have a potential, we can define a gauge covariant derivative as follows:

$$\nabla_\mu = \partial_\mu + e A_\mu$$

The derivative ∇ will act on sections of the vector bundle. These can be seen as functions on $\mathbb{R}^4$ with values in $\mathbb{C}^n$, where n is the dimension of the representation chosen for G. The transformation rule for A is chosen so that the derivative ∇ is in fact covariant, that is, so that $\nabla_\mu g\psi = g\nabla_\mu \psi$, with ψ an arbitrary (differentiable) vector function on $\mathbb{R}^4$.

The gauge field $F_{\mu\nu}$ associated with the potential A_μ is defined by

$$F_{\mu\nu} = \frac{1}{e}[\nabla_\mu, \nabla_\nu] = (\partial_\mu A_\nu) - (\partial_\nu A_\mu) + e[A_\mu, A_\nu].$$

its behaviour under gauge transformations is given by

$$F_{\mu\nu} \to g^{-1} F_{\mu\nu} g$$

which follows from the transformation rule for A. Note that F commutes with gauge transformations, in correspondence with the requirement that physical observables must be independent of the chosen gauge.

We will also need a covariant derivative D_μ on matrix valued functions which behave like $F_{\mu\nu}$ under gauge transformations. This is defined as follows:

$$D_\mu \Phi = [\nabla_\mu, \Phi] = \partial_\mu \Phi + e[A_\mu, \Phi]$$

Covariance of this derivative is easily checked: one has $D_\mu(g\Phi g^{-1}) = g D_\mu \Phi g^{-1}$ so that $D_\mu \Phi$ transforms the way Φ does, as desired.

Now that we have introduced all this machinery that is common to all gauge theories, We can define an action for YM-theory: [1]

$$S = -\frac{1}{2} \int d^4x \, \mathrm{Tr} \left\{ F_{\mu\nu} F^{\mu\nu} \right\} \tag{2.1}$$

[1] Note that, for convenience, we will set $e = 1$ in the rest of this chapter.

By the cyclic property of the trace this action is invariant under gauge transformations, so the physics of the theory is independent of the choice of gauge. The field equations are now

$$D_\mu F_{\nu\sigma} + D_\nu F_{\sigma\mu} + D_\sigma F_{\mu\nu} = 0 \tag{2.2}$$
$$D_\mu F_{\mu\nu} = 0 \tag{2.3}$$

The first of these equations is just the Bianchi identity of differential geometry and it follows easily from the definition of $F_{\mu\nu}$. and D in terms of ∇ that this equation is satisfied independent of the form of the action. The second equation is specific to this theory and is therefore called the Yang-Mills equation. In the next section, we will study the relation between the Yang Mills equation and self-(anti-)duality of the field F.

2.2 Self-Duality

If we define $^*F$, the dual of F, as follows:

$$^*F_{\mu\nu} = \epsilon^{\mu\nu\rho\sigma} F_{\rho\sigma} \tag{2.4}$$

then (2.3) can be rewritten as:

$$D_\mu {}^*F_{\nu\sigma} + D_\nu {}^*F_{\sigma\mu} + D_\sigma {}^*F_{\mu\nu} = 0$$

thus assuming the same form as equation (2.2), only for $^*F$ in stead of F. Now if we can find a potential A so that F is self-dual or anti self-dual (that is, $^*F = F$ or $^*F = -F$ resp.), then both field equations will be satisfied; the first because it is a consequence of the way F is constructed from A and the second because it follows from the first with (anti) self-duality. It turns out that if a Euclidean metric is used on $\mathbb{R}^4$, it is also possible to prove a kind of converse to this result for finite action solutions to the field equations. I will proceed to do this.

First, it is of interest to know under what conditions on the potential A the action integral converges. It is obvious that a necessary condition for convergence is that $F(x)$ tends to zero as x tends to infinity. It is also clear that this condition will be satisfied if $A(x)$ and its first derivatives tend to zero for x to infinity. However, because the condition that F tend to zero is gauge invariant, we can also take A tending to any gauge transform of to zero instead of zero itself. Thus the we demand:

$$A_\mu(x) \to g^{-1}(x)\partial_\mu g(x) \qquad (x \to \infty) \tag{2.5}$$

It can actually be proved that this is implied by the requirement that F approach zero on the edge of space, so it follows that this condition on A is necessary for convergence of the action. Several interesting remarks can be made about this condition.

First of all, the gauge transformation g mentioned in it does not necessarily have to be defined for all values of x, but only for large values, that is values outside a three-sphere of large radius, say R. In fact, if we restrict g to a sphere of radius larger than R, then we get a function of S^3 into the gauge group G, which will have an integer topological invariant k analogous to the degree of a map of S^3 into S^3. This invariant will have the same value on any sphere of radius greater than R, because all these restrictions of g, when pulled back to the sphere of radius $R+1$ by orthogonal projection will give continuous deformations of the restriction of g to this particular sphere. Therefore we can say that g itself has an integer topological invariant k (continuous transformations of g induce continuous transformations of the restrictions of g to the aforementioned spheres). It can be shown that g can be globally defined on $\mathbb{R}^4$ if and only if k is equal to zero.

Second, the condition (2.5) can be incorporated in the condition that A_μ be extendable to a potential on the four dimensional sphere obtained by adding to $\mathbb{R}^4$ a point at infinity. In fact, if we use the inverse of stereographic projection to map $\mathbb{R}^4$ into S^4 then it is not difficult to see that the requirement of finite action on $\mathbb{R}^4$ is equivalent to the requirement that the integral over S^4 of the volume form corresponding to the action be finite. Because S^4 is compact, this will certainly be the case if this form can be extended to an integrable volume form on all of S^4. Of course, in principle, one could still have potentials which give finite action on $\mathbb{R}^4$, but can not be continuously extended to S^4. However, it turns out that this does not happen. For a proof, see [14]. In the rest of this chapter we shall therefore assume our potential to be continuous and defined on S^4; this is equivalent to requiring finiteness of the action.

On S^4 there is a theorem relating the integer invariant k mentioned above to an integral expression in the field F. For the group $SU(2)$, this is

$$8\pi^2 k = \int_{S^4} \mathrm{Tr}\,\{F \wedge F\} \tag{2.6}$$

For other groups one has to take suitable multiples of k. The wedge product in the equation can be calculated by taking for the antisymmetric tensor F the antisymmetric differential form associated with it. In local coordinates, this gives:

$$(F \wedge F)_{\mu\nu\rho\sigma} = \frac{1}{4}\epsilon^{\mu\nu\rho\sigma} F_{\mu\nu} F_{\rho\sigma} \tag{2.7}$$

Using the same wedge product, we can rewrite the action (2.1) as follows:

$$S = -\int_{S^4} \mathrm{Tr}\,\{F \wedge {}^*F\} \tag{2.8}$$

If we now rewrite F in terms of its self-dual and its anti-self-dual part:

$$F = F^+ + F^- = (F + {}^*F) + (F - {}^*F)$$

and if we note that the trace of the wedge of a self-dual and an anti-self-dual operator is equal to zero, then we can rewrite the action (2.8) to give

$$S = \int d^4x\, \mathrm{Tr}\left\{F^- \wedge F^- - F^+ \wedge F^+\right\}$$

We can also rewrite the equation (2.6) for the instanton number in terms of F^+ and F^-. This yields

$$8\pi^2 k = \int \mathrm{Tr}\left\{F^+ \wedge F^+ + F^- \wedge F^-\right\}$$

Combining these formulae, we find

$$S = 8\pi^2 k - 2\int d^4x\, \mathrm{Tr}\left\{F^+ \wedge F^+\right\} = -8\pi^2 k + 2\int d^4x\, \mathrm{Tr}\left\{F^- \wedge F^-\right\}.$$

Since both of the trace terms that occur in this equation are positive, it follows that we have

$$S \geq 8\pi^2 |k|,$$

with equality if ${}^*F = -\mathrm{sign}(k)F$. We thus find that if we fix any value for k, a potential with this k which makes the action attain its minimal value of $8\pi^2|k|$ will give rise to a self-dual field for k negative and an anti-self-dual field for k positive.

2.3 Constructing Potentials by Embedding

In this section I will describe a method to construct potentials for vector bundles that applies to many interesting cases. In the next section, I will apply this method to construct self-dual potentials on S^4.

Suppose that we have a vector bundle E of rank n over a manifold X, which can be embedded in the trivial bundle $X \times \mathbb{R}^N$ $(N \geq n)$. With this, we mean that at each point x of X, we can identify the fibre E_x over x with an n-dimensional subspace of $\mathbb{R}^N$ in a way which depends differentiably on x. In this situation, a section f of E can be identified with a function F of X into $\mathbb{R}^N$. A connection ∇ on E can then be defined through:

$$(\nabla_\mu f)(x) = P(x)\partial_\mu F \tag{2.9}$$

Where for every x, $P(x)$ is a linear projection of $\mathbb{R}^N$ onto the subspace of $\mathbb{R}^N$ that corresponds to the fibre over x. Of course, $P(x)$ will also be required to vary smoothly with x. If no further conditions are imposed on P, then ∇ will just be a $GL(n, \mathbb{Q})$ connection, where $\mathbb{Q}$ corresponds to the real numbers, the complex numbers, or the quaternions [2], depending on whether the manifold X is real, complex or quaternionic. If we do impose further requirements on P however, then we can get connections and potentials for subgroups of $GL(n, \mathbb{Q})$. In particular, if we require that $P(x)$ is orthogonal projection for every x, then we get an $O(n)$, $U(n)$, or $Sp(n)$ potential, depending on $\mathbb{Q}$. The last statement can be explicitly checked if we write the connection in terms of a gauge for the bundle E, as we will now do.

[2] see the appendix for a definition and some more useful information about these

A gauge $u(x)$ for E is a choice of basis for the fibre over x that varies differentiably with x. In this case, this means that for every x, we can take $u(x)$ to be an injective linear map from $\mathbb{R}^n$ into $\mathbb{R}^N$ that has the fibre over x as its image. If we require $u(x)$ to fix the inner product, then the orthogonal projection $P(x)$ will be given by $P = uu^*$, while $u^*u = 1$. A section f of the bundle E can now be written in the form $f(x) = u(x)g(x)$, where g is a function on X with values in $\mathbb{R}^n$. This makes it possible for us to compute the potential A_μ corresponding to our connection. We have:

$$\nabla_\mu(ug) = uu^*\partial_\mu(ug) = u(\partial_\mu(g) + u^*(\partial_\mu u)g)$$

Which shows that the gauge potential is given by

$$A_\mu = u^*\partial_\mu u \tag{2.10}$$

Mind that this formula does not imply that A_μ is gauge equivalent to zero, unless $N = n$. If we now note that $u^*u = 1$ implies that $u^*\partial_\mu u = -\partial_\mu(u^*)u = -\partial_\mu(u)^*u$, then we see that $A^* = -A$ and thus that A_μ does indeed belong to the desired Lie algebra.

From the formula for the potential we find for the field:

$$F_{\mu\nu} = \partial_\mu u^*\partial_\nu u - \partial_\nu u^*\partial_\mu u + [u^*\partial_\mu u, u^*\partial_\nu u] \tag{2.11}$$

This formula gives the field in terms of $n \times n$ matrices, as one would expect for the field of a bundle of rank n. However, because we have embedded E in $X \times \mathbb{R}^N$, it is really more natural in this case to work with fields and potentials in terms of the embedding, i.e. in terms of $N \times N$ matrices. That way, we can actually do all necessary calculations without ever choosing a gauge. I will now show how to do this, by recalculating the gauge field.

Let $Q = I - P$ be the complementary projection to P and define a potential A' and corresponding covariant derivative ∇' on $X \times \mathbb{R}^N$ by

$$\begin{aligned} A'_\mu &= Q\partial_\mu Q \\ \nabla'_\mu &= \partial_\mu + A_\mu \end{aligned}$$

If we let ∇' work on a vector function f which takes values in the fibres of E, i.e. for which $Pf = f$ and thus $Qf = 0$, then we see that this covariant derivative is equal to ∇ on E:

$$\begin{aligned} \nabla'_\mu f &= \partial_\mu f + Q\partial_\mu Qf = \partial_\mu f + Q\partial_\mu Qf - Q\partial_\mu(Qf) \\ &= \partial_\mu f - Q^2\partial_\mu f = (1-Q)\partial_\mu f = P\partial_\mu f = \nabla_\mu f \end{aligned} \tag{2.12}$$

The expression for the field $F'_{\mu\nu}$ corresponding to A' is now calculated in the usual way, through the formula

$$F'_{\mu\nu} = (\partial_\mu A'_\nu) - (\partial_\nu A'_\mu) + [A'_\mu, A'_\nu].$$

This calculation is considerably simplified if one realises that the commutator term vanishes. This can be shown using $\partial_\mu Q = Q\partial_\mu Q + (\partial_\mu Q)Q$, which in turn follows from $Q^2 = Q$. The final result of the calculation is given by:

$$F'_{\mu\nu} = \partial_\mu Q \partial_\nu Q - \partial_\nu Q \partial_\mu Q \tag{2.13}$$

We will now check explicitly that this gives the same field as the equation (2.11). On the way, we will obtain a useful formula for the gauge field which we will need later. To do our check, let us look at the restriction F of F' to the sub-bundle of $X \times \mathbb{R}^n$ which we have identified with E. We can multiply F' by P both on the left and on the right without changing anything. Multiplying on the right changes nothing because every vector function f which takes values "in E" will have $Pf = f$. Multiplying on the left nothing changes because on E, we have

$$\begin{aligned} PF_{\mu\nu}f &= P[\nabla'_\mu, \nabla'_\nu]f = P(\nabla'_\mu \nabla_\nu f - \nabla'_\nu \nabla_\mu f) = P(\nabla_\mu \nabla_\nu f - \nabla_\nu \nabla_\mu f) \\ &= P(P\partial_\mu(\nabla_\nu f) - P\partial_\nu(\nabla_\mu f)) = P\partial_\mu(\nabla_\nu f) - P\partial_\nu(\nabla_\mu f) = [\nabla_\mu, \nabla_\nu]f = F_{\mu\nu}f \end{aligned}$$

where we have used (2.12) in the second and third step and $P^2 = P$ in the fifth. Now that we have established that $F = PFP$, we can write

$$F_{\mu\nu} = P(\partial_\mu Q \partial_\nu Q - \partial_\nu Q \partial_\mu Q)P = P(\partial_\mu P \partial_\nu P - \partial_\nu P \partial_\mu P)P \tag{2.14}$$

using (2.13) and $Q = I - P$. At this point we can choose a gauge u. If we then substitute $P = (uu^*)$ in (2.14), we get

$$F_{\mu\nu} = u\,[u^*(\partial_\mu(uu^*)\partial_\nu(uu^*) - \partial_\nu(uu^*)\partial_\mu(uu^*))u]\,u^*$$

and after some algebra, the expression in the square brackets reduces to (2.11) as expected.

The useful formula for the field that I promised we would obtain during this calculation is the formula (2.14). The reason that it is useful, is the fact that it makes it easy to write the field in terms of a gauge for $E^\perp$, the bundle complementary to E. This is by definition the bundle that X as its base space and for which the fibre over $x \in X$ is the orthogonal complement in $\mathbb{Q}^n$ of the fibre over x in E. If v is an orthogonal gauge for $E^\perp$, then we will have $Q = vv^*$. Substituting this into the second member of (2.14), we get

$$F_{\mu\nu} = P(\partial_\mu(vv^*)\partial_\nu(vv^*) - \partial_\nu(vv^*)\partial_\mu(vv^*))P.$$

Using $Pv = 0$ and $v^*P = 0$, this can be simplified to

$$F_{\mu\nu} = P(\partial_\mu v \partial_\nu v^* - \partial_\nu v \partial_\mu v^*)P$$

This formula only holds if the gauge used is orthogonal. However, if we have a linear, but not orthogonal gauge v, then we can still obtain an analogous formula, using a polar decomposition of v. Such a decomposition is given by

$$v = \alpha\rho$$

where α is an $N \times k$ matrix which satisfies $\alpha^*\alpha = I_k$. Any matrix whose columns contain an orthonormal base for the column space of v will fulfill this criterion and therefore, we can find the columns for a good α by orthogonalising the columns of v. Once we have a good α, it follows automatically that $\rho^2 = v^*v$. Also, α will in fact be an orthogonal gauge for the bundle V and so we will have $Q = \alpha\alpha^* = v\rho^{-2}v^*$. As before, we can now substitute for Q in (2.14) to obtain

$$F_{\mu\nu} = P(\partial_\mu v\rho^{-2}\partial_\nu v^* - \partial_\nu v\rho^{-2}\partial_\mu v^*)P \tag{2.15}$$

This formula will be crucial for the construction of self-dual fields in the next section. Note that we do not need the explicit form of the matrices ρ and α to calculate F; knowledge of v suffices.

2.4 Construction of $Sp(n)$ self-dual potentials

In this section, we will apply the general construction of potentials which we have just introduced to the case where the base manifold X of the bundle is S^4, or equivalently, $P_1(\mathbb{H})$. This will lead to a method for the construction of self-dual potentials for the gauge groups $Sp(n)$ and for arbitrary instanton number k. This method has been named the ADHM-construction, after Atiyah, Hitchin, Drinfeld and Manin, who discovered it [15]. At the end of this section we will formulate a theorem that states that the given construction actually yields all possible self-dual field configurations. The material in this section assumes some familiarity with the quaternions $\mathbb{H}$. Readers unfamiliar with these can consult appendix A for some useful information about these.

Let us start by defining a matrix-valued function v on $\mathbb{H}^2$:

$$v : (x, y) \mapsto Cx + Dy \tag{2.16}$$

Here (x, y) are the quaternionic coordinates on $\mathbb{H}^2$, while C and D are constant $(k+n) \times k$ matrices of quaternions. Let us now make the following assumption:

$$\text{Rank}(v(x, y)) = k \quad \text{for all } (x, y) \neq (0, 0) \tag{2.17}$$

If this is satisfied, then the columns of $v(x, y)$ will span a k-dimensional subspace $V_{x,y}$ of $\mathbb{H}^{k+n}$. It is clear from the definition of v that this subspace depends only on the quotient xy^{-1}. This means we can see v as a way of associating a k dimensional quaternionic vector space to each point in $P_1(\mathbb{H})$. We simply take (x, y) to be the collective coordinates of a point in $P_1(\mathbb{H})$ and associate to this point the vector space $V_{x,y}$. This recipe thus gives us a quaternionic k-vector bundle V over $P_1(\mathbb{H})$ and with that a gauge, given by the columns of v. Moreover, V is automatically embedded into $P_1(\mathbb{H}) \times \mathbb{H}^{n+k}$ and thus we also have the orthogonal bundle $E = V^\perp$, which naturally has quaternionic dimension n.

Let us now give E the covariant derivative induced by the orthogonal projection P, as in (2.9). After choosing a gauge for V, the curvature F can then be calculated using (2.15). Producing such a gauge is not difficult: In all points (x, y) of $P_1(\mathbb{H})$ which have $y \neq 0$, the matrix $v(x, 1)$ will do fine. Substituting this in (2.15) gives

$$F = PC\left\{(\partial_\mu x)\rho^{-2}\partial_\nu \bar{x} - (\partial_\nu x)\rho^{-2}\partial_\mu \bar{x}\right\} C^* P \tag{2.18}$$

where $\rho^2 = v^* v = (\bar{x}C^* + D^*)(Cx + D)$ and the bar denotes quaternionic conjugation. Now suppose that ρ^2 is a matrix with only real entries. In other words, suppose

$$(\bar{x}C^* + D^*)(Cx + D) \;\; \text{is a real matrix} \tag{2.19}$$

In that case, its inverse ρ^{-2} will also be a real matrix and hence it will commute with the scalar quaternions $\partial_\mu x$ and $\partial_\nu x$ in (2.18). This means we can write

$$F_{\mu\nu} = PC\rho^{-2}\left\{(\partial_\mu x)\partial_\nu \bar{x} - (\partial_\nu x)\partial_\mu \bar{x}\right\} C^* P. \tag{2.20}$$

It is now easy to see that F's behaviour under duality will be completely determined by the factor in curly brackets. But this term is self dual ! Hence, it follows that any v which is defined as in (2.16) and which satisfies the conditions (2.17) and (2.19) will give rise to a multi-instanton for the group $Sp(n)$.

At this point, we have the field for an instanton solution, but we don't have the potential yet. Also, the field F is still given in terms of $n + k \times n + k$ matrices, instead of $n \times n$ matrices. Both these problems can be resolved by choosing an orthogonal gauge for the bundle E. This is just a $k + n \times n$ matrix $u(x)$ which satisfies the conditions

$$\begin{aligned} u^* v &= 0 \\ u^* u &= I_{n\times n}. \end{aligned} \tag{2.21}$$

Once we have such a u, we can write the potential and field in terms of $n \times n$ matrices using formulae (2.10) and (2.11)

Of course, we still have to verify that the instantons we have found will have the desired value k for the instanton number. Though it is possible to check this by direct calculation, it is more aesthetically pleasing to look at the question from a higher mathematical viewpoint. The topological invariant we call the instanton number can be proved to be additive for direct sums of bundles (see for example [16]). We will not do this here, but we do notice that it can be easily seen from the integral formula (2.6) in the special case where $F_{\mu\nu}$ is block diagonal with the curvatures of the summand bundles as blocks. Now notice that the direct sum of the bundles E and $V = E^\perp$ is just the trivial bundle $S^4 \times R^{4(n+k)}$. Therefore, checking that E has instanton number k is equivalent to checking that V has instanton number $-k$. But V is by definition a direct sum of k quaternion line bundles: Their fibres over the point (x, y) of $P_1(\mathbb{H})$ are just the lines given by all quaternion multiples of the columns of $v(x, y)$. Now each of these line bundles can be identified with the standard line bundle over $P_1(\mathbb{H})$ which associates to (x, y) the line of all scalar (quaternion) multiples of (x, y). This bundle has instanton number ± 1, depending on conventions. In our case, it has to be -1, which leads us to conclude that V has instanton number $-k$ and therefore, E has instanton number k.

Let us now look back at the definition of the matrix $v(x,y)$ in (2.16). Though this definition was a convenient starting point for the construction of instantons, it also introduces a lot of unnecessary parameters. These correspond to the fact that there are many different choices of the matrices C and D that will make the columns of v span the same subspace of $\mathbb{H}^{n+k}$. In other words: the parameters in our construction do not only account for the gauge freedom of the bundle g, but also for the gauge freedom on its complement V. The extra parameters can be removed by choosing v to be of the following form:

$$v(x) = \begin{pmatrix} \Lambda \\ B - xI_{k\times k} \end{pmatrix} \tag{2.22}$$

where Λ is a constant $n \times k$ matrix and B is a constant $k \times k$ matrix. It is shown in [17] that this can be done without loss of generality. With v in this form, the reality condition (2.19) takes the form

$$\Lambda^*\Lambda + B^*B - B^*x - \bar{x}B + \bar{x}x \quad \text{is a real matrix.}$$

But, taking into account that $\bar{x}x$ is always real and that the above condition has to hold for all x, including $x = 0$, we see that this is equivalent to the two conditions

$$\begin{aligned} \Lambda^*\Lambda + B^*B \qquad & \text{is a real matrix} && (2.23) \\ B^*x + \bar{x}B \qquad & \text{is a real matrix for all } x \in \mathbb{H} && (2.24) \end{aligned}$$

But one easily sees that the last of these conditions is equivalent to requiring:

$$B \text{ is a symmetric matrix} \tag{2.25}$$

The regularity condition (2.17) for v can also be written in terms of Λ and B. This gives

$$((B - x)\xi = 0) \wedge (\Lambda\xi = 0) \Rightarrow \xi = 0 \tag{2.26}$$

where ξ is a k vector of quaternions.

Now it turns out that if we have v in the form (2.22) we can actually find a u which satisfies the conditions (2.21) explicitly. We start from the following ansatz for u :

$$u(x) = \begin{pmatrix} -I_{n\times n} \\ U \end{pmatrix} \sigma \tag{2.27}$$

Here, U is a $k \times n$ matrix and σ is a self adjoint $n \times n$ matrix. With u in this form, the equations (2.21) become

$$\begin{aligned} -\Lambda + U^*(B - xI) &= 0 \\ \sigma^{-2} &= I + U^*U \end{aligned} \tag{2.28}$$

As long as $B - xI$ is invertible, we can solve the first equation for U^*. This gives

$$U^* = \Lambda(B - xI)^{-1}. \tag{2.29}$$

Of course, this then also gives us U and σ and hence, we have determined u completely. Substituting this u into (2.10), we obtain A_μ in terms of U and σ :

$$A = \sigma U^* \partial_\mu (U\sigma) + \sigma^{-1}\partial_\mu \sigma. \tag{2.30}$$

This expression will have singularities at points where $B - xI$ is not invertible, but these singularities are not physical. They arose because we forced the gauge u to be of the form (2.27) and we can remove them by an appropriate gauge transformation.

We now have a construction of instanton solutions for all symplectic gauge groups and arbitrary instanton number. However, from the above, it is absolutely not clear whether or not we will find all instantons in this way. In fact, at first sight, this would seem very unlikely. Still, it is true, as was proved by Atiyah, Hitchin, Drinfeld and Manin [15]. We will not give the proof here, as it would require mathematics beyond the scope of this thesis. [3] We will however state the result for $Sp(1) = SU(2)$ [15, 18, 19] in the form of a theorem, as promised at the start of this section. This theorem also states when instanton solutions derived from different matrices v are gauge equivalent.

Theorem 1 *Every k-instanton for $SU(2)$ arises from quaternionic parameters (Λ, B), where Λ is a k vector and B is a symmetric $k \times k$ matrix. Λ and B are required to satisfy the reality condition (2.23) and the regularity condition (2.26). In a certain gauge, the potential corresponding to (Λ, B) will be given by*

$$A_\mu = \frac{\mathrm{Im}(U^*\partial_\mu U)}{1 + U^*U} \tag{2.31}$$

where U is equal to $[\Lambda(B - xI)^{-1}]^*$. The singularities in A_μ, which arise when $(B - xI)$ is singular, can be removed by a gauge transformation. The potentials defined by (Λ, B) and (Λ', B') are gauge equivalent if and only if $\Lambda' = q\Lambda T$ and $B' = T^{-1}BT$, with $q \in Sp(1)$ and $T \in O(k)$.

The formula for A_μ given in this theorem can be easily derived from (2.30), if one realizes that, for the case of $Sp(1) = SU(2)$, σ is actually just a real number. From the fact that A_μ can be written as an imaginary quaternion, we can see explicitly that $A_\mu(x) \in su(2)$, as we would want for an $SU(2)$ gauge potential. [4]

2.5 Some explicit instanton solutions

After all the hard work we did to describe the general construction of instantons in the previous section, it would be a shame not to use it. Let us start with the easiest case

[3] The interested reader may find a description of the proof which is somewhat less technical than the original articles in [17]

[4] see appendix A for details about the identification of $su(2)$ with the imaginary quaternions

to which the construction applies: that of the SU(2) instanton with $k = 1$. In this case, the parameters Λ and B are both just (scalar) quaternions. This means that the reality condition (2.23) is automatically satisfied. Moreover, we see from theorem 1 that we can use the gauge ambiguity to choose Λ to be a real number. Taking Λ to be a real number unequal to zero, we also automatically satisfy the regularity condition (2.26). We now have five parameters left, which describe all solutions with instanton number one. It is easy to see that four of these, the ones contained in B, just correspond to a translation of the space time origin. Also, we see that the parameter Λ is just an overall scale factor, which will determine the size of the region of space time in which the influence of the instanton will be felt. Therefore, let us restrict to the solution with $B = 0$ and $\Lambda = 1$. $U(x)$ will then be given by

$$U(x) = -(x^{-1})^* = -\frac{x}{|x|^2} \tag{2.32}$$

For the gauge potential, this gives

$$A_\mu = -(\frac{\partial_\mu \bar{x})x - \bar{x}(\partial_\mu x)}{2|x|^2(|x|^2+1)} = \mathrm{Im}\left[\frac{x^{-1}\partial_\mu x}{1+|x|^2}\right] \tag{2.33}$$

As expected, we have a singularity at $x = 0$. We have said before that we would be able to remove any such singularities by a gauge transformation. In this case, this can easily be shown explicitly. Using the quaternionic identity

$$\mathrm{Im}[f(x)\partial_\mu f(x)] = \left(\frac{f(x)}{|f(x)|}\right)\partial_\mu\left(\frac{f(x)}{|f(x)|}\right),$$

we see that A_μ is asymptotically equal to a gauge transform of zero at $x = 0$. We have

$$A_\mu \sim \phi(x)^{-1}\partial_\mu\phi(x)$$

with $\phi(x) = x/|x|^2$. If we now perform the gauge transformation given by $\phi^{-1} = \bar{\phi}$, we find the following form for A : [5]

$$A_\mu = \frac{\mathrm{Im}[x\partial_\mu x]}{1+|x|^2} \tag{2.34}$$

and we see that A is now regular everywhere on $\mathbb{R}^4$. The field F also takes a nice form in this gauge:

$$F_{\mu\nu} = \frac{\partial_\mu x\partial_\nu\bar{x} - \partial_\nu x\partial_\mu\bar{x}}{(1+|x|^2)^2} \tag{2.35}$$

[5]The one instanton formulae (2.34) and (2.35) we give here were actually found by Belavin et al. before the existence of the the ADHM-construction [20].

We can easily see that this F is indeed self-dual, everywhere regular (even at infinity) and that the Lagrangian density corresponding to F is maximal at the space time origin, where we had said the instanton would be located.

Now let us go on to multisolitons. We will choose the k vector Λ to be real and the $k \times k$ matrix B to be diagonal,i.e.

$$\Lambda = (\lambda_1, \ldots, \lambda_k), \quad B = \mathrm{diag}(b_1, \ldots, b_k), \tag{2.36}$$

where the λ_i are real numbers and the b_i are quaternions. With this choice of B and Λ, the reality condition (2.23) is automatically satisfied. If all the b_i are distinct and all the λ_i are non zero, then the regularity condition (2.26) is also satisfied. Moreover, the number of free parameters involved in the choice of the λ_i and b_i stays the same if we work modulo gauge transformations (see theorem (1)). It follows that the present choice of Λ and B will give us a $5k$ parameter family of instantons. These are the so called *'t Hooft instantons.*

Now let us calculate the gauge potential, using formula (2.31). The vector U in this formula can be easily calculated; we have:

$$U_i = -\lambda_i \frac{x_i - b_i}{|x_i - b_i|^2}. \tag{2.37}$$

We see that each of the U_i is just the same as the function U which determines the potential for a one instanton solution with scale λ_i at location b_i. This has very interesting consequences for the potential A. This is given by

$$A_\mu = \sum_{i+1}^{k} \left\{ \frac{\mathrm{Im}\,[U_i \partial_\mu U_i]}{1 + |U_1|^2 + \ldots + |U_k|^2} \right\} \tag{2.38}$$

Now if we are at a space time location x that is far removed from all the locations b_i, then all the U_i will be small. This means that all the terms $|U_i|^2$ in the denominator of 2.38 will be small and we have

$$A_\mu \sim \sum_{i+1}^{k} \{\mathrm{Im}\,[U_i \partial_\mu U_i]\} \qquad (x \to \infty)$$

from which we see that, asymptotically, the potential A is just the sum of the potentials of k basic instantons, located at b_i and with scales λ_i. Of course, instead of taking x to infinity, we could also have taken the scales λ_i to zero. This confirms our previous statement that the scale of an instanton determines the size of the space time area it influences.

Now let us look near one of the instanton locations b_i. Letting x approach b_i, we see that U_i becomes very large compared to the U_k with $k \neq i$. This means that we have, asymptotically:

$$A_\mu \sim \frac{\mathrm{Im}\,[U_i \partial_\mu U_i]}{1 + |U_i|^2}$$

but this is just the field of one instanton located at b_i. Concluding, we can say that near each of the locations b_i, the field is like that of an instanton located at b_i. Of course this property will be valid in a limited area of space time, which will be largest if the instantons are far apart and/or have small scale parameters, i.e. if the overlap between the instantons is as small as possible.

The remarkable properties shown above confirm the intuitive expectation that one should be able to build multi-instantons by superposition of instantons asymptotically far removed from each other. However, from theorem 1, we can deduce that the family of solutions given here does not nearly cover all of the possibilities. A count of the parameters involved in the general construction gives $8k-3$ parameters for the general k instanton (modulo gauge transformations) and we have only found a $5k$ parameter family here. There seems to be no reason to expect that the other solutions will also be approximate superpositions of single instantons. As we do not want to make Yang-Mills instantons the main focus of this thesis, we shall not go into this question here, but go on to the subject of magnetic monopoles.

Chapter 3

BPS-monopoles and Nahm's construction

3.1 The Georgi-Glashow model

Consider again a gauge theory, this time defined on $\mathbb{R}^4$ with the Minkowskian metric, with Lagrangian density:

$$\mathcal{L} = -\frac{1}{2}\mathrm{Tr}(F_{\mu\nu}F^{\mu\nu}) + \mathrm{Tr}(D^{\mu}\Phi D_{\mu}\Phi) - U(\Phi) \tag{3.1}$$

This is called the Yang-Mills-Higgs Lagrangian density. The gauge group is taken to be $SU(2)$. F and D are defined as in chapter 2, Φ a Lie algebra valued function that behaves like F under gauge transformations and U is a gauge invariant potential. For the potential U, we take

$$U(\Phi) = -\frac{\lambda}{2}\left(\mathrm{Tr}(\Phi^2) - v^2\right)^2 \tag{3.2}$$

with λ and v positive real parameters. The equations of motion are:[1]

$$\begin{aligned} D_{\mu}F^{\mu\nu} &= e[\Phi, D^{\nu}\Phi] & (3.3)\\ D^{\mu}D_{\mu}\Phi &= -\lambda\Phi\mathrm{Tr}(\Phi^2 - v^2). & (3.4) \end{aligned}$$

One of these equations, the analogue of Gauss' law in electromagnetism, is not a real dynamic equation, but rather a constraint on the initial conditions for the theory. We can see this by writing this equation in the following form:

$$D_iD_iA_0 + e[\Phi, [\Phi, A_0]] = D_i\dot{A}_i + e[\Phi, \dot{\Phi}] \tag{3.5}$$

Here, the dot denotes time differentiation. Now one can show that the operator $D_iD_i + e[\Phi, [\Phi,]]$, which works on A_0, is invertible on functions that go to zero sufficiently rapidly

[1] note that we have reinstated the fundamental electric charge e

at the edge of space, if the fields are such that the energy is finite. Therefore, if A_0 satisfies this condition, it is uniquely determined in terms of the other fields and their derivatives. Thus we see that A_0 is not really a dynamic variable independent of the other fields. We could also have found this by calculating the conjugate momentum to A_0. This is identically zero.

Now if we want to choose an initial field configuration for this theory, we can't just take any configuration, but we have to take a configuration which satisfies Gauss' equation above. Configurations which do not satisfy this equation are unphysical. The physical configuration space of the theory is thus the space of field configurations which satisfy Gauss' equation, rather than the space of all field configurations.

Next to the equations of motion, we have the Bianchi identity

$$D_\mu {}^*F^{\mu\nu} = 0 \tag{3.6}$$

which follows from the definition of F.

