The ASD Equations in Split Signature and Hypersymplectic Geometry

Markus Karl Röser

New College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2012

Abstract:

The ASD Equations in Split Signature and Hypersymplectic Geometry

Markus Karl Röser, New College, University of Oxford,
submitted for D.Phil., Trinity Term 2012

This thesis is mainly concerned with the study of hypersymplectic structures in gauge theory. These structures arise via applications of the hypersymplectic quotient construction to the action of the gauge group on certain spaces of connections and Higgs fields.

Motivated by Kobayashi-Hitchin correspondences in the case of hyperkähler moduli spaces, we first study the relationship between hypersymplectic, complex and paracomplex quotients in the spirit of Kirwan's work relating Kähler quotients to GIT quotients. We then study dimensional reductions of the ASD equations on $\mathbb{R}^{2,2}$. We discuss a version of twistor theory for hypersymplectic manifolds, which we use to put the ASD equations into Lax form.

Next, we study Schmid's equations from the viewpoint of hypersymplectic quotients and examine the local product structure of the moduli space. Then we turn towards the integrability aspects of this system. We deduce various properties of the spectral curve associated to a solution and provide explicit solutions with cyclic symmetry.

Hitchin's harmonic map equations are the split signature analogue of the self-duality equations on a Riemann surface, in which case it is known that there is a smooth hyperkähler moduli space. In the case at hand, we cannot expect to obtain a globally well-behaved moduli space. However, we are able to construct a smooth open set of solutions with small Higgs field, on which we then analyse the hypersymplectic geometry. In particular, we exhibit the local product structures and the family of complex structures. This is done by interpreting the equations as describing certain geodesics on the moduli space of unitary connections. Using this picture we relate the degeneracy locus to geodesics with conjugate endpoints.

Finally, we present a split signature version of the ADHM construction for so-called split signature instantons on $S^2 \times S^2$, which can be given an interpretation as a hypersymplectic quotient.

Acknowledgements

First of all, I would like to thank my supervisor Professor Andrew Dancer for proposing this project and for introducing me to this and many other exciting areas of mathematics. Without his encouragement and both mathematical and personal support and guidance over the past three years, this thesis would not exist.

I am grateful to Professor Nigel Hitchin and Professor Lionel Mason, who were the examiners in my Transfer and Confirmation of Status exams, for many insightful comments. I thank Nuno Romão for helpful discussions on Schmid's equations.

Professor Anton Deitmar, Professor Franz Pedit, Professor David Wharam and Dr. Ivan Smith have encouraged me to apply to Oxford to undertake this research project and I thank them for their support.

Thanks are also due to Jesus College for giving me a fantastic opportunity to gain teaching experience as a Retained Lecturer over the past two years and for offering me a quiet place in the Fellows' Library, where the final editing of this thesis was done over the past few months.

My studies in Oxford were partially funded by an EPSRC DPhil studentship.

On a more personal note, I want to thank all my friends at the Mathematical Institute, Jesus College and New College who made the last years such an enjoyable experience. Finally, I would like to thank my family for their support and, in particular, Heike for her love, patience and tolerance over the years.

Contents

Introduction **1**

1 General Background **9**
- 1.1 Quotients in Riemannian and Symplectic Geometry 9
  - 1.1.1 The Quotient Metric 10
  - 1.1.2 Moment Maps and Symplectic Quotients 12
  - 1.1.3 The Kähler Quotient 16
  - 1.1.4 The Hyperkähler Quotient 19
- 1.2 Gauge Theory 25
  - 1.2.1 The Anti-Self-Dual Yang-Mills Equations 25
  - 1.2.2 Analytical Tools 27
  - 1.2.3 The Space of Unitary Sobolev Connections 29
  - 1.2.4 The Action of the Gauge Group 29
  - 1.2.5 The ASD Moduli Space on a Hyperkähler Manifold is Hyperkähler 37

2 Hypersymplectic Geometry **44**
- 2.1 Hypersymplectic Manifolds 44
- 2.2 The Hypersymplectic Quotient 47
- 2.3 Kirwan-Type Theorems 51
  - 2.3.1 Complex Quotients 51
  - 2.3.2 Paracomplex Quotients 57
  - 2.3.3 Application to Hypersymplectic Quotients 60
  - 2.3.4 Examples coming from Linear Torus Actions 62
- 2.4 The ASD Equations on $\mathbb{R}^{2,2}$ and their Dimensional Reductions 75
  - 2.4.1 ASD Connections on Hypersymplectic Manifolds 75
  - 2.4.2 The Lax Pair Formalism 78
  - 2.4.3 Remarks on Twistors in Split Signature 84

3 Schmid's Equations **87**
3.1 The Hypersymplectic Setup . . . 87
3.2 The Spectral Curve . . . 92
3.2.1 The Underlying Real Curve . . . 99
3.3 Conserved Quantities . . . 100
3.4 Group actions on the moduli space $\mathcal{M}_{Sch}$ . . . 102
3.5 Complex Structures and Product Structures . . . 106
3.5.1 Complex Structures . . . 106
3.5.2 Product Structures . . . 112
3.6 Explicit Solutions . . . 115
3.6.1 Review of the $\mathfrak{su}(2)$-Case . . . 116
3.6.2 Solutions with Cyclic Symmetry . . . 117

4 Harmonic Maps **130**
4.1 Rewriting the Equations on $\mathbb{R}^2$ . . . 131
4.1.1 Degeneracies . . . 134
4.2 The Hypersymplectic Setup for the Equations on a Compact Riemann Surface . . . 136
4.3 Two Applications of the Implicit Function Theorem . . . 139
4.4 A Different Interpretation of the Harmonic Map Equations in Terms of Geodesics . . . 143
4.4.1 The Space of Geodesics on $\mathcal{A}$ . . . 145
4.4.2 Conjugate Points and the Degeneracy Locus . . . 152
4.5 The Circle Action . . . 153
4.6 Product Structures and Complex Structures . . . 155
4.6.1 Product Structures . . . 155
4.6.2 The Complex Structures . . . 159
4.7 Harmonic Tori from Schmid's Equations . . . 162

5 Split Signature Instantons on $S^2 \times S^2$ **167**
5.1 ASD Connections on $\mathbb{R}^{2,2}$ . . . 168
5.2 The Twistor Correspondence for ASD Connections on $S^2 \times S^2$ . . . 168
5.3 Monads, ADHM Data and Hypersymplectic Quotients . . . 171
5.3.1 ADHM Data in Split Signature . . . 172
5.3.2 The moduli space of ADHM Data as a Hypersymplectic Quotient 175
5.3.3 An ADHM Description for Framed Bundles on $\mathbb{CP}^1 \times \mathbb{CP}^1$ . . 178
5.3.4 Open Questions . . . 180

Introduction

The main topic of this thesis is hypersymplectic manifolds. These can be thought of as pseudo-Riemannian analogues of hyperkähler manifolds. Historically, hypersymplectic structures have their origin in the study of integrable systems. They arose naturally in Hitchin's work surrounding the gauge-theoretic equations describing harmonic maps from a 2-torus into compact Lie groups ([29], [30]) during the late 1980's. To motivate the definition, we give a brief overview over the interactions of hyperkähler geometry and gauge theory, which will be discussed in more detail in chapter 1.

It is known that many moduli spaces of solutions to (dimensional reductions of) the anti-self-dual Yang-Mills equations, or ASD equations for short, on Euclidean space $\mathbb{R}^4$ carry natural hyperkähler metrics. Examples include the moduli space of framed instantons on $\mathbb{R}^4$, [43], moduli spaces of magnetic monopoles [5], moduli spaces of solutions to Nahm's equations [38], [39], [10], [37], and, maybe most importantly in the present context, the moduli space of Higgs bundles on a compact Riemann surface of genus $g > 1$, [28]. These hyperkähler structures arise via an adaption of the Marsden-Weinstein symplectic quotient construction to the category of hyperkähler manifolds, [31], applied to the infinite-dimensional situation of the gauge group acting on the space of connections and possibly Higgs fields on a bundle with compact structure group. The space of connections acquires a hyperkähler structure essentially because $\mathbb{R}^4$ is a hyperkähler manifold, for as a Euclidean vector space, it is isomorphic to the quaternions $\mathbb{H}$. The ASD equations are then precisely the condition that the hyperkähler moment map associated to the action of the gauge group vanishes. On the one hand, knowing the existence of such a rich geometry has led to advances in understanding these ASD moduli spaces which play an important role in mathematical physics. Information about the Riemannian geometry of the moduli spaces is valuable, since for example low-energy scattering of magnetic monopoles is implemented by geodesic motion in the moduli space , see [44], [5].

But also from the viewpoint of pure mathematics these moduli spaces are very interesting objects, as they provide many examples of hyperkähler metrics on known non-compact complex symplectic manifolds, such as certain spaces of rational maps [21], the cotangent bundle and co-adjoint orbits of complex semi-simple Lie groups ([38], [39], [10], [37]), moduli spaces of complex representations of surface groups [28]. Knowing the existence of a hyperkähler metric then leads to new insights in the study of the algebro-geometric objects parametrised by the moduli space. The theorems identifying the above moduli spaces with ASD moduli spaces are called Kobayashi-Hitchin correspondences. They are essentially infinite-dimensional versions of the following fact due to Kirwan. For Hamiltonian actions of compact Lie groups on finite-dimensional Kähler manifolds, it was discovered by Kirwan in [35], that there is a homeomorphism between the associated Kähler quotient and the topological quotient of the subset of so-called stable points in the manifold by the action of the complexified group. Using Kirwan's theorem, we may interpret hyperkähler quotients as symplectic quotients in the holomorphic category. In infinite dimensions, this theorem does not apply automatically and one has to prove that the moduli spaces may be identified with moduli spaces of the respective complex-geometric objects modulo the action of the complexified gauge group.

In this context, the subject of hypersymplectic geometry has as a starting point the observation that also on $\mathbb{R}^{2,2}$ the Hodge star operator on 2-forms is an involution and so the ASD equations make sense. In fact, many interesting partial differential equations in mathematical physics arise as dimensional reductions of the ASD equations in split signature. The most prominent examples are the KdV and non-linear Schrödinger equations [46], and the equations for a harmonic map from a domain in $\mathbb{R}^2$ into a compact Lie group [29], [27]. It then turns out that in analogy to the above identification $\mathbb{R}^4 \cong \mathbb{H}$, one should think of $\mathbb{R}^{2,2}$ as the algebra of real 2×2-matrices $\mathbb{B} = \mathfrak{gl}(2, \mathbb{R})$. This then comes equipped with three symplectic forms satisfying certain algebraic relations such that the ASD equations on $\mathbb{R}^{2,2}$ again admit a moment map interpretation. In making this observation for the equations describing harmonic maps from Riemann surfaces into compact Lie groups, Hitchin was led to the definition of hypersymplectic structures and to the study of hypersymplectic quotients [30]. Thus, hypersymplectic geometry should be the natural geometry on moduli spaces of solutions of the ASD equations and their dimensional reductions in split signature. This is the main topic with which this thesis is concerned.

More formally, hypersymplectic manifolds are special kinds of complex symplectic manifolds which carry a pseudo-Riemannian metric, the holonomy of which is contained in the split real form $\mathrm{Sp}(2n, \mathbb{R})$ of the complex Lie group $\mathrm{Sp}(2n, \mathbb{C})$. In this sense they are cousins of hyperkähler manifolds, whose holonomy lies in the compact real form $\mathrm{Sp}(n)$ of $\mathrm{Sp}(2n, \mathbb{C})$. In analogy to the case of hyperkähler manifolds, which are to be thought of as being modelled on a vector space over the quaternions $\mathbb{H}$, the geometry of hypersymplectic manifolds is naturally associated to the algebra $\mathbb{B} = \mathfrak{gl}(2, \mathbb{R})$ of real 2×2-matrices.

As an algebra, $\mathbb{B}$ is given by generators $1, i, s, t$ with relations $i^2 = -1$, $s^2 = 1 = t^2$ and $is = t = -si$. It carries a natural indefinite inner product g of signature $(2, 2)$ given by $g(q, q) = q\bar{q} = q_0^2 + q_1^2 - q_2^2 - q_3^2$, where $q = q_0 + iq_1 + sq_2 + tq_3 \in \mathbb{B}$. Then the group $\mathrm{Sp}(n, \mathbb{B})$ of $\mathbb{B}$-linear isometries of $\mathbb{B}^n$ can be shown to be isomorphic to $\mathrm{Sp}(2n, \mathbb{R})$. Thus, a hypersymplectic manifold is a pseudo-Riemannian manifold M of dimension $4n$ with a metric g of neutral signature and parallel skew-adjoint endomorphisms I, S, T of the tangent bundle satisfying the above relations. These then give rise to three symplectic forms $\omega_I, \omega_S, \omega_T$ which satisfy appropriate algebraic relations, from which one can recover the endomorphisms and the metric. Since S, T square to the identity and are parallel, their ± 1-eigensubbundles of TM are integrable in the sense of Frobenius. So hypersymplectic manifolds can locally be written as a product of two $2n$-dimensional manifolds. Since a hyperkähler structure is defined by a triple of complex structures, this feature is not present in hyperkähler geometry. In this sense, there are different viewpoints from which one can study hypersymplectic manifolds. One can look at them either as pseudokähler manifolds, i.e. putting the emphasis on the complex structures, or from the point of view of parakähler geometry where the product structures play the main role. The third viewpoint puts emphasis on the three symplectic forms and is the one which seems to be the one fitting best into the spirit of integrable systems.

Analogously to the hyperkähler situation [31], there is a hypersymplectic quotient construction [30]. However, since the metric is not positive definite, this quotient construction is less well-behaved than its hyperkähler analogue. In order to obtain a well-defined hypersymplectic structure on the quotient, one has to worry about points at which the restriction of the metric to the tangent space to the G-orbit becomes degenerate, a pathology not occurring in the hyperkähler framework due to the positive definiteness of the Riemannian metric.

After its introduction in 1990, hypersymplectic geometry has been an active area of mathematical research. Apart from their link with the ASD equations in split signature, hypersymplectic manifolds are also split signature examples of Calabi-Yau manifolds and therefore appear in the study of certain supersymmetric sigma-models in string theory [15], [33].

In real dimension four, in analogy to the hyperkähler situation, hypersymplectic structures correspond to metrics of signature $(2,2)$ that are Ricci-flat and self-dual. Kamada in [34] has provided a complete classification of the compact examples. As complex manifolds with respect to the complex structure I, they are either complex tori or primary Kodaira surfaces, i.e. certain topologically non-trivial elliptic principal bundles over an elliptic curve. Such primary Kodaira surfaces have their first Betti number equal to 3 and hence cannot be Kähler. One of the key observations involved in the proof is the fact that the volume forms given by the squares of ω_I and ω_S have different signs. The class of compact four-manifolds admitting two symplectic forms inducing opposite orientations can then be studied in terms of Seiberg-Witten invariants. There are also many examples of hypersymplectic structures known on Lie groups, see [2].

Using the hypersymplectic quotient construction, Dancer and Swann constructed a large class of examples from linear torus actions on a hypersymplectic vector space [19]. More recently, applying the hypersymplectic quotient construction in an infinite-dimensional setting, Matsoukas in [48] has analysed the hypersymplectic structure on the moduli space of solutions to the $\mathfrak{su}(2)$-Schmid equations, an example we will return to in chapter 3.

Following Hitchin's original motivation, we study in this thesis the geometry of moduli spaces of solutions to gauge-theoretic equations that can be interpreted as the vanishing of a hypersymplectic moment map. In particular, we study the anti-self-dual Yang-Mills equations for connections on vector bundles with compact structure group G over $\mathbb{R}^{2,2} \cong \mathbb{B}$ and their dimensional reductions.

The thesis starts in chapter 1 with a survey on quotient constructions in Riemannian and symplectic geometry and a review of the analytical tools involved in studying gauge theory. In particular, we review the hyperkähler quotient construction and then move on to give a fairly self-contained proof of the fact that the moduli space of instantons on a $K3$-surface is hyperkähler, a theorem well-known to experts for which a detailed proof exploiting the interpretation as a hyperkähler quotient has not been available in the literature to our knowledge.

Next, we give an introduction to hypersymplectic geometry with special emphasis on the hypersymplectic quotient construction, which displays pathologies which do not play a role in its hyperkähler analogue. The proofs of the Kobayashi-Hitchin correspondences mentioned above give rise to an interpretation of hyperkähler quotients as complex symplectic quotients. In finite dimensions, this interpretation relies on Kirwan's work on the relation between Kähler quotients and GIT quotients [35]. This is our motivation to discuss the relation between pseudokähler quotients and complex quotients as well as parakähler quotients and paracomplex quotients, generalising Kirwan's work to the case of pseudokähler and parakähler manifolds in section 2.3. Let G be a compact Lie group with complexification $G^{\mathbb{C}}$ acting on a compact Kähler manifold (M, g, ω) preserving the Kähler structure with moment map μ. Assume that the action can be extended to a holomorphic action of $G^{\mathbb{C}}$. Kirwan's original result shows that there is a homeomorphism $\mu^{-1}(0)/G \cong (G^{\mathbb{C}}.\mu^{-1}(0))/G^{\mathbb{C}}$. Due to the change of signature in the metric and the use of gradient flow techniques in the proof of the original theorem, one cannot expect such a theorem to hold in general for pseudokähler manifolds. Therefore, we expect that $G^{\mathbb{C}}$-orbits intersect $\mu^{-1}(0)$ in more than one point modulo the action of G.

However, we are able to prove a version of Kirwan's theorem under the assumption that the restriction of the pseudokähler metric to the orbits of the compact group action should be positive or negative definite. Moreover, dropping the definiteness assumption, we find that away from the degeneracy locus the set of points in a $G^{\mathbb{C}}$-orbit that lie in $\mu^{-1}(0)$ has to be discrete modulo the action of G.

We then discuss ASD connections on hypersymplectic four-manifolds. We provide a general degeneracy criterion for the induced hypersymplectic structure on the moduli space of ASD connections, which gives a general conceptual interpretation for the degeneracy results of Matsoukas [48] and Hitchin [30] concerning moduli spaces of solutions to Schmid's equations and harmonic map equations, respectively. An important step in the study of the ASD equations on $\mathbb{R}^{2,2}$ and their dimensional reductions is to develop a version of twistor theory for hypersymplectic manifolds, which enables us to rewrite the equations in Lax form.

The following two chapters are then devoted to studying equations arising from dimensional reductions of the ASD equations on $\mathbb{R}^{2,2} \cong \mathbb{B}$. We start in chapter 3 with Schmid's equations, which are the split signature analogue of Nahm's equations, and are obtained by imposing translation invariance with respect to the i, s and t directions. These are thus a system of non-linear ODEs:

$$\begin{aligned}
\dot{T}_1 + [T_0, T_1] &= -[T_2, T_3], \\
\dot{T}_2 + [T_0, T_2] &= [T_3, T_1], \\
\dot{T}_3 + [T_0, T_3] &= [T_1, T_2],
\end{aligned}$$

where $T_i : I \to \mathfrak{g}$ are Lie algebra-valued functions on an interval $I \subset \mathbb{R}$. The Lax pair formalism implies that these equations are equivalent to the single Lax equation $\dot{T} = [T_+, T]$ involving functions $T(t,\zeta)$, $T_+(t,\zeta)$, that are time-dependent matrix polynomials of degree 2 with respect to the spectral parameter $\zeta \in \mathbb{CP}^1$. It follows that the spectral curve $S = \{\det(\eta + T(t,\zeta)) = 0\} \subset T^*\mathbb{CP}^1$ is a conserved quantity of the system, and a solution to the equations gives rise to a linear flow on the Jacobian torus of the curve S. We deduce various properties of the spectral curve. In particular, we calculate the genus in the generic case. The spectral curve carries an anti-holomorphic involution induced by inversion with respect to the unit circle in $\mathbb{CP}^1 \cong \mathbb{C} \cup \{\infty\}$, the fixed-point set of which is a real algebraic curve $S_{\mathbb{R}} \subset S$, which is not present in the Nahm case and which we investigate, too.

We study the hypersymplectic geometry of the moduli space and discuss in particular the families of complex structures and local product structures by trying to apply our results on the relation between complex, paracomplex and hypersymplectic quotients in this infinite dimensional setting.

There is a natural action of $\mathrm{SL}(2,\mathbb{R})$ on the moduli space which induces an action on spectral curves, and it is natural to look for solutions that are invariant under a cyclic subgroup of $\mathrm{SL}(2,\mathbb{R})$. We first discuss spectral curves with cyclic symmetry. It turns out that the spectral curve carries an action of the same cyclic group and that, if the action is free, the quotient curve is hyperelliptic. Modifying an ansatz of Sutcliffe [55] for cyclically invariant solutions to Nahm's equations, we are able to produce a family of explicit solutions for $\mathfrak{g} = \mathfrak{su}(n)$, where $n \geq 3$, expressible in terms of Jacobi elliptic functions and their integrals.

In chapter 4, we study Hitchin's gauge theoretic equations for harmonic maps from a Riemann surface M into a compact Lie group G. These are given by

$$R^{\nabla} = [\Phi \wedge \Phi^*], \qquad \bar{\partial}^{\nabla}\Phi = 0,$$

where (∇, Φ) is a pair consisting of a G-connection ∇ on a G-vector bundle E, and a Higgs field $\Phi \in \Gamma(M, \mathfrak{g}^{\mathbb{C}} \otimes K)$, where K is the canonical bundle and $\mathfrak{g} \subset \mathrm{End}(E)$ is the associated bundle of Lie algebras with fibre $\mathrm{Lie}(G)$. They can be obtained from the ASD equations on $\mathbb{R}^{2,2}$ by imposing translation invariance with respect to the s and t directions, and they are the split signature analogue of Hitchin's self-duality equations on a Riemann surface [28]. These equations describe harmonic sections of flat $G \times G$-bundles. Due to the different signature, we cannot expect to have a smooth global hypersymplectic moduli space in this situation, since many of the analytical arguments do not carry over from [28]. However, if we are looking for solutions with zero Higgs field, the sign change does not play a role and an argument involving the implicit function theorem enables us to produce a well-behaved neighbourhood of the moduli space of flat connections inside the moduli space of solutions to Hitchin's harmonic map equations. On this neighbourhood we can study the hypersymplectic geometry of the moduli space.

Reinterpreting the equations as the equation for a horizontal geodesic on the space $\mathcal{A}/\mathcal{G}$ of G-connections modulo gauge transformations, whose endpoints are flat connections (see[29]), we find that this neighbourhood is locally diffeomorphic to a neighbourhood of the diagonal in the product of the moduli space of flat connections with itself. Moreover, geodesics whose endpoints are conjugate can be related to elements of the degeneracy locus for the hypersymplectic structure. In classical terminology of Riemannian geometry we find that the degeneracy locus of the hypersymplectic structure is contained in the *cut locus* of the infinite-dimensional Riemannian manifold $\mathcal{A}/\mathcal{G}$. The local product structure of this neighbourhood is the one induced by the endomorphism S of the hypersymplectic structure on the moduli space, and corresponds to assigning to a geodesic its endpoints.

The moduli space furthermore carries a circle action which preserves the complex structure I and rotates the product structures. We calculate the moment map of this action, which is given by the L^2-norm of the Higgs field and, using Uhlenbeck's compactness theorem, we deduce that this moment map is proper. Thinking of the equations in terms of geodesics, this moment map corresponds to the energy of the geodesic associated to a solution.

Turning more to the integrability aspects of the equations, we are able to produce explicit formulae for harmonic maps from tori into Lie groups by interpreting Schmid's equations as a dimensional reduction of the harmonic map equations. In the case of $G = \mathrm{SU}(2)$, using Matsoukas' explicit solutions of Schmid's equations for $\mathfrak{g} = \mathfrak{su}(2)$ [48], we find formulae for Gauss maps of constant mean curvature surfaces in $\mathbb{R}^3$, which we are able to express in terms of Jacobi elliptic functions.

Finally, in chapter 5 we study a special class of ASD connections on the pseudokähler manifold $S^2 \times S^2$, which are called split signature instantons. It is shown by Mason in [45] that these correspond to stable bundles on the twistor space $\mathbb{CP}^3$, satisfying a certain reality condition. We provide a split signature version of the ADHM construction for such stable bundles on twistor space trivialised on a real line and show that it admits an interpretation as the vanishing condition of a hypersymplectic moment map associated to the action of the unitary group on a certain finite-dimensional space of matrices. This is analogous to Donaldson's construction in [20], relating the moduli space of framed instantons on $\mathbb{R}^4$ to stable holomorphic bundles on $\mathbb{CP}^2$ and is a first step towards establishing an ADHM construction in split signature.

Chapter 1

General Background

1.1 Quotients in Riemannian and Symplectic Geometry

We establish some basic background about group actions on manifolds, mostly to fix our notation, but also to give the reader a feel for the general context of the quotient constructions discussed later. Let

$$\mu : G \times M \to M, \qquad (a, p) \mapsto \mu(a, p) = a.p$$

be a smooth (left) action of a Lie group G on a manifold M. We then have two maps associated to this action. Firstly, there is the *orbit map*

$$\beta_p : G \to M, \qquad \beta_p(a) = a.p,$$

whose image is the orbit $\mathcal{O}_p = G.p$ of the group action through the point $p \in M$. Secondly, to each $a \in G$ we have an associated diffeomorphism of M

$$\lambda_a : M \to M, \qquad \lambda_a(p) = a.p,$$

which preserves orbits in the sense that $\lambda_a(p) \in \mathcal{O}_p$ for all $p \in M$. Observe that β_p and λ_a are related via

$$\lambda_a \circ \beta_p = \beta_p \circ L_a,$$

where $L_a : G \to G$ is the diffeomorphism given by left translation by $a \in G$. From this formula it follows in particular that β_p has constant rank. Therefore, the level set

$$G_p = \beta_p^{-1}(p) \subset G$$

is a smooth closed submanifold of G, which can easily be seen to be a subgroup. Hence, it is a closed Lie subgroup of G, called the *stabiliser of* p. Moreover, β_p induces a diffeomorphism

$$G/G_p \cong \mathcal{O}_p.$$

Let $\mathfrak{g} = T_1G$ be the Lie algebra of G and for $p \in M$ let $\mathfrak{g}_p \subset \mathfrak{g}$ be the Lie algebra of the stabiliser of p. Then for each $\xi \in \mathfrak{g}$ we get a vector field X^ξ on M via

$$X_p^\xi = (\mathrm{d}\beta_p)_1(\xi) = \frac{d}{dt}|_{t=0}\beta_p(\exp(t\xi)),$$

called the *fundamental vector field associated to* ξ *(and* μ*).* The map

$$\mathfrak{g} \to \Gamma(T\mathcal{O}_p) \qquad \xi \mapsto X^\xi$$

has kernel $\mathfrak{g}_p$ and induces an isomorphism of vector bundles

$$T\mathcal{O}_p \cong \mathcal{O}_p \times \mathfrak{g}/\mathfrak{g}_p.$$

If G acts freely and properly, then the orbit space M/G is naturally a smooth manifold of dimension $\dim M/G = \dim M - \dim G$ with tangent bundle given by

$$T(M/G) \cong TM/\mathfrak{g}.$$

That is

$$T_{\pi(p)}(M/G) \cong T_pM/T_p\mathcal{O}_p,$$

where $\pi : M \to M/G$ is the natural projection. In other words M, becomes a principal G-bundle over the quotient space M/G.

By definition, the action is proper whenever G is compact, which we will always assume in what follows, unless stated otherwise. We would like to know what happens if we start not just with a smooth manifold M, but with a smooth manifold carrying additional structure, such as a Riemannian metric, a Kähler structure, symplectic structure etc. Is there a way to construct a quotient which inherits this structure?

1.1.1 The Quotient Metric

We answer this question first for Riemannian manifolds. Let (M, g) be a Riemannian manifold and let G be a compact Lie group, that acts freely on M by isometries. We construct a natural metric on the orbit space M/G.

At each $p \in M$ we get an orthogonal decomposition of the tangent space into the tangent space to the orbit through p and its orthogonal complement with respect to g:

$$T_pM = T_p\mathcal{O}_p \oplus H_p,$$

where we wrote $H_p = T_p\mathcal{O}_p^{\perp}$. For reasons explained below, it is customary to call H_p the *horizontal subspace*. The derivative $\mathrm{d}\pi$ of the natural projection $\pi : M \to M/G$ gives an isomorphism

$$T_{\pi(p)}(M/G) \cong H_p.$$

Thus, for $p \in M$ every tangent vector $X \in T_{\pi(p)}(M/G)$ has a unique *horizontal lift* $\tilde{X} = (\mathrm{d}\pi_p)^{-1}(X) \in H_p$. We now define a metric $\tilde{g}$ on M/G via

$$\tilde{g}_{\pi(p)}(X, Y) = g_p(\tilde{X}, \tilde{Y}).$$

We have to check that this is well-defined, i.e. independent of the choice of point in the orbit through p. First observe that since λ_a preserves orbits, we have that if $\tilde{X} \in H_p$ is the horizontal lift of $X \in T_{\pi(p)}(M/G)$, then $(d\lambda_a)_p(\tilde{X}) \in H_{a.p}$ is the horizontal lift of X in $H_{a.p}$. Put another way, the family $\{H_p\}_{p\in M}$ defines a connection on the principal bundle M, hence we call these subspaces horizontal. Thus, we have to require that

$$g_p(\tilde{X}, \tilde{Y}) = g_{a.p}((d\lambda_a)_p(\tilde{X}), (d\lambda_a)_p(\tilde{Y})),$$

which is precisely the statement

$$\lambda_a^* g = g.$$

So the assumption that G acts by isometries ensures that our construction works. Observe in particular that the flow associated to a fundamental vector field is an isometry, which implies

$$\mathcal{L}_{X^\xi} g = 0 \quad \text{for all } \xi \in \mathfrak{g},$$

where $\mathcal{L}$ is the Lie derivative, i.e. fundamental vector fields are *Killing*. In this way we have put a Riemannian metric on the quotient space M/G. We remark that the above construction does not need any assumptions on the dimension of M and so in fact works for any Hilbert manifold as soon as the quotient and suitable splittings of T_M exist.

1.1.2 Moment Maps and Symplectic Quotients

The next class of manifolds we want to consider are symplectic manifolds. A good reference for symplectic geometry is [13]. A manifold M is called *symplectic* if it possesses a two-form $\omega \in \Omega^2(M)$ which is closed and non-degenerate. This immediately implies that the dimension of M is even.

Example. For any vector space V the space $V \oplus V^*$ carries a canonical symplectic form ω given by

$$\omega((v,\alpha),(w,\beta)) = \beta(v) - \alpha(w).$$

Example. The classical example of a symplectic manifold is the total space of the cotangent bundle T^*M of any smooth manifold M. The manifold T^*M carries a canonical symplectic form ω which arises as follows. Firstly, on T^*M there is the canonical one-form θ. If X is a vector field on T^*M, then at a point $\alpha \in T^*M$ it is given by

$$\theta_\alpha(X) = \pi^*\alpha(X),$$

where $\pi : T^*M \to M$ is the bundle projection. Then ω is defined to be

$$\omega = -\mathrm{d}\theta$$

and it can be shown that this is non-degenerate (by definition, it is closed). Identifying the tangent space $T_p(T^*M)$ at a point $p \in T^*M$ with $T_{\pi(p)}M \oplus T_{\pi(p)}M^*$, we recover the first example.

Symplectic geometry has its origins in Hamiltonian mechanics, where M plays the role of the *configuration space* and T^*M is called the *phase space* of a mechanical system.

By Darboux' theorem, every symplectic manifold is locally symplectomorphic to a cotangent bundle. So in this sense, this example is universal.

Let now G be a compact Lie group which acts on (M,ω) by *symplectomorphisms*, i.e.

$$\lambda_a^*\omega = \omega.$$

If the dimension of G is odd, then the dimension of the quotient M/G is odd too, and so cannot be symplectic again. And if the dimensions work out correctly, then still the quotient M/G will in general not be a symplectic manifold again, since if we restrict the symplectic form $\omega_p \in \Lambda^2(T_pM^*)$ to the tangent space of the quotient $T_p(M/G) = T_pM/T_p\mathcal{O}_p$, we cannot a priori conclude that it will define a symplectic form, as it may become *degenerate*.

In order to avoid these difficulties, we have to find a way to make sure that our quotient manifold is actually symplectic: The idea is to find a submanifold $N \subset M$, such that at each point $p \in N$ the restriction of ω to T_pN is degenerate precisely along the G-orbits, hence non-degenerate on $T_p(N/G) = T_pN/T_p\mathcal{O}_p$. Such a submanifold N is cut out by a *moment map*.

The action of G preserving the symplectic form implies that the fundamental vector fields for this action are *symplectic*. That is

$$\mathcal{L}_{X^\xi}\omega = 0 \quad \text{for all } \xi \in \mathfrak{g}.$$

Since ω is closed, Cartan's formula

$$\mathcal{L}_X\omega = \mathrm{d}i_X\omega + i_X\mathrm{d}\omega$$

implies that the one-form $i_{X^\xi}\omega$ is closed for any $\xi \in \mathfrak{g}$. The additional assumption we have to make is that each of these one-forms is also exact, i.e. can be written in the form

$$i_{X^\xi}\omega = \mathrm{d}H^\xi,$$

for a smooth function $H^\xi \in C^\infty(M)$, called the *Hamiltonian associated to* ξ. Moreover, we require the existence of a simultaneous equivariant choice of these Hamiltonians, i.e. there should exist a smooth *moment map*,

$$\mu : M \to \mathfrak{g}^*,$$

such that

$$\mu(p)(\xi) = H^\xi(p),$$

which is *equivariant* with respect to the action of G on M and the coadjoint action of G on $\mathfrak{g}^*$. Such a moment map always exists if, for example, G is compact and M is simply connected. In general, the obstruction for the existence of a moment map may be phrased in terms of the Lie algebra cohomology of $\mathfrak{g}$, see [13] for details.

We phrase the result as the following theorem.

Theorem 1.1.1 (Marsden-Weinstein Reduction). *If G acts freely on $N = \mu^{-1}(c)$, where c lies in the centre of $\mathfrak{g}^*$, then*

$$N/G = \mu^{-1}(c)/G$$

is in a natural way a symplectic manifold.

Proof. Let $c \in Z(\mathfrak{g}^*)$ be a central element. By equivariance of μ, G acts on the level set $\mu^{-1}(c)$ and we assume this action is free. The crucial observation is that for $p \in \mu^{-1}(c)$

$$\mathrm{im}(\mathrm{d}\mu)_p = \mathfrak{g}_p^0 = \{\alpha \in \mathfrak{g}^* \mid \alpha(\xi) = 0 \text{ for all } \xi \in \mathfrak{g}_p\}.$$

This can be proved as follows. First, we note the formula

$$(\mathrm{d}\mu)_p(Y)(\xi) = (\mathrm{d}\mu(\xi))_p(Y) = \omega_p(X_p^\xi, Y).$$

Moreover, we already know that the evaluation map ev_p fits into the following exact sequence

$$0 \to \mathfrak{g}_p \to \mathfrak{g} \to T_p\mathcal{O}_p \to 0,$$

where the last map is given by $\xi \mapsto \mathrm{ev}_p(X^\xi)$, which is surjective onto $T_p\mathcal{O}_p$. Hence, from the above formula for $(\mathrm{d}\mu)_p$ we see that

$$\ker(\mathrm{d}\mu)_p = \{Y \in T_pM \mid \omega(X^\xi, Y) = 0 \text{ for all } \xi \in \mathfrak{g}\} = (T_p\mathcal{O}_p)^{\omega_p}.$$

Here we use the notation $(T_p\mathcal{O}_p)^{\omega_p}$ to denote the symplectic complement of $T_p\mathcal{O}_p$ inside the symplectic vector space (T_pM, ω_p), i.e. $(T_p\mathcal{O}_p)^{\omega_p} = \{Y \in T_pM \mid \omega(X, Y) = 0 \text{ for all } X \in T_p\mathcal{O}_p\}$. Now again using the expression for $(\mathrm{d}\mu)_p$, we observe that $\mathrm{im}(\mathrm{d}\mu)_p$ must be contained in $\mathfrak{g}_p^0$, since $\mathfrak{g}_p = \{\xi \in \mathfrak{g} \mid X^\xi = 0\}$. Now we count dimensions to conclude that the two spaces are equal.

In our case, G acts freely on $\mu^{-1}(c)$. Hence, at $p \in \mu^{-1}(c)$ we get that $\mathfrak{g}_p = 0$, which means that $(\mathrm{d}\mu)_p$ is surjective. Thus, c is a regular value of μ, and so the level set

$$N = \mu^{-1}(c) \subset M$$

is a smooth submanifold of dimension $\dim M - \dim G$.

In particular, as G acts on $\mu^{-1}(c)$, the fundamental vector fields are tangent to $\mu^{-1}(c)$ and therefore lie in the kernel of $\mathrm{d}\mu$. Hence,

$$T_p\mathcal{O}_p \subset (T_p\mathcal{O}_p)^{\omega_p},$$

i.e. $T_p\mathcal{O}_p$ is isotropic. Since G acts freely on $\mu^{-1}(c)$, the orbit space $\mu^{-1}(c)/G$ is a smooth manifold of dimension $\dim M - 2\dim G$ with tangent space

$$T_{G.p}(\mu^{-1}(c)/G) \cong \ker(\mathrm{d}\mu)_p / T_p\mathcal{O}_p = (T_p\mathcal{O}_p)^{\omega_p} / T_p\mathcal{O}_p,$$

where the first identification is given by the derivative of the canonical projection $(\mathrm{d}\pi)_p : T_p(\mu^{-1}(c)) \to T_{G.p}(\mu^{-1}(c)/G)$. We claim that the restriction of ω_p to $(T\mathcal{O}_p)^{\omega_p}$ induces a 2-form $\tilde{\omega}$ on this quotient, given by

$$\tilde{\omega}_{\pi(p)}([X], [Y]) = \omega_p(X, Y).$$

In other words

$$\pi^*\tilde{\omega} = i^*\omega,$$

where $i : \mu^{-1}(c) \to M$ is the inclusion map. We check that this indeed defines a symplectic form as desired. First, we show that it is well-defined. Let $X + A, Y + B$, $A, B \in T_p\mathcal{O}_p$ be two different representatives of $[X], [Y]$. Then

$$\omega_p(X + A, Y + B) = \omega_p(X, Y) + \omega_p(X, B) + \omega_p(A, Y) + \omega_p(A, B).$$

The two terms in the middle vanish, as $X, Y \in (T_p\mathcal{O}_p)^{\omega_p}$ and the last one is zero, as $T_p\mathcal{O}_p$ is isotropic. Thus,

$$\omega_p(X + A, Y + B) = \tilde{\omega}_{\pi(p)}([X], [Y]),$$

and since G acts by symplectomorphisms, an analogous argument as above in the Riemannian case shows that this does not depend on the choice of point in $\pi^{-1}(p)$. For the non-degeneracy assume we are given $[X]$ such that

$$\tilde{\omega}_{\pi(p)}([X], [Y]) = 0 \qquad \text{for all } [Y].$$

This means

$$\omega_p(X, Y) = 0 \qquad \text{for all } Y \in (T_p\mathcal{O}_p)^{\omega_p},$$

i.e. $X \in ((T_p\mathcal{O}_p)^{\omega_p})^{\omega_p} = T_p\mathcal{O}_p$, which is precisely the statement

$$[X] = 0.$$

To see that $\tilde{\omega}$ is closed, we compute

$$\pi^*(\mathrm{d}\tilde{\omega}) = \mathrm{d}\pi^*\tilde{\omega} = \mathrm{d}i^*\omega = i^*\mathrm{d}\omega = 0.$$

Therefore, since π is a submersion, i.e. π^* is injective on forms, $\tilde{\omega}$ has to be closed. □

Remark. The point of the above proof is to show that

$$\ker(i^*\omega)_p \subset T_p\mathcal{O}_p.$$

So on the quotient the induced form $\tilde{\omega}$ is non-degenerate. This observation will become useful later, when we discuss hypersymplectic quotients.

Remark. This theorem carries over to Banach manifolds, and in this context we will use it when we look at moduli spaces of gauge theoretic equations. However, in the above proof we implicitly use the finite-dimensionality of the manifold to prove that the differential of μ is surjective for all $p \in \mu^{-1}(c)$. So in the statement of the theorem for Banach manifolds, we have to include the assumption that $\mathrm{d}\mu_p$ is surjective for all $p \in \mu^{-1}(c)$. See [36] for a complete proof and discussion, and also [40] for facts about Banach manifolds.

1.1.3 The Kähler Quotient

A Riemannian manifold (M, g) is *Kähler* if there exists a complex structure $I \in \Gamma(\mathrm{End}(TM))$, which is parallel with respect to the Levi-Civita connection ∇^g and compatible with the metric g, that is

$$\nabla^g I = 0 \qquad \text{and} \qquad g(IX, IY) = g(X, Y) \text{ for all } X, Y \in \Gamma(TM).$$

These two conditions imply that the associated two-form

$$\omega = g(I-, -) \in \Omega^2(M)$$

is closed and non-degenerate, in particular (M, ω) is a *symplectic manifold* ([9]), and ω is called the associated *Kähler form*.

Example. The easiest example of a Kähler manifold is $\mathbb{C}^n$ with its standard hermitian inner product coming from the Euclidean inner product under the isomorphism $\mathbb{R}^{2n} \cong \mathbb{C}^n$. More generally, we could take any complex Hilbert space of possibly infinite dimension, see the chapter on gauge theory. For details about exterior differentiation on Banach manifolds, see [40].

Example. Complex submanifolds of Kähler manifolds are Kähler.

Example. Complex projective space $\mathbb{CP}^n$ is a Kähler manifold with the *Fubini-Study metric*. This can be interpreted as a Kähler quotient, as we will see at the end of this chapter. Therefore, projective manifolds, i.e. those complex manifolds that can be realised as complex submanifolds of $\mathbb{CP}^n$ for some n, are Kähler.

Consider now a compact Lie group G acting freely on M preserving the Kähler structure (g, ω, I) with moment map μ. We claim that in this situation the quotient metric on the symplectic quotient $\mu^{-1}(0)/G$ is a Kähler metric with respect to the induced symplectic form and the complex structure $\tilde{I} = \pi_* I$, where again $\pi : \mu^{-1}(0) \to \mu^{-1}(0)/G$ is the natural projection. Let $i : \mu^{-1}(0) \to M$ be the inclusion map. Then we get a metric h on $\mu^{-1}(0)$ by pullback, i.e.

$$h = i^* g.$$

On $TM|_{\mu^{-1}(0)} = i^*TM$ the Levi-Civita connection induces a connection given by $i^*\nabla^g$, which we will again denote by ∇^g. The Levi-Civita connection ∇^h on $\mu^{-1}(0)$ is then related to the Levi-Civita connection ∇^g on M via

$$\nabla^h_X Y = \mathrm{pr}_{T(\mu^{-1}(0))}(\nabla^g_X Y),$$

where

$$\mathrm{pr}_{T(\mu^{-1}(0))} : TM|_{\mu^{-1}(0)} \cong T(\mu^{-1}(0)) \oplus \mathcal{N} \to T(\mu^{-1}(0))$$

is the orthogonal projection. Here we write $\mathcal{N}$ for the normal bundle associated to the submanifold $\mu^{-1}(0)$. On the quotient, we form the quotient metric $\tilde{h}$ described earlier. Recall also the decomposition

$$T_p(\mu^{-1}(0)) \cong \mathfrak{g} \oplus H_p$$

and that the natural projection gives an isomorphism $\mathrm{d}\pi : H \to T(\mu^{-1}(0)/G)$ covering π, such that $\tilde{h} = \pi_* h$. Our aim now is to exhibit the Levi-Civita connection $\nabla^{\tilde{h}}$ on $\mu^{-1}(0)/G$ and to show that the quotient metric is Kähler with respect to the complex structure $\tilde{I} = \pi_* I$. First of all, we check that this definition of the induced complex structure makes sense, i.e. we have to convince ourselves that I preserves the decomposition $TM|_{\mu^{-1}(0)} = (\mathcal{N} \oplus \mathfrak{g}) \oplus H$. The normal bundle of $\mu^{-1}(0) \subset M$ is spanned by the vectors

$$\mathrm{grad}_{\mathrm{g}}(\mu_i) \qquad i = 1, \ldots, \dim G,$$

here μ_i are the component functions of $\mu : M \to \mathfrak{g}^* \cong \mathbb{R}^{\dim G}$, where we identified $\mathfrak{g}$ with $\mathbb{R}^{\dim G}$ by choosing some basis $\{\xi_1, \ldots, \xi_{\dim G}\}$. In other words:

$$\mu_i(p) = \mu(p)(\xi_i) = H^{\xi_i}(p).$$

Moreover, the fundamental vector fields X^{ξ_i} yield an explicit identification

$$T_p\mathcal{O}_p \cong \mathfrak{g}.$$

So the orthogonal complement of H in TM is spanned by $\{\mathrm{grad}(\mu_i), X^{\xi_i} \mid i = 1, \ldots, \dim G\}$. Now for any tangent vector Y we calculate

$$g(\mathrm{grad}(\mu_i), Y) = d\mu_i(Y) = \omega(X^{\xi_i}, Y) = g(IX^{\xi_i}, Y),$$

which means

$$\mathrm{grad}\mu_i = IX^{\xi_i}.$$

Thus, the orthogonal complement of H is in fact a complex vector space, hence H itself has to be a complex vector space. H being complex means that the orthogonal projection $\mathrm{pr}_H : TM|_{\mu^{-1}(0)} \cong H \oplus \mathfrak{g} \oplus N \to H$ is complex linear, therefore $I|_H$ is parallel with respect to the connection $\tilde{\nabla}^h$ on H given by

$$\tilde{\nabla}^h_X Y = \mathrm{pr}_H(\nabla^h_X Y).$$

We check that this connection induces the Levi-Civita connection $\nabla^{\tilde{h}}$ on the quotient $\mu^{-1}(0)/G$. First of all, $\tilde{\nabla}^h$ preserves the metric h on H. To see this, let $X, Y, Z \in \Gamma(H)$ be horizontal vector fields, then

$$Xh(Y,Z) = h(\nabla^h_X Y, Z) + h(Y, \nabla^h_X Z) = h(\mathrm{pr}_H(\nabla^h_X Y), Z) + h(Y, \mathrm{pr}_H(\nabla^h_X Z)).$$

The second equality holds, since all vectors orthogonal to H will be sent to zero by $h(-, Z)$ and $h(Y, -)$. Thus, $\tilde{\nabla}^h$ preserves h on H. Now we check that it is also torsion-free for horizontal vector fields. In order to verify this, we remark that a vector field $X \in \Gamma(T(\mu^{-1}(0)/G))$ and its horizontal lift $\tilde{X} \in \Gamma(H)$ are by definition *π-related*, that is

$$\mathrm{d}\pi(\tilde{X}) = X \circ \pi.$$

It follows from general theory, that given vector fields $X, Y \in \Gamma(T(\mu^{-1}(0)/G))$ with horizontal lifts $\tilde{X}, \tilde{Y} \in \Gamma(H)$, their Lie brackets are also π-related,

$$\mathrm{d}\pi([\tilde{X}, \tilde{Y}]) = [X, Y] \circ \pi,$$

that is

$$\widetilde{[X,Y]} = \mathrm{pr}_H([\tilde{X}, \tilde{Y}]).$$

With this, we easily get for horizontal vector fields $X, Y \in \Gamma(H)$

$$\tilde{\nabla}^h_X Y - \tilde{\nabla}^h_Y X = \mathrm{pr}_H(\nabla^h_X Y - \nabla^h_Y X) = \mathrm{pr}_H([X, Y]),$$

which is precisely the statement that

$$\nabla^{\tilde{h}} = \mathrm{d}\pi \circ \tilde{\nabla}^h \circ (\mathrm{d}\pi)^{-1}$$

will be torsion-free. Altogether, the above discussion proves the following theorem.

Theorem 1.1.2. *Suppose a Lie group G acts on the Kähler manifold (M, ω, I) preserving the Kähler structure with moment map μ. If G acts freely and properly on the level set $\mu^{-1}(c)$ for an element c in the centre of $\mathfrak{g}^*$, then*

$$\mu^{-1}(c)/G$$

is naturally a Kähler manifold.

We remark that, as in the symplectic case, this theorem carries over to Banach manifolds under the assumptions that 0 is a regular value of μ, and that $\mu^{-1}(0)/G$ is a manifold. We illustrate this theorem by considering a simple example.

Example. Consider $\mathbb{C}^{n+1}$ with its standard Euclidean inner product and Kähler form given by

$$\omega = \sum_i \mathrm{d}z_i \wedge \mathrm{d}\bar{z}_i.$$

Let $G = \mathrm{U}(1) \subset \mathbb{C}$ act by scalar multiplication. This action cleary preserves the metric and the Kähler form ω. The Lie algebra of U(1) is $i\mathbb{R} \cong \mathbb{R}$ and the fundamental vector fields of this action are given by

$$X^\xi(z) = \frac{\mathrm{d}}{\mathrm{d}t}|_{t=0} e^{it\xi} z = i\xi z \frac{\partial}{\partial z} - i\xi \bar{z} \frac{\partial}{\partial \bar{z}} = i\xi z.$$

The moment map for the action has to satisfy

$$\mathrm{d}\mu_z(X)(i\xi) = \omega(X^\xi, X) = \omega(i\xi z, X),$$

and it is easy to show that it is given by

$$\mu(z) = \frac{i}{2}|z|^2 + ic,$$

for an arbitrary real constant c, since U(1) is abelian. Thus, the level set for $c = -\frac{1}{2}$ is given by the sphere $S^{2n+1} \subset \mathbb{C}^{n+1}$. Thus, the quotient can be identified with $\mathbb{CP}^n$ and the induced metric turns out to be the Fubini-Study metric.

1.1.4 The Hyperkähler Quotient

In this section, we give a short introduction to the vast subject of hyperkähler geometry. We will discuss basic properties of hyperkähler manifolds from different viewpoints. Finally, we prove a theorem analogous to the Kähler quotient discussed above: the hyperkähler quotient construction, which was introduced in [31].

There are various ways to look at hyperkähler manifolds. We give here the definition, which will prove most useful later on, when we will do gauge theory.

Definition 1.1.1. A Riemannian manifold (M, g) is *hyperkähler* if there exist three skew adjoint endomorphisms of the tangent bundle $I, J, K \in \Gamma(\mathrm{End}(TM))$, such that

- I, J, K are all parallel with respect to the Levi-Civita connection ∇^g on M, so

$$\nabla^g I = \nabla^g J = \nabla^g K = 0,$$

- I, J, K obey the algebraic identities of the generators of the imaginary quaternions, i.e.

$$IJ = K = -JI, \qquad I^2 = J^2 = K^2 = -\mathrm{id}_{TM}.$$

In terms of holonomy groups these conditions say that parallel transport by ∇^g is a quaternionic linear isometry, that is, the holonomy is contained in $\mathrm{Sp}(k)$. We see that this implies that the tangent space at each point in M is a quaternionic vector space and hence the dimension of M is a multiple of 4, $\dim M = 4k$.

The first condition means that the metric is Kähler with respect to each of the complex structures I, J, K. In particular the three non-degenerate, in fact positive, two-forms

$$\omega_I = g(I-,-), \qquad \omega_J = g(J-,-), \qquad \omega_K = g(K-,-)$$

are all parallel and hence closed. Each of them gives M the structure of a symplectic manifold. M is thus a Kähler manifold in many different ways. Indeed, we have a whole family of complex structures parametrised by the two-sphere

$$S^2 \subset \mathbb{R}^3 \cong \mathrm{Im}(\mathbb{H}),$$

for which the metric is Kähler. That is, for each $x = (x_1, x_2, x_3) \subset S^2$ we get a complex structure

$$I_x = x_1 I + x_2 J + x_3 K.$$

The relations satisfied by I, J, K immediately imply that I_x squares to $-\mathrm{id}_{TM}$. Moreover, from its definition as a linear combination of parallel endomorphisms it is clearly parallel.

Example. The basic example of a hyperkähler manifold is a quaternionic vector space endowed with a flat Euclidean metric compatible with i, j, k, e.g. the quaternionic left-module $\mathbb{H}^n$. The complex structures I, J, K are given by $-i, j, k$ acting on the right. These are $\mathbb{H}$-linear, since scalar multiplication acts on the left.

Proposition 1.1.3. *Let (M^{4k}, g, I, J, K) be a hyperkähler manifold. Then the form $\omega_I^{\mathbb{C}} = \omega_J + i\omega_K$ is a non-degenerate holomorphic, in fact parallel, $(2,0)$-form with respect to the complex structure I. Thus, every hyperkähler manifold is naturally a complex symplectic manifold. By cyclically permuting I, J, K in the above formula, we obtain analogous forms $\omega_J^{\mathbb{C}}$ and $\omega_K^{\mathbb{C}}$.*

Proof. We only give the proof for $\omega_I^{\mathbb{C}}$. Let $Y \in \Gamma(TM \otimes \mathbb{C})$ such that $IY = -iY$, i.e. $Y \in \Gamma(TM^{0,1})$ with respect to I. Then if $X \in \Gamma(TM \otimes \mathbb{C})$ is any complex vector field, we compute using the algebraic relations satisfied by I, J, K,

$$\begin{aligned} \omega_I^{\mathbb{C}}(Y,X) &= \omega_J(Y,X) + i\omega_K(Y,X) \\ &= g(JY,X) + ig(IJY,X) \\ &= g(JY,X) - ig(JIY,X) \\ &= g(JY,X) - g(JY,X) \\ &= 0. \end{aligned}$$

Thus, $\omega_I^{\mathbb{C}}$ is of type $(2,0)$. Since it is parallel, it is holomorphic. The non-degeneracy is easy to check by writing any $(1,0)$-vector field in the form $X - iIX$, where $X \in \Gamma(TM \otimes \mathbb{C})$. □

Corollary 1.1.4. *Hyperkähler manifolds are Ricci-flat.*

Proof. We have just seen that the canonical bundle $K_M = \Omega^{2k,0}(M)$ admits a non-vanishing parallel global section given by the *complex volume form* $\Omega = (\omega_I^{\mathbb{C}})^k$. Thus, the connection on the canonical bundle induced by the Levi-Civita connection is trivial and so in particular flat. Therefore, the corollary follows from the observation that the curvature of this connection is given by the *Ricci-form* $\rho = \mathrm{Ric}(I-,-)$, which is proven for example in [9]. □

We see that hyperkähler manifolds are examples of Ricci-flat Kähler manifolds, i.e. they are *Calabi-Yau*. This can also be shown by considering holonomy groups and the inclusion $\mathrm{Sp}(k) \subset \mathrm{SU}(2k)$.

Hyperkähler Four-Manifolds

In dimension four, Calabi-Yau manifolds are the same as hyperkähler manifolds, because $\mathrm{SU}(2) = \mathrm{Sp}(1)$. The only compact examples in four dimensions are Ricci-flat complex tori T^2 and $K3$ surfaces. We now discuss some properties of hyperkähler four-manifolds, which will be important when we do gauge theory on them later on.

Firstly, observe that the Hodge star operator on two-forms $* : \Lambda^2 T^*M \to \Lambda^2 T^*M$ satisfies $*^2 = \mathrm{id}$. So we get a splitting into three-dimensional subbundles

$$\Lambda^2 T^*M = \Lambda^+ \oplus \Lambda^-,$$

given by the ± 1-eigenspaces of $*$. The main observation we want to note is the following lemma.

Lemma 1.1.5. *At any point $p \in M$ the space Λ_p^+ is spanned by $\omega_i(p)$ for $i \in I, J, K$.*

Thus, for $p \in M$, the two-sphere of compatible complex structures on T_pM can be identified with the two-sphere in the three-dimensional space Λ^+_p.

Proof. Pick a holomorphic chart $(z_1 = x_1 + iy_1, z_2 = x_2 + iy_2)$ for I about $p \in M$, such that the metric osculates to order two from the standard metric on $\mathbb{C}^2$. In particular, T_pM is identified with standard $\mathbb{C}^2$ with its hermitian inner product and ω_I corresponds to the standard Kähler form

$$\omega_I = -\frac{i}{2}(\mathrm{d}z_1 \wedge d\bar{z}_1 + \mathrm{d}z_2 + \mathrm{d}\bar{z}_2) = \mathrm{d}x_1 \wedge \mathrm{d}y_1 + \mathrm{d}x_2 \wedge \mathrm{d}y_2.$$

But in this situation the claim is easy to verify. Also the space $\Lambda^{2,0}(M)_p$ is spanned just by

$$\omega_I^{\mathbb{C}} = \mathrm{d}z_1 \wedge \mathrm{d}z_2 = \mathrm{d}x_1 \wedge \mathrm{d}x_2 - \mathrm{d}y_1 \wedge \mathrm{d}y_2 + i(\mathrm{d}x_1 \wedge \mathrm{d}y_2 + \mathrm{d}y_1 \wedge \mathrm{d}x_2).$$

Now the real and the imaginary part of this give the other two Kähler forms ω_J and ω_K and it is easy to check that they are indeed self-dual. □

Note that moreover $\omega_i \wedge *\omega_j = \omega_i \wedge \omega_j = 0$ if $i \neq j$. Hence, the basis $\{\omega_I, \omega_J, \omega_K\}$ is orthogonal. If we now complexify, we see that point-wise $\Lambda^{2,0}$ and $\Lambda^{0,2}$ are spanned by $\omega_J + i\omega_K$ and $\omega_J - i\omega_K$, respectively. Thus,

$$\Lambda^+ \otimes \mathbb{C} = \Lambda^{2,0} \oplus \Lambda^{0,2} \oplus \mathbb{C}\omega_I.$$

This implies that

$$\Lambda^- \otimes \mathbb{C} = \Lambda^{1,1} \cap (\omega_I^\perp) \otimes \mathbb{C}.$$

In particular, a two-form is anti-self-dual if and only if it is of type $(1,1)$ with respect to any of the complex structures.

The Hyperkähler quotient

Going back to general hyperkähler manifolds, we now prove a theorem due to Hitchin et. al [31], which provides a hyperkähler analogue of the Marsden-Weinstein symplectic quotient. Let G be a Lie group acting on a hyperkähler manifold (M^n, g, I, J, K) preserving the hyperkähler structure. That is, G acts by isometries and by symplectomorphisms with respect to all three Kähler forms. We assume that this action admits a moment map with respect to each of the three Kähler forms and we group these together to get a vector-valued *hyperkähler moment map*

$$\mu = (\mu_I, \mu_J, \mu_K) : M \to \mathfrak{g}^* \otimes \mathbb{R}^3.$$

Then we can prove the following theorem.

Theorem 1.1.6 ([31]). *Let (M, g, I, J, K) be a hyperkähler manifold and let G be a compact Lie group acting on M preserving the hyperkähler structure with hyperkähler moment map $\mu : M \to \mathfrak{g}^* \otimes \mathbb{R}^3$. Then the quotient metric on*

$$\mu^{-1}(c_1, c_2, c_3)/G = (\mu_I^{-1}(c_1) \cap \mu_J^{-1}(c_2) \cap \mu_K^{-1}(c_3))/G$$

is hyperkähler if all the c_i lie in the centre of $\mathfrak{g}^*$ and if G acts freely and properly on $\mu^{-1}(c)$.

Proof. Our proof is slightly different from the one given in [31], making it easier to compare the hyperkähler quotient construction to its hypersymplectic analogue, which we will discuss later.

Firstly, we check that $c = (c_1, c_2, c_3)$ is a regular value of μ if G acts freely on the level set. The image of $\mathrm{d}\mu$ inside $\mathfrak{g}^* \otimes \mathbb{R}^3$ is given by

$$\{\alpha \in \mathfrak{g}^* \otimes \mathbb{R}^3 | \; \exists X \in T_pM : \alpha(\xi) = (\omega_I(X^\xi, X), \omega_J(X^\xi, X), \omega_K(X^\xi, X)) \; \forall \xi \in \mathfrak{g}\}.$$

So if $\mathfrak{g}_p$ is the Lie algebra of the stabiliser of $p \in M$, we see immediately that

$$\mathrm{im}(\mathrm{d}\mu_p) \subset \mathfrak{g}_p^0 \otimes \mathbb{R}^3,$$

where $\mathfrak{g}_p^0 \subset \mathfrak{g}^*$ is the annihilator of $\mathfrak{g}_p$.

On the other hand, the kernel of $\mathrm{d}\mu$ is given by

$$\begin{aligned} \ker(\mathrm{d}\mu_p) &= \{X \in T_pM \mid \omega_i(X^\xi, X) = 0 \text{ for all } \xi \in \mathfrak{g}, \; i \in \{I, J, K\}\} \\ &= \{X \in T_pM \mid 0 = g(IX^\xi, X) = g(JX^\xi, X) = g(KX^\xi, X) \text{for all} \xi \in \mathfrak{g}\} \\ &= [I(T_p\mathcal{O}_p) + J(T_p\mathcal{O}_p) + K(T_p\mathcal{O}_p)]^\perp. \end{aligned}$$

Since G acts on $\mu^{-1}(c)$, $T_p\mathcal{O}_p$ is tangent to this level set and hence in the kernel of $\mathrm{d}\mu_p$. This means that $T_p\mathcal{O}_p$ is orthogonal to each of the spaces $I(T_p\mathcal{O}_p), J(T_p\mathcal{O}_p), K(T_p\mathcal{O}_p)$. But since all the complex structures are compatible with the metric, these spaces are in fact all mutually perpendicular. We briefly explain this by showing that $J(T_p\mathcal{O}_p), K(T_p\mathcal{O}_p)$ are orthogonal. The other cases are analogous. Let $JX \in J(T_p\mathcal{O}_p)$ and $KY \in K(T_p\mathcal{O}_p)$. Then, since $T_p\mathcal{O}_p, I(T_p\mathcal{O}_p)$ are perpendicular,

$$g(JX, KY) = -g(JX, JIY) = -g(X, IY) = 0.$$

If G acts freely on the level set of the moment map, then $\mathfrak{g}_p^0 = \mathfrak{g} \otimes \mathbb{R}^3$ for each $p \in \mu^{-1}(c)$. Moreover, since all the spaces that occur in the above sum are pairwise orthogonal, it follows from the positivity of the metric that their sum is direct. So on $\mu^{-1}(c)$,

$$\dim \ker \mathrm{d}\mu = \dim M - 3 \dim G.$$

Hence, by counting dimensions, we see that on $\mu^{-1}(c)$

$$\dim \mathrm{im}(\mathrm{d}\mu) = 3 \dim G = \dim \mathfrak{g} \otimes \mathbb{R}^3.$$

Thus, c is a regular value of μ and hence $\mu^{-1}(c)$ is a well-defined submanifold of M with tangent space $\ker \mathrm{d}\mu = [I(T_p\mathcal{O}_p) + J(T_p\mathcal{O}_p) + K(T_p\mathcal{O}_p)]^\perp$. The tangent space of the quotient $\mu^{-1}(0)/G$ at a point $G.p$ will be given by $[(T_p\mathcal{O}_p) + I(T_p\mathcal{O}_p) + J(T_p\mathcal{O}_p) + K(T_p\mathcal{O}_p)]^\perp$ and is hence again quaternionic. In particular, the orthogonal projection onto this subspace is quaternionic linear. Then we show, like in the proof of the Kähler quotient, that the Levi-Civita connection on the quotient is given by composition with the orthogonal projection onto this tangent space. Hence, the induced complex structures on the quotient are all parallel and so the quotient metric is Kähler with respect to each of them, that is, hyperkähler. □

An alternative way of proving this theorem is given by observing that the complex moment map $\mu_I^{\mathbb{C}} = \mu_J + i\mu_K : M \to \mathfrak{g}^* \otimes \mathbb{C}$ is holomorphic with respect to the complex structure I and has $c_2 + ic_3$ as a regular value under the above assumptions. Thus, $\tilde{M} = (\mu_I^{\mathbb{C}})^{-1}(c)$ is a Kähler manifold on which G acts preserving the Kähler structure. The moment map with respect to this action is then just the moment map μ_I restricted to $\tilde{M}$. Therefore, the quotient

$$(\mu_I^{-1}(c_1) \cap \tilde{M})/G = \mu^{-1}(c)/G$$

is Kähler with respect to the complex structure I. Applying the same argument to the other complex structures by cyclically permuting I, J, K, we get another proof of the theorem.

So we have reduced the proof of the hyperkähler quotient to that of the Kähler quotient. In particular, if we again assume that c is a regular value of the hyperkähler moment map, the theorem carries over to Banach manifolds.

