Infinite dimensional GIT and moment maps in differential geometry

Author(s):
Trautwein, Samuel

Publication Date:
2018

Permanent Link:
https://doi.org/10.3929/ethz-b-000281862

DISS. ETH NO. 25225

Infinite dimensional GIT and moment maps in differential geometry

A thesis submitted to attain the degree of

DOCTOR OF SCIENCES of ETH ZURICH

(Dr. sc. ETH Zurich)

presented by

SAMUEL TRAUTWEIN

MSc ETH Mathematics

born on 20.12.1987

citizen of Germany

accepted on the recommendation of

Prof. Dietmar Salamon, ETH Zürich, examiner

Prof. Oscar García-Prada, ICMAT Madrid, co-examiner

Prof. Paul Biran, ETH Zürich, co-examiner

2018

Abstract

The starting point of this thesis is the following observation of Atiyah and Bott [4]: The curvature of a connection on a bundle over a surface can be understood as a moment map for the action of the gauge group. Moreover, the moduli space of flat connections, or more generally of Yang–Mills connections, is closely related to the moduli space of holomorphic bundles obtained from geometric invariant theory. We discuss the various implications of this observation to the Yang–Mills equations and the symplectic vortex equations over Riemann surfaces. As main results, we obtain the analogue of the Ness uniqueness theorem, the Kempf-Ness theorem, the Hilbert-Mumford criterion and the moment-weight inequality in both settings. The main technical ingredients are long-time existence and convergence of the Yang–Mills and the Yang–Mills–Higgs heat flow. These are the parabolic flows associated to the corresponding moment map squared functionals in both setups.

Donaldson introduced in [35, 38] various extensions of the Atiyah–Bott picture to actions of the diffeomorphism group. We begin with a self-contained exposition of his moment map framework in [38] and its applications to Teichmüller theory. This is the starting point for the following three projects, discussed in the remainder of this thesis.

The first one generalizes Donaldson's construction of Teichmüller space to the moduli spaces of tuples of holomorphic differentials of mixed degree. These moduli spaces are closely related to Hitchin's higher Teichmüller components [59]. A distant hope is, that this might lead to a new construction of the Hitchin component using the diffeomorphism group instead of the gauge group.

The second project is joint work with Dietmar Salamon and Oscar García-Prada. We show that the Ricci form yields a moment map for the action of the group of exact volume preserving diffeomorphisms on the space of almost complex structures. This yields an extended Weil–Petersson symplectic form on the Calabi–Yau Teichmüller space of isotopy classes of complex structures with first Chern class zero and nonempty Kähler cone.

The third project is joint work with Oscar García-Prada, Luis Álvarez-Consul and Mario Garcia-Fernandez. We investigate variants of the Hitchin equations [58] where the complex structure is not fixed and the gauge group is extended by the Hamiltonian diffeomorphism group. This leads to moduli spaces which naturally fiber over Teichmüller space with fibre being the corresponding Hitchin moduli space.

Zusammenfassung

Der Ausgangspunkt dieser Arbeit ist die folgende Beobachtung von Atiyah und Bott [4]: Die Krümmung eines Zusammenhangs auf einem Bündel über einer Fläche kann als Momentum-Abbildung für die Wirkung der Eichgruppe verstanden werden. Zudem ist der Modulraum der flachen Zusammenhänge, oder allgemeiner der Yang–Mills Zusammenhänge, eng mit dem Modulraum der holomorphen Bündel aus der geometrischen Invariantentheorie verbunden. Wir diskutieren die Auswirkungen dieser Beobachtung auf die Yang–Mills Gleichungen und die symplektischen Wirbelgleichungen über Riemannschen Flächen. Als Hauptresultate erhalten wir Varianten des Ness–Eindeutigkeitssatzes, des Kempf–Ness Theorems, des Hilbert–Mumford Kriteriums und der Momentum–Gewichts Ungleichung in beiden Fällen. Ein zentrales technisches Resultat ist die Existenz und Konvergenz des Yang–Mills und Yang–Mills–Higgs Wärmeflusses. Diese sind die parabolischen Differentialgleichungen, welche aus der Gradientengleichung der normquadrierten Momentum-Abbildungen hervorgehen.

Donaldson führte in [35, 38] verschiedene Erweiterungen für Wirkungen der Diffeomorphismengruppe ein. Wir beginnen mit einer eigenständigen Darstellung seiner Momenten-Abbildung in [38] und deren Anwendungen auf die Teichmüller Theorie. Dies ist der Ausgangspunkt für die folgenden drei Projekte, welche im restlichen Teil der Arbeit diskutiert werden.