We want to take the model described by (3.1) a bit more seriously than the physically rather uninteresting models we have studied up to now. For a start, we would like to derive the particle spectrum. To do this, we first rewrite the Lagrangian in terms of coordinates on $su(2)$. [2] We will use Roman indices for these, to distinguish them from the space time coordinates. The Lagrangian density then takes the form

$$\mathcal{L} = -\frac{1}{4}F^a_{\mu\nu}F^{a\mu\nu} + \frac{1}{2}D^\mu\Phi^a D_\mu\Phi^a - \frac{\lambda}{4}(\Phi^a\Phi^a - v^2)^2 \tag{3.7}$$

where the action of D_μ is described by

$$(D_\mu\Phi)^c = \partial_\mu\Phi + e\epsilon^{abc}A^b_\mu\Phi^c. \tag{3.8}$$

A change of gauge can be viewed as a spacetime dependent rotation of the vectors Φ and $F_{\mu\nu}$.

To find the particle spectrum of our model, we have to look at the second order variation of the Lagrangian about the vacuum. If the vacuum corresponded to the configuration with all fields equal to zero, we would have three massless vector bosons corresponding to the A^a and three massive scalars corresponding to the Φ^a. However, if we look closer, we see that the configuration with zero fields does not minimise the energy and can therefore not be identified with the vacuum. To write down an equation for the energy, we first define fields E and B in the usual way.

$$E^{ai} = F^{ai0} \tag{3.9}$$
$$B^{ai} = -\frac{1}{2}\epsilon^{ijk}F_{ajk} \tag{3.10}$$

where the spatial indices i, j, k run from one to three. The energy density is then given by

[2] for details, see section A.2 in the appendix

$$\mathcal{M} = \frac{1}{2}\left\{E_i^a E_i^a + B_i^a B_i^a + D_0\Phi^a D_0\Phi^a + D_i\Phi D_i\Phi\right\} + U(\Phi). \tag{3.11}$$

This expression is the 00-component of the energy-momentum tensor for this theory. The energy-momentum tensor itself can be computed most easily by varying the metric (around the Minkowskian metric) and taking as the energy-momentum tensor the variation of the action as a consequence of this procedure. For a proof that this works, see e.g. [21]

We see that any field configuration which has the A_μ equal to zero and Φ constant and of length v will minimise the energy and therefore be suitable as a vacuum. Of course, any gauge transform of these vacua will also be a vacuum. For now, let us choose the vacuum to be given by $A_\mu^{vac} = 0$ and $\Phi^{vac} = (0, 0, v)$. If we now write

$$\Phi = \Phi^{vac} + \phi$$

and choose the gauge in which $\phi_1 = \phi_2 = 0$, then we can rewrite the Lagrangian as

$$\begin{aligned} \mathcal{L} \;\; = \;\; & -\frac{1}{4}F_{\mu\nu}^a F^{a\mu\nu} + \frac{(ev)^2}{2}A^{\mu 1}A_\mu^1 + \frac{(ev)^2}{2}A^{\mu 2}A_\mu^2 \\ & +\frac{1}{2}\partial^\mu\phi^3\partial_\mu\phi^3 - \frac{1}{2}\lambda v^2(\phi^3)^2 \end{aligned} \tag{3.12}$$

to second order in the fields and their derivatives. From this we conclude that, in stead of three massless vector bosons and three massive scalars, we really have two massive vector bosons, one massless vector boson and a massive scalar. In our present gauge, in which Φ is forced to point in a fixed direction, the massive vector bosons are associated with the components of A normal to Φ. They have classical mass equal to ev. We will refer to these particles as gauge bosons. The massless vector boson is associated with the component of A in the direction of Φ. We will call this particle the photon. The massive scalar associated to Φ is called the Higgs boson. its classical mass is equal to $v\sqrt{\lambda}$. With this particle spectrum, it will not come as a surprise that our model first turned up as an attempt to describe unified electromagnetic and weak interactions. The first article (ref. [22]) in which it appears is by Georgi and Glashow, which is how the model got its name.

3.2 Embedding electromagnetism; monopoles appear

Like the Yang-Mills theory considered in chapter 2, the Georgi-Glashow model is a generalisation of electromagnetism. We have an $su(2)$ valued vector potential A_μ and we have defined $su(2)$ valued E and B fields in (3.9). However, this time we also want to embed electromagnetism itself into the theory. We would prefer to do so by giving an expression for a real number valued electromagnetic vector potential a_μ, as the vector potential is the most fundamental object in electromagnetism. In gauges in which Φ points in the same internal direction everywhere, we have an obvious candidate for this vector potential. We would like to define a_μ as the component of A_μ in the direction of Φ. This would make sense, because we had previously associated this component of A_μ with the photon. However, we can not define a_μ this way for any arbitrary gauge, because this would make the real number valued gauge field $f_{\mu\nu}$ derived from a_μ gauge dependent. This is undesirable, because the electric and magnetic fields contained in f should be physically observable and all observables should be gauge independent. Because of this difficulty, we will not work directly with an electromagnetic vector potential. In stead we will give a gauge invariant expression for an electromagnetic field tensor $f_{\mu\nu}$ which, in a gauge with Φ pointing in a fixed direction, equals the expression we get from antisymmetric differentiation of a_μ. 't Hooft proposed the following form for $f_{\mu\nu}$:

$$f_{\mu\nu} = \mathrm{Tr}\left(\hat{\Phi} F_{\mu\nu}\right) - \frac{1}{e}\mathrm{Tr}\left(\hat{\Phi}[D_\mu\hat{\Phi}, D_\nu\hat{\Phi}]\right) \tag{3.13}$$

where $\hat{\Phi}$ is equal to $\Phi/\mathrm{Tr}(\Phi^2)$. It is easy to see that this definition is gauge invariant and that in regions where $D_\mu\Phi$ vanishes, f is exactly the component of F in the direction of Φ. With some effort, we can also see that

$$f_{\mu\nu} = \partial_\mu(A^a_\nu\hat{\Phi}^a) - \partial_\nu(A^a_\mu\hat{\Phi}^a) - \frac{1}{e}\epsilon^{abc}\hat{\Phi}^a\partial_\mu\hat{\Phi}^b\partial_\nu\hat{\Phi}^c \tag{3.14}$$

Note that the last term is just the volume of the parallellepipid spanned by $\hat{\Phi}, \partial_\mu\hat{\Phi}$ and $\partial_\nu\hat{\Phi}$, divided by e. It follows that this term is zero if Φ is confined to a single plane. Thus, if Φ points in a single direction, then the field tensor f will certainly be equal to the quantity obtained by antisymmetric differentiation of a_μ, as we required.

From (3.13), we can now define the electric and magnetic fields $\mathbf{E}$ and $\mathbf{B}$ in the usual way:

$$\begin{aligned} \mathbf{E}^i &= f^{i0} && (3.15)\\ \mathbf{B}^i &= -\frac{1}{2}\epsilon^{ijk}f_{jk} && (3.16) \end{aligned}$$

We can also define a magnetic current density j_μ and the electric current density j^e_μ in terms of $f_{\mu\nu}$. We take

$$\begin{aligned} j_\mu &= \partial^*_\nu f_{\nu\mu} = -\frac{1}{2}\epsilon_{\mu\nu\rho\sigma}\partial^\nu f^{\rho\sigma} && (3.17)\\ j^e_\mu &= -\partial_\nu f_{\nu\mu} && (3.18) \end{aligned}$$

or equivalently,

$$\begin{aligned} j_0 &= \nabla\cdot\mathbf{B} \qquad & j_i &= -(\nabla\times\mathbf{E})_i - \partial_t\mathbf{B}_i\\ j^e_0 &= \nabla\cdot\mathbf{E} \qquad & j^e_i &= (\nabla\times\mathbf{B})_i - \partial_t\mathbf{E}_i \end{aligned}$$

from which we see that j and j^e are defined exactly so that $\mathbf{E}$ and $\mathbf{B}$ satisfy the Maxwell equations (0.1) for an electromagnetic system which includes magnetic charges. In an ordinary electromagnetic system, these magnetic charges would automatically be zero, because the field tensor would be the antisymmetric derivative of a potential. Here, this is not necessarily the case. If we look at $f_{\mu\nu}$ as given in (3.14), then we see that the last term in this equation may give rise to a non-zero magnetic charge/current density. Substituting (3.14) into (3.17) yields

$$j_\mu = -\frac{1}{2e}\epsilon_{\mu\nu\rho\sigma}\epsilon^{abc}\partial_\nu\hat{\Phi}^a\partial_\rho\hat{\Phi}^b\partial_\sigma\hat{\Phi}^c \tag{3.19}$$

In this expression, all terms with multiple differentiations of one field have disappeared because of the antisymmetry of $\epsilon_{\mu\nu\rho\sigma}$. Using only this same antisymmetry, it is easy to verify that j_μ is a conserved current:

$$\partial^\mu j_\mu = 0.$$

Note that we do not need any information on the time evolution of the field Φ to derive this result, except that it is differentiable. We might therefore expect j_μ to have some topological meaning. In particular, we may hope that the total magnetic charge g is constant on homotopy classes of the field Φ. To prove this, we have to show that the total flux of the magnetic current density j_i through the sphere at infinity is zero. To do this, let us us recall some basic facts:

1. If the field f is the antisymmetric derivative of a potential, then the magnetic charge/current density is zero.

2. In regions of space where Φ points in a fixed direction, f is indeed the antisymmetric derivative of a potential (see formula (3.14 for this))

3. The magnetic charge/current density is gauge independent (this can be read of from (3.17))

4. for every point x in space where $\Phi(x)$ is non-zero, there is a continuous gauge transformation which will make the field Φ point in a fixed direction on an open neighbourhood of x. (If $\Phi(x)$ is zero, the gauge transformation which achieves this may have to be discontinuous)

Combining these facts, we see that the magnetic charge/current density can only be nonzero at locations in space where Φ vanishes. This is an interesting result in itself, but it does not help us to show that g is constant on homotopy classes unless we restrict to the space of field configurations which keep the energy finite. From (3.11) we see immediately that in order for the energy of a field configuration to be finite, $v(\Phi)$ will have to approach zero on the edge of space. More explicitly, we will have to have

$$\Phi^a\Phi^a \to v^2 \qquad (x \to \infty) \tag{3.20}$$

This means that on the "sphere at infinity", $|\Phi|^2(x)$ has to be equal to the non-zero constant v^2 everywhere. Obviously, this implies that Φ is nowhere zero on the sphere at infinity, from which it follows that j_i will be zero on this sphere. Hence, if finite energy is required, the total magnetic charge g will be constant on homotopy classes of the field Φ. In particular, this implies that field configurations with different total magnetic charge will be topologically distinct.

The above analysis does not seem to depend essentially on the fact that the gauge group is $SU(2)$. As long as we can construct an Abelian field tensor which satisfies the condition 2 and as long as Φ can be gauge transformed to point in a fixed direction locally, we will still find that the magnetic charge g is homotopy invariant.

Now when we do make use of the fact that Φ is an element of $su(2)$, which can be interpreted as a 3-vector, we can actually give the homotopy invariant g a nice geometrical interpretation. It follows from (3.20) that $\hat{\Phi}(x)$ will be a continuous function everywhere on the sphere at infinity. So we see that, for every field configuration with finite energy, we will have a corresponding continuous map of spheres given by $x \to \hat{\Phi}(x)$. Such a map will have an integral winding number or degree N, which can not be changed by a continuous deformation of $\hat{\Phi}$. Roughly speaking, N is the number of times the map the surface of the image is covered by the surface of the original (taking orientation into account). It can be shown (see e.g. [23]) that N is equal to the integral over space of the 0-component of a topological current density k, given by

$$k_\mu = \frac{1}{8\pi}\epsilon_{\mu\nu\rho\sigma}\epsilon^{abc}\partial_\nu\hat{\Phi}^a\partial_\rho\hat{\Phi}^b\partial_\sigma\hat{\Phi}^c \tag{3.21}$$

But one sees easily that k is exactly equal to $\frac{e}{4\pi}$ times the magnetic current density j. Thus, there is an intimate connection between the occurrence of magnetic monopoles in this theory and the winding number N: For the total magnetic charge g of a field configuration, we have

$$g = \frac{4\pi N}{e}$$

We will therefore call the winding number N the *monopole number* and if we can find a solution that has monopole number N, we will say that it is in the N-monopole sector.

The idea that magnetic monopoles could arise as topologically non-trivial solitons of gauge theories was first proposed by [3] and [4]. This is why the monopoles of the Georgi-Glashow model are called *'t Hooft-Polyakov monopoles.*

3.3 Bogomol'nyi's equation

In the previous section, we found clues that there may exist solutions to the Georgi-Glashow model which have non-zero magnetic charge. On the other hand, we have not found any such solutions yet. Generally, this is very difficult, because of the complexity of the equations of motion (3.3). In stead of dealing with these directly, we will first derive a lower bound on the energy of a solution with a given electric charge q and magnetic charge g. In the static case, this will also give us an equation for the minimal energy solution with these charges, which turns out to be much easier to solve than (3.3).

We will start by deriving the energy bound. Recalling the expression for the energy given in (3.11), we have for the energy M:

$$M = \frac{1}{2}\int d^3x \left\{E_i^aE_i^a + B_i^aB_i^a + (D_0\Phi)^2 + D_i\Phi^aD_i\Phi^a + 2U(\Phi)\right\}. \tag{3.22}$$

It is easy to see that we have

$$M \geq \frac{1}{2}\int \left\{d^3x E_i^a E_i^a + B_i^a B_i^a + +D_i\Phi^a D_i\Phi^a\right\}, \tag{3.23}$$

with equality if U and $D_0\Phi$ are equal to zero. Rewriting the right hand side, we find

$$\begin{aligned} M &\geq \frac{1}{2}\int d^3x \left\{(E_i^a - \sin\theta D_i\Phi^a)^2 + (B_i^a - \cos\theta D_i\Phi^a)^2\right\} \\ &\quad + \sin\theta \int d^3x \left\{E_i^a D_i\Phi^a\right\} + \cos\theta \int d^3x \left\{B_i^a D_i\Phi^a\right\} \\ &\geq \sin\theta \int d^3x \left\{E_i^a D_i\Phi^a\right\} + \cos\theta \int d^3x \left\{B_i^a D_i\Phi^a\right\} \end{aligned} \tag{3.24}$$

Where θ can be any angle. Now let us concentrate first on the last term in this equation. Using partial integration and the definition of the covariant derivative (see f.e. (3.8)), we find

$$\int d^3x \left\{B_i^a D_i\Phi^a\right\} = -\int d^3x \left\{\Phi^a D_i B_i^a\right\} + \int_{S^2_\infty} d^2S_i \left\{\Phi^a B_i^a\right\} \tag{3.25}$$

where S^2_∞ denotes the sphere at infinity. Now the Bianchi identity (3.6) implies that $D_i B_i^a = 0$ and therefore, the first term on the right hand side disappears. If we require the energy to be finite, then, in the second term, we can write $\Phi = v\hat{\Phi}$, since we are looking at Φ on the sphere at infinity. This leaves us with

$$\int d^3x \left\{B_i^a D_i\Phi^a\right\} = v\int_{S^2_\infty} d^2S_i \left\{\hat{\Phi}^a B_i^a\right\} \tag{3.26}$$

Now let us look back at the original equation for the energy (3.22). We see that finiteness of the energy requires that the integral of $D^i\Phi^a D_i\Phi^a$ over large spheres converges to zero. Together with the fact that $|\Phi| = v$ on S^2_∞, this implies that in the right hand side of (3.26), we can change $\Phi^a B_i^a$ to b_i without changing the result. Thus we finally find

$$\int d^3x \left\{B_i^a D_i\Phi^a\right\} = v\int_{S^2_\infty} b_i d^2S_i = vg \tag{3.27}$$

where g is the total magnetic charge of the field configuration.

If we substitute this result into (3.24) and take $\theta = 0$, we get $U \geq vg$. For negative g, this is of course trivial. However, if we take $\theta = \pi$, we get $U \leq vg$. This means that generally, we can write

$$U \geq v|g| \tag{3.28}$$

This bound is known as the *Bogomol'nyi bound,* as it was Bogomol'nyi who first derived it in his article [24] From the above, it is clear that we have equality in (3.28) if and only if $v(\Phi)$ and $D_0\Phi$ are zero and the *Bogomol'nyi equation*

$$B_i = \mathrm{sign}(g) D_i \Phi \tag{3.29}$$

is satisfied.

Note that at this point, we have not used the equations of motion (3.3) yet. This means that the bound (3.28) has to be satisfied for all field configurations with finite energy, whether they solve the equations of motion or not. This makes it plausible that the Bogomol'nyi bound (3.28) will be satisfied quantum mechanically as well as classically.

If we do use the equations of motion, we can get an even better estimate for the energy, though of course it is not clear a priori whether this will still hold for the quantum mechanical version of the theory. Looking back at the second to last term in equation (3.24), we see that we can rewrite this

$$\int d^3x \, \{E_i^a D_i \Phi^a\} = -\int d^3x \, \{\Phi^a D_i E_i^a\} + \int_{S^2_\infty} d^2 S_i \, \{\Phi^a E_i^a\} \tag{3.30}$$

again, using partial integration and the definition of the covariant derivative. Now the equation of motion (3.3) for F gives us

$$D_i E_i^a = e\epsilon^{abc}(D_0 \Phi_b)\Phi_c \tag{3.31}$$

and using this and the antisymmetry of ϵ^{abc}, we find that the first term in (3.30) is equal to zero. Reasoning in the same way as before equation (3.27), we find that the second term in (3.30) is equal to the total electric charge q, leaving us with

$$\int d^3x \, \{E_i^a D_i \Phi^a\} = vq \tag{3.32}$$

If we now substitute (3.32) and (3.27) into (3.24), we find the bound

$$M \geq vq \sin\theta + vg \cos\theta \tag{3.33}$$

for arbitrary θ. Of course we want to make the choice of theta that maximises the right hand side. It is not difficult to see that this is the angle θ for which we have $\cos\theta = g/\sqrt{q^2+g^2}$ and $\sin\theta = q/\sqrt{q^2+g^2}$, yielding

$$M \geq v\sqrt{q^2+g^2} \tag{3.34}$$

This generalised version of the Bogomol'nyi bound was first found by Coleman et al. [25]. It is clear that in this equation, we have equality if and only if $U(\Phi)$ is zero and the fields satisfy the following equations:

$$\begin{aligned} D_0 \Phi^a &= 0 \\ E_i^a &= \sin(\theta) D_i \Phi^a \\ B_i^a &= \cos(\theta) D_i \Phi^a \end{aligned} \tag{3.35}$$

with θ as above.

Because these equations are easier to deal with than the field equations, it makes sense to look at the version of our theory with U set to zero. In this limit, we can hope to find a solution to the equations of motion which attains the bound (3.34) by solving the generalised Bogomol'nyi equations (3.35). In fact, it is easy to see that any solution to the equations (3.35) which has $\cos(\theta) \neq 0$ will automatically solve the equation of motion for the field Φ and satisfy Gauss' law; both essentially reduce to the Bianchi identity (3.6). It is less easy to see whether a solution to (3.35) will also satisfy the other equations of motion. This will be the case if the E-field is zero. With the E-field equal to zero, the Lagrangian is equal to minus the energy and thus at a maximum. Thus, at least all the solutions to (3.35) which have $cos(\theta) = \pm 1$ will also solve the full field equations. Generally, what we have to do is solve the Bogomol'nyi equations and then check that the obtained solution also satisfies the field equations (3.3). For all solutions we will encounter, this will be the case.

Before we go ahead and solve the Bogomol'nyi equations, a remark about working with U equal to zero is in order: We should ask ourselves whether the idea of monopole sectors still applies in this limit. There would seem to be reason for concern, because the field configurations with nonzero magnetic charge arose as a result of the boundary condition (3.20), which in turn was originally needed to ensure that the potential U (3.2) would approach zero at infinity. Removing U, we would also remove the necessity for this boundary condition. Classically, this does not present a problem; we can just impose the condition. Quantum mechanically, there are problems. First of all, one can not just set U zero; there will be quantum corrections to the classical theory which will create a new, nonzero U [26]. I will come back to this later. Secondly, it is not always possible to impose boundary conditions on quantum fields like Φ for all times. Time evolution can cause a change in the expectation value of a field Φ if there are paths of finite action between configurations with different expectation values for this field. Fortunately, this problem does not occur in our case, because any path that connects configurations with different expectation values for $\mathrm{Tr}(\Phi^2)$ at infinity will automatically have infinite action, even with $v = 0$.

3.4 Solving Bogomol'nyi's equations: Part 1

We have taken some trouble to arrive at the equations (3.35) and it is now natural to ask whether these equations have any (non-trivial) solutions and if so, whether we can describe their most important properties or even construct them explicitly. In this section, we will apply elementary techniques to get some general results and finally a family of dyons with magnetic charge 1. We will restrict ourselves to static solutions.

In the static case, the first of the equations (3.35) reduces to

$$\epsilon^{abc} A_0^b \Phi^c = 0 \tag{3.36}$$

This will be satisfied if an only if we choose A_0 so that it is parallel (or anti-parallel) to Φ. That is:

$$A_0^a(x) = r(x)\Phi^a(x).$$

Substituting this into the the second of the equations (3.35), we obtain the following equation for $r(x)$:

$$(\partial_i r)\Phi^a = (\sin(\theta) - r)\, D_i\Phi^a$$

This is solved by taking $r(x)$ equal to the constant $\sin(\theta)$. Thus, given a time independent solution (A_i, Φ) to the last of the generalised Bogomol'nyi equations (3.35), we can solve the first two equations as well by choosing

$$A_0^a(x) = \sin(\theta)\Phi^a(x). \tag{3.37}$$

The task that remains then, is solving the last of the equations (3.35). Now recall the original Bogomol'nyi equation (3.29) for positive charge:

$$B^i = D^i\Phi \tag{3.38}$$

It is easy to see that a field configuration (A, Φ) will solve this equation if and only if the configuration $(A, \Phi/\cos(\theta))$ solves the equation

$$B^i = \cos(\theta)D^i\Phi \tag{3.39}$$

Thus we only have to look for the solutions of (3.38) to find the solutions of (3.39) for arbitrary θ. There is one subtlety here: the configuration $(A, \Phi/\cos(\theta))$ will no longer satisfy the condition that $|\Phi|$ has to approach v at infinity. However, since we have set the potential U to zero, the value of v can really be chosen arbitrarily. Therefore we can use the configuration $(A, \Phi/\cos(\theta))$, if we take (A, Φ) to be a solution to (3.38) for which $|\Phi|$ goes to $v\cos(\theta)$ at infinity.

At this point, we have reduced the problem of solving the equations (3.35) to that of solving the equation (3.38) for given monopole charge g. Before we go ahead with this, let us look back at the equations of motion (3.3). In the previous section, we noted that a solution to the Bogomol'nyi equations with $\cos(\theta) = \pm 1$ would also solve the equations of motion. Here, we have given a method to construct a solution to Bogomol'nyi's equations with arbitrary given θ from a solution with $\theta = 0$. Now suppose we have a solution with $\theta = 0$. With hardly any algebra, one can then show that the associated solutions to the generalised Bogomol'nyi equations (3.35) will also satisfy the equations of motion. [3] In other words: if we have monopoles of minimal energy, we also automatically have dyons of minimal energy !

[3] Recall the equations of motion for the electromagnetic field:

$$D_\mu F^{\mu\nu} = [\Phi, D_\mu\Phi]$$

Suppose that, for some given $\Phi(x)$, these are solved by $B_i = D_i\Phi$, $E_i = 0$. Now define $\tilde{\Phi}$ by $\tilde{\Phi}(x) = \Phi(x)/\cos(\theta)$. If we change the fields as described above, so that we have $\tilde{B}_i = \cos(\theta)D_i\tilde{\Phi}$, $\tilde{E}_i = \sin(\theta)D_i\tilde{Phi}$, then the contribution of the B-field to the right hand side of the equation of motion will change to $\cos^2(\theta)[\Phi, D_\mu\Phi]$, while at the same time a contribution from the E-field, equal to $\sin^2(\theta)[\Phi, D_\mu\Phi]$ is introduced

These dyons have a remarkable property that allows us to extract some information on their asymptotic behaviour without actually solving Bogomol'nyi's equations. Substituting the generalised Bogomol'nyi equations (3.35) into the energy density (3.11), we see that we have

$$\begin{aligned} \mathcal{M} &= \frac{1}{2}\{E_i^a E_i^a + B_i^a B_i^a + D_i\Phi^a D_i\Phi^a\} \\ &= \frac{1}{2}\left\{\sin^2(\theta) D_i\Phi^a D_i\Phi^a + \cos^2(\theta) D_i\Phi^a D_i\Phi^a + D_i\Phi^a D_i\Phi^a\right\} \\ &= D_i\Phi^a D_i\Phi^a \\ &= \partial_i\Phi^a\partial_i\Phi^a + \sum_i \Phi^a\partial_i^2\Phi^a - \Phi^a(D_iD_i\Phi)^a = \sum_{i=1}^{3}\partial_i^2(|\Phi|^2) \end{aligned} \tag{3.40}$$

where the last equality uses the equation of motion for Φ (i.e. $D_iD_i\Phi = 0$). It follows that the energy density of the dyons depends only on the length of the Higgs field ! Let us now calculate the total energy of a configuration with Higgs field $\Phi(x)$. We have

$$\begin{aligned} M &= \frac{1}{2}\int d^3x\left\{\Delta(|\Phi|^2)\right\} \\ &= \frac{1}{2}\lim_{R\to\infty}\int_{S^2(R)} dS^2\left\{\nabla(|\Phi|^2)\right\} \\ &= \frac{1}{2}\lim_{R\to\infty}\int_{S^2(R)} dS^2\left\{\partial_r(|\Phi|^2)\right\} \\ &= \frac{1}{2}\lim_{R\to\infty} 4\pi R^2\partial_r\overline{|\Phi|^2} \end{aligned} \tag{3.41}$$

where $S^2(R)$ denotes the sphere of radius R and centre at the space origin and where $\overline{|\Phi|^2}$ is the average of $|\Phi^2|$ over this sphere. Filling in the mass formula for a dyon from (3.34) we find that for large r, we will have to have

$$2\pi r^2\partial_r\overline{|\Phi|^2} = v\sqrt{q^2+g^2}$$

and hence, after integration, we find:

$$\overline{|\Phi(r)|^2} \sim v^2 - \frac{v\sqrt{q^2+g^2}}{2\pi r} \tag{3.42}$$

Now it is possible (see [27]) to show that the angular derivatives of $|\Phi(r)|$ are of order r^{-2} as r goes to infinity and from this, it follows that the asymptotics for the average of $|\Phi|^2$ which we have found here are actually good for $|\Phi|^2$ itself (to order r^{-2}). We conclude that for a solution to Bogomol'nyi's equations with charges (g, q), we will have

$$|\Phi(r)| \sim v\left(1 - \frac{\sqrt{q^2+g^2}}{4\pi v r}\right) \qquad (r\to\infty) \tag{3.43}$$

for a pure monopole solution, this reduces to

$$|\Phi(r)| \sim v\left(1-\frac{N}{ver}\right) \qquad (r\to\infty) \tag{3.44}$$

where N is the monopole number. This asymptotic behaviour of the Higgs field turns out to be very important for the dynamics of BPS-monopoles. The $\frac{1}{r}$-term in the asymptotic expansion signals the fact that the Higgs field mediates a long range force. This is possible because we have taken the limit of vanishing potential and in this limit the Higgs particle is massless (in tree approximation, that is). For monopoles, this long range force turns out to be equal in magnitude to the coulomb force, but always attractive [28]. This means that it doubles the force between two monopoles of opposite magnetic charge, while cancelling the force between monopoles of like magnetic charge. Without this cancellation of long range forces, it would be impossible to build multimonopole solutions from well separated monopoles of charge 1. With it, this does turn out to be possible. We will give a more detailed discussion of the forces between monopoles/dyons in section 4.5

Now we have seen how we can obtain dyons from monopoles and we have given an important result on the asymptotic behaviour of the Higgs field, but at we still haven't seen any actual, real magnetic monopoles yet. To find these, we would have to solve the equation (3.38) and this in itself is still a formidable problem, especially for high monopole numbers. For the one monopole sector, the situation is somewhat simpler. Therefore, the rest of this section will be devoted to a brief sketch of the construction of a one monopole solution. We will come back to the construction of multimonopole solutions in section 3.5.

We start the construction by introducing the following "ansatz" for the fields:

$$\Phi^a(\bar{x}) = \delta_i^a \frac{x^i}{er^2} H(r) \tag{3.45}$$

$$A_i^a(\bar{x}) = \epsilon_{ij}^a \frac{x^j}{er^2} K(r) \tag{3.46}$$

where $r = |\vec{x}|$ and H and K are required to fulfill the following boundary conditions:

$$\begin{aligned} \frac{H(r)}{r} &\to v & \qquad (r\to\infty) \\ K(r) &\to 0 & \qquad (r\to\infty) \end{aligned}$$

where v is the asymptotic length of the Higgs field as in (3.20).

We use the above ansatz because it has several nice features:

- First, a little calculation shows that the ansatz (3.46) for A and the corresponding boundary condition result in a magnetic field that approaches that of a monopole with charge $\frac{4\pi}{e}$ as x approaches infinity, as we would expect for a solution of the field equations in the one monopole sector.
- Secondly, from (3.45) it is immediately clear that Φ will have winding number N equal to one; the ansatz for Φ^a is just so that, for any vector y on a large sphere

around the origin, the internal vector $\Phi(y)$ will point in the same direction as y (if we identify the internal and space coordinates). We can visualise the map from the "sphere at infinity" to the internal sphere by thinking of a frightened hedgehog rolled up into a ball, all its pins pointing outward.

- Furthermore, (3.45) and (3.46) are invariant under the transformations performed by combining of "the same" rotations in internal and ordinary space. This means that, for all physical purposes, our ansatz is rotationally symmetric; if we rotate it, we can undo the effect of the rotation by a gauge transformation. We perceive this to be a nice feature of our ansatz, because in general, one often finds that field configurations that minimise the energy have a lot of symmetry and vice versa.
- Finally, the undetermined functions H and K depend only on the single variable r. It is this feature that is especially nice in the following, because it allows us to turn partial differential equations into ordinary ones.

Substituting (3.45) and (3.46) into the Bogomol'nyi equation (3.29), we get:

$$\begin{aligned} v^2e^2r\frac{d}{dr}K &= -KH \\ v^2e^2r\frac{d}{dr}H &= H-(K^2-1). \end{aligned}$$

One of these equations follows from equating the diagonal ($i=a$) elements of B_i^a and $D_i\Phi^a$, the other from equating the off-diagonal ones.

The equations (3.47) can be solved explicitly [4] to give

$$\begin{aligned} H(r) &= \frac{ver}{\tanh(ver)}-1 \\ K(r) &= \frac{ver}{\sinh(ver)} \end{aligned}$$

It can be shown that this solution is unique up to spatial translation and gauge transformations and that it satisfies the equations of motion (3.3). The monopole associated with it is called the BPS-monopole after Bogomol'nyi, Prasad and Sommerfield. [5]

This monopole is part of a family of dyons with magnetic charge $g=\pm\frac{e}{4\pi}$ and arbitrary electric charge q. It follows from the previous discussion that the fields for these are given by

$$\Phi^a(\bar{x}) = \delta_i^a\frac{\tilde{x}^i}{e\tilde{r}^2}H(\tilde{r}) \tag{3.47}$$

$$A_i^a(\bar{x}) = \epsilon_{ij}^a\frac{\tilde{x}^j}{e\tilde{r}^2}K(\tilde{r}) \tag{3.48}$$

$$A_0^a(\bar{x}) = \delta_i^a\frac{\tilde{x}^i}{e\tilde{r}^2}H(\tilde{r})\sin(\theta) \tag{3.49}$$

[4] for a very explicit account of how to do this, see [29]

[5] Prasad and Sommerfield [30] were the first to obtain this solution. They actually did so by substituting the ansätze (3.46) and (3.45) into the original, second order equations of motion, requiring appropriate boundary conditions and (in their own words) "shimmying a bit"

where, as usual, $\cos(\theta) = g/\sqrt{q^2+g^2}$ and $\sin(\theta) = q/\sqrt{q^2+g^2}$ and where $\tilde{x}^i = x^i \cos\theta$. Note that replacing x^i by $\tilde{x}^i$ has the same effect as replacing v by $v\cos(\theta)$.

These dyons are called Julia-Zee dyons, because Julia and Zee [31] were the first to propose their existence. [6] The pure monopole is just the special case where $\theta = 0$.

One easily verifies that the Higgs field of the monopole solution does indeed have the asymptotic behaviour given in (3.44). Moreover, it is not difficult to see that it approaches zero at the space time origin. As an illustration, $|\Phi|^2$ is plotted in figure 3.1 below.

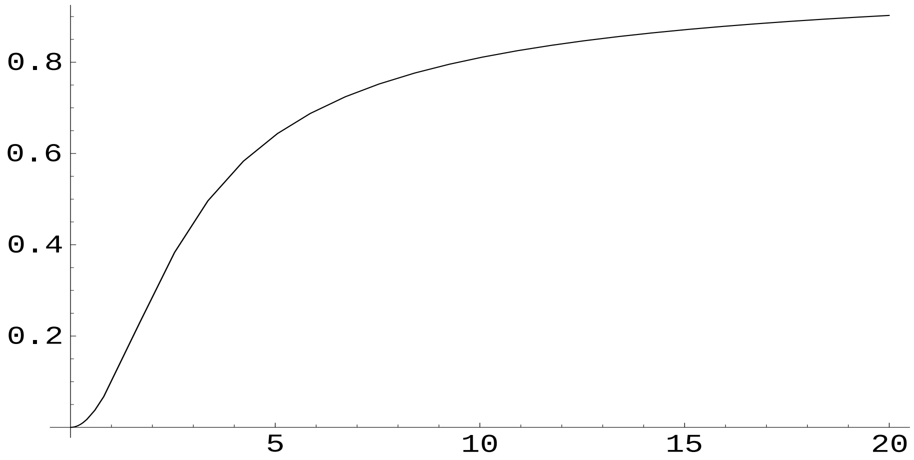


Figure 3.1: graph of $|\phi|^2$

As we can see in the figure, the space origin is actually the only point where the monopole's Higgs field is zero. But from the discussion in section 3.2, it then follows that all the monopole's magnetic charge is concentrated in this one point. Because the magnetic charge was defined in such a way that 't Hooft's abelian field tensor $f_{\mu\nu}$ satisfies the maxwell equations (with magnetic charge), it then follows that the Abelian magnetic field of the monopole is actually just that of a standard point monopole of charge $\frac{4\pi}{e}$ located at the origin.

The energy density of the monopole can most easily be derived using the formula (3.40). We will not give the resulting (long) expression here. In stead, we have made a graph of the energy density, which is shown in figure 3.2

This figure shows beautifully how the energy of the monopole is concentrated at the space origin.

We can see from (3.47) that the Higgs field Φ of an arbitrary dyon can be obtained from that of the monopole just by rescaling the space coordinates x^i with a factor of $\cos(\theta)$. It follows that all dyons will have a unique zero of the Higgs field at $\bar{x} = 0$ and thus that all dyons have the same magnetic charge distribution and B-field as the monopole. The energy density of the dyons is also simply related to that of the monopoles. We have

$$\mathcal{M}_{\text{dyon}}(r) = \cos^2(\theta)\mathcal{M}_{\text{monopole}}(r/\cos(\theta)).$$

We see that the peek of the energy density is flattened and broadened when we add electric charge. In other words: the higher a dyon's electric charge, the more its energy get spread out over space.