This alternative proof gives another interpretation of the hyperkähler quotient. Fix a compatible complex structure I. Assume that G is compact and that the G-action extends to an action of the complexification $G^{\mathbb{C}}$. This action will preserve the complex symplectic form $\omega_I^{\mathbb{C}}$ and we observe that $\mu_I^{\mathbb{C}}$ is then a holomorphic moment map for the $G^{\mathbb{C}}$-action. We have seen in the proof that for $c \in Z(\mathfrak{g}^*) \otimes \mathbb{C}$ the level sets $(\mu_I^{\mathbb{C}})^{-1}(c)$ are Kähler manifolds on which $G^{\mathbb{C}}$ acts holomorphically. Then we may identify the quotient $(\mu_I^{\mathbb{C}})^{-1}(c)^{stable,c_1}/G^{\mathbb{C}}$ with the Kähler quotient $((\mu_I^{\mathbb{C}})^{-1}(c) \cap \mu_I^{-1}(c_1))/G$, where $(\mu_I^{\mathbb{C}})^{-1}(c)^{stable,c_1}$ is the set of stable points, i.e. points p such that $\mu_I^{-1}(c_1) \cap G^{\mathbb{C}}.p \neq \emptyset$. This follows from the work of Frances Kirwan [35]. In particular, if $(\mu_I^{\mathbb{C}})^{-1}(c)$ is projective and if a suitable linearisation (depending on the real level $c_1 \in Z(\mathfrak{g}^*)$) of the $G^{\mathbb{C}}$-action is chosen, her result identifies this complex quotient with (an open dense set inside) the GIT quotient associated to this linearisation. In this sense, if we fix one of the complex structures, and view M as a complex symplectic manifold, the hyperkähler quotient may be thought of as the symplectic quotient in the holomorphic category. We will discuss Kirwan's work in more detail in section 2.3.

1.2 Gauge Theory

This chapter provides an introduction to gauge theory. Moduli spaces of solutions of equations coming from the anti-self-dual (ASD) Yang-Mills equations are an interesting source of examples of hyperkähler manifolds. The hyperkähler structures are obtained by viewing these moduli spaces as hyperkähler quotients in an infinite-dimensional setting. We will start by describing briefly the origin of the equations. Then we introduce the analytical technology involved in the proof of our main result: The ASD moduli space on a unitary bundle over a compact hyperkähler four-manifold is again hyperkähler. This result is of course by no means new, but it seems that there is no self-contained proof available in the literature.

Under suitable decay assumptions this also applies to the non-compact hyperkähler manifold $\mathbb{R}^4 = \mathbb{H}$, yielding a hyperkähler structure on the moduli space of *framed* instantons, see [43]. Considering dimensional reductions of the ASD equations and their moduli spaces, one can produce interesting examples of hyperkähler manifolds and also hyperkähler metrics on well-known complex symplectic manifolds. In this way it was shown by Kronheimer in [38] that the cotangent bundle of a complex Lie group admits a hyperkähler metric and that there exist hyperkähler metrics on semi-simple and nilpotent coadjoint orbits of complex Lie groups [39]. Later these results have been extended to general orbits, see [10], [37]. See also [18] for a survey of these results.

1.2.1 The Anti-Self-Dual Yang-Mills Equations

Let M^4 be a compact oriented Riemannian four-manifold, and let $E \to M$ be a complex vector bundle with a hermitian metric h. In other words E has structure group $\mathrm{U}(n)$, where n is the rank of E. Let ∇ be a connection on E such that $\nabla h = 0$. We could consider vector bundles whose structure group is an arbitrary compact matrix group G, but since every compact Lie group can be embedded into $\mathrm{U}(n)$ for some n, we do not lose any generality if we just work with $\mathrm{U}(n)$. We call ∇ *anti-self-dual* if its curvature two-form $R^\nabla \in \Omega^2(M, \mathfrak{u}(E))$ is anti-self-dual. Here $\mathfrak{u}(E)$ is the subbundle of $\mathrm{End}(E)$ of bundle endomorphism which are skew adjoint with respect to h. So for a unitary connection the anti-self-duality condition is expressed in the non-linear partial differential equation

$$R^\nabla = - * R^\nabla.$$

Let now $u \in \Gamma(M, \mathrm{End}(E))$ be such that $u^* u = \mathrm{id}_E$, i.e. u is a unitary bundle automorphism of E. We will often write $u \in \Gamma(M, \mathrm{U}(E))$. Then we can use u to pull-back the connection ∇ to get a new unitary connection

$$u.\nabla = u^{-1} \circ \nabla \circ u = \nabla + u^{-1} \mathrm{d}^\nabla u,$$

where we denote by d^∇ the induced connection on $\mathrm{End}(E)$. We call u a *gauge transformation*. The curvature then transforms according to

$$R^{u.\nabla} = u^{-1} R^\nabla u,$$

and we see that the anti-self-duality is in fact a gauge invariant notion. If we denote by $\mathcal{A}$ the space of unitary connections and by $\mathcal{G}$ the *gauge group*, that is

$$\mathcal{G} = \{u \in \Gamma(M, \mathrm{End}(E)) \mid u^* u = \mathrm{id}_E\} = \Gamma(M, \mathrm{U}(E)),$$

then $\mathcal{G}$ acts on $\mathcal{A}$ preserving the ASD condition. Thus, the object we would like to understand is the *moduli space of ASD connections*:

$$\mathcal{M} = \{\nabla \in \mathcal{A} \mid R^\nabla = - * R^\nabla\} / \mathcal{G}.$$

In the following section we will discuss in some detail why $\mathcal{M}$ is in fact a finite dimensional manifold, which is hyperkähler if the underlying base-manifold M is hyperkähler. But before we are in a position to do so, we have to introduce the necessary analytical framework.

1.2.2 Analytical Tools

In this short paragraph, we will briefly write down the main analytical tools that will be involved in our discussion. These are Sobolev spaces on vector bundles, the Sobolev embedding theorems and the multiplication properties of Sobolev spaces. We also need the basic theorems for elliptic operators on compact manifolds. The reason why we wish to work with Sobolev spaces is that they are Banach spaces and therefore allow us to use the inverse and implicit function theorems. Using these techniques, we can show that an appropriate Sobolev completion of $\mathcal{G}$ is in fact a smooth Hilbert Lie group. Following this, we will consider more closely the action of the gauge group on an appropriate space $\mathcal{A}^*$ of irreducible unitary Sobolev connections. A good textbook which contains a lot of the analytical machinery needed in gauge theory is [25], see also [41] and [36].

Let $(E, h, \nabla) \to (M, g)$ be a hermitian vector bundle over a compact Riemannian manifold. For a given positive integer k we define the L^2_k-norm on $\Gamma(E)$ by

$$||s||_{L^2_k} = \left(\sum_{i=1}^{k} \int_M |\nabla^{(i)} \dots \nabla^{(2)} \nabla s|^2 \right)^{\frac{1}{2}},$$

where $\nabla^{(i)} : \Omega^{i-1}(M, E) \to \Omega^i(M, E)$ is the connection on E-valued alternating multilinear forms induced by ∇ on E and the Levi-Civita connection ∇^g on M. We denote by $L^2_k(E)$ the Hilbert space completion of $\Gamma(E)$ in the L^2_k norm. This is the space of sections of E whose first k derivatives lie in $L^2(E)$. In our situation where the base manifold is compact, we have the following theorem.

Theorem 1.2.1. *Let the base manifold M be compact, then $L^2_k(E)$ does not depend on the choice of ∇.*

Proof. We only do the case $k = 1$. Let ∇^1, ∇^2 be two unitary connections on (E, h). Then, since the space of unitary connections is an affine space modelled on $\Omega^1(M, \mathfrak{u}(E))$, we can find a $\mathfrak{u}(E)$-valued one-form A such that $\nabla^1 = \nabla^2 + A$. But then, using the Cauchy-Schwarz and Young's inequality:

$$\begin{aligned}
||s||^2_{L^2_1,\nabla^1} &= ||s||^2_{L^2} + ||\nabla^1 s||^2_{L^2} \\
&= ||s||^2_{L^2} + ||\nabla^2 s + As||^2_{L^2} \\
&\leq ||s||^2_{L^2} + ||\nabla^2 s||^2_{L^2} + ||A||^2_{L^2}||s||^2_{L^2} + 2||\nabla^2 s||_{L^2}||A||_{L^2}||s||_{L^2} \\
&\leq ||s||^2_{L^2} + ||\nabla^2 s||^2_{L^2} + ||A||^2_{L^2}||s||^2_{L^2} + ||\nabla^2 s||^2_{L^2} + ||A||^2_{L^2}||s||^2_{L^2} \\
&\leq C(||s||^2_{L^2} + ||\nabla^2 s||^2_{L^2}) \\
&= C||s||^2_{L^2_1,\nabla^2}
\end{aligned}$$

The same argument works with ∇^1 and ∇^2 interchanged. Therefore, ∇_1 and ∇_2 yield equivalent norms on $L^2_k(E)$. □

More generally, one could define spaces L^p_k by replacing the L^2-norm above with the L^p norm and letting k be any real number. But this is a direction we will not pursue here. We now come to the Sobolev embedding and multiplication theorems which we will use heavily in our proofs later on. For proofs we refer to [41].

Theorem 1.2.2 (Sobolev Embedding Theorem)**.** *Let M be a compact Riemannian n-manifold and let $E \to M$ be a hermitian vector bundle endowed with a unitary connection ∇. Then there are continuous embeddings*

- $L^2_k(E) \subset L^2_l(E)$, *if $k > l$ and this embedding is* compact. *This is known as* Rellich's Lemma.
- *More generally, there is a bounded embedding $L^p_k(E) \to L^q_l(E)$, whenever $k > l$ and $k - n/p \geq l - n/q$.*
- *Also $L^2_k(E) \subset C^l(E)$, $k > l + n/2$.*

Theorem 1.2.3 (Sobolev Multiplication Theorem)**.** *If $\dim M = 4$ and $k > 2$, or if $\dim M = 2$ and $k > 1$, then $L^2_k(\mathrm{End}(E))$ is an algebra and if $k > l$, then $L^2_k(E)$ and $L^2_k(\mathrm{End}(E))$ are $L^2_l(\mathrm{End}(E))$-modules.*

A differential operator D of order m acting on sections of E gives rise to a linear map $L^2_k(E) \to L^2_{k-m}(E)$. We call D *elliptic*, if its principal symbol $\sigma(D)(\xi, p) : E_p \to E_p$ is an isomorphism for each $p \in M$. We have the following fundamental theorem.

Theorem 1.2.4. *Let M be compact and let $E \to M$ be a hermitian vector bundle. Let D be an elliptic differential operator of order $m > 0$ acting on $\Gamma(E)$. Then*

- $\ker D$ *and* $\ker D^*$ *are finite dimensional and consist of smooth sections, where D^* is the formal adjoint of D.*
- *The equation $Ds = t$, where $t \in L^2_k(E)$ has a solution if and only if t is L^2-orthogonal to the kernel of D^*. Moreover, there is a constant C_k such that $||s||_{L^2_k} \leq C_k(||u||_{L^2_{k-m}} + ||Ds||_{L^2_{k-m}})$, in particular, if t is smooth, then so is s.*
- *If D is self-adjoint, we have the following spectral decomposition:*

$$L^2(E) = \bigoplus_{\lambda} \mathrm{Eig}(D, \lambda),$$

where $\lambda \in \mathbb{R}$ are the eigenvalues of D with eigenspaces $\mathrm{Eig}(D, \lambda)$*. The set of eigenvalues is discrete in $\mathbb{R}$, in particular we can always find a smallest positive eigenvalue.*

1.2.3 The Space of Unitary Sobolev Connections

The space $\mathcal{A}$ of unitary connections on the hermitian vector bundle E introduced above is naturally an affine space modelled on $\Omega^1(M, \mathfrak{u}(E))$. So picking any unitary connection $\nabla_0 \in \mathcal{A}$, we can write

$$\mathcal{A} = \nabla_0 + \Omega^1(M, \mathfrak{u}(E)).$$

Therefore, the tangent space of $\mathcal{A}$ at a connection ∇ can be identified with $\Omega^1(M, \mathfrak{u}(E))$. This space is naturally endowed with the L^2 *inner product.* For two tangent vectors $A, B \in \Omega^1(M, \mathfrak{u}(E))$, we put

$$g_{L^2}(A, B) = -\int_M \mathrm{tr}(A \wedge *B).$$

If we just look at smooth connections, we see that $(\mathcal{A}, g_{L^2})$ is not a Hilbert manifold - the tangent spaces are not complete, i.e. not Banach spaces. We will instead work with the Sobolev completion of $\mathcal{A}$ and thus define for a given integer $k > 0$ the *Sobolev space of unitary connections* $\mathcal{A}_k$ to be

$$\mathcal{A}_k = \nabla_0 + \Omega^1_k(M, \mathfrak{u}(E)),$$

where we wrote $\Omega^1_k(M, \mathfrak{u}(E)) = L^2_k(\mathrm{Hom}(TM, \mathfrak{u}(E))$. Since on a compact manifold $\Omega^1_k(M, \mathrm{u}(E))$ is independent of the choice of connection, in the sence that different connections yield equivalent norms, we obtain the following proposition.

Proposition 1.2.5. *If the base manifold M is compact, $\mathcal{A}_k$ does not depend on the choice of background connection ∇_0.*

1.2.4 The Action of the Gauge Group

With the analytical technology introduced in the last section at hand, we shall now consider the group $\mathcal{G}$ and its action on $\mathcal{A}$ more closely. This section is inspired by [49]. Recall

$$\mathcal{G} = \{u \in \Gamma(\mathrm{End}(E)) \mid u^* u = \mathrm{id}_E\} = \Gamma(\mathrm{U}(E)).$$

Our starting point in this section is the following construction.

Theorem 1.2.6. *Let* $\dim M = 4$ *and* $k > 2$*, or* $\dim M = 2$ *and* $k > 1$*. Then the closure of* $\mathcal{G}$ *in* $L^2_k(\mathrm{End}(E))$*, denoted by* $\mathcal{G}_k$*, is an (infinite-dimensional) smooth Hilbert Lie group with Lie algebra* $L^2_k(\mathfrak{u}(E)) = \{\xi \in L^2_k(\mathrm{End}(E)) \mid u^* = -u\}$*. Charts are provided by the pointwise exponential map induced from the exponential map on* $\mathrm{U}(E_p, h_p) \cong U(n)$*, where* $n = \mathrm{rk}(E)$*, i.e.*

$$\exp : L^2_k(\mathfrak{u}(E)) \to \mathcal{G}_k,$$

$$\exp(\xi)(p) = \exp(\xi(p)).$$

Proof. Note first, that since $\dim M = 4$ and $k > 2$ (resp. $\dim M = 2$ and $k > 1$), we are in a situation where $L^2_k(\mathrm{End}(E))$ is actually an algebra consisting of continuous sections and the multiplication operation on $L^2_k(\mathrm{End}(E))$, being a bounded bilinear map, is a smooth map of Hilbert spaces.

$$\exp : \xi \mapsto \exp(\xi) = \sum_{n=0}^{\infty} \frac{\xi^n}{n!}$$

is thus well defined as a map $L^2_k(\mathrm{End}(E)) \to L^2_k(\mathrm{End}(E))$ by the multiplication theorem and the usual estimate $||\exp(\xi)||_{L^2_k} \leq \exp(||\xi||_{L^2_k})$. If we restrict to $L^2_k(\mathfrak{u}(E))$, then the image of exp lies in $\mathcal{G}_k$.

Let ρ be the injectivity radius of $\mathrm{U}(n)$, where n is the rank of E. Then, since we are in a range where L^2_k-sections of $\mathrm{End}(E)$ are in fact continuous by the Sobolev embedding theorem, we get that exp is injective on the ball of radius ρ about 0 (in the sup-norm) in $L^2_k(\mathfrak{u}(E))$. Also the derivative of exp at 0 is given by the identity. To see this, let $\xi \in L^2_k(\mathfrak{u}(E))$ then

$$\frac{1}{t}(\exp(t\xi) - \mathrm{id}_E - t\xi) = \sum_{n=1}^{\infty} \frac{\xi^n}{n!} t^{n-1} - \xi = \sum_{n=2}^{\infty} \frac{\xi^n}{n!} t^{n-1} = t \sum_{n=2}^{\infty} \frac{\xi^n}{n!} t^{n-2} \to 0 \text{ as } t \to 0.$$

Along these lines we prove that exp is smooth. So by the inverse function theorem exp is a diffeomorphism on a small open ball about $0 \in L^2_k(\mathfrak{u}(E))$ of radius $\delta > 0$, say. We consider the ball $B_\delta(0) \subset L^2_k(\mathfrak{u}(E))$. Given any $u \in \mathcal{G}_k$, we define a chart about u in the following way. Let $U_u = u \exp(B_\delta)$. This is an open neighbourhood of u in $\mathcal{G}_k$. Define a chart

$$\phi_u : U_u \to B_\delta$$

by

$$\phi_u(u \exp(\xi)) = \xi.$$

To show that $\mathcal{G}_k$ is a manifold, we have to check that the transition functions are smooth. So let $u, v \in \mathcal{G}_k$ be such that $U_u \cap U_v \neq \emptyset$. Let $\xi \in \phi_u(U_u \cap U_v)$, then

$$\phi_v \circ \phi_u^{-1}(\xi) = \phi_v(u \exp(\xi)) = \phi_v(v \exp(\eta_{uv}) \exp(\xi)),$$

for a unique η_{uv} such that $\exp(\eta_{uv}) = u^{-1}v \in \exp(B_\delta(0))$. By the Baker-Campbell-Hausdorff formula we can write

$$\exp(\eta_{uv}) \exp(\xi) = \exp(f_{uv}(\xi)),$$

where $f_{uv} : L^2_k(\mathrm{End}(E)) \to L^2_k(\mathrm{End}(E))$ is a power series in brackets of ξ and η_{uv} and is thus smooth in ξ. □

Proposition 1.2.7. *The action of $\mathcal{G}_k$ on $\mathcal{A}_{k-1}$ is smooth.*

Proof. Fix a background connection $\nabla_0 \in \mathcal{A}_{k-1}$. Then we get local co-ordinates on $\mathcal{A}_{k-1}$ by writing any connection $\nabla \in \mathcal{A}_{k-1}$ as $\nabla = \nabla_0 + (\nabla - \nabla_0)$ with $\nabla - \nabla_0 = A \in \Omega^1_k(M, \mathfrak{u}(E))$. So explicitly the chart is given by $\nabla \mapsto A$. It is enough to show that the action of $\mathcal{G}_k$ is smooth for $u \in \mathcal{G}_k$ close to the identity. Let $u = \exp(\xi) \in \exp(B_\delta)$. Recall

$$u.\nabla = \nabla + u^{-1}\mathrm{d}^\nabla u = \nabla_0 + u^{-1}Au + u^{-1}\mathrm{d}^{\nabla_0}u,$$

which, using $u = \exp(\xi)$, is equal to

$$\nabla_0 + \exp(-\xi)A\exp(\xi) + \mathrm{d}^{\nabla_0}\xi + \rho(A, \xi),$$

where ρ is a power series in brackets of A,ξ and $\mathrm{d}^\nabla\xi$, similar to the Baker-Campbell-Hausdorff formula, which is of order $||\xi||^2_k$.

So in the co-ordinates on $\mathcal{A}_{k-1}$ just defined and the co-ordinates on $\mathcal{G}_k$ given in the proof of the above theorem, the action reads

$$\mu(\xi, A) = \xi.A = \exp(-\xi)A\exp(\xi) + \mathrm{d}^{\nabla_0}\xi + \rho(A, \xi).$$

We see that, since d^{∇_0} maps $L^2_k(\mathfrak{u}(E))$ to $\Omega^1_{k-1}(M, \mathfrak{u}(E))$, we lose one derivative, and so this formula explains why $\mathcal{G}_k$ acts on $\mathcal{A}_{k-1}$ and not on $\mathcal{A}_k$. We check that μ is smooth in the $L^2_k(\mathfrak{u}(E))$ component at zero, with first derivative given by $\mathrm{d}^\nabla : L^2_k(\mathfrak{u}(E)) \to \Omega^1_{k-1}(M, \mathfrak{u}(E))$. In fact, up to terms of order one in t we have

$$\frac{1}{t}(\mu(t\eta, A) - \mu(0, A) - t(\mathrm{d}^{\nabla_0}\eta + [A, \eta])) = \frac{1}{t}(\exp(-t\eta)A\exp(t\eta) - A) + \mathrm{d}^{\nabla_0}\eta - \mathrm{d}^{\nabla_0}\eta - [A, \eta].$$

This is equal to

$$\frac{1}{t}((\exp(-t\eta)A\exp(t\eta) - A)) - [A, \eta],$$

which converges to zero in the L^p_k-norm by the same reasoning as in the case of the adjoint action of a finite dimensional Lie group on its Lie algebra.

Now we prove smoothness in the second factor at ∇_0. Up to terms of order one in t we obtain

$$\frac{1}{t}(\mu(\xi, tA) - \mu(\xi, 0)) = \exp(-\xi)A\exp(\xi) + \mathrm{d}^{\nabla_0}\xi - \mathrm{d}^{\nabla_0}\xi,$$

which is equal to

$$\exp(-\xi)A\exp(\xi).$$

The smoothness at arbitrary points in $\mathcal{G}_k$ and $\mathcal{A}_{k-1}$ can be deduced from this using again the Baker-Campbell-Hausdorff formula as in the proof of the theorem above. □

From our proof, we deduce the following important corollary, giving an explicit formula for the infinitesimal action of $\mathcal{G}_k$ on $\mathcal{A}_{k-1}$.

Corollary 1.2.8. *Let $\xi \in L^2_k(\mathfrak{u}(E)) = \mathrm{Lie}(\mathcal{G}_k)$. Then the fundamental vector field of the action of $\mathcal{G}_k$ on $\mathcal{A}_{k-1}$ associated to ξ is given by*

$$X^\xi(\nabla) = \mathrm{d}^\nabla \xi.$$

Proposition 1.2.9. *Let $k > 3$, then the* curvature map *given by*

$$\begin{aligned} R : \mathcal{A}_{k-1} &\to \Omega^2_{k-2}(M, \mathfrak{u}(E)), \\ \nabla &\mapsto R^\nabla \end{aligned}$$

is a smooth map of Hilbert manifolds with first derivative

$$\mathrm{d}R_\nabla(A) = \mathrm{d}^\nabla A.$$

Moreover, R is equivariant with respect to the action of $\mathcal{G}_k$ on $\mathcal{A}_{k-1}$ and the adjoint action of $\mathcal{G}_k$ on $\Omega^2_{k-2}(M, \mathfrak{u}(E))$. That is, for each $u \in \mathcal{G}_k$

$$R^{u.\nabla} = u^{-1}R^\nabla u.$$

Proof. We fix a background connection $\nabla_0 \in \mathcal{A}_k$. Then the curvature map is given by

$$R^\nabla = R^{\nabla_0 + A} = R^{\nabla_0} + \mathrm{d}^{\nabla_0}A + \frac{1}{2}[A \wedge A] = R(A).$$

Now

$$\frac{1}{t}(R(tA) - R(0) - t\mathrm{d}^{\nabla_0}A) = \frac{1}{2}t[A \wedge A].$$

But by the Sobolev multiplication theorem together with the embedding theorem we get that the L^2_{k-2}-norm of the right hand side is bounded by

$$tC||A||^2_{L^2_{k-1}},$$

which tends to zero as $t \to 0$. We have shown that R is smooth at the point ∇_0, which was arbitrary. Therefore, R is smooth everywhere. The equivariance property is easily checked by writing the action of $\mathcal{G}_k$ on ∇ as $u^{-1} \circ \nabla \circ u$. Then we get for the curvature

$$R^{u.\nabla} = u^{-1}\mathrm{d}^\nabla u u^{-1}\mathrm{d}^\nabla u = u^{-1}R^\nabla u.$$

□

Later on, when we construct the ASD moduli space, we want to take the quotient of the set of ASD connections by the action of the gauge group. If we just do this naively, we do not obtain a well-behaved space. Therefore, we will now discuss the *reduced gauge group* and the set of *irreducible connections.*

The centre of $U(n)$ consists of matrices of the form $e^{i\theta}\mathrm{id}$, where $\theta \in \mathbb{R}$. Therefore, the centre of the gauge group $\mathcal{G}_k$ is given by $L^2_k(M, \mathrm{U}(1))$. Now using the fact that the identity section $\mathrm{id}_E \in \mathcal{G}_k$ is parallel with respect to any connection on E we immediately arrive at the following lemma.

Lemma 1.2.10. *Let $\nabla \in \mathcal{A}_k$. The stabiliser subgroup $\mathcal{G}_{k,\nabla} = \{u \in \mathcal{G}_k \mid u.\nabla = \nabla\}$ always contains $\{e^{i\theta}\mathrm{id}_E \mid \theta \in \mathbb{R}\} \cong Z(\mathrm{U}(n))$.*

Proof. We consider the formula

$$u.\nabla = \nabla + u^{-1}\mathrm{d}^\nabla u,$$

from which we read off

$$\mathcal{G}_{k,\nabla} = \{u \in \mathcal{G}_k \mid \mathrm{d}^\nabla u = 0\}.$$

□

Thus, a priori, the action of $\mathcal{G}_k$ is nowhere free. In order to obtain a smooth moduli space, we restrict our attention to connections with minimal isotropy group and we divide $\mathcal{G}_k$ by the centre of $\mathrm{U}(n)$.

Definition 1.2.1. The *reduced gauge group* $\mathcal{G}^*_k$ is defined to be

$$\mathcal{G}^*_k = \mathcal{G}_k/Z(\mathrm{U}(n)).$$

We call a connection $\nabla \in \mathcal{A}_{k-1}$ *irreducible* if

$$\mathcal{G}_{k,\nabla} = Z(\mathrm{U}(n)).$$

The set of irreducible connections will be denoted by $\mathcal{A}^*_{k-1}$.

Note that the Lie algebra of $\mathcal{G}^*_k$ is given by $L^2_k(\mathfrak{u}(E))/i\mathbb{R}\mathrm{id}_E$.

Corollary 1.2.11. *The reduced gauge group $\mathcal{G}^*_k$ acts freely on the set of irreducible connections $\mathcal{A}^*_{k-1}$.*

We end this section by proving a *slice theorem* for the action of the gauge group on the space of connections. This will allow us to identify the tangent space of a connection ∇ in the moduli space $\mathcal{A}^*_{k-1}/\mathcal{G}^*_k$ of all irreducible connections with the orthogonal complement of the tangent space to the orbit through ∇. Moreover, these slices provide local models for the moduli space $\mathcal{A}^*_{k-1}/\mathcal{G}^*_k$. This is analogous to the situation at the very beginning of this thesis when we discussed finite dimensional quotient manifolds. In the proof that the quotient space of a manifold M by a smooth, free and proper Lie group action is again smooth, we construct so called *slice charts* which locally give a set of representatives for the G-orbits. This does not work immediately in the infinite-dimensional case, and we have to make sure that such a slice exists at every point.

Theorem 1.2.12 (Slice Theorem)**.** *Let $k > 2$ and let $\nabla \in \mathcal{A}^*_{k-1}$ be an irreducible connection, then there exists a constant $\epsilon = \epsilon(\nabla) > 0$, such that for any connection $\tilde{\nabla} = \nabla + A \in \mathcal{A}^*_{k-1}$ with $||A||_{L^2_{k-1}} < \epsilon$ there exists a unique gauge transformation $u \in \mathcal{G}^*_k$ satisfying*

$$(\mathrm{d}^\nabla)^*(\nabla - u.\tilde{\nabla}) = 0.$$

We say $\tilde{\nabla}$ is in Coulomb gauge with respect to ∇.

Proof. This is an application of the implicit function theorem in Banach spaces. We identify $\mathcal{A}_{k-1} = \nabla + \Omega^1_{k-1}(M, \mathfrak{u}(E)) \cong \Omega^1_{k-1}(M, \mathfrak{u}(E))$ and consider the map

$$F : \mathcal{G}^*_k \times \mathcal{A}_{k-1} \to \mathrm{im}((\mathrm{d}^\nabla)^*) \subset L^2_{k-2}(M, \mathfrak{u}(E)),$$

given by

$$F(u, \tilde{\nabla}) = (\mathrm{d}^\nabla)^*(\nabla - u.\tilde{\nabla}).$$

Writing $\tilde{\nabla} = \nabla + A$, we have $u.\tilde{\nabla} = \nabla + u^{-1}Au + u^{-1}\mathrm{d}^\nabla u$ and so

$$F(u, \tilde{\nabla}) = F(u, A) = (\mathrm{d}^\nabla)^*(u^{-1}Au + u^{-1}\mathrm{d}^\nabla u).$$

We want to show that F is smooth and that the derivative of F at $(\mathrm{id}_E, \nabla) = (\mathrm{id}_E, 0)$ in the $\mathcal{G}_k$-direction is surjective, so that by the implicit function theorem we can write the equation $F(u, \tilde{\nabla}) = 0$ on a small neighbourhood of (id_E, ∇) in the form

$$F(u(\tilde{\nabla}), \tilde{\nabla}) = 0$$

for some smooth function $u : \mathcal{A}_{k-1} \to \mathcal{G}_k^*$.

Firstly, we observe that F is in fact differentiable since $\mathcal{G}_k$ acts smoothly and $(\mathrm{d}^\nabla)^* : \Omega^1_{k-1}(M, \mathfrak{u}(E)) \to L^2_{k-2}(\mathfrak{u}(E))$ is a bounded linear map. To compute the partial derivative $D_1 F$ at $(\mathrm{id}_E, 0)$, we calculate with $u_t = \exp(t\xi)$:

$$\frac{1}{t}(F(\exp(t\xi), 0) - F(\mathrm{id}_E, 0)) = \frac{1}{t}((\mathrm{d}^\nabla)^*(\exp(-t\xi)\mathrm{d}^\nabla \exp(t\xi)) \to (\mathrm{d}^\nabla)^*\mathrm{d}^\nabla \xi,$$

as $t \to 0$. Thus, we get for the partial derivative in the $\mathcal{G}_k^*$-direction of F at $(\mathrm{id}_E, 0)$

$$D_1 F|_{(\mathrm{id}_E, 0)} = (\mathrm{d}^\nabla)^*\mathrm{d}^\nabla : L^2_k(\mathfrak{u}(E))/i\mathbb{R}\mathrm{id}_E \to \mathrm{im}((\mathrm{d}^\nabla)^*) \subset L^2_{k-2}(\mathfrak{u}(E)),$$

which is an elliptic operator. The fundamental theorem for elliptic operators says that the equation

$$(\mathrm{d}^\nabla)^*\mathrm{d}^\nabla \xi = \eta$$

admits a solution ξ if and only if η is orthogonal to the kernel of $((\mathrm{d}^\nabla)^*\mathrm{d}^\nabla)^* = (\mathrm{d}^\nabla)^*\mathrm{d}^\nabla$. But this is just the kernel of d^∇, which is equal to $i\mathbb{R}\mathrm{id}_E$ by irreducibility. Since we can identify the image of $(\mathrm{d}^\nabla)^*$ with the orthogonal complement to the kernel of d^∇ in $L^2_k(\mathfrak{u}(E))$, the partial derivative $D_1 F$ is surjective and thus an isomorphism at $(\mathrm{id}_E, 0)$ and the implicit function theorem in Banach spaces gives us our solution $u(A)$ for A sufficiently close to zero. Moreover, u depends smoothly on A.

To prove the uniqueness of u, suppose we have a connection ∇ and a gauge transformation u such that both ∇_1 and $\nabla_2 = u.\nabla_1$ are in Coulomb gauge with respect to ∇. We want to show that u has to be a constant multiple of the identity. Write $\nabla_1 = \nabla + A$ and $\nabla_2 = \nabla + B$. Then $u.\nabla_1 = \nabla_2$ means

$$\mathrm{d}^\nabla u = uB - Au.$$

Now apply $(\mathrm{d}^\nabla)^*$ to this equation. This gives

$$\begin{aligned}
(\mathrm{d}^\nabla)^*\mathrm{d}^\nabla u &= - * \mathrm{d}^\nabla * (uB - Au) \\
&= - * (\mathrm{d}^\nabla u \wedge *B + u \wedge \mathrm{d}^\nabla * B - ((\mathrm{d}^\nabla) * A)u - *A \wedge \mathrm{d}^\nabla u) \\
&= - * (\mathrm{d}^\nabla u \wedge *B) + u \wedge (\mathrm{d}^\nabla)^* B - ((\mathrm{d}^\nabla)^* A)u + *(*A \wedge \mathrm{d}^\nabla u) \\
&= - * (\mathrm{d}^\nabla u \wedge *B) + *(*A \wedge \mathrm{d}^\nabla u),
\end{aligned}$$

since by the Coulomb gauge condition $(\mathrm{d}^\nabla)^*A = 0 = (\mathrm{d}^\nabla)^*B$. Since $(\mathrm{d}^\nabla)^*\mathrm{d}^\nabla$ is elliptic, self-adjoint and positive, we have an eigenspace decomposition

$$L^2(M, \mathrm{End}(E)) \cong \ker((\mathrm{d}^\nabla)^*\mathrm{d}^\nabla) \oplus \bigoplus_{\lambda>0} \mathrm{Eig}((\mathrm{d}^\nabla)^*\mathrm{d}^\nabla, \lambda),$$

where the eigenvalues λ are real and strictly positive. We may therefore decompose $u = u_0 + u_1$ according to this decomposition, i.e. $u_0 \in \ker((\mathrm{d}^\nabla)^*\mathrm{d}^\nabla)$ and $u_1 \in \bigoplus_{\lambda>0} \mathrm{Eig}((\mathrm{d}^\nabla)^*\mathrm{d}^\nabla, \lambda)$. In particular, we get an estimate

$$||u_1||^2_{L^2} \leq \lambda_{min}^{-1}||\mathrm{d}^\nabla u_1||^2_{L^2} = \lambda_{min}^{-1}||\mathrm{d}^\nabla u||^2_{L^2},$$

where λ_{min} is the smallest non-zero eigenvalue. Now we compute, using the expression for $(\mathrm{d}^\nabla)^*\mathrm{d}^\nabla u$ found above,

$$\begin{aligned}
||\mathrm{d}^\nabla u_1||^2_{L^2} &= g_{L^2}((\mathrm{d}^\nabla)^*\mathrm{d}^\nabla u_1, u_1)\\
&= g_{L^2}(- * (\mathrm{d}^\nabla u_1 \wedge *B) + *(*A \wedge \mathrm{d}^\nabla u_1), u_1)\\
&= -\int_M \mathrm{tr}(-(\mathrm{d}^\nabla u_1 \wedge *B)u_1 + (*A \wedge \mathrm{d}^\nabla u_1)u_1)\\
&= -\int_M \mathrm{tr}(-(\mathrm{d}^\nabla u_1) \wedge (*Bu_1) - (\mathrm{d}^\nabla u_1) \wedge *u_1 A)\\
&= -g_{L^2}(\mathrm{d}^\nabla u_1, Bu_1 + u_1 A)\\
&\leq ||\mathrm{d}^\nabla u_1||_{L^2}(||Bu_1||_{L^2} + ||u_1 A||_{L^2}) \text{ (Cauchy-Schwartz, triangle inequality)}\\
&\leq ||\mathrm{d}^\nabla u_1||_{L^2}||u_1||_{L^4}(||B||_{L^4} + ||A||_{L^4}) \text{ (Hölder inequality)}\\
&\leq C||u_1||^2_{L^2_1}(||B||_{L^4} + ||A||_{L^4}) \text{ (Sobolev-embedding } L^2_1 \subset L^4)\\
&= C(||u_1||^2_{L^2} + ||\mathrm{d}^\nabla u_1||^2_{L^2})((||B||_{L^4} + ||A||_{L^4})\\
&\leq C(1 + \lambda_{min}^{-1})||\mathrm{d}^\nabla u_1||^2_{L^2}(||B||_{L^4} + ||A||_{L^4}) \text{ (by the above estimate).}
\end{aligned}$$

So if we choose $\epsilon < (2C(1 + \lambda_{min}^{-1}))^{-1}$, then $\mathrm{d}^\nabla u_1 = 0$, which means that $u_1 = 0$ and so u is parallel, i.e. $u \in Z(\mathrm{U}(n))$, since ∇ is irreducible. □

This theorem shows that we can view $\mathcal{A}^*_{k-1}$ as a smooth principal $\mathcal{G}^*$-bundle over the moduli space $\mathcal{B}^*_{k-1} = \mathcal{A}^*_{k-1}/\mathcal{G}^*_k$ of all irreducible unitary connections. The tangent space of $\mathcal{B}^*_{k-1}$ at a point ∇ can be identified with the orthogonal complement of the tangent space to a $\mathcal{G}^*_k$-orbit $\mathcal{O}_\nabla$. Since the fundamental vector fields are of the form $\mathrm{d}^\nabla\xi$ for $\xi \in L^2_k(\mathfrak{u}(E))$, we obtain

$$\begin{aligned}
T_\nabla\mathcal{B}^*_{k-1} &= (T_\nabla\mathcal{O}_\nabla)^\perp\\
&= \{A \in \Omega^1_{k-1}(M, \mathfrak{u}(E)) \mid g_{L^2}(\mathrm{d}^\nabla\xi, A) = 0 \ \forall\xi \in L^2_k(\mathfrak{u}(E))\}\\
&= \{A \in \Omega^1_{k-1}(M, \mathfrak{u}(E)) \mid g_{L^2}(\xi, (\mathrm{d}^\nabla)^*A) = 0 \ \forall\xi \in L^2_k(\mathfrak{u}(E))\}\\
&= \ker(\mathrm{d}^\nabla)^* \subset \Omega^1_{k-1}(M, \mathfrak{u}(E)).
\end{aligned}$$

Moreover, the theorem implies that the canonical projection $\pi : \mathcal{A}^*_{k-1} \to \mathcal{B}^*_{k-1}$ gives a diffeomorphism between a neighbourhood of $\nabla \in \mathcal{B}^*_{k-1}$ and $B_{\epsilon(\nabla)}(0) \cap (T_\nabla \mathcal{O}_\nabla)^\perp \subset \Omega^1_{k-1}(M, \mathfrak{u}(E)) = T_\nabla \mathcal{A}^*_{k-1}$ and so provides charts for $\mathcal{B}$. This will be useful in the following section.

1.2.5 The ASD Moduli Space on a Hyperkähler Manifold is Hyperkähler

From now on, we assume that our base four-manifold M is hyperkähler, i.e. either a four-torus or a $K3$-surface, and we look at the action of the gauge group in this case. Our aim in this section is to show that the space $\mathcal{A}_k$ inherits a hyperkähler structure from M, which is preserved by the action of $\mathcal{G}_k$. The main result is then that the moment map of this action vanishes exactly on the set of ASD connections. Therefore, the finite dimensional ASD moduli space may be interpreted as a hyperkähler quotient.

Proposition 1.2.13. *Suppose the base manifold is a compact hyperkähler four-manifold (M, g, I, J, K). Then $\mathcal{A}_k$ is a flat hyperkähler Hilbert manifold with complex structures given by the action of $-I, J, K$ on one-forms. The associated Kähler forms $\tilde{\omega}_i$ are given by*

$$\tilde{\omega}_I(A, B) = \int_M \operatorname{tr}(A \wedge B) \wedge \omega_I, \qquad \tilde{\omega}_i(A, B) = -\int_M \operatorname{tr}(A \wedge B) \wedge \omega_i, \qquad i = J, K,$$

where $A, B \in \Omega^1_k(M, \mathfrak{u}(E))$.

Proof. We only give the proof for the complex structure I. Pick normal co-ordinates about an arbitrary point $p \in M$ such that ω_I is then given at p by

$$\omega_I(p) = \mathrm{d}x_1(p) \wedge \mathrm{d}x_2(p) + \mathrm{d}x_3(p) \wedge \mathrm{d}x_4(p).$$

Now let $A, B \in \Omega^1(M, \mathfrak{u}(E))$. At p we can write them as

$$A(p) = \sum_i A_i \otimes \mathrm{d}x_i(p) \qquad B(p) = \sum_i B_i \otimes \mathrm{d}x_i(p).$$

Then, at p, we get

$$\begin{aligned}(\operatorname{tr}(A \wedge B) \wedge \omega_I)(p) &= \sum_{i,j} \operatorname{tr}(A_i B_j) \mathrm{d}x_i \wedge \mathrm{d}x_j \wedge (\mathrm{d}x_1 \wedge \mathrm{d}x_2 + \mathrm{d}x_3 \wedge \mathrm{d}x_4) \\ &= \operatorname{tr}(A_3 B_4 - A_4 B_3 + A_1 B_2 - A_2 B_1) \mathrm{d}x_1 \wedge \mathrm{d}x_2 \wedge \mathrm{d}x_3 \wedge \mathrm{d}x_4 \\ &= \operatorname{tr}(A_3 B_4 - A_4 B_3 + A_1 B_2 - A_2 B_1) \mathrm{vol}_g(p).\end{aligned}$$

On the other hand, we have at p

$$
\begin{aligned}
\operatorname{tr}(-IA \wedge *B)(p) &= \sum_{i,j} \operatorname{tr}(A_i B_j) g_p((-I)^*_p \mathrm{d}x_i(p), \mathrm{d}x_j(p)) \mathrm{vol}_g(p) \\
&= \operatorname{tr}(A_1 B_2 - A_2 B_1 + A_3 B_4 - A_4 B_3) \mathrm{vol}_g(p) \\
&= (\operatorname{tr}(A \wedge B) \wedge \omega_I)(p).
\end{aligned}
$$

Thus, the integrands coincide at every point, so the integrals also have to be equal.

To show that $\tilde{\omega}_I$ is non-degenerate, suppose $\tilde{\omega}_I(A, -)$ is the zero map. However, if we take $B = - * (A \wedge \omega_I)$, then

$$
\tilde{\omega}_I(A, - * (A \wedge \omega_I)) = - \int_M \operatorname{tr}(A \wedge *(A \wedge \omega_I)) \wedge \omega_I = ||A \wedge \omega_I||^2_{L^2},
$$

and so $A = 0$ since $L_I = \omega_I \wedge - : \Omega^1(M) \to \Omega^3(M)$ is injective.

The metric g_{L^2}, the complex structures $-I, J, K$ and the Kähler forms do not depend on $\nabla \in \mathcal{A}$. Thus, g and the $\tilde{\omega}_i$'s are bounded bilinear maps on $\Omega^1_k(M, \mathfrak{u}(E))$, which are constant. Thus, g is a smooth Riemannian metric and the $\tilde{\omega}_i$'s are smooth 2-forms, which are constant in local co-ordinates. Hence, the metric is flat and the Kähler forms are all closed. Thus, we have given $\mathcal{A}_k$ the structure of an infinite dimensional hyperkähler manifold. For details on the exterior derivative on Banach manifolds, see [40]. □

Proposition 1.2.14. *The action of $\mathcal{G}_k$ on $\mathcal{A}_{k-1}$ preserves the hyperkähler structure.*

Proof. We have seen in the last section that the induced action of $\mathcal{G}_k$ on the tangent space of $\mathcal{A}_{k-1}$ is given by the adjoint action. Since the trace is invariant under conjugation, it follows that $\mathcal{G}_k$ leaves invariant the L^2 metric as well as the three Kähler forms. □

Lemma 1.2.15. *There is a non-degenerate bilinear pairing*

$$
\langle -, - \rangle : L^2(\mathfrak{u}(E)) \times \Omega^4_{L^2}(\mathfrak{u}(E)) \to \mathbb{R}
$$

given by

$$
\langle \xi, \alpha \rangle = \int_M \operatorname{tr}(\xi \alpha).
$$

Proof. Given $\xi \in L^2(\mathfrak{u}(E))$, simply take $\alpha = - * \xi$. Then

$$
\langle \xi, - * \xi \rangle = ||\xi||^2_{L^2},
$$

which is non-zero unless ξ vanishes identically. □

Recall that there is a Sobolev embedding $L^2_k(\mathfrak{u}(E)) \subset L^2(\mathfrak{u}(E))$. Thus, the lemma allows us to embed the dual of $\mathrm{Lie}(\mathcal{G}^*_k)$ into $\Omega^4(\mathfrak{u}(E))$. In fact, if we choose an irreducible connection ∇, we may identify $\mathrm{Lie}(\mathcal{G}^*_k)^*$ as the annihilator of $\ker \mathrm{d}^\nabla \subset L^2_k(\mathfrak{u}(E))$, which is isomorphic to $\mathrm{im}\,\mathrm{d}^\nabla : \Omega^3_k(M, \mathfrak{u}(E)) \to \Omega^4_{L^2}(M, \mathfrak{u}(E))$.

Proposition 1.2.16. *Under the identification above, the hyperkähler moment map*

$$\mu = (\mu_I, \mu_J, \mu_K) : \mathcal{A}_{k-1} \to \Omega^4_{L^2}(M, \mathfrak{u}(E)) \otimes \mathbb{R}^3$$

for the action of $\mathcal{G}_k$ on $\mathcal{A}_{k-1}$ is (up to sign) given by

$$\mu_i(\nabla) = R^\nabla \wedge \omega_i.$$

Proof. The proof is analogous to the original construction of a symplectic structure on the moduli space of flat connections on a bundle over a Riemann surface by Atiyah and Bott, see [4]. We only do the proof for the complex structure I. By definition, we have at any point $\nabla \in \mathcal{A}_{k-1}$ for a tangent vector $A \in T_\nabla \mathcal{A}_{k-1} = \Omega^1_{k-1}(M, \mathfrak{u}(E))$

$$(\mathrm{d}\mu_I)_\nabla(A)(\xi) = \tilde{\omega}_I(X^\xi, A) = \tilde{\omega}_I(\mathrm{d}^\nabla \xi, A).$$

Now we integrate by parts, using Stokes' theorem, the closedness of ω_I and the fact that the trace is parallel and obtain

$$\begin{aligned} \tilde{\omega}_i(\mathrm{d}^\nabla \xi, A) &= \int_M \mathrm{tr}(\mathrm{d}^\nabla \xi \wedge A) \wedge \omega_I \\ &= \int_M \mathrm{d}(\mathrm{tr}(\xi A) \wedge \omega_I) - \mathrm{tr}(\xi \mathrm{d}^\nabla A) \wedge \omega_I \\ &= \langle \xi, -\mathrm{d}^\nabla A \wedge \omega_I \rangle. \end{aligned}$$

This should be equal to $\langle \xi, (\mathrm{d}\mu_I)_\nabla(A) \rangle$. So using 1.2.9, we deduce

$$\mu_I(\nabla) = -R^\nabla \wedge \omega_I.$$

The equivariance has already been shown in 1.2.9. □

Using the fact that locally Λ^+ is spanned by the ω_i's and that the decomposition $\Lambda^2 = \Lambda^+ \oplus \Lambda^-$ is orthogonal with respect to the wedge product, we see that the vanishing of the hyperkähler moment map is equivalent to $R^\nabla \in \Omega^-(M, \mathfrak{u}(E))$, i.e. the curvature has to be anti-self-dual. Thus, formally, the moduli space of ASD-connections is given as a hyperkähler quotient. However, we now work in the context of infinite-dimensional Hilbert manifolds. In order to make sure that the hyperkähler quotient construction applies, we have to check that 0 is a regular value of the hyperkähler moment map.

Lemma 1.2.17. *If $\nabla \in \mathcal{A}^*_{k-1}$ is an ASD connection, then*

$$0 \to L^2_k(\mathfrak{u}(E)) \to \Omega^1_{k-1}(M, \mathfrak{u}(E)) \to \Omega^+_{k-2}(M, \mathfrak{u}(E)) \to 0$$

is an elliptic complex.

Proof. For $p \in M$ and $0 \neq \alpha \in T_pM^*$ we have to check that the symbol sequence

$$0 \to \mathfrak{u}(E_p) \to T^*M \otimes \mathfrak{u}(E_p) \to \Lambda^+ \otimes \mathfrak{u}(E_p) \to 0$$

is exact. The first map is given by $\sigma(\mathrm{d}^\nabla)(p, \alpha) = \alpha \otimes -$ and the second is given by $\sigma((\mathrm{d}^\nabla)^+)(p, \alpha) = \mathrm{pr}_+ \circ (\alpha \wedge -)$. The proof of exactness is easy linear algebra and we do not present it here in detail. □

Thus, this complex is elliptic and therefore its cohomology spaces are finite dimensional. In particular, we observe that H^1 of this complex, which is given by

$$H^1 = \ker((\mathrm{d}^\nabla)^+)/\mathrm{im}(\mathrm{d}^\nabla),$$

is isomorphic to the orthogonal complement of $\mathrm{im}(\mathrm{d}^\nabla)$ in $\ker((\mathrm{d}^\nabla)^+)$. Recalling that $\mathrm{im}(\mathrm{d}^\nabla)$ is the tangent space to the $\mathcal{G}_k$-orbit through ∇ and $\ker((\mathrm{d}^\nabla)^+)$ is exactly the kernel of the derivative of the moment map, we conclude that the tangent space to the ASD moduli space is given by H^1.

As explained in [23], chapter 4, the Atiyah-Singer Index theorem computes the index of $\mathcal{D} = \mathrm{d}^\nabla \oplus (\mathrm{d}^{\nabla +})^*$, which equals minus the Euler characteristic of this complex, in terms of topological data. The following theorem shows that, if we restrict attention to irreducible ASD connections, in fact it computes the dimension of the ASD moduli space.

Theorem 1.2.18. *If the base manifold M is compact hyperkähler and ∇ is an irreducible ASD connection, then the cohomology groups H^0 and H^2 of the elliptic complex*

$$0 \to L^2_k(\mathfrak{u}(E)) \to \Omega^1_{k-1}(M, \mathfrak{u}(E)) \to \Omega^+_{k-2}(M, \mathfrak{u}(E)) \to 0$$

satisfy $h^0 = \dim H^0 = 1$ and $h^2 = \dim H^2 = 3$. In particular

$$\dim H^1 = \mathrm{ind}(\mathcal{D}) + \dim H^0 + \dim H^2 = \mathrm{ind}(\mathcal{D}) + 4.$$

The proof is built on the following remark.

Remark. An ASD connection ∇ on a hermitian vector bundle $E \to M$ has curvature of type $(1,1)$, i.e. the associated operators $\bar{\partial}_I^\nabla$, $\bar{\partial}_J^\nabla$ and $\bar{\partial}_K^\nabla$ all square to zero. Thus, ∇ endows E with the structure of a holomorphic vector bundle with respect to each of the three complex structures on the base manifold M.

Proof. First note that $h^0 = 1$ is just an equivalent way of saying that ∇ is irreducible. By definition, we have $H^2 = \ker((\mathrm{d}^\nabla)^*|_{\Omega^+(M,\mathfrak{u}(E))})$. Now $\Omega^+(M, \mathfrak{u}(E))$ as a $L^2_{k-2}(M, \mathfrak{u}(E))$-module is spanned by the three Kähler forms ω_I, ω_J and ω_K. So let $\Phi = \sum_i \phi_i \omega_i$ be a self-dual two-form. Using the Lefschetz operators $L_i = \omega_i \wedge -$, we write this as

$$\Phi = \sum_i L_i \phi_i.$$

Now

$$(\mathrm{d}^\nabla)^* \Phi = \sum_i (\mathrm{d}^\nabla)^* L_i \phi_i = \sum_i [(\mathrm{d}^\nabla)^*, L_i] \phi_i.$$

But M is compact Kähler and (E, ∇) is a holomorphic vector bundle with respect to each complex structure I, J, K. So for $j \in \{I, J, K\}$ we can use the Kähler identity

$$[(\mathrm{d}^\nabla)^*, L_j] = -i(\partial_j^\nabla - \bar{\partial}_j^\nabla).$$

Recalling that on sections of a vector bundle with connection over a complex manifold (M, I) we have by definition

$$\bar{\partial}^\nabla = \frac{1}{2}(\mathrm{d}^\nabla + iI\mathrm{d}^\nabla) \qquad \partial^\nabla = \frac{1}{2}(\mathrm{d}^\nabla - iI\mathrm{d}^\nabla),$$

we get that

$$-i(\partial_j^\nabla - \bar{\partial}_j^\nabla) = -I\mathrm{d}^\nabla.$$

Therefore

$$-(\mathrm{d}^\nabla)^* \Phi = I\mathrm{d}^\nabla \phi_I + J\mathrm{d}^\nabla \phi_J + K\mathrm{d}^\nabla \phi_K = IX^{\phi_I} + JX^{\phi_J} + KX^{\phi_K},$$

by the definition of the fundamental vector fields of the action. The three terms on the right hand side of the equation are mutually orthogonal with respect to the L^2 inner product, as can be seen by the following integration by parts argument (using that ω_K is closed).

$$\begin{aligned} g_{L^2}(I\mathrm{d}^\nabla\phi_I, J\mathrm{d}^\nabla\phi_J) &= g_{L^2}(K\mathrm{d}^\nabla\phi_I, \mathrm{d}^\nabla\phi_J) \\ &= \int_M \mathrm{tr}(\mathrm{d}^\nabla\phi_I \wedge \mathrm{d}^\nabla\phi_J) \wedge \omega_K \\ &= \int_M \mathrm{dtr}(\phi_I \mathrm{d}^\nabla\phi_J) \wedge \omega_K - \mathrm{tr}(\phi_I \mathrm{d}^\nabla\mathrm{d}^\nabla\phi_J) \wedge \omega_K \\ &= \int_M \mathrm{tr}(\phi_I(R^\nabla \wedge \omega_K)\phi_J) \\ &= 0 \end{aligned}$$

by the moment map equations (This computation should be compared to the one presented in the proof of the hyperkähler quotient construction). So if $(\mathrm{d}^\nabla)^*\Phi = I\mathrm{d}^\nabla\phi_I + J\mathrm{d}^\nabla\phi_J + K\mathrm{d}^\nabla\phi_K = 0$ each of the terms on the right hand side has to vanish already. But this implies that the ϕ_i's lie in the kernel of d^∇ which is one-dimensional since ∇ is irreducible. Thus,

$$\dim H^2 = 3\dim\ker\mathrm{d}^\nabla = 3\dim H^0 = 3,$$

as desired. □

We remark, that the hyperkähler property, i.e. the fact that the ω_i's are parallel, really lies at the heart of the above proof. This theorem implies that 0 is a regular value of the moment map μ restricted to $\mathcal{A}^*$. To see this, we use the Lefschetz operators $L_i = \omega_i \wedge -$ and identify as in the proof of the theorem

$$\Omega^+_{k-2}(M, \mathfrak{u}(E)) = \bigoplus_{i\in\{I,J,K\}} L_i(L^2_k(\mathfrak{u}(E))).$$

Now recall from section 3.1.1 that the Kähler forms $\omega_I, \omega_J, \omega_K$ are orthogonal and since they are self-dual, they satisfy $\omega_i \wedge \omega_J = 0$ if $i \neq j$. Therefore, since the ω_i's are non-degenerate, the map

$$(L_I, L_J, L_K) : \Omega^+_{k-2} \to \Omega^4_{k-2}(M, \mathfrak{u}(E)) \otimes \mathbb{R}^3$$

is an isomorphism. Under this isomorphism $\mu : \mathcal{A}_{k-1} \to \mathrm{Lie}(\mathcal{G}^*_k) \subset \Omega^4_{L^2}(M, \mathfrak{u}(E)) \otimes \mathbb{R}^3$ corresponds to $\mathrm{pr}^+ \circ R$, where R is the curvature map, and so $\mathrm{d}\mu : \Omega^1_{k-1}(M, \mathfrak{u}(E)) \to \Omega^4_{k-2}(M, \mathfrak{u}(E)) \otimes \mathbb{R}^3$ corresponds to the operator $(\mathrm{d}^\nabla)^+$ appearing in the elliptic complex, which as a map to $\mathrm{Lie}(\mathcal{G}^*_k)^* \otimes \mathbb{R}^3 = \mathrm{im}(\mathrm{d}^\nabla) \subset \Omega^4_{L^2}(\mathfrak{u}(E)) \otimes \mathbb{R}^3$ is surjective.

Corollary 1.2.19. *Under the above identification, 0 is a regular value of the hyperkähler moment map μ. Hence, $\mu^{-1}(0)$ is a smooth submanifold of $\mathcal{A}_{k-1}$.*

Theorem 1.2.20. *The moduli space of unitary irreducible ASD connections over a hyperkähler manifold is hyperkähler.*

Proof. By corollary 1.2.19 we see that the space $\mu^{-1}(0)$ of all ASD connections is a smooth submanifold of the space $\mathcal{A}_{k-1}$ of all connections of Sobolev class $k-1$. It is proved in [36], chapter VII, proposition 1.14, that the action of $\mathcal{G}_k$ on $\mathcal{A}_{k-1}$ is proper. In paragraph VII, 5 one can also find a proof of the symplectic quotient in the context of Banach manifolds. In the same manner, the hyperkähler quotient carries over to Banach manifolds. Our work in this section shows that all the necessary assumptions are satisfied in our case. Thus, the ASD moduli space $\mu^{-1}(0)/\mathcal{G}_k$ is hyperkähler.

□

The fact, that the moduli space of self-dual connections on some vector bundle with compact structure group is a smooth manifold, was first proven in [6] in the case when the base manifold has positive scalar curvature. The authors use this assumption for proving the vanishing theorem 1.2.18 by a Weitzenböck argument. In our situation, where the base manifold is hyperkähler, the scalar curvature is zero by Ricci-flatness. We presented a slightly different proof in this situation which relies on the fact that a hermitian vector bundle with an ASD connection is holomorphic. Note however, that since the Kähler forms ω_i are parallel on M, the original arguments still work.

Chapter 2

Hypersymplectic Geometry

In this chapter we introduce hypersymplectic manifolds and study their geometry. In particular, we prove the hypersymplectic quotient construction. Then we investigate the relation between hypersymplectic, complex and paracomplex quotients. We introduce the ASD equations in split signature and, after discussing twistorial aspects of hypersymplectic manifolds, we rewrite the equations in Lax form.

2.1 Hypersymplectic Manifolds

Hypersymplectic and hyperkähler manifolds are closely related and formally similar at first sight. However, there are also some crucial differences, which we will discuss later on. We start with the definition.

Definition 2.1.1. A *hypersymplectic manifold* is a manifold M^{4k} together with a pseudo-Riemannian metric g of signature $(2k, 2k)$ and three skew adjoint endomorphisms (with respect to g) of the tangent bundle $I, S, T \in \Gamma(\mathrm{End}(TM))$ such that

- $IS = T = -SI$ and $I^2 = -\mathrm{id}_{TM}$, $S^2 = T^2 = \mathrm{id}_{TM}$, and
- $\nabla^g I = \nabla^g S = \nabla^g T = 0$, where ∇^g is the Levi-Civita connection associated to g.

In analogy to the hyperkähler situation this means that the metric is pseudokähler with respect to the complex structure I and also that the two-forms $\omega_S = g(S-,-)$ and $\omega_T = g(T-,-)$ are in fact symplectic, i.e. g is *parakähler* with respect to S, T. Note that since S and T are parallel, their ±1-eigensubbundles, denoted by $TM^{\pm}$, are integrable in the sense of Frobenius. This implies that locally a hypersymplectic manifold can be written as a product of two $2k$-dimensional submanifolds given by the corresponding integral submanifolds. Therefore, such an endomorphism S is also sometimes called a product structure. We note that these integral submanifolds are actually symplectic with symplectic form given by the restriction of ω_I. Indeed, vectors in TM^+ are of the form $X + SX$, for $X \in TM$. Then one can check that

$$\omega_I(X + SX, Y + SY) = 2\omega_I(X + SX, Y).$$

From this it follows that the restriction of ω_I to TM^+ is non-degenerate. For if $\omega_I(X + SX, Y) = 0$ for all Y, then $X + SX = 0$, i.e. $X \in TM^-$.

An important observation is that, like hyperkähler manifolds, hypersymplectic manifolds are complex symplectic: The complex two-form

$$\omega_I^{\mathbb{C}} = \omega_S + i\omega_T$$

is a holomorphic $(2, 0)$-form with respect to I. In the language of holonomy groups, this can be phrased as follows. The holonomy of a hypersymplectic metric lies in $\mathrm{Sp}(2n, \mathbb{R})$, which is the split-real form of $\mathrm{Sp}(2n, \mathbb{C})$ and so in particular non-compact. Hyperkähler manifolds have holonomy contained in the compact real form $\mathrm{Sp}(n)$. So in that sense hypersymplectic manifolds are pseudo-Riemannian cousins of hyperkähler manifolds, and we may view both as special kinds of holomorphic symplectic manifolds.

We have remarked above that we can think of hyperkähler manifolds as manifolds modelled on a quaternionic inner product space in the same way as we think of Kähler manifolds as modelled on a complex inner product space. A similar statement holds for hypersymplectic manifolds. They are modelled on vector spaces which are modules over the algebra of *split quaternions* with their natural inner product of neutral signature.

The algebra $\mathfrak{gl}(2, \mathbb{C})$ of complex 2×2 matrices of trace zero has two real forms, that is, two real Lie subalgebras whose complexification it is. One is given by the quaternions $\mathbb{H}$ under the map

$$\mathrm{Im}(\mathbb{H}) \to \mathfrak{su}_2,$$

$$i \mapsto \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \qquad j \mapsto \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} \qquad k \mapsto \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix},$$

sending 1 just to the identity matrix.

There is another real form, the *split* real form, given by $\mathfrak{gl}(2, \mathbb{R})$, i.e. just real 2×2-matrices. We can view this as a unital algebra generated by

$$i = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \qquad s = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad t = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}.$$

So formally, we define the algebra $\mathbb{B}$ of *split quaternions* to be the four dimensional algebra over $\mathbb{R}$ generated by the elements $1, i, s, t$ subject to the relations

$$i^2 = -1 = -s^2 = -t^2 \qquad is = t = -si.$$

For an element $q = q_1 + q_2 i + q_3 s + q_4 t \in \mathbb{B}$ we define its *conjugate* by

$$\bar{q} = q_1 - q_2 i - q_3 s - q_4 t.$$

This gives $\mathbb{B}$ a natural metric g of signature $(2,2)$ given by

$$g(q,q) = (q\bar{q}) = q_1^2 + q_2^2 - q_3^2 - q_4^2.$$

Note that under the above matrix representation $\mathrm{Im}(\mathbb{B}) \cong \mathfrak{sl}(2,\mathbb{R})$, which is the split real form of $\mathfrak{sl}(2,\mathbb{C})$, whereas $\mathrm{Im}(\mathbb{H}) = \mathfrak{su}(2)$ is the compact real form.

We define three $\mathbb{B}$-linear endomorphisms I, S, T of $\mathbb{B}$ by multiplication on the right by $-i, s, t$ respectively (again scalars act on the left). With respect to the basis $\{e_1, e_2, e_3, e_4\} = \{1, i, s, t\}$, these are given by the following matrices:

$$I = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \qquad S = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \qquad T = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

A direct computation shows that they again obey the same relations as i, s, t. Next, we define the three symplectic forms $\omega_I, \omega_S, \omega_T$ by

$$\begin{aligned} \omega_I &= g(I-,-) = \mathrm{d}x_2 \wedge \mathrm{d}x_1 + \mathrm{d}x_4 \wedge \mathrm{d}x_3, \\ \omega_S &= g(S-,-) = \mathrm{d}x_3 \wedge \mathrm{d}x_1 + \mathrm{d}x_4 \wedge \mathrm{d}x_2, \\ \omega_T &= g(T-,-) = \mathrm{d}x_4 \wedge \mathrm{d}x_1 + \mathrm{d}x_2 \wedge \mathrm{d}x_3. \end{aligned}$$

This makes $\mathbb{B} \cong \mathbb{R}^{2,2}$ into a hypersymplectic manifold. Note that again all three symplectic forms are self-dual with respect to the Hodge-star operator defined by the split signature metric g and the orientation given by the basis $\{1, i, s, t\}$. In this way, we may think of hypersymplectic manifolds as being modelled on a vector space over $\mathbb{B}$.

But as noted earlier, there are also distinct differences between hyperkähler and hypersymplectic manifolds, one of which we will discuss in more detail in the next section. On a hypersymplectic manifold the algebraic setup is less symmetric than in the hyperkähler situation. While we deal with a collection of compatible complex structures on a hyperkähler manifold, one can either consider the complex structure I

or the product structures S and T. In fact, we have a two-sheeted hyperboloid of compatible complex structures on a hypersymplectic manifold given by endomorphisms of the form

$$I_q = q_1 I + q_2 S + q_3 T,$$

where $q_1^2 - q_2^2 - q_3^2 = 1$. On the other hand, we have a one-sheeted hyperboloid of product structures given by endomorphisms of the form

$$S_q = q_1 I + q_2 S + q_3 T,$$

where now $q_1^2 - q_2^2 - q_3^2 = -1$. So one can study hypersymplectic manifolds from the point of view of complex, or in fact pseudokähler, geometry, but also from the point of view of paracomplex, or parakähler geometry.