Das erste Projekt verallgemeinert Donaldson's Konstruktion des Teichmüller-Raums zu den Modulräumen von Tupeln von holomorphen Differentialen gemischten Grades. Diese Modulräume sind eng mit Hitchin's Teichmüller-Komponenten [59] verbunden. Eine entfernte Hoffnung ist, dass dies zu einer neuen Konstruktion der Hitchin-Komponente führen könnte, die die Diffeomorphismengruppe anstelle der Eichgruppe verwendet. Das zweite Projekt ist eine gemeinsame Arbeit mit Dietmar Salamon und Oscar García-Prada. Wir zeigen, dass die Ricci Form eine Momentum-Abbildung für die Wirkung der Gruppe der exakten volumenerhaltenden Diffeomorphismen auf dem Raum der fastkomplexen Strukturen liefert. Dies führt zu einer erweiterten Weil–Petersson symplektischen Form auf dem Calabi–Yau Teichmüller-Raum von Isotopieklassen komplexer Strukturen mit verschwindender reeller erster Chern–Klasse und nicht leerem Kähler-Kegel. Das dritte Projekt ist eine gemeinsame Arbeit mit Oscar García-Prada, Luis Álvarez-Consul and Mario Garcia-Fernandez. Wir untersuchen Varianten der Hitchin Gleichung [58], bei denen die komplexe Struktur auf der Fläche nicht festgehalten wird und die Eichgruppe um die Gruppe der Hamiltonischen Diffeomorphismen erweitert wird. Dies liefert Modulräume, welche auf natürliche Weise Faserungen über dem Teichmüller-Raums bilden. Als Fasern erhält man dabei die entsprechenden Hitchin-Modulräumen.

Acknowledgements

First and foremost, I want to thank my advisor Dietmar Salamon. I am deeply grateful for his constant support, patience and for sharing his knowledge and passion for mathematics with me over the last years. Without his guidance, this thesis would not have been possible.

Furthermore, I am thankful for the financial support provided by the Swiss National Science Foundation (grant number 200021-156000).

Besides my advisor, I would like to express my great appreciation to my co-examinators Paul Biran and Oscar García-Prada for their effort and interest in my work. I am particular thankful to Oscar for his hospitality when I visited the ICMAT in Madrid. During my stay, I benefited from fruitful discussions with Luis Álvarez-Consul, Mario Garcia-Fernandez and Nigel Hitchin.

I wish to thank my friends and colleagues from ETH, especially all the people from assistant group 6, for making the PhD such an enjoyable experience. A special word of thanks goes to my office mate and academic brother Charel for his friendship and the many stimulating discussions we had.

To all my friends and family, thank you for your thoughts and moral support. I am especially grateful to my mother, father, brother and his wife for their love and encouragement to follow my dreams. Last but not least, thank you Lilly for being by my side throughout this whole journey.

Contents

Chapter 1

Introduction

The concept of moment or momentum maps in symplectic geometry has a long history. They appeared first in Hamiltonian mechanics as conservation laws associated to symmetries and provided a formalism to reduce the degrees of freedom. Since then they turned out to be a powerful conceptual framework for many modern differential geometry questions and as such have led to various spectacular results, see [83] for an historic overview. In modern language, a moment map, associated to the action of a Lie group G with Lie algebra $\mathfrak{g}$ on a symplectic manifold X, is a map

$$\mu : X \to \mathfrak{g}^* \tag{1.1}$$

such that for every $\xi \in \mathfrak{g}$ the Hamiltonian function $H_\xi := \langle \mu(\cdot), \xi \rangle$ generates its infinitesimal action. In addition, one often requires that μ is equivariant for the coadjoint action. Marsden and Weinstein [82] observed that the quotient $\mu^{-1}(0)/G$, assuming that it is again a manifold, carries a canonical symplectic structure. This process is called symplectic reduction and $\mu^{-1}(0)/G$ is the Marsden–Weinstein quotient of X for the action of G.

Many important questions in differential geometry admit a description in terms of moment maps. Atiyah and Bott [4] observed that the curvature can be viewed as a moment map on the space of connections for the action of the gauge group. This observation has been extended to various gauge theoretical moduli problems, including the study of Hermitian Yang–Mills connections [31, 32, 118], Hitchin's equation [58, 102], Bradlow pairs [12] and the symplectic vortex equations [89]. Another beautiful observation due to Quillen and then generalized by Fujiki [48] and Donaldson [34] shows, that the scalar curvature provides a moment map on the space of compatible almost complex structures on a symplectic manifold for the action of the Hamiltonian diffeomorphism group. This is by no means intended to be a complete list and there are many more situations where moment maps occur naturally. All these examples have in common that the underlying symplectic manifold and the acting symmetry group are infinite dimensional. While some of them are very different in flavour, they rest on a conceptually unified framework provided by the moment map point of view which comes along with a package of standard theory.