[6] The actual solutions were first found by Prasad and Sommerfield [30], using the same method as for the monopole

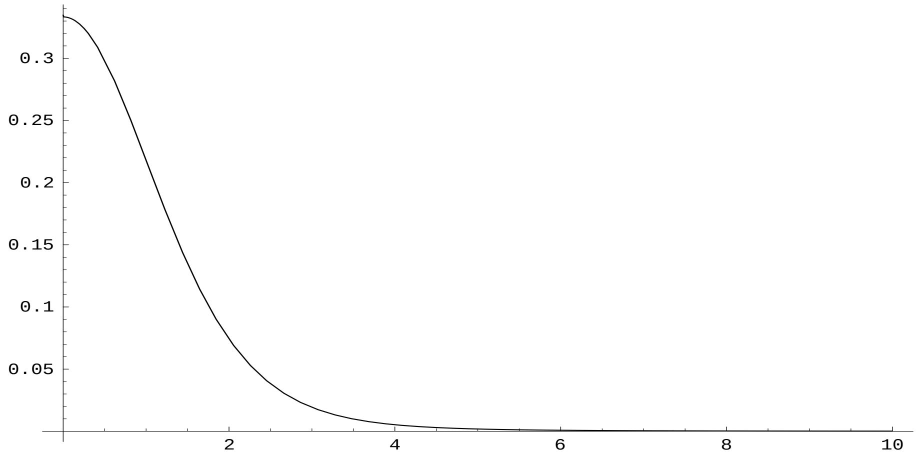


Figure 3.2: graph of the BPS energy density

3.5 Finding multimonopoles: Nahm's construction

In this section, I will say something about finding multimonopoles, that is, solutions to the equations of motion with monopole numbers higher than one. Again, we will restrict ourselves to the time independent case. Three different methods have been developed to attack this problem:

- There is a method based on techniques that stem from the study of lower dimensional solitons, due to Forgács, Horváth and Palla [32] [33]
- There is a method that uses twistor theory, due to Atiyah and Ward [34] and later extended by Corrigan and Goddard [35] and Hitchin [36]
- Finally, there is a method based on the ADHM-construction for instantons, due to Nahm [37] [38]. This is also called *Nahm's construction*, or also the *ADHMN-construction.*

Each of these methods has proved successful in providing explicit multimonopole solutions (see the above references and too many others to mention), but the most successful and currently most used method is Nahm's construction. One reason for this is, that the solutions yielded by this construction are automatically regular. In the other frameworks, regularity of the solutions is not a priori guaranteed. Another reason is, that Nahm's construction gives us a very good view of the parameters on which the solutions depend. We will therefore choose to describe Nahm's construction here and not the other two methods. Our treatment of the construction will follow that of Hitchin [39].

Nahm's construction for monopoles is inspired on the ADHM-construction for self-dual instantons which we described in chapter 2. To see the link between these instantons and the monopoles we are working on now, let us introduce a new notation. We define

$$A_4 := \Phi.$$

Because we are only interested in time independent field configurations here, we can ignore Minkowskian time and introduce a new, Euclidian time dimension, labelled with 4. In this way we go from a Minkowskian situation with a gauge potential and a Higgs field to a Euclidean situation with only a gauge potential $(A_1, \dots, A_4)$. Now we have introduced the extra coordinate x_4 to get some similarity with the Yang-Mills theory of chapter 2, but since we really want to describe monopoles, we don't want our fields to depend on x_4 in any essential way. Our first intuition might then be to require the Euclidean potential A_μ to be independent of the 4-direction. However, we do not have to be so restrictive. Because of the gauge symmetry of the Euclidean theory, we only have to require that if we shift all the fields in the x_4-direction, we end up with a field configuration that is gauge equivalent to the one we had before the shift. It is not difficult to see that the gauge transformations which are to compensate shifts in x_4 will form a subgroup S of the group of all gauge transformations. The map which sends a shift to the corresponding gauge transformation will then give a representation U of the group of shifts (which is just $\mathbb{R}$) and the image of this representation will of course just be S In the situation sketched above, all physical quantities will be invariant under a change of x_4. However, the fields themselves probably won't be x_4-independent. This can be remedied by setting x_4 equal to an arbitrary constant after all calculations in the Euclidean model have been performed (this is just a partial choice of gauge).

In the new setting we have introduced, the Bogomol'nyi equation takes a particularly nice form. We have

$$B_i = D_i\Phi = D_i A_4 = E_i, \tag{3.50}$$

where E denotes the electric field corresponding to the Euclidean gauge potential. Now the equation $B = E$ is just another way of saying that the electromagnetic field is self dual and thus we have found that every time independent solution to the Bogomol'nyi equation corresponds to a self dual gauge potential on Euclidean $\mathbb{R}^4$ which does not depend on Euclidean time. This means that we have to solve the same equations as in chapter 2, but with very different boundary conditions.

Let us recall how we went to work to solve the self duality equations in chapter 2. We introduced a matrix v, which depended on the space time coordinates in the following simple way

$$v(x) = Cx + D \tag{3.51}$$

where C and D were constant quaternionic $k+1 \times k$ matrices (for gauge group $SU(2)$). [7] We then demanded that v^*v be a real matrix and that v have maximal rank. Once we had found such a v, we could express the gauge potential and field in terms of a normalised element of the kernel of v^*, that is: a $k+1$ vector u such that

$$v^*u = 0, \quad u^*u = 1.$$

The potential would then be given by

[7] in chapter 2 we worked with homogeneous coordinates on S^4, but this does not make sense for the present boundary conditions and therefore we will work with the restriction of v to $\mathbb{R}^4$

$$A_\mu = u^* \partial_\mu u. \tag{3.52}$$

This procedure would yield a potential with instanton number k. Now notice that the only way in which the construction depends on the boundary conditions of chapter 2, is by the size of the matrix v. For instanton number k, v has to be a $k+1 \times k$ matrix. Here we are hoping to construct self dual field configurations which are independent of time. These will obviously make the action for Euclidean Yang-mills theory diverge. Intuitively, we could say this means that they have infinite instanton number. Therefore, if we want to still use the ADHM-construction, we have to use an "infinitely large matrix", or more correctly phrased, an operator between infinite dimensional spaces V and W. We will call this operator Δ. We want to construct Δ so that it satisfies the following conditions:

1. Δ is quaternionic linear
2. Δ is of the form $Cx + D$, where C and D do not depend on the space coordinate x.
3. $\Delta^*\Delta$ is real
4. $\Delta^*\Delta$ is invertible
5. The kernel of Δ^* has quaternionic dimension 1.
6. $\Delta(x_4 + \alpha) = U(\alpha)^{-1}\Delta(x)U(\alpha)$, where $\alpha \mapsto U(\alpha)$ is a representation of $\mathbb{R}$ in the group of quaternionic unitary transformations of V.

The first 5 conditions are there to insure that our construction will yield a self dual $SU(2)$ gauge field. Conditions 3 through 5 are just the infinite dimensional analogues of the conditions we had posed on v. Note that, in the finite dimensional case, conditions 4 and 5 were both incorporated in the condition that v should have maximal rank. This is why we did not have to impose them separately before. In the infinite dimensional case, condition 3 insures that the gauge field constructed from Δ will be regular (through the analogue of formula (2.15)), while condition 5 insures that the gauge field we get is actually an $sp(1) = su(2)$ field, and not, for example, an $sp(2)$ or $sp(3)$ field.

Condition 6 makes sure that the gauge field produced by the construction will in fact be independent of Euclidean time. We can see this as follows. From condition 5, we see that the equation $\Delta^* f = 0$ has only a single normalised solution (modulo multiplication with a unit quaternion). If we denote this solution with u, then it follows from condition 6 that the solution to $\Delta^*(x_4 + \alpha) f = 0$ will be given by $U^*(\alpha)u$. So we see that shifting the fields along the x_4 direction corresponds to a change of gauge, as we had required it should.

Generally, the gauge field that corresponds to a solution $u(x)$ to the equation $\Delta^* u = 0$ depends only on the quaternionic line that $u(x)$ spans in the vector space v. This is easy to see if we recall from chapter 2 that u is really an orthogonal gauge for a quaternionic line bundle over $\mathbb{R}^4$ which is embedded into $\mathbb{R}^4 \times v$. This bundle depends on the quaternionic lines $L(x)$ in V which we use to represent the fibres, not on the choice of the elements $u(x)$ which we use to represent these lines. A gauge transformation is just a different choice of normalised base vectors $u(x)$ for the lines $L(x)$. But all possible representative vectors can be obtained from each other by multiplication with a unit quaternion. Therefore, performing a gauge transformation on the field configuration that corresponds to u is the same as sending $u(x)$ to $U(x)u(x)$, where $U(x)$ is a quaternionic unitary transformation of V.

Let us now make some definitions. Let H^0 be the space of complex-valued square integrable functions on the interval $[-a, a]$, where we leave a to be an arbitrary positive real number for the moment. On H^0, we can define a real structure [8] σ by

$$\sigma(f)(z) := \bar{f}(-z)$$

We can now define the vector space V by

$$V = H^0 \otimes \mathbb{C}^k \otimes \mathbb{C}^2.$$

We identify the factor $\mathbb{C}^2$ in this product with the quaternions and take $\mathbb{C}^k$ to have some arbitrary real structure σ'. Denoting the quaternionic structure on $\mathbb{C}^2$ by j, we have natural quaternionic structures J and $\mathcal{J}$ on $\mathbb{C}^k$ and V, given by

$$\begin{aligned} J(v_1 \otimes v_2) &= \sigma'(v_1) \otimes j(v_2) \\ \mathcal{J}(f \otimes v_1 \otimes v_2) &= \sigma(f) \otimes \sigma'(v_1) \otimes j(v_2) \end{aligned}$$

We will choose to let these structures denote multiplication by j *on the right.* Now let us define the vector space W. We take

$$W = \left\{ f \in H^1 \otimes \mathbb{C}^k \,|\, f(-a) = f(a) = 0 \right\}.$$

where we take H^1 to be the Sobolev space of functions whose derivatives are in H^0. We give H^1 the same real structure as H^0. $\mathbb{C}^k$ also gets the same real structure as before.

At this moment, we note that an element f of W can be embedded into $W \otimes \mathbb{C}^2$ by the identification $f \equiv f \otimes (1, 0)$. Using this embedding, it makes sense to think of the quaternionic structure $\mathcal{J}$ of V as working on W as well.

For the map $\Delta : W \to V$, we now take a first order differential operator of the form

$$\Delta(x)f = (x_4 + x_1 e_1 + x_2 e_2 + x_3 e_3)\, f + i\partial_z f + i\, (T_1(z)e_1 + T_2(z)e_2 + T_3(z)e_3)\, f, \quad (3.53)$$

where e_1, e_2 and e_3 denote the action of left multiplication by the quaternions i, j, k (see section A.5 in the appendix for details), and where the T_i are complex $k \times k$ matrices which depend analytically on $z \in (-a, a)$. The T_i are called the *Nahm data.*

We see immediately that this Δ satisfies condition 2, with $C = I$. Condition 1 is equivalent with the requirement that Δ commute with the quaternionic structure $\mathcal{J}$. For the first term in Δ, this is trivial: left multiplication by a quaternion commutes with right multiplication. The second term is quaternionic linear because of our choice of real

[8] For details on real and quaternionic structures, see A.5

structure on V and W. We can see this by writing the quaternionic structure $\mathcal{J}$ on V (or $W \otimes \mathbb{C}^2$) in terms of the quaternionic structure J on $\mathbb{C}^k \otimes \mathbb{C}^2$. We have

$$\begin{aligned}
\mathcal{J}(f \otimes v_1 \otimes v_2)(z) &= \bar{f}(-z) \otimes \sigma'(v_1) \otimes j(v_2) \\
&= 1 \otimes \bar{f}(-z)\sigma'(v_1) \otimes j(v_2) \\
&= 1 \otimes \sigma'(f(-z)v_1) \otimes j(v_2) = 1 \otimes J(f(-z) \otimes v_2)
\end{aligned}$$

Now we had already identified the element $f \otimes v$ of W with the element $f \otimes v \otimes (0,1)$ of $W \otimes \mathbb{C}^2$. Now we can obviously also identify the element $1 \otimes v_1 \otimes v_2$ of $W \otimes \mathbb{C}^2$ with the element $v_1 \otimes v_2$ of $\mathbb{C}^k \times \mathbb{C}^2$ and using both identifications, we can rewrite the result we got above as

$$(\mathcal{J}f)(z) = J(f(-z)) \tag{3.54}$$

Filling this in in the second term of Δ, we get

$$\begin{aligned}
i\partial_z(\mathcal{J}f) &= i\partial_z(Jf(-z)) \\
&= -Ji\partial_z(f(-z)) \\
&= Ji(\partial_z f)(-z) = \mathcal{J}i\partial_z f
\end{aligned} \tag{3.55}$$

From which we see that this term is indeed quaternionic linear, thanks to our choice of real structures.

The last term in (3.53) gives

$$\begin{aligned}
i\sum_{n=1}^{3} T_n(z)e_n(\mathcal{J}f) &= i\sum_{n=1}^{3} T_n(z)e_n Jf(-z) \\
&= -Ji\sum_{n=1}^{3} \bar{T}_n(z)e_n f(-z) = \mathcal{J}i\sum_{n=1}^{3}(-\bar{T}_n)(z)e_n f(-z)
\end{aligned} \tag{3.56}$$

So if we require that

$$T_n(z) = -\bar{T}_n(-z), \tag{3.57}$$

then Δ is quaternionic linear.

Now let us consider the reality condition 3. We have

$$\begin{aligned}
\Delta^*\Delta &= \left(\bar{x} + i\partial_z + i\sum_{n=1}^{3} T_n^* e_n\right)\left(x + i\partial_z + i\sum_{n=1}^{3} T_n e_n\right) \\
&= -\partial_z^2 + \left(2ix_4 - \sum_{n=1}^{3}(T_n + T_n^*)e_n\right)\partial_z \\
&\quad + |x|^2 - \sum_{n=1}^{3}\partial_z T_n e_n + i\bar{x}\sum_{n=1}^{3} T_n e_n + i\sum_{n=1}^{3} T_n^* e_n x + \sum_{m,n=1}^{3} T_m^* T_n e_m e_n
\end{aligned} \tag{3.58}$$

From the first order term in ∂_z, we require $T_n + T_n^* = 0$ for reality and when we substitute this into the zero order term, we obtain the additional condition

$$\mathrm{Im}\left[\sum_{n=1}^{3} \partial_z T_n e_n\right] = \mathrm{Im}\left[\sum_{m,n=1}^{3} T_m T_n e_m e_n\right]. \tag{3.59}$$

where the imaginary part signifies that the only the coefficients of e_1, e_2 and e_3 in the left and right hand side have to be equal (The coefficient of 1 may differ). We can rewrite the above equation more explicitly as

$$\begin{aligned} \partial_z T_1 &= [T_2, T_3] \\ \partial_z T_2 &= [T_3, T_1] \\ \partial_z T_3 &= [T_1, T_2] \end{aligned} \tag{3.60}$$

$$\tag{3.61}$$

These equations are called *Nahm's equations.* Using these and $T_i = -T_i^*$, we may write

$$\Delta^* \Delta = -\partial_z^2 + 2ix_4 \partial_z + |x|^2 - \sum_{n=1}^{3} T_n^2 \tag{3.62}$$

It is now fairly simple to prove that $\Delta^* \Delta$ is invertible. First, note that, for $f \in W$, we have

$$\langle f, \partial_z f \rangle = 0,$$

where $< \cdot, \cdot >$ denotes the $\mathcal{L}^2$ inner product on the interval $[-a, a]$. This follows easily if one uses partial integration and the fact that $f(-a) = f(a)$ (which was conveniently built into the definition of W. Using this and (3.62), we can write

$$< \Delta^* \Delta f, f > = < \partial_z f, \partial_z f > + < (|x|^2 + \sum_{n=1}^{3} T_n^* T_n) f, f >, \tag{3.63}$$

using partial integration on the first term in (3.62). Now note that the right hand side of (3.63) is positive, unless f is identically zero. It follows that $\Delta^* \Delta$ is injective and hence invertible.

At this point, we have to check that the kernel of Δ^* has quaternionic dimension 1. Now if the T_n are analytic everywhere on the closed interval $[-a, a]$, then we would be unable to satisfy this condition for $k > 1$. In this situation, the T_i would automatically be Lipschitz-continuous on $[-a, a]$ and thus, the first order linear equation $\Delta^* v = 0$ would have $2k$ (complex) linearly independent regular (and thus normalisable) solutions on the interval $[-a, a]$. We see that, for $k > 1$, we will have to get rid of some of these solutions. At the same time, we will still want the T_i to be differentiable on the open interval $(-a, a)$, as this will insure differentiability of the solutions to $\Delta^* u = 0$. Therefore, we will require the T_i to have simple poles at $z = -a$ and $z = a$. Note that we have already accounted for this possibility by taking $f(-a) = f(a) = 0$ in the definition of W. Let us denote the residue matrices of the pole at $z = -a$ by t_i. Because we have $\bar{T}_i(-z) = T_i(z)$ and $T_i^*(z) = -T_i(z)$, the residues of the pole at $z = a$ will then be given by t_i^t. It is easy to see, using Nahm's equations (3.60), that the t_i will determine a k-dimensional representation of $SU(2)$.

We will now calculate the number of solutions to $\Delta^* u = 0$. We will do this by calculating the *index* of the operator Δ. This quantity is defined as follows

$$\mathrm{index}(\Delta) = \dim(\ker(\Delta)) - \dim(\ker(\Delta^*)). \tag{3.64}$$

Since we have already shown that $\Delta^*\Delta$ is invertible, we know that the equation $\Delta u = 0$ does not have any (regular) solutions and therefore, the index of Δ will give us (minus) the number of solutions to $\Delta^* u = 0$. The reason we do not calculate the number of solutions to $\Delta^* u = 0$ directly, but work with the index instead, is that the index of an operator is a kind of topological invariant on "operator space" [9] This means that in stead of calculating the index of Δ, we can calculate the index of a simpler operator $\tilde{\Delta}$, which can be "continuously deformed" into Δ and still get the right answer. Therefore, let us define

$$\begin{aligned} \tilde{\Delta} f &:= i\left\{\partial_z + \frac{\sum_{k=1}^3 t_i e_i}{a+z} + \frac{\sum_{k=1}^3 t_i^t e_i}{a-z}\right\} f \\ &:= i\left\{\partial_z + \frac{A}{a+z} - \frac{A^t}{a-z}\right\} f \end{aligned} \tag{3.65}$$

Where A is a linear map on $\mathbb{C}^k \times \mathbb{C}^2$. It is now natural to take the following ansatz for solutions to the equation $\tilde{\Delta} u = 0$:

$$u = v_{\alpha,\beta}(a+z)^\alpha (a-z)^\beta$$

where $v_{\alpha,\beta}$ is a constant vector in $\mathbb{C}^k \times \mathbb{C}^2$. We see that this will give a solution if α is an eigenvalue of A, β is an eigenvalue of A^t and $v_{\alpha,\beta}$ is an eigenvector of both A and A^t with the eigenvalues α and β. Thus we find $2k$ complex linearly independent solutions given by

$$u_i = v_i(a^2 - z^2)^{\alpha_i}$$

where the v_i are independent eigenvectors of A (and hence also of A^t, since A can be diagonalised). To see how many of these solutions are normalisable, we have to find the eigenvalues of A. Computing A^2, we find

[9]The relevant index theorem for this situation can be found in [40]

$$
\begin{aligned}
A^2 &= \left(\sum_{i=1}^{3} t_i \otimes e_i\right)^2 = \sum_{i,j=1}^{3} t_i t_j \otimes e_i e_j \\
&= \sum_{i,j=1}^{3} t_i t_j \otimes (\delta_{ij} + \epsilon^{ijk} e_k) = \sum_{i,j=1}^{3} t_i^2 \otimes I_{2\times 2} + \sum_{i,j=1}^{3} \epsilon^{ijk} t_i t_j \otimes e_k \\
&= \sum_{i,j=1}^{3} t_i^2 \otimes I_{2\times 2} + \sum_{k=1}^{3} -t_k \otimes e_k = \sum_{i,j=1}^{3} t_i^2 \otimes I_{2\times 2} - A
\end{aligned}
$$

so that we can conclude that $A^2 + A$ has the same eigenvalues as $\sum_{i,j=1}^{3} t_i^2$. But this is just the Casimir operator (or total spin) for the k-dimensional representation of $SU(2)$ determined by the t_i. If we now require that this representation is the unique irreducible k-dimensional representation, then it follows that $\sum_{i,j=1}^{3} t_i^2$ is equal to the constant $(\frac{k-1}{2})(\frac{k-1}{2}+1) = \frac{1}{4}(k^2-1)$. For any eigenvalue α of A, we will then have

$$
\alpha^2 + \alpha = \frac{1}{4}(k^2 - 1)
$$

from which it follows that the eigenvalues of A will be $\alpha_1 = -\frac{1}{2} + \frac{k}{2}$ and $\alpha_2 = -\frac{1}{2} - \frac{k}{2}$. The first value will give rise to a normalisable solution of $\tilde{\Delta} u = 0$, while the second will give rise to a non-normalisable one. The multiplicities of the eigenvalues can be determined using the fact that the trace of A is zero. We find that α_1 occurs $k-1$ times, while α_2 occurs $k+1$ times. Thus we have found that the equation $\tilde{\Delta} u = 0$ has $k-1$ normalisable solutions. In a completely analogous way, one can show that $\tilde{\Delta}^* u = 0$ has $k+1$ normalisable solutions. From this, we find that the index of $\tilde{\Delta}$, and hence of Δ, is -2. This proves that the kernel of Δ^* has complex dimension 2, and hence quaternionic dimension 1.

We are now left with our last condition, the one that has to insure the time independence of our monopole solutions. this condition, too, is satisfied. We have

$$
\Delta(x + x_4) f = e^{ix_4(z)} \Delta(x) e^{-ix_4(z)} f, \tag{3.66}
$$

which gives us the required unitary representation of R, and

$$
\mathcal{J}(e^{ix_4(z)} f(z)) = J(e^{ix_4(-z)} f(-z)) = e^{ix_4(z)} J(f(-z)) = e^{ix_4(z)} \mathcal{J} f, \tag{3.67}
$$

which shows that this representation is quaternionic.

We have now proved that Nahm's construction produces solutions to Bogomol'nyi's equations (if it produces anything at all). However, we have not said anything yet about the way in which these are produced from the zero modes of Δ. Of course, we will have to use some analogue of the formula (3.52) which we used to calculate instanton potentials. Since this formula uses the standard quaternionic scalar product, one would expect that, to have a valid analogue, we would first have to make an identification of the complex vector space V on which we are working with a certain quaternionic vector space and then define the monopole potential through the quaternionic scalar product on this space.

In the appendix, it is indicated how such an identification could proceed. However, in practice, one does not use identifications of the type discussed in the appendix to calculate monopole potentials in terms of quaternions. In stead, one uses the following formula to calculate them directly in terms of $SU(2)$ matrices; suppose $\{u_1, u_2\}$ is an orthonormal basis for the kernel of Δ^*, then we have

$$A_\mu^{ij} = \langle u_i^* \partial_\mu u_j \rangle \, . \tag{3.68}$$

It would be interesting to see if this equation can also be obtained through some identification of V with a quaternionic space, but we will not prove or disprove this and go on with the above formula.

Though we have now shown how to obtain solutions to Bogomol'nyi's equations using Nahm's construction, we have not yet said anything about the magnetic charge of these solutions. We will now show that this is equal to k (the size of the matrices T_i). We can deduce this by studying the asymptotic properties of the solutions of $\Delta^* u = 0$ as the space variable x goes to infinity. In this limit, the only appreciable contribution of the T_i to $\Delta^* u$ will come from the poles. Therefore, instead of Δ itself, we can study the operator $\hat{\Delta}$ defined by

$$\hat{\Delta} f := i \left\{ \partial_z - i(x_4 + x_1 e_1 + x_2 e_2 + x_3 e_3) + \frac{A}{a+z} - \frac{A^t}{a-z} \right\} f \tag{3.69}$$

where A is the same matrix as above. We have proved above that the equation $\Delta^* u = 0$ has exactly two linearly independent solutions. For large $\vec{x}$, we expect that this will also hold for the equation $\hat{\Delta}^* u = 0$. But we can actually find the essential z-dependence of these solutions fairly easily ! This works as follows: One can easily check that the operators $-i(x_4 + x_1 e_1 + x_2 e_2 + x_3 e_3)$ and A commute. Therefore, we can find simultaneous eigenvectors of these operators. Let us denote such a simultaneous eigenvector by v and let's call the corresponding eigenvalues of A and $-i(x_4 + x_1 e_1 + x_2 e_2 + x_3 e_3)$ α and λ respectively. We then see that the function u given by $u(z) = (z^2 - a^2)^\alpha e^{\lambda z}$ will be a solution to $\Delta^* u = 0$. Now we have already found the eigenvalues of A above. There are two: $\alpha_\pm = \frac{\pm k - 1}{2}$, but only α_+ gives rise to a normalisable solution, so we can forget about α_-. The two eigenvalues of $-i(x_4 + x_1 e_1 + x_2 e_2 + x_3 e_3)$ are easily computed. They are given by $\lambda_\pm = -ix_4 \pm r$. Thus we see that the two linearly independent solutions to $\hat{\Delta}^* u = 0$ are given by

$$u_+ = v_+ (z^2 - a^2)^{\frac{k-1}{2}} e^{r - ix_4} \tag{3.70}$$

$$u_- = v_- (z^2 - a^2)^{\frac{k-1}{2}} e^{-r - ix_4} \tag{3.71}$$

where v_+ and v_- are the appropriate simultaneous eigenvectors of A and $-i(x_4 + x_1 e_1 + x_2 e_2 + x_3 e_3)$. To normalise these solutions, we have to calculate the integrals

$$I_\pm := \int_{-a}^{a} dz \left\{ u_\pm^* u_\pm \right\} = |v_\pm|^2 \int_{-a}^{a} dz \left\{ (z+a)^{k-1} (z-a)^{k-1} e^{\pm 2rz} \right\} \, .$$

These can be done by iterated partial integration. It is possible to calculate the exact answer, but we will only be interested in the first two orders in $\frac{1}{r}$. Now the boundary terms obtained after n partial integrations will be of order $\frac{1}{r^n}$, so we want to look at the first two non-vanishing boundary terms. It is easy to see that these are the boundary terms we get after k and $k+1$ partial integrations (fewer partial integrations would leave positive powers of both $z-a$ and $z+a$ in the boundary terms). Moreover, we are only interested in the $z=a$ term for I_+ and in the $z=-a$ term for I_-, as the other terms will go exponentially to zero as r goes to infinity. This leaves us with

$$I_\pm = \pm|v_\pm|^2 \frac{(k-1)!(2a)^{k-1}e^{2ra}}{2r^k}\left(1 - \frac{k(k-1)}{4ar} + O(\frac{1}{r^2})\right)$$

and imposing the requirement that $I_\pm = 1$, we find

$$|v_\pm|^2 = \pm\frac{2r^k}{(k-1)!(2a)^{k-1}e^{2ra}}\left(1 + \frac{k(k-1)}{4ar} + O(\frac{1}{r^2})\right). \tag{3.72}$$

According to the formula (3.68) above, the $++$ and $--$ components of the asymptotic Higgs field corresponding to u will now be given by

$$\Phi_{\pm\pm} = \int_{-a}^{a} dz\left\{u_\pm^*\partial_{x_4}u_\pm\right\} = i|v_\pm|^2\int_{-a}^{a} dz\left\{(z+a)^{k-1}(z-a)^{k-1}ze^{\pm 2rz}\right\} \tag{3.73}$$

(the off-diagonal elements Φ_{+-} and Φ_{-+} are easily seen to be zero). The calculation of the integrals is similar to that of the integrals $I_\pm$ above and we find

$$\int_{-a}^{a} dz\left\{(z+a)^{k-1}(z-a)^{k-1}ze^{\pm 2rz}\right\} = \frac{a(k-1)!(2a)^{k-1}}{(2r)^k}\left(1 - \frac{1}{2ar} - \frac{k(k-1)}{4ar}\right)$$

Filling this and the result (3.72) in in (3.73), we find that we have

$$\Phi(x) \to i\begin{pmatrix} a(1-\frac{k}{2ar}) & 0 \\ 0 & -a(1-\frac{k}{2ar}) \end{pmatrix} \tag{3.74}$$

If we now set $a = v$, then we see that

$$|\Phi(r)| \sim v\left(1 - \frac{N}{vr}\right) \qquad (r\to\infty) \tag{3.75}$$

but this is just the formula (3.44) for the asymptotic behaviour of a k-monopole (with the fundamental electric charge e set to 1). Thus we can conclude that the monopole number of a solution produced by Nahm's construction is indeed equal to the dimension k of the matrices T_i. As a bonus, we can read off the eigenvalues of the asymptotic Higgs field from (3.74).

We have now come to a point in our discussion of Nahm's construction at which it would be the natural to try and construct some actual, concrete monopole solutions. For

$k = 1$, this is actually very simple. First, we have to solve Nahm's equations (3.60). Since the T_i are just complex numbers (1×1-matrices), the commutators in these equations vanish and we see that the T_i are independent of z. In other words, the T_i are just constant imaginary numbers. Looking at the equation $\Delta^* u = 0$, we see that these constants define a shift of the space coordinates x_i. Therefore we can limit ourselves to the case $T_i = 0$ and shift the solution we find this way in space to obtain the complete three parameter family of solutions. Solving the equation $\Delta^* u = 0$ for the case $T_i = 0$ is not difficult; we are dealing with a first order linear equation with constant coefficients. The solution is just the BPS-monopole we discussed before (possibly in a different gauge). A nice treatment that reobtains the equations (3.47) in an economical way can be found in [41].

For $k = 2$, Nahm's equations are a lot more difficult to solve than for $k = 1$, but it is still possible. The general solution can be given in terms of Jacobi elliptic functions (see [37]). However, solving the linear equation $\Delta^* u = 0$ with these Nahm data turns out to be very difficult and as far as I know, it has only been done in the special case where the resulting monopole configuration is axially symmetric. (see [42]). For higher monopole numbers, the situation is comparable. Solutions with special symmetries have been found and continue to be studied, both numerically and analytically (see e.g. [43, 44, 45] and also [1], from which I took the figure on the front page), but a general solution does not seem to be near.

Still, though explicit solutions are only available in special cases, it has been proved long ago [39] that *Nahm's construction does in fact yield all static solutions to Bogomol'nyi's equation (3.29).* As in the case of the ADHM-construction for instantons, a truly "logical"/intuitive proof requires advanced mathematical methods beyond the scope of this thesis. In [46], Corrigan and Goddard give a proof that doesn't preserve all the intuition, but uses only mathematics familiar to most physicists. Given a solution (A, Φ) to the Bogomol'nyi equations, they study a massless Dirac equation with this solution as a background field. From the solutions to this equation, they construct a set of matrices T_i which satisfy Nahm's equations and all the other requirements for Nahm's construction. Using these matrices as Nahm data, they then apply Nahm's construction and regain the solution (A, Φ) they started with, thus showing all static solutions to Bogomol'nyi's equation can be produced by Nahm's construction.

As an aside, I want to mention that Corrigan and Goddard's proof indicates a curious kind of "duality" (they call it reciprocity) between the fields (A, Φ) of a solution to the Bogomol'nyi equation and the Nahm data T_i. While Corrigan and Goddard show that the Nahm data can be extracted from solutions to the massless Dirac equations in the background of (A_i, Φ), our description of Nahm's construction shows that the fields (A_i, Φ) can be constructed from a massless Dirac equation in the background of the T_i (the equation $\Delta^* u = 0$). Moreover, while we have shown that the Bogomol'nyi equation is a self duality equation for a gauge field on Euclidean $\mathbb{R}^4$ which is independent of one space time direction, it can also be shown that the Nahm equations are the self duality equations for a gauge field on $\mathbb{R}^4$ which is independent of three of the space time directions. A nice generalisation of this interesting phenomenon of reciprocity can be found in [47]

Now let's go back to the main thread of our story. Though we don't want to get into the construction of monopoles with special symmetries, we would like to end this section with some application of Nahm's construction and preferably with a physically interesting one. Therefore, what we shall do is to sketch briefly how one can calculate the number of parameters involved in the general k-monopole solution, modulo gauge invariance. To do this calculation, we have to study perturbations of a solution Nahm's equations which leave us in the space of solutions. These perturbations will be solutions to the linearised version of Nahm's equations. We will only look at traceless perturbations, as the trace of a perturbation can always be removed by a spatial translation. Of course, this means we will have to add the three parameters for spatial translations by hand at the end of the computation.

To find the linearised equations, let us first write Nahm's equations in a more compact form:

$$\partial_z T_i = \frac{1}{2}\epsilon_{ijk}[T_j, T_k] \tag{3.76}$$

Now let $\hat{T}_i(z)$ be a solution to this equation. Substituting $T_i = \hat{T}_i + \delta_i$, we see easily that if T_i is to be a solution as well, we will have to have

$$\partial_z \delta_i = \epsilon_{ijk}[\hat{T}_j, \delta_k] \tag{3.77}$$

These are the linearised equations we were looking for. The solutions are just the zero modes of the operator P defined by

$$P\delta_i = \partial_z \delta_i - \epsilon_{ijk}[\hat{T}_j, \delta_k] \tag{3.78}$$

Thus, we are looking for the number of zero modes of P. However, since we are only interested in physically relevant parameters, we want to exclude the zero modes which are due to gauge transformations from the counting. It can be shown that if one works modulo these gauge modes, then PP^* is an invertible operator, so that the number of zero modes of the operator P is equal to its index. This means that, as in the calculation of the index of the operator Δ^*, we can simplify the operator P, calculate the index of the simplified operator and draw conclusions about the zero modes of P. Near the pole at $z = -a$, we can take a simplified operator $\tilde{P}$ defined by

$$\tilde{P}f := \left\{\partial_z + \frac{p}{a+z}\right\} f \tag{3.79}$$

where the linear operator p is given by

$$(p\delta)_i = \epsilon_{ijk}\left[\hat{T}_j, \delta_k\right]$$

The next step in the calculation is now to find the eigenvalues of p with their multiplicities. These are found in a way which is very similar to the way we found the eigenvalues of the matrix A in the calculation of the index of Δ^* above. Again, one shows that p is traceless, while $p^2 + p$ is equal to the Casimir operator for a certain representation of $SU(2)$. The representation in question is given in its infinitesimal form by the adjoint action of the residue matrices t_i on the complexification of the Lie algebra $su(k)$. One can show (for example by calculating the character) that this representation splits into $k-1$ irreducible representations of complex dimensions $d_1 = 3, d_2 = 5, d_3 = 7, \ldots, d_k = 2k+1$. The ith of these irreducible representations then gives us $d_i - 2$ (real independent) solutions that are non-normalisable and $d_i + 2$ solutions that are normalisable at least in the neighbourhood of $z = -a$. Thus, the pole at $z = -a$ makes $\sum_i \{d_i - 2\}$ zero modes of $\tilde{P}$ non normalisable. At $z = a$, we have the same situation and another $\sum_i \{d_i - 2\}$ zero modes of $\tilde{P}$ become non normalisable. This means we can now count the number of real parameters we are left with. We started with $3(k^2 - 1)$ real parameters for the three $su(k)$ matrices δ_i. Of these parameters, $k^2 - 1$ disappear because we work modulo gauge transformations. This leaves us with $2(k^2 - 1)$ parameters. Subtracting the number of parameters for non normalisable solutions and adding tree for the spatial translations, we obtain

$$2(k^2 - 1) - 2\sum_{i-1}^{k-1}(d_i - 2) + 3 = 4k - 1$$

real parameters.

Chapter 4

Dynamics of BPS-monopoles

4.1 Monopole dynamics; a geodesic approximation

In the previous chapter, we have looked into the problem of finding classical field configurations for time independent monopoles of minimal energy. We found that for monopole charge k, there is a $4k-1$ parameter family of such solutions, modulo gauge transformations. In this chapter, we want to study the low energy dynamics of these monopoles. To this end, we would like to use a geodesic approximation like the one we treated in section 1.6. That is, we want to describe monopole motion as a geodesic motion on the manifold of static minimal energy monopoles, where the metric on this manifold is the metric induced from the $\mathcal{L}^2$-metric on the space of field configurations. This means that the metric at any point z of the manifold of minimal energy configurations is given by the $\mathcal{L}^2$ inner product of the zero modes of the monopole configuration that corresponds to z.