2.2 The Hypersymplectic Quotient

A third possibility is to adopt a fairly symplectic point of view, as we did when we discussed hyperkähler quotients. There is a similar but less well-behaved quotient construction for hypersymplectic manifolds which will be the subject of this section. In contrast to the hyperkähler situation, this construction is more problematic: It is not automatic that the two-forms on the quotient coming from the symplectic forms of the original manifold will be symplectic, i.e. non-degenerate again.

The reason for this pathology is that our arguments in the Kähler and hyperkähler constructions crucially used the fact that the metric there is Riemannian, i.e. positive definite. In particular, we used the definiteness to conclude that if two vector subspaces are orthogonal then automatically their sum has to be direct.

In the hypersymplectic case, however, our metric has neutral signature. Therefore, this conclusion is no longer immediate. However, under certain additional assumptions, we can still construct a hypersymplectic structure on the quotient. More precisely, one has the following theorem.

Theorem 2.2.1 (Hypersymplectic Quotient, [30]). *Let G be a Lie group acting on a hypersymplectic manifold (M, g, I, S, T) preserving the hypersymplectic structure such that there are moment maps μ_I, μ_S, μ_T associated to each of the three symplectic forms $\omega_I, \omega_S, \omega_T$, respectively. Define the hypersymplectic moment map:*

$$\mu = (\mu_I, \mu_S, \mu_T) : M \to \mathfrak{g}^* \otimes \mathbb{R}^3.$$

If

- *$c = (c_1, c_2, c_3) \in Z(\mathfrak{g}^*) \otimes \mathbb{R}^3$ is a regular value of μ,*
- *G acts freely and properly on $\mu^{-1}(c)$ and*
- *$T_p\mathcal{O}_p \cap T_p\mathcal{O}_p^\perp = \{0\}$ for all $p \in \mu^{-1}(c)$,*

then the quotient

$$\mu^{-1}(c)/G$$

is in a natural way a hypersymplectic manifold of dimension $\dim M - 4 \dim G$.

Proof. Firstly, we observe that in contrast to the hyperkähler case it is not clear that c is a regular value of μ despite the fact that G acts freely on the level set. The image of $\mathrm{d}\mu$ inside $\mathfrak{g}^* \otimes \mathbb{R}^3$ is given by

$$\{\alpha \in \mathfrak{g}^* \mid \exists X \in T_pM : \alpha(\xi) = (\omega_I(X^\xi, X), \omega_S(X^\xi, X), \omega_T(X^\xi, X)) \ \forall \xi \in \mathfrak{g}^*\}.$$

So if $\mathfrak{g}_p$ is the stabiliser of $p \in M$, we see immediately that

$$\mathrm{im}(\mathrm{d}\mu_p) \subset \mathfrak{g}_p^0.$$

On the other hand, the same reasoning as in the hyperkähler situation shows that the kernel of $\mathrm{d}\mu$ is given by

$$\ker(\mathrm{d}\mu_p) \;\; = \;\; [I(T_p\mathcal{O}_p) + S(T_p\mathcal{O}_p) + T(T_p\mathcal{O}_p)]^\perp.$$

These spaces are mutually orthogonal. So if we were dealing with a Riemannian, i.e. positive definite, metric, we could now argue in just the same way as in the hyperkähler quotient construction to see that c is a regular value of the moment map. However, we now work with an indefinite metric, and there is no reason why the orthogonality of the vector spaces should ensure that their sum is direct. So we cannot automatically carry out our “dimension count” argument from the proof of the hyperkähler quotient construction to show that the freeness of the action on $\mu^{-1}(c)$ guarantees that c is a regular value of μ. Moreover, for the same reason we will in general not be able to identify the tangent space at p of the quotient with the metric orthogonal complement of $T_p\mathcal{O}_p$, the tangent space of the orbit through p. So in general, we cannot prove the smoothness of the quotient in this way, but we have to take it as an assumption.

We also have to worry about the non-degeneracy of the induced symplectic forms on the quotient. In the construction of the symplectic quotient and the Kähler quotient, we noted that $\ker i^*\omega \subset T_p\mathcal{O}_p$, which ensures that the symplectic structure on the quotient will be non-degenerate in these situations. Of course, we get the same statement, when we restrict ω_i to $\mu_i^{-1}(c_i)$. But if we restrict further to $\mu^{-1}(c) = \bigcap_i \mu_i^{-1}(0)$, this degeneracy subspace might become larger. In the case of a hypersymplectic manifold, we have the following explicit description of the degeneracy spaces of the restricted symplectic forms on $\mu^{-1}(c)$, obtained by Hitchin in [30].

Lemma 2.2.2 ([30])**.** *The degeneracy subspaces of the restrictions of the three symplectic forms to $\mu^{-1}(c)$ at $p \in \mu^{-1}(c)$ are given by*

$$\begin{aligned}
\ker(\omega_I)_p &= T_p\mathcal{O}_p + S(T_p\mathcal{O}_p \cap T_p\mathcal{O}_p^{\perp}) + T(T_p\mathcal{O}_p \cap T_p\mathcal{O}_p^{\perp}),\\
\ker(\omega_S)_p &= T_p\mathcal{O}_p + I(T_p\mathcal{O}_p \cap T_p\mathcal{O}_p^{\perp}) + T(T_p\mathcal{O}_p \cap T_p\mathcal{O}_p^{\perp}),\\
\ker(\omega_T)_p &= T_p\mathcal{O}_p + S(T_p\mathcal{O}_p \cap T_p\mathcal{O}_p^{\perp}) + I(T_p\mathcal{O}_p \cap T_p\mathcal{O}_p^{\perp}).
\end{aligned}$$

Proof. We only prove the statement for ω_I. The other two assertions are proved analogously. Let $p \in \mu^{-1}(c)$ and let $X \in \ker(\omega_I)_p \subset T_p\mu^{-1}(c)$, i.e

$$(\omega_I)_p(X, -) : T_p\mu^{-1}(c) \to \mathbb{R}$$

is the zero map. Let $\{\xi_i \mid 1 \le i \le \dim \mathfrak{g}\}$ be a basis for $\mathfrak{g}$. Then the fundamental vector fields $X_i = X_p^{\xi_i}$ are a basis for $T_p\mathcal{O}_p$. Reinterpreting $\mathrm{d}\mu$ as a map

$$\mathrm{d}\mu : \mathfrak{g} \to T_pM^* \otimes \mathbb{R}^3 \cong \mathrm{Hom}(T_pM, \mathbb{R}^3),$$

we can write

$$T_p\mu^{-1}(c) = \bigcap_{\xi \in \mathfrak{g}} \ker \mathrm{d}\mu(\xi) = \bigcap_{\xi_j, 1 \le j \le \dim \mathfrak{g}; i \in \{I,S,T\}} \ker \mathrm{d}\mu_i(\xi_j).$$

Now, since all the ω_i's are non-degenerate two-forms on T_pM and since the action of G is free on $\mu^{-1}(c)$, we see that if $\xi \neq 0 \in \mathfrak{g}$ then

$$\mathrm{d}\mu(\xi) = (\omega_I(X^\xi, -), \omega_S(X^\xi, -), \omega_T(X^\xi, -)) \neq 0 \in T_pM^* \otimes \mathbb{R}^3.$$

So $\mathrm{d}\mu$ is injective. Now complete the system $\{\mathrm{d}\mu_i(\xi_j)\}$ to a generating system $\{\mathrm{d}\mu_i(\xi_j), \alpha_k\}$ of $T_pM^* \otimes \mathbb{R}^3$ and let $V_k \in T_pM \otimes \mathbb{R}^3$ be the elements dual to the α_k's. Then we see from the definition of the dual basis that

$$T_p\mu^{-1}(c) = \mathrm{span}\{V_k\}.$$

Conversely, since $\omega_I(X, -)$ vanishes on $T_p\mu^{-1}(c)$, this means that as a one-form on T_pM, we have

$$(\omega_I)_p(X, -) \in \mathrm{span}\{\mathrm{d}\mu_i(\xi_j)\},$$

which means

$$(\omega_I)_p(X,-) = \sum_{j=1}^{\dim \mathfrak{g}} a_j \mathrm{d}\mu_I^j + b_j \mathrm{d}\mu_S^j + c_j \mathrm{d}\mu_T^j,$$

where we used the short-hand notation $\mathrm{d}\mu_i^j = (\mathrm{d}\mu_i)_p(\xi_j)$. Now by definition of the ω_i's and μ, this is the same as

$$g(IX,-) = \sum_{j=1}^{\dim \mathfrak{g}} a_j g(IX_j,-) + b_j g(SX_j,-) + c_j g(TX_j,-),$$

where we wrote $X_j = X^{\xi_j}$. Thus, since g is non-degenerate

$$IX = \sum_{j=1}^{\dim \mathfrak{g}} a_j IX_j + b_j SX_j + c_j TX_j.$$

This is equivalent to

$$X = \sum_{j=1}^{\dim \mathfrak{g}} a_j X_j - b_j TX_j + c_j SX_j.$$

Since X is tangent to $\mu^{-1}(c)$, we must have $\mathrm{d}\mu_i^j(X) = 0$ for all i,j. Hence, noting that X_j is also tangent to $\mu^{-1}(c)$ as G acts on $\mu^{-1}(c)$, we get

$$0 = \mathrm{d}\mu_T^k(X) = g(TX_k, X) = \sum_{j=1}^{\dim \mathfrak{g}} a_j g(TX_k, X_j) - b_j g(TX_k, TX_j) + c_j g(TX_k, SX_j).$$

The first and the third terms equal $\sum_j a_j \mathrm{d}\mu_T^k(X_j)$ and $\sum_j c_j \mathrm{d}\mu_I^k(X_j)$ respectively and hence both vanish. Thus, the expression simplifies to

$$0 = \mathrm{d}\mu_T^k(X) = g(\sum_j b_j X_j, X_k),$$

which means

$$\sum_j b_j X_j \in T_p\mathcal{O}_p \cap T_p\mathcal{O}_p^{\perp}.$$

Taking $\mathrm{d}\mu_S^k$ instead of $\mathrm{d}\mu_T^k$ we get

$$0 = \mathrm{d}\mu_S^k(X) = -g(\sum_j c_j X_j, X_k),$$

and hence

$$\sum_j c_j X_j \in T_p\mathcal{O}_p \cap T_p\mathcal{O}_p^{\perp}.$$

So we see

$$X \in T_p\mathcal{O}_p + S(T_p\mathcal{O}_p \cap T_p\mathcal{O}_p^{\perp}) + T(T_p\mathcal{O}_p \cap T_p\mathcal{O}_p^{\perp}).$$

That is,

$$\ker(\omega_I)_p \subset T_p\mathcal{O}_p + S(T_p\mathcal{O}_p \cap T_p\mathcal{O}_p^\perp) + T(T_p\mathcal{O}_p \cap T_p\mathcal{O}_p^\perp).$$

The converse inclusion is obvious, as $X \in T_p\mathcal{O}_p + S(T_p\mathcal{O}_p \cap T_p\mathcal{O}_p^\perp) + T(T_p\mathcal{O}_p \cap T_p\mathcal{O}_p^\perp)$ just means that $(\omega_I)_p(X, -) \in \operatorname{span}\{\mathrm{d}\mu_i^j\}$ and hence vanishes on $T_p\mu^{-1}(0)$. □

This lemma therefore finishes the proof of the theorem. The assumption $T_p\mathcal{O}_p \cap T_p\mathcal{O}_p^\perp = \{0\}$ guarantees that the induced forms on the quotient are non-degenerate and hence symplectic. □

We remark that along these lines one can prove analogous reduction theorems for pseudokähler and parakähler manifolds. In both cases the additional assumptions are the same.

2.3 Kirwan-Type Theorems

Recall from the discussion of the hyperkähler quotient, that this may be interpreted as a symplectic quotient in the holomorphic category. The key ingredient was the link between Kähler quotients and GIT quotients discovered by Kirwan in [35]. A natural question to ask is whether hypersymplectic quotients admit a similar interpretation.

In this section we therefore investigate to what extent Kirwan-type theorems remain true, if we do not assume that we are dealing with a positive definite Kähler metric. We also try to answer this question for parakähler manifolds and then apply both results to hypersymplectic quotients.

2.3.1 Complex Quotients

Let (M, g, I, ω) be a pseudokähler manifold, i.e. g is a pseudo-Riemannian metric of signature (p, q) and the form $\omega = g(I-, -)$ is symplectic, i.e. closed and non-degenerate.

Suppose a compact Lie group G acts on M preserving the pseudokähler structure with moment map $\mu : M \to \mathfrak{g}^*$ and fundamental vector fields X^ξ for each $\xi \in \mathfrak{g}$. We write as usual $\mathcal{O}_p$ for the G-orbit through a point $p \in M$.

Assume further that the action can be extended to a global holomorphic action of the complexification of G, which we denote by $G^\mathbb{C}$. For this to be possible, we have to assume in contrast to the Riemannian case, that the sum $T\mathcal{O}_p + IT\mathcal{O}_p$ is direct. Secondly, we need that the vector fields IX^ξ are complete, that is, the associated flows exist for all times. We write $\mathcal{O}_p^\mathbb{C}$ for the $G^\mathbb{C}$-orbit through p.

We now define the following subset of M (In GIT this would be called the set of *stable points.*):

$$\tilde{M} = \{p \in M \mid \mathcal{O}_p^{\mathbb{C}} \cap \mu^{-1}(0) \neq \emptyset\} = G^{\mathbb{C}}.\mu^{-1}(0).$$

That is, $\tilde{M}$ is the subset of points in M whose $G^{\mathbb{C}}$-orbit contains a zero of the moment map associated to the G action. Kirwan's theorem then asserts that if we are dealing with a Riemannian Kähler metric, then there is a homeomorphism

$$\mu^{-1}(0)/G \cong \tilde{M}/G^{\mathbb{C}}.$$

The question we consider in this section is: Under what conditions does this carry over to the case of a pseudokähler metric of signature (p, q) with both p and q non-zero?

In this subsection we answer this question by proving a general theorem. Recall that in order to make pseudokähler reduction work, one has to make a non-degeneracy assumption which in the Kähler case is automatically satisfied as soon as G acts freely on the zero-set of the moment map. So one would hope to be able to prove a Kirwan-type result under this non-degeneracy condition. However, under this assumption we are able to find a proof only for the special case of circle actions. For groups of higher rank, it seems necessary to impose a stronger condition, namely *definiteness*, which for circle actions is equivalent to non-degeneracy, see the corollary at the end of the subsection.

Theorem 2.3.1. *Let (M, g, ω) be a connected pseudokähler manifold with a Hamiltonian action by a compact Lie group G and associated moment map μ for which $0 \in \mathfrak{g}^*$ is a regular value such that $\mu^{-1}(0)$ is compact. Assume that the G-orbits are definite submanifolds of M, i.e. g restricts to a positive (or negative) definite metric on each orbit. Suppose further that G acts freely on $\mu^{-1}(0)$ and that the action may be extended to a holomorphic action of $G^{\mathbb{C}}$, the complexification of G. Then there is a homeomorphism*

$$\mu^{-1}(0)/G \cong \tilde{M}/G^{\mathbb{C}}.$$

Proof. Switching to $-g$ if necessary, we may assume that the restriction of g to any G-orbit is positive definite. We follow closely the original proof from chapter 7 in [35] and divide the proof into two lemmas.

Lemma 2.3.2 (Uniqueness)**.** *Let $p \in \mu^{-1}(0)$. Then under the hypotheses of the theorem*

$$\mathcal{O}_p^{\mathbb{C}} \cap \mu^{-1}(0) = \mathcal{O}_p.$$

Moreover, if G acts freely on $\mu^{-1}(0)$, then $G^{\mathbb{C}}$ acts freely on $\tilde{M}$.

Proof. We have from polar decomposition $G^{\mathbb{C}} = \exp(i\mathfrak{g})G$. Thus, we can write any $\gamma \in G^{\mathbb{C}}$ in the form $\gamma = \exp(i\xi)g_0$ for some $g_0 \in G$. Since G preserves the zero set of the moment map by construction, we may assume that $g_0 = \mathrm{id}$, i.e. that γ lies in $\exp(i\mathfrak{g})$.

Thus, assume $\mu(p) = \mu(\exp(i\xi).p) = 0$ and we want to show that this cannot occur for non-zero ξ. Consider the smooth function $f : \mathbb{R} \to \mathbb{R}$ given by

$$f(t) = \mu(\exp(it\xi).p)(\xi).$$

Then f vanishes at zero and at 1. Hence, it must have a critical point at some $t \in (0,1)$, i.e. $f'(t) = 0$. Now we compute, writing $q = \exp(it\xi).p$,

$$0 = f'(t) = \mathrm{d}\mu_q(IX_q^\xi)(\xi) = \omega_q(X_q^\xi, IX_q^\xi) = g_q(X_q^\xi, X_q^\xi) \geq 0.$$

We conclude $X_q^\xi = 0$ and so the point q is a fixed point of the action of the 1-parameter family $\exp(i\mathbb{R}\xi)$. So in particular $p = q = \exp(i\xi).p$ and we get our uniqueness result and also the freeness property. □

The above lemma establishes a set-theoretic bijection between $\tilde{M}/G^{\mathbb{C}}$ and $\mu^{-1}(0)/G$. In the next lemma we take into account the topology of $\tilde{M}/G^{\mathbb{C}}$.

Lemma 2.3.3 (Hausdorffness)**.** *Two distinct G-orbits in $\mu^{-1}(0)$ can be separated by $G^{\mathbb{C}}$-invariant neighbourhoods in $\tilde{M}$.*

Proof. Let $\mathcal{O}_p$ and $\mathcal{O}_x$ be two distinct G-orbits in $\mu^{-1}(0)$. Now by basic properties of the manifold topology, there is a compact G-invariant neighbourhood of p in $\mu^{-1}(0)$, V say, not containing x. So it is enough to show that $G^{\mathbb{C}}.V$ is a neighbourhood of p in M not containing x in its closure.

There are a couple of things to check. Firstly, we have to check that V actually is a neighbourhood in M of $\mathcal{O}_p$. Thus, we have to show that it contains an open subset of M. This is proven most easily by considering the map

$$\sigma : i\mathfrak{g} \times \mu^{-1}(0) \to \tilde{M} \qquad \sigma(i\xi, p) = \exp(i\xi).p.$$

This a smooth map and the dimensions of the source and target manifold are equal. Thus, all we have to show is, that the differential of σ is an isomorphism at any point in $\{0\} \times \mu^{-1}(0)$. Now observe that the derivative of σ at $(0, p)$ is given by

$$\mathrm{d}\sigma_{(0,p)}(i\xi, X) = \partial_1\sigma_{(0,p)}i\xi + X.$$

Note that by definition we have

$$\partial_1 \sigma_{(0,p)} i\xi = \frac{\mathrm{d}}{\mathrm{d}t}|_{t=0} \sigma(ti\xi, p) = \frac{\mathrm{d}}{\mathrm{d}t}|_{t=0} \exp(ti\xi).p = IX^\xi_p.$$

Furthermore, IX^ξ cannot be contained in $T_p\mu^{-1}(0)$ as shown by the following computation.

$$\mathrm{d}\mu(\xi, IX^\xi) = \omega(X^\xi, IX^\xi) = ||X^\xi||^2,$$

which is non-zero and positive unless $\xi = 0$. Thus, since the map $\xi \mapsto X^\xi$ is injective, the rank of $\mathrm{d}\sigma$ is maximal. As a consequence,

$$W = \{\exp i\xi \mid \xi \in \mathfrak{g},\ ||\xi|| \leq 1\}.V$$

is a compact neighbourhood of $\mathcal{O}_p$ in $\tilde{M}$. Put

$$\epsilon = \inf\{||X^\xi_q||^2 \mid ||\xi|| = 1, \xi \in \mathfrak{g}, q \in W\},$$

which is *positive*, since we assume the action is free and g is positive definite along the G-orbits. For $\xi \in \mathfrak{g}$ of unit length and $q \in V$, we consider again the function

$$f(t) = \mu(\exp it\xi.q)(\xi),$$

whose derivative at $t \in [0, 1]$ satisfies

$$f'(t) = ||X^\xi_{\exp it\xi.q}||^2 \geq \epsilon > 0.$$

Note that $f(0) = 0$, so in particular, we get that

$$f(t) \geq \epsilon, \qquad \text{if } t \geq 1.$$

Moreover, since ξ is of unit length, the norm of $\mu(\exp it\xi.q)$ thought of as a linear form on $\mathfrak{g}$ satisfies

$$||\mu(\exp it\xi.q)|| \geq \epsilon \qquad \text{if } t \geq 1.$$

Therefore,

$$||\mu|| \geq \epsilon \qquad \text{on } (G^{\mathbb{C}}.V)\backslash W.$$

Now W is compact and therefore an intersection of W with a small neighbourhood of $\mu^{-1}(0)$ will be closed in that neighbourhood. Therefore, together with $x \notin (G^{\mathbb{C}}).V$, we get that $x \notin \overline{(G^{\mathbb{C}}).V}$. $\square$

Putting these two lemmas together, we obtain that the natural continuous map

$$P : \mu^{-1}(0)/G \to \tilde{M}/G^{\mathbb{C}},$$

sending the G-equivalence class of a point to its $G^{\mathbb{C}}$-equivalence class is a homeomorphism: It is surjective by definition of $\tilde{M}$ and injective by the first lemma. By the second lemma it is a continuous bijection from a compact space to a Hausdorff space, hence a homeomorphism. □

As already mentioned at the beginning of this subsection, the theorem has a simpler form if we specialise to circle actions in which case the definiteness condition reduces to non-degeneracy.

Corollary 2.3.4. *Let (M, g, ω) be a connected pseudokähler manifold with a Hamiltonian circle action and associated moment map μ for which $0 \in \mathfrak{u}(1)^*$ is a regular value such that $\mu^{-1}(0)$ is compact. Assume that the associated vector field is nowhere null. Suppose further that G acts freely on $\mu^{-1}(0)$ and that the action may be extended to a global holomorphic action of $\mathbb{C}^*$. Then there is a homeomorphism*

$$\mu^{-1}(0)/\mathrm{U}(1) \cong \tilde{M}/\mathbb{C}^*.$$

There is actually a deeper reason why the Kirwan proof works in our situation. Her construction makes substantial use of the Riemannian gradient flow of the function $f = ||\mu||^2$. This leads to convexity statements which imply that $G^{\mathbb{C}}$ orbits in $\tilde{M}$ are closed and intersection points with the critical set $\mu^{-1}(0)$ are unique up to the action of G. In dynamical systems terms, $f = ||\mu||^2$ is a *Liapunov function* for the gradient system associated to f, i.e. it is strictly increasing (or decreasing, depending on sign conventions) along flow lines. An easy computation using the defining properties of the moment map shows that

$$\mathrm{grad} f_p = I X^{2\mu(p)}.$$

Indeed, identifying $\mathfrak{g} \cong \mathfrak{g}^*$ by means of the bi-invariant metric and choosing an orthonormal basis $\{\xi_k\}$, we can write

$$\mu = \sum_{k=1}^{\dim \mathfrak{g}} \mu(\xi_k)\xi_k.$$

Thus,

$$\mathrm{d}f = \mathrm{d}||\mu||^2 = \sum_{k=1}^{\dim \mathfrak{g}} 2\mu(\xi_k)\mathrm{d}\mu(\xi_k).$$

And so

$$\mathrm{d}f(X) = \sum_{k=1}^{\dim \mathfrak{g}} 2\mu(\xi_k)\mathrm{d}\mu(\xi_k)(X) = \sum_{k=1}^{\dim \mathfrak{g}} 2\mu(\xi_k)\omega(X^{\xi_k}, X) = \sum_{k=1}^{\dim \mathfrak{g}} 2\mu(\xi_k)g(IX^{\xi_k}, X),$$

proving our claim. Thus, the gradient flow of f preserves the $G^{\mathbb{C}}$ orbits. This computation is of course also true if g is indefinite. However, a priori we cannot say anything about any monotonicity of the pseudo-Riemannian gradient flow since f is no longer a Liapunov for the indefinite gradient system.

On the contrary, using the Liapunov property, one can show in the Riemannian case that one has the following variational description of $\tilde{M}$.

$$\tilde{M} = \{p \in M \mid \text{the gradient flow line of } f \text{ through } p \text{ contains a limit point in } \mu^{-1}(0)\}.$$

The Liapunov property of a Riemannian gradient flow thus lies at the heart of Kirwan's proof.

While one expects in the Riemannian case that gradient flow lines converge to critical points of f, one can actually have functions whose gradient flow with respect to an indefinite metric is periodic. Consider for example the function $F(x, y) = xy$ on $\mathbb{R}^{1,1}$, whose gradient vector field is given by $(-y, x)^T$ and so the flow lines are standard circles.

Our definiteness assumption in the theorem then precisely ensures that f is still a Liapunov function for the indefinite gradient system and therefore such periodicity behaviour cannot occur.

The positivity condition we have to impose on the G-orbits seems to be a rather strong one and we expect that in most examples, the theorem will not hold. However, we can still say something about the extent to which the theorem fails, if we drop the definiteness assumption.

Proposition 2.3.5. *Away from the degeneracy locus in $\mu^{-1}(0)$, the fibres of the map $\mu^{-1}(0)/G \to G^{\mathbb{C}}.\mu^{-1}(0)/G^{\mathbb{C}}$ are discrete.*

Proof. Let $p \in \mu^{-1}(0)$ and suppose we have a one-parameter family $\exp(i\xi(t))$, with t in some real interval and $\xi(0) = 0$, such that

$$\mu(\exp(i\xi(t)).p) = 0 \qquad \forall t.$$

Then, writing $\dot{\xi} = \frac{\mathrm{d}}{\mathrm{d}t}|_{t=0}\xi(t)$, we get for arbitrary $\eta \in \mathfrak{g}$,

$$\begin{aligned} 0 &= \frac{\mathrm{d}}{\mathrm{d}t}|_{t=0}\mu(\exp(i\xi(t)).p)(\eta) \\ &= \mathrm{d}\mu_p(IX_p^{\dot{\xi}})(\eta) \\ &= \omega_p(X^\eta, IX^{\dot{\xi}}) \\ &= g(X^\eta, X^{\dot{\xi}}). \end{aligned}$$

Thus, $X^{\dot{\xi}}$ lies in the degeneracy space. □

2.3.2 Paracomplex Quotients

We may of course ask the same questions about paracomplex actions. In [48], Matsoukas has defined the *paracomplexification* of a compact Lie group G to be the Lie group associated to the paracomplexification of the Lie algebra $\mathfrak{g}^{\mathbb{A}} = \mathfrak{g} \oplus s\mathfrak{g}$. We can decompose $\mathfrak{g}^{\mathbb{A}}$ into the ± 1-eigenspaces of the endomorphism S given by left multiplication by s, so that as a Lie algebra $\mathfrak{g}^{\mathbb{A}} \cong \mathfrak{g}^+ \oplus \mathfrak{g}^- \cong \mathfrak{g} \times \mathfrak{g}$, where the summands are copies of $\mathfrak{g}$ given by $\mathfrak{g}^{\pm} = \{\frac{1}{2}(\xi \pm s\xi) \mid \xi \in \mathfrak{g}\}$. Note that $\mathfrak{g}$ and $s\mathfrak{g}$ sit inside $\mathfrak{g}^{\mathbb{A}}$ as the diagonal and the anti-diagonal, respectively. It follows that the paracomplexification of G, which we denote by $G^{\mathbb{A}}$, is (at least locally) isomorphic to $G \times G$, with G being embedded diagonally and paracomplex structure S given by multiplication by 1 and -1 on the first and second factor, respectively.

In analogy to the complex situation, we consider a parakähler manifold (M, g, S, ω) with an action of a compact Lie group preserving this structure with moment map μ. Assume that the sum $T\mathcal{O} + S(T\mathcal{O})$ is direct and that the vector fields SX^ξ are complete for each $\xi \in \mathfrak{g}$. This then gives an action of the paracomplexification $G^{\mathbb{A}}$ on M and we write again $\tilde{M}$ for the set of points in M whose paracomplex orbit contains a zero of the moment map.

We begin by showing that any element of $G^{\mathbb{A}}$ admits a "polar decomposition" which, however, is not unique.

Lemma 2.3.6. *Let G be a compact Lie group equipped with a bi-invariant Riemannian metric. Then every element $(g, h) \in G^{\mathbb{A}} = G \times G$ can be written in the form $(g, h) = (x, x) \cdot (y, y^{-1})$ for some $x, y \in G$, not necessarily unique. So since the exponential map of G is surjective, we may write any element of $G^{\mathbb{A}}$ in the form $g \exp(s\xi)$, where g is an element of the diagonal in $G \times G$ and $s\xi$ belongs to the anti-diagonal in $\mathfrak{g} \times \mathfrak{g}$.*

Proof. Suppose we have found such x, y, i.e.

$$g = xy \qquad h = xy^{-1}.$$

Then we get

$$h^{-1}g = y^2.$$

To show existence, we recall that the one-parameter subgroups of G are the geodesics and any two points may be linked by a geodesic since G is compact. Taking the midpoint of the geodesic linking a given element with the identity, we see that every element in G has square root, which is not necessarily unique. So if we put

$$y = \sqrt{h^{-1}g} \qquad x = h\sqrt{h^{-1}g},$$

then $xy = g$ and $xy^{-1} = h$ as desired. □

With this lemma at hand, we can now prove a theorem analogous to the complex case discussed above.

Theorem 2.3.7. *Let (M, g, ω) be a connected parakähler manifold with a Hamiltonian action by a compact Lie group G and associated moment map μ for which $0 \in \mathfrak{g}^*$ is a regular value such that $\mu^{-1}(0)$ is compact. Assume that the G-orbits are definite submanifolds of M, i.e. g restricts to a positive (or negative) definite metric on each orbit. Suppose further that G acts freely on $\mu^{-1}(0)$ and that the action may be extended to a paraholomorphic action of $G^{\mathbb{A}}$, the paracomplexification of G. Then there is a homeomorphism*

$$\mu^{-1}(0)/G \cong \tilde{M}/G^{\mathbb{A}}.$$

Proof. Switching to $-g$ if necessary, we may assume that the restriction of g to any G-orbit is positive definite. In contrast to the discussion of complex quotients, we take advantage of the compactness of $G^{\mathbb{A}}$ and shorten the proof substantially.

Lemma 2.3.8 (Uniqueness)**.** *Let $p \in \mu^{-1}(0)$. Then under the hypotheses of the theorem*

$$\mathcal{O}_p^{\mathbb{A}} \cap \mu^{-1}(0) = \mathcal{O}_p.$$

Moreover, if G acts freely on $\mu^{-1}(0)$, then $G^{\mathbb{A}}$ acts freely on $\tilde{M}$.

Proof. Using polar decomposition, we can write any $\gamma \in G^{\mathbb{A}}$ in the form $\gamma = \exp(s\xi)g_0$ for some $g_0 \in G$. Since G preserves the zero set of the moment map by construction, we may assume that $g_0 = \mathrm{id}$, i.e. that γ lies in $\exp(s\mathfrak{g})$. Thus, assume $\mu(p) =$

$\mu(\exp(st\xi).p) = 0$. Without loss of generality, after rescaling ξ if necessary, we may suppose that $t = 1$, so we consider the following situation:

$$\mu(p) = \mu(\exp(s\xi).p) = 0,$$

and we want to show that this cannot occur for non-zero ξ.

Consider the smooth function $f : \mathbb{R} \to \mathbb{R}$ given by

$$f(t) = \mu(\exp(st\xi).p)(\xi).$$

Then f vanishes at zero and at 1. Hence, it must have a critical point at some $t \in (0,1)$, i.e. $f'(t) = 0$. Now we compute, writing $q = \exp(st\xi).p$,

$$0 = f'(t) = \mathrm{d}\mu_q(SX^\xi_q)(\xi) = \omega_q(X^\xi_q, SX^\xi_q) = -g_q(X^\xi_q, X^\xi_q) \geq 0.$$

We conclude $X^\xi_q = 0$ and so the point q is a fixed point of the action of the 1-parameter family $\exp(s\mathbb{R}\xi)$. So in particular $p = q = \exp(s\xi).p$ and we get our uniqueness result and also the freeness property. □

The above lemma establishes a set-theoretic bijection between $\tilde{M}/G^{\mathbb{A}}$ and $\mu^{-1}(0)/G$. Now, in contrast to the situation of a $G^{\mathbb{C}}$-action, we do not have to worry about non-Hausdorff behaviour. This is because, unlike $G^{\mathbb{C}}$, the paracomplexification $G^{\mathbb{A}}$ is compact if (and only if) G is compact. □

Proposition 2.3.9. *Away from the degeneracy locus in $\mu^{-1}(0)$, the fibres of the map $\mu^{-1}(0)/G \to G^{\mathbb{A}}.\mu^{-1}(0)/G^{\mathbb{A}}$ are discrete and therefore finite, since $G \times G$ is compact.*

Proof. Let $p \in \mu^{-1}(0)$ and suppose we have a one-parameter family $\exp(s\xi(t))$, with t in some real interval and $\xi(0) = 0$, such that

$$\mu(\exp(s\xi(t)).p) = 0 \qquad \forall t.$$

Then, writing $\dot{\xi} = \frac{\mathrm{d}}{\mathrm{d}t}|_{t=0}\xi(t)$, we get for arbitrary $\eta \in \mathfrak{g}$,

$$\begin{aligned} 0 &= \frac{\mathrm{d}}{\mathrm{d}t}|_{t=0}\mu(\exp(s\xi(t)).p)(\eta) \\ &= \mathrm{d}\mu_p(SX^{\dot{\xi}}_p)(\eta) \\ &= \omega_p(X^\eta, SX^{\dot{\xi}}) \\ &= g(X^\eta, X^{\dot{\xi}}). \end{aligned}$$

Thus, $X^{\dot{\xi}}$ lies in the degeneracy space. □

2.3.3 Application to Hypersymplectic Quotients

One of the consequences of Kirwan's original theorem is that it gives an interpretation of hyperkähler quotients as *holomorphic symplectic quotients.*

The crucial point in the hyperkähler situation is that $(\mu_I^{\mathbb{C}})^{-1}(0)$ is automatically a *Kähler* manifold since $\mu_I^{\mathbb{C}}$ is I-holomorphic with 0 as a regular value.

We now wish to apply the same reasoning to the hypersymplectic setting, where we have essentially the same setup with holomorphic symplectic form $\omega_I^{\mathbb{C}} = \omega_S + i\omega_T$ and associated holomorphic moment map $\mu_I^{\mathbb{C}} = \mu_S + i\mu_T$. Assuming $0 \in \mathfrak{g}^* \otimes \mathbb{C}$ is a regular value of $\mu_I^{\mathbb{C}}$, we want to know under which circumstances the complex submanifold $M_0 = (\mu_I^{\mathbb{C}})^{-1}(0)$ is pseudokähler and the split signature metric on M restricts to a positive metric on the G-orbits in M_0.

Proposition 2.3.10. *At any $p \in (\mu_I^{\mathbb{C}})^{-1}(0)$, the kernel of the restriction of ω_I to $(\mu_I^{\mathbb{C}})^{-1}(0)$ is given by*

$$(\ker \omega_I|_{(\mu_I^{\mathbb{C}})^{-1}(0)})_p = S((T_p\mathcal{O}_p + IT_p\mathcal{O}_p) \cap (T_p\mathcal{O}_p + IT_p\mathcal{O}_p)^{\perp})$$

Proof. The proof is very similar to the proof of the lemma needed to prove the hypersymplectic quotient construction. Choose an orthonormal basis $\{\xi_i, i = 1, \dots, \dim \mathfrak{g}\}$ of $\mathfrak{g}$. Let $Y \in T_p(\mu_I^{\mathbb{C}})^{-1}(0)$ such that $\omega_I(Y, -) : T_p(\mu_I^{\mathbb{C}})^{-1}(0) \to \mathbb{R}$ is the zero map. That is, $Y \in \ker \omega_I|_{(\mu_I^{\mathbb{C}})^{-1}(0)}$, which is equivalent to

$$\omega_I(Y, -) \in \operatorname{span}\{\mathrm{d}\mu_S(\xi_i), \mathrm{d}\mu_T(\xi_i), i = 1, \dots, \dim \mathfrak{g}\}.$$

Thus, we can find scalars $a_i, b_i \in \mathbb{R}$, $i = 1, \dots, \dim \mathfrak{g}$ such that, using the shorthand notation $\mathrm{d}\mu_j^i = \mathrm{d}\mu_j(\xi), j = S, T$, we can write

$$\begin{aligned} g(IY, -) &= \omega_I(Y, -) \\ &= \sum_{i=1}^{\dim \mathfrak{g}} a_i \mathrm{d}\mu_S^i + b_i \mathrm{d}\mu_T^i \\ &= \sum_{i=1}^{\dim \mathfrak{g}} a_i \omega_S(X^{\xi_i}, -) + b_i \omega_T(TX^{\xi_i}, -) \\ &= \sum_{i=1}^{\dim \mathfrak{g}} a_i g(SX^{\xi_i}, -) + b_i g(TX^{\xi_i}, -), \end{aligned}$$

which gives

$$IY = \sum_{i=1}^{\dim \mathfrak{g}} a_i SX^{\xi_i} + b_i TX^{\xi_i},$$

i.e.

$$Y = \sum_{i=1}^{\dim \mathfrak{g}} -a_i T X^{\xi_i} + b_i S X^{\xi_i} = S(\sum_{i=1}^{\dim \mathfrak{g}} a_i I X^{\xi_i} + b_i X^{\xi_i}).$$

Now, we apply $\mathrm{d}\mu_S^j$ to Y, which vanishes since Y is tangent to $(\mu_I^{\mathbb{C}})^{-1}(0)$, so

$$0 = \mathrm{d}\mu_S^j(Y) = \omega_S(X^{\xi_j}, Y) = -g(X^{\xi_j}, SY).$$

Similarly, applying $\mathrm{d}\mu_T^j$ yields

$$0 = \mathrm{d}\mu_T^j(Y) = \omega_T(X^{\xi_j}, Y) = g(ISX^{\xi_j}, Y) = g(IX^{\xi_j}, SY).$$

Thus,

$$SY \in (T_p\mathcal{O}_p + IT_p\mathcal{O}_p) \cap (T_p\mathcal{O}_p + IT_p\mathcal{O}_p)^\perp.$$

□

Corollary 2.3.11. *Assume $0 \in \mathfrak{g}^* \otimes \mathbb{C}$ is a regular value of $\mu_I^{\mathbb{C}}$. Then the complex submanifold $((\mu_I^{\mathbb{C}})^{-1}(0), \omega_I)$ is pseudokähler if and only if*

$$(T_p\mathcal{O}_p + IT_p\mathcal{O}_p) \cap (T_p\mathcal{O}_p + IT_p\mathcal{O}_p)^\perp = \{0\},$$

that is, g restricts to a non-degenerate metric on the $G^{\mathbb{C}}$-orbits.

The corollary takes again a simpler form for circle actions.

Corollary 2.3.12. *Let (M, g, I, S, T) be a hypersymplectic manifold with a circle action that preserves the hypersymplectic structure. Assume $0 \in \mathbb{C}$ is a regular value of $\mu_I^{\mathbb{C}}$. Then the complex submanifold $((\mu_I^{\mathbb{C}})^{-1}(0), \omega_I)$ is pseudokähler if and only if the vector field associated to circle action is nowhere null on $(\mu_I^{\mathbb{C}})^{-1}(0)$.*

We are now in a position to formulate how our Kirwan-type result applies to hypersymplectic quotients.

Theorem 2.3.13. *Let (M, g, I, S, T) be a hypersymplectic manifold with a compact Lie group action that preserves the hypersymplectic structure and admits a hypersymplectic moment map $\mu : M \to \mathfrak{g} \otimes \mathbb{R}^3$. Assume that the action can be extended to an action of the complexification $G^{\mathbb{C}}$.*

Suppose that g restricts to a non-degenerate inner product on the $G^{\mathbb{C}}$ orbits in $M_0 = (\mu_I^{\mathbb{C}})^{-1}(0)$ and is positive definite on the G-orbits in $\tilde{M}_0 = G^{\mathbb{C}}.(\mu_I^{-1}(0) \cap M_0) = G^{\mathbb{C}}.(\mu_I^{-1}(0) \cap (\mu_I^{\mathbb{C}})^{-1}(0))$.

Then there is a homeomorphism

$$\mu^{-1}(0)/G \cong \tilde{M}_0/G^{\mathbb{C}} = G^{\mathbb{C}}.(\mu_I^{-1}(0) \cap M_0)/G^{\mathbb{C}}.$$

Of course, we may ask the same questions starting from the paracomplex viewpoint, and consider the S-paraholomorphic symplectic form

$$\omega_S^{\mathbb{A}} = \omega_I + s\omega_T.$$

Then the above discussion gives analogous conditions to ensure that the level set of the paracomplex moment map $\mu_S^{\mathbb{C}} = \mu_I + s\mu_T$ is parakähler with respect to S. The proofs will go through, since in the above we do not use the fact that I is a complex structure. We only use the relations satisfied by I, S, T and the fact that they are skew-adjoint with respect to the metric. Thus, after permuting I, S, T appropriately in the above proofs, we arrive at the following statement.

Theorem 2.3.14. *Assume $0 \in \mathfrak{g}^* \otimes \mathbb{A}$ is a regular value of $\mu_S^{\mathbb{A}}$. Then the paracomplex submanifold $((\mu_S^{\mathbb{A}})^{-1}(0), \omega_S)$ is parakähler if and only if*

$$(T_p\mathcal{O}_p + ST_p\mathcal{O}_p) \cap (T_p\mathcal{O}_p + ST_p\mathcal{O}_p)^\perp = \{0\},$$

that is, g restricts to a non-degenerate metric on the $G^{\mathbb{A}}$-orbits.

If $G = \mathrm{U}(1)$, then $((\mu_S^{\mathbb{A}})^{-1}(0), \omega_I)$ is parakähler if and only if the vector field associated to circle action is nowhere null on $(\mu_S^{\mathbb{A}})^{-1}(0)$.

Suppose in addition that g is positive definite on the G-orbits in $\tilde{M}_0 = G^{\mathbb{A}}.(\mu_S^{-1}(0) \cap M_0) = G^{\mathbb{A}}.(\mu_S^{-1}(0) \cap (\mu_S^{\mathbb{A}})^{-1}(0))$.

Then there is a homeomorphism

$$\mu^{-1}(0)/G \cong \tilde{M}_0/G^{\mathbb{A}} = G^{\mathbb{A}}.(\mu_S^{-1}(0) \cap M_0)/G^{\mathbb{A}}.$$

2.3.4 Examples coming from Linear Torus Actions

In this section we study a class of rather explicit examples, namely linear torus actions on a hypersymplectic vector space. We first collect some notation from [19], which is our main background reference for this section.

We consider closed abelian subgroups $N \subset T^n$ acting on $\mathbb{C}^{n,n} = \mathbb{C}^n \oplus \mathbb{C}^n$ with its standard pseudokähler structure. Using the fact that we can write down everything explicitly, we will produce some examples in which our general Kirwan-type theorem fails. However, thanks to the explicitness, we can still obtain some results concerning the extent to which things go wrong in this case.

Of course, $\mathbb{C}^{n,n} \cong \mathbb{B}^n$ and so is actually hypersymplectic. In this subsection, however, we focus on the pseudokähler structure and ignore hypersymplectic aspects. We will come back to the hypersymplectic picture in the next subsection.

Basic Setup

Consider $\mathbb{C}^{n,n} = \mathbb{C}^n \oplus \mathbb{C}^n$ with co-ordinates z_i on the first and w_i on the second factor. The standard split signature metric is then given by

$$g = \mathrm{Re}\left(\sum_{k=1}^{n} \mathrm{d}z_k \mathrm{d}\bar{z}_k - \mathrm{d}w_k \mathrm{d}\bar{w}_k\right).$$

As a block matrix we write the complex structure as

$$I = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix},$$

so that the pseudokähler form is given by

$$\omega = g(I-,-) = \frac{1}{2i}\sum_{k=1}^{n} \mathrm{d}z_k \wedge \mathrm{d}\bar{z}_k + \mathrm{d}w_k \wedge \mathrm{d}\bar{w}_k.$$

On $\mathbb{C}^{n,n}$ the torus T^n acts in the standard way, simply by thinking of T^n as complex diagonal matrices with entries of modulus 1, i.e.

$$T^n = \{\mathrm{diag}(e^{i\theta_1}, \ldots, e^{i\theta_n}) \mid \theta_i \in \mathbb{R}\}.$$

Then T^n acts on $\mathbb{C}^{n,n}$ by the ordinary action of matrices on vectors on each $\mathbb{C}^n$-factor:

$$\mathrm{diag}(e^{i\theta_1}, \ldots, e^{i\theta_n}).(z,w) = ((\mathrm{diag}(e^{i\theta_1}, \ldots, e^{i\theta_n})z, \mathrm{diag}(e^{i\theta_1}, \ldots, e^{i\theta_n})w),$$

that is,

$$(z_k, w_k) \mapsto (e^{i\theta_k} z_k, e^{i\theta_k} w_k).$$

The fundamental vector fields associated to this action are given by

$$X^{\xi}_{(z,w)} = \sum_{k=1}^{n} i\theta_k z_k \frac{\partial}{\partial z_k} - i\theta_k \bar{z}_k \frac{\partial}{\partial \bar{z}_k} + i\theta_k w_k \frac{\partial}{\partial w_k} - i\theta_k \bar{w}_k \frac{\partial}{\partial \bar{w}_k},$$

where we wrote $\xi = (i\theta_1, \ldots, i\theta_n) \in \mathfrak{t} = \mathrm{Lie}(T^n) \cong \mathbb{R}^n$. This action clearly preserves the metric and the Kähler form. The metric on a T^n-orbit is given by

$$||X^{\xi}||^2 = \sum_{k=1}^{n} \theta_k^2(|z_k|^2 - |w_k|^2).$$

We now compute the complexification of the action. For the vector field IX^{ξ} one easily finds

$$IX^{\xi}_{(z,w)} = \sum_{k=1}^{n} -\theta_k z_k \frac{\partial}{\partial z_k} - \theta_k \bar{z}_k \frac{\partial}{\partial \bar{z}_k} + \theta_k w_k \frac{\partial}{\partial w_k} + \theta_k \bar{w}_k \frac{\partial}{\partial \bar{w}_k} = X^{i\xi}_{(z,w)}.$$

Thus, integrating this vector field, we find that $\lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n) \in (\mathbb{C}^*)^n = (T^n)^{\mathbb{C}}$ acts via

$$(z_k, w_k) \mapsto (\lambda_k z_k, \bar{\lambda}_k^{-1} w_k),$$

and clearly the vector fields IX^ξ are complete.

Just as for the circle action on $\mathbb{C}^2$, which gives rise to the construction of $\mathbb{CP}^1$ as a symplectic quotient, it is easy to compute the moment map associated to the T^n action, which is given by

$$\begin{aligned} \mu : \mathbb{C}^{n,n} &\rightarrow \mathfrak{t}^* \cong (\mathbb{R}^n)^*, \\ \mu(z, w) &= \frac{1}{2}\sum_{k=1}^{n}(|z_k|^2 + |w_k|^2 + c_k)\epsilon_k, \end{aligned}$$

where $c = (c_1, \ldots, c_n) \in \mathbb{R}^n$ and we write $\{\epsilon_i\}$ for the dual of the standard basis $\{e_i\}$ of $\mathbb{R}^n$, which we identify with the Lie algebra of T^n.

We now want to consider the restriction of this action to an abelian subgroup of T^n. To do this, we first need a good description of such a subgroup, and it turns out that for our purposes it is most convenient to define N in terms of its Lie algebra, which we realise as the kernel of a surjective linear map

$$\beta : \mathfrak{t} \rightarrow \mathbb{R}^m,$$

which maps the standard basis of $\mathfrak{t} = \mathbb{R}^n$ into the standard lattice $\mathbb{Z}^m \subset \mathbb{R}^m$, i.e.

$$\beta(e_i) = u_i,$$

where the set $\{u_i\}$ generates $\mathbb{R}^m$ as a vector space. We then define the Lie algebra $\mathfrak{n}$ of N by the short exact sequence

$$0 \rightarrow \mathfrak{n} \xrightarrow{\alpha} \mathfrak{t} \cong \mathbb{R}^n \xrightarrow{\beta} \mathbb{R}^m \rightarrow 0,$$

where we write α for the inclusion map. Requiring that β should map the standard basis into $\mathbb{Z}^m$ ensures that the definition

$$N = (\exp \circ \alpha)(\mathfrak{n})$$

makes sense. Now N acts on $\mathbb{C}^{n,n}$ by restricting the standard T^n action to N. To compute the associated moment map, we first write down the exact sequence dual to the one considered above:

$$0 \rightarrow \mathbb{R}^m \xrightarrow{\beta^*} \mathfrak{t}^* \xrightarrow{\alpha^*} \mathfrak{n}^* \rightarrow 0.$$

Then the moment map μ_N associated to this action of N is given by pulling back the moment map μ associated to the T^n action above to $\mathfrak{n}^*$:

$$\begin{aligned}\mu_N : \mathbb{C}^{n,n} &\rightarrow \mathfrak{n}^*,\\ \mu_N(z,w) &= \alpha^*\mu(z,w) = \frac{1}{2}\sum_{k=1}^{n}(|z_k|^2 + |w_k|^2 + c_k)\alpha_k,\end{aligned}$$

where we denote by α_k the pull-back of the ϵ_k's by α. That is, $\alpha_k = \alpha^*\epsilon_k$. So the moment map $\mu_N(z,w)$ is zero if and only if $\mu(z,w)$ lies in the kernel of α^*. Now note that by exactness we have that $\ker\alpha^* = \mathrm{im}\beta^*$. Therefore, a point (z,w) lies in $\mu_N^{-1}(0)$ if and only if there exists $a \in (\mathbb{R}^m)^*$ such that

$$\beta^* a = \mu(z,w).$$

This means

$$a(u_k) = \mu(z,w)(e_k), \qquad \text{for all } k = 1,\dots,n.$$

Examples and Some General Results

With the general setup established, we now try in this subsection to answer the following questions.

- Which $N^{\mathbb{C}}$-orbits intersect the zero set of the moment map?
- If an $N^{\mathbb{C}}$-orbit contains a zero of the moment map, to what extent is it unique, that is, what can we say about the fibers of the map $P : \mu^{-1}(0)/G \rightarrow \tilde{M}/G^{\mathbb{C}}$?

To get an intuition for the situation we study some examples first.

Example (The Standard Circle Action)**.** In the hypersymplectic setup, to which we shall come shortly, this has been discussed in [48] and in [19]. We choose $m = n-1$ and put $u_k = e_k$ for $k = 1,\dots,n-1$ and $u_n = -(e_1 + \cdots + e_n)$. So β is the map

$$\beta : \mathbb{R}^n \rightarrow \mathbb{R}^{n-1}, \qquad \sum_{i=1}^{n}\theta_i e_i \mapsto \sum_{i=1}^{n-1}(\theta_i - \theta_n)e_i,$$

which means

$$\mathfrak{n} = \mathbb{R}(e_1 + \cdots + e_n),$$

and α is just the inclusion. Therefore, $N \subset T^n$ is the diagonal circle. The moment map $\mu : \mathbb{C}^{n,n} \rightarrow \mathbb{R}$ associated to this action is given by

$$\mu_N(z,w) = \alpha^*\mu(z,w) = \frac{1}{2}\sum_{k=1}^{n}(|z_k|^2+|w_k|^2+c_k)\alpha_k = \frac{1}{2}(|z|^2+|w|^2)+c, \quad c = \sum_{k=1}^{n} c_k \in \mathbb{R},$$

where we used $\alpha^*\epsilon_k((\frac{1}{2}(|z_k|^2+|w_k|^2)+c_k)(e_1+\cdots+e_n)) = \frac{1}{2}(|z_k|^2+|w_k|^2)+c_k$ and identified $\mathfrak{n}\cong\mathbb{R}$. The induced $\mathbb{C}^*$-action is given by

$$(z,w)\mapsto(\lambda z,\bar{\lambda}^{-1}w).$$

What is $\tilde{M}$ in this case? We see that $U(1)$ leaves the moment map invariant, so in order to check whether a $\mathbb{C}^*$-orbit meets $\mu^{-1}(0)$, we have to solve the following equation for λ real and positive,

$$\lambda^2|z|^2+\lambda^{-2}|w|^2+2c=0.$$

In other words

$$\lambda^4|z|^2+\lambda^2 2c+|w|^2=0,$$

which, writing $x=\lambda^2$, gives

$$x_{1,2}=\frac{|c|\pm\sqrt{c^2-|z|^2|w|^2}}{|z|^2}.$$

Thus, in this case,

$$\tilde{M}=\{(z,w)\in\mathbb{C}^{n,n}\backslash\{(0,0)\}\mid |z||w|\leq|c|\}.$$

Note that this condition is preserved by the $\mathbb{C}^*$-action. Thus, we see that every $\mathbb{C}^*$-orbit will hit $\mu^{-1}(0)$ in two points unless

- it lies on the boundary of $\tilde{M}$, i.e. consists of points such that $|z||w|=|c|$

  or

- $z=0$ or $w=0$, but not both.

In the first case, we can write down the associated x in a simpler form, namely $x=\frac{|w|}{|z|}$, where we use that $|c|=-c$, i.e. $c\leq 0$, in order to ensure that $\mu^{-1}(0)$ is non-empty. In other words, the map

$$P:\mu^{-1}(0)/N\cong\mathbb{CP}^{2n-1}\to\tilde{M}/N^{\mathbb{C}}$$

is a $2:1$ cover branched over $\{|z||w|=|c|\}\cup\{z=0,w\neq 0\}\cup\{w=0,z\neq 0\}$.

Note that the fundamental vector field associated to $\xi=i\theta\in\mathfrak{u}(1)$ is given by

$$X^{\xi}_{(z,w)}=i\theta\left(\sum_{k=1}^{n}z_k\frac{\partial}{\partial z_k}-\bar{z}_k\frac{\partial}{\partial\bar{z}_k}+w_k\frac{\partial}{\partial w_k}-\bar{w}_k\frac{\partial}{\partial\bar{w}_k}\right),$$

and its length with respect to the hermitian split signature inner product is therefore

$$||X^{\xi}_{(z,w)}||^2 = \theta^2(|z|^2 - |w|^2),$$

which is null at $(z,w) \in \tilde{M}$ if $|z| = |w|$ and so the assumptions of the theorem are not satisfied. Note further that each $\mathbb{C}^*$-orbit intersects the set $\{|z| = |w|\}$ on which the vector field is null. To see this, we have to solve the equation $\lambda|z| = \lambda^{-1}|w|$, i.e. we can simply take

$$\lambda = \sqrt{\frac{|w|}{|z|}},$$

which is unique modulo U(1), in accordance with the above observations. Note that the conditions $|z| = |w|$ and $\mu(z,w) = 0$ imply that $|z|^2|w|^2 = c^2$. In particular, $\mathbb{C}^*$ does not preserve the inequalities $|z| < |w|$ or $|z| > |w|$.

Example ($N = T^{n-1} \subset T^n$)**.** In this example we choose $m = 1$, $\beta = (0, \ldots, 0, 1) : \mathbb{R}^n \to \mathbb{R}$, that is $u_i = 0$ for $i = 1, \ldots, n-1$ and $u_n = 1$, so that $\mathfrak{n} = \mathbb{R}^{n-1} \times \{0\} \subset \mathbb{R}^n$ and

$$N = \{\mathrm{diag}(e^{i\theta_1}, \ldots, e^{i\theta_{n-1}}, 1)\} \cong T^{n-1} \subset T^n.$$

The moment map is therefore given by

$$\mu_N(z,w) = \sum_{k=1}^{n-1} (\frac{1}{2}(|z_k|^2 + |w_k|^2) + c_k)\epsilon_k.$$

So in oder to see if a $N^{\mathbb{C}}$-orbit meets $\mu_N^{-1}(0)$, we have to find positive real numbers $\{\lambda_1, \ldots, \lambda_{n-1}\}$ solving the following system of equations

$$\frac{1}{2}(\lambda_k^2|z_k|^2 + \lambda_k^{-2}|w_k|^2) + c_k = 0 \qquad k = 1, \ldots, n-1.$$

Since the equations are independent, we can just solve each equation individually and find similarly to the previous example, writing again $x_k = \lambda_k^2$,

$$x_{k;1,2} = \frac{|c_k| \pm \sqrt{c_k^2 - |z_k|^2|w_k|^2}}{|z_k|^2}.$$

This gives a real solution if $c_k^2 - |z_k|^2|w_k|^2 \geq 0$. Altogether, combining the solutions of the individual equations in every possible way, we see that we get 2^D-many solutions, where D is the number of cases in which the inequality $c_k^2 \geq |z_k|^2|w_k|^2$ is strict and z_k and w_k are not both zero, i.e. the number of cases in which we get 2 solutions.

In the above notation, we have

$$\tilde{M} = \{(z,w) \in \mathbb{C}^{n,n} \mid c_k^2 - |z_k|^2|w_k|^2 \geq 0 \;\; \forall k = 1, \ldots, n-1\}.$$

Again the map

$$P : \mu_N^{-1}(0)/N \to \tilde{M}/N^{\mathbb{C}}$$

restricts to a $2^D : 1$ cover on the "D-strata" of $\tilde{M}$, i.e. the sets of points in $(z, w) \in \tilde{M}$ for which there are two solutions in precisely D cases, as D ranges over $0, 1, \ldots, n-1$.

The fundamental vector fields are of the form

$$X^{\xi}_{(z,w)} = \sum_{k=1}^{n-1} i\theta_k z_k \frac{\partial}{\partial z_k} - i\theta_k \bar{z}_k \frac{\partial}{\partial \bar{z}_k} + i\theta_k w_k \frac{\partial}{\partial w_k} - i\theta_k \bar{w}_k \frac{\partial}{\partial \bar{w}_k},$$

and the metric on fundamental vector fields is given by

$$||X^{\xi}_{(z,w)}||^2 = \sum_{k=1}^{n-1} \theta_k^2 (|z_k|^2 - |w_k|^2),$$

which is degenerate if $|z_k| = |w_k|$ for some k. As above, it is easy to show that any $N^{\mathbb{C}}$-orbit hits the degeneracy locus, therefore again, the assumptions of the theorem are not satisfied.

Note however, that in this case there are actually higher-dimensional fibres over points where $(z_k, w_k) = (0, 0)$ for at least one k. For such a point to lie in the zero-locus of the moment map, we need to require $c_k = 0$. If this is satisfied, we can then choose λ_k arbitrary. Note also, that in accordance with the general results, such a point will lie in the degeneracy locus. We observe also, that the topology of the fibres changes if we let the c_k vary, i.e. different level sets of the moment map display different behaviour.

The above two examples were fairly well-behaved and therefore relatively special. The next example is more complicated, and we think that it will contain the typical features that occur for general $N \subset T^n$.

Example. Consider the case, $m = 1$, $\beta = (1, 1, \ldots, 1) : \mathbb{R}^n \to \mathbb{R}$. So $u_k = 1$, $k = 1, \ldots n$. Thus,

$$N = \{\text{diag}(e^{i\theta_1}, \ldots, e^{i\theta_n}) \mid \sum_{k=1}^{n} \theta_k = 0\}.$$

So $(z, w) \in \mathbb{C}^{n,n}$ lies in the zero set of the moment map, if we can find a number $a \in \mathbb{R}$ such that

$$a = \mu(z, w)(e_k) \qquad k = 1, \ldots, n.$$

That is, in order to see if the $N^{\mathbb{C}}$-orbit through (z, w) contains a zero of the moment map, and if so, how many, we have to find $\lambda_k \in (\mathbb{C}^*)^n$, which we may assume to be real and positive, and $a \in \mathbb{R}$ such that

$$\begin{aligned} a &= \frac{1}{2}(\lambda_k^2|z_k|^2 + \lambda_k^{-2}|w_k|^2) + c_k, \qquad k = 1, \dots, n, \\ 1 &= \lambda_1\lambda_2 \dots \lambda_n. \end{aligned}$$

The last equation is the condition that $\mathrm{diag}(\lambda_1, \dots, \lambda_n) \in N^{\mathbb{C}}$. Viewing a as a parameter, we can just solve each equation separately, provided $|z_k||w_k| \leq |c_k - a|$, giving $\lambda_{k;1,2}(a)$ or $\lambda_k(a)$ depending on whether the inequality is strict or not, or whether one of $|z_k|$ or $|w_k|$ is zero. If $(z_k, w_k) = (0, 0)$, then we can take any $\lambda_k \in \mathbb{C}^*$ we like. The last equation can then be viewed as a condition on a. If we write it down explicitly, we get, writing $a_k = a - c_k$

$$1 = \prod_{k:z_k\neq 0\neq w_k} \frac{a_k \pm \sqrt{a_k^2 - |z_k|^2|w_k|^2}}{|z_k|^2} \prod_{k:w_k=0\neq z_k} \sqrt{\frac{2a_k}{|z_k|^2}} \prod_{k:z_k=0\neq w_k} \sqrt{\frac{|w_k|^2}{2a_k}} \prod_{k:w_k=0=z_k} \lambda_k \tag{2.1}$$

We see further that a has to satisfy certain additional constraints. Namely,

$$a_k^2 \geq |z_k|^2|w_k|^2 \qquad \text{if } z_k \neq 0 \neq w_k, \tag{2.2}$$

$$a_k \geq 0 \qquad \text{if } z_k = 0 \text{ or } w_k = 0, \text{ but } (z_k, w_k) \neq 0, \tag{2.3}$$

$$a_k = 0 \qquad \text{if } (z_k, w_k) = 0. \tag{2.4}$$

Thus, if we have more than one k with $(z_k, w_k) = (0, 0)$ then there will be no solution if the corresponding c_k's are different. If there exist at least two k's with $(z_k, w_k) = 0$ and $a_k = 0$, then we obtain again an at least one-dimensional fibre, since then the associated λ_k's can again be chosen arbitrarily, subject to condition (2.1).

We see again, that generically we can have at most 2^{n-1} solutions up to the action of N and possibly, depending on the choice of level c, higher-dimensional fibres for some points in the degeneracy locus. This time however, it is quite complicated to describe either the set $\tilde{M}$ in more detail. We can realise it as the image of a subset $\check{M} \subset \mathbb{C}^{n,n} \times \mathbb{R}$ under the projection onto the $\mathbb{C}^{n,n}$-factor. In fact, we put

$$\begin{aligned} \check{M} = \{(z, w, a) \in \mathbb{C}^{n,n} \times \mathbb{R} \mid \ & (2.1)\text{-}(2.4) \text{ hold} \\ & (\text{for some } \lambda_k \in \mathbb{C}^* \text{ in (2.1) in case } (z_k, w_k) = (0, 0))\}. \end{aligned}$$

A Theorem for General N

We now want to use the intuition gained from the above examples to extract some general statements about the extent to which our Kirwan theorem fails in this situation.

Theorem 2.3.15. *Let $N \subset T^n$ act on $\mathbb{C}^{n,n}$ and let $(z,w) \in \tilde{M}$. Then generically for fixed a we have at most $2^{\dim N}$ possible solutions $\lambda \in N^{\mathbb{C}}$ such that $\mu_N(\lambda.(z,w)) = 0$. Given $\lambda \in N^{\mathbb{C}}$ with this property, it uniquely determines an $a \in (\mathbb{R}^m)^*$ by requiring $\beta^* a = \mu(\lambda.(z,w))$.*

Proof. We have pretty much given the proof in the third example from above. Now we just have to do the same computation for general N.