Another important observation is the close connection between symplectic reduction and geometric invariant theory. This became apparent in the work of Atiyah and Bott [4] and has subsequently been explored in greater detail by Ness [92] and Kirwan [69]. Suppose G acts on a Kähler manifold X by Kähler isometries and the action is generated by a moment map $\mu : X \to \mathfrak{g}^*$. When the action of G extends to a holomorphic action of its complexification G^c then $\mu^{-1}(0)/G$ is naturally isomorphic to the GIT quotient $X//G^c$ introduced by Mumford [88]. The orbit space X/G^c is generally very badly behaved and the construction of the GIT quotient relies on a suitable notion of stability. It is defined as

$$X//G^c := (X^{ss}/G^c)/\sim \tag{1.2}$$

where $X^{ss} \subset X$ is the locus of semistable points and two orbits $G^c(x_1) \sim G^c(x_2)$ are identified if and only if $\overline{G^c(x_1)} \cap \overline{G^c(x_2)} \cap X^{ss} \neq \emptyset$. Here, the key observation is that every semistable orbit contains a unique G-orbit of solutions to the equation $\mu(x) = 0$ in its closure.

There are many remarkable infinite dimensional examples of geometric invariant theory. In these situations, the equation $\mu(x) = 0$ usually corresponds to difficult partial differential equations and one is interested in finding suitable stability criteria which characterize the existence of solutions. Examples are the Donaldson–Uhlenbeck–Yau correspondence [31, 32, 118] which relates stable holomorphic vector bundles to Hermitian Yang–Mills connections, the Kobayashi–Hitchin correspondence [89], or the recent work of Donaldson–Chen–Sun [20, 21, 22] relating K-stability to the existence of Kähler Einstein metrics on Fano manifolds.

1.1 Outline

The first part of this thesis investigates the GIT picture for the Yang–Mills and symplectic vortex equations over Riemann surfaces. In the second part, we provide an expository account on Donaldson's moment map framework for the diffeomorphism group [38] and its application to Teichmüller theory. We then present three projects based on these ideas: A construction of the moduli space of tuples of holomorphic differentials fibering over Teichmüller space, a new construction of the extended Weil–Petersson symplectic form on Calabi–Yau Teichmüller space, and a construction of universal Hitchin moduli spaces.

In the following, we give a detailed outline of this thesis. All theorems mentioned in this overview are new results proven in this thesis with the following exceptions: Theorem 1, Theorem 2, and Theorem 3 in Chapter 2 are mostly known, at least in special cases, and we present new proofs for these theorems which arise from a new perspective on the subject. Theorem 8, Theorem 9 and Theorem 10 in Chapter 4 are due to Donaldson [38] and Uhlenbeck [117]. We include detailed proofs of these results in order to have precise references when we consider generalizations of these constructions and for completeness of the exposition. The proofs of Theorem 14, Theorem 15, and Theorem 16 in Chapter 6 are only sketched in this thesis and full details are given in the joint paper [50]. Chapter 7 reports on work in progress and most of the material has not yet been explored in full detail.

A more comprehensive and detailed introduction of the various results and a review of the relevant existing literature can be found at the beginning of the respective chapters.

Chapter 2: The GIT picture for the Yang–Mills equation over Riemann surfaces

This chapter provides a self-contained exposition of the Atiyah–Bott picture for the Yang–Mills equation over Riemann surfaces [4]. In this overview we formulate three theorems that are mostly known (at least in the unitary case) and can be found in various places in the literature and were proved by different authors in different degrees of generality. The purpose of this chapter is to develop a new unified approach based on GIT and in the course of this we also give new proofs of the theorems. The fact that the Atiyah–Bott moment weight inequality in Theorem 3 can be proved along these lines was hinted at by Donaldson in [39]. We carefully study the semistable and unstable orbits and the main results are an analogue of the Ness uniqueness theorem for Yang-Mills connections, the Hilbert-Mumford criterion, and a sharp moment weight inequality. A central ingredient in our discussion is the Yang–Mills flow for which Råde [97] proved longtime existence and convergence. This chapter has been published in [114].

Let Σ be a Riemann surface, G a compact Lie group and $P \to \Sigma$ a principle G-bundle. We fix a Riemannian structure on Σ and an invariant inner product on the Lie algebra $\mathfrak{g}$ of G. Atiyah and Bott [4] observed that the curvature yields a moment map for the action of the gauge group $\mathcal{G}(P)$ on the space of connections $\mathcal{A}(P)$. By Chern–Weil theory, there is a unique central element $\tau \in Z(\mathfrak{g})$ with

$$\int_\Sigma \alpha(\tau) = \int_\Sigma \alpha(F_A) \qquad \text{for all } \alpha \in \mathfrak{g}^* \text{ and } A \in \mathcal{A}(P). \tag{1.3}$$

In particular, it follows that every projectively flat connection $A \in \mathcal{A}(P)$ has constant central curvature τ. We call this the central type of P and the Marsden–Weinstein quotient associated to the shifted moment $*F_A - \tau$ yields the moduli space of projectively flat connections on P. The complexified gauge group $\mathcal{G}(P)^c := \mathcal{G}(P^c)$ is defined as the gauge group of the complexified bundle. Its action on the space $\mathcal{A}(P)$ can be understood by identifying the space of connections $\mathcal{A}(P)$ with the space $\mathcal{J}(P^c)$ of holomorphic structures on the complexified bundle.