In principle, there are two things to do before application of this geodesic approximation is possible. First, we have to find coordinates that describe the manifold of static monopoles and then we have to calculate the metric from the zero modes that correspond to these coordinates. Now there is a problem with this approximation here, which did not occur in the discussion in chapter 1: Since gauge transformations leave the energy invariant, the manifold of minimal energy monopoles will be infinite dimensional, as any monopole will have infinitely many zero modes which correspond to infinitesimal gauge transformations. We do not want to incorporate all these gauge modes into our geodesic approximation, because most of them do not correspond to physical properties of the monopole system. Rather, they correspond to a redundancy in our description of the system. Therefore, we want to somehow get rid of most of these gauge modes, so that we will (hopefully) be left with a finite number of physically interesting zero modes to work with.

Note that we do not want to get rid of all gauge modes, as some (or actually:one) will turn out to have physical significance. Therefore, what we should really do at this point, is determine which class of gauge transformations describes real physical symmetries of the monopole system. The rest of the gauge transformations would then be redundant and we could look at the configuration space modulo these transformations.

It is argued in [48] (and references therein), that the redundant gauge transformations are the ones that approach unity at infinity and that are connected to the identity transformation. We will call these transformations *small gauge transformations.* and the remaining ones *large gauge transformations.* One can identify the small gauge transformations by requiring that Gauss' equation (3.5) is satisfied at all times. As we have argued at the start of chapter 3, this equation should be satisfied for all physical configurations. We will not repeat the argument given in [48] in detail here, but we will give a partial proof (which requires less work and fits more easily into our treatment) in section 4.2. In particular, we will make the extra assumption that we can get of all time dependent gauge modes through small gauge transformations. Time dependent gauge modes do not fit comfortably into the zero mode methods we have described in chapter 1, because there, we always assumed zero modes to be time independent. The easiest way to get rid of them is to require that A_0 is equal to zero at all times. This can always be assured by a suitable gauge transformation, but here, we assume that it is actually possible to do this through a small gauge transformation. Note that choosing A_0 equal to zero eliminates only the time dependent zero modes. All time independent gauge transformations keep the condition $A_0 = 0$ invariant and are thus still allowed. In section 4.2, we will show that requiring the fields to satisfy the $A_0 = 0$ version of Gauss' law effectively divides out the time independent small gauge transformations. That is, we will see that as time goes by, fields which satisfy Gauss' law only vary in directions orthogonal to the flow of small gauge transformations. For the geodesic approximation, this means that we restrict the geodesic motion on the manifold of static monopoles to a sub-manifold orthogonal to this flow. The whole analysis then proceeds in terms of coordinates for this sub-manifold and zero modes which lie along it, or in other words, which are orthogonal to small gauge modes. In this way, we are effectively working in the space of static monopole configurations of minimal energy modulo small gauge transformations. because this is rather a mouthful, we will call it the *moduli space* for short. There is a moduli space for each monopole number and we will call the n-monopole moduli space $\mathcal{M}_n$

There is one subtlety about dividing out small gauge transformations which I should mention here. In the literature on monopoles, one usually does not just divide out the group of small gauge transformations, but in stead the bigger group of transformations which do tend to unity at infinity, but are not necessarily connected to the identity transformation. For us, this difference is probably not important. In our semi-classical approximation, we will be unable to see gauge transformations which are not homotopic to the unity transformation, as these transformations will take us out of the neighbourhood of the classical solution we are using as a starting point for the approximation. In fact, we will only deal with gauge transformations which differ by an infinitesimal amount from unity. Therefore, I expect that using only semi-classical methods, we will not be able to distinguish between dividing out only the small gauge transformations and dividing out all gauge transformation which tend to unity at spatial infinity.

Though we have now indicated how a geodesic approximation for monopole dynamics can be set up, we have not given any arguments as to why geodesic motion along the moduli space should be a good approximation to the exact monopole dynamics at low energy. If there would be a gap in the spectrum of the BPS-monopoles, then we would be able to use estimates like the ones we made in section 1.6 to make this plausible, but already in the first article ([12],1982) on this subject, Manton remarks that this is not the case and that there are modes of arbitrary low non-zero energy, corresponding to low frequency electromagnetic waves. These should be radiated by the monopoles when they they are in accelerated motion. However, Manton also says that the electromagnetic radiation emitted by the monopoles should approach zero as the monopole speeds go to zero. This gives hope that, even though there is no gap, not too much of the kinetic energy of the monopoles will be transferred to the non zero modes and the speed of motion along the moduli space will stay approximately constant. In a later article [49], Manton actually calculates the amount of energy that the monopoles emit as radiation for two relatively simple scattering processes and shows that it does indeed become negligible compared to the total kinetic energy as the monopole speeds go to zero. However, it was not until 1994, that Stuart finally managed to rigorously prove the validity of the geodesic approximation for monopole dynamics [50]. The idea of the proof is similar to that of our discussion at the end of section 1.6, but the estimates involved are so complicated that including the proof in this thesis would probably double it's length. Stuart's main result can be (sloppily) stated as follows: If we start with a field configuration (A, Φ) which at time $t = 0$ lies in configuration space at a distance of order ϵ^2 from the space of minimal energy monopoles and which has an initial speed along this space of order ϵ and in all other directions of order ϵ^2, then, for ϵ small enough, there will be a time interval of length $O(\frac{1}{\epsilon})$ during which (A, Φ) will stay within a distance of order ϵ from a geodesic on the moduli space.

On the one hand, we see from this that the geodesic approximation for monopoles becomes asymptotically exact if we let the monopole start on the moduli space and let it's initial speed go to zero. On the other hand, we see that when we want to do classical scattering calculations in practice, i.e. with some fixed initial speed, we can expect the approximation to break down after some finite time. In fact, we can always expect to describe a stretch of the monopole's trajectory with length of order one, but not necessarily the whole trajectory.

4.2 General information on monopole moduli spaces

In this paragraph, we will make much of the material we dealt with in the previous paragraph much more explicit. At the same time, we will obtain some general information about monopole moduli spaces. We will work in $A_0 = 0$ gauge, thereby eliminating all time dependent gauge transformations and leaving all time independent ones.

The easiest way to approach monopole moduli spaces is to look at the zero modes of monopole configurations. To find these, we would normally linearise the equations of motion around the monopole in question and then solve the linearised equations to give the zero modes. In the case of the BPS monopole, there is an easier method: we can opt to solve the linearised Bogomol'nyi equations in stead of the linearised equations of motion. The reason for this is, that if we deform a monopole in such a way that we keep the energy fixed, then the deformed monopole will still saturate Bogomol'nyi's bound and will therefore still be a solution to Bogomol'nyi's equations. Suppose now that $(\tilde{A}, \tilde{\Phi})$ is a monopole solution and look at the deformation of this solution given by $(\tilde{A} + a, \tilde{\Phi} + \phi)$. If we now require that, to first order in a and ϕ, this is still a solution to Bogomol'nyi's equation $B_i = D_i\Phi$, then we find that (a, ϕ) will have to satisfy the following linear equation.

$$-\frac{1}{2}\epsilon^{ijk}(\tilde{D}_j a_k - \tilde{D}_k a_j) = \tilde{D}_i\phi + e[a_i, \tilde{\Phi}] \tag{4.1}$$

Next to this equation, we have an extra condition on zero modes, which follows from Gauss' law. With $A_0 = 0$, Gauss' law is just a constraint on the time derivatives of the fields (A_i, Φ). This constraint easily follows from (3.5). It is

$$D_i \dot{A}_i + e[\Phi, \dot{\Phi}] = 0. \tag{4.2}$$

Now we want to view the dynamics of monopoles as motion along the manifold of static minimal energy monopoles. The time derivatives of the fields in such a motion will obviously have to lie along tangent vectors to this manifold. These tangent vectors are just the monopole zero modes, i.e. the solutions to the equation (4.1). However, not all of these are allowed if we have to satisfy Gauss' law at all times. It is easy to see from (4.2) that the only zero modes which are physically relevant are the ones which satisfy

$$\tilde{D}_i a_i + e[\tilde{\Phi}, \phi] = 0. \tag{4.3}$$

Such zero modes are said to be in *background gauge.* Now let us show that this condition on the zero modes effectively divides out the small gauge transformations. From (4.3), it trivially follows that we have

$$\int d^3x \left\{ (\tilde{D}_i a_i)^a \Lambda^a + (\frac{1}{2} e \epsilon^{abc} \tilde{\Phi}^b \phi_c) \Lambda^a \right\} = 0 \tag{4.4}$$

for all Lie algebra valued functions Λ. In fact, (4.3) and (4.4) are equivalent, even if we restrict ourselves to compactly supported Λ. If we assume Λ to be compactly supported, we can use partial integration to rewrite (4.4) as

$$\begin{aligned} \int d^3x \left\{ a_i^a \tilde{D}_i \Lambda^a + \phi^a (\frac{1}{2} e \epsilon^{abc} \tilde{\Phi}^b \Lambda^c) \right\} &=: \\ \Big\langle (a_i, \phi); (\tilde{D}_i \Lambda, [e\tilde{\Phi}, \Lambda]) \Big\rangle &= 0 \end{aligned} \tag{4.5}$$

where we have defined a natural inner product $< \cdot ; \cdot >$ on Lie algebra valued vector functions on $\mathbb{R}^3$ through a combination of contracting internal and spatial indices and taking the $\mathcal{L}^2$ inner product on $\mathbb{R}^3$.

Now note that, for infinitesimally small Λ, the expression $(\tilde{D}_i \Lambda, [e\tilde{\Phi}, \Lambda])$ is just the variation of the fields that would be caused by the gauge transformation $g(x) = e^{e\Lambda(x)}$. Note that this is a small gauge transformation, because it is clearly connected to the unity transformation and we had required Λ to approach zero at spatial infinity, a requirement which was needed to enable us to do the partial integration that brought us from (4.4) to (4.5). Thus we see that the requirement (4.3) on the zero modes is equivalent to the requirement that zero modes be orthogonal to small gauge transformations, with respect to the inner product $< \cdot ; \cdot >$. In other words, by imposing Gauss' law on the fields, we are effectively dividing out small gauge transformations.

The inner product $< \cdot ; \cdot >$ on zero modes is also the inner product which gives us the metric which determines the dynamics on the moduli space. We will now show that this follows from the physics of the theory, combined with the assumption that the motion of the fields through configuration space is such that the fields stay on the manifold of BPS-solutions at all times. To this end, let us assume that we have a set of coordinates

z_k that describe the moduli space. Now if we have a field configuration (A, Φ) that moves along a path on the manifold of BPS-solutions and satisfies Gauss' law (4.2) at all times, then we can identify the path along which the configuration moves with a path on the moduli space. This means we can express the time dependence of the configuration in terms of the coordinates z_i. That is, we have $(A(t), \Phi(t)) = (A(z(t)), \Phi(z(t)))$, where $z(t)$ describes some path through the moduli space. The time derivatives of the fields can now be written as

$$(\dot{A}_i(z), \dot{\Phi}(z)) = \dot{z}_k(\delta_k A_i(z), \delta_k \Phi(z)) \tag{4.6}$$

where $(\delta_k A_i(z), \delta_k \Phi(z))$ is the zero mode of $(A_i(z), \Phi(z))$ that corresponds to an infinitesimal change of z_k. Note that $(\delta_k A_i(z), \delta_k \Phi(z))$ does not have to be equal to $(\partial_{z_k} A_i, \partial_{z_k} \Phi)$. This last expression may have a nonzero inner product with some small gauge transformations. $(\delta_k A_i(z), \delta_k \Phi(z))$ is just the orthogonal projection of $(\partial_{z_k} A_i, \partial_{z_k} \Phi)$ on the vector space of zero modes which are orthogonal to small gauge transformations. This means that in general, we can write

$$(\delta_k A_i, \delta_k \Phi) = (\partial_{z_k} A_i, \partial_{z_k} \Phi) + (D_i \epsilon, [\Phi, \epsilon]) \tag{4.7}$$

where the gauge parameter $\epsilon(x)$ is some Lie algebra valued function of x that goes to zero at spatial infinity. If $(\partial_{z_k} A_i, \partial_{z_k} \Phi)$ is known then ϵ can be uniquely determined by requiring that the right hand side of the above equation satisfies Gauss' law.

We are now in a position to derive an effective Lagrangian for the restricted monopole dynamics we are considering here. First, let us write down the full Lagrangian (cf. formula (3.1)) for the gauge $A_0 = 0$. This is given by

$$L = \frac{1}{2} \int d^3x \mathrm{Tr} \left\{ \dot{A}_i \dot{A}_i + \dot{\Phi}\dot{\Phi} + \frac{1}{2} F_{ij} F_{ij} + D_i \Phi D_i \Phi \right\}. \tag{4.8}$$

We see that, in this gauge, there is a clear distinction between the first two terms, which represent the kinetic energy of the configuration, and the last two terms, which represent (minus) the potential energy. Because we have restricted our configurations to lie in the manifold of configurations with minimal potential energy, the last two terms are actually constant in time. This means that they are not relevant to the dynamics of our restricted system and we can leave them out of our effective Lagrangian. The first two terms can be rewritten using (4.6) and this leads to the following effective Lagrangian:

$$\begin{aligned} L &= \frac{1}{2} \int d^3x \left\{ \dot{A}_i^a \dot{A}_i^a + \dot{\Phi}^a \dot{\Phi}^a \right\} \\ &= \frac{1}{2} \int d^3x \left\{ \dot{z}_k \delta_k A_i^a \, \dot{z}_l \delta_l A_i^a + \dot{z}_k \delta_k \Phi^a \, \dot{z}_l \delta_l \Phi^a \right\} \\ &= \frac{1}{2} \langle (\delta_k A, \delta_k \Phi); (\delta_l A, \delta_l \Phi) \rangle \, \dot{z}_k \dot{z}_l =: \frac{1}{2} g_{kl}(z) \dot{z}_k \dot{z}_l \end{aligned} \tag{4.9}$$

and we see that this is just the Lagrangian for geodesic motion on the moduli space endowed with the metric g defined through the inner product $< \cdot ; \cdot >$ of the zero modes, as we expected. The Hamiltonian that follows from this Lagrangian can be easily calculated. It is given by

$$H = \frac{1}{2} g^{-1}_{kl}(z) p_k p_l \tag{4.10}$$

where p_k is the momentum conjugate to z_k. Note the analogy between this formula and the formula (1.39) we found in section 1.5.3 for the lowest order effective Hamiltonian for a theory with only scalar fields and without gauge modes. Just as there, we see that, if the zero mode corresponding to a change of z_k is non-normalisable, then the coordinate z_k is not physically relevant. Here, we can see this most easily from the effective Lagrangian; if any non-normalisable mode is excited, then the Lagrangian, and hence the kinetic energy, diverges. Thus, for physical applications, we are only interested in normalisable zero modes.

Note that, though the above argument does give the metric on the moduli space, it certainly does not prove that the geodesic approximation for monopole dynamics works. In fact, by looking at configurations which were restricted to lie on the manifold of solutions with minimal potential energy, we artificially prevented that any energy would be transferred into the modes orthogonal to this manifold. Therefore, it is not surprising that we ended up with a Lagrangian which describes a geodesic motion. The whole point of the proof for the geodesic approximation (see [50]) is to show that, even when there are no artificial constraints on the monopoles' dynamics, the energy that flows into the orthogonal modes becomes negligible as the monopole speeds go to zero. Once one has given such a proof, one can say in retrospect that the monopole are effectively confined to the manifold of configurations with minimal potential energy and still use the above argument to find the effective action and the metric on the moduli space.

At this point we are really back at the end of the previous section, with this difference, that we now have some explicit formulae to work with. Let us now do something really new and calculate the dimension of the moduli space. We have already done most of the work for this at the end of section 3.5. There, we found that the most general monopole of charge k depends on $4k-1$ parameters, if we divide out all gauge transformations. From this, we can conclude that a k-monopole moduli space will have at least $4k-1$ zero modes and that any zero modes that we have not counted yet have to be gauge modes. Of course, we are looking modulo small gauge modes, so we are left with the task to find the number of large gauge modes that are orthogonal to all small gauge modes, or equivalently, all large gauge modes that satisfy the condition (4.3).

Now one checks easily that, if $(\tilde{A}, \tilde{\Phi})$ is a monopole solution, then all time independent gauge modes around it will be given by

$$(a, \phi) = (\tilde{D}_i \Lambda, e[\tilde{\Phi}, \Lambda])$$

for some Lie algebra valued function Λ of the space coordinates. Substituting this into the condition (4.3), we find the following condition on the function Λ :

$$\tilde{D}_i \tilde{D}_i \Lambda + e^2 \left[\tilde{\Phi}, [\tilde{\Phi}, \Lambda] \right] = 0 \tag{4.11}$$

Now there is a one dimensional space of solutions to this equation which can be easily identified. It is given by

$$\Lambda(x) = \frac{\chi\tilde{\Phi}(x)}{v} \tag{4.12}$$

where χ is a real parameter. The division by v is necessary to make Λ dimensionless and also makes sure that the length of $\Lambda(x)$ at infinity is just $|\chi(x)|$. Substituting the above into (4.11), we see that the second term on the left hand side is trivially equal to zero, while the first term becomes zero using $\tilde{D}_i\tilde{\Phi} = \tilde{B}_i$ and $\tilde{D}_i\tilde{B}_i = 0$. Of course, the question is now: are there any other physically relevant solutions ?

To answer this question, let us first note once more that any physically relevant zero mode has to be normalisable. For the gauge mode corresponding to Λ, this means that we have to have

$$\int d^3x \mathrm{Tr}\left\{\tilde{D}_i\Lambda\tilde{D}_i\Lambda + [\tilde{\Phi},\Lambda][\tilde{\Phi},\Lambda]\right\} < \infty$$

Since both terms in the integrand are non-negative, it follows that we have to have

$$\begin{aligned} \tilde{D}_i\Lambda &= o(\tfrac{1}{x^{3/2}}) \quad (x \to \infty) \\ \left[\tilde{\Phi},\Lambda\right] &= o(\tfrac{1}{x^{3/2}}) \quad (x \to \infty) \end{aligned} \tag{4.13}$$

It follows from the second of these that we have to have

$$\Lambda(x) = c(x)\tilde{\Phi}(x) + o(\frac{1}{x^{3/2}}).$$

Where $c(x)$ is an arbitrary real function of the spatial variables. Substituting this into the first of the equations (4.13), we find the following equation for $c(x)$

$$\partial_i(\ln(c(x))) = -\left(\tilde{D}_i\tilde{\Phi}\right)\tilde{\Phi}^{-1}. \tag{4.14}$$

Now $\tilde{D}_i\tilde{\Phi}$ has to be at least of order $\frac{1}{x\sqrt{x}}$ to make the energy of the configuration $(\tilde{A},\tilde{\Phi})$ converge and similarly, Φ^{-1} has to be of order 1, so we find

$$c(x) = o(e^{\frac{1}{\sqrt{x}}}) \qquad (x \to \infty)$$

and we see that, at infinity, $c(x)$ is equal to a constant plus terms that decrease to zero. Effectively, this means that any normalisable gauge mode will be equal to the mode described by (4.12), plus possibly some small gauge modes. Since we are working modulo small gauge modes, it follows that the mode given by (4.12) is the sole physically interesting gauge mode.

If we now add the large gauge mode we have found to the $4k-1$ zero modes we already had, we see that we have $4k$ independent zero modes around any k-monopole. We can thus conclude that the k-monopole moduli space is $4k$-dimensional. Weinberg [51] was the first to (informally) derive this dimensionality. A formal proof can be found in Taubes' work [52]. In this article, Taubes applies index theoretical methods directly to the equations (4.1) and (4.3). We have avoided this by going through Nahm's transformation.

Let us now give a physical interpretation to the gauge mode we have found. This is relatively easy, because we can readily identify the collective coordinate to which this mode corresponds. The gauge mode is just the generator of the one parameter group of gauge transformations given by $\chi \mapsto e^{e\frac{\chi}{v}\tilde{\Phi}}$ and we see that χ can be used as the collective coordinate corresponding to the gauge mode. Note that χ is a periodic coordinate with period $\frac{2\pi}{ev}$, because $\tilde{\Phi}$ has eigenvalues $\pm iv$ at infinity (away from infinity, we can neglect the effect of a change in χ, because we are working modulo small gauge transformations).

Though it is not clear if and how one should physically distinguish between configurations with different values of χ, it is easy to see what happens as χ becomes time dependent. If we gauge transform the fields $(\tilde{A}, \tilde{\Phi})$ by $e^{e\frac{\chi(t)}{v}\tilde{\Phi}}$ at time t, then we effectively turn on an E-field proportional to the B-field. More explicitly, in $A_0 = 0$ gauge, we have $E_i = \dot{A}_i$ and the deformation in $\tilde{\dot{A}}_i$ due to the gauge transformation $e^{e\frac{\chi(t)}{v}\tilde{\Phi}}$ is given by

$$\delta_\chi \dot{A}_i = \partial_t(\tilde{D}_i(\frac{\chi}{v}\tilde{\Phi})) = \frac{\dot{\chi}}{v}\tilde{D}_i\tilde{\Phi} = \frac{\dot{\chi}}{v}\tilde{B}_i \tag{4.15}$$

We can now find the electric charge of the deformed configuration by calculating the flux of its Abelian electric field $\mathbf{E}$ through the sphere at infinity. However, it is clear that this flux will be just $\frac{\dot{\chi}}{v}$ times the flux of the Abelian magnetic field $\mathbf{B}$. This last flux would give us the magnetic charge and so we find without any serious calculation, that the electric charge of the deformed configuration will be $\frac{\dot{\chi}}{v}g$, where g is the magnetic charge. Thus, we see that the time dependence of the collective coordinate χ determines the electric charge. Of course, as the electric field is turned on, the energy is also raised and we go from a situation with only a potential energy of vg to a situation with the same potential energy, but with an added kinetic energy of $\frac{g\dot{\chi}^2}{v}$.

Next to χ, there are three more collective coordinates which are are easy to identify and present in all the monopole moduli spaces. These are the coordinates corresponding to the position of the centre of mass of the monopole configuration. We will call these X_1, X_2 and X_3. We have already seen these arise explicitly in our description of Nahm's construction as the traces of the matrices T_i. Generally, one can show [53] that the n-monopole moduli space has the form $\mathcal{M}_n = \mathbb{R}^3 \times (S^1 \times \mathcal{M}_n^0)/\mathbb{Z}_n$. Here, the $\mathbb{R}^3$-factor corresponds to the centre of mass coordinates X_i and the S^1 is just the circle parametrised by χ. $\mathcal{M}_n^0$ is a simply connected $4(n-1)$-dimensional hyperkähler manifold which is sometimes called the moduli space of "strongly centred" monopoles, because it describes the dynamics of monopoles in the coordinate system for which the centre of mass of the monopole configuration is located at the space origin at all times. The n-fold cover of $\mathcal{M}_n$, $\mathbb{R}^3 \times S^1 \times \mathcal{M}_n^0$, is metrically a direct product. That is, the metric on $\mathbb{R}^3 \times S^1$ is independent of the coordinates on $\mathcal{M}_n^0$ and vice versa. This makes it pleasant to work with coordinates for this n-fold cover in stead of directly with coordinates for $\mathcal{M}_n$. When we do this, we have to implement the identification modulo $\mathbb{Z}_k$ by identifying certain sets of coordinates, which describe identical field configurations. We will see an explicit example of this in the two monopole case which is discussed in section 4.4.

I will not give a detailed proof that $\mathcal{M}_k^0$ is a Kähler manifold, but I will indicate the three anti-commuting complex structures that go with this manifold. These can be seen most easily in the notation of section 3.5, i.e. in a Euclidean setting with a "dummy coordinate" x_4 and with $A_4 := \Phi$. In this notation, the linearised Bogomol'nyi equation (4.1) and the gauge condition (4.3), which together determine the tangent vectors of $\mathcal{M}_n$, take the form

$$\begin{aligned} \tilde{D}_1 a_1 + \tilde{D}_2 a_2 + \tilde{D}_3 a_3 + \tilde{D}_4 a_4 &= 0 \\ -\tilde{D}_1 a_4 - \tilde{D}_2 a_3 + \tilde{D}_3 a_2 + \tilde{D}_4 a_1 &= 0 \\ \tilde{D}_1 a_3 - \tilde{D}_2 a_4 - \tilde{D}_3 a_1 + \tilde{D}_4 a_2 &= 0 \\ -\tilde{D}_1 a_2 + \tilde{D}_2 a_1 - \tilde{D}_3 a_4 + \tilde{D}_4 a_3 &= 0. \end{aligned} \tag{4.16}$$

The top equation in this set is the gauge condition, the rest are the three component equations of the linearised Bogomol'nyi equation. These four equations can be restated as one equation if we combine the four components of $a(x)$ into a single element of $su(2)\otimes\mathbb{H}$. Let us then write

$$(a_1, a_2, a_3, a_4) \equiv a_4 \otimes 1 + a_1 \otimes i + a_2 \otimes j + a_3 \otimes k.$$

Similarly, let us define an operator $\bar{D}$ on $su(2)\otimes\mathbb{H}$-valued fields by

$$\bar{D} := \tilde{D}_4 \otimes 1 - \tilde{D}_1 \otimes i - \tilde{D}_2 \otimes j - \tilde{D}_3 \otimes k$$

With this definition, we can write down the four equations (4.16) in the single quaternionic equation

$$\bar{D}a = 0. \tag{4.17}$$

Now note that the operator $\bar{D}$ commutes with quaternionic right multiplication. It follows that, for any quaternion q, we have $\bar{D}(aq) = (\bar{D}a)q = 0$. Since the solutions to $\bar{D}a = 0$ can be identified with tangent vectors to $\mathcal{M}_n$ at $(\tilde{A}, \tilde{\Phi})$, it follows that the tangent spaces of $\mathcal{M}_n$ are all closed under the action of quaternionic right multiplication. In particular, multiplication on the right by the quaternions i, j and k gives $\mathcal{M}_n$ three anti-commuting almost complex structures, which we will call e_1, e_2 and e_3. These structures are given explicitly by

$$\begin{aligned} e_1(a_1, a_2, a_3, \phi) &= (\phi, a_3, -a_2, -a_1) \qquad (4.18) \\ e_1(a_1, a_2, a_3, \phi) &= (-a_3, \phi, a_1, -a_2) \\ e_1(a_1, a_2, a_3, \phi) &= (a_2, -a_1, \phi, -a_3) \end{aligned} \tag{4.19}$$

We will not prove that these almost complex structures are actually complex structures, but we do want to note that they leave the inner product $< \cdot ; \cdot >$ and hence the infinitesimal metric g invariant.

To end this section, I want to mention that the moduli spaces $\mathcal{M}_n$ have an alternative description in terms of certain moduli spaces which can be defined on the space of Nahm data T_i. We have already seen that the Nahm data have the right number of parameters modulo gauge transformations and that their traces correspond to the collective coordinates X_i, but the correspondence goes much further. A similar construct to the one described above gives the moduli spaces of Nahm data three anti-commuting complex structures and one can in fact show that they are hyperkähler manifolds. Moreover, Nakajima [54] has proved that the moduli space of Nahm data for $SU(2)$ monopoles of charge n, with its natural $\mathcal{L}^2$ metric, is isometric to the n-monopole moduli space $\mathcal{M}_n$. This makes these moduli spaces the same for all practical purposes.

4.3 Metric and Dynamics on $\mathcal{M}_1$

In this section, we will calculate the metric on the one monopole moduli space $\mathcal{M}_1$ and discuss the resulting classical and quantum dynamics. This will be a fairly brief exercise, as most of the work was already done in the previous section.

From the information presented in section 4.2, we can easily draw the conclusion that, as a manifold $\mathcal{M}_1$ is just $\mathbb{R}^3 \times S^1$, where the factor $\mathbb{R}^3$ corresponds to the three parameters X_i for spatial translations and the factor S^1 to the parameter χ whose time dependence determines the electric charge. To calculate the metric, we need to know the explicit forms of the zero modes that correspond to χ and the X_i. For χ, we derived in section 4.2 that we have

$$\left(\delta_\chi \tilde{A}_i, \delta_\chi \tilde{\Phi}\right) = \left(\frac{1}{v}\tilde{D}_i\tilde{\Phi}, 0\right). \tag{4.20}$$

For the zero modes corresponding to spatial translations, we have to have

$$\left(\delta_{X_j}\tilde{A}_i, \delta_{X_j}\tilde{\Phi}_j\right) = \left(\partial_j\tilde{A}_i, \partial_j\tilde{\Phi}_j\right) + \left(\tilde{D}_i\Lambda_j, e[\Phi, \Lambda_j]\right) \tag{4.21}$$

for certain $su(2)$-valued functions Λ_j which approach zero at infinity. The first term on the right hand side of this equation is just the change in the fields (A, Φ) due to a translation in the j-direction. This is just the zero mode we would expect. The second term is an infinitesimal small gauge transformation, which is needed to bring this zero mode into background gauge.

We could find the functions Λ_j, and thereby the explicit form of the translational zero modes, by substituting the above equation into the gauge condition (4.3), but there is a more elegant way. We can obtain the translational zero modes from the gauge mode (4.20) by applying the three complex structures e_1, e_2 and e_3 given in (4.18) to it. This gives

$$\begin{aligned}
e_1\left(\delta_\chi\tilde{A}_i, \delta_\chi\tilde{\Phi}\right) &= \left(0, \tilde{D}_3\tilde{\Phi}, -\tilde{D}_2\tilde{\Phi}, -\tilde{D}_1\tilde{\Phi}\right) &&= -\partial_1\left(\tilde{A}_i, \tilde{\Phi}\right) + \left(\tilde{D}_iA_1, e[\tilde{\Phi}, \tilde{A}_1]\right)\\
e_2\left(\delta_\chi\tilde{A}_i, \delta_\chi\tilde{\Phi}\right) &= \left(-\tilde{D}_3\tilde{\Phi}, 0, \tilde{D}_1\tilde{\Phi}, -\tilde{D}_2\tilde{\Phi}\right) &&= -\partial_2\left(\tilde{A}_i, \tilde{\Phi}\right) + \left(\tilde{D}_i\tilde{A}_2, e[\tilde{\Phi}, \tilde{A}_2]\right)\\
e_3\left(\delta_\chi\tilde{A}_i, \delta_\chi\tilde{\Phi}\right) &= \left(\tilde{D}_2\tilde{\Phi}, -\tilde{D}_1\tilde{\Phi}, 0, -\tilde{D}_3\tilde{\Phi}\right) &&= -\partial_3\left(\tilde{A}_i, \tilde{\Phi}\right) + \left(\tilde{D}_i\tilde{A}_3, e[\tilde{\Phi}, \tilde{A}_3]\right),
\end{aligned} \tag{4.22}$$

where we have used Bogomol'nyi's equation $\tilde{B}_i = \tilde{D}_i\tilde{\Phi}$ in going from the second to the last members of these equations. Comparing the last members to equation (4.21), we can see that we have indeed obtained the translational zero modes (just take $\Lambda_i = -\tilde{A}_i$). The zero modes as we have given them in (4.22) are automatically in background gauge, because the complex structures e_i preserve this condition (it is just the first of the preserved equations (4.16)) and the gauge mode (4.20) was already in background gauge. Of course, one can also easily verify directly that the background gauge condition is satisfied.

Now that we have all the zero modes in background gauge, calculating the metric is easy. The off-diagonal elements all vanish, while the diagonal elements are independent of $\vec{X}$ and χ and are given by

$$g_{X_iX_i} = v^2 g_{\chi\chi} = \int d^3x \left\{ \tilde{D}_i\tilde{\Phi}\tilde{D}_i\tilde{\Phi} \right\} = M, \tag{4.23}$$

where M is the total energy of the configuration $(\tilde{A}, \tilde{\Phi})$, i.e the mass of one monopole. Thus, the low energy dynamics of one monopole are governed by the Hamiltonian

$$H = \frac{1}{2M}\left(P_1^2 + P_2^2 + P_3^2\right) + \frac{v^2}{2M}P_\chi^2, \tag{4.24}$$

where the P_i are the momenta conjugate to the X_i and P_χ is the momentum conjugate to χ. Needless to say, we have $P_i = \dot{X}_i$ and $P_\chi = \dot{\chi}$ and these quantities are constants of the motion. Thus, classically, the monopole moves along a straight line in $\mathbb{R}^3$ at a constant speed and it also has a constant speed along the S^1 factor of the moduli space, which means that its electric charge is conserved. The values of the momenta (P_i, P_χ) can be chosen arbitrarily and thus the monopole can move at arbitrary speeds and have an arbitrary electric charge.

When we look at the quantum mechanics of the Hamiltonian (4.24), we get a fairly similar picture. The momenta P_i and P_χ are still constants of the motion and each eigenstate of H is uniquely labelled by the quantum numbers p_i and p_χ which represent the eigenvalues of P_i and P_χ. Explicitly, we have

$$|\vec{p}, p_\chi >= e^{i(\vec{p}\cdot\vec{X} + p_\chi\chi)}.$$

Here, the p_i can still be chosen arbitrarily, but p_χ, and thus the electric charge, is now quantised in units of e, because $|\vec{p}, p_\chi >$ has to be single valued when we vary the $\frac{2\pi}{e}$-periodic coordinate χ.

Note that all we have done here for one monopole is also valid as a description of the dynamics of the centre of mass of a configuration of n monopoles. The only difference is that the mass M of a one monopole configuration will have to be replaced with the mass of the n monopole configuration in this case.

Note also that we can see from the results of this section that the geodesic approximation is really a low velocity approximation. Since we know the exact field configuration for a static monopole, we can produce the exact field configurations for a moving monopole by applying a Lorentz boost to these. Especially at high velocities, such a boost would cause a Lorentz contraction of the monopole, thus making the fields at any given instant quite different from the uncontracted fields of a static monopole. We also know the field configurations for a static Julia-Zee dyon. [1] We could produce these from monopoles by turning on an E-field, but we had to rescale the fields after doing this in order to still satisfy the boundary conditions. In the geodesic approximation, we can produce a static dyon by a uniform motion on the S^1 factor. Here we see once more that the approximation works better at lower speeds; The geodesic approximation "forgets" to rescale the fields, but this effect becomes negligible at low electric charges, i.e. at low speed on S^1.

4.4 Metric and Dynamics on $\mathcal{M}_2$

In this section, we will treat the low energy dynamics of a system of two monopoles. To do this, we need the metric on the moduli space $\mathcal{M}_2^0$ of two monopole configurations. If all two monopole configurations and their zero modes were known explicitly, then we would be able to find this metric by direct calculation, as we did for the one monopole case. However, this is not the case. [2] Of course, we can still find the metric on the $\mathbb{R}^3 \times S^1$-factor of the double cover of the moduli space in the same way as in section 4.3, but this still leaves us with the task to determine the metric on the moduli space $\mathcal{M}_2^0$ of strongly centred two monopole configurations. For this manifold we cannot calculate the metric directly. Still, Atiyah and Hitchin [53] managed to find the exact metric on $\mathcal{M}_2^0$ by exploiting the fact that this moduli space is hyperkäler and admits an isometric $SO(3)$ action.

This $SO(3)$ action on the moduli space is induced from the action of $SO(3)$ on the space of field configurations. Since this action just rotates field configurations in space, it leaves their $\mathcal{L}^2$ inner products fixed. Hence, it also fixes the metric on the moduli space, as this is induced from the $\mathcal{L}^2$ metric on the space of field configurations.