Let $\mathfrak{n}$ be given as $\ker(\beta : \mathfrak{t} \to \mathbb{R}^m)$ with

$$\beta(e_k) = u_k \in \mathbb{Z}^m \subset \mathbb{R}^m \qquad k = 1, \dots n.$$

Without loss of generality, after relabelling the basis vectors if necessary, we may assume that $\{u_1, \dots, u_m\}$ are linearly independent and hence form a basis of $\mathbb{R}^m$. Therefore, there exists a matrix $A \in \mathrm{Mat}_{n-m,m}(\mathbb{Q})$ such that

$$u_{m+i} = \sum_{k=1}^{m} A_{ik} u_k, \qquad i = 1, \dots, n-m.$$

Then $(z,w) \in \mathbb{C}^{n,n}$ lies in $\mu_N^{-1}(0)$ if there exists $a \in (\mathbb{R}^m)^*$ such that

$$a(u_k) = \mu(z,w)(e_k) = \frac{1}{2}(|z_k|^2 + |w_k|^2) + c_k, \qquad k = 1, \dots, n.$$

Using the expression of u_{m+i} in terms of the first m u_k's, we can rewrite the equation for u_{m+i} as

$$a(u_{m+i}) = \sum_{k=1}^{m} A_{ik} a(u_k) = \sum_{k=1}^{m} A_{ik}(\frac{1}{2}(|z_k|^2 + |w_k|^2) + c_k).$$

Thus, we obtain a system of m equations determining a and $m-n$ consistency conditions coming from the two equivalent formulae for $a(u_{m+i})$:

$$\begin{aligned} a(u_k) &= \frac{1}{2}(|z_k|^2 + |w_k|^2) + c_k, \quad k = 1, \dots, m \\ \frac{1}{2}(|z_{m+i}|^2 + |w_{m+i}|^2) + c_{m+i} &= \sum_{k=1}^{m} A_{ik}(\frac{1}{2}(|z_k|^2 + |w_k|^2) + c_k), \quad i = 1, \dots, n-m. \end{aligned}$$

Thus, deciding whether the $N^{\mathbb{C}}$-orbit through (z,w) contains a zero of the moment map μ_N amounts to finding $a \in (\mathbb{R}^m)^*$and $\lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_n) \in N^{\mathbb{C}}$ such that $\lambda.(z,w)$ satisfies the above equations. Equivalently, we want to find $a \in (\mathbb{R}^m)^*$ and scalars $\lambda_k \in \mathbb{R}_{>0} \subset \mathbb{C}^*$ such that

$$\begin{aligned}
a(u_k) &= \frac{1}{2}(\lambda_k^2|z_k|^2 + \lambda_k^{-2}|w_k|^2) + c_k, \quad k = 1, \ldots, m, \\
\frac{1}{2}(\lambda_{m+i}^2|z_{m+i}|^2 + \lambda_{m+i}^{-2}|w_{m+i}|^2) + c_{m+i} &= \sum_{k=1}^{m} A_{ik}(\frac{1}{2}(\lambda_k^2|z_k|^2 + \lambda_k^{-2}|w_k|^2) + c_k), \\
& \quad \text{for} \quad i = 1, \ldots, n-m, \\
1 &= \lambda_k \prod_{i=1}^{n-m} (\lambda_{m+i})^{A_{ik}}, \quad k = 1, \ldots, m.
\end{aligned}$$

The last equation is the condition that $\lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ lies in $N^{\mathbb{C}} \subset (\mathbb{C}^*)^n$. It is obtained by writing $\lambda_k = e^{\theta_k}$ for some $\theta_k \in \mathbb{R}$ and observing that

$$\mathrm{diag}(\lambda_1, \ldots, \lambda_n) \in N^{\mathbb{C}} \iff 0 = \sum_{k=1}^{n} \theta_k u_k = \sum_{k=1}^{m} \left(\theta_k + \sum_{i=1}^{n-m} \theta_{m+i} A_{ik} \right) u_k,$$

where we used the fact the we can express u_{m+i} in terms of $\{u_1, \ldots, u_m\}$ and the matrix A. The conditions follow then from the assumption that $\{u_1, \ldots, u_m\}$ are linearly independent and by exponentiating the so obtained constraint

$$\theta_k + \sum_{i=1}^{n-m} \theta_{m+i} A_{ik} = 0 \qquad k = 1, \ldots, m.$$

So altogether we have $(n+m)$ equations for $(n+m)$ unknowns (a, λ) and we expect a discrete set of solutions. We introduce the following notation:

$$a_k = a(u_k) - c_k.$$

Thinking of a as a parameter, we can view, after multiplication by λ_k^2, the first two sets of equations as quadratic equations in λ_k^2 which admit real solutions if and only if

$$a_k^2 - |z_k|^2|w_k|^2 \geq 0 \qquad k = 1, \ldots, n.$$

These are given by

$$\begin{aligned}
\lambda_k &= \frac{a_k \pm \sqrt{a_k^2 - |z_k|^2|w_k|^2}}{|z_k|^2} \qquad \text{if } z_k \neq 0 \neq w_k \\
\lambda_k &= \sqrt{\frac{2a_k}{|z_k|^2}} \qquad \text{if } w_k = 0 \neq z_k \\
\lambda_k &= \sqrt{\frac{|w_k|^2}{2a_k}} \qquad \text{if } z_k = 0 \neq w_k \\
\lambda_k &\in \mathbb{R}_{>0} \subset \mathbb{C}^* \text{ arbitrary} \qquad \text{if } w_k = 0 = z_k.
\end{aligned}$$

Now using the above definitions of a_k and λ_k,define the set $\check{M} \subset \mathbb{C}^{n,n} \times (\mathbb{R}^m)^*$,

$$\begin{aligned}\check{M} = \{((z,w),a) \in \mathbb{C}^{n,n} \times (\mathbb{R}^m)^* \quad | \quad & a_k^2 - |z_k|^2|w_k|^2 \geq 0 \qquad k = 1,\dots,n; \\ & 1 = \lambda_k \prod_{i=1}^{n-m} (\lambda_{m+i})^{A_{ik}} \qquad k = 1,\dots,m; \\ & a_k = 0, \text{ if } (z_k, w_k) = 0\}.\end{aligned}$$

Then

$$\tilde{M} = \mathrm{pr}_{\mathbb{C}^{n,n}}(\check{M}).$$

What are the fibres of the map $P : \mu^{-1}(0) \to \tilde{M}/N^{\mathbb{C}}$? We see from above that for fixed $a \in (\mathbb{R}^m)^*$ there are at most $2^{\dim N}$ possible solutions $\lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_n) \in N^{\mathbb{C}}$ to the moment map equations up to the action of N. On the other hand, once we are given some $a \in (\mathbb{R}^m)^*$ and $\lambda \in N^{\mathbb{C}}$, solving the moment map equations, then a is determined by λ by the relation

$$a(u_k) = \frac{1}{2}(\lambda_k^2|z_k|^2 + \lambda_k^{-2}|w_k|^2) + c_k, \qquad k = 1,\dots,m.$$

Finally, we note that we could have higher-dimensional fibres for special choices of parameters such that $a_k = 0$ and $(z_k, w_k) = (0,0)$ for at least $\mathrm{codim}(N) + 1$ many k's. □

Toric Hypersymplectic Quotients

We are now going to take the general result from the last section and derive some simple consequences for toric hypersymplectic quotients. First of all, we have to extend our setup. Details can again be found in [19]. Identifying $\mathbb{C}^{n,n} \cong \mathbb{B}^n$, we get a flat hypersymplectic structure upon defining S, T to be given by

$$S = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad T = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix},$$

giving parakähler forms

$$\begin{aligned}\omega_S &= g(S-,-) = \frac{1}{2}\sum_{k=1}^{n} \mathrm{d}z_k \wedge \mathrm{d}\bar{w}_k - \mathrm{d}w_k \wedge \mathrm{d}\bar{z}_k, \\ \omega_T &= g(T-,-) = \frac{1}{2i}\sum_{k=1}^{n} \mathrm{d}z_k \wedge \mathrm{d}\bar{w}_k + \mathrm{d}w_k \wedge \mathrm{d}\bar{z}_k.\end{aligned}$$

The holomorphic symplectic form $\omega_I^{\mathbb{C}}$ is then given by

$$\omega_I^{\mathbb{C}} = \omega_S + i\omega_T = \sum_{k=1}^{n} \mathrm{d}z_k \wedge \mathrm{d}\bar{w}_k.$$

Our standard T^n action clearly preserves this structure, and a simple computation gives the moment maps

$$\mu_I(z,w): \mathbb{C}^{n,n} \to \mathfrak{t}^* \qquad \mu_I(z,w) = \frac{1}{2}\sum_{k=1}^{n}(|z_k|^2 + |w_k|^2 + c_k^{(1)})\epsilon_k,$$

$$\mu_I^{\mathbb{C}}(z,w) = \mu_S + i\mu_T : \mathbb{C}^{n,n} \to \mathfrak{t}^* \otimes \mathbb{C} \qquad \mu_I^{\mathbb{C}}(z,w) = \sum_{k=1}^{n}(iz_k\bar{w}_k + c_k^{(2)} + ic_k^{(3)})\epsilon_k,$$

where $c_k^{(i)} \in \mathbb{R}$. We will often write $c_k^{(2)} + ic_k^{(3)} = C_k$. Now we come back to the three examples, discussed before from the point of view of pseudokähler geometry, and consider them now from the hypersymplectic point of view.

Example (The Standard Circle Action [48], [19])**.** We choose $m = n-1$ and put $u_k = e_k$ for $k = 1, \ldots, n-1$ and $u_n = -(e_1 + \cdots + e_n)$. So β is the map

$$\beta : \mathbb{R}^n \to \mathbb{R}^{n-1} \qquad \sum_{i=1}^{n}\theta_i e_i \mapsto \sum_{i=1}^{n-1}(\theta_i - \theta_n)e_i.$$

So

$$\mathfrak{n} = \mathbb{R}(e_1 + \cdots + e_n),$$

and α is just the inclusion. Therefore, $N \subset T^n$ is the diagonal circle. The moment map $\mu_I : \mathbb{C}^{n,n} \to \mathbb{R}$ associated to this action is given by $\alpha^*\mu$, i.e.

$$\mu_I(z,w) = \frac{1}{2}\sum_{k=1}^{n}(|z_k|^2 + |w_k|^2 + c_k)\alpha_k = \frac{1}{2}(|z|^2 + |w|^2) + c^{(1)}, \quad c^{(1)} = \sum_{k=1}^{n} c_k^{(1)} \in \mathbb{R}.$$

The induced $\mathbb{C}^*$-action is given by

$$(z,w) \mapsto (\lambda z, \bar{\lambda}^{-1}w),$$

and the complex moment map is

$$\mu_I^{\mathbb{C}} : \mathbb{C}^{n,n} \to \mathbb{C} \qquad \mu_I^{\mathbb{C}}(z,w) = \sum_{k=1}^{n} iz_k\bar{w}_k + c^{(2)} + ic^{(3)},$$

where $c^{(i)} \in \mathbb{R}, i = 1, 2$.

We would like to know which $\mathbb{C}^*$-orbits in $(\mu_I^{\mathbb{C}})^{-1}(0)$ meet the zero set of the real moment map. Recall that in order to check whether the $\mathbb{C}^*$-orbit through (z,w) meets $\mu_I^{-1}(0)$, we have to solve the following equation for λ real and positive.

$$\lambda^2|z|^2 + \lambda^{-2}|w|^2 + 2c = 0.$$

In other words

$$\lambda^4|z|^2 + \lambda^2 2c + |w|^2 = 0,$$

which, writing $x = \lambda^2$, gives

$$x_{1,2} = \frac{|c| \pm \sqrt{c^2 - |z|^2|w|^2}}{|z|^2}.$$

In this case we assume $(z, w) \in \mu_I^{\mathbb{C}}$, so we get the additional constraint

$$|z|^2|w|^2 = (c^{(2)})^2 + (c^{(3)})^2.$$

Thus, we see that, choosing the levels of the real and complex moment maps in such a way that $(c^{(2)})^2 + (c^{(3)})^2 = (c^{(1)})^2$, we get essential uniqueness:

$$\tilde{M} = \{(z, w) \in \mathbb{C}^{n,n} \backslash \{(0,0)\} \mid |z||w| = |c|\},$$

i.e. we single out the boundary of the previous $\tilde{M}$. On the other hand, if we choose the levels such that $0 \neq (c^{(2)})^2 + (c^{(3)})^2 < (c^{(1)})^2$, we get precisely two solutions every time.

Example $(N = T^{n-1} \subset T^n)$**.** In this example we choose $m = 1$, $\beta = (0, \ldots, 0, 1)$, that is $u_i = 0$ for $i = 1, \ldots, n-1$ and $u_n = 1$, so that $\mathfrak{n} = \mathbb{R}^{n-1} \times \{0\} \subset \mathbb{R}^n$ and

$$N = \{\mathrm{diag}(e^{i\theta_1}, \ldots, e^{i\theta_{n-1}}, 1)\} \cong T^{n-1} \subset T^n.$$

The moment maps are therefore given by

$$\begin{aligned} \mu_N(z,w) &= \sum_{k=1}^{n-1} (\frac{1}{2}(|z_k|^2 + |w_k|^2) + c_k^{(1)})\epsilon_k \\ \mu_I^{\mathbb{C}}(z,w) &= \sum_{k=1}^{n} (iz_k\bar{w}_k + c_k^{(2)} + ic_k^{(3)})\epsilon_k. \end{aligned}$$

Again, for each k we get an individual system of two equations which we just solve writing $x_k = \lambda_k^2$,

$$x_{k;1,2} = \frac{|c_k| \pm \sqrt{(c_k^{(1)})^2 - |z_k|^2|w_k|^2}}{|z_k|^2},$$

with the additional constraint $|z_k|^2|w_k|^2 = |C_k|^2$. This is solvable if $(c_k^{(1)})^2 - |z_k|^2|w_k|^2 \geq 0$, giving the condition that the levels should satisfy $(c_k^{(1)})^2 \geq |C_k|^2$. Altogether, combining the solutions of the individual equations in every possible way, we see that generically we get 2^D many solutions, where D is the number of cases in which the inequality $c_k^2 \geq |z_k|^2|w_k|^2$ is strict and both z_k and w_k are non-zero, i.e. the number of cases in which we get two solutions. Note that we obtain essential uniqueness, if we choose the parameters such that $|C_k|^2 = c_k^2$. In the above notation, we have

$$\tilde{M} = \{(z,w) \in \mathbb{C}^{n,n} \mid c_k^2 - |z_k|^2|w_k|^2 \geq 0 \ \forall k = 1, \dots, n-1\},$$

And again, the map

$$P : \mu_N^{-1}(0)/N \to \tilde{M}/N^{\mathbb{C}}$$

restricts to a $2^D : 1$ cover on the "D-strata" of $\tilde{M}$, i.e. the sets of points in $(z,w) \in \tilde{M}$ where there exist two solutions in precisely D cases, as D ranges over $0, 1, \dots, n-1$.

These examples suggest that due to the freedom in choosing the level sets, toric hypersymplectic quotients should be better behaved in general than toric pseudokähler quotients. In particular, for special choices of the levels of the moment maps, the interpretation of hypersymplectic quotients as holomorphic symplectic quotients is valid.

2.4 The ASD Equations on $\mathbb{R}^{2,2}$ and their Dimensional Reductions

2.4.1 ASD Connections on Hypersymplectic Manifolds

We now try to carry over the gauge theoretic constructions from section 1.2.5 to the situation where the base four-manifold is hypersymplectic instead of hyperkähler. As already observed in the finite dimensional situation, the hypersymplectic quotient construction is more subtle than its hyperkähler analogue. We have to worry about degeneracies. The discussion of these in a general setup is included in this section and we present a general formal degeneracy criterion.

In the subsequent sections we restrict our attention to $\mathbb{R}^{2,2}$ and discuss various special cases. We show that our degeneracy criterion gives a natural geometric interpretation for explicit formulae for degeneracy spaces found earlier by other authors.

If we work on a hypersymplectic four-manifold, we have again that $*^2 = \mathrm{id}$ on two-forms and that the three Kähler forms $\omega_I, \omega_S, \omega_T$ span the space of self-dual two-forms. Since we are now no longer working with a positive definite metric but instead in split signature, we cannot apply all the analytical machinery introduced in the last chapter. However, at least formally, the construction can be imitated. So we get again an induced hypersymplectic structure on the space $\mathcal{A}$ of unitary connections given by $-I, S, T$ acting on one-forms. The three Kähler forms are again (up to sign) given by the formulae

$$\tilde{\omega}_i(A, B) = \int_M \operatorname{tr}(A \wedge B) \wedge \omega_i.$$

The gauge group $\mathcal{G}$ acts as seen above preserving this structure with hypersymplectic moment map

$$\mu = (\mu_I, \mu_S, \mu_T) : M \to \Gamma(\mathfrak{su}(E))^* \otimes \mathbb{R}^3$$

given by

$$\mu_i(\nabla) = R^\nabla \wedge \omega_i \qquad i \in \{I, S, T\},$$

where we used again the embedding $\Gamma(\mathfrak{u}(E))^* \to \Omega^4(M, \mathfrak{u}(E))$. However, in order to apply the hypersymplectic quotient construction now, we have to worry about degeneracy spaces.

Proposition 2.4.1 (Degeneracy Criterion)**.** *In the situation described above we have at an irreducible connection* $\nabla \in \mathcal{A}$

$$T_\nabla \mathcal{O}_\nabla \cap T_\nabla \mathcal{O}_\nabla^\perp = \ker((\mathrm{d}^\nabla)^* \mathrm{d}^\nabla : \Gamma(M, \mathfrak{u}(E)) \to \Gamma(M, \mathfrak{u}(E))),$$

where the adjoint is taken with respect to the L^2 *inner product on* $\Omega^1(M, \mathfrak{u}(E))$ *coming from the split signature metric* g*. Thus, this could be a non-trivial space in general.*

Proof. As seen earlier the tangent space to an orbit of the gauge group action is given by

$$T_\nabla \mathcal{O}_\nabla = \{\mathrm{d}^\nabla \xi \mid \xi \in \Gamma(M, \mathfrak{u}(E))\}.$$

Its orthogonal complement is given by

$$\begin{aligned} T_\nabla \mathcal{O}_\nabla^\perp &= \{A \in \Omega^1(M, \mathfrak{u}(E)) \mid g_{L^2}(A, \mathrm{d}^\nabla \xi) = 0 \ \forall \xi \in \Gamma(M, \mathfrak{u}(E))\} \\ &= \{A \in \Omega^1(M, \mathfrak{u}(E)) \mid g_{L^2}((\mathrm{d}^\nabla)^* A, \xi) = 0 \ \forall \xi \in \Gamma(M, \mathfrak{u}(E))\} \\ &= \ker((\mathrm{d}^\nabla)^*). \end{aligned}$$

Combining these two observations, we obtain the result. □

The point is, that since our metric is not positive definite, the operator $(\mathrm{d}^\nabla)^* \mathrm{d}^\nabla$ is not elliptic, but instead has the same symbol as a *ultrahyperbolic wave operator* and so its kernel is not automatically the same as the kernel of d^∇. In fact, it could be infinite-dimensional. There may be non-parallel sections $\xi \in \Gamma(M, \mathfrak{u}(E))$, whose covariant derivative is a null vector at every point. Since our connection is assumed to be irreducible, they precisely comprise the degeneracy space.

Notice also that the deformation complex

$$0 \to L^2_k(\mathfrak{u}(E)) \to \Omega^1_{k-1}(M, \mathfrak{u}(E)) \to \Omega^+_{k-2}(M, \mathfrak{u}(E)) \to 0,$$

is no longer elliptic and so its cohomology spaces need not be finite dimensional. We therefore expect the moduli space to be of infinite dimension as is suggested by the discussion of the ASD moduli space on the pseudokähler manifold $S^2 \times S^2$ in [45].

ASD Connections on $\mathbb{R}^{2,2}$

In the following chapters we take a closer look at the ASD equations on $\mathbb{R}^{2,2} \cong \mathbb{B}$, the standard example of a hypersymplectic four-manifold. This is a non-compact manifold and so the methods of the previous section do not apply directly. It seems rather complicated to impose correct decay conditions in order to make the constructions work. The problem is that the equations are not elliptic. And because of the neutral signature of the metric, there are distinguished directions and in particular the conformal compactification of $\mathbb{R}^{2,2}$ does not coincide with the one-point compactification as in the case of Euclidean signature. We rather have to add in a whole hypersurface at infinity, see the discussion in chapter 5 for more details. Therefore, in this section we just derive the equations and discuss hypersymplectic aspects of the moduli space later.

On $\mathbb{R}^{2,2} = \mathbb{R}^4$ every vector bundle is trivial and we will assume in the following that our unitary vector bundle is already trivialised, $E \cong \mathbb{R}^{2,2} \times \mathbb{C}^n$, such that the metric h maps to the standard hermitian inner product on $\mathbb{C}^n$. So we get a preferred choice of background connection, the trivial connection d, given by the exterior derivative acting component-wise on vector-valued functions. Hence,

$$\mathcal{A} = \mathrm{d} + \Omega^1(\mathbb{R}^{2,2}, \mathfrak{u}(n)).$$

In our standard orthonormal co-ordinates and under the above isomorphism for $\mathcal{A}$ we can write any connection ∇ on E in the form

$$\nabla = \mathrm{d} + \sum_{i=1}^{4} A_i \mathrm{d}x_i.$$

Given a connection $\nabla = \mathrm{d} + \sum_i A_i \mathrm{d}x_i$, its curvature is then given by

$$R^\nabla = \sum_{i<j} R_{ij} \mathrm{d}x_i \wedge \mathrm{d}x_j.$$

Using the notation $\nabla_i = \frac{\partial}{\partial x^i} + A_i$, we can write the components of R^∇ in the form

$$R_{ij} = [\nabla_i, \nabla_j].$$

Then, formally, the moment map equations $R^\nabla \wedge \omega_i = 0$ for $i = I, S, T$, i.e. the anti-self-dual Yang-Mills equations, read

$$\begin{aligned} R_{12} &= -R_{34} \\ R_{13} &= -R_{24} \\ R_{14} &= R_{23}. \end{aligned}$$

Also, we write down explicitly our degeneracy criterion found in the last section for ASD connections on $\mathbb{R}^{2,2}$.

Corollary 2.4.2. *On $\mathbb{R}^{2,2}$ we have on $\Gamma(\mathfrak{u}(n))$*

$$-(\mathrm{d}^\nabla)^* \mathrm{d}^\nabla = \nabla_1\nabla_1 + \nabla_2\nabla_2 - \nabla_3\nabla_3 - \nabla_4\nabla_4 = \sum_{i=1}^4 g_{ii}\nabla_i\nabla_i.$$

An element $\xi \in \mathrm{Lie}(\mathcal{G})$ lies in the degeneracy space if and only if

$$\sum_{i=1}^4 g_{ii}\nabla_i(\nabla_i\xi) = 0.$$

We can consider the equations on a bounded domain $U \subset \mathbb{R}^{2,2}$. We then only allow gauge transformations that are the identity on ∂U. The Lie algebra of the gauge group is then given by sections of $\mathfrak{u}(E)$ that vanish on ∂U. The degeneracy space then consists of connections $\mathrm{d} + \sum_i A_i \mathrm{d}x_i$ defined on U such that there is a solution $\xi \in \Gamma(U, \mathfrak{u}(E))$ of the boundary value problem

$$\sum_{i=1}^4 g_{ii}\left(\frac{\partial^2\xi}{\partial x_i^2} + 2\left[A_i, \frac{\partial \xi}{\partial x_i}\right] + \left[\frac{\partial A_i}{\partial x_i}, \xi\right] + [A_i, [A_i, \xi]]\right) = 0, \qquad \xi \equiv 0 \text{ on } \partial U.$$

The ASD equations and their dimensional reductions will be the subject of the further chapters of this thesis. We will discuss their moduli spaces and the degeneracy loci of the induced hypersymplectic structures.

2.4.2 The Lax Pair Formalism

In this section, we develop a complex-geometric viewpoint on the equations and derive a Lax pair formulation. We begin with a construction closely analogous to Hitchin's construction of the twistor space of a hyperkähler manifold in [31], but adapted to the neutral signature geometry of the split quaternions.

On any hypersymplectic manifold M we have a two-sheeted hyperboloid of complex structures with respect to which the metric is pseudokähler:

$$C = \{x_1 I + x_2 S + x_3 T \mid x_1^2 - x_2^2 - x_3^2 = 1\}.$$

We could say that in both the hyperkähler and the hypersymplectic situation we have a set of complex structures given by the unit sphere in $\mathrm{Im}\mathbb{H}$ and $\mathrm{Im}\mathbb{B}$ with respect to the respective induced natural metrics. Now we consider the compactification of C, that is we embed C as $\bar{C}$ into $\mathbb{RP}^3$. $\bar{C}$ is then given by the homogenised equation, that is,

$$\bar{C} = \{x = [x_1, x_2, x_3, x_4] \in \mathbb{RP}^3 \mid x_1^2 - x_2^2 - x_3^2 - x_4^2 = 0\}.$$

Observe that both for points in C and $\bar{C}$ we always have $x_1 \neq 0$. Therefore, we can rewrite this as

$$\bar{C} = \{x = [1, x_2, x_3, x_4] \in \mathbb{RP}^3 \mid x_2^2 + x_3^2 + x_4^2 = 1\}.$$

This equivalent representation shows that $\bar{C}$ is in fact the same as S^2 embedded into $\mathbb{RP}^3$ via $x \mapsto [1, x]$. We can think of $\bar{C}$ as the hyperboloid C together with a circle added at infinity. The circle is given by points in $\bar{C}$ with $x_4 = 0$, i.e. points of the form $x = [1, \sin\theta, \cos\theta, 0]$. These points do not correspond to complex structures on M, since the associated endomorphism of TM, given by $I + \sin\theta S + \cos\theta T$, squares to zero. Note, however, that the circle at infinity parametrises the circle of compatible product structures $S_\theta = \cos\theta S - \sin\theta T$, which anti-commute with the complex structure I.

We now introduce a complex parameter on $\bar{C}$, and so on C, via the identification $S^2 \cong \mathbb{CP}^1$ in the following way:

$$[1, x_2, x_3, x_4] = \left[1, \frac{\zeta + \bar{\zeta}}{1 + \zeta\bar{\zeta}}, \frac{i(\zeta - \bar{\zeta})}{1 + \zeta\bar{\zeta}}, \frac{1 - \zeta\bar{\zeta}}{1 + \zeta\bar{\zeta}}\right].$$

Where $x_4 \neq 0$, that is on C, we can write

$$x = [x_1, x_2, x_3, x_4] = \left[\frac{1 + \zeta\bar{\zeta}}{1 - \zeta\bar{\zeta}}, \frac{\zeta + \bar{\zeta}}{1 - \zeta\bar{\zeta}}, \frac{i(\zeta - \bar{\zeta})}{1 - \zeta\bar{\zeta}}, 1\right].$$

In this picture, the circle at infinity corresponds to the circle $|\zeta| = 1$ and the two sheets of the hyperboloid correspond to $\{|\zeta| < 1\}$ and $\{|\zeta| > 1\}$. We can form a new space Z as the product of M and C,

$$Z = M \times C.$$

For each $\zeta \in C$ we get a complex structure on M, given by

$$I_\zeta = \frac{1+\zeta\bar{\zeta}}{1-\zeta\bar{\zeta}} I + \frac{\zeta+\bar{\zeta}}{1-\zeta\bar{\zeta}} S + \frac{i(\zeta-\bar{\zeta})}{1-\zeta\bar{\zeta}} T.$$

So we put on Z the complex structure $\mathbf{I}$ given at a point (p,ζ) by $\mathbf{I}_{p,\zeta} = (I_\zeta \oplus i) \in \mathrm{End}(T_pM \oplus T_\zeta C)$. In the following lemma we show that $\mathbf{I}$ is integrable and compute its $(1,0)$ forms.

Lemma 2.4.3. *Let (M^{4k}, g, I, S, T) be a hypersymplectic manifold and let $\phi \in \Omega^1(M, \mathbb{C})$ be of type $(1,0)$ with respect to the complex structure I. Then*

$$\theta = \phi + \zeta T\phi$$

is of type $(1,0)$ with respect to the complex structure $\mathbf{I}$ on Z and moreover $\mathbf{I}$ is integrable.

Proof. We imitate the proof of the analogous assertion for the twistor space of a hyperkähler manifold given in [31]. Firstly we use $I\phi = i\phi$ and the algebraic identities satisfied by I, S, T to calculate

$$\begin{aligned} I\theta &= i\phi - i\zeta T\phi \\ S\theta &= iT\phi - i\zeta\phi \\ T\theta &= T\phi + \zeta\phi. \end{aligned}$$

Using this, we compute using $S = TI$

$$\begin{aligned} (1-\zeta\bar{\zeta})\mathbf{I}\theta &= (1+\zeta\bar{\zeta})I\theta + (\zeta+\bar{\zeta})S\theta + i(\zeta-\bar{\zeta})T\theta \\ &= (1+\zeta\bar{\zeta})(i\phi - i\zeta T\phi) + (\zeta-\bar{\zeta})(iT\phi - i\zeta\phi) + i(\zeta-\bar{\zeta})(T\phi+\zeta\phi) \\ &= i\phi + i\zeta\bar{\zeta}\phi - i\zeta T\phi - i\zeta^2\bar{\zeta}T\phi + i\zeta T\phi - i\bar{\zeta}T\phi - i\zeta^2\phi \\ &\quad + i\zeta\bar{\zeta}\phi + i\zeta T\phi - i\bar{\zeta}T\phi + i\zeta^2\phi - i\bar{\zeta}\zeta\phi \\ &= i\phi + i\zeta T\phi - i\zeta\bar{\zeta}\zeta T\phi - i\zeta\bar{\zeta}\phi \\ &= i(1-\zeta\bar{\zeta})(\phi + \zeta T\phi) \\ &= i(1-\zeta\bar{\zeta})\theta. \end{aligned}$$

To show that $\mathbf{I}$ is integrable, we have to check that the exterior derivative maps $(1,0)$-forms to $\Omega^{2,0} \oplus \Omega^{1,1}$. In other words, we have to show that $d\theta$ does not have a component of type $(0,2)$.

$$\mathrm{d}\theta = \mathrm{d}(\phi + \zeta T\phi) = \sum_i \mathrm{d}x_i \wedge \nabla_{\frac{\partial}{\partial x^i}}(\phi + \zeta T\phi) + d\zeta \wedge T\phi.$$

The second term is easy to handle. It is the wedge product of the $(1,0)$-form $\mathrm{d}\zeta$ with another one-form and hence of the desired type. The first sum also does not have a $(0,2)$ component, as $\mathbf{I}$ is parallel with respect to the Levi-Civita connection ∇ and thus $\mathbf{I}$ and ∇ commute:

$$\mathbf{I}\nabla_{\frac{\partial}{\partial x^i}}\theta = \nabla_{\frac{\partial}{\partial x^i}}\mathbf{I}\theta = i\nabla_{\frac{\partial}{\partial x^i}}\theta.$$

Hence, the first term is also given by the wedge product of a $(1,0)$-form with another one-form. So the complex structure $\mathbf{I}$ is, in fact, integrable and Z is thus a complex manifold of complex dimension $2k+1$. □

In fact (M, I_ζ) is a complex symplectic manifold. To see this, we recall that (M, I) carries a holomorphic symplectic form $\omega^{\mathbb{C}} = \omega_S + i\omega_T$. Picking a local Darboux basis for the $(1,0)$-forms, we can write

$$\omega^{\mathbb{C}} = 2\sum_{i=1}^{2k} \phi_i \wedge \phi_{2k+i}.$$

This is a holomorphic $(2,0)$-form with respect to the complex structure I, which corresponds to $\zeta = 0$. So we guess that the form

$$\omega_\zeta = 2\sum_{i=1}^{2k}(\phi_i + \zeta T\phi_i) \wedge (\phi_{2k+i} + \zeta T\phi_{2k+i})$$

will be a holomorphic symplectic form with respect to $\mathbf{I}$. Again, this is completely analogous to the twistor space construction from [31].

Proposition 2.4.4. *For each ζ in C ω_ζ is a holomorphic symplectic form on M with respect to the complex structure I_ζ. In terms of $\omega_I, \omega_S, \omega_T$ it may be written as*

$$\omega_\zeta = \omega_S + i\omega_T - 2\zeta\omega_I + \zeta^2(\omega_S - i\omega_T).$$

Proof. Using the expression for ω_ζ given above, we compute

$$\begin{aligned} \frac{1}{2}\omega_\zeta &= \sum_i (\phi_i + \zeta T\phi_i) \wedge (\phi_{2k+i} + \zeta T\phi_{2k+i}) \\ &= \sum_i \phi_i \wedge \phi_{2k+i} + \zeta((T\phi_i) \wedge \phi_{2k+i} + \phi_i \wedge T\phi_{2k+i}) + \zeta^2 T\phi_i \wedge T\phi_{2k+i}. \end{aligned}$$

Now we just check that the coefficients of ζ and ζ^2 coincide with the ones that appear in the statement of the proposition. The term constant in ζ is equal to $\omega^{\mathbb{C}} = \omega_S + i\omega_T$ by our choice of the ϕ_i's. Consider the term linear in ζ. Given two vector fields

$X, Y \in \Gamma(M, TM)$, we compute using the algebraic identities satisfied by I, S, T (suppressing the summation symbol):

$$\begin{aligned} 2((T\phi_i)\wedge\phi_{2k+i}+\phi_i\wedge T\phi_{2k+i})(X,Y) &= \phi_i(TX)\phi_{2k+i}(Y)+\phi_i(X)\phi_{2k+i}(TY) \\ &\quad -\phi_i(TY)\phi_{2k+i}(X)-\phi_i(Y)\phi_{2k+i}(TX) \\ &= \omega^{\mathbb{C}}(TX,Y)+\omega^{\mathbb{C}}(X,TY) \\ &= g(STX,Y)+ig(X,Y)+g(SX,TY) \\ &\quad +ig(TX,TY) \\ &= -g(IX,Y)+ig(X,Y)-g(IX,Y) \\ &\quad -ig(X,Y), \\ &= -2\omega_I(X,Y), \end{aligned}$$

where in the last two lines we used the skew-adjointness of I, S and T. Finally, we treat the quadratic term and obtain

$$\begin{aligned} 2T\phi_i\wedge T\phi_{2k+i}(X,Y) &= \phi_i(TX)\phi_{2k+i}(TY)-\phi_i(TY)\phi_{2k+i}(TX) \\ &= \omega^{\mathbb{C}}(TX,TY) \\ &= g(STX,TY)+ig(TTX,TY) \\ &= -g(TSX,TY)-ig(TX,Y) \\ &= g(SX,Y)-ig(TX,Y) \\ &= (\omega_S-i\omega_T)(X,Y). \end{aligned}$$

This expression shows that ω_ζ is actually parallel, hence holomorphic, and moreover it depends holomorphically on ζ. The non-degeneracy is clear from this expression together with the non-degeneracy of $\omega^{\mathbb{C}}$. □

This quadratic dependence of ω_ζ shows that the twistor space Z carries a twisted holomorphic symplectic form ω with values in the line bundle $\mathrm{pr}_C^*\mathcal{O}(2)$ given by

$$\omega_{(p,\zeta)} = \omega_\zeta(p).$$

Here $\mathrm{pr}_C : Z = M \times C \to C \to \mathbb{CP}^1$ is the projection followed by the inclusion. In other words, ω is a section of the bundle $\Lambda^2 TZ^* \otimes \mathrm{pr}_C^*\mathcal{O}(2)$.

Note that the twistor space Z is, in contrast to its hyperkähler analogue, disconnected, since $C = \mathbb{CP}^1 \setminus \{|\zeta| = 1\}$ is. We observe that Z carries a natural involution τ induced by inversion with respect to the unit circle on $\mathbb{CP}^1$,

$$\tau(m,\zeta) = (m, \frac{1}{\bar{\zeta}}).$$

This is anti-holomorphic, fixes the circle $\{|\zeta| = 1\}$, i.e. the circle at infinity and interchanges the two sheets of the hyperboloid of complex structures. The fixed points of τ, i.e. the unit circle in the above co-ordinates, correspond, as we have seen, to endomorphisms that square to zero. Alternatively, if we fix the complex structure I, then this real circle parametrises the product structures that anti-commute with I and are skew-adjoint with respect to the hypersymplectic metric.

If we now have a group G acting on M preserving the hypersymplectic structure, it also preserves ω_ζ for all $\zeta \in C$ and we get an associated holomorphic moment map

$$\mu_\zeta = \mu_S + i\mu_T - 2\zeta\mu_I + \zeta^2(\mu_S - i\mu_T).$$

We see that the condition $\mu_\zeta = 0$ for all ζ is equivalent to the simultaneous vanishing of all three moment maps.

We now apply these ideas in the infinite dimensional context of the ASD equations on $\mathbb{R}^{2,2}$. We found for the moment maps

$$\mu_i(\nabla) = R^\nabla \wedge \omega_i.$$

So our complex moment map is given by

$$\mu_\zeta(\nabla) = R^\nabla \wedge \omega_\zeta.$$

Earlier on, we discussed the action of I, S, T on $\mathbb{B}$ with respect to the canonical basis and obtained explicit formulae for the matrices corresponding to I, S, T and the Kähler forms $\omega_I, \omega_S, \omega_T$. Note in particular, that complex co-ordinates with respect to the complex structure I are given by $z = x_1 - ix_2$ and $w = x_3 + ix_4$. Writing $Z = \nabla_1 - i\nabla_2$ and $W = \nabla_3 + i\nabla_4$ with their adjoints $Z^* = -\nabla_1 - i\nabla_2$, and $W^* = -\nabla_3 + i\nabla_4$, we arrive at

$$\begin{aligned} \mu_I(\nabla) &= \frac{i}{2}([Z, Z^*] - [W, W^*]) \\ (\mu_S + i\mu_T)(\nabla) &= [Z, W] \\ (\mu_S - i\mu_T)(\nabla) &= [Z^*, W^*]. \end{aligned}$$

We remark that the vanishing of the three moment maps μ_I, μ_S, μ_T is equivalent to the condition $\mu_I = 0$ and $\mu^{\mathbb{C}} = \mu_S + i\mu_T = 0$. This viewpoint will play a role later on and we call the first equation the *real equation* and the second one the *complex equation*.

Putting the terms together, we obtain for μ_ζ

$$\mu_\zeta(\nabla) = [Z, W] - i\zeta([Z, Z^*] - [W, W^*]) + \zeta^2[Z^*, W^*],$$

which is the same as

$$\mu_\zeta(\nabla) = [Z - i\zeta W^*, W - i\zeta Z^*].$$

So $\mu_\zeta = 0$ is equivalent to the condition $[Z - i\zeta W^*, W - i\zeta Z^*] = 0$ for all $\zeta \in C$. Or equivalently, we can modifiy this to

$$[Z - i\zeta W^*, W - i\zeta(Z + Z^*) - \zeta^2 W^*] = 0.$$

It will be useful for us to note the explicit expressions

$$\mu_\zeta(\nabla) = [\nabla_1 - i\nabla_2 + i\zeta(\nabla_3 - i\nabla_4), \nabla_3 + i\nabla_4 + i\zeta(\nabla_1 + i\nabla_2)]$$

and the modified one

$$[\nabla_1 - i\nabla_2 + i\zeta(\nabla_3 - i\nabla_4), \nabla_3 + i\nabla_4 - 2\zeta\nabla_2 + \zeta^2(\nabla_3 - i\nabla_4)] = 0.$$

We remark that these equations make sense for all $\zeta \in \mathbb{CP}^1$. But only for $|\zeta| \neq 1$ we get a holomorphic symplectic form ω_ζ as the endomorphism I_ζ squares to zero if $|\zeta| = 1$. Nevertheless, we can consider the equation $\mu_\zeta = 0$ as a formal combination of moment maps. For later purposes we therefore define $\bar{Z} = M \times \mathbb{CP}^1$, a compactified twistor space. The above complex structure on $Z = \bar{Z} \backslash \{|\zeta| = 1\}$ does not, however, extend to $\bar{Z}$.

2.4.3 Remarks on Twistors in Split Signature

The construction described above is suitable to obtain Lax pair formulations for PDEs that arise as hypersymplectic moment maps and was very much inspired by the construction of the twistor space of a hyperkähler manifold [31]. Originally, the aim of twistor theory is to encode information about the conformal geometry of a smooth manifold via its twistor space into complex-algebraic data, so that then questions concerning the (conformal differential) geometry of M, but also the study of linear and non-linear field equations on M, can be investigated using powerful complex (algebraic-)geometric machinery on the twistor space [47],[3],[58].

In the Riemannian case in real dimension four, the twistor space Z is constructed as follows. Given a four-dimensional oriented Riemannian manifold (M, g), the twistor space is defined to be the bundle of maximal isotropic subspaces of $TM \otimes \mathbb{C}$ with respect to the complex-bilinear extension of the metric g to $TM \otimes \mathbb{C}$. In the case of a Riemannian, i.e. positive definite, metric every such maximal isotropic subspace $V \subset T_pM \otimes \mathbb{C}$ satisfies $V \cap \bar{V} = \{0\}$ and thus defines an almost complex structure I_V orthogonal with respect to g by declaring that V should be its $+i$-eigenspace. Such complex structures come in two disjoint families - either they are compatible with the given orientation on M or they induce the opposite orientation. One then picks the family which is compatible with the orientation and defines the twistor space to be the bundle of such null planes, called β-planes.

Thus, in the Riemannian case, we can view the twistor space Z as the bundle of compatible complex structures, and this bundle has fibre S^2. Our construction above was largely inspired by this interpretation. We identify the tangent bundle to the fibre with the tangent bundle of $\mathbb{CP}^1$. Note that we have a connection on the bundle Z induced from the metric on M, so we can associate to a β-plane in $T_pM \otimes \mathbb{C}$ its horizontal lift, which thus gives rise to a complex structure on the horizontal spaces. Thus, we have put an almost complex structure on the manifold Z and in [6] it is shown that the integrability of this almost complex structure on the twistor space is equivalent to the metric on M being self-dual, i.e. having self-dual Weyl tensor.

In split signature, there are also twistor space constructions, some of which follow a route similar to ours above, but there are also different ones. Inspired by the twistor theory for quaternionic Kähler manifolds [52], Blair et al. in [11] have given a general twistor space construction for split-quaternionic Kähler manifolds, i.e. manifolds (M, g) with a subbundle of $\mathrm{End}(TM)$ locally spanned by a triple $I, S, T \in \Gamma(M, \mathrm{End}(M))$ satisfying the split-quaternionic relations, which is moreover preserved by the Levi-Civita connection of the split-signature metric g. Here the twistor space is a bundle over M with fibre given by (one sheet of) the hyperboloid of complex structures as we have discussed above. In their paper it is shown that in contrast to the quaternionic situtation, where one has a twistor space with fibre $\mathbb{CP}^1$, these *hyperbolic twistor spaces* carry many global non-constant holomorphic functions. In split signature, there are actually two twistor spaces one can associate with a split-quaternionic Kähler manifold by either considering compatible complex structures or compatible product structures. See also the paper [1].

In the case of a split-signature zollfrei metric on a four-manifold, for example the standard scalar flat split-signature metric on $S^2 \times S^2$, LeBrun and Mason in [42] were able to construct a compact twistor space as follows. The metric g being of signature $(2, 2)$ means that the tangent bundle of M splits into two-dimensional subbundles $T_+ \oplus T_-$ on which g is positive, respectively negative, definite. The manifold (M, g) is said to be space-time orientable, if these two subbundles are orientable. Due to the neutral signature of the metric, there exist complex β-planes V_p which are the complexification of *real* null planes. In this case, $V_p \cap \bar{V}_p$ is of real dimension two, so that the interpretation in terms of compatible complex structures breaks down.

Moreover, as we have seen above, the total space of the bundle of complex non-real β-planes, i.e. the bundle of compatible complex structures, has two connected components (we have a two-sheeted hyperboloid in each fibre) distinguished by the induced orientation on T_+. It follows that in this situation we obtain a compact disc bundle $\mathcal{Z}_+ \to M$. With suitable conventions, the fibres are given by the loci $\{|\zeta| \leq 1\}$ in the picture described above. Let $F = \partial\mathcal{Z}_+$ be the bundle of real β-planes. It turns out that the self-duality of the metric and the zollfrei condition together impliy that the distribution on F given by the horizontal lifts is integrable and that the space P of leaves is a manifold diffeomorphic to $\mathbb{RP}^3$. So $F \to P$ is a fibration with fibres given by embedded two-spheres. The twistor space Z is then given by collapsing the boundary $\partial\mathcal{Z}_+$ along this fibration.

The original manifold M may then be interpreted as the parameter space for a family of holomorphic disks in $\mathbb{CP}^3$ whose boundary lies on $\mathbb{RP}^3$. We will come back to this setup in chapter 5, when we discuss split-signature instantons on $S^2 \times S^2$.

Twistor spaces in split signature arise for example in the study of the relation between real integral transforms and the Pensrose transform, [7],[8]. Motivated by this, Eastwood and Graham [24] have considered the case of $M = \mathbb{R}^{2,2}$, whose conformal compactification is given by the Grassmannian $\mathrm{Gr}(2, \mathbb{R}^4)$ (see the discussion in chapter 5), in the context of *involutive structures*. An involutive structure on a manifold Z is a distribution $V \subset TZ \otimes \mathbb{C}$ which is formally integrable in the sense that its space of sections is closed under the Lie bracket. The almost complex structure on the twistor space Z, i.e. the bundle of complex β-planes, can be viewed as a involutive structure on the complement of the submanifold $Z_{\mathbb{R}}$ in Z corresponding to the real β-planes. In [24] it is shown that on the real blow-up B of Z along $Z_{\mathbb{R}}$ the pull-back of this involutive structure extends to all of B and satisfies a special property called *hypocomplexity*.

Chapter 3

Schmid's Equations

We now discuss a dimensional reduction of the ASD equations called *Schmid's equations*. Originally, they were introduced by W. Schmid in [53], where they arose in the study of singularities of families of algebraic manifolds.

We think of Schmid's equations as the hypersymplectic analogue of *Nahm's equations*, a system of non-linear ODEs, which play an important role in hyperkähler geometry. Nahm moduli spaces are hyperkähler and can be used, for instance, to construct various hyperkähler structures associated to Lie groups, see [18]. So it is an interesting question to ask whether Schmid's equations could play a similar role in hypersymplectic geometry. There is also a way to analyse them more from the point of view of integrable systems, and we are going to study both aspects of the equations in this section.

3.1 The Hypersymplectic Setup

Schmid's equations correspond to ASD connections $\nabla = \mathrm{d} + \sum_{i=1}^{4} A_i \mathrm{d}x_i = \sum_i \nabla_i \mathrm{d}x_i$ on $\mathbb{R}^{2,2}$, such that the connection matrices A_i are independent of x_2, x_3, x_4 and so effectively are defined on the real line $\mathbb{R} \cong \mathbb{R}e_1$. We recall that the ASD equations on $\mathbb{R}^{2,2}$ are

$$\begin{aligned} [\nabla_1, \nabla_2] &= -[\nabla_3, \nabla_4], \\ [\nabla_1, \nabla_3] &= -[\nabla_2, \nabla_4], \\ [\nabla_1, \nabla_4] &= [\nabla_2, \nabla_3]. \end{aligned}$$

As introduced earlier, we wrote

$$\nabla_i = \frac{\partial}{\partial x_i} + A_i, \quad i = 1, \dots, 4,$$

for the covariant partial derivatives in this equation. We think of x_1 as time and hence relabel it as t. Since we assume that the A_i only depend on t, the equations become

$$\begin{aligned}
[\frac{\mathrm{d}}{\mathrm{d}t} + A_1, A_2] &= -[A_3, A_4], \\
[\frac{\mathrm{d}}{\mathrm{d}t} + A_1, A_3] &= -[A_2, A_4], \\
[\frac{\mathrm{d}}{\mathrm{d}t} + A_1, A_4] &= [A_2, A_3].
\end{aligned}$$

We relabel $T_i = A_{i+1}$ for $i = 0, \ldots, 3$ as is customary for Nahm's equations and abbreviate differentiation with respect to t by a dot. So we arrive at the following definition.

Definition 3.1.1. Let $U \subset \mathbb{R}$ be a closed interval, possibly infinite. A quadruple of $\mathfrak{u}(n)$-valued functions (T_0, T_1, T_2, T_3), $T_i : U \to \mathfrak{u}(n)$ is said to satisfy *Schmid's equations* if

$$\begin{aligned}
\dot{T}_1 + [T_0, T_1] &= -[T_2, T_3], \\
\dot{T}_2 + [T_0, T_2] &= [T_3, T_1], \\
\dot{T}_3 + [T_0, T_3] &= [T_1, T_2].
\end{aligned}$$

We will now define the moduli space of solutions to Schmid's equations and identify the degeneracy locus of the induced hypersymplectic structure. We first have to develop a general hypersymplectic setup adapted to this situation.

Denote by $\mathcal{A}$ the space of all unitary connections on a trivial hermitian vector bundle E of rank n over $U \times \mathbb{R}^{1,2} \subset \mathbb{R}^{2,2}$. Instead of the space $\mathcal{A}$ we consider the subspace $\mathcal{A}_{Sch}$, consisting of connections whose connection matrices only depend on $t \in U$. This is again an affine space with tangent space isomorphic to the space of $\mathfrak{u}(E)$-valued one-forms on $U \times \mathbb{R}^{1,2}$ depending on $t = x_1 \in U$ only.

More invariantly, we consider the subspace of $\mathcal{A}$ which is invariant under the additive group $(\mathbb{R}^3, +)$ acting on $U \times \mathbb{R}^{1,2} \subset \mathbb{R}^{2,2}$ by translation in the last three variables. The subspace $\mathcal{A}_{Sch}$ is not preserved by the full gauge group $\mathcal{G}$ but by the subgroup $\mathcal{G}_{Sch}$ of gauge transformations invariant under the above action of $\mathbb{R}^3$ satisfying appropriate boundary conditions. This is the subgroup of gauge transformations only depending on $t \in U$, which are equal to the identity on ∂U, i.e.

$$\mathcal{G}_{Sch} = \{u \in \Gamma(U, \mathrm{U}(E)) \mid u|_{\partial U} = \mathrm{id}\}.$$

Hence,

$$\mathrm{Lie}(\mathcal{G}_{Sch}) = \{\xi \in \Gamma(U, \mathfrak{u}(E)) |\ \xi|_{\partial U} = 0\}.$$

The endomorphisms I, S, T are constant in all variables, so in particular invariant under the action of $\mathbb{R}^3$ and thus induce endomorphisms I, S, T on $T\mathcal{A}_{Sch}$. We endow $\mathcal{A}_{Sch}$ with an indefinite metric given by

$$g_{Sch}(X, Y) = -\int_U \sum_{i=1}^{4} g_{ii}\mathrm{tr}(X_i Y_i)\ \mathrm{d}t,$$

for tangent vectors $X = \sum_i X_i \mathrm{d}x_i, Y = \sum_i Y_i \mathrm{d}x_i \in \Omega^1(\mathbb{R}^{2,2}, \mathrm{u}(E))$ depending only on t. This gives $\mathcal{A}_{Sch}$ the structure of a flat infinite-dimensional hypersymplectic manifold. Note that this is automatic only for a bounded interval U. In the case that U is (semi-)infinite, we have to impose correct decay conditions to make the metric finite. However, we will now proceed formally and from the next section onwards we will always work on bounded intervals. The gauge group $\mathcal{G}_{Sch}$ acts on $\mathcal{A}_{Sch}$ in the usual way, i.e.

$$u.(T_0, T_1, T_2, T_3) = (u^{-1}T_0 u + u^{-1}\dot{u}, u^{-1}T_1 u, u^{-1}T_2 u, u^{-1}T_3 u).$$

The vanishing of the associated moment map μ_{Sch} is then the same as Schmid's equations. It is at this point where the condition that gauge transformations should be the identity on ∂U comes in. If we do not impose these boundary conditions, we would pick up boundary terms coming from integration by parts when we compute the moment maps. In earlier discussions this issue did not play a role, since the manifolds considered did not have any boundary.

In this way we get the moduli space of solutions to Schmid's equations:

$$\mathcal{M}_{Sch} = \mu_{Sch}^{-1}(0)/\mathcal{G}_{Sch}.$$

The fundamental vector fields associated to the action of $\mathcal{G}_{Sch}$ are given by

$$X_\nabla^\xi = \mathrm{d}^\nabla \xi \qquad \xi \in \mathrm{Lie}(\mathcal{G}_{Sch}), \nabla \in \mathcal{A}_{Sch}.$$

Explicitly, if $\nabla = \mathrm{d} + \sum_{i=1}^4 T_{i-1}\mathrm{d}x_i$, then

$$X_\nabla^\xi = (\dot{\xi} + [T_0, \xi])\mathrm{d}x_1 + \sum_{i=2}^{4} [T_{i-1}, \xi]\mathrm{d}x_i.$$

We naturally identify the tangent space of $\mathcal{M}_{Sch}$ at a solution ∇ with the kernel of the derivative of the moment map intersected with the g_{Sch}-orthogonal complement of the

$\mathcal{G}_{Sch}$-orbit through ∇. This means a tangent vector $X = \sum_{i=1}^4 X_{i-1}\mathrm{d}x_i \in T_\nabla \mathcal{M}_{Sch}$ will be defined by the equations

$$\begin{aligned}
\dot{X}_1 + [T_0, X_1] + [X_0, T_1] + [T_2, X_3] + [X_2, T_3] &= 0 \\
\dot{X}_2 + [T_0, X_2] + [X_0, T_2] - [T_3, X_1] - [X_3, T_1] &= 0 \\
\dot{X}_3 + [T_0, X_3] + [X_0, T_3] - [T_1, X_2] - [X_1, T_2] &= 0 \\
\\
\dot{X}_0 + [T_0, X_0] + [T_1, X_1] - [T_2, X_2] - [T_3, X_3] &= 0.
\end{aligned}$$

The last equation says that X is orthogonal to the $\mathcal{G}_{Sch}$ orbit as can be seen by an integration by parts argument using the metric g_{Sch}. A tangent vector lies in the degeneracy space if it satisfies these equations and is of the form X^ξ_∇. So ξ has to satisfy

$$\frac{\mathrm{d}^2\xi}{\mathrm{d}t^2} + [\dot{T}_0, \xi] + 2[T_0, \dot{\xi}] + [T_0, [T_0, \xi]] + [T_1, [T_1, \xi]] - [T_2, [T_2, \xi]] - [T_3, [T_3, \xi]] = 0.$$

This is a recent result found in 2010 by Matsoukas in his DPhil thesis [48]. It fits well with our general criterion, which here would say (see 2.4.2)

$$\sum_{i=1}^4 g_{ii} \left(\frac{\partial^2 \xi}{\partial x_i^2} + 2\left[A_i, \frac{\partial \xi}{\partial x_i}\right] + \left[\frac{\partial A_i}{\partial x_i}, \xi\right] + [A_i, [A_i, \xi]] \right) = 0.$$

Using our notation with the T_i's and the fact that T_i and ξ only depend on t, this simplifies to

$$\frac{\mathrm{d}^2\xi}{\mathrm{d}t^2} + 2[T_0, \dot{\xi}] + [\dot{T}_0, \xi] + [T_0, [T_0, \xi]] + \sum_{i=2}^4 g_{ii}[T_{i-1}, [T_{i-1}, \xi]] = 0.$$

So we can say that ξ defines a solution to the ultrahyperbolic wave equation $(\mathrm{d}^\nabla)^*\mathrm{d}^\nabla \xi = 0$ on $\mathbb{R}^{2,2}$, which depends on $t = x_1$ only, satisfying the boundary condition $\xi \equiv 0$ on ∂U. This gives a natural interpretation for Matsoukas' result.

The points where the hypersymplectic structure on the moduli space is degenerate are those for which there is a solution to the boundary value problem

$$\frac{\mathrm{d}^2\xi}{\mathrm{d}t^2} + 2[T_0, \dot{\xi}] + [\dot{T}_0, \xi] + \sum_{i=1}^4 g_{ii}[T_{i-1}, [T_{i-1}, \xi]] = 0 \qquad \xi \equiv 0 \text{ on } \partial U$$

for a given solution $\mathcal{T} = (T_0, T_1, T_2, T_3)$.

To what extent do the moduli space $\mathcal{M}_{Sch}$ and the degeneracy space depend on the choice of the (possibly infinite) interval $U \subset \mathbb{R}$ on which the equations are considered?

The next two propositions tell us that all that the behaviour of the moduli space only depends on the topological type of the interval, i.e. whether the interval is finite, semi-infinite or infinite.

Proposition 3.1.1. *Let $U = [a,b]$ be a bounded closed interval and consider the affine transformation*

$$\lambda : [0,1] \to [a,b] \qquad \lambda(t) = (b-a)t + a.$$

Let $\mathcal{T} = (T_0, T_1, T_2, T_3)$ be a solution to Schmid's equations over $[a,b]$. Then we get a solution $\mathcal{P} = (P_0, P_1, P_2, P_3)$ over $[0,1]$ by

$$P_i = (b-a)T_i \circ \lambda.$$

Moreover, if ξ lies in the degeneracy space at $\mathcal{T}$, then $\eta = (b-a)\xi \circ \lambda$ lies in the degeneracy space at $\mathcal{P}$. If $\mathcal{T}$ and $\mathcal{T}'$ are gauge equivalent solutions on $[a,b]$ related by a gauge transformation u, then the corresponding solutions $\mathcal{P}$ and $\mathcal{P}'$ are gauge equivalent via the gauge transformation $u \circ \lambda$. Thus, if the interval U is finite, $\mathcal{M}_{Sch}$ does not depend on U in a non-trivial way.

Proof. Both statements are proved by straight-forward calculations. For example,

$$\begin{aligned} \dot{P}_1(t) + [P_0(t), P_1(t)] &= (b-a)^2(\dot{T}_1(\lambda(t)) + [T_0(\lambda(t)), T_1(\lambda(t))]) \\ &= (b-a)^2(-[T_2(\lambda(t)), T_3(\lambda(t))]) \\ &= -[P_2(t), P_3(t)], \end{aligned}$$

and analogously for the other two equations. The fact that gauge equivalent solutions remain gauge equivalent is a direct consequence of the definition of the action of the gauge group.

The proof that η satisfies the degeneracy equation works in exactly the same fashion. □

There is a similar statement if the interval is semi-infinite.

Proposition 3.1.2. *Let $a, b \in \mathbb{R}$. If $\mathcal{T}$ is a solution to Schmid's equations over $U = [a, \infty)$, we get a solution $\mathcal{P}$ over $[0, \infty)$ by setting $P_i(t) = T_i(t-b)$. Moreover, if ξ lies in the degeneracy space at $\mathcal{T}$, then $\eta(t) = \xi(t-b)$ lies in the degeneracy space at $\mathcal{P}$.*

If $\mathcal{T}$ is a solution to Schmid's equations over $U = (-\infty, b]$, we get a solution $\mathcal{P}$ over $[0, \infty)$ by setting $P_i(t) = -T_i(-t-b)$. Moreover, if ξ lies in the degeneracy space at $\mathcal{T}$, then $\eta(t) = -\xi(-t-b)$ lies in the degeneracy space at $\mathcal{P}$.

Proof. We only prove the second assertion, the first one is obvious.

$$\begin{aligned}\dot{P}_1(t) + [P_0(t), P_1(t)] &= (-1)^2\dot{T}_1(-t-b) + (-1)^2[T_0(-t-b), T_1(-t-b)]) \\ &= -[T_2(-t-b), T_3(-t-b)]) \\ &= -[P_2(t), P_3(t)],\end{aligned}$$

and analogously for the other two equations. The proof that η defines an element in the degeneracy space at P works analogously. □

Therefore, if we consider the equations over a closed bounded interval, we will from now on always work over $[0,1]$. Also, if the interval is semi-infinite, i.e. unbounded from above or below but not the whole real line we may assume to work over $[0,\infty)$. Thus, actually there are only three different situations we have to analyse: $U=[0,1]$, $U=[0,\infty)$ or $U=\mathbb{R}$. Later in this chapter we will consider the case $U=[0,1]$ more closely.

We now discuss two different viewpoints from which the Schmid equations can be investigated.

3.2 The Spectral Curve

Firstly, we exploit the Lax pair perspective which we developed in section 2.4.2. This will lead to a way of solving the equations (in principle) explicitly by linearising them on the space of line bundles over the so called *spectral curve*. This section follows closely the analogous discussion of Nahm's equations in [27] and [26].

Let U be one of the possible choices, $U=[0,1],[0,\infty),\mathbb{R}$. Looking at our modified Lax pair from section 2.4.2, we obtain for Schmid's equations

$$[\frac{\mathrm{d}}{\mathrm{d}t} + T_0 - iT_1 + i\zeta(T_2 - iT_3), T_2 + iT_3 - 2\zeta T_1 + \zeta^2(T_2 - iT_3)] = 0.$$

With the short-hand notation

$$T_+ = -iT_0 - T_1 + \zeta(T_2 - iT_3) \qquad T = T_2 + iT_3 - 2\zeta T_1 + \zeta^2(T_2 - iT_3)$$

we can write this as

$$[\frac{\mathrm{d}}{\mathrm{d}t} + iT_+, T] = 0.$$

Or

$$\dot{T} = [T, iT_+],$$

where as usual the dot denotes differentiation with respect to t. From this form of the equations it is easy to check that for all n, $\mathrm{tr}(T^n)$ is independent of t, which implies that the coefficients of the characteristic polynomial of T are independent of t. Thus, the set of eigenvalues of T,

$$S = \{(\zeta, \eta) \in \mathbb{C} \times \mathbb{C} \mid \det(\eta + T) = 0\},$$

is a conserved quantity of the system. Setting

$$T_- = \frac{1}{\zeta}(T_2 + iT_3) - T_1 + iT_0,$$

we get

$$\frac{1}{\zeta}T = T_+ + T_-,$$

and so

$$[\frac{\mathrm{d}}{\mathrm{d}t} - iT_-, \frac{1}{\zeta^2}T] = [\frac{\mathrm{d}}{\mathrm{d}t} + iT_+ - \frac{i}{\zeta}T, \frac{1}{\zeta^2}T] = \frac{1}{\zeta^2}[\frac{\mathrm{d}}{\mathrm{d}t} + iT_+, T] = 0.$$

We thus get a Lax equation in $\tilde{\zeta} = \frac{1}{\zeta}$ and the associated set of eigenvalues satisfies

$$\tilde{S} = \{\det(\tilde{\eta} + \frac{1}{\zeta^2}T) = 0\} = \{\frac{1}{\zeta^{2n}}\det(\eta + T) = 0\},$$

if $\tilde{\eta} = \frac{\eta}{\zeta^2}$. In this manner we associate to Schmid's equations an algebraic curve S called the *spectral curve*, which lies in the total space of the line bundle $\mathcal{O}(2)$, i.e. the complex line bundle over $\mathbb{CP}^1$ with transition function $\frac{1}{\zeta^2}$.

Before we go on, we examine the object T. By the above discussion, η is a local coordinate on the line bundle $\mathcal{O}(2) \to \mathbb{CP}^1$, so for fixed t the object $\eta + T$ defines a holomorphic section of the vector bundle

$$\mathrm{Hom}(\mathrm{pr}_1^* E, \mathrm{pr}_1^* E \otimes \mathrm{pr}_{\mathbb{CP}^1}^* \mathcal{O}(2)) \to \{t\} \times \mathcal{O}(2).$$

Here $\mathrm{pr}_1 : \mathbb{R}^{2,2} \times \mathcal{O}(2) \to \mathbb{R}^{2,2}$ is the projection onto the first factor and $\mathrm{pr}_{\mathbb{CP}^1} : \mathbb{R}^{2,2} \times \mathcal{O}(2) \to \mathcal{O}(2) \to \mathbb{CP}^1$ is the projection onto the second factor followed by the bundle projection $\mathcal{O}(2) \to \mathbb{CP}^1$.

If we now assume that $T(t)$ only has one-dimensional eigenspaces, i.e. that the curve S is smooth, we get a holomorphic line bundle $L(t)$ on S given by $L(t)_{(\zeta,\eta)} = \ker(\eta + T(t))$. $L = L(t)$ will be given by a transition function $a(t, \zeta, \eta)$ depending smoothly on t. Moreover, since the operators $\frac{\mathrm{d}}{\mathrm{d}t} + iT_+$ and $\frac{\mathrm{d}}{\mathrm{d}t} - iT_-$ commute with T, respectively $\frac{1}{\zeta^2}T$, they preserve $L(t)$. So we try to trivialise L by sections in the respective kernels of these operators. Let ψ and ϕ be local trivialisations of L over

$\zeta \neq \infty$ and $\tilde{\zeta} \neq \infty$, which in addition lie in the kernel of $\frac{\mathrm{d}}{\mathrm{d}t} + iT_+$ and $\frac{\mathrm{d}}{\mathrm{d}t} - iT_-$, respectively. That is, they satisfy the following set of equations

$$\begin{aligned}\psi &= a\phi, \\ (\eta + T)\psi &= 0 = (\eta + T)\phi, \\ \dot{\psi} &= -iT_+\psi, \\ \dot{\phi} &= iT_-\phi.\end{aligned}$$

We now want to use these equations to determine the t-dependence of a. That is, we wish to determine the evolution of the line bundle $L(t)$ viewed as a curve in the *Jacobian*, i.e. the space of line bundles, over S. We compute

$$-iT_+\psi = \dot{\psi} = \dot{a}\phi + a\dot{\phi} = \dot{a}\phi + aiT_-\phi.$$

With $\phi = \frac{\psi}{a}$ this gives

$$0 = \frac{\dot{a}}{a}\psi + iT_-\psi + iT_+\psi.$$

But $T_+ + T_- = \frac{1}{\zeta}T$ and $T\psi = -\eta\psi$, so we arrive at

$$\frac{\dot{a}}{a}\psi - i\frac{\eta}{\zeta}\psi = 0.$$

That is,

$$(\dot{a} - i\frac{\eta}{\zeta}a)\psi = 0.$$

And since ψ is non-vanishing, we obtain the following ODE for a:

$$\dot{a} = i\frac{\eta}{\zeta}a,$$

with general solution given by

$$a(t, \zeta, \eta) = b(\zeta, \eta)\exp(\frac{i\eta}{\zeta}t).$$

This means, that on the Jacobian $\mathrm{Jac}(S) = H^1(S, \mathcal{O})/H^1(S, \mathbb{Z}) \cong \mathbb{C}^g/\mathbb{Z}^{2g}$, the equations are linearised, since this description of the Jacobian comes from the exponential sequence, i.e. the long exact sequence associated to the short exact sequence of sheaves on S

$$0 \to \mathbb{Z} \to \mathcal{O} \to \mathcal{O}^* \to 0.$$

The map $\mathcal{O} \to \mathcal{O}^*$ is given by $\exp(2\pi i -)$. In the long exact sequence, we get

$$\cdots \to 0 \to H^1(S, \mathbb{Z}) \to H^1(S, \mathcal{O}) \to H^1(S, \mathcal{O}^*) \to H^2(S, \mathbb{Z}) \cong \mathbb{Z} \to 0 = H^2(S, \mathcal{O}).$$

The Cech cohomology group $H^1(S, \mathcal{O}^*)$ parametrises isomorphism classes of holomorphic line bundles. The map $H^1(S, \mathcal{O}) \to H^1(S, \mathcal{O}^*)$ is given again by exp. So $\frac{i\eta}{\zeta}t$ describes a straight line on the torus $\mathrm{Jac}(S)$.

Note in particular the factor i appearing in the exponent of the time-dependent transition functions. In the analogous computation for Nahm's equations, this is not present. So we observe that, at least if the spectral curve is elliptic, the directions of the Schmid and Nahm flows differ by an angle of $\frac{\pi}{2}$.

A natural question to ask is which curves in the total space of $\mathcal{O}(2)$ arise as spectral curves associated to a solution to Schmid's equations, i.e. if we can characterise them algebraically. The next proposition shows that the curves have to satisfy a reality condition.

Recall that we obtained our Lax pair formulation of Schmid's equations by twistor methods and then considered the equations for $\zeta \in \mathbb{CP}^1$, i.e. we compactified the two-sheeted hyperboloid, on which ζ lives, and ignored the complex structure. We noted that inversion in the unit circle gives an antiholomorphic involution, or real structure, τ and it turns out that the spectral curve is *real* with respect to this involution. Since an analogous statement holds for Nahm's equations, where the involution is given by the antipodal map and corresponds to reversing the orientation of straight lines in $\mathbb{R}^3$ where the associated monopoles live, this is not entirely surprising.

Proposition 3.2.1. *Write*

$$\det(\eta + T) = \eta^n + \sum_{i=1}^{n} a_i(\zeta)\eta^{n-i},$$

with a_i a polynomial in ζ of degree $\leq 2i$. Then

$$a_i(\zeta) = (-1)^i \zeta^{2i} \overline{a_i(\frac{1}{\bar{\zeta}})}.$$

We remark that a_i is indeed of degree $\leq 2i$, since the matrix entries of T are quadratic polynomials in ζ.

Proof. This is a direct computation, using the fact that

$$\det((\eta + T)^*) = \overline{\det(\eta + T)}.$$

Now,

$$\begin{aligned} T^*(\zeta) &= (T_2 + iT_3 - 2\zeta T_1 + \zeta^2(T_2 - iT_3))^* \\ &= -T_2 + iT_3 + 2\bar{\zeta}T_1 - \bar{\zeta}^2(T_2 + iT_3) \\ &= -\bar{\zeta}^2(T_2 + iT_3 - 2\frac{1}{\bar{\zeta}}T_1 + \frac{1}{\bar{\zeta}^2}(T_2 - iT_3)) \\ &= -\bar{\zeta}^2 T(\frac{1}{\bar{\zeta}}). \end{aligned}$$

Thus,

$$\det((\eta + T(\zeta)^*) = \det\left(\bar{\zeta}^2(\frac{\bar{\eta}}{\bar{\zeta}^2} - T(\frac{1}{\bar{\zeta}}))\right) = (-1)^n\bar{\zeta}^{2n}\det\left(\frac{-\bar{\eta}}{\bar{\zeta}^2} + T(\frac{1}{\bar{\zeta}})\right).$$

Therefore, the relation

$$\det((\eta + T)^*) = \det(\bar{\eta} + T^*) = \overline{\det(\eta + T)}$$

implies

$$\overline{\det(\eta + T(\zeta))} = (-1)^n\bar{\zeta}^{2n}\det\left(\frac{-\bar{\eta}}{\bar{\zeta}^2} + T(\frac{1}{\bar{\zeta}})\right),$$

which, if we expand both sides as polynomials in η, gives

$$\begin{aligned} \bar{\eta}^n + \sum_{i=1}^{n-1}\overline{a_i(\zeta)}\bar{\eta}^{n-i} &= (-1)^n\bar{\zeta}^{2n}\left(\frac{(-1)^n\bar{\eta}^n}{\bar{\zeta}^{2n}} + \sum_{i=1}^{n-1}a_i(\frac{1}{\bar{\zeta}})\frac{(-1)^{n-i}\bar{\eta}^{n-i}}{\bar{\zeta}^{2(n-i)}}\right) \\ &= \bar{\eta}^n + \sum_{i=1}^{n-1}a_i(\frac{1}{\bar{\zeta}})\frac{(-1)^i\bar{\eta}^{n-i}}{\bar{\zeta}^{-2i}} \\ &= \bar{\eta}^n + \sum_{i=1}^{n-1}\left((-1)^i\bar{\zeta}^{2i}a_i(\frac{1}{\bar{\zeta}})\right)\bar{\eta}^{n-i}. \end{aligned}$$

That is,

$$a_i(\zeta) = (-1)^i\zeta^{2i}\overline{a_i(\frac{1}{\bar{\zeta}})}.$$

□

Corollary 3.2.2. *Suppose* $a_i(\zeta) = \sum_{k=0}^{2i}\alpha_k\zeta^k$*, then the coefficients* $\alpha_k \in \mathbb{C}$ *satisfy the relation*

$$\alpha_{2k-i} = (-1)^i\bar{\alpha}_i.$$

Thus, we have at most $n^2 + 2n$ *real degrees of freedom in defining the spectral curve.*

Proof. This follows directly upon expanding the relation

$$a_i(\zeta) = (-1)^i \zeta^{2i} \overline{a_i(\frac{1}{\bar{\zeta}})}.$$

This yields

$$\sum_{k=0}^{2i} \alpha_k \zeta^k = \sum_{k=0}^{2i} (-1)^i \bar{\alpha}_k \zeta^{2i-k},$$

and the corollary follows by equating coefficients. In particular we get that $\alpha_i = (-1)^i \bar{\alpha}_i$, so α_i is either real or purely imaginary leaving us with $2i$ real degrees of freedom to define the first i coefficients $\{\alpha_k\}_{k=0}^{i-1}$. Now the spectral curve is given by the equation

$$\eta^n + \sum_{i=1}^{n} a_i(\zeta)\eta^{n-i} = 0.$$

So we have $\sum_{i=1}^{n}(2i+1) = n^2 + 2n$ real degrees of freedom in the definition of the polynomial. Of course, there might be further constraints coming from the induced flow on the Jacobian, so all we can say is that we have *at most* $n^2 + 2n$ degrees of freedom. □

Proposition 3.2.3. *The genus of S is equal to*

$$g(S) = (n-1)^2.$$

Proof. Recall that $g(S) = \dim H^1(S, \mathcal{O})$. Earlier, we gave a description of S as a subvariety of $\mathcal{O}(2)$ with charts induced by the standard charts of $\mathcal{O}(2)$. It is therefore easy to describe $H^1(S, \mathcal{O})$ in terms of Cech cohomology. Elements of $H^1(S, \mathcal{O})$ are represented by holomorphic functions $f(\zeta, \eta)$ on $\mathbb{C}^* \times \mathbb{C}$ modulo holomorphic functions that can be extended to $\mathbb{C}$ or $(\mathbb{C}\backslash\{0\}) \cup \{\infty\}$ and modulo the equation $\det(\eta + T(\zeta)) = \eta^n + \sum_{i=1}^{n} a_i(\zeta)\eta^{n-i} = 0$. So we are considering functions of the form

$$f(\zeta, \eta) = \sum_{k\in\mathbb{Z}} \sum_{l=0}^{\infty} c_{kl} \zeta^k \eta^l,$$

modulo

$$\sum_{k=0}^{\infty}\sum_{l=0}^{\infty} c_{kl}\zeta^k \eta^l, \qquad \sum_{k=1}^{\infty}\sum_{l=0}^{\infty} c_{kl}\zeta^{-k-2l}\eta^l \qquad \text{and} \qquad \eta^n + \sum_{i=1}^{n} a_i(\zeta)\eta^{n-i} = 0.$$

Such functions can be represented in the form

$$\begin{aligned} f(\zeta,\eta) &= \sum_{l=0}^{n-1}\sum_{k=1}^{2l-1} c_{kl}\zeta^{-k-2l}\eta^l \\ &= \frac{1}{\zeta^{2n-2}}\sum_{l=0}^{n-1}\sum_{k=1}^{2l-1} c_{kl}\zeta^{-k}\zeta^{2(n-1-l)}\eta^l \\ &= \frac{1}{\zeta^{2n-2}}\sum_{l=0}^{n-1} P_l(\zeta^{-1})\zeta^{2(n-1-l)}\eta^l, \end{aligned}$$

where P_l is a polynomial of degree at most $2l-1$, with zero constant term. Thus, the dimension of $H^1(S,\mathcal{O})$ is given by

$$\dim H^1(S,\mathcal{O}) = \sum_{l=0}^{n-1}(2l-1) = -n+1+2\sum_{l=1}^{n-1} l = -n+1+n(n-1) = n^2-2n+1 = (n-1)^2,$$

as required. □

We have seen above that τ acts on the spectral curve S. We observe that in addition the line bundle L is real with respect to τ.

Proposition 3.2.4. *The line bundle $L(t)$ with transition function $a(t,\zeta,\eta) = \exp(\frac{i\eta}{\zeta}t)$ satisfies*

$$\tau^*L(t) \cong \bar{L}(t)^*.$$

Proof. The transition function of the line bundle τ^*L is given by $a(t,\tau(\zeta,\eta)) = \exp(\frac{-i\bar{\eta}}{\bar{\zeta}}t) = \bar{a}(t,\eta,\zeta)$. The claim then follows by observing that the real structure τ interchanges the two trivialising neighbourhoods. □

Recall that in general a solution to Schmid's equations defines a flow on the Jacobian of the spectral curve given by transition functions $a(t,\zeta,\eta) = b(\zeta,\eta)\exp(\frac{i\eta}{\zeta}t)$, i.e. the flow is of the form $B\otimes L(t)$, where B is the line bundle determined by the initial conditions. Thus, in the $\mathfrak{su}(2)$-case, when the spectral curve is expected to be elliptic, we obtain that the flow lines $B\otimes L(t)$ are along the real direction in the complex one-dimensional Jacobian. The set of real line bundles forms a collection of disjoint circles inside the Jacobian. Thus this flow is constrained to the real circles translated by B and so therefore we expect solutions be periodic, although in principle it could happen that a solution takes infinitely long time to go round the circle. This checks with a result by Matsoukas [48], who solved the equations in this case in terms of elliptic functions. We will discuss his result later, when we produce explicit solutions with cyclic symmetry. We note this purely algebraic explanation of the periodicity of the solutions in the $\mathfrak{su}(2)$-case in the following corollary.

Corollary 3.2.5. *The solutions to Schmid's equations with values in $\mathfrak{su}(2)$ are generically periodic.*

For higher genus spectral curves, the set of real line bundles forms a real torus of real dimension $g > 1$ inside the Jacobian and there is no reason why the flow lines should close up as they could have irrational slope.

3.2.1 The Underlying Real Curve

We have seen earlier that the spectral curve S is acted upon by the antiholomorphic involution $\tau : (\zeta, \eta) \mapsto (\bar{\zeta}^{-1}, -\bar{\eta}/\bar{\zeta}^2) = (\bar{\bar{\zeta}}, -\bar{\bar{\eta}})$. Note that this interchanges the two standard open sets which cover S. In the Nahm case, the involution is induced from the antipodal map and hence acts freely on S. In our situation the involution fixes the unit circle in $\mathbb{CP}^1$, and thus will have fixed points on S. Therefore, in contrast to the Nahm case, we have embedded in S a real curve

$$S_{\mathbb{R}} = \{(\zeta, \eta) \in S \mid \tau(\zeta, \eta) = (\zeta, \eta)\} \subset S.$$

First, we will describe the fixed point set of τ on $\mathcal{O}(2) \cong T\mathbb{CP}^1$. We necessarily need ζ to lie on the unit circle, i.e $\zeta = e^{\imath\theta}$. Then η should satisfy $\eta = -\bar{\bar{\eta}} = -\bar{\eta}e^{2i\theta}$. So $\eta = iae^{i\theta}$ for some $a \in \mathbb{R}$. Thus, the set of real points may be identified with the trivial real line bundle over S^1. In fact, if we change to co-ordinates (λ, μ) via

$$\zeta = \frac{\lambda + i}{\lambda - i},$$

then the involution τ becomes simply

$$(\lambda, \mu) \mapsto (\bar{\lambda}, \bar{\mu}),$$

which exhibits the fixed-point set as the tangent bundle of $\mathbb{RP}^1 \cong S^1$ sitting inside the tangent bundle of $\mathbb{CP}^1$.

We would like to deduce some properties of the real curve $S_{\mathbb{R}}$. A classical result is Harnack's bound on the number of connected components, a proof of which can be found for example in [22].

Proposition 3.2.6. *Let X be a compact Riemann surface of genus g with an antiholomorphic involution $\sigma : X \to X$ and denote by $X_{\mathbb{R}}$ the real algebraic curve given by the fixed points of σ. Then the number r of connected components of $X_{\mathbb{R}}$ satisfies*

$$r \leq g + 1.$$

The following example shows that in the $\mathfrak{su}(2)$-case equality is achieved.

Example (The $\mathfrak{su}(2)$-curve)**.** In the $\mathfrak{su}(2)$-case it follows from work by Matsoukas [48] that up to an SO(1, 2)-action, which we will describe later, the spectral curve may be chosen to be of the form

$$P(\zeta,\eta) = \eta^2 + A\zeta^2 + B(1+\zeta^4) = 0.$$

This will be discussed in more detail in the section on explicit solutions. From the above discussion we see that real points are points of the form $(\zeta,\eta) = (e^{i\theta}, iae^{i\theta})$, with $a \in \mathbb{R}$, satisfying $P(\zeta,\eta) = 0$. Rewriting this as an equation for (θ, a) this becomes

$$-a^2 + A + 2B\cos(2\theta) = 0.$$

Thus,

$$a = \pm\sqrt{A + 2B\cos(2\theta)}.$$

Now it turns out that A and B have the following form:

$$A = \alpha + \beta \qquad B = \frac{1}{2}(\alpha - \beta),$$

where $\alpha, \beta \in \mathbb{R}$ are non-negative. They both vanish simultaneously if and only if we are dealing with the trivial solution $T_i = 0$. Therefore,

$$A + 2B\cos(2\theta) = \alpha(1+\cos(2\theta)) + \beta(1-\cos(2\theta))$$

is positive, unless we are considering the trivial solution. In particular, the real curve is parametrised by

$$(\zeta,\eta) = (e^{i\theta}, e^{i(\theta\pm\pi/2)}\sqrt{A+2B\cos(2\theta)}),$$

which is a trivial double covering of the unit circle $\{|\zeta| = 1\}$.

3.3 Conserved Quantities

We have seen that we have many conserved quantities associated to a solution of Schmid's equations. We obtained a family $Q_n = \operatorname{tr}(T^n)$. Let us look at the case $n = 2$. We compute

$$\begin{aligned}
\operatorname{tr}(T^2) &= \operatorname{tr}((T_2+iT_3)^2 - 4\zeta T_1(T_2+iT_3) + \zeta^2(4T_1^2 + 2(T_2-iT_3)(T_2+iT_3)) \\
&\quad -4\zeta^3 T_1(T_2-iT_3) + \zeta^4(T_2-iT_3)^2) \\
&= \operatorname{tr}(T_2^2 - T_3^2 + 2iT_2T_3 + \zeta(-4T_1T_2 - 4iT_1T_3) + \zeta^2(4T_1^2 + 2T_2^2 + 2T_3^2) \\
&\quad +\zeta^3(-4T_1T_2 + 4iT_1T_3) + \zeta^4(T_2^2 - T_3^2 - 2iT_2T_3)).
\end{aligned}$$

Viewing this expression as a polynomial in ζ, its time-derivative can only be zero if all the coefficients are already constant in t. So we deduce that the following gauge-invariant quantities are conserved:

$$\mathrm{tr}(T_2^2 - T_3^2), \qquad \mathrm{tr}(T_2T_3), \qquad -\mathrm{tr}(T_1T_2), \qquad \mathrm{tr}(T_1T_3), \qquad \mathrm{tr}(2T_1^2 + T_2^2 + T_3^2).$$

The last expression $-\mathrm{tr}(2T_1^2 + T_2^2 + T_3^2)$ is particularly interesting. It is the sum of non-negative quantities and since it is constant in time, we deduce that each of the terms has to be bounded. Also $-\mathrm{tr}(AB)$ is the natural inner product on $\mathfrak{u}(n)$. Thus, we obtain Matsoukas' result that every solution to Schmid's equations exists automatically for all times [48]. This follows from the fact that if a solution to an ODE only exists for finite time, it has to leave every compact set.

This observation allows us to identify the moduli space of solutions on $U = [0,1] \subset \mathbb{R}$. We will see in particular that it is finite-dimensional. Let $\mathcal{T} = (T_0, T_1, T_2, T_3)$ be a solution to Schmid's equations on U. Firstly, we observe that we can choose a smooth function $u \in C^\infty(U, \mathrm{U}(n))$ such that $u.T_0 = u^{-1}T_0u + u^{-1}\dot{u} = 0$. To achieve this, we just have to solve the linear matrix ODE

$$\dot{u} = -T_0u.$$

If we fix an initial value, we get a unique solution. Let u_0 be the unique solution of the above ODE with initial value $u(0) = \mathrm{id}$. Note that in general u_0 does not lie in $\mathcal{G}_{Sch}$ as this would mean that $u_0 = \mathrm{id}$ on ∂U. In our situation, where $U = [0,1]$, we also need $u_0(1) = \mathrm{id}$, which may be impossible. Then $\mathcal{T}$ defines an element in $\mathrm{U}(n) \times \mathfrak{u}(n) \times \mathfrak{u}(n) \times \mathfrak{u}(n)$ via

$$\mathcal{T} \mapsto (u_0(1), T_1(1), T_2(1), T_3(1)).$$

Note that this map is indeed $\mathcal{G}_{Sch}$-invariant and so descends to a map on the moduli space. If we fix the values of the T_i's at $t = 1$, then according to the existence theorem for ODEs we get a unique locally defined solution to Schmid's equations with these values at $t = 1$. By the discussion of conserved quantities above, such a solution automatically extends to the whole real line. Hence the above map is surjective. It is injective by the uniqueness-part of the existence theorem of solutions to ODEs, and hence defines an isomorphism, at least on the level of sets.

If we consider the equations on a (semi-)infinite interval, we have to make sure that a solution $\mathcal{T}$ satisfies the right decay conditions. Thus, all we can hope for, is that in this case the moduli space will be in bijection with some open set in $\mathrm{U}(n) \times \mathfrak{u}(n) \times \mathfrak{u}(n) \times \mathfrak{u}(n)$.

The results of this paragraph remain true, if we consider the equations with the T_i taking values in an arbitrary compact Lie algebra $\mathfrak{g}$. In this case we have a bi-invariant Euclidean inner product on $\mathfrak{g}$, which plays the role of the inner product given by the trace tr.

Recall that, compared to the Nahm equations, the direction of the flow associated on the Jacobian of the spectral curve is multiplied by i in our situation. It is known that solutions of Nahm's equations will always have poles, which correspond to points where the flow hits the Θ-divisor in $\mathrm{Jac}(S)$. So we suspect that the observed regularity of solutions to Schmid's equations could be related to the fact that the direction of the flow on the Jacobian of the spectral curve is multiplied by i.

3.4 Group actions on the moduli space $\mathcal{M}_{Sch}$

If we work on the interval $U = [0, 1]$ and do not impose any boundary conditions, we see that the equations are invariant not only under the group $\mathcal{G}_{Sch}$ but under the whole group $\mathcal{G} = \Gamma(U, \mathrm{U}(E))$. We have subgroups $\mathcal{G}_L$ and $\mathcal{G}_R$ consisting of gauge transformations that are the identity at 0 and 1, respectively. Note that $\mathcal{G}_{Sch} = \mathcal{G}_L \cap \mathcal{G}_R$.

Lemma 3.4.1. *$\mathcal{G}_{Sch}$ is contained in $\mathcal{G}$ as a normal subgroup. The quotient $\mathcal{G}/\mathcal{G}_{Sch}$ can be canonically identified with $\mathrm{U}(n) \times \mathrm{U}(n)$.*

Proof. We only have to check that if $v \in \mathcal{G}_U$, then $v^{-1}uv \in \mathcal{G}_{Sch}$ for all $u \in \mathcal{G}_{Sch}$. But this is fairly obvious. Let $x \in \{0, 1\}$, then

$$v^{-1}uv(x) = v^{-1}(x)u(x)v(x) = v^{-1}(x)\mathrm{id}v(x) = \mathrm{id}.$$

For the identification of the quotient, consider the evaluation map

$$\mathrm{ev}_0 \times \mathrm{ev}_1 : \mathcal{G}_U \to \mathrm{U}(n) \times \mathrm{U}(n), \qquad \mathrm{ev}_0 \times \mathrm{ev}_1(v) = (v(0), v(1)).$$

This is a surjective group homomorphism with kernel $\mathcal{G}_{Sch}$. □

In fact, we can write $U(n) \times U(n) = \mathcal{G}/\mathcal{G}_L \times \mathcal{G}/\mathcal{G}_R =: G_L \times G_R$.

Lemma 3.4.2. *The actions of $\mathcal{G}_{Sch}$ and $\mathcal{G}_U$ on $\mu_{Sch}^{-1}(0)$ commute modulo $\mathcal{G}_{Sch}$. Therefore the action of $\mathcal{G}_U$ descends to an action of $\mathcal{G}_U/\mathcal{G}_{Sch} = \mathrm{U}(n) \times \mathrm{U}(n)$ on the moduli space $\mathcal{M}_{Sch}$.*

Proof. Viewing a solution $\mathcal{T}$ as a connection ∇ the action of $\mathcal{G}$ is given simply by conjugation. Then the assertion follows from the almost trivial observation that for $u \in \mathcal{G}_{Sch}$ and $v \in \mathcal{G}$ we have

$$uv = vv^{-1}uv = vu(u^{-1}v^{-1}uv),$$

and $u^{-1}v^{-1}uv$ lies in $\mathcal{G}_{Sch}$ since u does and $\mathcal{G}_{Sch}$ is a normal subgroup of $\mathcal{G}$. □

$\mathcal{G}_U$ acts on $\mathcal{A}_{Sch}$ preserving the hypersymplectic structure. Note that this is not true if U is (semi-)infinite, as then the T_i's have to satisfy the right decay conditions.

Now consider the object $T = T_2 + iT_3 - 2\zeta T_1 + \zeta^2(T_2 - iT_3)$ from the last paragraph. We see that any gauge transformation $v \in \mathcal{G}_U$ acts on T simply by conjugation. In particular, two points $\mathcal{T}, \mathcal{T}' \in \mathcal{M}_{Sch}$ that lie in the same $\mathrm{U}(n) \times \mathrm{U}(n) = \mathcal{G}_U/\mathcal{G}_{Sch}$ orbit will have the same spectral curve. Using that T_+ and T_- transform under gauge transformations according to

$$u - T_\pm = u^{-1}T_\pm u \mp iu^{-1}\dot{u},$$

it is easy to check that we obtain isomorphic line bundles $L(t)$. This is because the transition functions remain unchanged, see the computation in the last section.

Furthermore, we have an action of the group $\mathrm{SL}(2,\mathbb{R}) = \mathrm{Spin}(1,2)$ on the moduli space. Let $\tilde{A} \in \mathrm{SL}(2,\mathbb{R})$ descend to $A \in \mathrm{SO}(1,2) = \mathrm{SO}(\mathrm{Im}(\mathbb{B}))$. Then

$$\tilde{A}.(T_0, T_1, T_2, T_3) = (T_0, \sum_{j=1}^{3} A_{1j}T_j, \sum_{j=1}^{3} A_{2j}T_j, \sum_{j=1}^{3} A_{2j}T_j).$$

In other words: A acts as $1 \oplus (1 \otimes A)$ on $\mathcal{T} = (T_0, T_1, T_2, T_3) \in \mathfrak{g} \oplus \mathfrak{g} \otimes \mathrm{Im}(\mathbb{B})$. It is easy, but slightly messy, to check by direct computation that this actually preserves the equations. We use the defining relation

$$A^T \mathrm{diag}(1,-1,-1)A = \mathrm{diag}(1,-1,-1)$$

to obtain

$$A^{-1} = \mathrm{diag}(1,-1,-1)A^T\mathrm{diag}(1,-1,-1),$$

and then expresses A^{-1} in terms of the matrix $A^\sharp$ whose (ij)-th entry is given by $(-1)^{i+j} \det A^{(i,j)}$, where $A^{(i,j)}$ is the matrix obtained from A by removing the i-th row and j-th column.

Notice that if we take the quotient of $\mathcal{G}_L$ by gauge transformations that are the identity at 0 by the group $\mathcal{G}_{Sch}$, we get an action of $G_L \cong G$ on the moduli space. We now wish to construct a map from the moduli space of solutions to Schmid's equations modulo the action of G_L to the space of self-adjoint traceless endomorphisms of $\mathbb{R}^{1,2}$, which will be moreover equivariant with respect to the actions of $\mathrm{SO}(1,2)$ on $\mathcal{M}_{Sch}/G_L$ and the adjoint action on $\mathrm{End}_0^{Sym}(\mathbb{R}^{1,2})$.

In fact, this map can be written down in a fairly natural way. The construction becomes clearer if we formulate it for an arbitrary compact Lie algebra. Let $\mathcal{T} = (T_1, T_2, T_3)$ be a triple of elements in $\mathfrak{g}$ and fix a bi-invariant inner product $<,>$. This then defines a linear map to $\mathbb{R}^{1,2}$ by contraction with the bi-invariant inner product. That is for $\xi \in \mathfrak{g}$, we define

$$\mathcal{T}(\xi) = (< T_1, \xi >, < T_2, \xi >, < T_3, \xi >)^t.$$

Now $\mathrm{SO}(1,2)$ acts on $\mathrm{Hom}(\mathfrak{g}, \mathbb{R}^{1,2})$ simply by matrix multiplication on $\mathbb{R}^{1,2}$, i.e. for $A \in \mathrm{SO}(1,2)$ we have

$$(A.\mathcal{T})(\xi) = A(\mathcal{T}(\xi)).$$

Let g denote the Lorentz inner product on $\mathbb{R}^{1,2}$ which we identify with its matrix $\mathrm{diag}(1,-1,-1)$. To $\mathcal{T}$, we have the dual map $\mathcal{T}^* \in \mathrm{Hom}((\mathbb{R}^{1,2})^*, \mathfrak{g}^*)$, i.e. for $\alpha = (a_1, a_2, a_3) \in (\mathbb{R}^{1,2})^*$ we have $\mathcal{T}^*(\alpha)(\xi) = \sum_{i=1}^3 a_i < T_i, \xi >$. Let ϕ be the isomorphism $\mathfrak{g}^* \cong \mathfrak{g}$ induced by the invariant inner product. Then to $\mathcal{T}$ we associate a trace-free self-adjoint endomorphism $\Psi(\mathcal{T})$ of $\mathbb{R}^{1,2}$, i.e. we define a map

$$\Psi : \mathrm{Hom}(\mathfrak{g}, \mathbb{R}^{1,2}) \to \mathrm{End}_0^{Sym}(\mathbb{R}^{1,2}),$$

by taking $\Psi(\mathcal{T})$ to be the trace-free part of the composition

$$\tilde{\Psi}(\mathcal{T}) : \mathbb{R}^{1,2} \to^g (\mathbb{R}^{1,2})^* \to^{\mathcal{T}^*} \mathfrak{g}^* \to^{\phi} \mathfrak{g} \to^{\mathcal{T}} \mathbb{R}^{1,2}.$$

Explicitly, for $v = (v_1, v_2, v_3)^t \in \mathbb{R}^{1,2}$ we have

$$\tilde{\Psi}(\mathcal{T})(v) = \sum_{i=1}^{3} \sum_{j=1}^{3} < T_i, g_{jj} v_j T_j > e_i.$$

Lemma 3.4.3. *Ψ is $\mathrm{SO}(1,2)$-equivariant, i.e. for any $A \in \mathrm{SO}(1,2)$ we have*

$$\Psi(A.\mathcal{T}) = A\Psi(\mathcal{T})A^{-1}.$$

Proof. Note that $A \in \mathrm{SO}(1,2)$ implies that

$$A^t g = g A^{-1}.$$

Therefore, we now compute with $v = (v_1, v_2, v_3)^t \in \mathbb{R}^{1,2}$

$$\begin{aligned}
\tilde{\Psi}(A.\mathcal{T})(v) &= \sum_{i=1}^{3} \sum_{j,k,l=1}^{3} < A_{ik}T_k, g_{jj}v_jA_{jl}T_l > e_i \\
&= \sum_{i=1}^{3} \sum_{k,l=1}^{3} < A_{ik}T_k, (A^t gv)_l T_l > e_i \\
&= \sum_{i=1}^{3} \sum_{k,l=1}^{3} < A_{ik}T_k, (gA^{-1}v)_l T_l > e_i \\
&= A\tilde{\Psi}(\mathcal{T})(A^{-1}v).
\end{aligned}$$

Together with the conjugation invariance of the trace this implies the claim. □

For Schmid matrices we can carry out this construction pointwise by taking $\mathfrak{g} = \mathrm{Lie}(\mathcal{G})$. This gives an $\mathrm{SO}(1,2)$-equivariant map

$$\Psi : C^\infty(U, \mathrm{Hom}(\mathfrak{g}, \mathbb{R}^{1,2})) \to C^\infty(U, \mathrm{End}_0^{Sym}(\mathbb{R}^{1,2})).$$

We now want to show that if $\mathcal{T}$ is a solution to Schmid's equations, this map actually takes values in $\mathrm{End}_0^{Sym}(\mathbb{R}^{1,2})$, i.e. $\Psi(\mathcal{T})$ is constant in time.

Proposition 3.4.4 ([48]). *Let $\mathcal{T} = (T_1, T_2, T_3)$ be a solution to Schmid's equations, then $\frac{\mathrm{d}}{\mathrm{d}t}\Psi(\mathcal{T}) = 0$.*

Proof. The prove is very simple. Let $v = (v_1, v_2, v_3)^t \in \mathbb{R}^{1,2}$. Then

$$\tilde{\Psi}(\mathcal{T})(v) = (< T_1, \sum_{i=1}^{3} g_{ii}v_iT_i >, < T_2, \sum_{i=1}^{3} g_{ii}v_iT_i >, < T_3, \sum_{i=1}^{3} g_{ii}v_iT_i >)^t.$$

This means that the matrix of $\Psi(\mathcal{T})$ is given by

$$B = \begin{pmatrix} < T_1, T_1 > & - < T_1, T_2, > & - < T_1, T_3 > \\ < T_2, T_1 > & - < T_2, T_2, > & - < T_2, T_3 > \\ < T_3, T_1 > & - < T_3, T_2 > & - < T_3, T_3 > \end{pmatrix} - \frac{1}{3}(\sum_{i=1}^{3} g_{ii} < T_i, T_i >)\mathrm{id}.$$

That is,

$$B = \begin{pmatrix} \alpha & - < T_1, T_2, > & - < T_1, T_3 > \\ < T_2, T_1 > & \beta & - < T_2, T_3 > \\ < T_3, T_1 > & - < T_3, T_2 > & \gamma \end{pmatrix},$$

where

$$\begin{aligned}
\alpha &= \frac{1}{3}(2 < T_1, T_1 > + < T_2, T_2 > + < T_3, T_3 >), \\
\beta &= \frac{1}{3}(- < T_1, T_1 > -2 < T_2, T_2 > + < T_3, T_3 >), \\
\gamma &= \frac{1}{3}(- < T_1, T_1 > + < T_2, T_2 > -2 < T_3, T_3 >),
\end{aligned}$$

which we recognise as the conserved quantities we found in the last paragraph, or linear combinations thereof. □

We remark that the bi-invariance of the inner product $<,>$ on $\mathfrak{g}$ implies that Ψ is invariant under gauge transformations and thus descends to a $G_L \times G_R$-invariant map from the Schmid moduli space to $\mathrm{End}_0^{Sym}(\mathbb{R}^{1,2})$.

3.5 Complex Structures and Product Structures

We now consider an infinite-dimensional application of our previous discussion on Kirwan-type theorems. Very often infinite dimensional versions of Kirwan's theorem yield identifications between the moduli spaces of solutions to gauge theoretic equations and moduli spaces of certain algebraic objects. A famous example is the interpretation of the moduli space of Higgs bundles on a Riemann surface Σ as complex representations of the fundamental groups of Σ.

3.5.1 Complex Structures

In this section, we present an approach analogous to Donaldson's way of treating Nahm's equations in [21], establishing a link with geometric invariant theory. He considers the real and complex equation introduced in the last section and shows that every gauge orbit of a solution to the complex equation under the action of the complexified gauge group contains an essentially unique solution to the real equation, and hence a zero of the - in his case hyperkähler - moment map. We try to see what happens if we carry over his constructions to our situation.

When we developed the Lax pair viewpoint, we noted that the vanishing of the hypersymplectic moment map is the same as saying that the complex equation $\mu^{\mathbb{C}} = 0$ and the real equation $\mu_I = 0$ are satisfied simultaneously. In the case of the ASD equations we had

$$\begin{aligned} \mu_I(\nabla) &= \frac{i}{2}([Z, Z^*] - [W, W^*]), \\ (\mu_S + i\mu_T)(\nabla) &= [Z, W]. \end{aligned}$$

We had $Z = \nabla_1 - i\nabla_2$ and $W = \nabla_3 + i\nabla_4$, which, when we look for solutions that depend on $t = x_1$ only, become

$$Z = \frac{\mathrm{d}}{\mathrm{d}t} + T_0 - iT_1 \qquad W = T_2 + iT_3.$$

In analogy to Donaldson's notation in [21], we write $\alpha = T_0 - iT_1$, $\beta = T_2 + iT_3$, so that

$$Z = \frac{\mathrm{d}}{\mathrm{d}t} + \alpha \qquad W = \beta.$$

Then the complex equation becomes

$$\dot{\beta} + [\alpha, \beta] = 0.$$

Noting that $Z^* = -\frac{\mathrm{d}}{\mathrm{d}t} + \alpha^*$, we obtain for the real equation

$$\mu_I(\alpha, \beta) = \dot{\alpha} + \dot{\alpha}^* + [\alpha, \alpha^*] - [\beta, \beta^*] = 0.$$

We now consider the equations on the interval $[0, 1] \subset \mathbb{R}$ and note that they are invariant under the action of the *compexified gauge group*

$$\mathcal{G}^{\mathbb{C}}_{Sch} = \{u : U = [0, 1] \to \mathrm{GL}(n, \mathbb{C}) \mid u(0) = \mathrm{id} = u(1)\}.$$

Here u acts on $\mathcal{T} = (T_0, T_1, T_2, T_3)$ in the usual way:

$$u.\alpha = u^{-1}\alpha u + u^{-1}\dot{u}, \qquad u.\beta = u^{-1}\beta u.$$

Again, since we do not specify any boundary conditions the T_i's should satisfy, we have an action of the full group $\mathcal{G}^{\mathbb{C}}_U$ of gauge transformations with arbitrary values at 0 and 1. We can easily write down the general solution of the complex equation. Let $u \in \mathcal{G}^{\mathbb{C}}_U$ be a complex gauge transformation such that $u.\alpha = 0$. That is, u solves the matrix ODE

$$\dot{u} + \alpha u = 0.$$

In this gauge we see from the complex equation that $u.\beta = u^{-1}\beta u$ has to be constant. We thus obtain a map

$$\begin{aligned} \Phi : \mathcal{M}_{Sch} &\to \mathrm{GL}(n, \mathbb{C}) \times \mathfrak{gl}(n, \mathbb{C}), \\ \mathcal{T} &\mapsto (u_0(1), \beta(0)), \end{aligned}$$

where u_0 is the unique complex gauge transformation gauging α to zero such that $u_0(0) = \mathrm{id}$. It can be checked that Φ is holomorphic with respect to the complex structure I and that it gives an isomorphism when viewed as a map from the moduli space of solutions to the complex equation modulo complex gauge transformations. We would like to show that it is an isomorphism when interpreted as a map on $\mathcal{M}_{Sch}$. Note that in this way, Kronheimer in [38] has constructed a hyperkähler metric on $T^*G^{\mathbb{C}}$ where G is compact Lie group. In fact, his proof shows that Φ pulls back the standard symplectic form on $T^*\mathrm{GL}(n,\mathbb{C}) \cong \mathrm{GL}(n,\mathbb{C}) \times \mathfrak{gl}(n,\mathbb{C})$ to the symplectic form $\omega_I^{\mathbb{C}}$.

In his paper [21], Donaldson shows that one can identify the space of solutions to the complex Nahm equation modulo complex gauge transformations with the moduli space of solutions to the full Nahm equations modulo unitary gauge transformations. He shows that every complex gauge orbit of a solution to the complex equation contains a solution to the real equation by solving a variational problem, which in his case is always possible. We have the analogous observation, but there is the question of solvability.

Proposition 3.5.1. *Let (α, β) be a solution to the complex equation. Consider the $\mathcal{G}$-invariant functional*

$$\mathcal{L} : \mathcal{G}^{\mathbb{C}} \to \mathbb{R}$$

given by

$$\mathcal{L}(u) = \int_0^1 |u.\alpha + (u.\alpha)^*|^2 - |u.\beta|^2.$$

Then u is a critical point of $\mathcal{L}$ if and only if $(u.\alpha, u.\beta)$ satisfies the real equation.

Proof. The proof is completely analogous to the proof of the corresponding statement in the Nahm case given in [21], lemma 2.3. □

We note that in fact everything we did so far is still $\mathcal{G}$-invariant. So what we actually want is to find a critical point of $\mathcal{L}$ which is only defined up to unitary gauge transformations, i.e. which takes values in the homogenous space $H = \mathrm{GL}(n,\mathbb{C})/\mathrm{U}(n)$. We identify the space H with the space of self-adjoint positive-definite matrices via $u \mapsto uu^* = h$. In the gauge where $\alpha = 0$, we then get for the Lagrangian $\mathcal{L}$, using $|u.\beta|^2 = \mathrm{tr}(u^{-1}\beta u(u^{-1}\beta u)^*) = \mathrm{tr}(h^{-1}\beta h\beta^*)$,

$$\mathcal{L}(h) = \int_{\mathbb{R}} \mathrm{tr}(h^{-1}\dot{h})^2 - \mathrm{tr}(h^{-1}\beta h\beta^*).$$

Since $\mathrm{tr}(h^{-1}\dot{h})^2$ is the natural Riemannian metric on the homogenous space H, we can equally well think of $\mathcal{L}$ as the action functional associated to a particle moving in the Riemannian manifold H under the influence of the positive potential $V(h) = \mathrm{tr}(h^{-1}\beta h\beta^*)$. In the Nahm case there is an analogous statement, but there the potential is $-V$. In this case, the functional $\mathcal{L}$ is positive definite and the potential is geodesically convex. Therefore, it can be shown that a minimiser always exists. Due to the change of signature, this argument does not work in our situation.

However, we will see that proposition 2.3.5 still holds in this case. To simplify computations, we define the following operators acting on $C^\infty([0,1],\mathbb{C}^n)$,

$$\begin{aligned}
\bar\partial_\alpha &= \frac{\mathrm{d}}{\mathrm{d}t} + \alpha \\
\partial_\alpha &= \frac{\mathrm{d}}{\mathrm{d}t} - \alpha^* \\
\bar\partial_\beta &= \beta \\
\partial_\beta &= -\beta^*.
\end{aligned}$$

Note that in the usual way, we get induced operators on $C^\infty([0,1],\mathrm{End}(\mathbb{C}^n))$ by putting

$$\begin{aligned}
\bar\partial_\alpha &= \frac{\mathrm{d}}{\mathrm{d}t} + [\alpha,-] \\
\partial_\alpha &= \frac{\mathrm{d}}{\mathrm{d}t} - [\alpha^*,-] \\
\bar\partial_\beta &= [\beta,-] \\
\partial_\beta &= -[\beta^*,-].
\end{aligned}$$

Lemma 3.5.2.

$$\mu_I(\alpha,\beta) = [\partial_\alpha,\bar\partial_\alpha] - [\partial_\beta,\bar\partial_\beta],$$

as operators on $C^\infty([0,1],\mathbb{C}^n)$.

Proof. This is a straight-forward computation,

$$[\partial_\alpha,\bar\partial_\alpha] = [\frac{\mathrm{d}}{\mathrm{d}t} - \alpha^*, \frac{\mathrm{d}}{\mathrm{d}t} + \alpha] = \dot\alpha + \dot\alpha^* + [\alpha,\alpha^*]$$

$$[\partial_\beta,\bar\partial_\beta] = -[\beta^*,\beta] = [\beta,\beta^*].$$

□

We would like to know how the real equation behaves under complex gauge transformations. First, we determine the transformation behaviour of the operators introduced above.

Lemma 3.5.3. *Let* $u \in \mathcal{G}^{\mathbb{C}}$ *be a complex gauge transformation. Then*

$$\begin{aligned}
\bar\partial_{u.\alpha} &= u^{-1} \circ \bar\partial_\alpha \circ u \\
\partial_{u.\alpha} &= u^* \circ \partial_\alpha \circ (u^*)^{-1} \\
\bar\partial_{u.\beta} &= u^{-1} \circ \bar\partial_\beta \circ u \\
\partial_{u.\beta} &= u^* \circ \partial_\beta \circ (u^*)^{-1}.
\end{aligned}$$

Proof. Direct calculation. □

Using this, we can now work out the transformation behaviour of the real equation.

Proposition 3.5.4. *Let $u \in \mathcal{G}^{\mathbb{C}}$ be a complex gauge transformation, put $h = uu^*$. Then*

$$u\mu_I(u.\alpha, u.\beta)u^{-1} = \mu(\alpha,\beta) - \bar{\partial}_\alpha(h(\partial_\alpha h^{-1})) + \bar{\partial}_\beta(h(\partial_\beta h^{-1})).$$

Proof. This is again a computation using the identities from the lemma,

$$\begin{aligned} u \circ \mu_I(u.\alpha, u.\beta) \circ u^{-1} &= u \circ [\partial_\alpha, \bar{\partial}_\alpha] \circ u^{-1} - u \circ [\partial_\beta, \bar{\partial}_\beta] \circ u^{-1} \\ &= u \circ \partial_{u.\alpha} \circ u^{-1} u \circ \bar{\partial}_{u.\alpha} \circ u^{-1} - u \circ \bar{\partial}_{u.\alpha} \circ u^{-1} u \circ \partial_{u.\alpha} \circ u^{-1} \\ &\quad - u \circ \partial_{u.\beta} \circ u^{-1} u \circ \bar{\partial}_{u.\beta} \circ u^{-1} + u \circ \bar{\partial}_{u.\beta} \circ u^{-1} u \circ \partial_{u.\beta} \circ u^{-1} \\ &= uu^* \circ \partial_\alpha \circ (u^*)^{-1} u^{-1} \circ \bar{\partial}_\alpha - \bar{\partial}_\alpha uu^* \partial_\alpha \circ (u^*)^{-1} u^{-1} \\ &\quad - uu^* \circ \partial_\beta \circ (u^*)^{-1} u^{-1} \circ \bar{\partial}_\beta + \bar{\partial}_\beta uu^* \partial_\beta \circ (u^*)^{-1} u^{-1} \\ &= [h \circ \partial_\alpha \circ h^{-1}, \bar{\partial}_\alpha] - [h \circ \partial_\beta \circ h^{-1}, \bar{\partial}_\beta] \\ &= [\partial_\alpha + h(\partial_\alpha h^{-1}), \bar{\partial}_\alpha] - [\partial_\beta + h(\partial_\beta h^{-1}), \bar{\partial}_\beta] \\ &= \mu_I(\alpha,\beta) - \bar{\partial}_\alpha(h(\partial_\alpha h^{-1})) + \bar{\partial}_\beta(h(\partial_\beta h^{-1})). \end{aligned}$$

□

By polar decomposition we can always write

$$u = pv,$$

where p is self-adjoint and positive definite and v is unitary. Then we get $h = p^2$. If we are given a solution (α, β) to the complex equation, we now want to find h self-adjoint and positive such that

$$\mu_I(\alpha,\beta) - \bar{\partial}_\alpha(h(\partial_\alpha h^{-1})) + \bar{\partial}_\beta(h(\partial_\beta h^{-1})) = 0.$$

This will then imply that the complex gauge transformation $u = h^{1/2}$ takes (α, β) to a solution of the real equation. As observed above, any solution to the complex equation is gauge-equivalent (under a complex gauge transformation) to a solution with $\alpha = 0$ and β constant. This implies

$$\mu_I(0,\beta) = -[\beta, \beta^*],$$

and the equation we want to solve becomes

$$-[\beta,\beta^*] - \frac{\mathrm{d}}{\mathrm{d}t}(h\dot{h}^{-1}) - [\beta, h[\beta^*, h^{-1}]] = 0.$$

A little manipulation gives

$$\frac{\mathrm{d}}{\mathrm{d}t}(\dot{h}h^{-1}) - [\beta, h\beta^* h^{-1}] = 0,$$

or more explicitly

$$\ddot{h}h^{-1} + \dot{h}\dot{h}^{-1} - [\beta, h\beta^* h^{-1}] = 0.$$

Thus, the solution theory of this equation will answer the question of *existence*. To treat *uniqueness*, we start with (α, β) solving both equations, i.e. a genuine solution to Schmid's equations, and we have to study the equation

$$-\bar{\partial}_\alpha(h(\partial_\alpha h^{-1})) + \bar{\partial}_\beta(h(\partial_\beta h^{-1})) = 0.$$

Expanding this yields a more explicit form:

$$-\frac{\mathrm{d}}{\mathrm{d}t}(h\dot{h}^{-1} - h\alpha^* h^{-1} - \alpha) - [\alpha, h\dot{h}^{-1} - h\alpha^* h^{-1}] - [\beta, h\beta^* h^{-1}] = 0.$$

Recall that our gauge group consists of gauge transformations that are equal to the identity at the endpoints of the interval. So the uniqueness question actually yields a boundary value problem. We have to require that $h(0) = h(1) = \mathrm{id}$. Now suppose that we have a one-parameter family

$$h(s,t) = \exp(\xi(s,t)), \qquad h(0,t) = \mathrm{id}.$$

We wish to compute the linearisation of the above equation. We denote the so-obtained linear operator by L. Write ξ' for the partial derivative of ξ with respect to s at $s = 0$. We want to compute $L\xi'$, that is,

$$\begin{aligned} L\xi' &= \frac{\mathrm{d}}{\mathrm{d}s}|_{s=0}\left(\frac{\mathrm{d}}{\mathrm{d}t}(-h\dot{h}^{-1} + h\alpha^* h^{-1} + \alpha) - [\alpha, h\dot{h}^{-1} - h\alpha^* h^{-1}] - [\beta, h\beta^* h^{-1}]\right) \\ &= \ddot{\xi}' - [\dot{\alpha}^*, \xi'] - [\alpha^*, \xi'] + [\alpha, \xi'] - [\alpha, [\alpha^*, \xi']] + [\beta, [\beta^*, \xi']]. \end{aligned}$$

Writing $\alpha = T_0 - iT_1$ and $\beta = T_2 + iT_3$, we obtain

$$\begin{aligned} L\xi' &= \ddot{\xi}' - [\dot{\alpha}^*, \xi'] - [\alpha^*, \xi'] + [\alpha, \xi'] - [\alpha, [\alpha^*, \xi']] + [\beta, [\beta^*, \xi']] \\ &= \ddot{\xi}' + [\dot{T}_0, \xi'] + i[\dot{T}_1, \xi'] + [T_0 + iT_1, \xi'] + [T_0 - iT_1, \xi'] + [T_0 - iT_1, [T_0 + iT_1, \xi']] \\ &\quad -[T_2 + iT_3, [T_2 - iT_3, \xi']] \\ &= \ddot{\xi}' + [\dot{T}_0, \xi'] - i[[T_2, T_3], \xi'] + 2[T_0, \xi'] + [T_0 - iT_1, [T_0, \xi'] + [iT_1, \xi']] \\ &\quad -[T_2 + iT_3, [T_2, \xi'] - i[T_3, \xi']] \\ &= \ddot{\xi}' + 2[T_0, \xi'] + [\dot{T}_0, \xi'] + [T_0, [T_0, \xi']] + [T_1, [T_1, \xi']] - [T_2, [T_2, \xi']] - [T_3, [T_3, \xi']]. \end{aligned}$$

That is, if $h(s,t)$ solves the uniqueness-equation for all s, then ξ' defines an element of the degeneracy space at $\mathcal{T} = (T_0, T_1, T_2, T_3)$. We sum up this observation in a proposition.

Proposition 3.5.5. *Away from the degeneracy locus, there is only a discrete set of solutions to the uniqueness-equation. The linearisation of the uniqueness-equation is the degeneracy equation. And so the degeneracy space can be thought of as the tangent space to the space of self-adjoint solutions to the uniqueness-equation.*

Of course, there may be obstructions to producing a genuine solution of the uniqueness-equation from an infinitesimal one. So the proposition only says, that if we have a solution to Schmid's equations that admits a smooth family of complex gauge transformations leaving the equation invariant, it has to be contained in the degeneracy locus. So away from the degeneracy locus we expect at most a discrete set of complex gauge transformations that preserve the full Schmid equations.

Remark. When studying the uniqueness equation we start with a solution (α, β) to the real and the complex equation, apply a complex gauge transformation, u say, and require that $u.(\alpha, \beta)$ should still solve the real equation. We view this as a condition on the self-adjoint gauge transformation $h = uu^*$. Note that the real equation is invariant under unitary gauge transformations, i.e. if u is a solution, then vu is also a solution for any unitary gauge transformation v. Since we can diagonalise h by a unitary gauge transformation v, i.e. vhv^{-1} is diagonal, we may, after replacing u by vu, assume that $h(t)$ is diagonal when trying to solve the uniqueness-equation. In the future, we hope to provide a more explicit analysis of the uniqueness-equation in the $\mathfrak{su}(2)$ case.

3.5.2 Product Structures

In this subsection we investigate to what extent we can identify the Schmid moduli space as a parakähler quotient in the spirit of our discussion of paracomplex Kirwan theorems earlier in this thesis. As a first step, we identify the configuration space $\mathcal{A}$ equipped with the product structure S as a paracomplex, in fact parakähler, manifold. Let

$$\mathcal{B} = \left\{ \frac{\mathrm{d}}{\mathrm{d}t} + A \;|A : U \to \mathfrak{u}(n) \right\}$$

be the space of unitary connections on the interval U. Then

$$T^*\mathcal{B} = \left\{ \left(\frac{\mathrm{d}}{\mathrm{d}t} + A, B \right) \;|A, B : U \to \mathfrak{u}(n) \right\}.$$

Proposition 3.5.6. *The map* $P : (\mathcal{A}, S) \to (T^*\mathcal{B} \times T^*\mathcal{B}, 1 \oplus (-1))$ *given by*

$$P(\frac{\mathrm{d}}{\mathrm{d}t} + T_0, T_1, T_2, T_3) = ((\frac{\mathrm{d}}{\mathrm{d}t} + T_0 + T_2, T_1 + T_3), (\frac{\mathrm{d}}{\mathrm{d}t} + T_0 - T_2, T_1 - T_3))$$

is a diffeomorphism respecting the product structures.

Proof. It is straight-forward to compute the derivative of P at a point $\mathcal{T} \in \mathcal{A}$. It can be represented by the following 4×4 matrix acting on $T_{\mathcal{T}}\mathcal{A} = C^\infty(U, \mathfrak{g})^4 \cong T(T^*\mathcal{B}) \oplus T(T^*\mathcal{B})$,

$$\mathrm{d}P = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix}.$$

The product structure S is given in this notation by

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

As 2×2 block matrices we may write

$$\mathrm{d}P = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \qquad S = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad 1 \oplus (-1) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Then it is straight-forward to check that

$$\mathrm{d}P \circ S = (1 \oplus (-1)) \circ \mathrm{d}P.$$

□

Now P is not just a diffeomorphism. We observe that it intertwines the symplectic form ω_I and the standard symplectic form on $T^*\mathcal{B} \times T^*\mathcal{B}$.

Proposition 3.5.7. *Let $\Omega = \omega_{T^*\mathcal{B}} \oplus \omega_{T^*\mathcal{B}}$ be the symplectic form on $T^*\mathcal{B} \times T^*\mathcal{B}$ formed by adding together the standard symplectic forms on each of the $T^*\mathcal{B}$ factors. Then*

$$\omega_I = P^*\Omega.$$

Proof. ω_I is given by

$$\omega_I(X, Y) = -g(X_0, Y_1) + g(X_1, Y_0) - g(X_2, Y_3) + g(X_3, Y_2),$$

where $X = (X_0, X_1, X_2, X_3), Y = (Y_0, Y_1, Y_2, Y_3) \in T\mathcal{A}$ are two tangent vectors. The symplectic form Ω is given by

$$\Omega(Z, W) = g(Z_0, W_1) - g(Z_1, W_0) + g(Z_2, W_3) - g(Z_3, W_2),$$

with $Z = (Z_0, Z_1, Z_2, Z_3), W = (W_0, W_1, W_2, W_3) \in T(T^*\mathcal{B} \times T^*\mathcal{B}) = T(T^*\mathcal{B}) \oplus T(T^*\mathcal{B})$. Now $\mathrm{d}P$ maps a tangent vector X on $\mathcal{A}$ to a tangent vector Z on $T^*\mathcal{B} \times T^*\mathcal{B}$ of the form

$$Z = (X_0 + X_2, X_1 + X_3, X_0 - X_2, X_1 + X_3).$$

Plugging Z, W of the above form for some X, Y tangent to $\mathcal{A}$ into the formula for Ω, gives the desired result, i.e.

$$\Omega(Z, W) = \omega_I(X, Y).$$

$\square$

In the co-ordinates on $\mathcal{B} \times \mathcal{B}$, Schmid's equations are then equivalent to the system

$$\begin{aligned} \left[\frac{\mathrm{d}}{\mathrm{d}t} + A_1, B_1\right] &= 0 \\ \left[\frac{\mathrm{d}}{\mathrm{d}t} + A_2, B_2\right] &= 0 \\ \left[\frac{\mathrm{d}}{\mathrm{d}t} + \frac{1}{2}(A_1 + A_2), A_1 - A_2\right] + [B_1, B_2] &= 0. \end{aligned}$$

The T_i's can be recovered from a solution to this system via

$$\begin{aligned} A_1 &= T_0 + T_2 \\ A_2 &= T_0 - T_2 \\ B_1 &= T_1 + T_3 \\ B_2 &= T_1 - T_3. \end{aligned}$$

In analogy to the situation of complex structures, we may view the first two equations above as a single "paracomplex" equation and the third equation as the real equation. Note that clearly the paracomplex system is invariant under paracomplex gauge transformations, i.e. elements of $\mathcal{G} \times \mathcal{G}$ acting componentwise:

$$(u_1, u_2).\left(\frac{\mathrm{d}}{\mathrm{d}t} + A_i, B_i\right)_{i=1,2} = \left(\frac{\mathrm{d}}{\mathrm{d}t} + u_i^{-1}A_iu_i + u_i^{-1}\dot{u}_i, u_i^{-1}B_iu_i\right)_{i=1,2}.$$

This allows us to construct the following map. Let us consider the moduli space $\mathcal{B}'$ of solutions to the Lax equation

$$[\frac{\mathrm{d}}{\mathrm{d}t} + A_1, B_1] = 0$$

modulo the group $\mathcal{G}_{00}$ of gauge transformations equal to the identity at the endpoints of the bounded interval U.

If we have a solution (A_1, B_1) and apply a gauge transformation $u \in \mathcal{G}$ (not necessarily satisfying the boundary conditions of $\mathcal{G}_{00}$) such that $u^{-1}A_1u + u^{-1}\dot{u} = 0$, i.e. $A_1 = -\dot{u}u^{-1}$, then we see that $u^{-1}Bu$ has to be constant. Thus any solution is gauge equivalent to a solution of the form $u.(0, B)$, with $B \in \mathfrak{u}(n)$ constant and $u \in \mathcal{G}$. This shows that we have a surjection

$$\mathcal{B}' \to \mathrm{U}(n) \times \mathfrak{u}(n), \qquad (A, B) \mapsto (u(1), B(1)),$$

where again u is the unique gauge transformation that gauges away A and satisfies $u(0) = \mathrm{id}$. This map is injective modulo $\mathcal{G}_{00}$. We may view $\mathcal{B}'$ as a symplectic quotient of $T^*\mathcal{B}$ by the action of the group of gauge transformations that vanish at the endpoints of the interval U. In fact, it is straight-forward to check that this map identifies $\mathcal{B}'$ and $T^*\mathrm{U}(n)$ as symplectic manifolds. By applying this map on each factor of $T^*\mathcal{B} \times T^*\mathcal{B}$ and composing with the map P from above, we obtain a map from the moduli space of solutions to Schmid's equations to $T^*\mathrm{U}(n) \times T^*\mathrm{U}(n)$, intertwining the symplectic structures ω_I and Ω.

So in order to identify the moduli space $\mathcal{M}_{Sch}$ as a paracomplex quotient, we would like to prove that for any solution (A_i, B_i) to the paracomplex equation there exist gauge transformations (u_1, u_2) such that $(A'_i, B'_i) = (u_i.A_i, u_i.B_i)$ solves the real equation. A natural first step is to fix a gauge in which the equations become simpler. Note that the gauge transformations u_1 and u_2 are independent.

We use u_2 to gauge away A_2 so that $B'_2 =: B$ is therefore constant. Then u_1 has to solve the second order equation

$$\dot{A}'_1 + [B'_1, B] = 0.$$

More explicitly,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(u_1^{-1}A_1u_1 + u^{-1}\dot{u}_1\right) + [u^{-1}B_1u_1, B] = 0.$$

Hence, if we want to identify the moduli space as a paracomplex quotient, we have to know to what extent a solution to this equation is unique. This is a question we would like to answer in the future.

3.6 Explicit Solutions

We have seen above that Schmid's equations form what is called an *integrable system*. There is no precise definition what this actually means, but there are some recognisable key features. The equations can be put in Lax form and we have proved that a solution defines a linear flow on the Jacobian of the spectral curve. In other words, we have translated the problem of solving this non-linear system of ordinary differential equations into a *linear* one in complex algebraic geometry. Since the solutions are described by curves on the Jacobian of the spectral curve they are, at least in principle, expressible in terms of θ-functions. Therefore, we might hope that at least in some fortunate situations we could be able to construct explicit solutions in terms of classical special functions.

In this section, we will see that this is possible and hence produce new examples of explicit solutions to Schmid's equations satisfying certain additional symmetry conditions. Our method will work in principle for any simple Lie algebra and we carry it out explicitly to obtain solutions with values in $\mathfrak{su}(n)$. We have seen above that $\mathrm{SO}(1,2)$ acts on the moduli space of solutions to Schmids equations and that we always have a map Ψ from the moduli space of solutions to Schmid's equations modulo the action of G_L to the set of self-adjoint traceless endomorphisms of $\mathbb{R}^{1,2}$. If $\Psi(\mathcal{T})$ is diagonalisable by an element $A \in \mathrm{SO}(1,2)$, we see that at every point the matrices in $A\mathcal{T}$ are mutually orthogonal in $\mathfrak{g}$.

3.6.1 Review of the $\mathfrak{su}(2)$-Case

In the case of $\mathfrak{g} = \mathfrak{su}(2)$, which has been completely worked out by Matsoukas in his thesis [48], this observation allows us to simplify the equations substantially. As a vector space $\mathfrak{su}(2)$ is three-dimensional. Thus, if we have a solution $\mathcal{T}$ to Schmid's equations with $T_0 = 0$ and the other T_i mutually orthogonal, the equations imply that the directions of the T_i's are constant in time. We may therefore assume that the T_i are of the form

$$T_i(t) = f_i(t)\sigma_i \qquad i = 1, 2, 3,$$

where $\{\sigma_1, \sigma_2, \sigma_3\}$ form a standard orthonormal basis of $\mathfrak{su}(2)$. Schmid's eqations then read

$$\begin{aligned} \dot{f}_1 &= -f_2 f_3 \\ \dot{f}_2 &= f_1 f_3 \\ \dot{f}_3 &= f_1 f_2. \end{aligned}$$

The general solution is given by

$$\begin{aligned} f_1(t) &= kD\mathrm{sn}_k(Dt + C) \\ f_2(t) &= kD\mathrm{cn}_k(Dt + C) \\ f_3(t) &= -D\mathrm{dn}_k(Dt + C), \end{aligned}$$

where D and C are arbitrary real constants, $k \in [0, 1]$, and sn_k, cn_k and dn_k are the Jacobi elliptic functions. Our main reference for the definitions and properties of special functions is [59].

3.6.2 Solutions with Cyclic Symmetry

We now wish to construct explicit solutions with values in Lie algebras of higher rank. We do not hope to solve the full equations in this case, but we may obtain explicit solutions in terms of classical special functions if we require that our solutions should be invariant under some discrete group of symmetries. For the solution of this problem in the Nahm case, see [32], [12]. Recall that we have an action of $\mathrm{PSL}(2,\mathbb{R}) = \mathrm{SO}(1,2)$ on the moduli space of solutions to Schmid's equations. We are looking for fixed points of this action. From now on, we will always fix a gauge such that $T_0 = 0$. Consider an element $A \in \mathrm{SO}(1,2)$ of the form

$$A_\theta = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix},$$

for some $\theta \in 2\pi\mathbb{Q}$. We wish to characterise solutions invariant under A_θ. So let $\mathcal{T} = (T_1, T_2, T_3)$ be a solution to Schmid's equations. $\mathcal{T}$ is invariant, if there exists a gauge transformation $u \in \mathcal{G}_{Sch}$ such that

$$A_\theta.\mathcal{T} = u.\mathcal{T}.$$

Observe that applying the gauge transformation u will in general make the T_0-term in $u.\mathcal{T}$ non-zero. We wish to translate this criterion therefore into a condition on the spectral curve associated to $\mathcal{T}$ which, as we have seen, is invariant under arbitrary gauge transformations, even if they do not satisfy the boundary conditions of $\mathcal{G}_{Sch}$.

The spectral curve is defined as the vanishing locus of the characteristic polynomial of $T = (T_2 + iT_3) - 2\zeta T_1 + \zeta^2(T_2 - iT_3)$. Now the matrix A_θ will act on $T_2 + iT_3$ via multiplication by $e^{i\theta}$ and will leave T_1 unchanged. Thus, the spectral curve of $A_\theta.\mathcal{T}$ is given by

$$\begin{aligned} 0 &= \det(\eta + e^{i\theta}(T_2 + iT_3) - 2\zeta T_1 + \zeta^2 e^{-i\theta}(T_2 - iT_3)) = 0 \\ &= \det(e^{i\theta}(e^{-i\theta}\eta + (T_2 + iT_3) - 2e^{-i\theta}\zeta T_1 + \zeta^2 e^{-2i\theta}(T_2 - iT_3))) \\ &= (e^{in\theta}\det(e^{-i\theta}\eta + (T_2 + iT_3) - 2e^{-i\theta}\zeta T_1 + \zeta^2 e^{-2i\theta}(T_2 - iT_3)). \end{aligned}$$

Thus, we arrive at the following observation.