It can be shown that two monopole configurations are never spherically symmetric and axially symmetric if and only if the monopoles' centres (the zeros of the Higgs field) are located at the same point. Therefore, the orbits of the $SO(3)$ action on the moduli space $\mathcal{M}_n^0$ of strongly centred monopoles are three dimensional, except in the one special case where the monopoles are both located at the space origin. As a result, we can parametrise $\mathcal{M}_n^0$ in terms of a variable r which determines the separation between the monopoles, two angular variables $\theta \in [0, \pi]$ and $\phi \in [0, 2\pi]$, which determine the orientation of the axis that joins the monopoles and an angular variable $\psi \in [0, 2\pi]$, which parametrises a rotation about this axis.[3] Note that the radial variable r does not actually have to be equal to the separation of the monopoles, as long as it determines this separation uniquely. This means we can redefine r to suit our purposes later. The angles θ, ϕ and ψ are called the Euler angles. They give a parametrisation of $SO(3)$. For a good review of their precise

[1] See the equations (3.47). Note that for comparison with the geodesic approximation, we either have to transfer these into $A_0 = 0$ gauge or apply the inverse transformation to the fields obtained from the geodesic approximation.

[2] Note that the exact Nahm data for all two monopole configurations are known and that in principle, it should be possible to calculate the metric by exploiting the isometry between the moduli space of Nahm data and that of monopole configurations. However, this was not the way the metric was first found. In fact, the isometry between the moduli spaces was found long after the two monopole metric

[3] In principle, we could also need an additional discrete index to label distinct $SO(3)$ orbits of configurations which do have equal separation between the monopole centres. It turns out that this is not necessary, but we will not prove this.

definition, see e.g. [55].

Though the orbits of the $SO(3)$ action are almost all three dimensional, they are not isomorphic to $SO(3)$, because the field configurations in the moduli space are invariant under a discrete group of symmetries. These are given by

$$\begin{array}{ll} s_1: & \theta \mapsto \pi - \theta, \;\; \phi \mapsto \pi + \phi, \;\; \psi \mapsto -\psi, \;\; \chi \mapsto \chi \\ s_2: & \theta \mapsto \theta, \;\; \phi \mapsto \phi, \;\; \psi \mapsto \pi + \psi, \;\; \chi \mapsto \pi + \chi \\ s_3: & \theta \mapsto \pi - \theta, \;\; \phi \mapsto \pi + \phi, \;\; \psi \mapsto \pi - \psi, \;\; \chi \mapsto \pi + \chi \end{array} \tag{4.25}$$

To get a one to one parametrisation of the moduli space, we have to identify sets of coordinates that are related by these symmetries. Note that while s_2 and s_3 involve a change in the coordinate χ on the S^1 factor of the moduli space, s_1 changes only the spatial location of the monopoles. To be precise, s_1 corresponds to point reflection in the origin and hence it exchanges the monopoles. The fact that the fields are invariant under this transformation tells us that monopoles are indistinguishable particles. In the double cover $\mathbb{R}^3 \times S^1 \times \mathcal{M}_2^0$ of $\mathcal{M}_2$, the identification corresponding to s_1 has already been performed. The symmetry s_2 is the symmetry that corresponds to the $\mathbb{Z}_2$ in $\mathbb{R}^3 \times (S^1 \times \mathcal{M}_2^0)/\mathbb{Z}_2$. The fact that s_2 leaves the fields invariant means that rotation of a monopole configuration by an angle of π radians around the axis joining the monopoles can be undone by a shift of π radians on the phase factor of the moduli space. It is likely that monopole configurations that differ only by a constant phase shift on the S^1 factor of the moduli space cannot be physically distinguished. Therefore, the symmetry s_2 may physically appear as a discrete rotational symmetry of two monopole configurations. The symmetry s_3 does not have to be discussed separately, as it is the product of s_1 and s_2

We now want to write down an expression for the metric on $\mathcal{M}_2^0$ that makes the $SO(3)$ invariance as clear as possible. To this end, we introduce three one-forms σ_1, σ_2 and σ_3, given by

$$\begin{aligned} \sigma_1 &= -\sin\psi \, d\theta + \cos\psi \sin\theta \, d\phi \\ \sigma_2 &= \cos\psi \, d\theta + \sin\psi \sin\theta \, d\phi \\ \sigma_3 &= d\psi + \cos\theta \, d\phi \end{aligned} \tag{4.26}$$

These one forms have an interesting physical meaning in the theory of rotations of rigid bodies which will be of use to us. The position of a rigid body in space can be described by the position of one reference point in the body (usually the centre of mass) and Euler angles θ, ϕ and ψ defined as above. If one fixes the reference point in the origin, the motion of the body is described by a set of time dependent Euler angles, i.e. by a path through $SO(3)$. The components of the instantaneous angular velocity of the body in space will then be given by one forms on $SO(3)$. Naturally, these one forms can be changed by rotating the inertial frame in which one chooses to study the body. However, if one fixes an orthogonal coordinate frame in the body (so that it rotates with the body), then the components of the angular velocity in this frame will be independent of the angle from which one chooses to look at the body in the laboratory. Now the one forms σ_i above are just the one forms which give the components of the angular velocity of a rigid body in an orthogonal frame that moves with the body (see for example [55],p.176). Hence they are $SO(3)$ invariant.

If we now write the metric in terms of dr and the $SO(3)$ invariant one forms σ_i, then the condition that the metric is $SO(3)$ invariant becomes just the condition that the coefficients of the metric are independent of the Euler angles. Moreover, one can show [56] that the metric takes the form

$$ds^2 = f^2 dr^2 + a^2\sigma_1^2 + b^2\sigma_2^2 + c^2\sigma_3^2, \tag{4.27}$$

where the functions $a,\ b,\ c$ and f depend on r only.

Before we determine the functions a, b, c and f, let us first write the Lagrangian and the equations of motion in terms of these. The Lagrangian for two monopole motion is

$$L = \frac{M}{2}(\dot{X}_i\dot{X}_i + \dot{\chi}^2) + \frac{\mu}{2}(f^2\dot{r}^2 + a^2\omega_1^2 + b^2\omega_2^2 + c^2\omega_3^2), \tag{4.28}$$

where μ is the reduced mass of the monopole system [4] and where we have written ω_i for the component of angular velocity that corresponds to σ_i, i.e.

$$\omega_1 = -\sin\psi\,\dot{\theta} + \cos\psi\sin\theta\,\dot{\phi}, \quad etc.$$

We see that the analogy with rigid body motion persists. If r is kept constant, this would be exactly the Lagrangian for the free motion of a rigid body with principal moments of inertia equal to μa^2, μb^2 and μc^2. It follows that the equations of motion for $a,\ b$ and c are the same as those for the rigid body. These are most conveniently expressed in terms of the angular momentum $\mathbf{L}$. In an inertial frame these will of course be constant, but in the coordinate system that moves with the body, we have (cf. [55],p.204)

$$\partial_t\mathbf{L} = \mathbf{L}\times\omega \tag{4.29}$$

For the axes of the moving system we will of course take the principal axes of the body and hence the components of $\mathbf{L}$ are given by

$$L_1 = \mu a^2\omega_1, \quad L_2 = \mu b^2\omega_2, \quad L_3 = \mu c^2\omega_3$$

Filling this in in (4.29) gives us the equations of motion for the ω_i (in terms of the L_i)

$$\begin{aligned}
\partial_t L_1 &= \left(\frac{1}{\mu b^2} - \frac{1}{\mu c^2}\right) L_2 L_3 \\
\partial_t L_2 &= \left(\frac{1}{\mu c^2} - \frac{1}{\mu a^2}\right) L_3 L_1 \\
\partial_t L_3 &= \left(\frac{1}{\mu a^2} - \frac{1}{\mu b^2}\right) L_1 L_2
\end{aligned} \tag{4.30}$$

[4]Note that, in this section, we use units in which e and v have the numerical value 1, following [57]. It follows that $M = 4\pi$ and $\mu = 2\pi$ and that we have χ^2 in stead of $\frac{\chi^2}{v^2}$

Of course, these equations are not really the equations for a rigid body in free motion, because r varies in time. This implies that the "moments of inertia" $\mu a^2(r)$, $\mu b^2(r)$ and $\mu c^2(r)$ change in time and hence the "body" changes shape during the motion. The equation of motion for r that follows from (4.28) is

$$\mu^2 f \partial_t (f \partial_t r) = \frac{\partial_r a}{a^3} L_1^2 + \frac{\partial_r b}{b^3} L_2^2 + \frac{\partial_r c}{c^3} L_3^2 + \tag{4.31}$$

Note that, whatever the time behaviour of r and whatever the form of the functions $a,\ b,\ c$, we can always solve the equations for the angular motion by taking two of the L_i zero and the third equal to a constant. These three kinds of special solutions correspond to rotations of the "body" around one of its principal axes. When the functions $a,\ b,\ c$ and f are known, the time dependence of r can then be determined from the equation (4.31) above.

The rigid body analogy also helps in finding the constants of the motion for the monopole system. Next to the centre of mass momentum and the total electric charge, they are the total energy of the relative motion, given by

$$E = \frac{1}{2}\mu(f^2\dot{r}^2 + a^2\omega_1^2 + b^2\omega_2^2 + c^2\omega_3^2) = \frac{1}{2}\mu f^2\dot{r}^2 + \frac{L_1^2}{2\mu a^2} + \frac{L_2^2}{2\mu b^2} + \frac{L_3^2}{2\mu c^2} \tag{4.32}$$

and the three components of the angular momentum, when measured in an inertial system. Note that these are not the L_i, as these are the components of the angular momentum in the frame that rotates with the body. Still, the length of the angular momentum is the same in the inertial and rotating frames and hence we do know that $L_1^2 + L_2^2 + L_3^2$ is a constant.

To go further, we need to determine the functions $a,\ b,\ c$ and f in the metric. This can be done by using the fact that the moduli space $\mathcal{M}_2^0$ is a hyperkäler manifold. It can be shown that the Riemann curvature tensor for a four dimensional hyperkähler manifold is necessarily self dual. That is, we have to have

$$R_{\mu\nu\rho\sigma} = \frac{1}{2}\epsilon_{\mu\nu\alpha\beta}R_{\alpha\beta\rho\sigma}.$$

This condition translates to a system of differential equations for the components of the metric. For the $SO(3)$-invariant metric we have given in (4.27), this system reduces to the following three equations for the functions $a,\ b,\ c$ and f.

$$\begin{aligned}
\frac{2bc}{f}\frac{da}{dr} &= (b-c)^2 - a^2 \qquad (4.33)\\
\frac{2ca}{f}\frac{db}{dr} &= (c-a)^2 - b^2\\
\frac{2ab}{f}\frac{dc}{dr} &= (a-b)^2 - c^2
\end{aligned}$$

(4.34)

Note that these equations can be obtained from each other by cyclicly permuting $\{a, b, c\}$. Note also that the function f in these equations is to a great extent at our disposal, because it can be changed by a redefinition of the radial coordinate r. Atiyah and Hitchin [53] found that, given f, there is an essentially unique solution to the equations (4.33) which gives a regular, complete metric and makes generic $SO(3)$ orbits three dimensional. [5] Atiyah and Hitchin found this solution by setting $f = abc$. Later, Gibbons and Manton [57] found the same solution in a somewhat more explicit form, which we will use.

In stead of $f = abc$, Gibbons and Manton chose $f = -b/r$. They then set

$$r = 2K\sin(\frac{1}{2}\beta), \tag{4.35}$$

where K is the elliptic integral given by

$$K(k) = \int_0^{\pi/2} d\tau\sqrt{1 - k^2\sin^2\tau}. \tag{4.36}$$

We see that $r(\beta)$ is strictly increasing (and hence invertible) for β between 0 and π. We also see that r has to take values between π and infinity. This implies that we cannot identify r completely with the separation between the monopoles, because there are monopoles which have zero separation. However, we will see that the separation between monopoles does at least increase with r and that, when r is large, it does make sense to think of r as (approximately) the monopole separation. Let us now write down the solution. This takes the form

$$\begin{aligned} w_1 &= -\sin(\beta)r\frac{\mathrm{d}r}{\mathrm{d}\beta} - \frac{1}{2}r^2(1 + \cos(\beta)) \\ w_2 &= -\sin(\beta)r\frac{\mathrm{d}r}{\mathrm{d}\beta} \\ w_3 &= -\sin(\beta)r\frac{\mathrm{d}r}{\mathrm{d}\beta} + \frac{1}{2}r^2(1 - \cos(\beta)) \end{aligned} \tag{4.37}$$

when we define

$$w_1 = bc, \quad w_2 = ca, \quad w_3 = ab$$

From the formulae given here, one can easily derive a formula for a, b and c in terms of the elliptic integral K and its inverse. We will not write this out explicitly, but in stead, we include a numerically obtained graph of the functions a, b and $-c$, taken from [57] (see figure 4.1).

[5] Note that there are also two solutions which do give a complete metric and generically three dimensional $SO(3)$ orbits, but do not result in a a regular manifold. One of these is the solution given by $f = -1$, $a = b = c = \frac{r}{2}$. This looks very much like the flat metric on $\mathbb{R}^4$, but because a, b and c are only equal to $\frac{r}{2}$ in stead of equal to r, there is in fact a conical singularity at $r = 0$. This singularity would be removed if we let the Euler angles parametrise $SU(2)$ in stead of $SO(3)$. The other solution is a Taub-NUT metric we will study later in this section. This, too would give a regular manifold if $SO(3)$ was replaced with $SU(2)$.

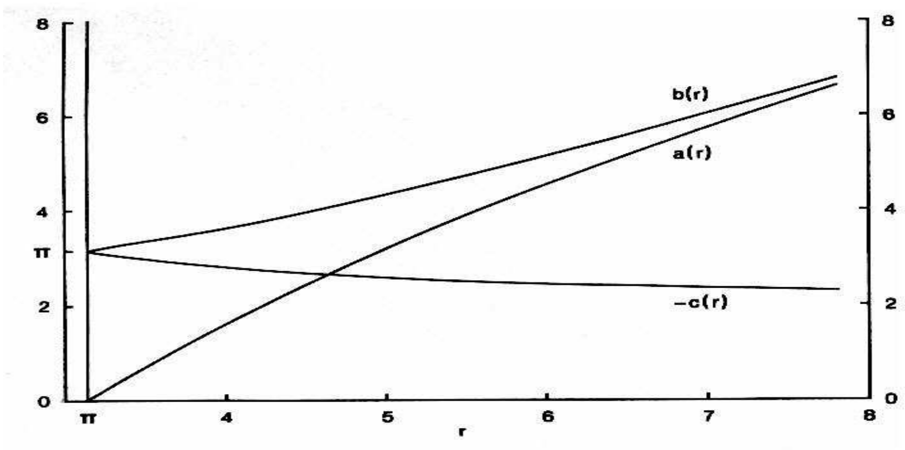


Figure 4.1: graph of the functions a,b and (minus) c that appear in the Atiyah-Hitchin metric (4.27). Taken from [57].

We see that $a = 0$ at $r = \pi$. This reflects the fact that configurations in which the monopoles are both located at the origin are axially symmetric and thus only two dimensional (in stead of three dimensional). We also see this reflected in a coordinate singularity, called a bolt, at $r = \pi$. On a path that goes through the bolt, the Euler angles can make a discontinuous jump without making the path discontinuous. Figure 4.2 illustrates how this comes about.

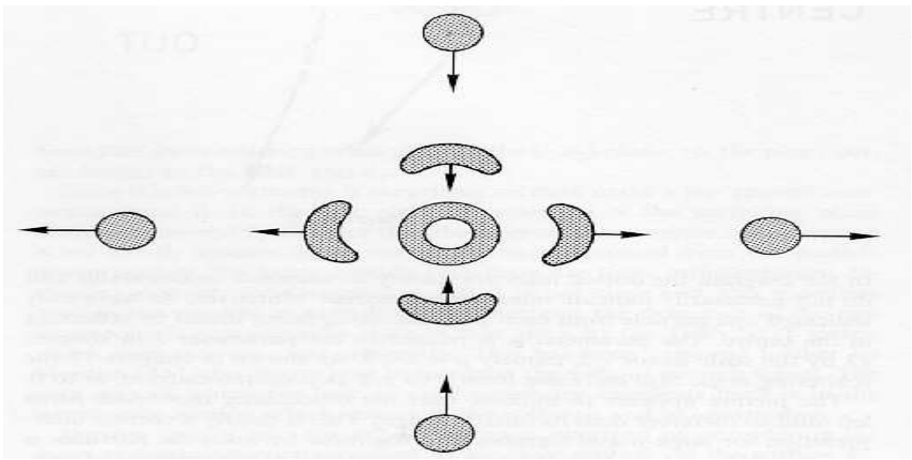


Figure 4.2: Schematic representation of a monopole scattering process. Taken from [53].

This figure shows a schematic representation of a two monopole scattering process. As the monopoles approach each other, they deform until, when they are coincident, their mass is mostly concentrated in a toroidal structure around the origin. [6] After this, the monopoles can separate again in a completely different direction, which creates a jump in the angles which describe the orientation of the axis which joins the monopoles. However, from the picture, it should be clear that this discontinuity in the coordinates does not correspond to a discontinuity in the monopoles' motion.

[6] Numerical analysis shows [58] that this is indeed the case for configurations of two monopoles whose centres coincide

Using an asymptotic expansion for the elliptic integral K, one can obtain an asymptotic expansion for the functions a, b and c as r becomes large. This is given by

$$
\begin{aligned}
a &= r\sqrt{1-\frac{2}{r}} - 4r^2\left(1-\frac{1}{2r^2}\right)e^{-r}+\dots && (4.38)\\
b &= r\sqrt{1-\frac{2}{r}} + 4r^2\left(1-\frac{2}{r}-\frac{1}{2r^2}\right)e^{-r}+\dots \\
c &= -2\sqrt{1-\frac{2}{r}}+\dots && (4.39)
\end{aligned}
$$

The neglected terms in a and b are of order $r^{-1}e^{-r}$, those in c are of order e^{-2r} times an algebraic function of r. We see that, as r becomes large, a and b become large compared to c. In terms of our rigid body analogy, this means that, as r increases, the body becomes long and thin, like a thin rod with the same orientation as axis that joins the monopoles. We also see that a and b become asymptotically equal at large r. It follows that the angular momentum component L_3 is conserved when the monopoles are well separated. It was first argued by Manton in [59] that this extra conserved quantity corresponds to the relative electric charge of the monopoles (we will present his argument in section 4.5). Thus, at large separation, the electric charges of the monopoles are individually conserved.

The exact form of the functions a, b and c is so complicated, that even though all the constants of the motion are known, we still cannot give the general solution to the equations of motion (4.30) and (4.31). However, Atiyah and Hitchin [53] have described the special solutions for which two of the "angular momenta" L_i vanish and the third is a constant (i.e. for which the "body" that corresponds to the monopole configuration rotates about one of its principal axes). We will now give a very brief description of the scattering processes they found.

For $L_2 = L_3 = 0$, we have two monopoles scattering. Their impact parameter and speed are determined by the constants L_1 and E. The time dependence of r can be determined most easily through the conservation law (4.32) for the energy E of the relative motion. Rewriting this, we get

$$\dot{r}^2 = \frac{1}{f^2}\left(\frac{2E}{\mu} - \frac{L_1^2}{\mu^2 a^2(r)}\right) \tag{4.40}$$

From this we see that, if we start out with the monopoles asymptotically far apart and moving towards each other, the relative speed $\dot{r}$ decreases monotonically from its asymptotic value of $\frac{2E}{\mu}$, i.e. the monopoles repel each other. Before their centres coincide, the monopoles reach a point of smallest separation, after which $\dot{r}$ changes sign and the monopoles move back out to infinity. The angle through which the monopoles scatter in this process can also be (numerically) calculated. As the impact parameter goes to zero, this scattering angle increases monotonically from 0 to $\pi/2$.

When L_1 and L_3 are zero and L_2 is a non zero constant, the equation that determines the radial motion becomes

$$\dot{r}^2 = \frac{1}{f^2}\left(\frac{2E}{\mu} - \frac{L_2^2}{\mu^2 b^2(r)}\right) \tag{4.41}$$

At large monopole separations, the motion that follows from this equation is the same as that in the case we discussed previously, because $a = b$ asymptotically. However, when the monopoles get near each other, the scattering can differ from what we saw before, because b, unlike a, does not go to zero at $r = 0$. In stead, it goes to π. This implies that we now have two different kinds of scattering. If L_2^2 is greater than $2\mu\pi^2 E$, then we will still have $\dot{r} = 0$ before the monopole centres coincide and the scattering is similar to that described above. If L_2^2 smaller than $2\mu\pi^2 E$, then the monopoles do move closer and closer until their centres coincide. When the monopole centres coincide, the axis that joins the monopoles is instantaneously rotated by $\pi/2$ and the monopoles separate again in the direction orthogonal to the plane in which they approached. Figure 4.2 can be seen as a schematic representation of this phenomenon if the approaching monopoles are taken to be at different heights in the plane vertical to the paper. In the process, L_2 and L_3 exchange roles, so that the outgoing monopoles have $L_2 = 0$ and $L_3 \neq 0$. This indicates that the monopoles' orbital angular momentum has been converted into a relative electric charge, so that the separating particles are dyons. These dyons attract, as can be seen from the new expression for $\dot{r}^2$ after the exchange of L_2 and L_3. We have

$$\dot{r}^2 = \frac{1}{f^2}\left(\frac{2E}{\mu} - \frac{L_3^2}{\mu^2 c^2(r)}\right) \tag{4.42}$$

and since c^2 diminishes as r grows (cf. figure abcplot), $\dot{r}$ also diminishes as r grows, i.e. the dyons attract. This attraction is quite remarkable, because it depends only on the relative charge of the dyons and not on their absolute charges. In fact, we see that the force between dyons without any orbital angular momentum is always attractive, even if the dyons are oppositely charged. Because of the attraction between these dyons, there are once more two possibilities. Either they have enough kinetic energy to escape to infinity or they don't. In the last case, the dyons come back to the origin, where L_2 and L_3 change roles again, and so they are converted back into monopoles. These then move out to infinity in the plane perpendicular to the plane in which the monopoles originally approached. Obviously, this last kind of scattering process may take arbitrarily long and it is the question whether the geodesic approximation will hold long enough to give an adequate description of the real scattering in all cases.

The case with $L_1 = L_2 = 0$ and $L_3 \neq 0$ is just the reverse of one the $L_1 = L_3 = 0$ processes. We start with two dyons which collide head on and we end up with two monopoles flying away in the plane perpendicular to the line along which the dyons approached.

The simplest scattering process is of course the case $L_1 = L_2 = L_3 = 0$. This is a limiting case of all the families of processes that we have described up to now. In this process, the incoming particles are monopoles which approach each other head on until their centres coincide. As soon as this happens, the axis that joins the monopoles is rotated through an angle of $\pi/2$ radians and the result is a configuration of two monopoles moving out back to back. Figure 4.2 can be seen as a representation of this process.

The monopole scattering orbits we have just described are the only ones which are known exactly. To study general scattering orbits, one has to rely on numerical methods. Some numerical results on classical two monopole scattering were presented by Temple-Raston in [60]. These results indicate that the general scattering has features that are similar to those of the special cases we have just described. That is, there is scattering with and without electric charge exchange and the scattered particles generally do not stay in the plane in which they approached. Temple-Raston also showed that the time behaviour of two monopole configurations can depend chaotically on the initial conditions [61].

We will not go into the numerical analysis of monopole dynamics. In stead, we will study the dynamics of well separated monopoles analytically. To study well separated monopoles (i.e. at large r), we can replace the Atiyah-Hitchin metric with its asymptotic form 4.38. In fact, we will even neglect the exponentially decreasing terms in 4.38 and so we will work with the metric given by

$$\begin{aligned} a = b = -rf &= r\sqrt{1-\frac{2}{r}} \\ -c &= \frac{2}{\sqrt{1-\frac{2}{r}}} \end{aligned} \tag{4.43}$$

This metric is identical in form to the so called Taub-NUT metric which appears in the study of gravitational instantons [62], but it has a negative mass parameter. One can easily check that this asymptotic metric is actually an exact solution of the equations (4.33) for the functions a, b and c. It can also be shown that it is geodesically complete and finite, but that the corresponding manifold has a conical singularity [53].

When one uses the Taub-NUT metric, it makes sense to interpret r as exactly the separation of the monopole centres. We will see this in section 4.5, where we rederive the asymptotic metric through a different method. At first sight, the interpretation of r as the inter-monopole distance may seem problematic, because the metric has a singularity at $r = 2$. However, this singularity has no physical significance, as the dynamics of monopoles can only be approximated through this metric when r is large.

The main reason to work with the Taub-NUT metric is that, for this metric, we have $a = b$. It follows from this that the angular momentum component L_3, which was only asymptotically constant in the Atiyah-Hitchin metric, is now an exact constant of the motion. Physically, this means that, in the system described by the Taub-NUT metric, the monopoles' electric charges are individually conserved. The fact that L_3 is conserved considerably simplifies the dynamics. In fact, it turns out that we can now find all classical orbits using only the constants of the motion. Before we do this, let us first write the metric in a more convenient form. When we substitute the explicit form of the one forms σ_i (see equations (4.26)) and the functions a, b and c into the equation (4.27) for the metric (noting that $f = -b/r$), we get

$$ds^2 = \left(1-\frac{2}{r}\right)(dr^2 + r^2 d\theta^2 + r^2\sin^2(\theta)d\phi^2) + 4\left(1-\frac{2}{r}\right)^{-1}(d\psi + \cos\theta d\phi)^2. \tag{4.44}$$

We recognise the first term in this equation as $1-\frac{r}{2}$ times the ordinary flat metric in $\mathbb{R}^3$, when this is expressed in spherical coordinates (r, θ, ϕ). We can thus simplify this term by introducing Cartesian coordinates (r_1, r_2, r_3) for the vector that connects the monopoles' centres. These will, as usual, be given by

$$
\begin{aligned}
r_1 &= r\cos(\phi)\sin(\theta) \\
r_2 &= r\sin(\phi)\sin(\theta) \\
r_3 &= r\cos(\theta)
\end{aligned}
$$

Note that taking $\theta \mapsto \pi - \theta$ and $\phi \mapsto \pi + \phi$ corresponds to taking $\vec{r} \mapsto -\vec{r}$, so one may also easily rewrite the coordinate identifications (4.25) in these new coordinates.

The metric now becomes

$$
ds^2 = \left(1 - \frac{2}{r}\right) d\vec{r} \cdot d\vec{r} + 4\left(1 - \frac{2}{r}\right)^{-1} (d\psi + \vec{w}(\vec{r}) \cdot d\vec{r})^2 \tag{4.45}
$$

where we have defined the (vector) function $\vec{w}$ through

$$
\vec{w}(\vec{r}) \cdot d\vec{r} = \cos(\theta) d\phi
$$

It is not difficult to find an explicit expression for $\vec{w}$ in terms of the Cartesian coordinates r_i, but the result is a bit messy and we will not need it. What we will need is the curl of $\vec{w}$. This is most easily computed in spherical coordinates. First, we have to give $\vec{w}$ itself in terms of these. One can easily check that $d\phi$ is equal to $\frac{d\vec{r}}{r\sin(\theta)}$ and hence that we have the following expressions for $\vec{w}$ and its curl:

$$
\begin{aligned}
\vec{w} &= \frac{\cos(\theta)}{r\sin(\theta)}\hat{\phi} & (4.46)\\
\nabla \times \vec{w} &= \frac{\hat{\mathbf{r}}}{r\sin(\theta)}\partial_\theta(\sin(\theta) w_\phi) + \frac{\hat{\theta}}{r}\partial_r(r w_\phi) \;=\; -\frac{\vec{r}}{r^3}. & (4.47)
\end{aligned}
$$

Thus we see that $\vec{w}$ is the vector potential for a magnetic monopole located at the space origin.

Let us now write down the Lagrangian for geodesic motion in the metric (4.45). This is given by

$$
L = \frac{\mu}{2}\left(1 - \frac{2}{r}\right)\dot{r}_i\dot{r}_i + 2\mu\left(1 - \frac{2}{r}\right)^{-1}(\dot{\psi} + w_i(\vec{r})\dot{r}_i)^2 \tag{4.48}
$$

The fact that L_3 is conserved can easily be rederived from this Lagrangian; it is just the equation of motion for the cyclic coordinate ψ. Because we will interpret this conserved quantity as the relative electric charge of the monopoles, we will call it q, i.e. we define

$$
q = -4\mu\left(1 - \frac{2}{r}\right)^{-1}(\dot{\psi} + w_i(\vec{r})\dot{r}_i). \tag{4.49}
$$

The equation of motion for ψ is then just $\dot{q} = 0$. Let us now derive the equations of motion for the r_i. We could find these directly from the equations of motion (4.30) and (4.31) we gave earlier, but it is actually easier just to derive them from the Lagrangian (4.48) above. The momenta π_i conjugate to the r_i are given by

$$\pi_k = \mu \left(1 - \frac{2}{r}\right) \dot{r}_k + q w_k \tag{4.50}$$

We see that π_k consists of a part that corresponds asymptotically to the usual linear momentum and a part due to the electromagnetic interaction of the dyons. To get more familiar formulae later on, it is useful to introduce the notation p_k for the first part of π_k, i.e. we define

$$p_k = \mu \left(1 - \frac{2}{r}\right) \dot{r}_k$$

The time derivative of π_k can then be written as

$$\dot{\pi}_k = \dot{p}_k - q \partial_j w_k \dot{r}_j, \tag{4.51}$$

where we have used the fact that q is time independent. To get the equations of motion, we also need $\frac{\partial L}{\partial r_k}$. Straightforward calculation yields

$$\frac{\partial L}{\partial r_k} = \mu \frac{r_k}{r^3} \left(\dot{r}_i \dot{r}_i - (\frac{q}{2\mu})^2\right) - q \partial_k w_i \dot{r}_i \tag{4.52}$$

combining this equation with the equation (4.51) for $\dot{\pi}_k$, we find that the equations of motion are

$$\dot{p}_k = \mu \frac{r_k}{r^3} \left(\dot{r}_i \dot{r}_i - (\frac{q}{2\mu})^2\right) - q(\partial_k w_i \dot{r}_i - \partial_j w_k \dot{r}_j).$$

We may now rewrite the last term in this equation in a much nicer form, as follows. We have

$$\begin{aligned} \partial_i w_j \dot{r}_j - \partial_j w_i \dot{r}_j &= (\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl})(\dot{r}_j \partial_l w_m) \\ &= \epsilon_{ijk} \dot{r}_j \epsilon_{klm} \partial_l w_m \\ &= (\dot{r} \times (\nabla \times w))_i \;\; = \;\; - \left(\frac{\dot{\mathbf{r}} \times \mathbf{r}}{r^3}\right)_i \end{aligned}$$

and hence the equations of motion become

$$\dot{\mathbf{p}} = \mu \frac{\mathbf{r}}{r^3} \left(\dot{\mathbf{r}} \cdot \dot{\mathbf{r}} - (\frac{q}{2\mu})^2\right) + q(\frac{\dot{\mathbf{r}} \times \mathbf{r}}{r^3}). \tag{4.53}$$

From this formula we see that dyons experience an attractive coulomb force proportional to q^2, a Lorentz force proportional to q and a repulsive force proportional to the square of their relative velocity, but independent of q. When the dyons are equally charged, only this last, repulsive force remains and thus we can expect equally charged dyons to have only scattering orbits. However, if the dyons' charges are not equal and the relative speed is low, then the attractive Coulomb force dominates the dynamics. Thus, we may expect that there will be bound orbits for unequally charged dyons, as well as scattering orbits. Note that we can also find solutions to the equations of motion that correspond to the static BPS-monopoles and Julia-Zee dyons we started with. These solutions are of course given by $q = 0$ and $\mathbf{r} = const.$

To get some more concrete information the orbits, let us write down the constants of the motion in Cartesian coordinates. Next to q, we have the total energy (4.32) of the relative motion, which is now given by

$$E = \frac{\mu}{2}\left(1 - \frac{2}{r}\right)\left(\dot{\mathbf{r}} \cdot \dot{\mathbf{r}} + \left(\frac{q}{2\mu}\right)^2\right). \tag{4.54}$$

Using this equation, we can write the equations of motion in the form

$$\dot{\mathbf{p}} = \frac{\mathbf{r}}{r^3}\left(\frac{2E}{1 - \frac{2}{r}} - \frac{q^2}{2\mu}\right) + q(\frac{\dot{\mathbf{r}} \times \mathbf{r}}{r^3}). \tag{4.55}$$

which shows that, when E is fixed, the central force between the monopoles becomes velocity independent. We also see from this formula that when r is large (much larger than 2), the force becomes just the sum of a Coulomb force and a Lorenz force. This Coulomb force will be attractive when E is smaller than $\frac{q^2}{4\mu}$ and repulsive otherwise. Thus we see that, for $E < \frac{q^2}{4\mu}$, we can expect to have bound states in which r is always large, whereas for $E > \frac{q^2}{4\mu}$, we can only expect scattering states (except if r is smaller than 2, but then the Taub-Nut approximation is not valid). Note also that pure monopoles (q=0) will not have bound states, as the force between them is always repulsive (for $E > 0$ and $r > 2$)

The next conserved quantities are the three components of the angular momentum vector $\mathbf{J}$ given by

$$\mathbf{J} = \mathbf{r} \times \mathbf{p} - q\hat{\mathbf{r}} \tag{4.56}$$

We will not derive this form of the angular momentum from first principles, but one may easily check that $\mathbf{J}$ is indeed conserved by computing its time derivative and substituting the equations of motion. The first term on the right hand side of the equation for $\mathbf{J}$ is just the ordinary orbital angular momentum, while the second is a term which appears due to the simultaneous presence of both magnetic and electric charges. Since this second term is orthogonal to the first, we see that the length of the orbital angular momentum is also conserved.

Next to q and the conserved quantities E and $\mathbf{J}$ which were also present for the exact Atiyah-Hitchin metric, there are three more conserved quantities, united in a conserved vector $\mathbf{K}$, which is analogous to the Runge-Lenz vector which occurs in a system of two particles that interact through a central $\frac{1}{r^2}$-potential [55].[7] We have

$$\mathbf{K} = \mathbf{p} \times \mathbf{J} + \left(2\mu E - \frac{1}{2}q^2\right)\hat{\mathbf{r}} \tag{4.57}$$

Again, one may easily check that $\mathbf{K}$ is conserved using the equations of motion.