Lemma 3.6.1. *Let $\mathcal{T}$ be a solution to Schmid's equations with spectral curve $P(\eta,\zeta) = 0$. Then the spectral curve of $A_\theta.\mathcal{T}$ is given by $P(e^{-i\theta}\eta, e^{-i\theta}\zeta) = 0$.*

We conclude that a solution $\mathcal{T}$ is cyclically invariant if and only if the $P(\eta,\zeta)$ and $P(e^{-i\theta}\eta, e^{-i\theta}\zeta)$ define the same curve. So if we look for solutions which are invariant under A_{θ_n}, where $\theta_n = 2\pi/n$, then the non-zero terms in the polynomials defining their spectral curves have to be of fixed degree k modulo n. Since the leading term of $P(\zeta,\eta)$ is of degree n, the only such k can be n itself. Thus, all non-zero terms in P have to be of degree zero modulo n. What can we say about the cyclically symmetric curves?

Proposition 3.6.2. $S = \{P(\zeta,\eta) = 0\}$ *is cyclically invariant if and only if P is of the form*

$$P(\zeta,\eta) = \eta^n + \sum_{i=1}^{n-1} a_i \zeta^i \eta^{n-i} + (-1)^n \bar{b} + b\zeta^{2n},$$

where $b \in \mathbb{C}$, a_i is real if i is even, and purely imaginary if i is odd. In particular, if $b \neq 0$ then the quotient curve S/C_n is smooth of genus $g(S/C_n) = (n-1)$.

Proof. This is just the reality condition satisfied by the spectral curve observed earlier together with the invariance assumption. $b \neq 0$ implies that the point $(\zeta,\eta) = (0,0)$ does not lie on S and hence the action of C_n is free on S. The genus of the quotient curve can then be obtained from the Riemann-Hurwitz formula:

$$\chi_S = n\chi_{S/C_n}.$$

We have seen earlier that $g(S) = (n-1)^2$ and so we get

$$2 - 2(n-1)^2 = n(2 - 2g(S/C_n)),$$

i.e.

$$-2n^2 + 4n = 2n - 2ng(S/C_n).$$

Hence, $g(S/C_n) = n - 1$, as desired. □

We will see that in fact more is true, see [12]. Write $b = |b|e^{i\alpha}$. We have seen that acting by $e^{i\theta}$ on a solution $\mathcal{T}$, corresponds to transforming the polynomial $P(\eta,\zeta)$ defining the spectral curve to $e^{in\theta}P(e^{-i\theta}\eta, e^{-i\theta}\zeta)$. Thus, after applying an auxiliary rotation by $\theta = \alpha/n$, we may assume that b is real. Then the equation defining the spectral curve may be written in the form

$$P(\zeta,\eta) = \eta^n + \sum_{i=1}^{n-1} a_i \zeta^i \eta^{n-i} + b((-1)^n + \zeta^{2n}).$$

Now consider the C_n-invariant quantities

$$z = \eta/\zeta \qquad \tilde{w} = b\zeta^n.$$

These give co-ordinates on the quotient curve S/C_n and we want to derive the equation satisfied by them, i.e. the equation of the quotient curve. In these co-ordinates, after multiplication by ζ^{-n}, the equation of S reads

$$z^n + \sum_{i=1}^{n-1} a_i z^{n-i} + \tilde{w} + \frac{b^2(-1)^n}{\tilde{w}} = 0.$$

Now put $w = \tilde{w} - \frac{b^2(-1)^n}{\tilde{w}}$. Then we may write

$$w^2 = (\tilde{w} + \frac{b^2(-1)^n}{\tilde{w}})^2 - (-1)^n 4b^2 = \left(z^n + \sum_{i=1}^{n-1} a_i z^{n-i} \right)^2 - (-1)^n 4b^2.$$

So we arrive at the following proposition.

Proposition 3.6.3. *After a rotation, the quotient curve by the C_n-action of the spectral curve associated to a cyclically symmetric solution $\mathcal{T}$ is hyperelliptic. That is, S/C_n is defined by a equation of the form*

$$w^2 = F(z).$$

Recall that we have an anti-holomorphic involution acting on the spectral curve with fixed points. In an earlier section, we have observed that the real points are given by points (ζ, η) on the curve of the form

$$(\zeta, \eta) = (e^{i\theta}, iae^{i\theta}), \qquad a \in \mathbb{R}.$$

Now since C_n acts as multiplication by $e^{i\theta_n}$, we come to the following proposition:

Proposition 3.6.4. *For a C_n-invariant solution, the real curve inside the spectral curve is preserved by the C_n-action. So if the spectral curve has cyclic symmetry, the real curve is also cyclically symmetric.*

Our aim in this section is to produce explicit solutions, which are invariant under the cyclic group C_n generated by A_{θ_n}.

Example. We revisit the case $\mathfrak{g} = \mathfrak{su}(2)$. Let

$$\sigma_1 = \frac{1}{2}\begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}, \qquad \sigma_2 = \frac{1}{2}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_3 = \frac{1}{2}\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}.$$

Then Matsoukas showed that up to the action of $\mathrm{SO}(1,2)$ and G_L the solutions are given by $T_i = f_i\sigma_i$, where $i = 1,2,3$. We compute the spectral curve of such a solution. In this case we get

$$T = f_2\sigma_2 + if_3\sigma_3 - 2\zeta f_1\sigma_1 + \zeta^2(f_2\sigma_2 - if_3\sigma_3).$$

And a direct computation gives

$$\det(\eta + T) = \eta^2 + \frac{1}{2}(2f_1^1 + f_2^2 + f_3^2)\zeta^2 + \frac{1}{4}(f_2^2 - f_3^2)(1+\zeta^4).$$

So this solution is invariant under the cyclic group generated by A_{θ_2}. Therefore, we get the following corollary of Matsoukas' analysis of the $\mathfrak{su}(2)$-case.

Corollary 3.6.5. *Any solution to Schmids equations with values in $\mathfrak{su}(2)$ may be gauged and rotated into a solution with C_2-symmetry.*

We will now rewrite this ansatz in a slightly different form, in which it will generalise to arbitrary compact simple Lie algebras. Namely, consider the complexification of $\mathfrak{su}(2)$ given by $\mathfrak{sl}(2,\mathbb{C})$. This has a standard basis of generators given by

$$E_1 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad F_1 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \qquad H_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

satisfying the standard $\mathfrak{sl}(2,\mathbb{C})$ relations

$$[H_1, E_1] = 2E_1, \qquad [H_1, F_1] = -2F_1, \qquad [E_1, F_1] = H_1.$$

Define

$$H_0 = -H_1, \qquad E_0 = F_1, \qquad F_0 = E_1.$$

Then they satisfy the same relations. Now make an ansatz as follows:

$$\begin{aligned} T_1(t) &= \frac{i}{2}\sum_{i=0}^{1} p_i(t)H_i \\ T_2(t) &= -\frac{1}{2}\sum_{i=0}^{1} q_i(t)(E_i - F_i) \\ T_3(t) &= \frac{i}{2}\sum_{i=0}^{1} q_i(t)(E_i + F_i), \end{aligned}$$

where the p_i and q_i have to be real-valued functions in order to make the T_i unitary. In fact, we are just rewriting Matsoukas' ansatz with $f_1 = p_1 - p_0$, $f_2 = q_1 - q_0$ and

$f_3 = q_0 + q_1$. Then these satisfy Schmid's equations if

$$\begin{aligned}\dot{p}_1 - \dot{p}_0 &= q_0^2 - q_1^2 \\ \dot{q}_1 - \dot{q}_0 &= (p_1 - p_0)(q_1 + q_0) \\ \dot{q}_1 + \dot{q}_0 &= (p_1 - p_0)(q_1 - q_0).\end{aligned}$$

Subtracting and adding the second and third equations, we arrive at

$$\begin{aligned}\dot{p}_1 - \dot{p}_0 &= q_0^2 - q_1^2 \\ \dot{q}_1 &= (p_1 - p_0)q_1 \\ \dot{q}_0 &= -(p_1 - p_0)q_0,\end{aligned}$$

which yields, upon rewriting $\phi_i = 2\log q_i$,

$$\begin{aligned}\ddot{\phi}_0 &= -2e^{\phi_0} + 2e^{\phi_1} \\ \ddot{\phi}_1 &= 2e^{\phi_0} - 2e^{\phi_1}.\end{aligned}$$

This can be written in the form

$$\begin{pmatrix} \ddot{\phi}_0 \\ \ddot{\phi}_1 \end{pmatrix} = -\begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix} \begin{pmatrix} e^{\phi_0} \\ e^{\phi_1} \end{pmatrix}.$$

Where we recognise $K = \begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix}$ as the generalised affine Cartan matrix of $\mathfrak{sl}(2,\mathbb{C})$, and the equations are the *affine Toda equations for the Lie algebra* $\mathfrak{sl}(2,\mathbb{C})$, except that in our situation a minus sign appears on the right-hand side.

Consider the two-sheeted cover of the unit circle given by the fixed point set of the anti-holomorphic involution and recall that the real curve is parametrised by

$$(\zeta, \eta) = (e^{i\theta}, e^{i(\theta \pm \pi/2)}\sqrt{A + 2B\cos(2\theta)}).$$

We can now read off from above that

$$A = \frac{1}{2}(2f_1^2 + f_2^2 + f_3^2), \qquad B = \frac{1}{4}(f_2^2 - f_3^2).$$

The C_2-action $(\zeta, \eta) \mapsto (-\zeta, -\eta)$ sends the point $(e^{i\theta}, e^{i(\theta\pm\pi/2)}\sqrt{A + 2B\cos(2\theta)})$ to $(e^{i(\theta+\pi)}, e^{i(\theta+\pi\pm\pi/2)}\sqrt{A + 2B\cos(2\theta)}) = (e^{i(\theta+\pi)}, e^{i((\theta+\pi)\pm\pi/2)}\sqrt{A + 2B\cos(2(\theta+\pi))})$, and so preserves each of the two sheets. Notice moreover, that the real curve is exactly given by $S \cap \{(\zeta,\eta) \mid |\zeta| = 1\}$, i.e. all points above the fixed circle in $\mathbb{CP}^1$ are real.

We now wish to extend this example to higher rank Lie algebras. In our construction we adapt Sutcliffe's ansatz for cyclically symmetric solutions to Nahm's equations [55]. So let $\mathfrak{g}$ be a compact simple real Lie algebra of rank n and let $\mathfrak{g}^{\mathbb{C}}$ be its complexification. Denote by $\mathfrak{h}$ the Cartan subalgebra of $\mathfrak{g}^{\mathbb{C}}$. Let Φ be the associated root system and choose a set of positive roots, Φ^+ and let $\Delta = \{\alpha_1, \ldots, \alpha_n\}$ be the set of simple roots. We can then find a set of *Chevalley generators* for $\mathfrak{g}^{\mathbb{C}}$, i.e. to each simple root α_i we associate a triple $(E_i, F_i, H_i) \in \mathfrak{g}_{\alpha_i} \times \mathfrak{g}_{-\alpha_i} \times \mathfrak{h}$, such that

$$\begin{aligned}
[E_i, F_j] &= \delta_{ij} H_j, \\
[H_i, E_j] &= C_{ji} E_j, \\
[H_i, F_j] &= -C_{ji} F_j, \\
\mathrm{ad}(E_i)^{-C_{ji}+1}(E_j) &= 0, \qquad \text{whenever } i \neq j \\
\mathrm{ad}(F_i)^{-C_{ji}+1}(F_j) &= 0, \qquad \text{whenever } i \neq j,
\end{aligned}$$

where $C_{ji} = (< \alpha_i, \alpha_j >) \in \mathrm{Mat}(n, \mathbb{Z})$ is the *Cartan matrix* of $\mathfrak{g}^{\mathbb{C}}$.

Now the affine Lie algebra associated to $\mathfrak{g}$ is given by extending the root system of $\mathfrak{g}$ by a set of generators asscociated to the *imaginary* root $\alpha_0 = -\sum_{i=1}^n a_i \alpha_i$, where $\theta = \sum_{i=1}^n a_i \alpha_i$ is the highest root of $\mathfrak{g}$, i.e. we add in the associated co-root $H_0 = -\sum_{i=1}^n c_i H_i$, $E_0 \in \mathfrak{g}_{\alpha_0}$ and $F_0 \in \mathfrak{g}_{\alpha_0}$ such that $[E_0, F_0] = H_0$. Then we will again produce an analogous set of relations, but we have to replace the Cartan matrix C by the generalised Cartan matrix $K \in \mathrm{Mat}(n+1, \mathbb{Z})$, from which C may be recovered by erasing the 0-th row and column.

However, while C was positive definite, K is only positive semi-definite and so the resulting affine Lie algebra associated to $\mathfrak{g}$ is infinite-dimensional. It can be realised as a certain extension of the infinite-dimensional Lie algebra of Laurent-polynomials with coefficients in $\mathfrak{g}$. For details on the theory of both finite-dimensional and affine Lie algebras we refer to [14].

With this technology introduced, we can now write down our ansatz for Schmid's equations with values in a general compact simple Lie algebra.

$$\begin{aligned}
T_1(t) &= \frac{i}{2} \sum_{i=0}^n p_i(t) H_i \\
T_2(t) &= -\frac{1}{2} \sum_{i=0}^n q_i(t)(E_i - F_i) \\
T_3(t) &= \frac{i}{2} \sum_{i=0}^n q_i(t)(E_i + F_i).
\end{aligned}$$

Then imposing Schmid's equations gives the following relations:

$$\begin{aligned}
\dot{T}_1 = -[T_2, T_3] \quad &\Longrightarrow \quad \sum_{i=0}^{n} \dot{p}_i H_i = \sum_{i=0}^{n} q_i^2 H_i, \\
\dot{T}_2 = [T_3, T_1] \quad &\Longrightarrow \quad \sum_{i=0}^{n} \dot{q}_i (E_i - F_i) = -\frac{1}{2} \sum_{i,j=0}^{n} q_i p_j K_{ij} (E_i - F_i), \\
\dot{T}_3 = [T_1, T_2] \quad &\Longrightarrow \quad \sum_{i=0}^{n} \dot{q}_i (E_i + F_i) = -\frac{1}{2} \sum_{i,j=0}^{n} q_i p_j K_{ij} (E_i + F_i).
\end{aligned}$$

Bearing in mind the definition of H_0, we read off the following set of equations

$$\begin{aligned}
\dot{p}_i - c_i \dot{p}_0 &= q_i^2 - c_i q_0^2 \qquad i = 1, \ldots, n, \\
\dot{q}_i &= -\left(\frac{1}{2} \sum_{j=0}^{n} K_{ij} p_j \right) q_i \qquad i = 0, \ldots, n.
\end{aligned}$$

Writing $\phi_i = 2 \log q_i$, this can be rewritten as

$$\ddot{\phi}_i = -\sum_{j=0}^{n} K_{ij} \dot{p}_j = -\sum_{j=0}^{n} K_{ij} (q_j^2 - c_j (q_0^2 - \dot{p}_0)) = -\sum_{j=0}^{n} K_{ij} (e^{\phi_j} - c_j (e^{\phi_0} - \dot{p}_0)),$$

where we set $c_0 = 1$ to give sense to the last sum. In general, this does not allow further simplification. However, if we suppose that the matrix K is symmetric, then general theory implies $\sum_{j=0}^{n} K_{ij} c_j = 0$, where we again take $c_0 = 1$, see for example [14]. In this case the equations become

$$\ddot{\phi}_i = -\sum_{j=0}^{n} K_{ij} e^{\phi_j},$$

and are therefore (up to the sign) equal to the affine Toda equations. In the Nahm case the above ansatz yields almost the same equations for the ϕ_i's, except that we have to replace the minus-sign with a plus-sign, [55]. We check, that this ansatz actually produces solutions with cyclic symmetry in the case $\mathfrak{g} = \mathfrak{su}(n)$. A simple direct calculation shows that we have in this case

$$T = T_2 + iT_3 - 2\zeta T_1 + \zeta^2 (T_2 - iT_3) = \sum_{j=0}^{n} q_j (E_j - \zeta^2 F_j) - i p_j \zeta H_j.$$

Proposition 3.6.6. *For $\mathfrak{g} = \mathfrak{su}(n)$, the spectral curve associated to a solution to Schmid's equation coming from the above ansatz is cyclically symmetric.*

Proof. Indeed, for $\mathfrak{su}(n)$, which has rank $n-1$, we are dealing with a symmetric Cartan matrix and so the above discussion applies. Let $\{\alpha_i \mid i = 1,\dots,n-1\}$ be the standard simple roots. The highest root θ is given by $\theta = \sum_i \alpha_i$. We choose the simple co-roots H_i and the associated E_i's and F_i's for $\mathfrak{su}(n)$ in the form

$$H_i = \begin{pmatrix} 0 & 0 & \cdots & & \cdots & 0 \\ 0 & \ddots & \cdots & & \cdots & 0 \\ \vdots & \vdots & 1 & 0 & & \vdots \\ & & 0 & -1 & & \\ \vdots & & & & \ddots & \vdots \\ 0 & 0 & \cdots & & \cdots & 0 \end{pmatrix}, \qquad E_i = \begin{pmatrix} 0 & 0 & \cdots & & \cdots & 0 \\ 0 & \ddots & \cdots & & \cdots & 0 \\ \vdots & \vdots & 0 & 1 & & \vdots \\ & & 0 & 0 & & \\ \vdots & & & & \ddots & \vdots \\ 0 & 0 & \cdots & & \cdots & 0 \end{pmatrix} = F_i^T,$$

where $i = \{1,\dots,n-1\}$. That is, the kl-th entry of H_i is given by $\delta_{ki}\delta_{li} - \delta_{k,i+1}\delta_{l,i+1}$, while E_i has as kl-th entry $\delta_{ki}\delta_{l,i+1}$. For $i = 0$, we get $H_0 = -\sum_{i=1}^{n-1} H_i$ and so

$$H_0 = \begin{pmatrix} -1 & 0 & \cdots & & \cdots & 0 \\ 0 & 0 & \cdots & & \cdots & 0 \\ \vdots & \vdots & \ddots & & & \vdots \\ & & & \ddots & & \vdots \\ \vdots & & & & 0 & 0 \\ 0 & 0 & \cdots & \cdots & 0 & 1 \end{pmatrix}, \qquad E_0 = \begin{pmatrix} 0 & & \cdots & & \cdots & 0 \\ & & \cdots & & \cdots & \\ \vdots & \vdots & \ddots & & & \vdots \\ \vdots & & & \ddots & & \vdots \\ 0 & & & & & \\ 1 & 0 & \cdots & \cdots & & 0 \end{pmatrix} = F_0^T.$$

We want to compute the characteristic polynomial of the Lax operator $T = \sum_{j=0}^{n} q_j(E_j - \zeta^2 F_j) - ip_j\zeta H_j$, i.e. the determinant of the matrix $\eta + T$, which is given by

$$\eta + T = \begin{pmatrix} \eta + i(p_1 - p_0)\zeta & q_1 & 0 & \cdots & & 0 & -q_0\zeta^2 \\ -q_1\zeta^2 & \eta + i(p_2 - p_1)\zeta & q_2 & 0 & \cdots & \cdots & 0 \\ 0 & -q_2\zeta^2 & \ddots & \ddots & 0 & \cdots & \vdots \\ \vdots & & & \ddots & & 0 & \\ & \ddots & \ddots & & \ddots & \ddots & \vdots \\ \vdots & & & & & \ddots & 0 \\ 0 & & & & & \ddots & q_{n-1} \\ q_0 & 0 & \cdots & 0 & & -q_{n-1}\zeta^2 & \eta + i(p_0 - p_{n-1})\zeta \end{pmatrix}.$$

This is a tri-diagonal matrix with corners, and there is a general formula for its determinant, a proof of which can be found in [50]. Explicitly, the formula reads in our situation

$$\det(\eta + T) = (-1)^{n+1}(\prod_{i=0}^{n-1} q_i(1 + (-1)^n\zeta^{2n}) + \operatorname{tr}\left(\prod_{i=1}^{n}\begin{pmatrix} \eta + (p_{n-i+1} - p_{n-i})\zeta & q_{n-i+1}^2\zeta^2 \\ 0 & 1 \end{pmatrix}\right),$$

where the indices are to be taken modulo n. Now an easy induction shows that

$$\operatorname{tr}\left(\prod_{i=1}^{n}\begin{pmatrix}\eta+(p_{n-i+1}-p_{n-i})\zeta & q_{n-i+1}^2\zeta^2\\ 0 & 1\end{pmatrix}\right)$$

is a polynomial in (η,ζ) such that all non-zero terms are of degree $0 \mod n$, whenever $n \geq 2$. $\square$

Having introduced this ansatz, it is an interesting question to what extent it is complete, i.e. can we obtain all cyclically symmetric solutions from it? It is known in the Nahm case, that an adapted version of the above ansatz gives all cyclically symmetric monopoles. We therefore conjecture that this might be also true in our situation. However, in order to prove this, we lack a bit of extra information. Since in the Nahm case the solutions corresponding to monopoles have poles at the endpoints of the interval on which they are defined, such that the residues form the standard irreducible n-dimensional representation of $\mathfrak{sl}(2,\mathbb{C})$, the action of C_n has to be modified in order to preserve these boundary conditions. An element $\omega \in C_n$ acts on a triple of Nahm matrices $\mathcal{T} = (T_1, T_2 + iT_3)$ via

$$\omega.\mathcal{T} = u(\omega)^{-1}(T_1, \omega(T_2 + iT_3))u(\omega),$$

where $u(\omega)$ is a gauge transformation such that $u(0) = u(1)$ is the image of ω under the irreducible representation. Viewing $\mathcal{T}$ as taking values in $\mathbb{R}^3 \otimes \mathfrak{su}(n)$, Hitchin, Manton and Murray [32] have determined this action of $\mathrm{SL}(2,\mathbb{C})$ in terms of irreducibles. Then using a theorem of Kostant, Braden in [12] was able to deduce that $u(\omega)$ has be of the form

$$u(\omega) = \operatorname{diag}(1, \omega, \omega^2, \ldots, \omega^{n-1}),$$

from which it follows that $\mathcal{T}$ is obtained from the above ansatz. In our situation we are lacking this extra bit of representation theoretic insight, but we still believe that the ansatz should yield a large class of cyclically symmetric solutions.

Conjecture 3.6.7. *Let $\omega_n = e^{\frac{2\pi i}{n}}$ and let $\mathcal{T} = (T_0, T_1, T_2, T_3)$ be a solution to Schmid's equations with cyclic symmetry, i.e. there exists a gauge transformation u such that*

$$u^{-1}T_0u + u^{-1}\dot{u} = T_0 \qquad u^{-1}T_1u = T_1 \qquad u^{-1}(T_2 + iT_3)u = \omega_n(T_2 + iT_3).$$

Then $\mathcal{T}$ is gauge equivalent to a solution obtained from the above ansatz.

A strategy of proof could be as follows. We diagonalise u by some other special unitary gauge transformation v:

$$u = vDv^{-1},$$

where $D \in \mathfrak{su}(n)$ is diagonal. Thus, since the SO(1, 2) action commutes with gauge transformations, we may assume without loss of generality that u is itself diagonal, i.e. $u = \mathrm{diag}(e^{i\theta_1}, \ldots, e^{i\theta_n})$, such that $\det u = 1$. Let $\{H_i, E_{ij}\}$ be the standard basis of $\mathfrak{sl}(n, \mathbb{C})$, i.e. the (ij)-th entry of E_{ij} is 1 while all the others are zero and H_i satisfies $(H_i)_{kl} = \delta_{ik}\delta_{il} - \delta_{i+1k}\delta_{i+1,l}$, where i ranges over $1, \ldots n-1$. Then we may write

$$T_2 + iT_3 = \sum_{i \neq j} a_{ij} E_{ij} + \sum_i b_i H_i.$$

Now u acts on $T_2 + iT_3$ by conjugation and satisfies therefore

$$u.(T_2 + iT_3) = \sum_{i,j} a_{ij} e^{i(\theta_i - \theta_j)} E_{ij} + \sum_i b_i H_i.$$

On the other hand, by assumption this should equal $\omega(T_2 + iT_3)$. So we conclude that

$$b_i = 0 \text{ for all } i = 1, \ldots, n-1, \qquad e^{i(\theta_i - \theta_j)} = \omega \text{ for all } i, j \text{ such that } a_{ij} \neq 0.$$

Expanding T_1 in the same fashion, i.e.

$$T_1 = \sum_{i,j} c_{ij} e^{i(\theta_i - \theta_j)} E_{ij} + \sum_i d_i H_i,$$

and using the condition that $u.T_1 = T_1$, we obtain

$$\theta_i = \theta_j \text{ for all } i, j \text{ such that } c_{ij} \neq 0.$$

We want to show that $u(t) = e^{if(t)}\mathrm{diag}(1, \omega, \omega^2, \ldots, \omega^{n-1})$, which would imply the claim. However, it is not clear how to proceed from this point. It is probably not possible to prove the result in full generality, and one has to find an appropriate regularity condition on the solution $\mathcal{T}$, to ensure that it is obtained from the above ansatz.

Explicit Solutions in the $\mathfrak{su}(n)$-Case

In the case where $\mathfrak{g} = \mathfrak{su}(n)$, we have a symmetric affine Cartan matrix, given by

$$K = \begin{pmatrix} 2 & -1 & 0 & \cdots & 0 & -1 \\ -1 & 2 & -1 & 0 & \cdots & 0 \\ 0 & -1 & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & & & 0 \\ 0 & & & & \ddots & -1 \\ -1 & 0 & \cdots & \cdots & -1 & 2 \end{pmatrix},$$

and thus the above construction gives

$$\ddot{\phi}_i = e^{\phi_{i+1}} + e^{\phi_{i-1}} - 2e^{\phi_i} \qquad i = 0, 1, \ldots, n-1,$$

where the indices are taken modulo n. Writing $f_j(t) = e^{\phi_j(t)} - 1$, we obtain the following system:

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\left[\log(1 + f_j(t))\right] = f_{j+1}(t) + f_{j-1}(t) - 2f_j(t).$$

To solve this, we need an elliptic functions identity, the proof of which is based on the proof of a similar identity that can be found in chapter 4.7 of [56].

Lemma 3.6.8.

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\log\left(\mathrm{dn}_k^2(t) - 1 + \frac{1}{\mathrm{sn}_k(v)}\right) = \mathrm{dn}_k^2(t+v) + \mathrm{dn}_k^2(t-v) - 2\mathrm{dn}_k^2(t)$$

Proof. We first prove a simpler identity, from which the lemma can then be deduced by an integration trick.

Lemma 3.6.9.

$$\mathrm{sn}_k^2(t+v) - \mathrm{sn}_k^2(t-v) = 2\frac{\mathrm{d}}{\mathrm{d}v}\frac{\mathrm{sn}_k(t)\mathrm{cn}_k(t)\mathrm{dn}_k(t)\mathrm{sn}_k^2(v)}{1 - k^2\mathrm{sn}_k^2(t)\mathrm{sn}_k^2(v)}$$

Proof. We have the following addition formula for the function sn_k:

$$\mathrm{sn}_k(t \pm v) = \frac{\mathrm{sn}_k(t)\mathrm{cn}_k(v)\mathrm{dn}_k(v) \pm \mathrm{sn}_k(v)\mathrm{cn}_k(t)\mathrm{dn}_k(t)}{1 - k^2\mathrm{sn}_k^2(t)\mathrm{sn}_k^2(v)}.$$

Therefore,

$$\mathrm{sn}_k^2(t+v) - \mathrm{sn}_k^2(t-v) = 4\frac{\mathrm{sn}_k(t)\mathrm{cn}_k(t)\mathrm{dn}_k(t)\mathrm{sn}_k(v)\mathrm{cn}_k(v)\mathrm{dn}_k(v)}{(1 - k^2\mathrm{sn}_k^2(t)\mathrm{sn}_k^2(v))^2}.$$

Moreover, from the relation $\frac{\mathrm{d}}{\mathrm{d}v}\mathrm{sn}_k(v) = \mathrm{cn}_k(v)\mathrm{dn}_k(v)$, it follows that

$$\frac{\mathrm{d}}{\mathrm{d}v}\mathrm{sn}_k^2(v) = 2\mathrm{sn}_k(v)\mathrm{cn}_k(v)\mathrm{dn}_k(v).$$

Thus,

$$\begin{aligned} 2\frac{\mathrm{d}}{\mathrm{d}v}\frac{\mathrm{sn}_k(t)\mathrm{cn}_k(t)\mathrm{dn}_k(t)\mathrm{sn}_k^2(v)}{1 - k^2\mathrm{sn}_k^2(t)\mathrm{sn}_k^2(v)} &= 4\frac{(1 - k^2\mathrm{sn}_k^2(t)\mathrm{sn}_k^2(v))\mathrm{sn}_k(t)\mathrm{cn}_k(t)\mathrm{dn}_k(t)\mathrm{sn}_k(v)\mathrm{cn}_k(v)\mathrm{dn}_k(v)}{(1 - k^2\mathrm{sn}_k^2(t)\mathrm{sn}_k^2(v))^2} \\ &\quad + 4\frac{k^2\mathrm{sn}_k(t)\mathrm{cn}_k(t)\mathrm{dn}_k(t)\mathrm{sn}_k^2(v)\mathrm{sn}_k^2(t)\mathrm{sn}_k(v)\mathrm{cn}_k(v)\mathrm{dn}_k(v)}{(1 - k^2\mathrm{sn}_k^2(t)\mathrm{sn}_k^2(v))^2}. \end{aligned}$$

The terms involving k^2 cancel and so we arrive at

$$\begin{aligned} 2\frac{\mathrm{d}}{\mathrm{d}v}\frac{\mathrm{sn}_k(t)\mathrm{cn}_k(t)\mathrm{dn}_k(t)\mathrm{sn}_k^2(v)}{1 - k^2\mathrm{sn}_k^2(t)\mathrm{sn}_k^2(v)} &= 4\frac{\mathrm{sn}_k(t)\mathrm{cn}_k(t)\mathrm{dn}_k(t)\mathrm{sn}_k(v)\mathrm{cn}_k(v)\mathrm{dn}_k(v)}{(1 - k^2\mathrm{sn}_k^2(t)\mathrm{sn}_k^2(v))^2} \\ &= \mathrm{sn}_k^2(t+v) - \mathrm{sn}_k^2(t-v), \end{aligned}$$

as desired. □

Using this, we can now prove the lemma. First note that $1 - \mathrm{dn}_k^2(t) = k^2\mathrm{sn}_k^2(t)$. Therefore, we may rewrite $1 - k^2\mathrm{sn}_k^2(t)\mathrm{sn}_k^2(v)$ as

$$1 - k^2\mathrm{sn}_k^2(t)\mathrm{sn}_k^2(v) = \mathrm{sn}_k^2(v)\left(\frac{1}{\mathrm{sn}_k^2(v)} - 1 + \mathrm{dn}_k^2(t)\right).$$

Multiplying the identity from the lemma by k^2 and using again that $1 - \mathrm{dn}_k^2(t) = k^2\mathrm{sn}_k^2(t)$, we get

$$\mathrm{dn}_k^2(t-v) - \mathrm{dn}_k^2(t+v) = 2k^2 \frac{\mathrm{d}}{\mathrm{d}v} \frac{\mathrm{sn}_k(t)\mathrm{cn}_k(t)\mathrm{dn}_k(t)}{\left(\frac{1}{\mathrm{sn}_k^2(v)} - 1 + \mathrm{dn}_k^2(t)\right)}.$$

Now integrate this with respect to v and define the function

$$E_k(t) = \int_0^t \mathrm{dn}_k^2(u)\mathrm{d}u.$$

This gives

$$E_k(t-v) + E_k(t+v) - 2E_k(t) = -2k^2 \frac{\mathrm{sn}_k(t)\mathrm{cn}_k(t)\mathrm{dn}_k(t)}{\left(\frac{1}{\mathrm{sn}_k^2(v)} - 1 + \mathrm{dn}_k^2(t)\right)}.$$

Now note that, since $\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{dn}_k^2(t) = -2k^2\mathrm{cn}_k(t)\mathrm{sn}_k(t)\mathrm{dn}_k(t)$, we can write

$$E_k(t-v) + E_k(t+v) - 2E_k(t) = \frac{\mathrm{d}}{\mathrm{d}t} \log\left(\frac{1}{\mathrm{sn}_k^2(v)} - 1 + \mathrm{dn}_k^2(t)\right).$$

Taking derivatives with respect to t on both sides then finishes the proof. □

With the identity $\frac{\mathrm{d}^2}{\mathrm{d}t^2} \log\left(\mathrm{dn}_k^2(t) - 1 + \frac{1}{\mathrm{sn}_k(v)}\right) = \mathrm{dn}_k^2(t+v) + \mathrm{dn}_k^2(t-v) - 2\mathrm{dn}_k^2(t)$ at hand, we may now write down explicit solutions for our system

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\left[\log(1 + f_j(t))\right] = f_{j+1}(t) + f_{j-1}(t) - 2f_j(t).$$

We observe that the right-hand side of the equation does not change if we add the same constant to all the f_j's. Let K denote the complete elliptic integral of the first kind. We know that on the real line dn_k is periodic with period $2K$. So we take an ansatz for f_j in the form

$$f_j(t) = \mathrm{dn}_k^2(t + j\frac{2K}{n} + C) + D,$$

where $C, D \in \mathbb{R}$ are constants. We see that if we put $D = \frac{1}{\mathrm{sn}_k^2(2K/n)} - 2$, then the equation we want to solve is transformed into precisely the identity just proved. Recall that by definition we have $f_j = e^{\phi_j} - 1 = q_j^2 - 1$. Thus,

$$q_j^2(t) = f_j(t) + 1 = \mathrm{dn}_k^2(t + j\frac{2K}{n} + C) + \frac{1}{\mathrm{sn}_k^2(2K/n)} - 1.$$

Note that since $\mathrm{dn}_k^2(t) = 1 - k^2\mathrm{sn}_k^2(t)$, we may write $q_j^2(t) = -k^2\mathrm{sn}_k^2(t + j\frac{2K}{n} + C) + \frac{1}{\mathrm{sn}_k^2(2K/n)}$. The first term in this expression is always less than or equal to 1, whereas the second term is always greater than 1. Hence, $q_j^2(t)$ is non-zero for all t. So, we may take the square root and obtain a smooth function

$$q_j(t) = \sqrt{\mathrm{dn}_k^2(t + j\frac{2K}{n} + C) + \frac{1}{\mathrm{sn}_k^2(2K/n)} - 1}.$$

To compute the p_j's, our general ansatz for $\mathfrak{su}(n)$, that is, we put $c_j = 1$ for all j, gives

$$\dot{p}_j - \dot{p}_0 = q_j^2 - q_0^2.$$

We can integrate this directly:

$$\begin{aligned} p_j(t) - p_0(t) &= \int q_j^2(t) - q_0^2(t)\mathrm{d}t = E_k(t + j\frac{2K}{n} + C) - E_k(t + C) \\ &= Z_k(t + j\frac{2K}{n} + C) - Z_k(t + C) + \frac{2jE}{n} + \frac{CE}{K}, \end{aligned}$$

where E is the complete elliptic integral of the second kind and $Z_k(t) = E_k(t) - \frac{E}{K}t$ is the *Zeta-function*, which is periodic with period $2K$.

Chapter 4

Harmonic Maps

In what follows, we will investigate how the ASD equations on $\mathbb{R}^{2,2}$ reduce, if we require the solutions to be independent of x_3 and x_4. We will first consider the equations on $\mathbb{R}^2$ and show that on $\mathbb{R}^2$ the degeneracy space at such a connection has to be non-trivial. The equations can be written in a conformally invariant form and we view them as the split signature analogue of Hitchin's self-duality equations on a Riemann surface [28], where the moduli space of solutions is known to be a smooth hyperkähler manifold. The main motivation for Hitchin to study the equations in [29] was that solutions satisfying an additional triviality condition give rise to harmonic maps from Riemann surfaces into the structure group G of the bundle on which they are defined. General solutions to the equations correspond to harmonic sections of flat $G \times G$-bundles. Our focus in this chapter will be to study the moduli space of solutions to the full gauge-theoretic equations, and we will only come back to harmonic maps at the end of the chapter when we discuss harmonic tori arising from Schmid's equations.

If we try to carry over the construction of the moduli space from [28] to our situation, we run into difficulties. Due to the change of signature, we are unable to take advantage of Weitzenböck arguments, since the Laplacians in the deformation complex are not necessarily positive operators. In general we will therefore not be able to prove a vanishing theorem analogous to 1.2.18. However, if we look at solutions with zero Higgs field, the signature does not play a role and we are just getting flat unitary connections and in this case the vanishing theorem holds. Using the implicit function theorem, we are therefore able to produce a smooth open set of irreducible solutions with small Higgs field. On this open set we can study the hypersymplectic geometry of the moduli space.

We also study the equations from an alternative viewpoint. Hitchin in [29] observed that the equations can be interpreted as describing geodesics on the moduli space of connections, whose endpoints are flat connections. We then show that every connection has a neighbourhood on which such geodesics are uniquely determined by their endpoints. Thus, we may identify the above open set with a neighbourhood of the diagonal in the product of the moduli space of flat connections with itself. And it turns out that this local product structure is precisely the one corresponding to the endomorphism S defining the hypersymplectic structure. Next, we investigate the complex structures on the moduli space and show that these complex structures correspond to interpreting the moduli space as the moduli space of so-called λ-connections.

If we think of the equations as defining geodesics, we can relate the degeneracy locus to the *cut locus* of the infinite-dimensional Riemannian manifold $\mathcal{A}/\mathcal{G}$ by observing that geodesics whose endpoints are conjugate must be contained in the degneracy locus. We also observe that we have a circle action, which preserves the pseudokähler structure given by I, admitting a proper moment map.

Finally, we produce some explicit formulae for harmonic tori in S^3 by taking Matsoukas' solutions to the $\mathfrak{su}(2)$ Schmid equations and viewing them as translation invariant solutions to the harmonic map equations on $\mathbb{R}^2$.

The main background references for this chapter are [29], [28] and [30].

4.1 Rewriting the Equations on $\mathbb{R}^2$

In order to interpret the equations as a moment map and to be able to define the moduli space of their solutions, we adjust our setting in a similar way as we did when we considered Schmid's equations. Let $(E, h) \to \mathbb{R}^{2,2}$ be a trivial unitary vector bundle of rank n. The configuration space $\mathcal{A}_H$ is given by unitary connections on E whose connection matrices do not depend on x_3 and x_4. We view this again as the subspace of the space $\mathcal{A}$ of all unitary connections on E that are invariant under translations in the x_3 and x_4 direction. Since I, S, T are constant, they preserve $\mathcal{A}_H$ making it into a hypersymplectic manifold with metric

$$g_H(X, Y) = -\int_{\mathbb{R}^2} \sum_{i=1}^{4} g_{ii} \mathrm{tr}(X_i Y_i) \mathrm{d}x_1 \mathrm{d}x_2,$$

where $X = \sum_{i=1}^4 X_i \mathrm{d}x_i$ and $Y = \sum_{i=1}^4 Y_i \mathrm{d}x_i$ are $\mathfrak{u}(E)$-valued one-forms independent of x_3, x_4. The subgroup $\mathcal{G}_H \subset \mathcal{G}$ of gauge transformations, that depend only on x_1, x_2 and converge to the identity as $|x| \to \infty$, acts on $\mathcal{A}_H$. Its Lie algebra is given by elements $\xi \in \Gamma(\mathbb{R}^{2,2}, \mathfrak{u}(E))$ depending on x_1, x_2 only and decaying to zero as $|x| \to \infty$.

The action of $\mathcal{G}_H$ preserves the hypersymplectic structure on $\mathcal{A}_H$, and its moment map μ_H is given by restricting the moment map associated to the action of $\mathcal{G}$ on $\mathcal{A}$ to $\mathcal{A}_H$. Thus, the zeroes of μ_H are the solutions to the ASD equations on $\mathbb{R}^{2,2}$ that are independent of x_3, x_4. We want to study the moduli space

$$\mathcal{M}_H = \mu_H^{-1}(0)/\mathcal{G}_H.$$

To do so, we first use our Lax pair formalism to write down the equations in a conformally invariant manner so that they make in fact sense on any Riemann surface M. Then we look at our degeneracy condition. We use it to give a new interpretation for Hitchin's result, that for any Riemann surface M the hypersymplectic structure on the $\mathcal{M}_H$ is everywhere degenerate in the case $M = \mathbb{R}^2$.

Observe that as we require the solutions $\nabla = \mathrm{d} + \sum_{i=1}^4 A_i \mathrm{d}x_i$ to be independent of x_3 and x_4, the ASD equations become

$$\begin{aligned} [\nabla_1, \nabla_2] &= -[A_3, A_4] \\ [\nabla_1, A_3] &= -[\nabla_2, A_4] \\ [\nabla_1, A_4] &= [\nabla_2, A_3]. \end{aligned}$$

In Lax pair form this reads

$$\mu_\zeta(\nabla) = [\nabla_1 - i\nabla_2 + i\zeta(A_3 - iA_4), A_3 + iA_4 + i\zeta(\nabla_1 + i\nabla_2)] = 0.$$

The coefficients of $1, \zeta, \zeta^2$ all have to be zero and give the following system:

$$\begin{aligned} [\nabla_1 - i\nabla_2, i(A_3 + iA_4)] &= 0 \\ [\nabla_1 - i\nabla_2, \nabla_1 + i\nabla_2] + [A_3 - iA_4, A_3 + iA_4] &= 0 \\ [\nabla_1 + i\nabla_2, i(A_3 - iA_4)] &= 0. \end{aligned}$$

We interpret $\nabla = \mathrm{d} + A_1\mathrm{d}x_1 + A_2\mathrm{d}x_2$ as a connection on $\mathbb{R}^2 \cong \mathbb{C}$ and $\phi_1 = A_3, \phi_2 = A_4$ as two auxiliary fields, which we use to define the $\mathfrak{u}(E) \otimes \mathbb{C}$-valued $(1,0)$-form

$$\Phi = i(\phi_1 - i\phi_2)\mathrm{d}z = \phi \mathrm{d}z,$$

called the *Higgs field*. Here we use the standard coordinate $z = x_1 + ix_2$. Then the last two equations are equivalent to the system

$$R^\nabla = [\Phi \wedge \Phi^*] \tag{4.1}$$

$$\bar{\partial}^\nabla \Phi = 0. \tag{4.2}$$

The first equation becomes $0 = \partial^\nabla \Phi^* = (\bar{\partial}^\nabla \Phi)^*$ since ∇ is unitary and Φ takes values in $\mathfrak{u}(E) \otimes \mathbb{C}$. Thus, the Lax pair equations are in fact equivalent to these two equations. The second equation says that Φ is holomorphic and the first one is a non-linear equation involving the curvature of the connection.

In this form, the equations are conformally invariant and so make sense on any Riemann surface. The Lax equation is then equivalent to (after factoring out a factor i on the right hand term in the bracket)

$$[\partial^\nabla + \zeta\phi, -\phi^* + \zeta\bar{\partial}^\nabla] = 0,$$

or equivalently

$$[\partial^\nabla + \zeta\phi, \bar{\partial}^\nabla - \zeta^{-1}\phi^*] = 0.$$

This says that the connection

$$\nabla^\zeta = \nabla + \zeta\Phi - \zeta^{-1}\Phi^*$$

is flat for all $\zeta \in \mathbb{CP}^1$. If $|\zeta| = 1$, then ∇^ζ is in addition unitary, since then $\zeta = e^{i\theta}$ and so $e^{i\theta}\Phi - e^{-i\theta}\Phi^*$ is skew-adjoint. Thus, choosing $\zeta = \pm 1$, we can associate a pair (∇^+, ∇^-) of flat unitary connections to a solution (∇, Φ) given by

$$\nabla^+ = \nabla + \Phi - \Phi^*, \qquad \nabla^- = \nabla - \Phi + \Phi^*.$$

This pair of flat connections establishes the link to harmonic maps: On $\mathbb{R}^2$ every flat connection is in fact *trivial*. So we obtain two global unitary trivialisations of the bundle E on which they are defined. These two frames are then related by a unitary bundle automorphism

$$u : M \to \mathrm{U}(E) \cong \mathrm{U}(n),$$

where $n = \mathrm{rk}E$. In other words, u is a gauge transformation gauging ∇^+ to ∇^-:

$$u^{-1}\mathrm{d}^{\nabla^+}u = 2(\Phi - \Phi^*),$$

which in the trivialisation defined by ∇^+ reads

$$u^{-1}\mathrm{d}u = 2(\Phi - \Phi^*).$$

The relation $\mathrm{d}^{\nabla^+} * (\Phi - \Phi^*) = 0$ then gives

$$\mathrm{d} * u^{-1}\mathrm{d}u = 0,$$

i.e. u is a harmonic map into $\mathrm{U}(n)$.

4.1.1 Degeneracies

We now show explicitly that, if we consider the equations on $\mathbb{R}^2 \subset \mathbb{R}^{2,2}$ formally as ASD equations on $\mathbb{R}^{2,2}$ as in 2.4.1, then the hypersymplectic structure is degenerate at every point (∇, Φ) of the moduli space. This result has been found by Hitchin in [30]. We now restrict ourselves to $\mathbb{R}^2$ and show that this fits well into our general formalism introduced earlier, relating degeneracies formally to the ultrahyperbolic wave operator associated to a connection on $\mathbb{R}^{2,2}$. We study the equations on $\mathbb{R}^2$ and for a solution (∇, Φ) to the harmonic map equations the operator induced by the ultrahyperbolic wave operator on endomorphism valued sections is given by

$$(\mathrm{d}^\nabla)^*\mathrm{d}^\nabla + (\mathrm{ad}(\phi_1))^2 + (\mathrm{ad}(\phi_2))^2.$$

We will call this operator the split-signature laplacian associated to the solution (∇, ϕ). In real co-ordinates on $\mathbb{R}^2$ our connections read

$$\begin{aligned}\nabla^+ &= \mathrm{d} + (A_1 + \phi_2)\mathrm{d}x_1 + (A_2 - \phi_1)\mathrm{d}x_2\\ \nabla^- &= \mathrm{d} + (A_1 - \phi_2)\mathrm{d}x_1 + (A_2 + \phi_1)\mathrm{d}x_2.\end{aligned}$$

Since these two connections are flat and $\mathbb{R}^2$ is simply connected, they must be *trivial*. Now consider a tangent vector $(\dot{B}, \dot{\psi}) = (B_1\mathrm{d}x_1 + B_2\mathrm{d}x_2, \psi_1\mathrm{d}x_1 + \psi_2\mathrm{d}x_2)$ representing an infinitesimal deformation of the solution $(\nabla, \phi) = (\mathrm{d} + A_1\mathrm{d}x_1 + A_2\mathrm{d}x_2, \phi_2\mathrm{d}x_1 - \phi_1\mathrm{d}x_2)$ on $\mathbb{R}^2$ through solutions. Then the trivial connections ∇^+ and ∇^- on $\mathbb{R}^2$ will remain trivial and hence change by infinitesimal gauge transformations ξ_+, ξ_-, respectively:

$$\begin{aligned}(B_1 + \psi_2)\mathrm{d}x_1 + (B_2 - \psi_1)\mathrm{d}x_2 &= \mathrm{d}^{\nabla^+}\xi_+\\ &= \mathrm{d}^\nabla\xi_+ + [\phi_2\mathrm{d}x_1 - \phi_1\mathrm{d}x_2, \xi_+]\end{aligned}$$

$$\begin{aligned}(B_1 - \psi_2)\mathrm{d}x_1 + (B_2 + \psi_1)\mathrm{d}x_2 &= \mathrm{d}^{\nabla^-}\xi_-\\ &= \mathrm{d}^\nabla\xi_- - [\phi_2\mathrm{d}x_1 - \phi_1\mathrm{d}x_2, \xi_-].\end{aligned}$$

From this, we obtain expressions for the B_i's and ψ_i's:

$$\begin{aligned}2(B_1\mathrm{d}x_1 + B_2\mathrm{d}x_2) &= \mathrm{d}^\nabla(\xi_+ + \xi_-) + [\phi_2\mathrm{d}x_1 - \phi_1\mathrm{d}x_2, \xi_+ - \xi_-]\\ 2(\psi_2\mathrm{d}x_1 - \psi_1\mathrm{d}x_2) &= \mathrm{d}^\nabla(\xi_+ - \xi_-) + [\phi_2\mathrm{d}x_1 - \phi_1\mathrm{d}x_2, \xi_+ + \xi_-].\end{aligned}$$

We read off:

$$\begin{aligned}
2B_1 &= \mathrm{d}_1^\nabla(\xi_+ + \xi_-) + [\phi_2, \xi_+ - \xi_-] = \mathrm{d}_1^\nabla(\xi_+ + \xi_-) + \mathrm{ad}(\phi_2)(\xi_+ - \xi_-)\\
2B_2 &= \mathrm{d}_2^\nabla(\xi_+ + \xi_-) - [\phi_1, \xi_+ - \xi_-] = \mathrm{d}_2^\nabla(\xi_+ + \xi_-) - \mathrm{ad}(\phi_1)(\xi_+ - \xi_-)\\
2\psi_1 &= -\mathrm{d}_2^\nabla(\xi_+ - \xi_-) + [\phi_1, \xi_+ + \xi_-] = -\mathrm{d}_2^\nabla(\xi_+ - \xi_-) + \mathrm{ad}(\phi_1)(\xi_+ + \xi_-)\\
2\psi_2 &= \mathrm{d}_1^\nabla(\xi_+ - \xi_-) + [\phi_2, \xi_+ + \xi_-] = \mathrm{d}_1^\nabla(\xi_+ - \xi_-) + \mathrm{ad}(\phi_2)(\xi_+ + \xi_-).
\end{aligned}$$

Therefore we obtain

$$\begin{aligned}
2\dot{B} &= (\mathrm{d}_1^\nabla(\xi_+ + \xi_-) + \mathrm{ad}(\phi_2)(\xi_+ - \xi_-))\mathrm{d}x_1 + (\mathrm{d}_2^\nabla(\xi_+ + \xi_-) - \mathrm{ad}(\phi_1)(\xi_+ - \xi_-))\mathrm{d}x_2\\
2\dot{\psi} &= (\mathrm{d}_1^\nabla(\xi_+ - \xi_-) + \mathrm{ad}(\phi_2)(\xi_+ + \xi_-))\mathrm{d}x_1 - (-\mathrm{d}_2^\nabla(\xi_+ - \xi_-) + \mathrm{ad}(\phi_1)(\xi_+ + \xi_-))\mathrm{d}x_2\\
&= X^{\xi_+ + \xi_-} + TX^{\xi_+ - \xi_-}.
\end{aligned}$$

Now we claim that $\xi = \xi_+ - \xi_-$ solves

$$\mathcal{D}_1^*\mathcal{D}_1(\xi) := (\mathrm{d}^\nabla)^*\mathrm{d}^\nabla(\xi) + (\mathrm{ad}(\phi_1))^2(\xi) + (\mathrm{ad}(\phi_2))^2(\xi) = 0.$$

Now we use the above equations for the B_i's and ψ_i's to express the derivatives of ξ and get that this equals

$$\begin{aligned}
\mathcal{D}_1^*\mathcal{D}_1(\xi) &= \mathrm{d}_1^\nabla(2\psi_2 - \mathrm{ad}(\phi_2)(\xi_+ + \xi_-)) - \mathrm{d}_2^\nabla(2\psi_1 - \mathrm{ad}(\phi_1)(\xi_+ + \xi_-))\\
&\quad + \mathrm{ad}(\phi_1)(2B_2 - \mathrm{d}_2^\nabla(\xi_+ + \xi_-)) - \mathrm{ad}(\phi_2)(2B_1 - \mathrm{d}_1^\nabla(\xi_+ + \xi_-)),
\end{aligned}$$

which simplifies to

$$2(\mathrm{d}_1^\nabla\psi_2 - \mathrm{d}_2^\nabla\psi_1 + \mathrm{ad}(\phi_1)(B_2) - \mathrm{ad}(\phi_2)(B_1)) + (-[\mathrm{d}_1^\nabla, \phi_2] + [\mathrm{d}_2^\nabla, \phi_1])(\xi_+ + \xi_-).$$

Now we treat the two terms separately. The term with the commutators is equal to zero, since (∇, ϕ) solves the equations. The other term

$$\mathrm{d}_1^\nabla\psi_2 - \mathrm{d}_2^\nabla\psi_1 + \mathrm{ad}(\phi_1)(B_2) - \mathrm{ad}(\phi_2)(B_1),$$

is the linearisation of the equation $\mathrm{d}_1^\nabla\phi_2 - \mathrm{d}_2^\nabla\phi_1 = 0$ applied to $(\dot{B}, \dot{\psi})$, which vanishes since by hypothesis $\dot{B}$ is tangent to $\mu^{-1}(0)$. Hence,

$$X^{\xi_+ - \xi_-} \in T_\nabla\mathcal{O} \cap T_\nabla\mathcal{O}^\perp.$$

And we see that $\xi = \xi_+ - \xi_-$ lies indeed in the kernel of the split-signature laplacian on $\mathbb{R}^2$ associated to the pair (∇, ϕ).

Remark. This is of course equivalent to the statement that if we consider the ASD connection on $\mathbb{R}^{2,2}$ from which (∇, ϕ) is derived, then ξ, viewed as a section on $\mathbb{R}^{2,2}$, solves the associated ultrahyperbolic wave equation.

4.2 The Hypersymplectic Setup for the Equations on a Compact Riemann Surface

From now on, we will work on a hermitian vector bundle E of rank n over a compact Riemann surface M. Since every compact Lie group may be embedded into $\mathrm{U}(n)$ for some n, this is not really a restriction and our proofs should work for arbitrary vector bundles with compact structure group.

Let M be a compact Riemann surface of genus g and let $E \to M$ be a hermitian vector bundle. Let $\mathcal{A}$ be the space of unitary connections on E. On $\mathcal{A}$ we have a natural complex structure I given by the Hodge-star operator acting on one-forms on M. Together with the L^2 inner product we get a Kähler structure on $\mathcal{A}$, just as we have seen in the 4-dimensional case earlier. Explicitly,

$$g(A,B) = -\int_M \mathrm{tr}(A \wedge *B), \qquad IA = *A, \qquad \omega(A,B) = -\int_M \mathrm{tr}(A \wedge B),$$

where in the definition of the Kähler form ω, we used that $*$ is an isometry. We now adopt a more complex geometric view point.

Since there are no $(2,0)$-forms on a Riemann surface, we see that in fact any partial connection defines a holomorphic structure on E. Moreover, any unitary connection is determined by its $(0,1)$-part. This is because for every form $A \in \Omega^1(M, \mathfrak{u}(E))$ we have the basic identity

$$(A^{0,1})^* = -A^{1,0}.$$

Thus, in the following we can equally well think of $\mathcal{A}$ as

$$\mathcal{A} = \{\bar\partial^\nabla : \Gamma(M,E) \to \Omega^{0,1}(M,E)\} = \bar\partial^{\nabla_0} + \Omega^{0,1}(M, \mathfrak{u}(E) \otimes \mathbb{C}).$$

The Higgs-field lives in $\Omega^{1,0}(M, \mathfrak{u}(E) \otimes \mathbb{C})$, and so the space we want to consider is

$$T^*\mathcal{A} = \mathcal{A} \times \Omega^{1,0}(M, \mathfrak{u}(E) \otimes \mathbb{C}).$$

Here we used the map

$$\Lambda : \Phi \mapsto -2i \int_M \mathrm{tr}(\Phi \wedge -)$$

to identify

$$\Omega^{1,0}(M, \mathfrak{u}(E) \otimes \mathbb{C}) \cong (\Omega^{0,1}(M, \mathfrak{u}(E) \otimes \mathbb{C}))^*.$$

Under this identification the complex structure induced by the Hodge-star operator is just multiplication by i. Thus, we get for the complex structure

$$I(A, \Phi) = (iA, i\Phi).$$

Under this interpretation of $T^*\mathcal{A}$ the indefinite metric reads

$$g((A,\Phi),(B,\Psi)) = \mathrm{Re}\left(2i\int_M \mathrm{tr}(A^*\wedge B - \Phi\wedge\Psi^*)\right).$$

Like on any complex cotangent bundle, we also have the canonical holomorphic symplectic form $\omega_I^{\mathbb{C}}$ given by

$$\omega_I^{\mathbb{C}}((A,\Phi),(B,\Psi)) = \Lambda(\Psi)(A) - \Lambda(\Phi)(B) = 2i\int_M \mathrm{tr}(\Phi\wedge B - \Psi\wedge A).$$

This clearly has type $(2,0)$ with respect to the complex structure I. We now define endomorphisms S and T of $T(T^*\mathcal{A})$ by taking the real and imaginary parts of $\omega_I^{\mathbb{C}}$. That is, we write

$$\omega_I^{\mathbb{C}} = g(S-,-) + ig(T-,-).$$

We first note that $\mathrm{Re}(\mathrm{tr}(\Phi\wedge A)) = \mathrm{Re}(\mathrm{tr}(A^*\wedge\Phi^*))$. Thus,

$$\begin{aligned}
\omega_I^{\mathbb{C}}((A,\Phi),(B,\Psi)) &= 2i\int_M \mathrm{tr}(\Phi\wedge B - \Psi\wedge A)\\
&= \mathrm{Re}(2i\int_M \mathrm{tr}(\Phi\wedge B - \Psi\wedge A)) + i\mathrm{Im}(2i\int_M \mathrm{tr}(\Phi\wedge B - \Psi\wedge A))\\
&= g((\Phi^*,A^*),(B,\Psi)) + i\mathrm{Re}(2i\int_M \mathrm{tr}(-i\Phi\wedge B - \Psi\wedge(-iA)))\\
&= g((\Phi^*,A^*),(B,\Psi)) + ig((i\Phi^*,iA^*),(B,\Psi)).
\end{aligned}$$

So we obtain two real structures S and T by

$$S(A,\Phi) = (\Phi^*,A^*) \qquad T(A,\Phi) = (i\Phi^*,iA^*) = IS(A,\Phi).$$

The gauge group acts on $T^*\mathcal{A}$ and the vanishing of the moment map is given by the harmonic map equations, which we can write again as a real and a complex equation, explicitly

$$\begin{aligned}
\mu_I^{\mathbb{C}}(\nabla,\Phi) &= \bar{\partial}^\nabla\Phi = 0,\\
\mu_I(\nabla,\Phi) &= R^\nabla - [\Phi\wedge\Phi^*] = 0.
\end{aligned}$$

We wish to characterise the degeneracy locus.

Proposition 4.2.1. *An element $(\nabla,\Phi)\in T^*\mathcal{A}$ lies in the degeneracy locus if and only if the kernel of $\mathcal{D}_1^\dagger\mathcal{D}_1$ is non-zero. Where $\mathcal{D}_1 : \Gamma(M,\mathfrak{u}(E)) \to \Omega^1(M,\mathfrak{u}(E))\oplus\Omega^1(M,\mathfrak{u}(E))$ is the operator defined by the infinitesimal gauge action and $\mathcal{D}_1^\dagger$ denotes the adjoint with respect to the split signature inner product on $\Omega^1(M,\mathfrak{u}(E))\oplus\Omega^1(M,\mathfrak{u}(E))$. This is a not necessarily positive, elliptic operator, given by*

$$\mathcal{D}_1^\dagger\mathcal{D}_1(\xi) = (\mathrm{d}^\nabla)^*\mathrm{d}^\nabla\xi + *[\phi\wedge *[\phi,\xi]],$$

where $\phi = \Phi - \Phi^*$.

Proof. We want to compute the intersection of the tangent space to a gauge orbit with its orthogonal complement. We work with real co-ordinates, $\phi = \Phi - \Phi^*$. The fundamental vector fields of the gauge action are given by

$$X^{\xi}_{(\nabla,\phi)} = (\mathrm{d}^{\nabla}\xi, [\phi, \xi]) = (\mathrm{d}^{\nabla}\xi, \mathrm{ad}(\phi)(\xi)) = \mathcal{D}_1\xi.$$

We compute the adjoint of $\mathcal{D}_1$ with respect to the neutral inner product defined above. Let $(A, \psi) \in \Omega^1(M, \mathfrak{u}(E)) \oplus \Omega^1(M, \mathfrak{u}(E))$. Then the adjoint is characterised by the property

$$g(\mathcal{D}_1\xi, (A, \psi)) = g(\xi, \mathcal{D}_1^{\dagger}(A, \psi)).$$

The only thing we actually have to compute is the adjoint of $\mathrm{ad}(\phi)(\xi)$ with respect to the ordinary L^2 inner product. Let $A \in \Omega^1(M, \mathfrak{u}(E))$.

$$\begin{aligned} g_{L^2}(\mathrm{ad}(\phi)(\xi), A) &= -\int_M \mathrm{tr}([\phi, \xi] \wedge *A) \\ &= -\int_M \mathrm{tr}((\phi\xi - \xi\phi) \wedge *A) \\ &= -\int_M \mathrm{tr}((\xi(- * A \wedge \phi - \phi \wedge *A)) \\ &= -\int_M \mathrm{tr}(\xi * (- * [\phi \wedge *A])) \\ &= g_{L^2}(\xi, - * [\phi \wedge *A]). \end{aligned}$$

So the adjoint is given by $\mathrm{ad}(\phi)^*(A) = - * \mathrm{ad}(\phi)(*A)$. With this we now compute for $\xi, \eta \in \Gamma(M, \mathfrak{u}(E))$:

$$\begin{aligned} g(\mathcal{D}_1\xi, \mathcal{D}_1\eta) &= g_{L^2}(\mathrm{d}^{\nabla}\xi, \mathrm{d}^{\nabla}\eta) - g_{L^2}(\mathrm{ad}(\phi)(\xi), \mathrm{ad}(\phi)(\eta)) \\ &= g_{L^2}((\mathrm{d}^{\nabla})^*\mathrm{d}^{\nabla}\xi, \eta) - g_{L^2}((\mathrm{ad}(\phi))^*\mathrm{ad}(\phi)(\xi), \eta) \\ &= g_{L^2}((\mathrm{d}^{\nabla})^*\mathrm{d}^{\nabla}\xi + *[\phi \wedge *[\phi, \xi]], \eta). \end{aligned}$$

Thus, we conclude that $\mathcal{D}_1\xi$ lies in the orthogonal complement of the tangent space to the gauge orbit through (∇, ϕ), i.e. in the kernel of $\mathcal{D}_1^{\dagger}$ if and only if

$$(\mathrm{d}^{\nabla})^*\mathrm{d}^{\nabla}\xi + *[\phi \wedge *[\phi, \xi]] = 0.$$

Since ϕ is skew-adjoint, we get that the operator $*[\phi \wedge *[\phi, -]]$ is self-adjoint with non-positive eigenvalues. Hence, the self-adjoint elliptic operator $(\mathrm{d}^{\nabla})^*\mathrm{d}^{\nabla}\xi + *[\phi \wedge *[\phi, \xi]]$ is in general not positive and might a priori have a non-trivial kernel. □

4.3 Two Applications of the Implicit Function Theorem

In this paragraph we aim to show that the harmonic map equations admit solutions with ∇ close to being flat and with small Higgs field. This will be done using the implicit function theorem in Banach spaces. We now work on a Riemann surface. Thus, we require $k > 1$ and consider the space $\mathcal{T}^*\mathcal{A}_k$ of pairs consisting of a unitary L^2_k-Sobolev connection and a Higgs field $\Phi \in \Omega^{1,0}_k(M, \mathfrak{u}(E) \otimes \mathbb{C})$. On this space we have the smooth Hilbert Lie group $\mathcal{G}_{k+1} = L^2_k(M, \mathrm{U}(E))$ acting by

$$u.(\nabla, \Phi) = (\nabla + u^{-1}\mathrm{d}^\nabla u, u^{-1}\Phi u),$$

for $u \in \mathcal{G}_{k+1}$. The fundamental vector fields of this action are given by

$$X^\xi(\nabla, \Phi) = (\mathrm{d}^\nabla \xi, [\Phi, \xi]) = \mathcal{D}_1(\xi).$$

Here

$$\mathcal{D}_1 = \mathrm{d}^\nabla \oplus [\Phi, -] : \Gamma(M, \mathfrak{u}(E)) \to \Omega^1(M, \mathfrak{u}(E)) \oplus \Omega^{1,0}(M, \mathfrak{u}(E) \otimes \mathbb{C}).$$

We consider $T^*\mathcal{A}$ with its standard *positive definite* L^2 inner product and start by proving a slice theorem analogous to theorem 1.2.12. Note that again the centre of $\mathrm{U}(n)$, i.e. constant gauge transformations of the form $e^{i\theta}\mathrm{id}_E$ are contained in the stabiliser group of any pair (∇, Φ). We say again that a pair $(\nabla, \Phi) \in T^{\mathcal{A}}$ is irreducible if its stabiliser is equal to $Z(\mathrm{U}(n))$ and define the reduced gauge group $\mathcal{G}^* = \mathcal{G}/Z(\mathrm{U}(n))$. The Lie algebra of $\mathcal{G}^*$ is given by $\Gamma(M, \mathfrak{u}(E))/i\mathbb{R}\mathrm{id}_E$.

Proposition 4.3.1. *Let $(\nabla, \Phi) \in T^*\mathcal{A}_k$ be irreducible. Then there exists a constant $\epsilon(\nabla, \Phi) > 0$, such that if $(\nabla + A, \Phi + \Psi) \in T^*\mathcal{A}_{L^2_1}$ with $||A||^2_{L^4} + ||\Psi||^2_{L^4} < \epsilon$, there exists a unique gauge transformation $u \in \mathcal{G}^*_{k+1}$ such that*

$$\mathcal{D}_1^*(u.(A, \Psi)) = 0.$$

Proof. We again use the implicit function theorem in Banach spaces. We put on $T^*\mathcal{A}$ the chart $(\tilde{\nabla}, \tilde{\Phi}) \mapsto (\tilde{\nabla} - \nabla, \tilde{\Phi} - \Phi)$. We therefore write $(\tilde{\nabla}, \tilde{\Phi}) = (A, \Psi)$ in this chart and consider the map

$$F : \mathcal{G}^* \times T^*\mathcal{A} \to \mathrm{im}\mathcal{D}_1^* \subset \Gamma(M, \mathfrak{u}(E)),$$

given by

$$F(u, (A, \Psi)) = \mathcal{D}_1^*(u.(\tilde{\nabla}, \tilde{\Phi}) - (\nabla, \Phi)) = \mathcal{D}_1^*(u^{-1}Au + u^{-1}\mathrm{d}^\nabla u, u^{-1}(\Phi + \Psi)u - \Phi).$$

This extends to a smooth map between the respective Sobolev completions of the above spaces. We are interested in the partial derivative of F in the $\mathcal{G}$-direction, and we want to show it is an isomorphism. Observe that trivially $F(\mathrm{id}, 0) = 0$. We now compute the partial derivative of F at this point:

$$\begin{aligned} D_1 F_{(\mathrm{id},0)}(\xi) &= \frac{\mathrm{d}}{\mathrm{d}t}|_{t=0} F(\exp(t\xi), 0) \\ &= \frac{\mathrm{d}}{\mathrm{d}t}|_{t=0} \mathcal{D}_1^*(\exp(-t\xi)\mathrm{d}^\nabla \exp(t\xi), \exp(-t\xi)\Phi\exp(t\xi) - \Phi) \\ &= ((\mathrm{d}^\nabla)^* \mathrm{d}^\nabla + \mathrm{ad}(\Phi)^* \mathrm{d}^\nabla)(\xi) \\ &= \mathcal{D}_1^*(\mathrm{d}^\nabla \xi, [\Phi, \xi]) \\ &= \mathcal{D}_1^* \mathcal{D}_1 \xi \end{aligned}$$

Note that this is an elliptic operator. The equation

$$\mathcal{D}_1^* \mathcal{D}_1 \xi = \eta$$

has a solution if and only if η is orthogonal to the kernel of the adjoint operator, which is again given by

$$\mathcal{D}_1^* \mathcal{D}_1.$$

By the usual integration by parts argument, the kernel of this operator is just the kernel of $\mathcal{D}_1$ which is zero in $\mathrm{Lie}(\mathcal{G}^*)$, since (∇, Φ) is assumed to be irreducible. Therefore, $D_1 F_{(\mathrm{id},0)}$ is an isomorphism and so the implicit function theorem applies to give a solution u to the equation $F(u, (A, \Psi)) = 0$ provided A and Ψ are sufficiently small in norm.

In order to prove uniqueness, suppose u_1 and u_2 are two gauge transformations such that

$$\mathcal{D}_1^*(u_i.(\nabla + A, \Phi_A)) = 0, \qquad i = 1, 2.$$

Without loss of generality, we may assume that u_1 is the identity, i.e that $(\nabla + A, \Phi_A)$ already is in Coulomb gauge and simply write $u = u_2$. Let us write $u.(\nabla + A, \Phi_A) = (\nabla + B, \Phi_B)$. Put $\Phi_A = \Phi + \Psi_A$ and $\Phi_B = \Phi + \Psi_B$. By assumption, we have

$$||A||_{L^4} + ||\Psi_A||_{L^4} < \epsilon, \qquad ||B||_{L^4} + ||\Psi_B||_{L^4} < \epsilon.$$

Now the equation $u.(\nabla + A, \Psi_A) = (\nabla + B, \Psi_B)$ implies

$$\mathrm{d}^\nabla u = uB - Au \qquad [\Phi, u] = u\Psi_B - \Psi_A u.$$

In other words,

$$\mathcal{D}_1 u = (\mathrm{d}^\nabla u, [\Phi, u]) = (uB - Au, u\Psi_B - \Psi_A u).$$

Apply $\mathcal{D}_1^*$ to this equation. Bearing in mind that $\mathcal{D}_1^*(A, \Psi_A) = (\mathrm{d}^\nabla)^* A + \mathrm{Re}([\Phi^* \wedge \Psi_A]) = 0$ and analogously for (B, Ψ_B), this gives

$$\begin{aligned}
\mathcal{D}_1^* \mathcal{D}_1 u &= \mathcal{D}_1^*(uB - Au, u\Psi_B - \Psi_A u) \\
&= - * \mathrm{d}^\nabla (u(*B) - (*A)u) + \mathrm{Re}([\Phi^* \wedge (u\Psi_B - \Psi_A u)]) \\
&= - * \mathrm{d}^\nabla u \wedge *B - u * \mathrm{d}^\nabla * B + (*\mathrm{d}^\nabla * A)u + *(*A \wedge \mathrm{d}^\nabla u) \\
&\quad + \mathrm{Re}([\Phi^* \wedge (u\Psi_B - \Psi_A u)]) \\
&= - * \mathrm{d}^\nabla u \wedge *B + u(\mathrm{d}^\nabla)^* B - (\mathrm{d}^\nabla)^* A u + *(*A \wedge \mathrm{d}^\nabla u) \\
&\quad + \mathrm{Re}([\Phi^* \wedge u\Psi_B] - [\Phi^* \wedge \Psi_A u]) \\
&= - * \mathrm{d}^\nabla u \wedge *B + u((\mathrm{d}^\nabla)^* B) - ((\mathrm{d}^\nabla)^* A)u + *(*A \wedge \mathrm{d}^\nabla u) \\
&\quad + \mathrm{Re}(u[\Phi^* \wedge \Psi_B] + [\Phi^*, u] \wedge \Psi_B - [\Phi^* \wedge \Psi_A]u + \Psi_A[\Phi^*, u]) \\
&= - * \mathrm{d}^\nabla u \wedge *B + *(*A \wedge \mathrm{d}^\nabla u) + \mathrm{Re}([\Phi^*, u] \wedge \Psi_B + \Psi_A[\Phi^*, u]).
\end{aligned}$$

Now, $\mathcal{D}_1^* \mathcal{D}_1$ is self-adjoint, (strongly) elliptic, having the same symbol as the Laplacian, and positive. Thus, we have a spectral decomposition $u = u_0 + u_1$, with $\mathcal{D}_1 u_0 = 0$, and if we denote by λ the first non-zero eigenvalue of $\mathcal{D}_1^* \mathcal{D}_1$, we get the estimate $||u_1||_{L^2}^2 \leq \lambda^{-1} ||\mathcal{D}_1 u_1||_{L^2}^2$. Hence, taking the inner product of the above equation with u_1 and using the usual Sobolev embedding $L_1^2 \subset L^4$ and Cauchy-Schwarz and Hölder inequalities, we get

$$\begin{aligned}
||\mathcal{D}_1 u_1||_{L^2}^2 &\leq \left[(||B|||_{L^4} + ||A||_{L^4})||\mathrm{d}^\nabla u||_{L^2} + (||\Psi_B|||_{L^4} + ||\Psi_A||_{L^4})||[\Phi, u]||_{L^2}\right] ||u_1||_{L^4} \\
&\leq C\left[(||B|||_{L^4} + ||A||_{L^4})||\mathrm{d}^\nabla u||_{L^2} + (||\Psi_B|||_{L^4} + ||\Psi_A||_{L^4})||[\Phi, u]||_{L^2}\right] ||u_1||_{L_1^2} \\
&\leq C\,[||B|||_{L^4} + ||A||_{L^4} + ||\Psi_B|||_{L^4} + ||\Psi_A||_{L^4}]\, ||\mathcal{D}_1 u||_{L^2} (||u_1||_{L^2}^2 + ||\mathrm{d}^\nabla u_1||_{L^2}^2)^{1/2} \\
&\leq C\,[||B|||_{L^4} + ||A||_{L^4} + ||\Psi_B|||_{L^4} + ||\Psi_A||_{L^4}]\, ||\mathcal{D}_1 u||_{L^2}^2 (1 + \lambda^{-1})^{1/2} \\
&\leq 2C^2 \epsilon (1 + \lambda^{-1})^{1/2} ||\mathcal{D}_1 u_1||_{L^2}^2.
\end{aligned}$$

Therefore, we choose $\epsilon < (2C^2(1+\lambda^{-1})^{1/2})^{-1}$ to conclude that $u_1 = 0$ and so $\mathcal{D}_1 u = 0$. Now since we assumed that (∇, Φ) is irreducible, this implies that u lies in $Z(\mathrm{U}(n))$ and hence is equal to the identity in $\mathcal{G}^*$. □

Now we write the moduli space of solutions to the harmonic map equations as the zero locus of a smooth section of a vector bundle on $\mathcal{B}_k^* = \mathcal{T}^* \mathcal{A}_k^* / \mathcal{G}_{k+1}^*$, i.e. the moduli space of irreducible pairs (∇, Φ). Recall that the gauge group acts on $\mathfrak{u}(E)$-valued two-forms by conjugation, i.e. by the adjoint action Ad. Considering the principal $\mathcal{G}^*$-bundle $T^* \mathcal{A}_k^* \to \mathcal{B}_k^*$, we can form the associated vector bundle

$$\mathcal{V} = T^* \mathcal{A}_k^* \times_{\mathrm{Ad}} (\Omega^2(M, \mathfrak{u}(E)) \oplus \Omega^2(M, \mathfrak{u}(E) \otimes \mathbb{C})).$$

Now we interpret the moduli space $\mathcal{M}_H$ as the zero locus of a section G of $\mathcal{V}$. In a local chart, G is defined as follows:

$$\begin{aligned} G : \ker \mathcal{D}_1^* &\rightarrow \Omega^2(M, \mathfrak{u}(E)) \oplus \Omega^2(M, \mathfrak{u}(E) \otimes \mathbb{C}) \\ G(\nabla, \Phi) &= (R^\nabla - [\Phi \wedge \Phi^*], \bar{\partial}^\nabla \Phi). \end{aligned}$$

Note that $G(u.\nabla, u.\Phi) = u^{-1}(G(\nabla, \Phi)u$, i.e. G is equivariant with respect to the actions of $\mathcal{G}^*$ on $T^*\mathcal{A}_k^*$ and $\Omega^2(M, \mathfrak{u}(E)) \oplus \Omega^2(M, \mathfrak{u}(E) \otimes \mathbb{C})$, thus it descends to define a section of $\mathcal{V}$. We compute the derivative of G at a point (∇, Φ). It follows easily from arguments similar to those given in section 1.2 that it is given by

$$\mathrm{d}G_{(\nabla,\Phi)}(A, \psi) = (\mathrm{d}^\nabla A - [\Phi \wedge \psi^*] - [\psi \wedge \Phi^*], \bar{\partial}^\nabla \psi + [A^{0,1} \wedge \Phi]).$$

Proposition 4.3.2. *Let $\nabla \in \mathcal{A}_k^*$ be a flat connection, then the differential of G at $(\nabla, 0)$ is surjective. Moreover, its kernel has dimension $(\dim \mathrm{U}(n))4(g-1) + 4 = 4(n^2(g-1) + 1)$.*

Proof. Since we have zero Higgs field, the derivative $\mathrm{d}G$ simplifies to

$$\mathrm{d}G_{(\nabla,0)}(A, \psi) = (\mathrm{d}^\nabla A, \bar{\partial}^\nabla \psi).$$

Let us denote this operator by $\mathcal{D}_2(A, \psi)$. Furthermore, we have

$$\mathcal{D}_2^*(\alpha, \beta) = ((\mathrm{d}^\nabla)^* \alpha, (\bar{\partial}^\nabla)^* \beta),$$

where $(\alpha, \beta) \in \Omega^2(M, \mathfrak{u}(E)) \oplus \Omega^2(M, \mathfrak{u}(E) \otimes \mathbb{C})$, and

$$\mathcal{D}_1(\xi) = (\mathrm{d}^\nabla \xi, 0) \qquad \xi \in \Gamma(M, \mathfrak{u}(E)).$$

Before we proceed, we make use of some elliptic theory. Since the domain of G is contained in $\ker \mathcal{D}_1^*$, we have $\mathcal{D}_2 = \mathcal{D}_2 + \mathcal{D}_1^*$. So we have to check that the kernel of the adjoint operator $\mathcal{D}_2^* + \mathcal{D}_1$ is zero.

We use the Hodge star operator to identify $\Omega^2(M, \mathfrak{u}(E)) \cong \Omega^0(M, \mathfrak{u}(E))$ and analogously for the complex forms. Under this identification, the operator $(\mathrm{d}^\nabla)^*$ corresponds to d^∇ and the $(\bar{\partial}^\nabla)^*$ corresponds to ∂^∇. Moreover we identify Ω^1 with $\Omega^{0,1}$ in the usual way and thus think of the operator

$$\mathcal{D}_2^* + \mathcal{D}_1 : \Omega^2(M, \mathfrak{u}(E)) \oplus \Omega^2(M, \mathfrak{u}(E) \otimes \mathbb{C}) \oplus \Omega^0(M, \mathfrak{u}(E)) \rightarrow \Omega^1(M, \mathfrak{u}(E)) \oplus \Omega^{1,0}(M, \mathfrak{u}(E)),$$

after putting $\Omega^0(M, \mathfrak{u}(E)) \oplus \Omega^0(M, \mathfrak{u}(E)) \cong \Omega^0(M, \mathfrak{u}(E) \otimes \mathbb{C})$, as the operator

$$\mathcal{D}_2^* + \mathcal{D}_1 : \Omega^0(M, \mathfrak{u}(E) \otimes \mathbb{C}) \oplus \Omega^0(M, \mathfrak{u}(E) \otimes \mathbb{C}) \rightarrow \Omega^1(M, \mathfrak{u}(E)) \oplus \Omega^{1,0}(M, \mathfrak{u}(E)),$$

given by

$$\begin{pmatrix} \mathrm{d}^\nabla & 0 \\ 0 & \partial^\nabla \end{pmatrix}.$$

That is, an element $(\xi, \eta) \in \Omega^0(M, \mathfrak{u}(E)) \oplus \Omega^0(M, \mathfrak{u}(E) \otimes \mathbb{C})$ lies in the kernel if and only if

$$\mathrm{d}^\nabla \xi = 0 \qquad \text{and} \qquad \partial^\nabla \eta = 0.$$

Now by irreducibility of ∇, we immediately conclude that $\xi = 0$. Moreover, η is also parallel as can be seen by the following integration by parts argument:

$$\begin{aligned} 0 &= ||\partial^\nabla \eta||^2_{L^2} \\ &= -2i \int_M \mathrm{tr}(\partial^\nabla \eta \wedge (\partial^\nabla \eta)^*) \\ &= -2i \int_M \mathrm{tr}(\partial^\nabla \eta \wedge -\bar{\partial}^\nabla (\eta^*)) \\ &= -2i \int_M \bar{\partial} \mathrm{tr}((\partial^\nabla \eta)\eta^*) - \mathrm{tr}((\bar{\partial}^\nabla \partial^\nabla \eta)\eta^*) \\ &= -2i \int_M \mathrm{tr}(-(\partial^\nabla \bar{\partial}^\nabla \eta)\eta^*) \qquad \text{(as } \nabla \text{ is flat)} \\ &= -2i \int_M \mathrm{tr}(\bar{\partial}^\nabla \eta \wedge (\bar{\partial}^\nabla \eta)^*) \\ &= ||\bar{\partial}^\nabla \eta||^2_{L^2}. \end{aligned}$$

Thus, since ∇ is assumed to be irreducible, it follows that $\xi \in i\mathbb{R}\mathrm{id}_E$ and $\eta \in \mathbb{C}\mathrm{id}_E$.

The statement about the dimension of the kernel of $\mathrm{d}G$ is obtained from the Atiyah-Singer-Index theorem. The details can be found in [28], section 5. This makes sense, since in the case $\Phi = 0$ both complexes here and in [28] reduce to the same elliptic complex. Note however, that from the above discussion we have $H^0 = 1$ and $H^2 = 3$ in the deformation complex. This is analogous to the situation considered in 1.2.18. □

Corollary 4.3.3. *Let ∇ be an irreducible flat connection, then on a sufficiently small neighbourhood of ∇ in $\mathcal{A}$ there exists a $4(n^2(g-1)+1)$ dimensional family of solutions to the harmonic map equations.*

4.4 A Different Interpretation of the Harmonic Map Equations in Terms of Geodesics

Let $\mathcal{A}$ be the space of unitary connections on the hermitian vector bundle E over the Riemann surface M. We can put a natural Riemannian metric on $\mathcal{A}$ by the L^2 inner product on $\mathfrak{u}(E)$-valued one-forms. The complex structure is given by the Hodge star operator. As seen in earlier sections, we have the gauge group $\mathcal{G}$ acting on $\mathcal{A}$ preserving this structure with fundamental vector fields

$$X^{\xi}_{\nabla} = \mathrm{d}^{\nabla}\xi.$$

According to the slice theorem 1.2.12, the tangent space to the quotient $\mathcal{A}/\mathcal{G}$ can be identified with the orthogonal complement of the space of fundamental vector fields.

$$T_{\nabla}\mathcal{A}/\mathcal{G} = \ker(\mathrm{d}^{\nabla})^* = \{A \in \Omega^1(M, \mathfrak{u}(E) \mid \mathrm{d}^{\nabla} * A = 0\},$$

that is, the space of horizontal one-forms. We have seen that to each solution to the harmonic map equations we can naturally associate a pair of flat connections (∇^+, ∇^-) such that the Higgs field Φ can be characterised in terms of their difference:

$$\nabla^+ - \nabla^- = 2(\Phi - \Phi^*).$$

The equations imply that $\phi = 2(\Phi - \Phi^*)$ is horizontal with respect to ∇:

$$\mathrm{d}^{\nabla} * \phi = 0.$$

But moreover

$$\mathrm{d}^{\nabla^{\pm}} * \phi = \mathrm{d}^{\nabla} * \phi \pm [\phi \wedge *\phi] = 0 \pm [\phi \wedge *\phi] = \pm(\phi \wedge *\phi + (*\phi) \wedge \phi) = 0.$$

With these observations we can give a natural geometric interpretation to the harmonic map equations. They can be viewed as the equation for a geodesic on the moduli space of unitary connections whose endpoints lie on the moduli space of flat unitary connections. Let us be more precise. The space $\mathcal{A}$ of unitary connections is an affine space modelled on $\Omega^1(M, \mathfrak{u}(E))$ with a flat metric given by the L^2 inner product. Therefore, geodesics starting at some connection $\nabla \in \mathcal{A}$ are just straight lines:

$$\gamma(t) = \nabla + t\alpha \qquad \alpha \in \Omega^1(M, \mathfrak{u}(E)).$$

Now by construction the canonical projection

$$\pi : \mathcal{A}^* \to \mathcal{A}^*/\mathcal{G}^*$$

is a Riemannian submersion. Here we denote again by $\mathcal{A}^*$ the open subset of irreducible connections and by $\mathcal{G}^*$ the reduced gauge group. In particular, geodesics on $\mathcal{A}^*/\mathcal{G}^*$ can (at least locally) be lifted to horizontal geodesics on $\mathcal{A}^*$. Under the identification of $T_\nabla \mathcal{A}^*/\mathcal{G}^*$ with the kernel of $(\mathrm{d}^\nabla)^*$, horizontal geodesics are given by straight lines as above such that

$$(\mathrm{d}^\nabla)^*\alpha = 0 \iff \mathrm{d}^\nabla * \alpha = 0.$$

So in fact our computation above shows that the geodesic

$$\gamma(t) = \nabla^+ - 2t(\Phi - \Phi^*)$$

is horizontal for all t and connects ∇^+ and ∇^-, thus establishing our correspondence.

4.4.1 The Space of Geodesics on $\mathcal{A}$

In this section, we consider the space of geodesics on $\mathcal{A}$ more closely. We choose a reference connection ∇_0, so that $\mathcal{A} = \nabla_0 + \Omega^1(M, \mathfrak{u}(E)) \cong \Omega^1(M, \mathfrak{u}(E))$. Then we can identify the space of all oriented geodesics, i.e. directed straight lines, in $\mathcal{A}$ with the tangent bundle to the unit sphere in $\Omega^1(M, \mathfrak{u}(E))$:

$$\mathrm{G}(\mathcal{A}) = TS(\Omega^1(M, \mathfrak{u}(E))).$$

In other words, since geodesics are straight lines, we can specify a geodesic by choosing a starting point $\nabla = \nabla_0 + A$ and a unit direction determined by ϕ with $g_{L^2}(A, \phi) = 0$. Then the geodesic is given by

$$\gamma(t) = \nabla_0 + A + t\phi.$$

We are looking for horizontal geodesics linking flat connections. We start with an observation.

Lemma 4.4.1. *Let $\gamma(t) = \nabla_0 + A + t\phi$ be a straight line on $\mathcal{A}$ containing three flat connections. Then the curvature of γ is zero for all t.*

Proof. Without loss of generality, we may assume that the flat points are given by $\gamma(0)$, $\gamma(1)$ and $\gamma(t_0)$ for some $t_0 \neq 0, 1$. We write $\nabla = \nabla_0 + A$. Then we have the following identities:

$$\begin{aligned} R^\nabla &= 0, \\ R^{\nabla+\phi} &= R^\nabla + \mathrm{d}^\nabla\phi + \frac{1}{2}[\phi \wedge \phi] = 0, \\ R^{\nabla+t_0\phi} &= R^\nabla + t_0\mathrm{d}^\nabla\phi + \frac{t_0^2}{2}[\phi \wedge \phi] = 0. \end{aligned}$$

This gives

$$\begin{aligned} \mathrm{d}^\nabla\phi + \frac{1}{2}[\phi\wedge\phi] &= 0, \\ \mathrm{d}^\nabla\phi + \frac{t_0}{2}[\phi\wedge\phi] &= 0. \end{aligned}$$

And therefore

$$(1-t_0)[\phi\wedge\phi] = 0.$$

That is,

$$[\phi\wedge\phi] = 0.$$

It follows that

$$\mathrm{d}^\nabla\phi = 0,$$

and so for all t the curvature of $\nabla + t\phi$ vanishes. □

We want to find horizontal geodesics which have exactly two flat points lying on them. Therefore, we now change our point of view slightly. We consider the space of oriented geodesic segments of finite length, i.e. straight lines with two endpoints, which we denote by $\mathrm{Geo}(\mathcal{A})$. That is, we look at the space of geodesics with a choice of two points on them. Any such line admits a standard parametrisation by $t \in [0,1]$. We write for the straight line linking ∇ and $\tilde{\nabla}$

$$\nabla + t(\tilde{\nabla} - \nabla).$$

We identify $\mathrm{Geo}(\mathcal{A})$ with $T^*\mathcal{A}$ via

$$(\bar{\partial}^\nabla, \Phi) \in T^*\mathcal{A} \mapsto (\nabla - (\Phi - \Phi^*) + 2t(\Phi - \Phi^*)) \in \mathrm{Geo}(\mathcal{A}).$$

Or in real co-ordinates

$$(\nabla, \phi) \mapsto \nabla - \phi + 2t\phi.$$

That is, we send a pair (∇, ϕ) to the oriented geodesic linking $\nabla - \phi$ and $\nabla + \phi$.

What do Jacobi fields on $\mathcal{A}$ look like? Recall that a Jacobi field is a tangent vector to the space of geodesics, i.e. a curve in $T\mathcal{A}$ obtained by differentiating a 1-parameter family of geodesics $\gamma_s(t)$ with respect to s at $s = 0$. Let $\gamma(t) = \nabla + t\phi$ be the geodesic on $\mathcal{A}$ linking ∇ and $\nabla + \phi$ and let

$$\gamma_s(t) = \nabla(s) + t\phi(s)$$

be a family of geodesics such that $\gamma = \gamma_0$, i.e. a curve through γ in $\mathrm{Geo}(\mathcal{A})$. We write $\nabla(s) = \nabla + A(s)$ where $A(0) = 0$ and $\phi(s) = \phi + \psi(s)$, where $\psi(0) = 0$. Then the corresponding Jacobi field is given by

$$X = \dot{A} + t\dot{\psi},$$

where the dot stands for $\frac{\mathrm{d}}{\mathrm{d}s}|_{s=0}$. Thus, the above isomorphism $T^*\mathcal{A} \cong \mathcal{A}\times\Omega^1(M, \mathfrak{u}(E)) \cong \mathrm{Geo}(\mathcal{A})$ gives that Jacobi fields are of the form

$$X_{(A,\phi)}(t) = A + (2t-1)\phi,$$

where $(A, \phi) \in \Omega^1(M, \mathfrak{u}(E)) \times \Omega^1(M, \mathfrak{u}(E))$. We see that every Jacobi field has two components, one corresponding to perturbing the start point of the geodesic, the other coming from perturbing the direction, i.e. the endpoint. We are now in a position to track through how the endomorphisms I, S, T act on Jacobi fields. In real co-ordinates their action on $\mathcal{A} \times \Omega^1(M, \mathfrak{u}(E)) \cong T^*\mathcal{A}$ is given by

$$I(A, \phi) = (*A, - * \phi), \qquad S(A, \phi) = (-\phi, -A), \qquad T(A, \phi) = (- * \phi, *A).$$

Thus, on a Jacobi field they act as follows

$$\begin{aligned} IX_{(A,\phi)}(t) &= X_{I(A,\phi)}(t) = *(A - (2t-1)\phi), \\ SX_{(A,\phi)}(t) &= X_{S(A,\phi)} = -\phi - (2t-1)A, \\ TX_{(A,\phi)}(t) &= X_{T(A,\phi)} = *(\phi - (2t-1)A). \end{aligned}$$

The L^2 metric on $\mathcal{A} \times \Omega^1(M, \mathfrak{u}(E))$ carries over to a metric on the space of Jacobi fields via

$$g(X, Y) = -\int_M \left[\int_0^1 \mathrm{tr}(X(t) \wedge *Y(t) + \frac{1}{4}\dot{X}(t) \wedge *\dot{Y}(t))\mathrm{d}t\right].$$

The space we are eventually interested in is the space of horizontal geodesics with flat endpoints. This is the subset of $\mathrm{Geo}(\mathcal{A}) = \mathcal{A} \times \Omega^1(M, \mathfrak{u}(E))$ cut out by the equations

$$\begin{aligned} R^\nabla - \mathrm{d}^\nabla\phi + \frac{1}{2}[\phi \wedge \phi] &= 0, \\ R^\nabla + \mathrm{d}^\nabla\phi + \frac{1}{2}[\phi \wedge \phi] &= 0, \\ \mathrm{d}^\nabla * \phi &= 0. \end{aligned}$$

Setting $\phi = \Phi - \Phi^*$, these become the harmonic map equations as we have seen in the last paragraph.

Our aim is to show that a subset in the moduli space of solutions to the harmonic map equations can be identified with a neighbourhood of the diagonal in the product of the space of flat connections with itself. To achieve this, we have to show that any point in $\mathcal{A}/\mathcal{G}$ has a neighbourhood in which any two points can be joined by a unique geodesic lying in this neighbourhood. We divide up the proof into several theorems which are all based on the implicit function theorem. The first step is basically a restatement of the slice theorem 1.2.12.

Theorem 4.4.2. *Let $k > 1$. Let ∇_1 and $\nabla_2 = \nabla_1 + A$ be two connections in $\mathcal{A}_{k-1}$. Then, there exists an $\epsilon > 0$ such that if $||A||_{L^2_{k-1}} < \epsilon$, there exists a unique gauge transformation $u \in \mathcal{G}^*_k$ such that*

$$(\mathrm{d}^{\nabla_1})^*(\nabla_1 - u.\nabla_2) = 0.$$

That is, the straight line

$$\gamma(t) = \nabla_1 + t(u.\nabla_2 - \nabla_1)$$

is a horizontal geodesic between ∇_1 and $u.\nabla_2$, i.e. a geodesic on $\mathcal{A}_{k-1}/\mathcal{G}^*_k$ linking the gauge equivalence classes of ∇_1 and ∇_2.

Proof. Exactly the same as 1.2.12. □

Corollary 4.4.3 (Existence of solutions with small Higgs field). *Let ∇_1 and $\nabla_2 = \nabla_1 + \tilde{\phi}$ be two flat connections in $\mathcal{A}_{k-1}$, such that the norm of $||\tilde{\phi}||_{L^2_{k-1}}$ is small enough for the above theorem to apply. Then defining Φ via*

$$\frac{1}{2}(\Phi - \Phi^*) = u.\nabla_2 - \nabla_1 = u^{-1}\tilde{\phi}u + u^{-1}\mathrm{d}^{\nabla_1}u,$$

where u is the gauge transformation from the theorem, defines a solution (∇, Φ) to the harmonic map equations. Here ∇ is given by $\nabla = \frac{1}{2}(\nabla_1 + u.\nabla_2)$.

Proof. This is clear from the correspondence between solutions to the harmonic map equations and horizontal geodesics discussed earlier. The flat connections ∇_1 and ∇_2 play the role of ∇^- and ∇^+. □

The next corollary rephrases the uniqueness statement of the theorem in terms of the geodesics picture.

Corollary 4.4.4 (Uniqueness). *Every connection $\nabla \in \mathcal{A}^*_{k-1}$ has a neighbourhood consisting of connections that are linked to ∇ by a unique horizontal geodesic.*

Hence, we have shown that any connection $\nabla \in \mathcal{A}^*_{k-1}$ has a neighbourhood on which horizontal geodesics starting at ∇ are uniquely determined by their endpoint. What we really want is to show that ∇ has a neighbourhood, such that *any* horizontal geodesic in this neighbourhood is determined by its endpoints. More precisely, we have the following theorem.

Theorem 4.4.5. *Let $\nabla \in \mathcal{A}^*_{k-1}$ be an irreducible connection and let $\nabla_i = \nabla + A_i$ for $i = 1, 2$ be two connections with ∇_1 in Coulomb gauge relative to ∇, i.e. $(\mathrm{d}^\nabla)^* A_1 = 0$. Then there exists a constant $C > 0$ depending on ∇, but which is independent of A_1 and A_2 such that if $||A_1||_{k-1} < C$ there exists a gauge transformation $u \in \mathcal{G}^*_k$ such that $u.\nabla_2$ is in Coulomb gauge with respect to ∇_1, provided the norm of A_2 is sufficiently small.*

Proof. The proof is similar to the proof of theorem 1.2.12 and uses the implicit function theorem. The equation we want to solve for a small $\xi \in L^2_k(\mathfrak{u}(E))$ is

$$(\mathrm{d}^{\nabla_1})^*(\nabla_1 - \exp\xi.\nabla_2) = 0.$$

In terms of ∇ and the connections matrices A_i this reads

$$(\mathrm{d}^{\nabla_1})^*(A_1 - A_2 - (\exp -\xi)\mathrm{d}^{\nabla_2}\exp\xi) = 0.$$

We view this as a map between Sobolev spaces:

$$F(A_2, \xi) = (\mathrm{d}^{\nabla_1})^*(A_1 - A_2 - (\exp -\xi)\mathrm{d}^{\nabla + A_2}\exp\xi) = 0,$$

where

$$F : \Omega^1_{k-1}(M, \mathfrak{u}(E)) \times L^2_k(M, \mathfrak{u}(E)) \to \mathrm{Im}((\mathrm{d}^{\nabla_1})^*) \subset L^2_{k-2}(M, \mathfrak{u}(E)).$$

By assumption, $F(0, 0) = 0$. We have to show that the partial derivative of F with respect to ξ at $(A_2, \xi) = (0, 0)$ is surjective. A computation analogous to the one in the proof of 1.2.12 shows that this is given by

$$D_2F(\eta) = (\mathrm{d}^{\nabla_1})^*\mathrm{d}^\nabla\eta.$$

We show that this is surjective if the norm of A_1 is sufficiently small. Suppose it is not surjective, then we can find χ which is orthogonal to the image of $(\mathrm{d}^{\nabla_1})^*\mathrm{d}^\nabla$, that is, for all η we have

$$g_{L^2}((\mathrm{d}^{\nabla_1})^*\mathrm{d}^\nabla\eta, \chi) = 0.$$

Put $\eta = \chi$ and compute

$$0 = g_{L^2}((\mathrm{d}^\nabla \chi, \mathrm{d}^{\nabla + A_1}\chi) = ||\mathrm{d}^\nabla \chi||^2_{L^2} + g_{L^2}(\mathrm{d}^\nabla \chi, [A_1, \chi]).$$

Recall that the kernel of d^∇ is zero, since ∇ is irreducible. Hence, it is injective as an operator $L^2_k(\mathfrak{u}(E)) \to \Omega^1_{k-1}(M, \mathfrak{u}(E))$. In particular, since its symbol is injective, it has a bounded left-inverse G, say. Thus, we can write $\chi = G\mathrm{d}^\nabla \chi$ and get the estimate

$$||\chi||_{L^2_1} \leq c||\mathrm{d}^\nabla \chi||_{L^2}.$$

Now we use this to finish the proof. We have seen above that $0 = ||\mathrm{d}^\nabla \chi||^2_{L^2} + g_{L^2}(\mathrm{d}^\nabla \chi, [A_1, \chi])$. Thus,

$$\begin{aligned}
||\mathrm{d}^\nabla \chi||^2_{L^2} &= |g_{L^2}(\mathrm{d}^\nabla \chi, [A_1, \chi])| \\
&\leq ||\mathrm{d}^\nabla \chi||_{L^2} ||[A_1, \chi]||_{L^2} \qquad \text{by Cauchy-Schwartz} \\
&\leq ||\mathrm{d}^\nabla \chi||_{L^2} ||A_1||_{L^4} ||\chi||_{L^4} \qquad \text{by Hölder's inequality} \\
&\leq ||\mathrm{d}^\nabla \chi||_{L^2} ||A_1||_{L^2_1} ||\chi||_{L^2_1} \qquad \text{by Sobolev inequality} \\
&\leq c||\mathrm{d}^\nabla \chi||^2_{L^2} ||A_1||_{L^2_1} \qquad \text{by the above estimate.}
\end{aligned}$$

Hence, if $||A_1||_{L^2_1} < C := \frac{1}{c}$, this gives $||\mathrm{d}^\nabla \chi||^2_{L^2} < ||\mathrm{d}^\nabla \chi||^2_{L^2}$, and thus we conclude

$$||\mathrm{d}^\nabla \chi||_{L^2} = 0,$$

and so by irreducibility

$$\chi = 0.$$

That is, the differential $D_2 F = (\mathrm{d}^{\nabla_1})^* \mathrm{d}^\nabla$ is surjective. Thus, if the norm of A_2 is sufficiently small, say less than some $\epsilon_2(\nabla_1) > 0$, the implicit function theorem guarantees the existence of a small gauge transformation which puts ∇_2 into Coulomb gauge with respect to ∇_1. Note that the constant C only depends on ∇ and neither on ∇_1 nor ∇_2. □

Corollary 4.4.6. *Every flat L^2_k-connection has a small neighbourhood on which a horizontal geodesic with flat endpoints is uniquely determined by its endpoints. In particular, an open subset of the moduli space of solutions to the harmonic map equations can be identified with a neighbourhood of the diagonal in the product of the moduli space of flat connections with itself.*

Proof. Let ∇ be a flat connection in $\mathcal{A}_k^*$ and let $\epsilon = \frac{1}{2}\min(\epsilon(\nabla), C)$, where $\epsilon(\nabla)$ is the constant from theorem 4.4.2 and C is the constant from theorem 4.4.5. We know from theorem 4.4.5 that we can cover the ball of radius ϵ about ∇, which we denote by $B_\epsilon(\nabla)$ by smaller balls on which horizontal geodesics are determined by their endpoints: For each flat connection $\nabla_1 \in B_\epsilon(\nabla)$ the theorem provides a small ball $B_{\epsilon_2(\nabla_1)}(\nabla_1)$ on which we can solve the Coulomb gauge equation. Let $\mathcal{F}_\epsilon = \{A_1 \in \Omega^1_{k-1} \mid R^{\nabla+A_1} = 0\ ;\ (\mathrm{d}^\nabla)^* A_1 = 0\ ;\ ||A_1||_{k-1} \leq \epsilon\ \}$. That is, we have produced a covering

$$B_\epsilon(\nabla) \cap \mathcal{F}_\epsilon \subset \bigcup_{\nabla_1 \in B_\epsilon(\nabla)\cap\mathcal{F}_\epsilon} B_{\epsilon_2(\nabla_1)}(\nabla_1).$$

Now recall that the moduli space of flat connections is actually a finite-dimensional manifold, and so $\mathcal{F}_\epsilon$ is therefore a compact set. Thus, we can choose a finite subcover. Now put

$$U_\nabla = \bigcap_{i=1}^{N} B_{\epsilon_2(\nabla_1^i)}(\nabla_1^i),$$

where we only allow such connections for which $\nabla \in B_{\epsilon_2(\nabla_1^i)}(\nabla_1^i)$. This is an open neighbourhood of ∇ which after shrinking we may assume to be a ball. This then has the desired properties.

We have seen that geodesics become unique once they are short enough. Thus, the space of horizontal geodesics with flat endpoints locally looks like a product of a small ball in the moduli space of flat connections with itself. This subset is open in the harmonic map moduli space, since we know that near a flat connection the moduli space has dimension $4(n^2(g-1)+1)$, which equals twice the dimension of the moduli space of flat unitary connections. $\square$

A few remarks are in order. Our argument shows that a subset in the moduli space of solutions to the harmonic map equations can be identified with a neighbourhood of the diagonal in the product of the moduli space of flat connections with itself. This is in general an open set in the moduli space of solutions to the harmonic map equations. Actually, we have shown that every two flat connections which lie in a small open ball determine a unique geodesic linking them *in this ball.* There could be many other geodesics linking them. Think for example of the round two sphere, where geodesics are great circles. There, any two points in a hemisphere are linked by a unique segment of a great circle, which lies completely in that hemisphere. However, the other segment of the great circle is also a geodesic. In this example, the hemisphere only sees half of the geodesics.

4.4.2 Conjugate Points and the Degeneracy Locus

Theorem 4.4.7. *Let $\gamma(t)$ be a horizontal geodesic with flat endpoints, such that the connection $\gamma(1/2)$ is irreducible. If the endpoints of γ are conjugate, the hypersymplectic structure on the moduli space of harmonic maps is degenerate at γ.*

The endpoints of γ being conjugate means that there exists a one-parameter family $\gamma(s,t)$ of horizontal geodesics with flat endpoints, such that $\gamma(t) = \gamma(0,t)$ and the endpoints are gauge-equivalent for all s. Equivalently, there exists a Jacobi field along γ which is tangent to the gauge orbits through the endpoints of γ .

Proof. Let $\nabla^- = \gamma(0,0)$ and $\nabla^+ = \gamma(0,1)$. Consider the Jacobi field

$$Y(t) = \frac{\partial}{\partial s}|_{s=0}\gamma(s,t).$$

Then by assumption, $Y(0)$ and $Y(1)$ are tangent to the gauge orbit through $\gamma(0)$ and $\gamma(1)$ respectively. This means, there are Lie algebra elements $\xi^\pm \in \Gamma(M, \mathfrak{u}(E))$ such that

$$Y(0) = \mathrm{d}^{\nabla^-}\xi^- \qquad Y(1) = \mathrm{d}^{\nabla^+}\xi^+.$$

Tracking through our correspondence, we write $\nabla = \frac{1}{2}(\nabla^+ + \nabla^-)$, $\phi = \frac{1}{2}(\nabla^+ - \nabla^-) = \Phi - \Phi^*$. In other words

$$\gamma(s,t) = \nabla(s) + (2t-1)\phi(s),$$

and so $\nabla = \nabla(0)$ and $\phi = \phi(0)$. We write $\nabla(s) = \nabla + A(s)$. In this notation

$$Y(0) = \dot{A} - \dot{\phi} \qquad Y(1) = \dot{A} + \dot{\phi}.$$

This gives

$$\begin{aligned} \dot{A} - \dot{\phi} &= \mathrm{d}^\nabla\xi^- - [\Phi - \Phi^*, \xi^-] \\ \dot{A} + \dot{\phi} &= \mathrm{d}^\nabla\xi^+ + [\Phi - \Phi^*, \xi^+]. \end{aligned}$$

It follows that

$$\begin{aligned} 2\dot{A} &= \mathrm{d}^\nabla(\xi^+ + \xi^-) + [\Phi - \Phi^*, \xi^+ - \xi^-] \\ 2\dot{\phi} &= \mathrm{d}^\nabla(\xi^+ - \xi^-) + [\Phi - \Phi^*, \xi^+ + \xi^-]. \end{aligned}$$

This means, that in complex co-ordinates $2(\dot{A}, \dot{\phi})$ is represented by the point

$$\begin{aligned} (2\dot{A}^{0,1}, 2\dot{\phi}^{1,0}) &= (\bar{\partial}^\nabla(\xi^+ + \xi^- - [\Phi^*, \xi^+ - \xi^-], \partial^\nabla(\xi^+ - \xi^-) + [\Phi, \xi^+ + \xi^-]) \\ &= (\bar{\partial}^\nabla(\xi^+ + \xi^-, [\Phi, \xi^+ + \xi^-]) + (-[\Phi^*, \xi^+ - \xi^-], \partial^\nabla(\xi^+ - \xi^-)) \\ &= X^{\xi^+ + \xi^-} + SX^{\xi^+ - \xi^-}. \end{aligned}$$

As in the proof of the fact that at a genuine harmonic map the hypersymplectic structure is degenerate, we see that $X^{\xi^+-\xi^-}$ defines an element of the degeneracy locus.

This shows that if the endpoints of a geodesic corresponding to a solutions to the harmonic map equations are *conjugate*, then the hypersymplectic structure is degenerate at this point of the moduli space.

□

On our nice open set the converse of this theorem is also true.

Proposition 4.4.8. *The hypersymplectic structure on the open set, on which geodesics are determined by their flat endpoints, is non-degenerate.*

Proof. Let (∇, ϕ) be a solution and suppose $||\mathrm{ad}(\phi)||^2_{L^2} < \lambda_1(\nabla)$, where $\lambda_1(\nabla)$ is the smallest non-zero eigenvalue of $(\mathrm{d}^\nabla)^*\mathrm{d}^\nabla$ acting on sections of $\mathfrak{u}(E)$. Assume further that $\eta \in L^2_2(M, \mathfrak{u}(E))$ is a solution to the elliptic degeneracy equation, i.e.

$$(\mathrm{d}^\nabla)^*\mathrm{d}^\nabla\eta - (\mathrm{ad}(\phi))^*\mathrm{ad}(\phi)(\eta) = 0.$$

In other words,

$$||\mathrm{d}^\nabla\eta||^2_{L^2} = ||[\phi, \eta]||^2_{L^2}.$$

This gives

$$\begin{aligned}\lambda_1(\nabla)||\eta||^2_{L^2} &\leq ||\mathrm{d}^\nabla\eta||^2_{L^2} \\ &= ||[\phi, \eta]||^2_{L^2} \\ &\leq ||\mathrm{ad}(\phi)||^2_{L^2}||\eta||^2_{L^2} \\ &< \lambda_1(\nabla)||\eta||^2_{L^2}.\end{aligned}$$

Therefore, η has to vanish in $\mathrm{Lie}(\mathcal{G}^*)$ and the point (∇, ϕ) does not belong to the degeneracy locus. □

4.5 The Circle Action

The harmonic map equations

$$R^\nabla = [\Phi \wedge \Phi^*] \tag{4.3}$$

$$\bar{\partial}^\nabla\Phi = 0 \tag{4.4}$$

are invariant under the action of the unit circle $U(1)$ given by

$$\Phi \mapsto e^{i\theta}\Phi.$$

This action preserves the split signature metric and is holomorphic with respect to the complex structure I, but does not preserve the two product structures S and T: It actually rotates them taking $S + iT$ to $e^{i\theta}(S + iT)$. We compute its pseudokähler moment map, which we denote by ν.

Proposition 4.5.1. *The moment map ν is given by*

$$\nu(\nabla, \Phi) = \frac{i}{2}||\Phi||^2_{L^2}.$$

Proof. We first compute the fundamental vector fields associated to the action. Let $\xi = i\theta \in \mathfrak{u}(1) = i\mathbb{R}$. Then

$$X^{\xi}_{(\nabla,\Phi)} = \frac{\mathrm{d}}{\mathrm{d}t}|_{t=0}(\nabla, e^{i\theta t}\Phi) = (0, \xi\Phi).$$

Now we compute

$$\omega_I(X^{\xi}, (A, \Psi)) = g(I(0, \xi\Phi), (A, \Psi)) = -g_{L^2}(-i\xi\Phi, \Psi) = ig_{L^2}(\xi\Phi, \Psi) = \mathrm{d}\nu(\xi)(\Psi).$$

□

Lemma 4.5.2. *Let (∇, Φ) be a solution to the harmonic map equations. Then we have the following estimate:*

$$||R^{\nabla}||_{L^2} \leq C||\Phi||^2_{L^2}.$$

Proof. Since Φ satisfies the equation $\bar{\partial}^{\nabla}\Phi = 0$, we have the elliptic estimate

$$||\Phi||_{L^2_1} \leq K(||\bar{\partial}\Phi||_{L^2} + ||\Phi||_{L^2}) = K||\Phi||_{L^2}.$$

Also, we will use the Sobolev embedding theorem which gives a bounded inclusion $L^2_1 \to L^4$. Combining these gives the following chain of estimates:

$$\begin{aligned} ||R^{\nabla}||_{L^2} &= ||[\Phi \wedge \Phi^*]||_{L^2} \\ &\leq 2||\Phi||^2_{L^4} \\ &\leq 2\tilde{C}||\Phi||^2_{L^2_1} \\ &\leq C||\Phi||^2_{L^2}. \end{aligned}$$

□

Corollary 4.5.3. *Let $k > 1$. Let $\mathcal{U}$ be the smooth neighbourhood of the diagonal inside the moduli space of pairs of flat L^2_k-connections, we showed to sit inside the moduli space of solutions to harmonic map equations in the previous section. The function*

$$\tilde{\nu} = 2i\nu : \mathcal{U} \to \mathbb{R}, \qquad (\nabla, \Phi) \mapsto ||\Phi||^2_{L^2}$$

is proper, i.e. preimages of compact sets are compact.

Proof. From the lemma, we see that all connections in the preimage of the interval $[0, a]$ satisfy the uniform curvature bound

$$||R^\nabla||_{L^2} < 2Ca^2.$$

Therefore, the properness follows from Uhlenbeck's compactness theorem (see chapter 2.3 in [23]) together with Rellich's Lemma. That is, every sequence in $\tilde{\nu}^{-1}([0, a])$ has a convergent subsequence. □

Notice that ν associates to a geodesic (∇, Φ) its energy, up to a scalar multiple. So in other words, we see that on the space of horizontal geodesics with flat endpoints the energy functional is *proper*. Thus, the fibres of ν are compact. In other words, the space of geodesics of a fixed length is compact. If (∇, Φ) gives rise to a genuine harmonic map, then the function ν associates to this pair the energy of the harmonic map it corresponds to.

4.6 Product Structures and Complex Structures

4.6.1 Product Structures

When we introduced hypersymplectic manifolds, we observed that there is a circle of product structures given by

$$S_\theta = \cos\theta S - \sin\theta T, \qquad T_\theta = \sin\theta S + \cos\theta T.$$

Or in more compact notation

$$S_\theta + iT_\theta = e^{i\theta}(S + iT).$$

These product structures are integrable, therefore a hypersymplectic manifold is locally a product of submanifolds given by integral submanifolds associated to the distributions given by the ±1-eigenspaces of S_θ (analogously for T_θ). In our situation, the hypersymplectic manifold in question is the cotangent bundle of the space of unitary connections and we get

$$S_\theta = \cos\theta S - \sin\theta T = \begin{pmatrix} 0 & -\cos\theta - \sin\theta * \\ -\cos\theta + \sin\theta * & 0 \end{pmatrix},$$

where $*$ is, as usual, the Hodge star operator acting on one-forms. Since $*$ squares to -1, a suggestive short-hand notation is

$$S_\theta = \begin{pmatrix} 0 & -e^{*\theta} \\ -e^{-*\theta} & 0 \end{pmatrix}.$$

Proposition 4.6.1. *The map*

$$P_\theta : T^*\mathcal{A} \to \mathcal{A} \times \mathcal{A} \qquad (\nabla, \phi) \mapsto (\nabla + e^{\theta *}\phi, \nabla - e^{\theta *}\phi)$$

identifies $(T^*\mathcal{A}, S_\theta)$ *with* $(\mathcal{A} \times \mathcal{A}, \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix})$ *as paracomplex manifolds.*

Proof. Let us write $\mathbf{s}$ for the paracomplex structure on $\mathcal{A} \times \mathcal{A}$. We take the derivative of P_θ and show that $\mathbf{s} \circ \mathrm{d}P_\theta = \mathrm{d}P_\theta \circ S_\theta$. The derivative of P_θ is given by

$$\mathrm{d}P_\theta = \begin{pmatrix} 1 & e^{\theta *} \\ 1 & -e^{\theta *} \end{pmatrix}.$$

Now

$$\mathrm{d}P_\theta \circ S_\theta = \begin{pmatrix} 1 & e^{\theta *} \\ 1 & -e^{\theta *} \end{pmatrix} \begin{pmatrix} 0 & -e^{*\theta} \\ -e^{-*\theta} & 0 \end{pmatrix} = \begin{pmatrix} -1 & -e^{*\theta} \\ 1 & -e^{\theta *} \end{pmatrix}.$$

On the other hand

$$\mathbf{s} \circ \mathrm{d}P_\theta = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & e^{\theta *} \\ 1 & -e^{\theta *} \end{pmatrix} = \begin{pmatrix} -1 & -e^{\theta *} \\ 1 & -e^{\theta *} \end{pmatrix}.$$

The inverse of P_θ is given by

$$P_\theta^{-1}(\nabla_1, \nabla_2) = (\frac{1}{2}(\nabla_1 + \nabla_2), \frac{e^{-\theta *}}{2}(\nabla_1 - \nabla_2)).$$

□

If we suppose that (∇, ϕ) is a solution to the equations, then for $\theta = 0$ this is just the map that associates to a solution the associated pair of flat connections (∇^+, ∇^-). Moreover, this map is gauge equivariant, so descends to a map on the respective moduli spaces. We explain below, that this induces the local product structure on the open set in the harmonic map moduli space we constructed earlier. First, we want to see more explicitly what kind of objects the moduli space parametrises.

In general, the pair of flat connections $\nabla^{\pm}$ will not give rise to a globally defined harmonic map into $\mathrm{U}(n)$, since these connections may have non-trivial holonomy. On a simply connected trivialising open set of M we can then still perform the construction explained earlier and get a harmonic map. But the harmonic map will be defined only locally. We now study its transformation behaviour under changes of trivialisations.

So let u_i be defined on an open set U_i for $i = 1, 2$. Let $s_i^{\pm}$ be some local frames on U_i, parallel with respect to $\nabla^{\pm}$ and let $g_{ij}^{\pm}$ be the corresponfing transition functions of E, defined on $U_i \cap U_j$. Then by definition

$$s_i^{\pm} = s_j^{\pm} g_{ji}^{\pm} \qquad \text{on } U_i \cap U_j.$$

On the other hand,

$$s_i^{+} = s_i^{-} u_i, \qquad i = 1, 2.$$

Then we compute on $U_i \cap U_j$,

$$s_i^{+} = s_i^{-} u_i = s_j^{-} g_{ji}^{-} u_i = s_j^{+} u_j^{-1} g_{ji}^{-} u_i = s_i^{+} g_{ij}^{+} u_j^{-1} g_{ji}^{-} u_i.$$

That is,

$$g_{ij}^{+} u_j^{-1} g_{ji}^{-} u_i = \mathrm{id},$$

and so

$$u_i = g_{ij}^{-} u_j g_{ji}^{+}.$$

Thus, if we take an open cover $\{U_i\}$ of our Riemann surface M, and produce local parallel frames for ∇^{+} and ∇^{-} respectively, we get a section $\{U_i, u_i\}$ of the bundle of groups with typical fibre $\mathrm{U}(n)$ and structure group $\mathrm{U}(n) \times \mathrm{U}(n)$ associated to the action $\mathrm{U}(n) \times \mathrm{U}(n) \to \mathrm{U}(n)$ given by $(a, b).u = a^{-1}ub$. So the map that assigns to a solution its associated pair of flat connections, is in general not injective. On our nice open set, where geodesics are determined by their endpoints, it actually is injective. Moreover, we have seen that if we have a continuous space of harmonic sections on a fixed flat bundle, the endpoints are conjugate and the hypersymplectic structure will be degenerate. Formally, we can rephrase this as follows:

Corollary 4.6.2. *Away from the intersection of the harmonic map moduli space with the cut locus of $\mathcal{A}/\mathcal{G}$, the product structure is a diffeomorphism.*

In this way, we may naturally think of the open set inside the moduli space of solutions to the harmonic map equations where geodesics are determined by their endopints as the paracomplexification of the moduli space of flat connections. This is the split signature analogue of the fact that the Higgs bundle moduli space in complex structure J is the moduli space of flat $G^{\mathbb{C}}$ connections.

This is a manifestation of proposition 2.3.9 in the infinite-dimensional setting of the harmonic map moduli space.

Lemma 4.6.3. *Let (∇, ϕ) be a solution to the harmonic map equations. If $\nabla^+_\theta = \nabla + e^{\theta *}\phi$ or $\nabla^-_\theta = \nabla - e^{\theta *}\phi$ are reducible, then the solution (∇, ϕ) lies in the degeneracy locus.*

Proof. We only prove the lemma in the case $\theta = 0$, the case of general θ is proven analogously. Suppose ∇^+ is reducible. Then there exists a section $\xi \in \Gamma(M, \mathfrak{u}(E))$ such that

$$0 = \mathrm{d}^{\nabla^+}\xi = \mathrm{d}^{\nabla}\xi + [\phi, \xi].$$

Now consider

$$\begin{aligned} 0 &= (\mathrm{d}^{\nabla^-})^* \mathrm{d}^{\nabla^+}\xi \\ &= - * \mathrm{d}^{\nabla^-} * \mathrm{d}^{\nabla^+}\xi \\ &= - * (\mathrm{d}^{\nabla} * \mathrm{d}^{\nabla^+}\xi + [\phi \wedge * \mathrm{d}^{\nabla^+}\xi]) \\ &= - * (\mathrm{d}^{\nabla} * \mathrm{d}^{\nabla}\xi - [*\phi \wedge \mathrm{d}^{\nabla}\xi] - [\phi \wedge * \mathrm{d}^{\nabla}\xi] - [\phi \wedge *[\phi, \xi]]) \\ &= (\mathrm{d}^{\nabla})^* \mathrm{d}^{\nabla}\phi + *[\phi \wedge *[\phi, \xi]], \end{aligned}$$

where in the last line we used the equation $\mathrm{d}^{\nabla}\xi = -[\phi, \xi]$ and the Jacobi identity. □

Using the circle action, we may identify all the product structures anti-commuting with I.

Proposition 4.6.4. *The circle action $(\nabla, \Phi) \mapsto (\nabla, e^{i\alpha}\Phi)$ induces a paraholomorphic diffeomorphism $(T^*\mathcal{A}, S_\theta) \cong (T^*\mathcal{A}, S_{\theta+\alpha})$.*

Proof. Let $\tau : \Omega^1(M, \mathrm{End}(E)) \to \Omega^1(M, \mathrm{End}(E))$ be the transposition map (or more invariantly the anti-linear involution induced by minus the Cartan involution) $\tau(\Phi) = \Phi^*$. The product structure $S_\theta = \cos\theta S - \sin\theta T$ may then be written as

$$S_\theta = \begin{pmatrix} 0 & \cos\theta\tau - i\sin\theta\tau \\ \cos\theta\tau - i\sin\theta\tau & 0 \end{pmatrix} = \begin{pmatrix} 0 & e^{-i\theta}\tau \\ e^{-i\theta}\tau & 0 \end{pmatrix}.$$

For fixed α, the derivative of the map $(\nabla, \Phi) \mapsto (\nabla, e^{i\alpha}\Phi)$ is given by

$$\begin{pmatrix} 1 & 0 \\ 0 & e^{i\alpha} \end{pmatrix}.$$

The proof is finished by a direct calculation:

$$\begin{pmatrix} 1 & 0 \\ 0 & e^{i\alpha} \end{pmatrix} \begin{pmatrix} 0 & e^{i\theta}\tau \\ e^{i\theta}\tau & 0 \end{pmatrix} = \begin{pmatrix} 0 & e^{i\theta+\alpha}\tau \\ e^{i\theta+\alpha}\tau & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & e^{i\alpha} \end{pmatrix}.$$

□

4.6.2 The Complex Structures

Recall that we have a two-sheeted hyperboloid of complex structures on the moduli space, parametrised by $\zeta \in \mathbb{CP}^1 \setminus \{|\zeta| = 1\}$:

$$I_\zeta = \frac{1}{1-|\zeta|^2}\left((1+|\zeta|^2)I + (\zeta+\bar{\zeta})S + i(\zeta-\bar{\zeta})T\right).$$

Note that in this notation $I = I_0$. With respect to the complex structure I_0, we have already seen that the map assigning to a unitary connection and a Higgs field the associated $\bar{\partial}$-operator and the $(1,0)$-component of the Higgs field is biholomorphic. It identifies $(\mathcal{A} \times \Omega^1(M, \mathfrak{u}(E)), I_0)$ with the holomorphic tangent bundle $T^*\mathcal{A}$ of the space of $\bar{\partial}$-operators. But how about the other complex structures?

Definition 4.6.1. Let $\lambda \in \mathbb{C}^*$. A *partial λ-connection* on a hermitian vector bundle (E, h) is a $\mathbb{C}$-linear map

$$\nabla^\lambda : \Gamma(E) \to \Omega^{1,0}(M, E),$$

such that

$$\nabla^\lambda(fs) = \lambda\partial f \otimes s + f\nabla^\lambda s,$$

for all $s \in \Gamma(E)$ and $f \in C^\infty(M)$.

To our knowledge, the definition of a λ-connection is due to Deligne and appeared first in Simpson's work on non-abelian Hodge theory, see for example [54].

We denote by $\mathcal{A}^\lambda$ the set of partial λ-connections on E, which is an affine space modelled on $\Omega^{1,0}(M, \mathrm{End}(E))$. We also observe that if $\lambda = 0$, we may think of a 0-connection as just a $\mathrm{End}(E)$-valued $(1,0)$-form. The complex gauge group acts on $\mathcal{A}^\lambda$ in a natural way by conjugation.

Proposition 4.6.5. *Let $\zeta \in \mathbb{C}$ with $|\zeta| \neq 1$. The map*

$$\begin{array}{rcl} F_\zeta : (T^*\mathcal{A}, I_\zeta) & \to & (\mathcal{A} \times \mathcal{A}^{-i\bar{\zeta}}, i \oplus i) \\ (\bar{\partial}^\nabla, \Phi) & \mapsto & (\bar{\partial}^\nabla - i\bar{\zeta}\Phi^*, -i\bar{\zeta}\partial^\nabla + \Phi) \end{array}$$

is a $\mathcal{G}$-equivariant holomorphic diffeomorphism.

Proof. Let τ be the transposition map introduced earlier. Then we may write $I_\zeta \in \mathrm{End}(\Omega^{1,0}(M, \mathrm{End}(E)) \oplus \Omega^{1,0}(M, \mathrm{End}(E)))$ schematically as the two by two matrix

$$I_\zeta = \frac{1}{1-|\zeta|^2}\begin{pmatrix} (1+|\zeta|^2)i & 2\bar{\zeta}\tau \\ 2\zeta\tau & (1+|\zeta|^2)i \end{pmatrix}.$$

The derivative of F_ζ is given by

$$\mathrm{d}F_\zeta(A,\Phi) = \begin{pmatrix} 1 & -i\bar{\zeta}\tau \\ -i\bar{\zeta}\tau & 1 \end{pmatrix}.$$

Now a direct computation, keeping in mind that τ is conjugate-linear, gives

$$\mathrm{d}F_\zeta \circ I_\zeta = i\mathrm{d}F_\zeta.$$

The equivariance is clear. □

Thinking of 0-connections as Higgs fields, we see that the map F_ζ is a direct generalisation of the map F_0 given in these co-ordinates by the identity.

In analogy to the case of Higgs bundles, it turns out that apart from $\pm I$ all other complex structures are equivalent.

Proposition 4.6.6. *The complex structures I_ζ, where $\zeta \neq 0, \infty$, are all equivalent.*

Proof. Let $\lambda, \zeta \in \mathbb{C}^*$. Then the map

$$\mathcal{A} \times \mathcal{A}^\lambda \to \mathcal{A} \times \mathcal{A}^\zeta,$$

given by the identity on the first factor and multiplication by $\zeta\lambda^{-1}$ on the second factor, gives the desired biholomorphism. □

Kirwan uniqueness

Modifying the proof of theorem 2.7 in [28] , we can show by a rearrangement argument that given two solutions with sufficiently small Higgs fields, which are gauge equivalent by a complex gauge transformation, they already are by unitary ones.

Proposition 4.6.7 (Local Uniqueness)**.** *Let (∇_i, Φ_i) be two solutions to the harmonic map equations defined on a hermitian vector bundle E. Suppose that there exists a complex gauge transformation $u \in L^2_k(M, \mathrm{GL}(E))$ such that*

$$(\bar{\partial}^{\nabla_1}, \Phi_1) = u.(\bar{\partial}^{\nabla_2}, \Phi_2).$$

Then (∇_1, Φ_1) and (∇_2, Φ_2) are gauge equivalent by a unitary gauge transformation, provided $\Phi = -\Phi_1 \otimes 1 + 1 \otimes \Phi_2$ satisfies

$$||\Phi \wedge \Phi^*||^2_{L^2} < \lambda_1(\nabla),$$

where ∇ is the induced connection on $\mathrm{E}^* \otimes E \cong \mathrm{End}(E)$ with ∇_1 acting on E^* and ∇_2 acting on E and $\lambda_1(\nabla)$ denotes the first non-zero eigenvalue of its associated Laplacian acting on sections of End(E).

Proof. It is straight-forward to check that (∇, Φ) satisfies the harmonic map equations. The map u being a complex gauge transformation transforming (∇_2, Φ_2) into (∇_1, Φ_1), means that

$$\bar{\partial}^{\nabla} u = 0,$$

if we view u as a section of $\mathrm{End}(E)$.

Moreover, as Φ_1 acts on E^* via $\Phi(\alpha)(v) = \alpha(\Phi(v))$, it follows that $\Phi u = -u\Phi_1 + \Phi_2 u = 0$, since $u^{-1}\Phi_2 u = \Phi_1$.