Using $\mathbf{J}$ and $\mathbf{K}$, we can now determine the orbits for two monopole scattering. First, we note that the inner product of $\mathbf{J}$ and $\hat{\mathbf{r}}$ is constant and equal to $-q$. It follows that the angle between $\mathbf{r}$ and the constant vector $\mathbf{J}$ is always equal to $\arccos(|q|/|J|)$. Hence $\mathbf{r}$ lies on a cone with axis $\mathbf{J}$ and opening angle $\arccos(|q|/|J|)$. Further information about the orbits can be obtained by computing the inner product of $\mathbf{K}$ and $\mathbf{r}$. We have

$$\mathbf{K} \cdot \mathbf{r} = (\mathbf{r} \times \mathbf{p}) \cdot \mathbf{J} + \left(2\mu E - \frac{1}{2}q^2\right) r.$$

When we write $\mathbf{r} \times \mathbf{p} = \mathbf{J} + q\hat{\mathbf{r}}$ and use $\mathbf{J} \cdot \hat{\mathbf{r}} = -q$, this equation can be rewritten as

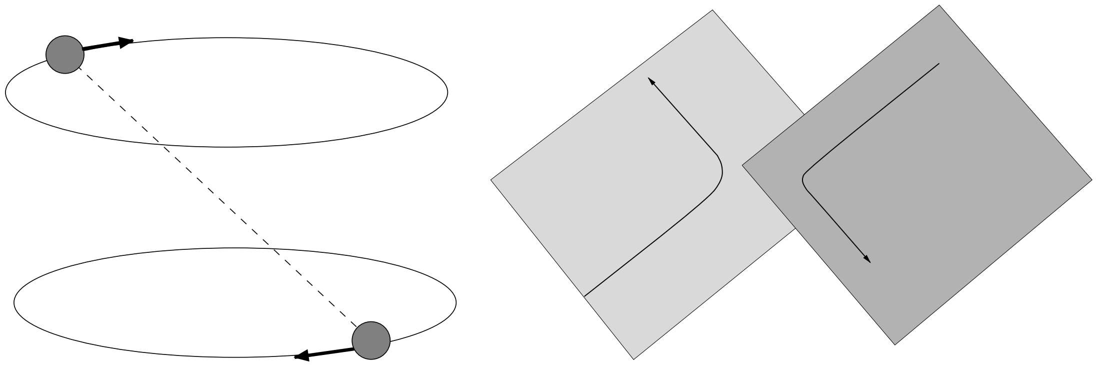
$$\left(\mathbf{K} + \frac{1}{q}(2\mu E - \frac{1}{2}q^2)\mathbf{J}\right) \cdot \mathbf{r} = J^2 - q^2, \tag{4.58}$$

from which we see that $\mathbf{r}$ is confined to a fixed plane. Combining this with the fact that r was already confined to a cone, we see that the orbits must be conic sections, i.e. points, (pairs of) lines, circles and ellipses, which correspond to bound states and parabolae and hyperbolae, which describe scattering processes. Note that our derivation that the orbits are conic sections is valid only of q is non zero. Thus, we have to treat the special case of pure monopole scattering separately. In this case, the orbits can be determined through the equation $\mathbf{J} \cdot \mathbf{r} = 0$, which shows that $\mathbf{r}$ lies in the plane orthogonal through $\mathbf{J}$ that contains the origin, and the equation $\mathbf{K} \cdot \mathbf{r} = J^2 + 2\mu E r$, which determines a cone-like surface. It turns out that all the orbits are hyperbolae, in accord with our earlier statement that monopoles have only scattering orbits.

Note that the fact that $\mathbf{r}$ moves in a fixed plane does not imply that the two particles whose separation $\mathbf{r}$ describes move in the same plane. In fact, it is easy to see that, in the centre of mass frame, the dyons have to move on conic sections in distinct parallel planes. The distance between these planes is just the distance from the plane given in (4.58) to the origin. In the case of pure monopoles, the particles do move in the same plane throughout their motion, which follows from the fact that $\mathbf{r}$ now lies in a plane through the origin, or alternatively, from the fact that the force between pure monopoles is central. Schematic pictures of a circular bound dyon orbit and a hyperbolic scattering orbit of two dyons are given in figure 4.4

The fact that dyons move in distinct parallel planes has consequences for the description of dyon scattering. Though dyons do not move in the same plane throughout their motion, they do of course move in the same plane asymptotically long before and after the actual scattering, when their tracks are just straight lines. However, these planes do not

[7] see also especially exercise 23 on page 124 of [55], where Goldstein treats the angular momentum and Runge-Lenz vector for a system that includes a Lorentz force



(a) (b)

Figure 4.3: (a) circular bound dyon orbit (b) hyperbolic scattering dyon orbit

have to be the same. We say that the orbits are twisted. The twist is described by a so called twist angle γ. Next to γ, we have the usual scattering angle between the incoming and outgoing tracks of the scattered particles/ See fig. 4.4 for a schematic picture of what happens.

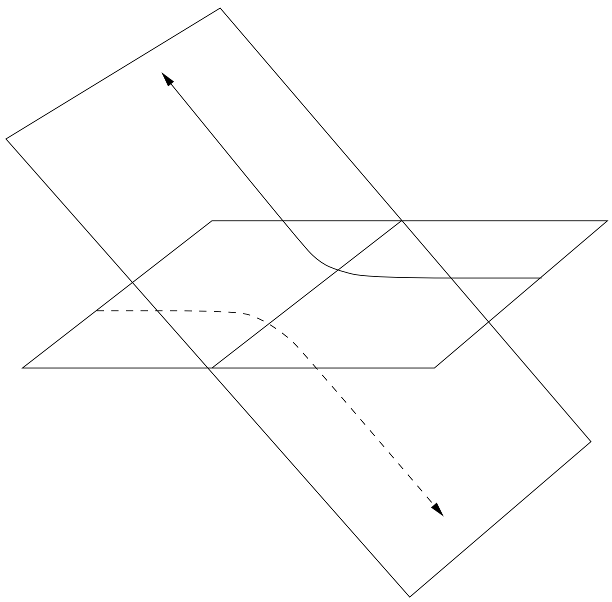


Figure 4.4: Schematic picture of a dyon scattering process. The initial and final planes of motion are shown to demonstrate the notion of a "twisted" orbit

Gibbons and Manton have given a complete treatment of dyon and monopole scattering. They find the twist and scattering angles in terms of speed, impact parameter and relative charge and give a differential cross section. Gibbons and Manton have also explicitly treated the circular orbits which correspond to bound states of dyons. For all this, we refer the reader to [57].

To end this section, I want to make some remarks about the quantum mechanics on $\mathcal{M}_2^0$. As for the case of $\mathcal{M}_1^0$, this can be treated by writing down the Hamiltonian for geodesic motion and converting this into an operator that works on functions on the moduli space. The resulting operator is just the covariant Laplacian on $\mathcal{M}_2$. There are some problems with this quantum mechanical version of the moduli space approximation which should be noted. First, the ground state energy of the non zero modes of a monopole will probably vary with the collective coordinates, thus causing an effective potential of order $\hbar$ on the moduli space, which is ignored in the geodesic approximation. This problem is already present in the classical theory, where the small oscillations orthogonal to the moduli space have an influence on the motion along the moduli space (cf. the discussion and graphs on page 31), but it is all the more pressing in the quantum theory, where it is fundamentally impossible even to choose initial conditions with no energy stored in the transverse modes. The second and more important problem with the moduli space approximation is purely a quantum effect. Classically, the mass of the Higgs field is exactly zero because the Higgs potential is taken equal to zero, but quantum mechanically this is not so. When we go beyond tree approximation, the Higgs field still has self interaction and this causes it to get a small mass. When the Higgs field has a non zero mass, it will also have a finite range and hence, it will no longer be able to mediate a long range force which cancels the repulsive Coulomb force between the monopoles. Therefore, it is unlikely that there will be static multimonopole configurations in the exact quantum theory. Static monopoles would repel each other. This problem can only be really resolved in supersymmetric generalisations of the Georgi-Glashow model, where supersymmetry ensures that the mass of the Higgs field stays equal to zero. However, in our version of the model, we may hope that the effect of monopole-monopole repulsion is small compared to the effects we have taken into account (being a purely quantum mechanical effect, it should be of order $\hbar$) and continue with the geodesic approximation for lack of a better method.

I do not have time to go into the quantum dynamics on $\mathcal{M}_2$ in detail, but I will state some general features and give references for the interested reader. The quantum problem exhibits many similarities with the classical one. The Schrödinger equation for the exact Atiyah-Hitchin metric has defied a general solution, but Gibbons and Manton [57] have solved the Schrödinger equation for the Taub-NUT metric explicitly in terms of confluent hypergeometric functions. Again, they found both bound states and scattering states and they have given binding energies for the first and cross sections for the last. From their results, they draw the conclusion that the exact metric will also exhibit bound and scattering states, but that some of the bound states in the Taub-NUT metric will become unstable and turn into resonances. Note that we could have expected resonances to occur from the exact classical dynamics, because there, we already found a type of dyon scattering process in which the scattering time could increase without bound. Bound states for the Atiyah-Hitchin metric were later studied numerically by Manton [63]. Their binding energies turned out to be in very good agreement with the Taub-NUT results. The quantum mechanics of dyon scattering in the Atiyah-Hitchin metric were extensively studied by Schroers [64], using a partial wave expansion and numerical methods.

4.5 Asymptotic metric and dynamics on $\mathcal{M}_n$

In this section, we will concern ourselves with the metric on the n-monopole moduli space $\mathcal{M}_n$. For $n > 2$, not many analytic results about this metric are available. Of course, we know from section 4.2 that $\mathcal{M}_n$ has a simply connected n-fold cover which splits as

$\mathbb{R}^3 \times S^1 \times \mathcal{M}_n^0$, where the factor $\mathbb{R}^3 \times S^1$ is flat, but generally, little is known about the exact metric on $\mathcal{M}_n^0$. As far as I am aware, there are only two explicit results for the exact metric on moduli spaces for more than two monopoles. The first of these is a proof that each of the $\mathcal{M}_n^0$ $(n > 2)$ contains a totally geodesic sub-manifold that consists of configurations of collinear monopoles [65, 66] and has the same metric as the Atiyah-Hitchin manifold. The second is a calculation of the exact metric on the space of tetrahedrally symmetric 4-monopoles [67]. Rather than to go into detail about these exactly known, but rather special cases, we will use this chapter to calculate an asymptotic form of the metric on $\mathcal{M}_k$ which is valid if all the monopoles are well separated. The calculation we will give was first done by Gibbons and Manton in 1995 [68], but the idea goes back on an article by Manton from 1985 [59], in which the same calculation is performed for the special case of $\mathcal{M}_2$. This led to the Taub-NUT metric we have studied in section 4.4. As we saw there, this metric approximates the exact metric exponentially well and this led Gibbons and Manton to conjecture that their asymptotic metric for n monopoles would also approach the real metric asymptotically fast when the monopoles' separations become large. Gibbons and Manton did not prove this and neither will we. The calculation that they gave (and that we will repeat) is based on physical, rather than mathematical arguments and it is not clear a priori that it is mathematically sound. However, in a recent preprint, Bielawski [69] has managed to prove that the Gibbons-Manton metrics are in fact asymptotically valid up to exponentially decreasing terms. His proof uses the mathematical apparatus of twistor theory and goes through the moduli space of Nahm data which we have briefly touched upon before. Using similar methods, Bielawski also found asymptotic metrics for monopoles with gauge group $SU(k)$ [70]. Unfortunately, we will not be able to go into details about Bielawski's results in this thesis, because it would take too much time and space to introduce the necessary mathematics. We will give the physical calculation of Gibbons and Manton, not only because it happens to give the right result, but also because it uses physical ideas that have recently been used and extended by Bak, Lee and Lee [71] to come to a promising new description of the asymptotic dynamics of BPS monopoles.

The idea on which the calculation is based is simple. BPS-dyons are subject to two kinds of interactions: long range interactions, mediated by the Higgs field and the (Abelian) electromagnetic fields and short range interaction mediated by the components of the vector potential that correspond to the heavy gauge bosons. Since we want to treat well separated dyons, we will ignore the short range interactions. This immediately implies that the electric charges of the individual dyons are conserved, because electric charge exchange can only take place through the short range interactions (quantum mechanically: through the exchange of charged gauge bosons). Our strategy is now to find a Lagrangian for the electromagnetic and Higgs interaction of n dyons with magnetic charge g and small, fixed electric charges $q_1, \dots, q_n$ and to extract a metric from this later, by interpreting the q_i as speeds. Throughout the treatment, we will treat the dyons as point particles, an idealisation which should be justified at large separations. Next to the usual electric and magnetic fields $\mathbf{E}$ and $\mathbf{B}$, the dyons carry a scalar field ϕ which represents the physical content of the Higgs field Φ. Note that most of the physical content of the Higgs field is in its length; its direction can be changed by gauge transformations. We can even have magnetic monopoles with the Higgs field in a constant direction if we allow singular gauge potentials. In the full theory, we did not like to have singularities in the gauge potential, but in the long distance approximation we are using now we can not even see the gauge field of the full theory and so we may as well work with a one component Higgs field ϕ, which has to be some function of the length of the original Higgs field Φ. We will define $\phi = v - |\Phi|$, because this gives convenient formulae. For example, using formula (3.43), we see that the scalar field of a Julia-Zee dyon at long distance is given by

$$\phi(r) = \frac{g_s}{4\pi r} \tag{4.59}$$

where we have defined the *scalar charge* g_s of the dyon through $g_s = \sqrt{q^2 + g^2}$

Let us now write down the Lagrangian for the n^{th} dyon, as it moves in the fields of all the other dyons. It is of the form

$$L_n = -\frac{1}{\gamma(\dot{r}_n)}\left(m_n - g_s\phi\right) + q_n\dot{\mathbf{r}}_{\mathbf{n}} \cdot \mathbf{A} - q_n A_0 + g\dot{\mathbf{r}}_{\mathbf{n}} \cdot \tilde{\mathbf{A}} - g\tilde{A}_0. \tag{4.60}$$

Here $\mathbf{A}$ and A_0 are the usual electromagnetic vector and scalar potentials at the location of dyon n due to all the other dyons, ϕ is the scalar Higgs field we have just defined and $\tilde{\mathbf{A}}, \tilde{A}_0$ are the so called dual vector and scalar potentials, which are defined so that we have

$$\begin{aligned} \nabla \times \tilde{\mathbf{A}} &= -\mathbf{E} \\ -\nabla \tilde{A}_0 - \frac{\mathrm{d}}{\mathrm{dt}}\tilde{\mathbf{A}} &= \mathbf{B} \end{aligned}$$

The reason we need these dual potentials is that the ordinary potentials couple only to the electric charge, thus giving only the ordinary Lorentz force due to the magnetic field and the ordinary Coulomb force due to the electric field of dyons 1 to $n-1$. When we couple the dual potentials to the magnetic charge, we also get the dual Coulomb force due to the magnetic field and the dual Lorentz force due to the electric field. The way in which the ordinary electromagnetic potentials appear in the Lagrangian is of course just the standard way. For a derivation that this is the right way, see for instance Jackson [72]. The dual potentials naturally have to appear in the same way as the ordinary ones (but with q_n replaced by g) to obtain the right dual Lorentz and Coulomb forces. Note that the dual and ordinary potentials are of course not independent; they both yield the same fields. In fact, one may obtain the dual potentials from the ordinary ones (up to gauge transformations) and vice versa. (see for example [71]) However, the resulting expressions are complicated and for our purposes, it is much easier to keep both types of potentials present.

The scalar field ϕ of dyons 1 to $n-1$ appears in the Lagrangian in such a way that its presence just modifies the rest mass of dyon n. One can show easily that, at least when the electromagnetic and dual potentials are zero, this is the only Lorentz invariant possibility ([72], ex. 12.3). However, the exact way in which the Higgs field should modify the particle's mass can not be deduced from Lorentz invariance. For this, one really needs to know the force that the particle feels due to the presence of the Higgs field.

The first calculation of this force is due to Manton. In his article [28], he calculated the force that works on a monopole in a constant magnetic field B in the monopole's instantaneous rest frame, as well as the total force that works between two well separated monopoles which are at rest at but have an acceleration along the axis that joins them. Indirectly, this gives the force the Higgs field exerts on a monopole as the difference between the total force and the force due to the (nearly constant) magnetic field of the other monopole.

The calculation of the forces proceeded through the following method. Starting from the exact BPS monopole solution, Manton constructed an ansatz for a moving solution by replacing the spatial variable $\mathbf{x}$ with $\mathbf{x}-\frac{1}{2}\mathbf{a}t^2$, where a is assumed to be small. In other words, he allowed the monopole to get a small non zero acceleration $\mathbf{a}$, but he did not allow it to change shape. He then substituted this ansatz into the full Yang-Mills-Higgs equations of motion and required that they were satisfied at least to first order in a. This requirement gives an equation from which $\mathbf{a}$ can be determined in terms of the fields. From the acceleration $\mathbf{a}$, one can then immediately calculate the force through Newton's formula $\mathbf{F}=m\mathbf{a}$

Manton only looked at the rather special cases described above, but this still provided him with enough information to find the way in which the Higgs field Φ has to appear in the Lagrangian. The electromagnetic and dual potentials were just put in in such a way that they would exert the forces we are used to from classical electrodynamics, even though the real forces had only been calculated for the special case of pure monopoles at rest. Though this is certainly the easiest and most intuitive way to handle the electromagnetic potentials, it would be prudent to do a calculation like the one Manton did to see if moving BPS dyons really respond to electromagnetic fields in the same way as ordinary dyons. Fortunately, we won't have to do this ourselves, because it has recently been done by Bak, Lee and Lee [71]. Using a refinement of Manton's method [8], they have calculated the forces that work on a dyon that moves in weak, asymptotically constant electric, magnetic and Higgs field strengths. They found that, in the dyon's instantaneous rest frame, these were given by the formula

$$M\mathbf{a}=g\mathbf{B}+q\mathbf{E}+g_s\mathbf{H}, \tag{4.61}$$

where $\mathbf{B}$ and $\mathbf{E}$ are the constant magnetic and electric fields, g_s is the so called scalar charge of the dyon, which is given by $g_s=\sqrt{q^2+g^2}$, and $\mathbf{H}$ is the Higgs field strength, which is defined through

$$H_\mu=-\hat{\Phi}^a D_\mu \Phi^a$$

They also generalised the force formula above to give the force law for a dyon that moves at an arbitrary velocity $\mathbf{v}$. This is

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\gamma(\mathbf{v})(M-g_s r_\mu H^\mu)\mathbf{v}\right)=g(\mathbf{B}-\mathbf{v}\times\mathbf{E})+q(\mathbf{E}+\mathbf{v}\times\mathbf{B})+\frac{g_s}{\gamma(\mathbf{v})}\mathbf{H} \tag{4.62}$$

From this force law, we want to extract a force law that involves the scalar Higgs field ϕ, rather than the Higgs field strength $\mathbf{H}$. To this end, note the following identity

[8] Unlike Manton's method, Bak, Lee and Lee's method also allows small deformations of the monopoles. For a short review, see section 4.6

$$H_\mu = -\hat{\Phi}^a D_\mu \Phi^a = -\frac{\Phi^a \partial_\mu \Phi^a}{|\Phi|} = -\partial_\mu |\Phi| = -\partial_\mu \phi$$

Using this, we get the following force equation in terms of the electromagnetic and dual potentials and the scalar Higgs field.

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\gamma(\mathbf{v})(M - g_s\phi)\mathbf{v}\right) = g(\mathbf{B} - \mathbf{v}\times\mathbf{E}) + q(\mathbf{E} + \mathbf{v}\times\mathbf{B}) - \frac{g_s}{\gamma(\mathbf{v})}\nabla\phi \tag{4.63}$$

But this equation is just the equation of motion for the Lagrangian (4.60) and so we see that Bak, Lee and Lee's work provides a fairly solid basis for the choice of Lagrangian we have made, at least if the fields which appear in the Lagrangian vary slowly or seem to vary slowly at the locations of the dyons, because the dyons move through them slowly.

The next thing we have to do is write down expressions for the potentials and scalar field caused by each of the dyons 1 to $n-1$, as observed at the location of dyon n. This is not an easy task. In fact, the simplest way to find the exact electromagnetic and Higgs fields for an arbitrary configuration of moving dyons may be to solve the full equations of motion for Yang-Mills-Higgs theory. Of course, we have found one particularly important solution to these equations in chapter 3: Prasad and Sommerfield's one monopole solution. The electric and magnetic fields for this solution were just the same as those for a classical point dyon, because the electric and magnetic charges were concentrated in one point. When we Lorentz boost the Prasad-Sommerfield solution, we still have an exact solution and from this, we can obtain the exact electric, magnetic and Higgs fields for a single monopole moving at a constant speed. Of course, the electric and magnetic charge densities for the Lorentz boosted solution are just those of a uniformly moving classical point dyon and hence the electric and magnetic fields will also be those of a uniformly moving point dyon. Now the fields of an arbitrarily moving electric point charge of strength q may be obtained from the Liénard-Wiechert potentials (see [72] for a derivation) given by

$$\begin{aligned} A_0(\mathbf{x},t) &= \frac{q}{4\pi(|\mathbf{x}-\mathbf{r}(t_0)| - \mathbf{v}(t_0)\cdot\mathbf{r}(t_0))} \\ \mathbf{A}(\mathbf{x},t) &= \frac{q\mathbf{v}(t_0)}{4\pi(|\mathbf{x}-\mathbf{r}(t_0)| - \mathbf{v}(t_0)\cdot\mathbf{r}(t_0))}, \end{aligned} \tag{4.64}$$

where $\mathbf{r(t)}$ is the position of the charge at time t and where t_0 is the so called retarded time, which can be determined from the condition that $\mathbf{r}(t_0)$ has to lie on the backward light cone from $\mathbf{x}$.

If we write these potentials as $(qA_0^{lw}, q\mathbf{A}^{lw})$ and their dual potentials (which yield the same fields) as $(q\tilde{A}_0^{lw}, q\tilde{\mathbf{A}}^{lw})$, then we can write the potentials for a moving dyon as

$$\begin{array}{rclcrcl} A_0 &=& qA_0^{lw} + g\tilde{A}_0^{lw} & \quad & \mathbf{A} &=& q\mathbf{A}_0^{lw} + g\tilde{\mathbf{A}}^{lw} \\ \tilde{A}_0 &=& gA_0^{lw} - q\tilde{A}_0^{lw} & \quad & \tilde{\mathbf{A}} &=& g\mathbf{A}_0^{lw} - q\tilde{\mathbf{A}}^{lw} \end{array} \tag{4.65}$$

In the special case that the velocity $\mathbf{v}$ in (4.64) is taken constant, these potentials give the exact fields for a moving BPS-dyon. If the velocity is not constant, then we can not be sure that a BPS-dyon will still behave exactly as an ordinary point dyon, because we can not be sure that the electric charge will stay concentrated in one point. However, the work of Bak, Lee and Lee [71] indicates that, at least when the acceleration of a BPS-dyon is small, its electromagnetic fields are still consistent with those of an ordinary point dyon. Therefore, it is likely that the equation for the potentials given above will be valid in more general cases than that of a uniformly moving dyon alone.

We shall not be so ambitious as to try and work with this general equation for the potentials. Our goal is just to find an effective Lagrangian in terms of the dyon velocities and positions which is correct to second order in the electric charges q_i and the velocities $\mathbf{v}_i$, when these velocities and charges are small. Looking back at the Lagrangian (4.60), we see that this means that we need to know $\mathbf{A}$, to order v^0, we need to know A_0 and $\tilde{\mathbf{A}}$ to order v and we need to know $\tilde{A}_0$ to order v^2. [9]

Let us then find an expansion of the Liénard-Wiechert potentials in powers of the velocity. When we assume that the velocity of the dyons is small and that their acceleration is negligible to order v^2, we may write

$$\mathbf{r}(t_0) = \mathbf{r} - \mathbf{v}|\mathbf{x} - \mathbf{r}(\mathbf{t_0})| + O(v^3).$$

The expansion of the Liénard-Wiechert potentials in powers of the velocity that we are looking for can be found by iterated substitution of the above formula in the formulae (4.64). Explicitly, we find that the denominator in the formulae (4.64) is given by

$$\begin{aligned} |\mathbf{x} - \mathbf{r}(t_0)| - \mathbf{v}(t_0)\cdot\mathbf{r}(t_0) &= |\mathbf{x} - \mathbf{r}(t)| + \frac{1}{2}\left(\frac{(\mathbf{v}\cdot(\mathbf{x} - \mathbf{r}(t)))^2}{|\mathbf{x} - \mathbf{r}(t)|} - v^2|\mathbf{x} - \mathbf{r}(t)|\right) + O(v^3) \\ &= |\mathbf{x} - \mathbf{r}(t)| + \frac{(\mathbf{v}\times(\mathbf{x} - \mathbf{r}(t)))^2}{2|\mathbf{x} - \mathbf{r}(t)|} + O(v^3) \end{aligned}$$

Using this result, it is easy to see that the potentials due to dyon 1 and evaluated at the position of dyon n must be given by

$$\begin{array}{rclrcl} A_0 &=& \frac{q_1}{4\pi r_{n1}} - \frac{g\mathbf{w}_{n1}\cdot\mathbf{v}_1}{4\pi} + O(v^2) & \mathbf{A} &=& -\frac{g\mathbf{w}_{n1}}{4\pi} + O(v) \\ \tilde{A}_0 &=& \frac{g}{4\pi r_{n1}}\left(1 + \frac{(\mathbf{v}_1\cdot\hat{\mathbf{r}}_{n1})^2}{2} - \frac{v^2}{2}\right) + \frac{q_1\mathbf{w}_{n1}\cdot\mathbf{v}_1}{4\pi} + O(v^3) & \tilde{\mathbf{A}} &=& \frac{g\mathbf{v}_1}{4\pi r_{n1}} + \frac{q_1\mathbf{w}_{n1}}{4\pi} + O(v^2). \end{array} \tag{4.66}$$

Here we have defined $\mathbf{r}_{ij} := \mathbf{x}_i - \mathbf{x}_j$ and we have introduced a Dirac vector potential $\mathbf{w}(\mathbf{x})$ which satisfies $\nabla\times\mathbf{w}(\mathbf{x}) = -\frac{\mathbf{x}}{x^3}$ and $\mathbf{w}(\mathbf{x}) = \mathbf{w}(-\mathbf{x})$. A possible choice for $\mathbf{w}$ is the potential we gave in formula (4.46), but we may also take a different potential, related to this by a gauge transformation. We have also defined $\mathbf{w}_{ij} := \mathbf{w}(\mathbf{r}_{ij})$ and obviously, we have $\mathbf{w}_{ij} = \mathbf{w}_{ji}$ The terms that involve $\mathbf{w}_{n1}$ are the lowest order terms in the dual Liénard-Wiechert potentials that appear in the general equation (4.65). The easiest way to check that they are correct is just to calculate the $\mathbf{E}$ and $\mathbf{B}$ fields that follow from these terms. This yields $\mathbf{E} = 0$ and $\mathbf{B} = \frac{g}{4\pi r_{n1}(t)}$, which are the correct fields for a moving monopole to lowest order in the velocity.

[9]Note that we consider terms like qv and q^2 to be of order v^2

Now that we have a low velocity expansion for the electromagnetic fields of a moving BPS-dyon, let us calculate one for the scalar field. The easiest way to do this is just to Lorentz transform the scalar field of a static dyon. To turn the static dyon into a dyon that moves with velocity $\mathbf{v}$, we have to apply the Lorentz transformation $L_{\mathbf{v}}$ given by

$$L_{\mathbf{v}}: \begin{array}{rcl} t & \mapsto & \gamma_v(t - \mathbf{v}\cdot\mathbf{r}) \\ \mathbf{r} & \mapsto & \mathbf{r} - \gamma_v \mathbf{v}t - \frac{(\gamma_v - 1)(\mathbf{v}\cdot\mathbf{r})\mathbf{v}}{v^2} \end{array} \tag{4.67}$$

Now let $\mathbf{x}$ be the point in space where we measure ϕ and let $\mathbf{r}$ be the location of the dyon which is the source for ϕ at the time of measurement. Because ϕ is a Lorentz scalar field, we have $\phi'(L_{\mathbf{v}}(\mathbf{x}), L_{\mathbf{v}}(\mathbf{r})) = \phi(\mathbf{x}, \mathbf{r})$, where ϕ' is the Lorentz transform of the field ϕ, i.e. where ϕ' is the scalar field of the moving dyon. This implies that we have

$$\phi'(\mathbf{r}) = \phi(L_{\mathbf{v}}^{-1}(\mathbf{x}), \mathbf{L}_{\mathbf{v}}^{-1}(\mathbf{r})) = \phi(\mathbf{L}_{-\mathbf{v}}(\mathbf{x}), \mathbf{L}_{-\mathbf{v}}(\mathbf{r})) = \frac{\mathbf{g_s}}{|\mathbf{L}_{-\mathbf{v}}(\mathbf{x}) - \mathbf{L}_{-\mathbf{v}}(\mathbf{r})|}$$

From this expression, we can easily extract the first order terms in q, v and we see that the scalar field due to dyon 1 at the location of dyon n will be given by

$$\phi = \frac{g}{4\pi r_{n1}}\left(1 + \frac{q_1^2}{2g^2} - \frac{(\mathbf{v}_1\cdot\hat{\mathbf{r}}_{n1})^2}{2}\right) + O(v^3) \tag{4.68}$$

Of course this equation is only valid when the effect of the acceleration of dyon 1 is negligible to this order in v, but we had already made this assumption in our derivation of the electromagnetic and dual potentials, so there is no further loss of generality here. We can now substitute the above equation and the equations (4.66) for the potentials into the Lagrangian (4.60) and, after some algebra, this gives us the following order v^2 effective Lagrangian for the motion of dyon n in the fields of dyon 1

$$L_n = -m_n + \frac{1}{2}m_n v_n^2 - \frac{g^2(\mathbf{v}_n - \mathbf{v}_1)^2}{8\pi r_{n1}} - \frac{g}{4\pi}(q_n - q_1)(\mathbf{v}_n - \mathbf{v}_1)\cdot\mathbf{w}_{n1} + \frac{(q_n - q_1)^2}{8\pi r_{n1}}$$

Of course, we really want to take the fields of all dyons into account and not just those of dyon 1. In classical electromagnetism, this would be easy enough, because there, we would just be able to add the fields of the individual dyons. Here, we can not expect things to be so simple. The Yang-Mills-Higgs equations that underly our effective theory are nonlinear and therefore we must expect the dyon fields to have nontrivial (=nonlinear) superposition. However, we may hope that the superposition becomes asymptotically linear at large distance from all the dyons. As far as I know, there is no direct conclusive evidence for this in the general situation, but it has been rigorously proved by Taubes [27] at least for configurations of static monopoles. Also, the work of Bak, Lee and Lee [71] shows that the electromagnetic and Higgs field strengths of a single slowly moving dyon may be superposed linearly on weak constant field strengths. This in particular gives much hope that we will have linear superposition for the fields of well separated dyons, because if dyons are far apart, then, at the position of one dyon, the field strengths of the other dyons should be weak and slowly varying (approximately constant). If we assume that linear superposition holds for our well separated dyons, we arrive at the following Lagrangian for dyon n when it moves in the fields of all the other dyons.

$$\begin{aligned} L_n &= \frac{1}{2}mv_n^2 - \frac{g^2}{8\pi}\sum_{i=1}^{n-1}\frac{(\mathbf{v}_n - \mathbf{v}_i)^2}{r_{ni}} \\ &\quad -\frac{g}{4\pi}\sum_{i=1}^{n-1}(q_n - q_i)(\mathbf{v}_n - \mathbf{v}_i)\cdot\mathbf{w}_{n1} + \frac{1}{8\pi}\sum_{i=1}^{n-1}\frac{(q_n - q_i)^2}{r_{ni}} \end{aligned}$$

In this expression, we have dropped the constant $-m_n$, which is allowed, since this does not change the equations of motion. We have also replaced m_n with $m := vg$ in the term $\frac{1}{2}mv_n^2$, which is allowed, since it makes no difference to order q^2. We see from the Lagrangian above that the interactions between dyon n and each of the other dyons are perfectly symmetric (note that we had chosen $\mathbf{w}$ so that $\mathbf{w}_{ij} = \mathbf{w}_{ji}$). It follows that, when we symmetrise the Lagrangian above (so that we get interaction terms for each pair of dyons) we will get a Lagrangian that gives the correct equations of motion for each of the dyons. This Lagrangian is

$$\begin{aligned} L &= \sum_{i=1}^{n}\frac{1}{2}mv_i^2 - \frac{g^2}{8\pi}\sum_{1\le i<j\le n}^{n}\frac{(\mathbf{v}_i - \mathbf{v}_j)^2}{r_{ij}} \\ &\quad -\frac{g}{4\pi}\sum_{1\le i<j\le n}^{n}(q_i - q_j)(\mathbf{v}_i - \mathbf{v}_j)\cdot\mathbf{w}_{ij} + \frac{1}{8\pi}\sum_{1\le i,j\le n}^{n}\frac{(q_i - q_j)^2}{r_{ij}}. \end{aligned} \tag{4.69}$$

Though this Lagrangian is not Lorentz invariant, it is invariant under "Galilei boosts," spatial rotations and translations, which is all we can hope for in a low velocity approximation.

The momentum π_k conjugate to the position of dyon k is given by

$$\pi_k = m\mathbf{v}_k - \frac{g^2}{4\pi}\sum_{i\neq k}\frac{\mathbf{v}_k - \mathbf{v}_i}{r_{ik}} - \frac{g}{4\pi}\sum_{i\neq k}(q_i - q_k)\mathbf{w}_{ik} \tag{4.70}$$

To write the equations of motion in a nice form, it is also useful to introduce a quantity $\mathbf{p}_k$, which is given by the first two terms on the right hand side of the equation for π_k above. That is,

$$p_k = m\mathbf{v}_k - \frac{g^2}{4\pi}\sum_{i\neq k}\frac{\mathbf{v}_k - \mathbf{v}_i}{r_{ik}} \tag{4.71}$$

In terms of the "momentum" p_j, the equation of motion for dyon j is given by

$$\dot{\mathbf{p}}_j = \frac{1}{8\pi}\sum_{i\neq j}\frac{(g^2(\mathbf{v}_j - \mathbf{v}_i)^2 - (q_j - q_i)^2)\mathbf{r}_{ji}}{r_{ji}^3} + \frac{g}{4\pi}\sum_{i\neq j}\frac{(q_j - q_i)(\mathbf{v}_j - \mathbf{v}_i)\times\mathbf{r}_{ji}}{r_{ji}^3} \tag{4.72}$$

We see immediately that these equations are solved by all configurations in which all the dyons have equal velocity $\mathbf{v}$ and electric charge q. These solutions correspond to (boosted) static BPS multi-dyons and the fact that we find solutions that correspond to these dyons is a first indication that the approximation we have made works.

A second, more important check on the validity of our approach can be done by comparing the special case where we have only two dyons with the results we found for the dynamics of two dyons in section 4.4. This yields a very interesting result indeed. We see that the equation for the relative motion of the dyons, which we get by subtracting the equations of motion for the individual dyons given above, is exactly the same as the equation (4.53) for geodesic motion in the Taub-NUT metric. Thus, for the case of two monopoles, we have done as well as we could ever hope to do, considering that we neglected all short range interactions.

In the general case of n dyons, we see that the total force that works on each dyon is just the sum of the Taub-NUT forces from the other dyons. That is, between each pair of dyons, there is a repulsive force proportional to the square of the relative speed, an attractive Coulomb force proportional to the square of the relative electric charge and a Lorentz force which also involves only the relative electric charge. Note that the word force as I use it here refers to the time derivative of the quantity $\mathbf{p}$, which is not quite the same as the linear momentum. However, from the definition (4.71) of the $\mathbf{p}_k$, we see immediately that these quantities do approach the linear momenta when the dyon separations $\mathbf{r}_{ij}$ become large.

Of course, we would like to solve the equations of motion, but since the case of one dyon is trivial and since we have already treated the case of two dyons in section 4.4, we are left with the case of three or more dyons, which is very complicated. In fact, it would be quite a surprise if a general solution to the equations of motion could be found for any number of dyons greater than two, because the problem of three particles interacting only through attractive Coulomb forces already does not have a general solution. It would be interesting to study certain specific phenomena which should occur in the dynamics of three or more dyons, like the scattering of dyons and Taub-NUT bound states, but unfortunately we do not have time to go into this.