Since the Laplacian Δ^{∇} associated to ∇ is elliptic, self-adjoint and positive, $L^2(M, E)$ decomposes into an orthogonal direct sum of its (finite-dimensional) eigenspaces. Decompose $u = u_0 + u_\perp$, with u_0 the orthogonal projection onto $\ker(\Delta^{\nabla}) = \ker(\mathrm{d}^{\nabla})$ and $u_\perp = u - u_0$. Let λ_1 be the smallest non-zero eigenvalue of Δ^{∇}. Now we apply a Weitzenböck argument.

$$\begin{aligned}
||\mathrm{d}^{\nabla} u||_{L^2}^2 &= ||\mathrm{d}^{\nabla} u_\perp||_{L^2}^2 \\
&= ||\partial^{\nabla} u_\perp||_{L^2}^2 \\
&= \int_M \mathrm{tr}(R^{\nabla} u_\perp \wedge * u_\perp) \\
&= \int_M \mathrm{tr}([\Phi \wedge \Phi^*] u_\perp \wedge * u_\perp) \\
&= \int_M \mathrm{tr}(\Phi \wedge \Phi^* u_\perp \wedge * u_\perp) \\
&\leq ||\Phi \wedge \Phi^*||_{L^2}^2 ||u_\perp||_{L^2}^2 \\
&< \lambda_1 ||u_\perp||_{L^2}^2.
\end{aligned}$$

On the other hand, we have that

$$\lambda_1 ||u_\perp||_{L^2}^2 \leq g_{L^2}(\Delta^{\nabla} u_\perp, u_\perp) = ||\mathrm{d}^{\nabla} u_\perp||_{L^2}^2 < \lambda_1 ||u_\perp||_{L^2}^2.$$

So we conclude that $u_\perp = 0$ and hence u is parallel with respect to ∇ and moreover

$$0 = ||\mathrm{d}^{\nabla} u||_{L^2}^2 = g_{L^2}(\Phi \wedge \Phi^* u, u) = ||\Phi^* u||^2,$$

so $\Phi^* u = 0$. Now we define the unitary gauge transformation

$$\tilde{u} = u(u^* u)^{-\frac{1}{2}}.$$

Then $\tilde{u}$ is also parallel and hence gauges ∇_2 to ∇_1. Furthermore, since $\Phi u = 0 = \Phi^* u$, it follows that $\Phi u^* = 0$ and hence $\Phi \tilde{u} = 0$, i.e

$$\tilde{u}^{-1} \Phi_2 \tilde{u} = \Phi_1.$$

□

This proposition shows that a small neighbourhood of the moduli space of flat connections in the moduli space of solutions to the harmonic map equations may be identified with an appropriate moduli space of $\lambda-$connections.

4.7 Harmonic Tori from Schmid's Equations

We now come back to the original interpretation of the equations as defining a harmonic map into the structure group G of the vector bundle E on which our connection and Higgs field is defined. We now specialise to the case of M being a torus, then we may view a solution to the harmonic map equations on M as *doubly periodic* solution on $\mathbb{R}^2 \subset \mathbb{R}^{2,2}$. So periodic solutions to Schmid's equations correspond to solutions that are constant along one of the coordinate directions in $\mathbb{R}^2$. In other words, we think of a periodic solution to Schmid's equations as an S^1-invariant solution to the harmonic map equations on a 2-torus.

Thus, we are now going to take the explicit solutions to Schmid's equations obtained earlier and produce the corresponding harmonic maps.

Given a solution $\mathcal{T} = (T_0, T_1, T_2, T_3)$ of Schmid's equations with values in $\mathfrak{g}$, we define a solution of the harmonic map equations on $\mathbb{R}^2$ via

$$\nabla = \mathrm{d} + T_0(x)\mathrm{d}x + T_1(x)\mathrm{d}y \qquad \Phi = i(T_2(x) - iT_3(x))\mathrm{d}z.$$

The pair of flat connections is then given by

$$\nabla^{\pm} = \mathrm{d} + (T_0 \pm T_3)\mathrm{d}x + (T_1 \mp T_2)\mathrm{d}y.$$

To calculate the associated harmonic map, we have to trivialise these two connections using doubly periodic parallel frames s^+ and s^-. That is, we want to find $s^{\pm} : \mathbb{R}^2 \to G$ such that

$$\nabla^{\pm} s^{\pm} = \left(\frac{\partial s^{\pm}}{\partial x} + (T_0 \pm T_3)s^{\pm} \right) \mathrm{d}x + \left(\frac{\partial s^{\pm}}{\partial y} + (T_1 \mp T_2)s^{\pm} \right) \mathrm{d}y = 0.$$

The harmonic map is then given by the automorphism $u : \mathbb{R}^2 \to G$ taking one trivialisation to the other:

$$u = (s^-)^{-1} s^+.$$

Without loss of generality, we may assume that $T_0 = 0$. Thus, we want to solve the following system of equations

$$\frac{\partial s^+}{\partial x} + T_3(x)s^+ = 0, \qquad \frac{\partial s^-}{\partial x} - T_3(x)s^- = 0,$$
$$\frac{\partial s^+}{\partial y} + (T_1(x) - T_2(x))s^+ = 0, \qquad \frac{\partial s^-}{\partial y} + (T_1(x) + T_2(x))s^- = 0.$$

Now note that the second line of equations is trivially solvable since the T_i do not depend on y:

$$s^+(x,y) = \exp(-(T_1 - T_2)y)A^+(x) \qquad s^-(x,y) = \exp(-(T_1 + T_2)y)A^-(x),$$

for some constants of integration $A^\pm : \mathbb{R} \to G$. The first line of equations together with Schmid's equations then gives ordinary differential equations for A^+ and A^-.

Harmonic Tori in SU(2)

We now carry out this procedure for $G = \mathrm{SU}(2)$, in which case Schmid's equations were solved explicitly by Matsoukas in [48]. Consider the standard generators of $\mathfrak{su}(2)$:

$$\sigma_1 = \frac{1}{2}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_2 = \frac{1}{2}\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} \qquad \sigma_3 = \frac{1}{2}\begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}.$$

Then we have seen before, that the non-constant solutions to Schmid's equations are essentially, i.e. up to the action of SU(2) and SO(1,2), given by

$$\mathcal{T} = (0, T_1, T_2, T_3), \qquad T_i = f_i\sigma_i, \quad i = 1, 2, 3,$$

with f_i given in terms of Jacobi elliptic functions satisfying

$$\dot{f}_1 = -f_2 f_3 \qquad \dot{f}_2 = f_3 f_1 \qquad \dot{f}_3 = f_1 f_2.$$

Our first observation is that the harmonic maps we are producing are going to be maps to totally geodesic 2-spheres in S^3, i.e. can be interpreted as Gauss maps of constant mean curvature surfaces. This relies on the following lemma proved by Hitchin.

Lemma 4.7.1 ([29], Prop. 1.9)**.** *A harmonic map $u : M \to \mathrm{SU}(2) \cong S^3$ with associated pair (∇, Φ) maps onto a totally geodesic 2-sphere if and only if there exists a gauge transformation g such that $g^2 = -1$ and $\Phi g = -g\Phi$.*

Corollary 4.7.2. *Harmonic maps obtained from Schmid's equations map tori onto totally geodesic 2-spheres in* SU(2)*.*

Proof. Since in this case $\Phi = (T_3 - iT_2)\mathrm{d}z$, this is just a reformulation of the fact observed in 3.6.2 that such a solution to Schmid's equations has C_2-symmetry. □

We make an ansatz for s^+ in the form

$$s^+(x,y) = \begin{pmatrix} a(x,y) & \bar{b}(x,y) \\ -b(x,y) & \bar{a}(x,y) \end{pmatrix}, \qquad \text{with } a, b : \mathbb{R}^2 \to \mathbb{C} \text{ and } |a(x,y)|^2 + |b(x,y)|^2 = 1.$$

Now $T_1(x) - T_2(x)$ is given explicitly by

$$T_1 - T_2 = f_1\sigma_1 - f_2\sigma_2 = \frac{1}{2}\begin{pmatrix} 0 & -f_1 - if_2 \\ f_1 - if_2 & 0 \end{pmatrix}.$$

The third matrix T_3 then reads

$$T_3 = \frac{1}{2}\begin{pmatrix} -if_3 & 0 \\ 0 & if_3 \end{pmatrix}.$$

We write a_x, a_y etc. to indicate partial differentiation with respect to x, y respectively. Then the equations we want to solve are

$$\begin{aligned} \frac{\partial s^+}{\partial x} + T_3(x)s^+ = 0 \quad &\Longrightarrow \quad a_x = \frac{1}{2}if_3 a \\ & \quad b_x = -\frac{1}{2}if_3 b \\ \frac{\partial s^+}{\partial y} + (T_1(x) - T_2(x))s^+ = 0 \quad &\Longrightarrow \quad a_y = -\frac{1}{2}(f_1 + if_2)b \\ & \quad b_y = \frac{1}{2}(f_1 - if_2)a. \end{aligned}$$

We first look at the derivatives with respect to y. Differentiating both equations once more with respect to y gives

$$a_{yy} = -\frac{1}{4}(f_1^2 + f_2^2)a \qquad b_{yy} = -\frac{1}{4}(f_1^2 + f_2^2)b,$$

At this point, we insert the explicit formulae for the f_i,

$$\begin{aligned} f_1(x) &= kD\mathrm{sn}_k(Dx + C), \\ f_2(x) &= kD\mathrm{cn}_k(Dx + C), \\ f_3(x) &= -D\mathrm{dn}_k(Dx + C). \end{aligned}$$

The relation

$$\mathrm{sn}_k^2 + \mathrm{cn}_k^2 = 1$$

implies that

$$f_1^2 + f_2^2 = k^2D^2.$$

So we arrive at

$$a_{yy} = -\frac{1}{4}k^2D^2 a \qquad b_{yy} = -\frac{1}{4}k^2D^2 b.$$

The general solution to this system is given by

$$\begin{aligned} a(x,y) &= A_1(x)\sin\left(\frac{kD}{2}y\right)+A_2(x)\cos\left(\frac{kD}{2}y\right), \\ b(x,y) &= B_1(x)\sin\left(\frac{kD}{2}y\right)+B_2(x)\cos\left(\frac{kD}{2}y\right). \end{aligned}$$

Now the equation

$$a_y = -\frac{1}{2}(f_1+if_2)b$$

implies

$$\begin{aligned} A_1(x) &= -(\mathrm{sn}_k(Dx+C)+i\mathrm{cn}_k(Dx+C))B_2(x), \\ A_2(x) &= (\mathrm{sn}_k(Dx+C)+i\mathrm{cn}_k(Dx+C))B_1(x). \end{aligned}$$

We introduce the *Jacobi amplitude function* $\varphi_k(x)$ which satisfies

$$e^{i\varphi_k(x)} = \mathrm{cn}_k(x)+i\mathrm{sn}_k(x), \qquad \varphi'(x) = \mathrm{dn}_k(x).$$

With this we can write

$$B_1(x) = -ie^{i\varphi_k(Dx+C)}A_2(x), \qquad B_2(x) = ie^{i\varphi_k(Dx+C)}A_1(x).$$

Our general solution becomes

$$\begin{aligned} a(x,y) &= A_1(x)\sin\left(\frac{kD}{2}y\right)+A_2(x)\cos\left(\frac{kD}{2}y\right), \\ b(x,y) &= -ie^{i\varphi_k(Dx+C)}A_2(x)\sin\left(\frac{kD}{2}y\right)+ie^{i\varphi_k(Dx+C)}A_1(x)\cos\left(\frac{kD}{2}y\right). \end{aligned}$$

In order to determine the A_i's, we look at the other equation:

$$a_x(x,y) = \frac{1}{2}if_3(x)a(x,y),$$

which implies

$$A_1'(x) = -\frac{i}{2}D\mathrm{dn}_k(Dx+C)A_1(x), \qquad A_2'(x) = -\frac{i}{2}D\mathrm{dn}_k(Dx+C)A_2(x),$$

whose solution is given by

$$A_1(x) = \alpha_1 e^{-\frac{i}{2}\varphi_k(Dx+C)}, \qquad A_2(x) = \alpha_2 e^{-\frac{i}{2}\varphi_k(Dx+C)},$$

with constants $\alpha_i \in \mathbb{C}$, $i=1,2$ to be determined by initial conditions. Thus, we get the general solution

$$\begin{aligned} a(x,y) &= \left[\alpha_1\sin\left(\frac{kD}{2}y\right)+\alpha_2\cos\left(\frac{kD}{2}y\right)\right]e^{-\frac{i}{2}\varphi_k(Dx+C)}, \\ b(x,y) &= \left[-\alpha_2\sin\left(\frac{kD}{2}y\right)+\alpha_1\cos\left(\frac{kD}{2}y\right)\right]ie^{\frac{i}{2}\varphi_k(Dx+C)}, \end{aligned}$$

where in order to ensure that the determinant of s^+ is equal to 1, we have to require that $|\alpha_1|^2 + |\alpha_2|^2 = 1$.

Similarly, the equations

$$\frac{\partial s^-}{\partial x} - T_3(x)s^- = 0, \qquad \frac{\partial s^-}{\partial y} + (T_1(x) + T_2(x))s^- = 0,$$

with s^- in the form

$$s^-(x,y) = \begin{pmatrix} c(x,y) & \bar{d}(x,y) \\ -d(x,y) & \bar{c}(x,y) \end{pmatrix},$$

where $c, d : \mathbb{R}^2 \to \mathbb{C}$ and $|c(x,y)|^2 + |d(x,y)|^2 = 1$, yield the system

$$\begin{aligned} \frac{\partial s^-}{\partial x} + T_3(x)s^- = 0 \quad &\Longrightarrow \quad c_x = -\frac{1}{2} i f_3 c, \\ & \qquad\;\; d_x = \frac{1}{2} i f_3 d, \\ \frac{\partial s^-}{\partial y} + (T_1(x) - T_2(x))s^- = 0 \quad &\Longrightarrow \quad c_y = -\frac{1}{2}(f_1 - i f_2) d, \\ & \qquad\;\; d_y = \frac{1}{2}(f_1 + i f_2) c. \end{aligned}$$

Solving this in a similar fashion as above, we get the general solution

$$\begin{aligned} c(x,y) &= \left[\gamma_1 \sin\left(\frac{kD}{2} y\right) + \gamma_2 \cos\left(\frac{kD}{2} y\right)\right] e^{\frac{i}{2}\varphi_k(Dx+C)}, \\ d(x,y) &= \left[\gamma_2 \sin\left(\frac{kD}{2} y\right) - \gamma_1 \cos\left(\frac{kD}{2} y\right)\right] i e^{-\frac{i}{2}\varphi_k(Dx+C)}, \end{aligned}$$

where again we have to require $|\gamma_1|^2 + |\gamma_2|^2 = 1$ as before. To arrive at our harmonic map, we now just have to perform a matrix multiplication.

$$u(x,y) = (s^-)^{-1}s^+ = (s^-)^* s^+ = \begin{pmatrix} \bar{c}a + \bar{d}b & \bar{c}\bar{b} - \bar{d}\bar{a} \\ da - cb & d\bar{b} + c\bar{a} \end{pmatrix} = \begin{pmatrix} u_1 & \bar{u}_2 \\ -u_2 & \bar{u}_1 \end{pmatrix},$$

where, after using some trigonometric identities and the relation $e^{i\varphi_k(x)} = \mathrm{cn}_k(x) + i\mathrm{sn}_k(x)$, we may write

$$\begin{aligned} u_1(x,y) &= (\cos(kDy)(\bar{\gamma}_2\alpha_2 - \bar{\gamma}_1\alpha_1) + \sin(kDy)(\bar{\gamma}_2\alpha_1 + \bar{\gamma}_1\alpha_2))\mathrm{cn}_k(Dx+C) \\ &\quad - i(\bar{\gamma}_1\alpha_1 + \bar{\gamma}_2\alpha_2)\mathrm{sn}_k(Dx+C), \\ u_2(x,y) &= (\cos(kDy)(i\alpha_1\gamma_2 + i\alpha_2\gamma_1) + \sin(kDy)(i\alpha_1\gamma_1 - i\alpha_2\gamma_2))\mathrm{cn}_k(Dx+C) \\ &\quad + (\alpha_2\gamma_1 - \alpha_1\gamma_2)\mathrm{sn}_k(Dx+C). \end{aligned}$$

Chapter 5

Split Signature Instantons on $S^2 \times S^2$

We conclude this thesis with a discussion of the full ASD equations on $\mathbb{R}^{2,2}$ subject to the boundary condition that the solutions should extend to $S^2 \times S^2$.

It is known that there exists a hyperkähler structure on the space of framed instantons on Euclidean space $\mathbb{R}^4$. The idea of the proof is essentially the same as in our construction of the ASD moduli space on compact hyperkähler four-manifolds carried out in section 1.2. Since $\mathbb{R}^4$ is not compact, appropriate decay conditions for the connection matrices have to be imposed, see [43] for details. In the presence of a compatible complex structure on $\mathbb{R}^4$, these decay conditions imply that the solution may be extended to S^4. Using the twistor correspondence, such instantons then correspond to stable bundles on $\mathbb{CP}^3$ trivial on a line l_∞, which can be described using monads and the ADHM construction. Choosing a complex structure on $\mathbb{R}^4$, we obtain a natural complex compactifiaction $\mathbb{CP}^2$. Moreover, viewing $\mathbb{CP}^3 \setminus l_\infty$ as the space of compatible complex structures on $\mathbb{R}^4$, the choice of complex structure on $\mathbb{R}^4$ gives rise to an embedding of $\mathbb{CP}^2$ as a plane through l_∞ in $\mathbb{CP}^3$. In this way, the framed instantons on S^4 can be pulled back to holomorphic bundles on $\mathbb{CP}^2$ trivial on a line, which are then given by ADHM data.

The ADHM equations allow for a moment map interpretation, as so the ASD moduli space can be constructed as a hyperkähler quotient of a certain finite-dimensional space of matrices by an appropriate unitary group action. We wish to carry out a similar program in the case of ASD connections on $\mathbb{R}^{2,2}$, respectively $S^2 \times S^2$. The main background references for this chapter are [45], [16], [20], [3].

5.1 ASD Connections on $\mathbb{R}^{2,2}$

In this chapter, we would like to study the ASD equations on $\mathbb{R}^{2,2}$ in the presence of a complex structure, which we then use to view $\mathbb{R}^{2,2}$ as an open set in $\mathbb{CP}^1 \times \mathbb{CP}^1$. A natural boundary condition is to require that solutions should extend to this compactification.

We begin with a lemma.

Lemma 5.1.1. *A unitary connection* ∇ *on* $\mathbb{R}^{2,2}$ *is ASD if and only if its curvature is of type* $(1,1)$ *with respect to every* I *in the two-sheeted hyperboloid of compatible complex structures.*

Proof. Let I_ζ be a compatible complex structure. We have seen when we discussed the Lax pair formalism that the ASD equations may be written in the form

$$[\nabla_1 - i\nabla_2 + i\zeta(\nabla_3 - i\nabla_4), \nabla_3 + i\nabla_4 + i\zeta(\nabla_1 + i\nabla_2)] \;=\; 0,$$

and this is precisely the $(0,2)$-part of the curvature of ∇. For if we write the complex co-ordinates with respect to the complex structure I as $z = x_1 - ix_2$, $w = x_3 + ix_4$, then

$$z_\zeta = z + i\zeta\bar{w} \qquad w_\zeta = w + i\zeta\bar{z}$$

are complex co-ordinates with respect to I_ζ. □

In other words, for any choice of compatible complex structure on $\mathbb{R}^{2,2}$, an ASD connection equips the bundle E with a holomorphic structure. This implies that if we pull-back the bundle E to the twistor space Z of $\mathbb{R}^{2,2}$, which as a smooth manifold is given by $Z = \mathbb{R}^{2,2} \times C$, where C is the two-sheeted hyperboloid of compatible complex structures, we obtain a holomorphic vector bundle on Z.

5.2 The Twistor Correspondence for ASD Connections on $S^2 \times S^2$

Locally, the above lemma 5.1.1 also applies to ASD connections on the conformal compactification $\mathbb{BP}^1 \cong \mathrm{Gr}(2, \mathbb{R}^4) \cong (S^2 \times S^2)/\{\pm 1\}$. The identification $\mathrm{Gr}(2, \mathbb{R}^4) \cong (S^2 \times S^2)/\{\pm 1\}$ is given by the Plücker embedding $\mathrm{Gr}(2, \mathbb{R}^4) \to \mathbb{RP}^5$, the image of which is the Klein quadric $\{\omega \in \mathbb{RP}^5 \mid \omega \wedge \omega = 0\}$, which is a real projective quadric of signature $(3,3)$. The space $\mathbb{BP}^1$ is defined to be the space of split-quaternionic lines in $\mathbb{B}^2$, i.e.

$$\mathbb{BP}^1 = \{(p,q) \in \mathbb{B}^2 \mid \nexists \lambda \in \mathbb{B} \setminus \{0\} \text{ such that } \lambda p = 0 \text{ and } \lambda q = 0.\}/\mathbb{B}^*,$$

where we avoided the subspaces spanned by two linearly dependent null-vectors, as these do not correspond to planes in $\mathbb{B} \cong \mathbb{R}^{2,2}$. On $\mathbb{BP}^1$ we have ordinary homogenous co-ordinates $[p:q]$ and $\mathbb{B}$ embeds as the set of points of the form $[p:1]$ for $p \in \mathbb{B}$. Note that unlike in the case of $\mathbb{CP}^1$ or $\mathbb{HP}^1$ we are adding in a whole "null-cone at infinity", given by points of the form $[q:p]$, where p is a null-vector in $\mathbb{B}$. Identifying $\mathbb{B}$ with the set $\mathfrak{gl}(2,\mathbb{R}) = \mathrm{End}(\mathbb{R}^2)$ as before, we therefore obtain a map to $\mathrm{Gr}(2,\mathbb{R}^4)$ given by identifying $[p:q]$ with the image of the linear map $p \oplus q : \mathbb{R}^2 \to \mathbb{R}^2 \oplus \mathbb{R}^2$, where we think of elements in $\mathbb{R}^2$ as row vectors. This map is an isomorphism.

Next we will see that the twistor space Z of $\mathbb{BP}^1$, viewed as the bundle of compatible complex structures, is given by $\mathbb{CP}^3 \setminus \mathbb{RP}^3$. We identify $\mathbb{C}^4$ with $\mathbb{B}^2$ via

$$(z_1, z_2, z_3, z_4) \mapsto (z_1 + z_2 s, z_3 + z_4 s).$$

Under this identification, $\mathbb{C}^4$ inherits a real structure, i.e. an anti-holomorphic involution τ, corresponding to left-multiplication by s on $\mathbb{B}^2$. In co-ordinates on $\mathbb{C}^4$ τ is given by

$$(z_1, z_2, z_3, z_4) \mapsto (\bar{z}_2, \bar{z}_1, \bar{z}_4, \bar{z}_3).$$

Since τ is conjugate-linear, it descends to an involution on $\mathbb{CP}^3$ and we may ask what its fixed points are. A point $[z_1 : z_2 : z_3 : z_4]$ is fixed if and only if

$$[z_1 : z_2 : z_3 : z_4] = [\bar{z}_2 : \bar{z}_1 : \bar{z}_4 : \bar{z}_3],$$

i.e.

$$(z_1, z_2, z_3, z_4) = \lambda(\bar{z}_2, \bar{z}_1, \bar{z}_4, \bar{z}_3), \qquad \text{for some } \lambda \in \mathbb{C}^*.$$

These four conditions imply that $|\lambda| = 1$, and that therefore $|z_1| = |z_2|$ and $|z_3| = |z_4|$. In other words, $z_1 + z_2 s$ and $z_3 + z_4 s$ should be *null-vectors* in $\mathbb{B} \cong \mathbb{C}^{1,1}$, which satisfy moreover $z_i = \lambda \bar{z}_{i-1}$ for $i = 2, 4$. In particular $(1 - \lambda s)(z_i + \lambda \bar{z}_i s) = 0$, for $i = 1, 3$. Thus, we obtain a well-defined surjection

$$\begin{aligned} \pi : \mathbb{CP}^3 \setminus \mathrm{Fix}(\tau) &\to \mathbb{BP}^1 \\ [z_1 : z_2 : z_3 : z_4] &\mapsto [z_1 + z_2 s : z_3 + z_4 s]. \end{aligned}$$

The fibre of π over a point $[p:1] \in \mathbb{BP}^1$ is given by the set of complex lines in $\mathbb{C}^4$ that lie in the copy of $\mathbb{B}$ spanned by the vector (p,q) and that are not fixed by τ. So

$$\begin{aligned} \pi^{-1}([p:1]) &= \{[z_1 : z_2 : z_3 : z_4] \in \mathbb{CP}^3 \mid [z_1 + z_2 s : z_3 + z_4 s] = [p:1]\} \\ &\cong \{(z_3, z_4) \in \mathbb{C}^2 \setminus \{0\} \mid |z_3| \neq |z_4|\}/\mathbb{C}^* \\ &\cong \mathbb{CP}^1 \setminus \{[z_3 : z_4] \mid |z_3| = |z_4| \ \} \\ &\cong (\mathbb{C} \cup \{\infty\}) \setminus \{|\zeta| = |z_3/z_4| = 1\}, \end{aligned}$$

which, as we have seen before, parametrises precisely the two-sheeted hyperboloid of complex structures. Notice that, by construction, τ preserves the fibres of π, we say the fibres of π are *real lines*, and τ corresponds to inversion with respect to the unit circle on $\mathbb{C} \cup \{\infty\}$. In other words, if we write $P = [z_1 : z_2 : z_3 : z_4]$, then the fibre of π over $\pi(P)$ is the line through P and $\tau(P)$ with the fixed points of τ removed. So $\mathbb{BP}^1$ parametrises the real lines in $\mathbb{CP}^3$. This observation also leads to an alternative twistor approach, by noting that the fixed point set of τ corresponds to the circle bundle of compatible product structures. This is a fact we have seen before, when we looked at twistor spaces of hypersymplectic manifolds.

In new co-ordinates $\{w_i\}$ on $\mathbb{C}^4$ given by

$$w_1 = z_1 + z_2 \qquad w_2 = i(z_1 - z_2) \qquad w_3 = z_3 + z_4 \qquad w_4 = i(z_3 - z_4),$$

τ becomes just ordinary complex conjugation and so the fixed point set of τ in $\mathbb{CP}^3$ is given by a copy of $\mathbb{RP}^3$. Hence, our twistor space becomes

$$Z = \mathbb{CP}^3 \setminus \mathbb{RP}^3.$$

Now $\mathbb{BP}^1 \cong \mathrm{Gr}(2, \mathbb{R}^4)$ comes with a natural double cover, given by the space of oriented planes in $\mathbb{R}^4$. If we view $q \in \mathbb{B}^*$ as a linear map from $\mathbb{R}^2$ to $\mathbb{R}^2$, then reversing the orientation on its graph corresponds to multiplying the matrix q by

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

We have interpreted $\mathrm{Gr}(2, \mathbb{R}^4)$ as a quadric inside $\mathbb{RP}^5$ via the Plücker embedding with topology $(S^2 \times S^2)/\{\pm 1\}$ and we see that changing the orientation of a two-plane corresponds to multiplying its associated bivector by -1. Thus, this double cover is topologically equal to $S^2 \times S^2$. With the picture established above, we obtain that the associated twistor space is again given by $\mathbb{CP}^3 \setminus \mathbb{RP}^3$. Since changing the orientation is essentially given by τ, which interchanges the two components of the fibre of π, the space $S^2 \times S^2$ parametrises holomorphic discs in $\mathbb{CP}^3$, the boundary circle of which lies on $\mathrm{Fix}(\tau) \cong \mathbb{RP}^3$.

With this setup, Mason in [45] gave a complete twistor description of instantons on $S^2 \times S^2$ with its split signature metric.

Theorem 5.2.1 ([45]). *There is a $1-1$ correspondence between gauge equivalence classes of smooth unitary ASD-connections on $S^2 \times S^2$ and isomorphism classes of pairs (E, H), where E is a holomorphic vector bundle on $\mathbb{CP}^3$ and H is a smooth positive definite hermitian metric on the restriction of E to the real $\mathbb{RP}^3$.*

If there exists an anti-linear isomorphism $\tilde{\tau}_E : E \to \bar{E}^*$ covering the real structure τ such that H is induced by $\tilde{\tau}_E$ via the formula $H(v, v) = \tilde{\tau}_E(v)(v)$, then the associated ASD-connection is one that is pulled back from $\mathbb{BP}^1$. Such solutions are called split signature instantons.

Note that in contrast to the case of Euclidean instantons this theorem shows in particular that the moduli space of ASD-connections is infinite-dimensional, since the space of hermitian metrics on $E|_{\mathbb{RP}^3}$ is infinite dimensional. So to an ASD-connection we have associated two types of data. On the one hand we have algebraic data, i.e. the holomorphic bundle E and on the other hand there is smooth data, given by the hermitian metric. Split signature instantons thus correspond to solutions purely given by algebraic data, since in this case the hermitian metric is determined by the real structure.

In the papers of Donaldson and Maciocia [20], [43], it is shown that the moduli space of framed instantons on $\mathbb{R}^4$, which as explained earlier may be viewed as the moduli space of stable holomorphic bundles on $\mathbb{CP}^2$ with a trivialisation on the line at infinity, carries a natural hyperkähler metric. Using the ADHM construction, this metric may be interpreted as coming from a finite-dimensional hyperkähler quotient construction, where $\mathrm{U}(n)$ acts on a certain finite-dimensional space of matrix data. The program we would like to carry out now, is to use an adaption of the ADHM construction to split signature and thus find a hypersymplectic structure inside the moduli space of ASD-connections on $S^2 \times S^2$ in the presence of a complex structure.

5.3 Monads, ADHM Data and Hypersymplectic Quotients

Split signature instantons correspond to stable bundles on $\mathbb{CP}^3$ satisfying an additional reality condition. In this section, we wish to construct such bundles that are in addition trivial when restricted to the line $\{z_3 = z_4 = 0\} \subset \mathbb{CP}^3$. It will turn out that these can be described in terms of matrix-data satisfying some quadratic equations. Similarly to writing Schmid's equations in complex co-ordinates, these matrix equations can be interpreted as the vanishing condition of a hypersymplectic moment map.

5.3.1 ADHM Data in Split Signature

Since split signature instantons correspond to stable bundles on $\mathbb{CP}^3$, see section 4 in [45], we know that they can be described in terms of *monads*, see also [3] for a discussion of the application of monads to construct Euclidean instantons. By a monad, we mean a collection of complex vector spaces V, W, U of dimensions $k, 2k+l$ and k respectively, with maps

$$A_z : V \to W \qquad B_z : W \to U,$$

of the form

$$A_z = \sum_{i=1}^{4} A_i z_i \qquad B_z = \sum_{i=1}^{4} B_i z_i,$$

for each $z = (z_1, z_2, z_3, z_4) \in \mathbb{C}^4$. These are required to form a complex, i.e.

$$B_z \circ A_z = 0 \qquad \text{for all } z \in \mathbb{C}^4,$$

with the non-degeneracy condition that A_z should be injective and B_z should be surjective for all $z \in \mathbb{C}^4$. The non-degeneracy conditions ensure that the collection of cohomology spaces

$$E_z = \ker(B_z)/\mathrm{im}(A_z)$$

forms a holomorphic vector bundle on $\mathbb{CP}^3$ of rank l and second Chern class k. We are interested in vector bundles that are trivial along the line $\{z_4 = 0 = z_3\}$. It is proved in [51], chapter IV, Lemma 4.2.3 and Remark 4.2.4, that for the monad this translates into the condition that the composite

$$B_2 \circ A_1 = -B_1 \circ A_2$$

should be an isomorphism. In particular, A_1 and A_2 as well as B_1, B_2 should have maximal rank k. Thus, we may choose bases of V, W and U such that A_1 and B_2 are given by the following block matrices

$$A_1 = \begin{pmatrix} 1_{k\times k} \\ 0_{k\times k} \\ 0_{l\times k} \end{pmatrix}, \qquad B_2 = \begin{pmatrix} -1_{k\times k} & 0_{k\times k} & 0_{k\times l} \end{pmatrix}.$$

We have the additional condition that

$$B_1 \circ A_1 = 0 = B_2 \circ A_2.$$

And this implies that we may modify the above bases so that

$$A_2 = \begin{pmatrix} 0_{k\times k} \\ 1_{k\times k} \\ 0_{l\times k} \end{pmatrix} \qquad B_1 = \begin{pmatrix} 0_{k\times k} & 1_{k\times k} & 0_{k\times l} \end{pmatrix}.$$

With respect to these bases, the remaining four linear maps can be written in the form

$$A_3 = \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} \qquad B_3 = \begin{pmatrix} -\beta & \alpha & \delta \end{pmatrix},$$

and

$$A_4 = \begin{pmatrix} \tilde{\alpha} \\ \tilde{\beta} \\ \tilde{\gamma} \end{pmatrix} \qquad B_4 = \begin{pmatrix} -\tilde{\beta} & \tilde{\alpha} & \tilde{\delta} \end{pmatrix},$$

where $\alpha, \tilde{\alpha}, \beta, \tilde{\beta}$ are $k \times k$-matrices and $\gamma, \tilde{\gamma}$ and $\delta, \tilde{\delta}$ are matrices of dimensions $l \times k$ and $k \times l$ respectively. The condition that $B_z \circ A_z = 0$ is then equivalent to the following system of equations

$$\begin{aligned} [\alpha, \beta] + \delta\gamma &= 0 \\ [\tilde{\alpha}, \tilde{\beta}] + \tilde{\delta}\tilde{\gamma} &= 0 \\ [\alpha, \tilde{\beta}] + [\tilde{\alpha}, \beta] + \delta\tilde{\gamma} + \tilde{\delta}\gamma &= 0. \end{aligned}$$

The above choice of bases can be rephrased as giving isomorphisms

$$W \cong (V \otimes \mathbb{C}^2) \oplus H, \qquad U \cong V,$$

where H is a complex vector space of dimension l. Then the monad is described by

$$\alpha, \tilde{\alpha}, \beta, \tilde{\beta} \in \mathrm{End}(V), \qquad \gamma, \tilde{\gamma} \in \mathrm{Hom}(V, H), \qquad \delta, \tilde{\delta} \in \mathrm{Hom}(H, V).$$

Recall that we are looking for bundles E on $\mathbb{CP}^3$ which are trivial along real lines and carry an isomorphism $\tilde{\tau}_E : \tau^* E \cong \bar{E}^*$ inducing a hermitian metric on $E|_{\mathbb{RP}^3}$. For our monad this means that we have an isomorphism to $V \cong \bar{V}^*$, inducing a hermitian metric and also a hermitian metric on H. We put on W the pseudo-hermitian metric obtained by interpreting this space as $(V \otimes \mathbb{C}^{1,1}) \oplus H$. Then we would like to describe the bundle E as the kernel of the map

$$R_z = A_z^* \oplus B_z : W \to V \oplus V \cong V \otimes \mathbb{C}^2.$$

Notice however, that this description of E is only valid for those $z \in \mathbb{CP}^3$, for which the restriction of the inner product on W to the image of A_z is non-degenerate. This condition is equivalent to saying that the matrix $A_z^* A_z$ should be invertible. After modifying the above bases once more to make them compatible with the real structures, R may be written as

$$R_z = \begin{pmatrix} A_1^* \bar{z}_1 + A_2^* \bar{z}_2 + A_3^* \bar{z}_3 + A_4^* \bar{z}_4 \\ B_1 z_1 + B_2 z_2 + B_3 z_3 + B_4 z_4 \end{pmatrix}.$$

On the space $V \otimes \mathbb{C}^2$ we have a natural action of the split-quaternions given by

$$I = \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix} \qquad S = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

and we require that R should be equivariant with respect to the two $\mathbb{B}$-actions, i.e.

$$R_{qz} = q.R_z.$$

This condition makes sure that the kernel of R_z remains constant along the real lines. Using the explicit form of A_1, A_2, B_1, B_2 derived above, this reality condition reduces to

$$\begin{pmatrix} B_3 z_3 + B_4 z_4 \\ A_3^* \bar{z}_3 + A_4^* \bar{z}_4 \end{pmatrix} = \begin{pmatrix} A_3^* z_4 + A_4^* z_3 \\ B_3 \bar{z}_4 + B_4 \bar{z}_3 \end{pmatrix}.$$

This means

$$A_3^* = B_4 \qquad A_4^* = B_3.$$

Keeping in mind the signature of our metric on W, we may write this explicitly as

$$A_3^* = \begin{pmatrix} -\alpha^* & \beta^* & \gamma^* \end{pmatrix} = \begin{pmatrix} -\tilde{\beta} & \tilde{\alpha} & \tilde{\delta} \end{pmatrix} = B_4,$$

and

$$A_4^* = \begin{pmatrix} -\tilde{\alpha}^* & \tilde{\beta}^* & \tilde{\gamma}^* \end{pmatrix} = \begin{pmatrix} -\beta & \alpha & \delta \end{pmatrix} = B_3,$$

which implies

$$\tilde{\alpha} = \beta^* \qquad \tilde{\beta} = \alpha^* \qquad \tilde{\gamma} = \delta^* \qquad \tilde{\delta} = \gamma^*.$$

So we may write the above monad equations purely in terms of $\alpha, \beta, \gamma, \delta$, and they become

$$\begin{aligned} [\alpha, \beta] + \delta\gamma &= 0, \\ -[\alpha^*, \alpha] + [\beta^*, \beta] + \gamma^*\gamma + \delta\delta^* &= 0. \end{aligned}$$

We say that $\alpha, \beta, \gamma, \delta$ form a system of *ADHM-data* for the bundle E and call the above equations the *split signature ADHM equations*. Note that we could have chosen different unitary bases for V and W and so we get an action of the group $U(V)$ on the set of ADHM-data. An element $u \in U(V)$ acts according to

$$u.(\alpha, \beta, \gamma, \delta) = (u^{-1}\alpha u, u^{-1}\beta u, \gamma u, u^{-1}\delta).$$

Two sets of ADHM-data that are related by this $U(V)$-action will give rise to isomorphic bundles. In this way, we are able to obtain (possibly singular) split signature instantons from an ADHM type construction.

5.3.2 The moduli space of ADHM Data as a Hypersymplectic Quotient

In this section, we are going to show that the split signature ADHM equations may be interpreted as a moment map for the action of $U(V)$ on the set of ADHM-data.

The Hypersymplectic Setup and the Moment Maps

Let V, H be complex hermitian vector spaces of dimension k, l, respectively. We have seen above that ADHM data are of the form $(\alpha, \beta, \gamma, \delta)$, where $\alpha, \beta \in \mathrm{End}(V)$, $\gamma \in \mathrm{Hom}(V, H)$ and $\delta \in \mathrm{Hom}(H, V)$. They are acted upon by the unitary group $U(V)$. We wish to put a hypersymplectic structure on the complex vector space

$$\mathcal{A} = \mathrm{End}(V) \oplus \mathrm{End}(V) \oplus \mathrm{Hom}(V, H) \oplus \mathrm{Hom}(H, V).$$

We do this as follows: Writing a tangent vector $T_{(\alpha,\beta,\gamma,\delta)}\mathcal{A}$ in the form (a, b, c, d), we define a split signature metric g in terms of the usual inner products given by the trace on the individual summands of $\mathcal{A}$, i.e.

$$g((a, b, c, d), (a, b, c, d)) = \mathrm{tr}_V(a^*a - b^*b - c^*c + dd^*).$$

The complex structure I is just multiplication by i,

$$I(a, b, c, d) = (ia, ib, ic, id),$$

and we define S to be

$$S(a, b, c, d) = (b^*, a^*, d^*, c^*),$$

where the adjoints are taken with respect to the hermitian inner products on V and H. Then by construction $IS = -SI$ and I and S are skew-adjoint with respect to g and parallel, as they do not depend on the point $(\alpha, \beta, \gamma, \delta)$. Thus, putting $T = IS$, we have defined a hypersymplectic structure on $\mathcal{A}$. The symplectic form associated to the complex structure I is given by

$$\omega_I((a,b,c,d),(a',b',c',d')) = \mathrm{Re}\left(\mathrm{tr}_V(-ia^*a' + ib^*b' + ic^*c' + idd'^*)\right).$$

The complex symplectic form $\omega_I^c = \omega_S + i\omega_T$ is then

$$\omega_I^c((a,b,c,d),(a',b',c',d')) = \mathrm{tr}_V(ba' - ab' - dc' + d'c).$$

We have seen before that we have an action by $U(V)$ on $\mathcal{A}$ given by

$$u.(\alpha,\beta,\gamma,\delta) = (u^{-1}\alpha u, u^{-1}\beta u, \gamma u, u^{-1}\delta),$$

where $u \in \mathrm{U}(V)$. For $\xi \in \mathfrak{u}(V)$, the associated fundamental vector field is given by

$$X^\xi_{(\alpha,\beta,\gamma,\delta)} = \frac{\mathrm{d}}{\mathrm{d}t}|_{t=0} \exp(t\xi).(\alpha,\beta,\gamma,\delta) = ([\alpha,\xi],[\beta,\xi],\gamma\xi,-\xi\delta).$$

With all the setup in place, we are now in a position to calculate the moment maps associated to this action.

Theorem 5.3.1. *We use the pairing given by* tr_V *to identify* $\mathfrak{u}(V)^* \cong \mathfrak{u}(V)$*, then we have* $\mathfrak{u}(V)^* \otimes \mathbb{C} \cong \mathfrak{u}(V) \otimes \mathbb{C} \cong \mathrm{End}(V)$*. Under these identifications, the moment maps of the action of* $U(V)$ *on* $\mathcal{A}$ *are given by*

$$\begin{aligned} \mu_I^c(\alpha,\beta,\gamma,\delta) &= [\alpha,\beta] + \delta\gamma, \\ \mu_I(\alpha,\beta,\gamma,\delta) &= \frac{-i}{2}\left(-[\alpha^*,\alpha] + [\beta^*,\beta] + \gamma^*\gamma + \delta\delta^*\right). \end{aligned}$$

In particular, the moduli space of ADHM-data modulo the action of $\mathrm{U}(V)$ *arises as the hypersymplectic quotient of* $\mathcal{A}$ *by the* $\mathrm{U}(V)$*-action and is expected to be of real dimension* $4\dim V \dim H$*.*

Proof. Let $\xi \in \mathfrak{u}(V)$ and let $X = (a,b,c,d) \in T_{(\alpha,\beta,\gamma,\delta)}\mathcal{A}$. We have

$$\begin{aligned} \omega_I^c(X^\xi, X) &= \mathrm{tr}_V([\beta,\xi]a - [\alpha,\xi]b + \xi\delta c + d\gamma\xi) \\ &= \mathrm{tr}_V([a,\beta]\xi - [b,\alpha]\xi + \delta c\xi + d\gamma\xi) \\ &= \frac{\mathrm{d}}{\mathrm{d}t}|_{t=0}\mathrm{tr}_V([\alpha+ta,\beta+tb]\xi + (\delta+td)(\gamma+tc)\xi) \\ &= (\mathrm{d}\mu_I^c)_{(\alpha,\beta,\gamma,\delta)}(\xi)(X), \end{aligned}$$

which gives the desired complex moment map. For the real moment map, we perform a similar calculation,

$$\begin{aligned}\omega_I(X^\xi, X) &= \mathrm{Re}(\mathrm{tr}_V(-i[\alpha,\xi]^*a + i[\beta,\xi]^*b + i(\gamma\xi)^*c - i\xi\delta d^*)) \\ &= \mathrm{Re}(\mathrm{tr}_V(-i[\alpha^*,\xi]a + i[\beta^*,\xi]b - i\xi\gamma^*c - i\xi\delta d^*)) \\ &= \mathrm{Re}(\mathrm{tr}_V(i[\alpha^*,a]\xi - i[\beta^*,b]\xi - i\gamma^*c\xi - i\delta d^*\xi)) \\ &= \frac{1}{2}(\mathrm{tr}_V(i[\alpha^*,a]\xi - i[\alpha,a^*]\xi + i[\beta,b^*]\xi - i[\beta^*,b]\xi)) \\ &\quad + \frac{1}{2}(\mathrm{tr}_V(-i\gamma^*c\xi - ic^*\gamma\xi - i\delta d^*\xi - id\delta^*\xi)) \\ &= (\mathrm{d}\mu_I)_{(\alpha,\beta,\gamma,\delta)}(\xi)(X).\end{aligned}$$

□

Degeneracies

In proving the hypersymplectic quotient construction, we saw that we can only expect a non-degenerate hypersymplectic structure on some open set in the quotient. In this paragraph, we give an explicit description of the degeneracy locus. At a solution $(\alpha,\beta,\gamma,\delta)$, the tangent space will be given by the space of solutions to the linearised equation intersected with the orthogonal complement to the tangent space to the orbit through $(\alpha,\beta,\gamma,\delta)$. That is, a tangent vector $X = (a,b,c,d) \in T_{(\alpha,\beta,\gamma,\delta)}\mathcal{A}$ has to satisfy

$$\begin{aligned}[a,\beta] + [\alpha,b] + \delta c + d\gamma &= 0 \\ [a^*,\alpha] + [\alpha^*,a] - [b^*,\beta] - [\beta^*,b] + c^*\gamma + \gamma^*c + d\delta^* + \delta d^* &= 0 \\ [\alpha^*,a] - [\beta^*,b] - \gamma^*c + \delta d^* &= 0,\end{aligned}$$

where the third equation is the condition that X should be perpendicular to the $U(V)$-orbit through $(\alpha,\beta,\gamma,\delta)$. Thus, the degeneracy locus is precisely given by solutions to the third equation that are of the form

$$X = X^\xi = ([\alpha,\xi],[\beta,\xi],\gamma\xi,-\xi\delta),$$

for some $\xi \in \mathfrak{u}(V)$. Hence, we arrive at the following description of the degeneracy locus.

Proposition 5.3.2. *A solution $(\alpha,\beta,\gamma,\delta)$ to the ADHM equations is contained in the degeneracy locus if and only if there exists a non-trivial solution $\xi \in \mathfrak{u}(V) \setminus \{0\}$ to the equation*

$$[\alpha^*,[\alpha,\xi]] - [\beta^*,[\beta,\xi]] - \gamma^*\gamma\xi + \delta\delta^*\xi = 0.$$

5.3.3 An ADHM Description for Framed Bundles on $\mathbb{CP}^1 \times \mathbb{CP}^1$

In the case of rank 2, the ADHM equations can also be interpreted in a different context. Recall that we want to produce vector bundles which are holomorphic with respect to each compatible complex structure on the base space $\mathbb{R}^{2,2}$ and which are framed at infinity. It is shown in [16] that holomorphic bundles of rank 2 with structure group $\mathrm{SL}(2,\mathbb{C})$ on $\mathbb{CP}^1 \times \mathbb{CP}^1$ trivial over $\{\infty\} \times \mathbb{CP}^1 \cup \mathbb{CP}^1 \times \{\infty\}$ can be obtained by a monad construction. Suppose we start with a collection of ADHM data $(\alpha, \beta, \gamma, \delta)$ defined as above, that is, satisfying

$$[\alpha, \beta] + \delta\gamma = 0, \qquad [\alpha^*, \alpha] - [\beta^*, \beta] + \gamma^*\gamma + \delta\delta^* = 0.$$

Let $z = x_1 - ix_2, w = x_3 + ix_4$ be complex co-ordinates on $\mathbb{R}^{2,2} \cong \mathbb{B} \cong \mathbb{C}^{1,1}$ with respect to the complex structure I. According to [16], we can define a holomorphic vector bundle via matrices

$$A_{(z,w)} : V \to V \otimes \mathbb{C}^{1,1} \oplus H \qquad B_{(z,w)} : V \otimes \mathbb{C}^{1,1} \oplus H \to V,$$

where

$$A_{(z,w)} = \begin{pmatrix} \alpha - w \\ \beta - z \\ \gamma \end{pmatrix}, \qquad B_{(z,w)} = \begin{pmatrix} -\beta + z & \alpha - w & \delta \end{pmatrix},$$

and V, H are hermitian vector spaces as before. For this to work we actually only need the complex equation to be satisfied. The bundle obtained in this way will be trivial on the two lines at infinity. The result we need here is Lemma 5 in section 2 of [16].

On $\mathcal{A}$ we have a two-sheeted hyperboloid of compatible complex structures and we have seen in the discussion of Lax pairs that the vanishing of the hypersymplectic moment map is equivalent to the vanishing of μ_ζ for all $\zeta \in \mathbb{CP}^1 \setminus \{|\zeta| = 1\}$, i.e. the complex equation should be satisfied for any choice of compatible complex structure.

Recall that for $|\zeta| \neq 1$ complex co-ordinates with respect to the associated complex structure I_ζ are given by

$$z_\zeta = z + i\zeta\bar{w} \qquad w_\zeta = w + i\zeta\bar{z}.$$

For fixed ζ we view these again as standard affine co-ordinates on $\mathbb{CP}^1 \times \mathbb{CP}^1$. That is, we embed $\mathbb{R}^{2,2}$ into $\mathbb{CP}^1 \times \mathbb{CP}^1$ as the subset $\{([1 : z_\zeta], [1 : w_\zeta]\}$. Analogously, we obtain complex co-ordinates on $\mathcal{A}$ via

$$\alpha_\zeta = \alpha - i\zeta\beta^* \qquad \beta_\zeta = \beta - i\zeta\alpha^* \qquad \gamma_\zeta = \gamma - i\zeta\delta^* \qquad \delta_\zeta = \delta - i\zeta\gamma^*.$$

In these co-ordinates the matrices describing the vector bundles are then given by

$$A_{(z_\zeta,w_\zeta)} : V \to V \otimes \mathbb{C}^{1,1} \oplus H \qquad B_{(z_\zeta,w_\zeta)} : V \otimes \mathbb{C}^{1,1} \oplus H \to V,$$

where

$$A_{(z_\zeta,w_\zeta)} = \begin{pmatrix} \alpha_\zeta - w_\zeta \\ \beta_\zeta - z_\zeta \\ \gamma_\zeta \end{pmatrix} \qquad B_{(z_\zeta,w_\zeta)} = \begin{pmatrix} -\beta_\zeta + z_\zeta & \alpha_\zeta - w_\zeta & \delta_\zeta \end{pmatrix}.$$

As before, the complex equation with respect to I_ζ is equivalent to the condition

$$B_{(z_\zeta,w_\zeta)} \circ A_{(z_\zeta,w_\zeta)} = 0.$$

Note that this equation only depends on the ADHM data and not on the co-ordinates (z_ζ, w_ζ), as they cancel out.

We summarise this discussion in the following proposition.

Proposition 5.3.3. *A quadruple $(\alpha, \beta, \gamma, \delta)$ satisfies the ADHM equations if and only if $B_{(z_\zeta,w_\zeta)} \circ A_{(z_\zeta,w_\zeta)} = 0$ for all $\zeta \in \mathbb{CP}^1 \setminus \{|\zeta| = 1\}$. Equivalently, the bundle defined by the cohomology of the monad is holomorphic with respect to any choice of compatible complex structure I_ζ.*

Thus, using lemma 5.1.1, we have produced an ASD connection on the bundle E restricted to the open subset over which the induced metric on the fibres of E is positive definite. As before, we identify the bundle E with the kernel of the map

$$R_{(z,w)} = A^*_{(z,w)} \oplus B_{(z,w)} : V \otimes \mathbb{C}^{1,1} \oplus H \to V \otimes \mathbb{C}^2.$$

The bundle E then acquires its connection from the trivial connection on $W = V \otimes \mathbb{C}^{1,1} \oplus H$ by orthogonal projection. At a point $p = (z, w)$, the projection map onto the kernel of A^*_p is given by

$$P_p = 1 - A_p(A^*_p A_p)^{-1} A^*_p.$$

The problem is now, that the metric on W is not positive definite and so $A^*_p A_p$ is not automatically invertible, so that this description breaks down at points where $A^*_p A_p$ has a non-trivial kernel. At such points the ASD connection will be singular. We also cannot guarantee that the induced metric on the kernel of the map R_p will be positive definite, which is necessary to obtain instantons with gauge group SU(2) (recall that the above description has only been developed in [16] for bundles of rank 2).

Bearing in mind the signature of the metric on W, A^*_p is given by

$$A^*_p = \begin{pmatrix} (\alpha - w)^* & -(\beta - z)^* & \gamma^* \end{pmatrix},$$

so that

$$A_p^* A_p = (\alpha - w)^*(\alpha - w) - (\beta - z)^*(\beta - z) + \gamma^*\gamma,$$

which is a self-adjoint, but unfortunately not automatically a positive $k \times k$ matrix, depending on $z, \bar{z}, w, \bar{w}$ in a polynomial manner. Its determinant $\det A_p^* A_p =: D(z, \bar{z}, w, \bar{w})$ is hence a polynomial in the four variables $z, \bar{z}, w, \bar{w}$, which is *real valued* since $A_p^* A_p$ is self-adjoined. Our ASD connection will be singular at points (z, w) where $D(z, \bar{z}, w, \bar{w}) = 0$, so we have to study this polynomial and and find out under which circumstances has a non-empty zero-locus. Explicitly, we see that as a matrix polynomial $A_p^* A_p$ is of the form

$$A_p^* A_p = (|w|^2 - |z|^2)\mathrm{id}_V + (\text{lower order terms}).$$

Its determinant D hence looks like

$$D(z, \bar{z}, w, \bar{w}) = (|w|^2 - |z|^2)^k + (\text{lower order terms}).$$

In particular, if k is even, then the leading term is always positive and we cannot say much about the zero set of D. However, if k is odd, we observe that if we fix $|z|$ and let $|w|$ become large, then certainly then value of D will be positive, for large enough $|w|$. If we fix $|w|$ and let $|z|$ grow, then, for large enough $|z|$, D will become negative. Hence, by continuity D must be zero for some (z, w). We summarise this result in the following proposition and remark that it has been found previously by Mason [45] using different methods.

Proposition 5.3.4 ([45]). *A non-singular split-signature instanton must have even second Chern number k.*

Thus, we have described a way to construct possibly singular ASD connections on $\mathbb{CP}^1 \times \mathbb{CP}^1$ directly in terms of ADHM data on this space rather than on the twistor space $\mathbb{CP}^3$.

5.3.4 Open Questions

There are several open questions in the above construction that we would like to answer in the future. The most obvious one is how to relate the two different ADHM constructions. We know that a choice of complex structure on $S^2 \times S^2$ corresponds to an embedding of $S^2 \times S^2$ into $\mathbb{CP}^3$ as a quadric with no real points, see appendix B in [45]. In this way stable bundles on $\mathbb{CP}^3$ give holomorphic bundles on $S^2 \times S^2$ by restriction. However, the quadric corresponding to the choice of complex structure in section 5.3.1 is given in the w-co-ordinates by the equation

$$\sum_{i=1}^{4} w_i^2 = 0,$$

which clearly has no real points. In the z-co-ordinates this equation reads

$$z_1 z_2 + z_3 z_4 = 0.$$

Under this embedding, the lines at infinity correspond to the union of lines $\{z_2 = 0 = z_3 z_4\}$, none of which is real, of course. However, when we derived the reality condition for the map R_z in 5.3.1, it seems that we needed the fact that the line $\{z_3 = z_4 = 0\}$ actually *is* real.

Another task is to understand the singularities of the ASD connection we produced above, i.e. can we characterise the points at which the description of the bundle and its hermitian structure in terms of the map R_z breaks down? What is the significance of ASD connections that come from ADHM data that lie in the degeneracy locus of the hypersymplectic structure? Notice also that we have not addressed the question of smoothness of the ADHM moduli space in the above discussion.

A full analysis of the above questions could then finally lead to the construction of a *Nahm transform*, with the hope of relating solutions to Schmid's equations to space-time monopoles, also known as solutions to Ward's chiral model, see [57] and [17]. See [26] for the correspondence between Nahm's equations and magnetic monopoles. This could be an interesting project to work on in the future.

Bibliography

[1] D. Alekseevsky and V. Cortés. The twistor spaces of a para-quaternionic kähler manifold. *Osaka J. Math.*, 45(1):215–251, 2008.

[2] A. Andrada. Hypersymplectic Lie algebras. *J. Geom. Phys.*, 56(10):2039–2067, 2006.

[3] M. F. Atiyah. *Geometry of Yang-Mills fields.* Scuola Normale Superiore Pisa, Pisa, 1979.

[4] M. F. Atiyah and R. Bott. The Yang-Mills equations over Riemann surfaces. *Philos. Trans. Roy. Soc. London Ser. A*, 308:523–615, 1982.

[5] M. F. Atiyah and N. J. Hitchin. *The Geometry and Dynamics of Magnetic Monopoles.* M. B. Porter Lectures. Princeton University Press, Princeton, NJ, 1988.

[6] M. F. Atiyah, N. J. Hitchin, and I. M. Singer. Self-duality in four-dimensional Riemannian geometry. *Proc. Roy. Soc. London Ser. A*, 362(1711):425–461, 1978.

[7] T. N. Bailey, M. G. Eastwood, A. R. Gover, and L. J. Mason. The funk transform as a penrose transform. *Math. Proc. Cambridge Philos. Soc.*, 125(1):67 – 81, 1999.

[8] T. N. Bailey, M. G. Eastwood, A. R. Gover, and L. J. Mason. Complex analysis and the funk transform. *J. Korean Math. Soc.*, 40(4):577 – 593, 2003.

[9] A. L. Besse. *Einstein Manifolds.* Number 10 in Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]. Springer-Verlag, Berlin, 1987.

[10] O. Biquard. Sur les équations de Nahm et la structure de Poisson des algèbres de Lie semi-simples complexes. *Math. Ann.*, 304(2):253–276, 1996.

[11] D. E. Blair, J. Davidov, and O. Muškarov. Hyperbolic twistor spaces. *Rocky Mountain J. Math.*, 35(5):1437–1465, 2005.

[12] H. W. Braden. Cyclic monopoles, affine Toda and spectral curves. *Comm. Math. Phys.*, 308(2):303–323, 2011.

[13] A. Cannas da Silva. *Lectures on Symplectic Geometry.* Number 1764 in Lecture Notes in Mathematics. Springer-Verlag, Berlin, 2001.

[14] R. W. Carter. *Lie Algebras of Finite and Affine Type.* Number 96 in Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, 2005.

[15] M. Carvalho, J. A. Helayël-Neto, and M. W. de Oliveira. Locally product structures and supersymmetry. *Lett. Math. Phys.*, 64(2):93–104, 2003.

[16] B. Charbonneau and J. Hurtubise. Calorons, Nahm's equations on S^1 and bundles over $\mathbb{P}^1 \times \mathbb{P}^1$. *Comm. Math. Phys.*, 280(2):315–349, 2008.

[17] B. Dai, C.-L. Terng, and K. K. Uhlenbeck. On the space-time monopole equation. In *Surv. Differ. Geom.*, volume 10, pages 1–30. Int. Press, Somerville, 2006.

[18] A.S. Dancer and A. Swann. Hyper-Kähler metrics associated to compact Lie groups. *Math. Proc. Cambridge Philos. Soc.*, 120(1):61–69, 1996.

[19] A.S. Dancer and A. Swann. Toric Hypersymplectic quotients. *Trans. Amer. Math. Soc.*, 359(3):1265–1284, 2007.

[20] S. K. Donaldson. Instantons and geometric invariant theory. *Comm. Math. Phys.*, 93(4):453–460, 1984.

[21] S. K. Donaldson. Nahm's equations and the classification of monopoles. *Comm. Math. Phys.*, 96(3):387–407, 1984.

[22] S. K. Donaldson. *Riemann surfaces.* Number 22 in Oxford Graduate Texts in Mathematics. Oxford University Press, Oxford, 2011.

[23] S. K. Donaldson and P. B. Kronheimer. *The Geometry of Four-Manifolds.* Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, 1990.

[24] M. G. Eastwood and C. R. Graham. The involutive structure on the blow-up of $\mathbb{R}^n$ in $\mathbb{C}^n$. *Comm. Anal. Geom.*, 7(3):609 – 622, 1999.

[25] D. S. Freed and K. K. Uhlenbeck. *Instantons and Four-Manifolds.* Number 1 in Mathematical Sciences Research Institute Publications. Springer-Verlag, New York, 1984.

[26] N. J. Hitchin. On the construction of monopoles. *Comm. Math. Phys.*, 89(2):145–190, 1983.

[27] N. J. Hitchin. *Monopoles, Minimal Surfaces and Algebraic Curves.* Number 105 in Séminaire de Mathématiques Supérieures [Seminar on Higher Mathematics]. Presses de l'Université de Montréal, Montreal, QC, 1987.

[28] N. J. Hitchin. The self-duality equations on a Riemann surface. *Proc. London Math. Soc. (3)*, 55(1):59–126, 1987.

[29] N. J. Hitchin. Harmonic maps from a 2-torus to the 3-sphere. *J. Differential Geom.*, 31(3):627–710, 1990.

[30] N. J. Hitchin. Hypersymplectic quotients. *Acta Acad. Sci. Tauriensis*, 124:169–180, 1990.

[31] N. J. Hitchin, A. Karlhede, U. Lindström, and M. Roček. Hyper-Kähler metrics and supersymmetry. *Comm. Math. Phys.*, 108(4):535–589, 1987.

[32] N. J. Hitchin, N. S. Manton, and M. K.Murray. Symmetric monopoles. *Nonlinearity*, 8(5):661 – 692, 1995.

[33] C. M. Hull. Actions for $(2, 1)$ sigma models and strings. *Nuclear Phys. B*, 509(1-2):252–272, 1998.

[34] H. Kamada. Neutral hyper-kähler structures on primary kodaira surfaces. *Tsukuba J. Math.*, 23(2):321–332, 1999.

[35] F. C. Kirwan. *Cohomology of Quotients in Symplectic and Algebraic Geometry.* Number 31 in Mathematical Notes. Princeton University Press, Princeton, NJ, 1984.

[36] S. Kobayashi. *Differential Geometry of Complex Vector Bundles.* Number 5 in Kanô Memorial Lectures. Princeton University Press, Princeton, NJ, 1987.

[37] A. G. Kovalev. Nahm's equations and complex adjoint orbits. *Quart. J. Math. Oxford Ser. (2)*, 47(185):41–58, 1996.

[38] P. Kronheimer. A Hyperkähler structure on the cotangent bundle of a complex Lie group. *MSRI preprint*, 1988.

[39] P. Kronheimer. A hyper-Kählerian structure on coadjoint orbits of a semisimple complex group. *J. London Math. Soc. (2)*, 42(2):193–208, 1990.

[40] S. Lang. *Fundamentals of differential geometry.* Number 191 in Graduate Texts in Mathematics. Springer-Verlag, New York, 1999.

[41] H. B. Lawson and M.-L. Michelsohn. *Spin Geometry.* Princeton University Press, 1989.

[42] C. LeBrun and L. J. Mason. Nonlinear gravitons, null geodesics, and holomorphic disks. *Duke Math. J.*, 136(2):205 – 273, 2007.

[43] A. Maciocia. Metrics on the moduli spaces of instantons over euclidean 4 -space. *Comm. Math. Phys.*, 135(3):467–482, 1991.

[44] N. S. Manton. A remark on the scattering of BPS monopoles. *Phys. Lett. B*, 110(1):54–56, 1982.

[45] L. J. Mason. Global anti-self-dual Yang-Mills fields in split signature and their scattering. *J. Reine Angew. Math.*, 597:105–133, 2006.

[46] L. J. Mason and G. A. Sparling. Twistor correspondences for the soliton hierarchies. *J. Geom. Phys.*, 8(1-4):243–271, 1992.

[47] L. J. Mason and N. M. J. Woodhouse. *Integrability, self-duality, and twistor theory.* Number 15 in London Mathematical Society Monographs. New Series, Oxford Science Publications. The Clarendon Press, Oxford University Press, New York, 1996.

[48] T. Matsoukas. *Hypersymplectic Quotients.* PhD thesis, University of Oxford, 2010.

[49] P. K. Mitter and C.-M. Viallet. On the bundle of connections and the gauge orbit manifold in Yang-Mills theory. *Comm. Math. Phys.*, 79(4):457–472, 1981.

[50] L. G. Molinari. Determinants of block tridiagonal matrices. *Linear Algebra Appl.*, 429(8-9), 2008.

[51] C. Okonek, M. Schneider, and H. Spindler. *Vector bundles on complex projective spaces*. Number 3 in Progress in Mathematics. Birkhäuser, Boston, Mass., 1980.

[52] S. Salamon. Quaternionic kähler manifolds. *Invent. Math.*, 67(1):143–171, 1982.

[53] W. Schmid. Variation of Hodge structure: the singularities of the period mapping. *Invent. Math.*, 22:211–319, 1973.

[54] C. T. Simpson. Nonabelian Hodge theory. In *Proceedings of the International Congress of Mathematicians, Vol. I, II (Kyoto, 1990)*, pages 747–756, Tokyo, 1991. Math. Soc. Japan.

[55] P. M. Sutcliffe. Seiberg-Witten theory, monopole spectral curves and affine Toda solitons. *Phys. Lett. B*, 381(1-3):129–136, 1996.

[56] M. Toda. Studies of a non-linear lattice. *Phys. Rep.*, 18C(1), 1975.

[57] R. S. Ward. Classical solutions of the chiral model, unitons, and holomorphic vector bundles. *Comm. Math. Phys.*, 128(2):319–332, 1990.

[58] R. S. Ward and R. O. Jr. Wells. *Twistor geometry and field theory*. Cambridge Monographs on Mathematical Physics. Cambridge University Press, Cambridge, 1990.

[59] E. T. Whitaker and G. N. Watson. *A Course of Modern Analysis*. Cambridge University Press, 4th edition, 1927.