What we will do is give the conserved quantities that result from the Galilei invariance of the Lagrangian. The invariance under spatial translations is reflected in the fact that the velocity of the centre of mass is conserved. We may show this explicitly in the Lagrangian (4.69) by writing

$$\sum_{i=1}^{n}\frac{1}{2}m\mathbf{v}_i^2 = \frac{m}{2n}\left(\sum_{i=1}^{n}\mathbf{v}_i\right)^2 + \frac{1}{n}\sum_{1\leq i<j\leq n}\frac{1}{2}m(\mathbf{v}_j-\mathbf{v}_i)^2, \tag{4.73}$$

from which we see that the centre of mass position is a cyclic coordinate.

The conserved quantity that arises as a result of the invariance under spatial rotations is a kind of generalised angular momentum, given by

$$\mathbf{J} = 4\pi\sum_{i=1}^{n}\mathbf{x}_i\times\mathbf{v}_j - \sum_{1\leq i,j\leq n}\frac{4\pi(\mathbf{x}_i-\mathbf{x}_j)\times(\mathbf{v}_i-\mathbf{v}_j)+(\mathbf{x}_i-\mathbf{x}_j)(q_i-q_j)}{r_{ij}} \tag{4.74}$$

This expression generalises the angular momentum for the case of two particles, which we gave in equation (4.56)

Let us now try to establish contact between the long distance approach we have described so far and the geodesic approximation. For the two particle case, we have already done this, because we saw that the equations of motion for the dyon positions in the present approach are the same as those obtained previously for geodesic motion in the Taub-NUT metric. There are only two differences between the two dyon system we have here and the Taub-NUT system. The first is that here, the electric charges of the dyons are constant parameters, whereas in the Taub-NUT case, they are constants of the motion, associated with the "internal coordinates" of the moduli space (i.e. the coordinates which are not associated with the locations of the dyons, but rather with their internal degrees of freedom). The second is, that in the geodesic approximation, dyons are indistinguishable particles, but in the point particle approach they are not. The second difference can be removed by identifying point particle configurations which differ only by a permutation of the particles. Removing the first takes a bit more effort. We will somehow have to interpret all the charges q_i as constants of the motion associated with internal coordinates. To do this, we first have to introduce the required internal coordinates. We will take these to be n phase angles $\theta_1, \dots, \theta_n$, one for each dyon. When we say that two field configurations differ only by a permutation of dyons, we mean that not only the dyons' positions, but also their phase angles are identified. The new configuration space we have constructed now is thus an n-torus bundle over the configuration space of the point particle system, modulo the action of the permutation group S_n on the particle indices of the coordinates $(\mathbf{x}_i, \theta_i)$ [10]

Let us now write down the Lagrangian for geodesic motion on this bundle. In general, this takes the form

$$\mathcal{L} = \frac{1}{2}a_{ij}^{lm}v_i^l v_j^m + \dot{\theta}_i c_{ij}^l v_j^l + \frac{1}{2}h_{ij}\dot{\theta}_i\dot{\theta}_j \tag{4.75}$$

Here, the bottom index in v_i^l denotes the particle number and the top index is a component index. We take a and h to be symmetric in the particle indices (this is possible without loss of generality). We want the momenta conjugate to the angles θ_i to be constants of the motion. Therefore, we will assume that the metric is T^n invariant, or in other words: that the coefficients a_{ij}^{lm}, c_{ij}^l and h_{ij} above do not depend on the θ_i. The θ_i are then cyclic coordinates and the associated conserved quantities are given by

$$p_{\theta_i} = h_{ij}\dot{\theta}_j + c_{ij}^l v_j^l \tag{4.76}$$

Since we want to associate these with the electric charges of the dyons, we will write

$$q_i = \kappa p_{\theta_i}$$

where κ is some constant which is to be determined later.

[10] strictly speaking, we should also mod out the action of a cyclic group $\mathbb{Z}_n$ to come as close to the moduli space $\mathcal{M}_n$ as possible, but we will not do this and work in the connected n-fold cover of $\mathcal{M}_n$.

What we want to do now is choose the coefficients in the Lagrangian (4.75) so that the equations of motion for the dyons' positions that follow from this Lagrangian are exactly the same as those given in equation (4.72). We could try to do this by direct comparison of the equations of motion, but there is a more elegant way, which also leads to nicer formulae later on. First, note that we can rewrite the Lagrangian above as follows

$$\mathcal{L} = \frac{1}{2}G_{ij}^{lm}v_i^l v_j^m + \frac{1}{2}h_{ij}(\dot{\theta}_i + W_{ik}^m v_k^m)(\dot{\theta}_j + W_{jl}^n v_l^n). \tag{4.77}$$

Here, we have defined G_{ij}^{kl} and W_{jl}^m through

$$\begin{aligned} W_{jl}^m &:= h_{jk}^{-1}c_{kl}^m \\ G_{ij}^{kl} &:= a_{ij}^{kl} - h_{mn}W_{mi}^k W_{nj}^l \end{aligned}$$

With these definitions, the last term in the Lagrangian is equal to $\frac{1}{2\kappa^2}h_{ij}^{-1}q_i q_j$. It is tempting to just fill this in and then to equate the Lagrangian above with the Lagrangian (4.69) for the point particles. However, this procedure does not give us the right metric, because in substituting for the last term in $\mathcal{L}$, we change the equations of motion for the dyons' positions. But all is not lost; we can replace the Lagrangian above with the following point particle Lagrangian without changing the equations of motion for the positions of the dyons:

$$\mathcal{L}_{eff} = \frac{1}{2}G_{ij}^{lm}v_i^l v_j^m + \frac{1}{\kappa}q_i W_{ij}^k v_j^k - \frac{1}{2\kappa^2}h_{ij}^{-1}q_i q_j \tag{4.78}$$

When we equate this with the Lagrangian (4.69), we see that we have to have $G_{ij}^{kl} = g_{ij}\delta^{kl}$, where g_{ij} is given by

$$\begin{aligned} g_{jj} &= m - \frac{g^2}{4\pi}\sum_{i\neq j}\frac{1}{r_{ij}} \quad \text{(no sum over j)} \\ g_{ij} &= \frac{g^2}{4\pi r_{ij}} \quad \text{(i} \neq \text{j)} \end{aligned} \tag{4.79}$$

The $\mathbf{W}_{ij}$ must be given by

$$\begin{aligned} \mathbf{W}_{jj} &= -\frac{\kappa g}{4\pi}\sum_{i\neq j}\mathbf{w}_{ij} \quad \text{(no sum over j)} \\ \mathbf{W}_{ij} &= \frac{\kappa g}{4\pi}\mathbf{w}_{ij} \quad \text{(i} \neq \text{j)} \end{aligned} \tag{4.80}$$

The matrix h^{-1} presents us with a problem. When we calculate it in the same way as g_{ij} and the $\mathbf{W}_{ij}$, we find

$$h_{ij}^{-1} = \frac{\kappa^2}{g^2}g_{ij} - m\delta_{ij}$$

But this matrix is not invertible. We could have seen this directly from the Lagrangian (4.69). The last term in this Lagrangian is just the quadric in "**q**-space" which is associated with the matrix h and we see that this has a flat direction along the line where all the electric charges q_i are equal. In other words: the vector $(1, 1, \ldots, 1)$ is a non trivial element of the kernel of h. One may also verify this explicitly form the expression for h given above. Fortunately, we can redefine h without changing the equations of motion. In fact, we can add an arbitrary constant matrix to h without changing the equations of motion that follow from the effective Lagrangian $\mathcal{L}_{eff}$. A logical choice is to add the matrix $m\delta_{ij}$. This gives

$$h_{ij} = \frac{\kappa^2}{g^2} g_{ij} \tag{4.81}$$

and one may easily check that the matrix g is invertible (the associated quadric is positive definite), so it follows that h is now invertible too.

The last thing we have to do before we can write down the metric is to determine the constant κ. Suppose we let the angles θ_i take values in the usual interval $[0, 2\pi]$. In that case, the conjugate momenta p_{θ_i} will take integer values when the theory is quantised. It follows that the electric charges q_i will be quantised in units of κ. Of course, we want the electric charges to be quantised in units of the elementary charge e. This means we have to choose $\kappa = e = \frac{4\pi}{g}$

We can now write down the metric. To facilitate comparison with the formulae we gave previously for the Taub-NUT metric and to obtain nice looking formulae, we will do this in the units of section 4.4. In these units, the fundamental magnetic charge g and the mass m are both equal to 4π. If we substitute this and also remove an overall factor of 4π, we have

$$ds^2 = g_{ij} d\vec{r}_i \cdot d\vec{r}_j + g_{ij}^{-1}(d\theta_i + \mathbf{W}_{ik} \cdot d\mathbf{r}_k)(d\theta_j + \mathbf{W}_{jl} \cdot d\mathbf{r}_l) \tag{4.82}$$

where, in these conventions, we have

$$\begin{array}{rclcrcl} g_{jj} & = & 1 - \sum_{i \neq j} \frac{1}{r_{ij}} & \quad & \mathbf{W}_{jj} & = & -\sum_{i \neq j} \mathbf{w}_{ij} \\ g_{ij} & = & \frac{1}{r_{ij}} & & \mathbf{W}_{ij} & = & \mathbf{w}_{ij} \end{array} \tag{4.83}$$

For the case of two dyons, this is just the sum of the Taub-NUT metric given in equation (4.45) and the flat metric on $\mathbb{R}^3 \times S^1$, of which we know that it approximates the true metric exponentially well. For an arbitrary number of dyons, we see that the metric is asymptotically flat, so that the monopoles are asymptotically free particles. Gibbons and Manton also showed that it is hyperkähler. Finally, Bielawski [69] showed that, for arbitrary numbers of dyons the metric given here approximates the true metric exponentially well at large dyon separations. This result rigorously establishes contact between the point particle approach we have treated in this section and the geodesic approximation. In particular, it proves indirectly that when the dyons are well separated, move at low speeds and have small electric charges and positive unit magnetic charge, the point particle approach yields mathematically correct results.

4.6 Beyond the geodesic approximation

Up to now, we have studied the dynamics of BPS-dyons in the framework of the geodesic approximation. However, there at least two kinds of interesting phenomena which we cannot study in this approximation. Explicitly:

1. We cannot study scattering and bound states of dyons which have unlike magnetic charges, because there are no static solutions with dyons of unlike magnetic charges.

2. We cannot study processes that involve electromagnetic or Higgs radiation, because the degrees of freedom associated with radiation are all contained in the non zero modes of the dyons, which are ignored in the geodesic approximation.

To study these phenomena, we will have to think of a description of BPS-dyons which captures a different part of the physics than the geodesic approximation. To find this, we don't have to look far; the point particle model we introduced in the previous section can easily be extended to include dyons of opposite magnetic charges. In fact, the only reason we did not do this straight away was that we wanted to establish contact between the point particle and the geodesic approximation as efficiently as possible.

Still, even with the extended point particle model, we cannot study radiative phenomena. The reason for this is, that in this approach the fields $\vec{A}$ and ϕ do not have their own dynamics; they depend only on the positions of the dyons. Recently, this shortcoming has been at least partially remedied by Bak, Lee and Lee [71]. Based on a study of the full Yang-Mills-Higgs field equations, they have proposed a modification of the point particle model in which the dyons are still point particles, but in which the fields through which they interact do have their own dynamics. Unfortunately, I do not have enough time to treat Bak, Lee and Lee's model in detail, but I will give their most important results here and describe briefly how they were obtained.

First, Bak, Lee and Lee study the deformation and acceleration of a dyon due to weak electric, magnetic and Higgs field strengths which are asymptotically constant at spatial infinity. To this end, they use the ansatz

$$\begin{aligned}
\Phi^a(\mathbf{x},t) &= \bar{\Phi}(\mathbf{x}-\mathbf{r}(t);\theta) + \Pi^a(\mathbf{x}-\mathbf{r}(t);\theta) \\
A_i^a(\mathbf{x},t) &= \bar{A}_i^a(\mathbf{x}-\mathbf{r}(t);\theta) - ta_i\bar{A}_0^a(\mathbf{x}-\mathbf{r}(t);\theta) + \alpha_i^a(\mathbf{x}-\mathbf{r}(t);\theta) \\
A_0^a(\mathbf{x},t) &= \bar{A}_0(\mathbf{x}-\mathbf{r}(t);\theta) - ta_i\bar{A}_i^a(\mathbf{x}-\mathbf{r}(t);\theta) + \alpha_0^a(\mathbf{x}-\mathbf{r}(t);\theta)
\end{aligned} \tag{4.84}$$

Here, $\mathbf{r}(t)$ is the position of the centre of the dyon at time t. This is assumed to be in uniformly accelerated motion with acceleration $\mathbf{a}$. If we look at the dyon in the frame in which it is instantaneously at rest in the space origin at $t=0$, we may thus write $\mathbf{r}(t) = \frac{1}{2}\mathbf{a}t^2$. The fields with the bars are the fields of a Julia-Zee dyon as given in equation (3.47); the angle θ is just the angle which gives the electric charge of the dyon ($q = g\tan(\theta)$). We see that the terms with the bars above represent the fields of a BPS-dyon that moves non-relativistically with velocity $\mathbf{a}t$. The fields α and Π are supposed to be $O(a)$. They are there to account for deformations in the dyon's fields due to interactions with the electric, magnetic and Higgs field strengths. In fact, these field strengths are implemented through the fields α and Π, by means of the boundary conditions

$$r \to \infty: \quad \mathbf{B}_i \to (\mathbf{B}_0)_i \;\; \mathbf{E}_i \to (\mathbf{E}_0)_i \;\; \mathbf{H}_i \to (\mathbf{H}_0)_i. \tag{4.85}$$

Note that the deformations α and Π are assumed to depend on time only through the time dependence of the position $\mathbf{r}$ of the dyon centre. It is not a priori clear that solutions to the full equations of motion which have this property exist. However, when we substitute the ansatz (4.84) into the the equations of motion, we will see that the resulting equation for α and Π does not involve t itself, but indeed involves only $\mathbf{r}$. Before we write this equation down, it is convenient to introduce the notation $\tilde{A}$ for the part of the A-field that includes the deformations due to interaction, but excludes those due to motion. Explicitly, we define

$$\tilde{A}^a_\mu := \bar{A}^a_\mu + \alpha^a_\mu$$

Correspondingly, we define $\tilde{D}$ and $\tilde{F}$ as the covariant derivative and field strength derived from the potential $\tilde{A}$.

If we now substitute the ansatz (4.84) above into the equations of motion (3.3) for the Georgi-Glashow model, we find the equations [73] [11]

$$\begin{aligned} \left(\tilde{D}_j + a_j\right) \tilde{F}^{ij} + e\left[\tilde{A}_0, \tilde{D}_i \tilde{A}_0\right] &= -e\left[\tilde{D}_i\tilde{\Phi}, \tilde{\Phi}\right] \\ \tilde{D}_i\left(\tilde{D}_i + a_i\right)\tilde{A}_0 + e^2\left[[\tilde{A}_0, \tilde{\Phi}], \tilde{\Phi}\right] &= 0 \\ \tilde{D}_i\left(\tilde{D}_i + a_i\right)\tilde{\Phi} + e^2\left[[\tilde{A}_0, \tilde{\Phi}], \tilde{A}_0\right] &= 0. \end{aligned} \tag{4.86}$$

The first two of these equations come from the equations of motion for A and the last one from that that for Φ. The above equations are rather difficult to work with, but it turns out that one can find a less complicated set of equations whose solutions all satisfy the above equations to $O(a)$. This set of equations is

$$\begin{aligned} \tilde{B}^a_i &= \mp(\tilde{D}_i + a_i)(\cos(\theta)\Phi \pm \tan(\theta)\alpha_0) \\ \bar{D}_i\bar{D}_i\alpha_0 &= -e^2\cos^2(\theta)\left[[\alpha_0, \bar{\Phi}], \bar{\Phi}\right] \end{aligned} \tag{4.87}$$

The situation we have here is somewhat analogous to the situation we had in chapter 3, where we studied the Bogomol'nyi equations (3.35) in order to find solutions to the Yang-Mills-Higgs equations (3.3). In fact, the analogy is closer than it seems at first, because when we look at the above equations in the special case for which the acceleration $\mathbf{a}$ and the deformations α and Π are zero, we see that they reduce to the single equation $B_i = \mp\cos(\theta)D_i\Phi$, which is one of the Bogomol'nyi equations for a dyon. Note also that one needs the fact that the unperturbed fields satisfy the Bogomol'nyi equations to prove that the equations (4.87) imply the equations (4.86).

The first of the equations (4.87) gives us the non-Abelian magnetic field. At infinity, we may obtain the Abelian magnetic field from this by projecting on the non-perturbed

[11] Note that we are working in the Prasad-Sommerfield limit, i.e. the parameter λ in the equations of motion (3.3) is taken to be zero

Higgs field $\bar{\Phi}$. Similarly, we may obtain the Abelian electric field at infinity from the following expression for the non Abelian electric field:

$$E_i = -ta_j \bar{F}_{ij} + \left(\tilde{D}_i + a_i\right) \tilde{A}_0 \tag{4.88}$$

By imposing the boundary condition (4.85), one may now determine the acceleration $\mathbf{a}$ from (4.88) and the first of the equations (4.87), if one makes the additional assumption that the deformation function α_0 is asymptotically of the following form:

$$\alpha_0^a(\mathbf{x}-\mathbf{r};\theta) = \frac{\cos(\theta)\mathbf{C}\cdot(\mathbf{x}-\mathbf{r})}{|\mathbf{x}-\mathbf{r}|}(\mathbf{x}-\mathbf{r}) \tag{4.89}$$

where $\mathbf{C}$ is a constant vector. Substituting this in (4.88) and the first of the equations (4.87) and applying the boundary conditions, we obtain two linear equations for $\mathbf{a}$ and $\mathbf{C}$ and these finally yield the result

$$\begin{aligned} \mathbf{C} &= \mp\frac{1}{v}\left[\sin(\theta)\mathbf{B}_0 - \cos(\theta)\mathbf{E}_0\right] \\ \mathbf{a} &= \mp\frac{1}{v}\left[\cos(\theta)\mathbf{B}_0 - \sin(\theta)\mathbf{E}_0 \mp \mathbf{H}_0\right] \end{aligned} \tag{4.90}$$

From the expression for $\mathbf{a}$, we get the force that works on the dyon in its instantaneous rest frame. Using $\cos(\theta) = \frac{g}{g_s}$, $\sin(\theta) = \frac{q}{g_s}$ and $M = vg_s$, we see that this is given by

$$M\mathbf{a} = g\mathbf{B}_0 + q\mathbf{E}_0 + g_s\mathbf{H}_0 \tag{4.91}$$

Remember that we have already used this result and its generalisation (4.62) to an arbitrary inertial frame as input for the point particle approximation we described in the previous section. However, one can do something extra in the current approach. Now that $\mathbf{a}$ and $\mathbf{C}$ are determined, it is actually possible to find an explicit solution to the equations (4.87) which is everywhere regular and can be expressed in terms of elementary functions [71]. We will not give this solution here, because it is rather a large and complicated expression, but we will give the asymptotic electromagnetic and Higgs field strengths that can be derived from it. In the frame in which the dyon is instantaneously at rest in the origin at $t = 0$, these are given by

$$\mathbf{B}(\mathbf{x},t) \sim \mathbf{B}_0 + \frac{g}{4\pi}\frac{\hat{\mathbf{R}}-\mathbf{v}}{(1-\hat{\mathbf{R}}\cdot\mathbf{v})^3R^2} - \frac{q}{4\pi}\frac{\hat{\mathbf{R}}\times\mathbf{v}}{R^2} + \frac{g}{4\pi}\frac{\mathbf{R}\times(\mathbf{R}\times\mathbf{a})}{R^3} - \frac{q}{4\pi}\frac{\mathbf{R}\times\mathbf{a}}{R^2} \tag{4.92}$$

$$\mathbf{E}(\mathbf{x},t) \sim \mathbf{E}_0 + \frac{q}{4\pi}\frac{\hat{\mathbf{R}}-\mathbf{v}}{(1-\hat{\mathbf{R}}\cdot\mathbf{v})^3R^2} + \frac{g}{4\pi}\frac{\hat{\mathbf{R}}\times\mathbf{v}}{R^2} + \frac{q}{4\pi}\frac{\mathbf{R}\times(\mathbf{R}\times\mathbf{a})}{R^3} + \frac{g}{4\pi}\frac{\mathbf{R}\times\mathbf{a}}{R^2} \tag{4.93}$$

$$\mathbf{H}(\mathbf{x},t) \sim \mathbf{H}_0 + \frac{g_s}{4\pi}\frac{\hat{\mathbf{R}}-\mathbf{v}}{(1-\hat{\mathbf{R}}\cdot\mathbf{v})^3R^2} + \frac{g_s}{4\pi}\frac{(\mathbf{R}\cdot\mathbf{a})\hat{\mathbf{R}}}{R^2} \tag{4.94}$$

$$H^0(\mathbf{x},t) \sim \mathbf{H}_0 + \frac{g_s}{4\pi}\frac{\hat{\mathbf{R}}\cdot\mathbf{v}}{(1-\hat{\mathbf{R}}\cdot\mathbf{v})^3R^2} + \frac{g_s}{4\pi}\frac{\mathbf{R}\cdot\mathbf{a}}{R^2} \tag{4.95}$$

Here we have defined $\mathbf{R} := \mathbf{x} - \mathbf{r}$. Since the dyon is in uniformly accelerated motion, we have $\mathbf{R}(t) = \mathbf{x} - \frac{1}{2}\mathbf{a}t^2$. Similarly, we have $\mathbf{v}(t) = \mathbf{a}t$. All the expressions on the right hand side of the formulae above should be evaluated at the retarded time t_0, which can be determined from the condition that $\mathbf{r}(t_0)$ has to lie on the backward light cone from $\mathbf{x}$. We see from the equations above that we have not only the usual near-zone fields for a dyon (the terms of order $\frac{1}{r^2}$), but also radiation fields. These are given by the terms which involve $\mathbf{a}$ explicitly. We can see that these terms represent radiation from the fact that they are of order $\frac{1}{r}$. In the point particle approach of the previous section, we did not have these radiation fields, because we assumed the acceleration to be negligible. Note that the electromagnetic fields (both near-zone and radiation fields) are of exactly the form we would get in duality invariant Maxwell theory. In fact, one sees immediately from the formulae above that all the field strengths are consistent with electric-magnetic duality. The same is of course true for the force law (4.91) above.

From the above, it is clear that electromagnetic and Higgs radiation are produced whenever a dyon is accelerated. Once such radiation has been emitted from a dyon, it can scatter with another dyon which lies in its path. Such scattering could influence dyon dynamics and could also be important for the detection of dyons. Thus, it makes sense to study processes in which light or Higgs radiation scatters off a dyon. Bak, Lee and Lee have done this, making use of a perturbative scheme analogous to the one we have just described.

This time, the dyon is supposed to be in oscillatory movement whose magnitude and direction are given by a vector $\mathbf{a}$, due to an incoming plane wave of electromagnetic and Higgs radiation which has wave vector $\mathbf{k}$ and angular frequency ω. These ingredients lead to the following ansatz for the fields

$$\begin{aligned} \Phi^a(\mathbf{x},t) &= \bar{\Phi}^a(\mathbf{x}-\mathbf{r}(t);\theta) + \mathrm{Re}\left\{\Pi^a(\mathbf{x}-\mathbf{r}(t);\theta)e^{-i\omega t}\right\} \\ A^a_\mu(\mathbf{x},t) &= \bar{A}^a_\mu(\mathbf{x}-\mathbf{r}(t);\theta) + \mathrm{Re}\left\{\alpha^a_\mu(\mathbf{x}-\mathbf{r}(t);\theta)e^{-i\omega t}\right\}, \end{aligned} \tag{4.96}$$

where the functions α and Π are again deformations of the dyon's fields which are supposed to be $O(a)$. Substituting the above ansatz into the equations of motion yields complicated equations for which one can again find solutions relatively easily by the use of an analogue of the Bogomol'nyi equation. After a lot of algebra, one then finds the following asymptotic Electromagnetic and Higgs radiation fields:

$$\begin{aligned} \mathbf{B} &= \mathrm{Re}\Big\{\frac{i\omega^2}{g_s}[g\hat{\mathbf{k}}\times(\hat{\mathbf{k}}\times\mathbf{a}) - q\hat{\mathbf{k}}\times\mathbf{a}]ve^{i\mathbf{k}\cdot\mathbf{x}-i\omega t} + 4\pi i\omega[g\hat{\mathbf{x}}\times(\hat{\mathbf{x}}\times\mathbf{a}) - q\hat{\mathbf{x}}\times\mathbf{a}]\frac{e^{i\mathbf{k}\cdot\mathbf{x}-i\omega t}}{x}\Big\} \\ \mathbf{E} &= \mathrm{Re}\Big\{\frac{i\omega^2}{g_s}[q\hat{\mathbf{k}}\times(\hat{\mathbf{k}}\times\mathbf{a}) + g\hat{\mathbf{k}}\times\mathbf{a}]ve^{i\mathbf{k}\cdot\mathbf{x}-i\omega t} + 4\pi i\omega[q\hat{\mathbf{x}}\times(\hat{\mathbf{x}}\times\mathbf{a}) + g\hat{\mathbf{x}}\times\mathbf{a}]\frac{e^{i\mathbf{k}\cdot\mathbf{x}-i\omega t}}{x}\Big\} \\ \mathbf{H} &= \mathrm{Re}\Big\{-i\omega^2(\mathbf{a}\cdot\hat{\mathbf{k}})\hat{\mathbf{k}}ve^{i\mathbf{k}\cdot\mathbf{x}-i\omega t} + \frac{4\pi i\omega^2}{g_s}(\mathbf{a}\cdot\hat{\mathbf{x}})\hat{\mathbf{x}}\frac{ve^{i\mathbf{k}\cdot\mathbf{x}-i\omega t}}{x}\Big\} \\ H_0 &= \mathrm{Re}\Big\{-i\omega^2(\mathbf{a}\cdot\hat{\mathbf{k}})ve^{i\mathbf{k}\cdot\mathbf{x}-i\omega t} + \frac{4\pi i\omega^2}{g_s}(\mathbf{a}\cdot\hat{\mathbf{x}})\frac{ve^{i\mathbf{k}\cdot\mathbf{x}-i\omega t}}{x}\Big\} \end{aligned} \tag{4.97}$$

In these expressions, we can clearly distinguish terms which represent the incoming plane wave and terms which represent an outgoing spherical wave. On comparing the above expressions with the radiation fields for a uniformly accelerated dyon given in the equations (4.92), we see that the radiation that the dyon emits due to electromagnetic and Higgs scattering is just the radiation it should emit due to the oscillatory motion induced by the incoming wave.

From the asymptotic fields given above, one may also derive differential cross sections for the scattering of electromagnetic and Higgs waves on dyons. These are given by

$$\begin{array}{rclrcl} \left(\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right)_{em\to em} & = & \left(\frac{g_s^2}{4\pi M}\right)^2 \sin^2\Phi & \left(\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right)_{em\to Higgs} & = & \left(\frac{g_s^2}{4\pi M}\right)^2 \cos^2\Phi \\ \left(\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right)_{Higgs\to em} & = & \left(\frac{g_s^2}{4\pi M}\right)^2 \sin^2\phi & \left(\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right)_{Higgs\to Higgs} & = & \left(\frac{g_s^2}{4\pi M}\right)^2 \cos^2\phi \end{array} \tag{4.98}$$

Here, Φ is the angle between $\mathbf{x}$ and $g\mathbf{B}_{\mathrm{inc}} + q\mathbf{E}_{\mathrm{inc}}$ (where "inc" stands for incoming) and ϕ is the angle between $\mathbf{r}$ and $\mathbf{H}_{\mathrm{inc}} \sim \hat{\mathbf{k}}$. We see immediately that the cross sections are invariant under electric magnetic duality, because they involve only g_s, which is just the symmetric combination of q and g

Bak, Lee and Lee have found that all the above results, which they obtained from the full theory through the use of classical perturbation theory, could also have been obtained from an effective theory in which the dyons are treated as point particles and the only fields are the (Abelian) electromagnetic fields and an isoscalar Higgs field ϕ, as in the point particle model of the previous section. However, in contrast to the model of the previous section, the fields in Bak, Lee and Lee's model are not just functions of the particles' positions and speeds, but they have their own dynamics and consequently, their own equations of motion. The equations of motion for both the fields and the dyon positions can be conveniently summarised by means of the following effective action:

$$\begin{aligned} S_{eff} \;=\; & \int \mathrm{d}^4x \Big\{ \frac{1}{4}F^{\mu\nu}F_{\mu\nu} - \frac{1}{2}F^{\mu\nu}(\partial_\mu A_\nu - \partial_\nu A_\mu) - \frac{1}{2}\partial_\mu\phi\partial^\mu\phi \\ & + \sum_{i=1}^{n}\Big[-\frac{1}{\gamma(\dot{\mathbf{r}}_i)}(M_i + (g_s)_i\phi)\delta^3(\mathbf{x}-\mathbf{r}_i) \\ & - q_i\delta^3(\mathbf{x}-\mathbf{r}_i)[A_0 - \dot{\mathbf{r}}\cdot\mathbf{A}] - g_i\delta^3(\mathbf{x}-\mathbf{r}_i)[\tilde{A}_0 - \dot{\mathbf{r}}\cdot\tilde{\mathbf{A}}]\Big]\Big\} \end{aligned}$$

Here, the r_i are the positions of the dyons and the q_i and g_i their electric and magnetic charges. $\tilde{A}$ is the dual vector potential we introduced in section 4.5. Up to a gauge transformation, we can define this as a functional of F and hence of A. This functional is non-local, which makes the value of the above action as a starting point for quantum physics rather doubtful. However, this should be no problem for classical physics, as long as we get the right Euler-Lagrange equations.

To obtain the equations of motion from the action, we have to view F and A as independent fields and vary them separately. When we do this, we obtain the generalisation of Maxwell's equations which includes magnetic charges. The charges in these equations are just the electric and magnetic point charges due to the dyons. Varying the field ϕ gives the equation of motion

$$\partial_\mu\partial^\mu\phi = \sum_{i=1}^{n}\frac{(g_s)_n}{\gamma(\dot{\mathbf{r}})}\delta^3(\mathbf{x}-\mathbf{r}_i(t)) \tag{4.99}$$

From the equations of motions for the fields, we may rederive the asymptotic fields of an accelerated dyon given in (4.92). The equations of motion for the dyon positions are just the equations (4.62) we used in the point particle approach of the previous section. We have already shown that these incorporate the force law (4.91) we derived in this section. In fact, all the results of this section and the previous can be recovered using only the equations of motion for the fields and the particles. One can even derive the generalisation of the Gibbons-Manton point particle model to the case where the magnetic charges of the dyons can be equal to plus or minus g and the electric charges are not necessarily very small. In the next section, we will study the dynamics of oppositely charged dyons by means of this generalised point particle model.

4.7 Dynamics of dyons of unlike magnetic charge

In this section, we will study the dynamics of a system of dyons which have magnetic charges $\pm g$. As in section 4.5, we will consider the dyons as point particles with Abelian electromagnetic fields and an isoscalar Higgs field. In the same way as in section 4.5, we may find an $O(v^2)$ effective Lagrangian for this theory. The only essential input we need in the derivation of this effective Lagrangian is an expansion of the electromagnetic and dual potential and the scalar Higgs field of a dyon to the appropriate orders in v. We do not want to use the expansions (4.66) and (4.68) given in section 4.5, because they treat the electric and magnetic charges in an asymmetrical way; the electric charges are assumed to be of order v. Although this is a logical thing to do from the point of view of the geodesic approximation, in which the charges are associated with speeds, we do not a priori want to restrict ourselves to the case of small electric charges in our current long-distance theory, where the charges are just constants. Thus, we have to modify the expansions (4.68) and (4.66). This is most easily done for the expansion of the Higgs field. Using the same derivation as in section 4.5, but keeping terms of order q^3 and higher, we see that the scalar Higgs field of dyon 1 at the location of dyon n is given by

$$\phi = \frac{g_s}{4\pi r_{n1}}\left(1 - \frac{(\mathbf{v}_1\cdot\hat{\mathbf{r}}_{n1})^2}{2}\right) + O(v^3) \tag{4.100}$$

To get the electromagnetic fields, we can again use the low velocity expansion of the Liénard-Wiechert potentials for the part which is due to the presence of electric charge, but to get the part caused by the magnetic charge as well, one has to use the form that the field tensor takes in this case and some formulae which connect A and $\tilde{A}$ to the field tensor F. For the details of the calculation I refer the reader to [71]. Here, I will just give the result, which is exactly what one would intuitively expect, namely:

$$\begin{aligned}
A_0 &= \tfrac{q_1}{4\pi r_{n1}}\left(1+\tfrac{(\mathbf{v}_1\cdot\hat{\mathbf{r}}_{n1})^2}{2}-\tfrac{v^2}{2}\right)-\tfrac{g\mathbf{w}_{n1}\cdot\mathbf{v}_1}{4\pi}+O(v^3) \qquad & \mathbf{A} &= \tfrac{q_1\mathbf{v}_1}{4\pi r_{n1}}-\tfrac{g\mathbf{w}_{n1}}{4\pi}+O(v^2)\\
\tilde{A}_0 &= \tfrac{g}{4\pi r_{n1}}\left(1+\tfrac{(\mathbf{v}_1\cdot\hat{\mathbf{r}}_{n1})^2}{2}-\tfrac{v^2}{2}\right)+\tfrac{q_1\mathbf{w}_{n1}\cdot\mathbf{v}_1}{4\pi}+O(v^3) \qquad & \tilde{\mathbf{A}} &= \tfrac{g\mathbf{v}_1}{4\pi r_{n1}}+\tfrac{q_1\mathbf{w}_{n1}}{4\pi}+O(v^2).
\end{aligned} \tag{4.101}$$

Using these expansions for the potentials and the expansion for the Higgs field, we may now proceed as in section (4.5) to derive an effective Lagrangian for the dyon system which is correct to order v^2. This yields the following expression:

$$
\begin{aligned}
L &= \sum_{i=1}^{n} \frac{1}{2} m_i v_i^2 - \frac{1}{8\pi} \sum_{1\le i<j\le n}^{n} (g_s)_i (g_s)_j \frac{(\mathbf{v}_i - \mathbf{v}_j)^2}{r_{ij}} \\
&\quad - \frac{1}{4\pi} \sum_{1\le i<j\le n}^{n} (q_i g_j - q_j g_i)(\mathbf{v}_i - \mathbf{v}_j)\cdot \mathbf{w}_{ij} + \frac{1}{4\pi} \sum_{1\le i,j\le n}^{n} \frac{(g_s)_i (g_s)_j - q_i q_j - g_i g_j}{r_{ij}} \\
&\quad - \frac{1}{8\pi} \sum_{1\le i,j\le n}^{n} ((g_s)_i (g_s)_j - q_i q_j - g_i g_j) \left\{ \frac{\mathbf{v}_i \cdot \mathbf{v}_j + (\hat{\mathbf{r}}_{ij}\cdot \mathbf{v}_i)(\hat{\mathbf{r}}_{ij}\cdot \mathbf{v}_j)}{r_{ij}} \right\} \qquad (4.102)
\end{aligned}
$$

We see immediately that this slow-motion effective Lagrangian is invariant under electric-magnetic duality. We may see this as an indication that the full theory is indeed also invariant under this duality, as conjectured by Montonen and Olive [6], but it certainly is not proof, because the classical approximation we are using here is very likely to break down at strong coupling and duality in the full theory is dependent on the strong coupling regime as much as on the weak coupling regime. If duality does indeed hold in the full theory then one may speculate that the above Lagrangian could give a description of the slow motion of dyons even at large coupling. However, this too would certainly need further investigation.

Let us now descend once more to the situation in which all the electric charges are small, i.e. $q_i = O(v)$ for all i. In that case, the effective Lagrangian above reduces to

$$
\begin{aligned}
L &= \sum_{i=1}^{n} \frac{1}{2} m v_i^2 - \frac{1}{8\pi} \sum_{1\le i<j\le n}^{n} g^2 \frac{(\mathbf{v}_i - \mathbf{v}_j)^2}{r_{ij}} \\
&\quad - \frac{1}{4\pi} \sum_{1\le i<j\le n}^{n} (q_i g_j - q_j g_i)(\mathbf{v}_i - \mathbf{v}_j)\cdot \mathbf{w}_{ij} + \frac{1}{8\pi} \sum_{1\le i,j\le n}^{n} \frac{2g^2 - 2g_i g_j - (q_i - q_j)^2}{r_{ij}} \\
&\quad - \frac{1}{16\pi} \sum_{1\le i,j\le n}^{n} (2g^2 - 2g_i g_j) \left\{ \frac{\mathbf{v}_i \cdot \mathbf{v}_j + (\hat{\mathbf{r}}_{ij}\cdot \mathbf{v}_i)(\hat{\mathbf{r}}_{ij}\cdot \mathbf{v}_j)}{r_{ij}} \right\} \qquad (4.103)
\end{aligned}
$$

It is easy to check that this reduces to the Lagrangian (4.69) for the Gibbons-Manton point particle model if we set all the magnetic charges equal to the unit charge g. In general, we see that the dyons still have only pairwise interaction and that the forces between two dyons of like charges are the same as those in the Gibbons-Manton model. The interactions we still need to study are those between dyons of opposite magnetic charges. To have a closer look at these, let study a system of two dyons with opposite magnetic charges g and $-g$ and electric charges q_1 and q_2. In this special case, the Lagrangian above reduces to

$$
\begin{aligned}
L &= \frac{1}{2} m(v_1^2 + v_2^2) + \frac{4g^2 - (q_1 - q_2)^2 - g^2(v_1^2 + v_2^2 + 2(\mathbf{r}_{12}\cdot \mathbf{v}_1)(\mathbf{r}_{12}\cdot \mathbf{v}_2))}{8\pi r_{12}} \\
&\quad + \frac{g}{4\pi}(q_1 + q_2)(\mathbf{v}_1 - \mathbf{v}_2)\cdot \mathbf{w}_{12} \qquad (4.104)
\end{aligned}
$$

The momentum π_1 and π_2 conjugate to $\mathbf{r}_1$ and $\mathbf{r}_2$ are now given by

$$
\begin{aligned}
\pi_1 &= m\mathbf{v}_1 - \frac{g^2\mathbf{v}_1 + g^2(\hat{\mathbf{r}}_{12}\cdot\mathbf{v}_2)\hat{\mathbf{r}}_{12}}{4\pi r_{12}} + g(q_1+q_2)\mathbf{w}_{12} \\
\pi_2 &= m\mathbf{v}_2 - \frac{g^2\mathbf{v}_2 + g^2(\hat{\mathbf{r}}_{12}\cdot\mathbf{v}_1)\hat{\mathbf{r}}_{12}}{4\pi r_{12}} - g(q_1+q_2)\mathbf{w}_{12}
\end{aligned}
\tag{4.105}
$$

As in the case of dyons of like magnetic charge, we can define quantities $\mathbf{p}_1$ and $\mathbf{p}_2$ which are parts of the canonical momenta π_1, π_2 and which approximate the linear momenta of the particles when the distance between the particles is large:

$$
\begin{aligned}
\mathbf{p}_1 &= m\mathbf{v}_1 - \frac{g^2\mathbf{v}_1 + g^2(\hat{\mathbf{r}}_{12}\cdot\mathbf{v}_2)\hat{\mathbf{r}}_{12}}{4\pi r_{12}} \\
\mathbf{p}_2 &= m\mathbf{v}_2 - \frac{g^2\mathbf{v}_2 + g^2(\hat{\mathbf{r}}_{12}\cdot\mathbf{v}_1)\hat{\mathbf{r}}_{12}}{4\pi r_{12}}
\end{aligned}
\tag{4.106}
$$

In terms of $\mathbf{p}_1$ and $\mathbf{p}_2$, the equations of motion take the form

$$
\begin{aligned}
\dot{\mathbf{p}}_1 + \dot{\mathbf{p}}_2 &= 0 \\
\dot{\mathbf{p}}_1 - \dot{\mathbf{p}}_2 &= \frac{\mathbf{r}_{12}}{4\pi r_{12}^3}\left\{-4g^2 - (q_1-q_2)^2 + g^2(v_1^2+v_2^2 - 2(\hat{\mathbf{r}}_{12}\cdot\mathbf{v}_1)(\hat{\mathbf{r}}_{12}\cdot\mathbf{v}_2))\right\} \\
&\quad -\frac{g(q_1+q_2)(\mathbf{v}_1-\mathbf{v}_2)\times\hat{\mathbf{r}}_{12}}{4\pi r_{12}^2} - \frac{2g^2}{4\pi r_{12}^3}\left\{(\mathbf{r}_{12}\cdot\mathbf{v}_2)\mathbf{v}_1 + (\mathbf{r}_{12}\cdot\mathbf{v}_1)\mathbf{v}_2\right\}
\end{aligned}
\tag{4.107}
$$

We see that the dominant ($O(v^0)$) force between the dyons is a magnetic Coulomb force, which is twice the Coulomb force we would have expected if the Higgs field had not been there. In other words: The Higgs force doubles the attractive Coulomb force in this case. The magnitude of the other forces that appear is of order v^2. We see that, as in the case of dyons of like charge, there are an attractive electric Coulomb force which is proportional to the square of the relative electric charge and a repulsive velocity dependent force. This last force is not exactly the same as the velocity dependent repulsive force we had before. In particular, it no longer depends only on the relative velocity of the dyons. We also see a Lorentz force, which is now proportional to the sum, rather than the difference of the electric charges. Finally, we see from the last term above that an extra non central velocity dependent force has appeared.

It would be interesting to try and solve the above equations of motion. To me, it seems unlikely that there will be enough constants of the motion to determine the orbits completely, at least I have not been able to find them. However, it should be possible to find good approximate solutions using perturbation theory, starting from the situation in which only the magnetic Coulomb force is present. Unfortunately, I do not have time to do this within this thesis, so for now, the above description of the forces between oppositely charged dyons will have to do. I think it should also be possible to find some special solutions to the equations of motions relatively easily and I am currently investigating this.

4.8 Discussion and Outlook

In this thesis, we have studied two different approaches that can be used to describe the dynamics of BPS dyons in the Georgi-Glashow model. These approaches have different regions of validity. Manton's geodesic approximation works well for dyons of like magnetic charges, as long as the speeds and electric charges of the dyons are small, but it does not describe the dynamics of oppositely charged dyons and it neglects the effect of electromagnetic and Higgs radiation. The effective field theory of Bak, Lee and Lee, which is also rooted in work of Manton's, can describe radiative processes and processes that involve dyons of opposite magnetic charge, but it is only valid if the dyons are far apart. Although these two approaches give us a fairly good grip on dyon dynamics, there are certainly still processes left which cannot be described by either of these models. The most prominent of these are probably the creation and annihilation of dyon-antidyon pairs. Thus, it is clear that further study into the dynamics of BPS dyons in the Georgi-Glashow model is still warranted.

It is also interesting to study BPS dyons in generalisations of the Georgi-Glashow model. These generalisations may go in several directions. One may for example add matter fields to the theory or one may change the gauge group from $SU(2)$ to some larger group. Much research has been done in this last direction and the current state of affairs for, say, the case of $SU(N)$ broken down to $U(n-1)$ by the Higgs effect is the same as that for $SU(2)$, as far as the available approaches are concerned. Concrete results, such as exact metrics on moduli spaces, or even geodesics are not so readily available in general. Much less is known about the situation in which the gauge group is broken down to a non Abelian subgroup. In this case, many interesting new phenomena occur. One may for instance have two dyon configurations that have internal properties that cannot be ascribed to either of the dyons separately.

Another generalisation of the model we have studied here can be obtained by adding supersymmetry. The theories which are obtained this way form a natural setting for the study of electric/magnetic duality. Also, the supersymmetry in these theories guarantees that the BPS bound can be saturated even quantum mechanically. The study of these supersymmetric extensions of Yang-Mills theory has recently become very popular, mostly because of the work of Seiberg and Witten [7].

Finally, it would be very interesting to study non BPS monopoles, that is, to study 'tHooft-Polyakov monopoles without taking the non-physical BPS limit. For very small Higgs potential (and hence very small Higgs mass), one may be able to do this by a slight adaptation of either of the methods we have treated in this thesis, but in the more realistic case where the Higgs particle gets a large mass. the geodesic approximation will certainly break down. The point particle approach may still be able to give some information even in this case, if the asymptotic Higgs field is modified to accommodate for the (now) finite range of the Higgs force, but of course, even if this was done satisfactorily, we would still be left without a theory to describe the short range interactions of the monopoles.

Concluding we may say that there are many interesting problems still to be adressed in the theory of magnetic monopoles. Looking back at past results, I cannot help but look forward to seeing and hopefully helping some of the solutions emerge over the next four years.

Appendix A

Quaternions and such

A.1 Basic definitions and identities

In this thesis, it will sometimes be convenient to make use of the skew field of quaternions $\mathbb{H}$. The classical way to define this is to adjoin to the real numbers the three symbols i, j and k which are taken to commute with real numbers and satisfy the following relations:

$$\begin{aligned} i^2 = j^2 = k^2 = -1 \\ ij = -ji = k, \quad jk = -kj = i, \quad ki = -ik = j \end{aligned} \tag{A.1}$$

A general quaternion will then be of the form

$$x = x_1 + x_2 i + x_3 j + x_4 k.$$

We can define a conjugation $x \mapsto \bar{x}$ by

$$\bar{x} = x_1 - x_2 i - x_3 j - x_4 k$$

We also have a natural norm on the quaternions, which can be conveniently written in terms of the conjugation:

$$|x| = \sqrt{\bar{x}x} = \sqrt{x\bar{x}} = \sqrt{x_1^2 + x_2^2 + x_3^2 + x_4^2}$$

The last identity can be easily verified using the relations (A.1). These relations can also be used to verify that we have

$$\begin{aligned} \bar{xy} &= \bar{y}\bar{x}, \quad \text{and hence} \\ |xy| &= |x||y|. \end{aligned}$$

Furthermore, one can see that if the quaternion x is unequal to zero, then it has a unique inverse:

$$x^{-1} = \frac{\bar{x}}{|x|^2}$$

From the above it follows that the quaternions defined as above are a division algebra or skew field and that the quaternions of unit norm form a subgroup of their multiplicative group. This group of unit quaternions is called $Sp(1)$

A.2 Identification of $Sp(1)$ with $SU(2)$

If we identify the symbol i used above with the complex number i, then we can take the complex numbers to be a sub-algebra of the quaternions. Moreover, we can identify the quaternions with $\mathbb{C}^2$ by writing

$$x = z_1 + z_2 j$$

with z_1 and z_2 uniquely determined by x. Quaternion multiplication can now be written in terms of complex two by two matrices. For example: multiplication on the right by x is given by the matrix

$$\begin{pmatrix} z_1 & z_2 \\ -\bar{z}_2 & \bar{z}_1 \end{pmatrix}$$

This way, we can identify the quaternions with a sub-algebra of the complex two by two matrices. The quaternions i, j and k then correspond to the Pauli matrices σ_1, σ_2 and σ_3, multiplied by a factor of $-i$ (the complex number!). [1] Using this identification, we see that the group Sp(1) of quaternions of unit norm corresponds to the Lie group $SU(2)$ and that the algebra of imaginary quaternions corresponds to its Lie algebra $su(2)$. In fact, the identification of (i, j, k) with the canonical basis $(i\sigma_1, i\sigma_2, i\sigma_3)$ for $su(2)$ gives us a quick way to derive some well known formulae: Let A and B be arbitrary elements of $su(2)$. Using our identification, we can then write

$$\begin{aligned} A &= A_1 i + A_2 j + A_3 k \\ B &= B_1 i + B_2 j + B_3 k. \end{aligned}$$

Using only the relations (A.1), we then easily find

$$[A, B]_i = \sum_{j,k=1}^{3} \epsilon^{ijk} A_j B_k \tag{A.2}$$

and similarly

$$\mathrm{Tr}(AB) \equiv -2\ \mathrm{Re}(AB) = 2 \sum_{i=1}^{3} A_i B_i. \tag{A.3}$$

where the factor of 2 appears as the trace of the 2×2 identity matrix.

[1] It is possible to take this as the definition of the symbols i, j and k, and physicists may prefer to do this.

A.3 Quaternionic linear algebra: $\mathbb{H}^n$ and $Sp(n)$.

The only fundamental difference between a field like $\mathbb{C}$ and the quaternions is the non-commutativity of quaternion multiplication. Therefore, it is possible to use the quaternions as scalars for a linear space, provided that one decides in advance whether to use left or right scalar multiplication. In particular, one has the quaternionic vector spaces $\mathbb{H}^n$ with right multiplication, analogously to the spaces $\mathbb{C}^n$ or $\mathbb{R}^n$. On $\mathbb{H}^n$, one can define an inner product analogously to the standard hermitian inner product on $\mathbb{C}^n$, but using quaternionic in stead of complex conjugation. The linear automorphisms of $\mathbb{H}^n$ that preserve this inner product form a Lie group that is denoted $Sp(n)$. If we identify a linear transformation with its matrix then we can view these groups as matrix groups and in particular, we see that this definition of $Sp(n)$ agrees with our previous definition of $Sp(1)$, just the group of unit quaternions. In general, $Sp(n)$ can be identified with the group of quaternion matrices A whose (quaternionic) conjugate transposed A^* is equal to their inverse (note the analogy with $U(n)$). A complex realisation of $Sp(n)$ can also be obtained for general n, by taking a complex representation of the quaternions, as we did above for the case of Sp(1). The Lie algebra of Sp(n) is the algebra of matrices A for which $A^* = -A$ (again, in analogy with $U(n)$).

A.4 Quaternionic projective spaces

Using the spaces $\mathbb{H}^n$, one can define the n-dimensional quaternionic projective space $P_n(\mathbb{H})$ to be the collection of all one dimensional subspaces of $\mathbb{H}^{n+1}$.

The quaternionic projective spaces can be described in terms of so called homogeneous coordinates. An $n+1$-tuple of homogeneous coordinates for an element x of $P_n(\mathbb{H})$ is just a non-zero element of the one dimensional subspace of $\mathbb{H}^{n+1}$ to which x corresponds. Of course, the same element x of $P_n(\mathbb{H})$ will be represented by many different sets of homogeneous coordinates. To be exact: if we choose to work with scalar multiplication on the right, then two $n+1$-tuples $(x_1, \dots, x_{n+1})$ and $(y_1, \dots, y_{n+1})$ of homogeneous coordinates will represent the same element of $P_n(\mathbb{H})$ if and only if there is a non-zero quaternion z, such that

$$(x_1, \dots, x_{n+1}) = (y_1, \dots, y_{n+1})z$$

From this we see that every element x of $P_1(\mathbb{H})$ can be uniquely represented by a set of homogeneous coordinates of the form $(1, y)$, unless x is the unique point represented by $(0, 1)$. Thus we find that the quaternionic projective line $P_1(\mathbb{H})$ can be seen as a copy of the quaternions, or equivalently $\mathbb{R}^4$, with an extra point added to it. Since $P_1(\mathbb{H})$ is also obviously (path)connected, this means we can identify $P_1(\mathbb{H})$ with S^4. In chapter 2, we use this correspondence to describe S^4 in terms of homogeneous coordinates.

As an aside, I want to say that it is possible to define and work with quaternionic manifolds in a manner completely analogous to real and complex manifolds. In fact, it is not difficult to see that the $P_n(\mathbb{H})$ are quaternionic manifolds of dimension n.

A.5 Structures and tensor product spaces

Sometimes, it is useful to have an action of the quaternions on a complex vector space V. Such an action can be defined by means of a so called *quaternionic structure.* This is a complex anti-linear map J from V into itself, whose square is minus the identity. That is, we have

$$J : V \to V \tag{A.4}$$
$$J(zv) = \bar{z}Jv \tag{A.5}$$
$$J^2 = -I \tag{A.6}$$

If we now define $K = iJ$, then we can easily verify that the maps (left multiplication by) i, J and K satisfy the defining relations (A.1) for the quaternions. Thus, if we define

$$\rho : \mathbb{H} \times V \to V,\ \rho(x_1 + x_2 i + x_3 j + x_4 k, v) = (x_1 + x_2 i + x_3 J + x_4 K)(v), \tag{A.7}$$

then ρ gives an action of the quaternions on V.

We can now define a quaternionic subspace of V as a subspace that is invariant under the action ρ, or equivalently, under the action of J. In particular, for each non-zero vector $v \in V$, we can define the quaternionic line spanned by v as the subspace of V spanned by v and Jv. We will call this line ℓ_v. We can easily see that these quaternionic lines are two dimensional, for suppose that, for some $v \in V$, the line ℓ_v is one dimensional, then we would have $Jv = cv$ for some complex number c. But this is impossible, since it would follow that

$$-v = J^2 v = J(cv) = \bar{c}Jv = \bar{c}cv = |c|^2 v.$$

Thus, ℓ_v is two dimensional and spanned by v and Jv. It is also easy to show that $w \in \ell_v$ implies $\ell_v = \ell_w$. In fact, $w \in \ell_v$ implies $w = \rho(q)v$ for some quaternion q, which in turn implies that $v = \rho(q^{-1})w$ and it follows that v and Jv are contained in ℓ_w, giving $\ell_w = \ell_v$. We can now easily deduce that any two quaternionic lines in V which have a non-zero element in common will be identical (both lines will be identical to the line spanned by the common element).

If V has complex dimension 2, then the above facts about quaternionic lines make it possible to identify V with the quaternions in a way that is consistent with the action of ρ. In fact, we can do this in infinitely many ways. These identifications work as follows: take an arbitrary non-zero element v of V. Since V is two dimensional, we have $V = \ell_v$ and a basis for V is given by v and Jv. This means that every element w of V can be written in the form $w = z_1 v + z_2 Jv$, where z_1 and z_2 are complex numbers. If we now identify this with a quaternion as follows

$$z_1 v + z_2 Jv \equiv z_1 + jz_2, \tag{A.8}$$

then we see that the action of $\rho(q)$ on the left hand side is just the same as right multiplication with the quaternion q on the right hand side. [2] Note that the identification

[2] If we had taken the identification $z_1 v + z_2 Jv \equiv z_1 + z_2$, then the action of $\rho(q)$ would correspond to left multiplication with the quaternion q, but we will choose to let ρ denote right multiplication here.

above also gives us an algebra of linear maps on V which correspond to left multiplication by quaternions. In particular, we have the maps corresponding to multiplication by the quaternions i, j, k and we will call these e_1, e_2, e_3 respectively.

If V has complex dimension higher than 2, things become a bit more complicated. Fortunately, we don't need to develop the general case, because we won't use it in this thesis. The interested reader may easily verify, using the properties of quaternionic lines proved above, that a three dimensional complex vector space cannot have a quaternionic structure and that a four dimensional complex vector space with a quaternionic structure can be identified with $\mathbb{H}^2$ much like a two dimensional one can be identified with $\mathbb{H}$.

We will now introduce the notion of a quaternionic linear map between vector spaces V and W with quaternionic structures J_V and J_W. We define this to be a linear map that commutes with quaternion multiplication as given by the actions ρ_V and ρ_W that correspond to J_V and JW. A necessary and sufficient condition for a map $L : V \to W$ to be quaternionic linear is, that it is complex linear and that we have $J_W L = L J_V$. It is easy to see that a quaternionic linear map L from V to W will send quaternionic subspaces of V onto quaternionic subspaces of W. In particular, the image of L will be a quaternionic subspace of W. Also, the preimage under L of a quaternonic subspace of W will be a quaternionic subspace of V and in particular, the kernel of L will be quaternionic.

Explicit examples of quaternionic linear maps are the maps $e_1, e_2, e_3 : V \to V$ defined above for an arbitrary two dimensional vector space V with a quaternionic structure. In fact, in this case, it is easy to show that we can write an arbitrary quaternionic linear operator L from V to V as

$$L = a_0 I_{2\times 2} + a_1 e_1 + a_2 e_2 + a_3 e_3,$$

where the a_i are arbitrary real numbers. This means that all quaternionic linear operators on a two dimensional space V with the quaternionic structure are just left multiplication with some quaternion. We will call such an operator real, if the quaternion in question is a real number.

Now let V again be a complex vector space with a quaternionic structure J and let W be some other complex vector space. We would like to define a quaternionic structure $\mathcal{J}$ on the tensor product $W \otimes V$. The easiest way to do this is to take

$$\mathcal{J}(w \otimes v) = \bar{w} \otimes (Jv)$$

This is a good quaternionic structure, but certainly not the only possible one. To define more general quaternionic structures, we need the notion of a *real structure.* This is a complex anti-linear map $\sigma : W \to W$, whose square is equal to the identity. That is, we have

$$\begin{aligned} \sigma &: W \to W \\ \sigma(zv) &= \bar{z}\sigma(v) \\ \sigma^2 &= I \end{aligned}$$

A real structure can be taken as an alternative definition of complex conjugation. Using the structure, one can define alternative real and imaginary parts $R(w)$ and $I(w)$ for any element w of W by

$$R(w) = \frac{w + \sigma(w)}{2}$$
$$I(w) = \frac{w - \sigma(w)}{2i} \tag{A.9}$$

These are permuted in the usual way when one multiplies w by i and we also have

$$\sigma(R(w) + iI(w)) = R(w) - iI(w), \tag{A.10}$$

which shows the analogy with complex conjugation explicitly.

An arbitrary complex vector space permits many different choices of real structure, next to the standard complex conjugation. Any one of these structures will permit us to define a good quaternionic structure on $W \otimes V$ by

$$\mathcal{J}(w \otimes v) = (\sigma w) \otimes (Jv).$$

Of course, we can easily generalise to the case of n spaces $W_1, \ldots W_n$ with real structures $\sigma_1 \ldots \sigma_n$. In that case we find that $\mathcal{J}$ given by

$$\mathcal{J}(w_1 \otimes \ldots \otimes w_n \otimes v) = (\sigma_1 w_1) \otimes \ldots \otimes (\sigma_n w_n) \otimes (Jv).$$

is a good quaternionic structure.

If V is two dimensional and identified with $\mathbb{H}$ in the way we described above, then we can identify $W \otimes V$ with $\mathbb{H}^n$, where n is the dimension of W. This is done as follows. If v is the element of V which we have identified with the quaternion 1, then we can write an arbitrary element u of $W \otimes V$ as $w_1 \otimes v + w_2 \otimes Jv$, where w_1 and w_2 are complex n-vectors. The first thing that comes to mind now, is to identify u with the quaternionic n-vector $w_1 + jw_2$. However, this is not the identification we were looking for, because with this identification, we would have

$$\begin{aligned} \mathcal{J}(w_1 \otimes v + w_2 \otimes Jv) &= -\sigma(w_2) \otimes v + \sigma(w_1) \otimes Jv, \text{ and} \\ (w_1 + jw_2)j &= -\bar{w}_2 + j\bar{w}_1 \end{aligned}$$

and the right hand sides do not in general correspond, unless the real structure σ is just the ordinary complex conjugation. For general σ, we have to do something slightly more complicated. First, we note that, if we look at the linear space W with real in stead of complex scalars, then the maps R and iI defined through (A.9) become linear projections. Moreover, we have $W = R(W) + iI(W)$ and $R(W) \bigcap iI(W) = 0$, [3] which gives $W = R(W) \oplus iI(W)$. We can now take bases in the real linear spaces $R(W)$ and $iI(W)$ - if we already had a base β on W, then we can take $R(\beta)$ and $iI(\beta)$ - and we can write $\hat{R}(w)$ and $\hat{I}(w)$ for the (real) coordinate vectors of $R(w)$ and $iI(w)$ with respect to

[3] We have $R(v) = iI(w) \Rightarrow v + \sigma(v) = w + \sigma(w) \Rightarrow w - v = \sigma(w + v) \Rightarrow \sigma(w - v) = w + v$. The last two expressions together imply $w = \sigma(w)$ and $v = -\sigma(v)$, giving $R(v) = iI(w) = 0$, as claimed.

the chosen bases. The identification of $W \otimes V$ with $\mathbb{H}^n$ that we were looking for is now given by

$$w_1 \otimes v + w_2 \otimes Jv \equiv \hat{R}(w_1) + i\hat{I}(w_1) + j\left(\hat{R}(w_2) + i\hat{I}(w_2)\right). \tag{A.11}$$

We have just shown that this identification is well defined and one easily checks (using (A.10)) that it makes the quaternionic structure coincide with quaternionic right multiplication, as desired.

As before, we can also define quaternionic left multiplication on $W \otimes V$ and the multiplications by i, j, k are given by the formula

$$e_i(w \otimes v) := w \otimes e_i(v)$$

(with a slight abuse of notation). Once more, one easily checks that the e_i are quaternionic linear with respect to the structure $\mathcal{J}$ on $W \otimes V$. Also, If A_1, A_2 and A_3 are real maps on W which commute with the real structure on W, then the map L defined by

$$L = A_0 \otimes I_{2\times 2} + A_1 \otimes e_1 + A_2 \otimes e_2 + A_3 \otimes e_3,$$

will be quaternionic linear. If W is finite dimensional, it can be identified with a quaternionic matrix, which will be real exactly if A_1, A_2 and A_3 are zero. Therefore, we will call a map like L real exactly if that condition is satisfied.

This is all we will need about quaternionic structures. A more formal treatment of real and quaternionic structures and their relation to Lie group actions and of some of the other subjects in this appendix can be found in [74]

Bibliography

[1] P.M. Sutcliffe. BPS monopoles. *Int. J. Mod. Phys.*, A12:4663–4706, 1997.

[2] P.A.M. Dirac. Quantised singularities in the electromagnetic field. *Proc. Roy. Soc.*, A133:60, 1931.

[3] G. 't Hooft. Magnetic monopoles in unified gauge theories. *Nucl. Phys.*, B79:276, 1974.

[4] A. M. Polyakov. Particle spectrum in quantum field theory. *JETP Lett.*, 20:194, 1974.

[5] S. Mandelstam. Vortices and quark confinement in nonabelian gauge theories. *Phys. Lett.*, 53B:476–478, 1975.

[6] C. Montonen and D. Olive. Magnetic monopoles as gauge particles ? *Phys. Lett.*, 72B:117, 1977.

[7] N. Seiberg and E. Witten. Electric-magnetic duality, monopole condensation, and confinement in n=2 supersymmetric Yang-Mills theory. *Nucl. Phys.*, B426:19, 1994.

[8] R. Rajaraman. *Solitons and Instantons.* North Holland, 1982.

[9] P. Morse and H. Feschbach. *Methods of Mathematical Physics.* McGraw-Hill Book Co., New York, 1953.

[10] R. F. Dashen, B. Hasslacher, and A. Neveu. *Phys. Rev.*, D10:4130, 1974.

[11] N. H. Christ and T. D. Lee. Quantum expansion of soliton solutions. *Phys. Rev.*, D12:1606, 1975.

[12] N.S. Manton. A remark on the scattering of BPS monopoles. *Phys. Lett.*, 110B(1):54, 1982.

[13] M. Shifman, editor. *Instantons in Gauge Theories.* World Scientific, 1994.

[14] K. Uhlenbeck. Removable singularities in Yang-Mills fields. *Bull. Amer. Math. Soc.*, 1:465–480, 1979.

[15] M. F. Atiyah, V.G. Drinfeld, N. J. Hitchin, and Y.I. Manin. Construction of instantons. *Phys. Lett.*, 65 A:185–187, 1978.

[16] M. Spivak. *Calculus on Manifolds.* Benjamin, New York, 1965.

[17] M.F. Atiyah. *Geometry of Yang-Mills Fields.* Lezioni Fermiane. Scuola Normale Superiore, Pisa, 1979.

[18] V.G. Drinfeld and Y.I. Manin. Instantons and sheaves on CP^3. *Funk. Analiz.*, 13:59–74, 1979.

[19] V.G. Drinfeld and Y.I. Manin. A description of instantons, 1978. preprint.

[20] A. Belavin, A. Polyakov, A. Schwartz, and Y. Tyupkin. Pseudoparticle solutions of the Yang-Mills equations. *Phys. Lett.*, 59B:85–87, 1975.

[21] G. 't Hooft. Introduction to general relativity. lecture notes of a course at the University of Utrecht, unpublished.

[22] H. Georgi and S. Glashow. Unified weak and electromagnetic interactions without neutral currents. *Phys. Rev. Lett.*, 28:1494, 1972.

[23] J. Arafune, P. Freund, and C. Goebel. Topology of Higgs fields. *J. of Math. Phys.*, 16:433, 1975.

[24] E. B. Bogomol'nyi. The stability of classical solutions. *Sov. J. Nucl. Phys.*, 24:449, 1976.

[25] S. Coleman, S. Parke, A. Neveu, and C. Sommerfield. Can one dent a dyon ? *Phys. Rev.*, D15:544, 1977.

[26] S. Coleman and E. Weinberg. *Phys. Rev.*, D7:1888, 1973.

[27] C. Taubes. Min-max theory for the Yang-Mills-Higgs equations. *Commun. Math. Phys.*, 97:473, 1985.

[28] N. Manton. The force between 't Hooft Polyakov monopoles. *Nucl. Phys.*, B126:525, 1977.

[29] P.J. Kwanten. BPS monopoles and their low energy scattering, 1995. student thesis, University of Utrecht.

[30] M. K. Prasad and C.H. Sommerfield. Exact classical solutions of the 't Hooft monopole and Julia-Zee dyon. *Phys. Rev. Lett.*, B35:760, 1975.

[31] B. Julia and A. Zee. Poles with both magnetic and electric charges in non-Abelian gauge theory. *Phys. Rev.*, D11:2227, 1975.

[32] P. Forgács, Z. Horváth, and L. Palla. Generating the BPS one monopole solution by a bäcklund transformation. *Phys. Rev. Lett.*, 45:505, 1980.

[33] P. Forgács, Z. Horváth, and L. Palla. *Ann. Phys.*, 136:371, 1981.

[34] M. F. Atiyah and R. S. Ward. Instantons and algebraic geometry. *Commun. Math. Phys.*, 55:117–124, 1977.

[35] E. Corrigan and P. Goddard. An n monopole solution with 4n-1 degrees of freedom. *Commun. Math. Phys.*, 80:575–587, 1981.

[36] N. Hitchin. Monopoles and geodesics. *Commun. Math. Phys.*, 83:579–602, 1982.

[37] W. Nahm. All self-dual monopoles for arbitrary gauge groups. In J. Hohnerkamp, editor, *Structural elements in particle physics*, page 301, 1981. NATO ASI.

[38] W. Nahm. The construction of all self-dual monopoles by the ADHM method. In N. S. Craigie, P. Goddard, and W. Nahm, editors, *Monopoles in quantum field theory*, pages 87–94, December 1981. Proceedings of the monopole meeting, Trieste, Italy.

[39] N. J. Hitchin. On the construction of monopoles. *Comm. Math. Phys.*, 89:145, 1983.

[40] C. Callias. Axial anomalies and index theorems on open spaces. *Commun. Math. Phys.*, 62:213, 1978.

[41] P. van Driel. The Nahm construction for self dual gaugefields, 1988. student thesis, University of Amsterdam.

[42] W. Nahm. *Phys. Lett.*, 93B:42, 1980.

[43] N.J. Hitchin, N.S. Manton, and M.K. Murray. Symmetric monopoles. *Nonlinearity*, 8:661–692, 1995.

[44] C.J. Houghton and P.M. Sutcliffe. Octahedral and dodecahedral monopoles. *Nonlinearity*, 9:385–401, 1996.

[45] T. Brzeziǹski and H. Merabet. Nahm's equations and root systems. hep-th/9709070.

[46] E. Corrigan and P. Goddard. Construction of instanton and monopole solutions and reciprocity. *Ann. Phys.*, 154:253, 1984.

[47] P.J. Braam and P. van Baal. Nahm's transformation for instantons. *Commun. Math. Phys.*, 122:267–280, 1989.

[48] E. van der Meer. Magnetic monopoles in gauge field theories, 1997. student thesis, University of Amsterdam.

[49] N. Manton. Radiation from monopole scattering. *Phys. Lett.*, B215:559, 1988.

[50] D. Stuart. The geodesic approximation for the Yang-Mills-Higgs equations. *Commun. Math. Phys.*, 166:149–190, 1994.

[51] E. J. Weinberg. Parameter counting for multimonopole solitons. *Phys. Rev.*, D20:936, 1979.

[52] C. H. Taubes. Stability in Yang-Mills theories. *Commun. Math. Phys.*, 91:235, 1983.

[53] M.F. Atiyah and N. Hitchin. *The Geometry and Dynamics of Magnetic Monopoles*. Princeton University Press, 1988.

[54] H. Nakajima. Monopoles and Nahm's equations. In Marcel Dekker, editor, *Einstein metrics and Yang-Mills connections*, New York, 1993.

[55] H. Goldstein. *Classical Mechanics*. Addison-Wesley, second edition, 1980.

[56] G. Gibbons and C. Pope. The positive action conjecture and asymptotically Euclidean metrics in quantum gravity. *Commun. Math. Phys.*, 66:267–290, 1979.

[57] G.W. Gibbons and N. S. Manton. Classical and quantum dynamics of BPS monopoles. *Nucl. Phys.*, B274:183, 1986.

[58] P. Forgács, Z. Horváth, and L. Palla. Nonlinear superposition of monopoles. *Nucl. Phys.*, 192B:141, 1981.

[59] N. S. Manton. Monopole interactions at long range. *Nucl. Phys.*, 154B:397, 1985. Erratum Phys. Lett. 110B, 54.

[60] M. Temple-Raston. BPS two monopole scattering. *Phys. Lett.*, B206:503, 1988.

[61] M. Temple-Raston. Two magnetic monopole Julia sets. *Phys. Lett.*, B213:168, 1988.

[62] S.W. Hawking. Gravitational instantons. *Phys. Lett.*, 60A:81, 1977.

[63] N. S. Manton. Monopole and Skyrmion bound states. *Phys. Lett.*, B198:226, 1987.

[64] B. J. Schroers. Quantum scattering of BPS monopoles at low energy. *Nucl. Phys.*, B367:177–214, 1991.

[65] C. J. Houghton and P. M. Sutcliffe. Inversion symmetric 3-monopoles and the Atiyah-Hitchin manifold. *Nonlinearity*, 1996. hep-th 9609136.

[66] R. Bielawski. The existence of closed geodesics on the moduli space of k-monopoles. *Nonlinearity*, 1996.

[67] H. W. Braden and P. M. Sutcliffe. A monopole metric. hep-th 9610141.

[68] G.W. Gibbons and N. S. Manton. The moduli space metric for well-separated BPS monopoles. *Phys. Lett.*, B356:32–38, 1995.

[69] R. Bielawski. Monopoles and the Gibbons-Manton metric. hep-th 9801091.

[70] R. Bielawski. Asymptotic metrics for $SU(N)$ monopoles with maximal symmetry breaking. hep-th 9801092.

[71] D. Bak, C. Lee, and K. Lee. Dynamics of BPS dyons: effective field theory approach. hep-th/9708149.

[72] J. D. Jackson. *Classical Electrodynamics*. John Wiley and sons, second edition, 1962.

[73] D. Bak and C. Lee. BPS dyon in a weak electromagnetic field: Equations of motion and radiation fields. *Nucl. Phys.*, B424:124–136, 1994.

[74] T. Bröcker and T. tom Dieck. *Representations of compact Lie groups*, volume 98 of *Graduate Texts in Mathematics*. Springer-Verlag New York Inc., New York, 1985.