A central theorem in this theory relates the algebraic geometric notion of stable holomorphic principle bundles (see Definition 2.3.2) to the existence of projectively flat connections in a given complexified gauge orbit.

Theorem 1 (Narasimhan-Seshadri, Ramanathan [91, 95]). *Every $A \in \mathcal{A}(P)$ determines a unique holomorphic structure J_A on the complexified bundle $P^c := P \times_G G^c$ and the following holds:*

1. *(P^c, J_A) is stable if and only if there exists $g \in \mathcal{G}^c(P)$ such that $*F_{gA} = \tau$ and the kernel of $L_A : \Omega^0(\Sigma, ad(P) \otimes \mathbb{C}) \to \Omega^1(\Sigma, ad(P))$, $L_A(\xi + \boldsymbol{i}\eta) := d_A\xi + *d_A\eta$, contains only constant central sections.*

2. *(P^c, J_A) is polystable if and only if there exists $g \in \mathcal{G}^c(P)$ with $*F_{gA} = \tau$.*

3. *(P^c, J_A) is semistable if and only if* $\inf_{g \in \mathcal{G}^c(P)} ||*F_{gA} - \tau||_{L^2} = 0$.

4. *(P^c, J_A) is unstable if and only if* $\inf_{g \in \mathcal{G}^c(P)} ||*F_{gA} - \tau||_{L^2} > 0$.

Proof. This is reformulated as Theorem C in the introduction of Chapter 2 and proved in Theorem 2.3.10. □

The stable case is a reformulation of a theorem of Narasimhan-Seshadri [91] in the case $G = U(n)$ and Ramanathan [95] in the general case. They formulated the theorem in terms of irreducible representations instead of projectively flat connections and used entirely algebraic geometric methods for the proof. Analytic proofs of the stable case were found by Donaldson [30] in the case $G = U(n)$ and by Bradlow [12] and Mundet [89] for more general moduli problems. The polystable case is deduced from the stable case by induction on the dimension of G. The unstable and semistable cases have not been explicitly formulated in the literature to the best of our knowledge, but they are certainly known to the experts. The proof given here for the semistable and unstable case in Theorem 1 is new and based on the Yang-Mills flow, which we discuss next. The Yang–Mills functional is defined by

$$\mathcal{YM} : \mathcal{A}(P) \to \mathbb{R}, \qquad \mathcal{YM}(A) := \frac{1}{2}\int_\Sigma |F_A|^2 \,\mathrm{dvol}_\Sigma. \tag{1.4}$$

Råde [97] showed that for every initial data $A_0 \in \mathcal{A}(P)$ the gradient flow

$$\partial_t A(t) = -\nabla\mathcal{YM}(A(t)) = -d^*_{A(t)}F_{A(t)}, \qquad A(0) = A_0 \tag{1.5}$$

has a unique solution which exists for all time, remains in a single complexified orbit and converges in the $W^{1,2}$-topology to a critical point, satisfying the Yang–Mills equation $d_A^* F_A = 0$. This is the key ingredient in proving the following analogue of the Ness-Uniqueness theorem for Yang–Mills connections:

Theorem 2 (**Uniqueness of Yang-Mills connections**). *Let $A_0 \in \mathcal{A}(P)$ and let A_∞ be the limit of the Yang-Mills flow (1.5) starting at A_0. Then*

$$\mathcal{YM}(A_\infty) = \inf_{g\in\mathcal{G}^c(P)} \mathcal{YM}(gA) =: m. \tag{1.6}$$

Moreover, for every connection $B \in \overline{\mathcal{G}^c(A_0)}$ in the $W^{1,2}$-closure of $\mathcal{G}^c(A_0)$ with $\mathcal{YM}(B) = m$, it holds $\mathcal{G}(B) = \mathcal{G}(A_\infty)$.

Proof. This is reformulated as Theorem A in the introduction of Chapter 2 and proved in Theorem 2.4.14 and Theorem 2.4.15. □

This has been proven by Daskalopoulos [27] in the case $G = U(n)$ using slightly different methods. The general case has not been established in the literature, although we believe that it should be possible to reduce it to the unitary case by algebraic methods. We present an alternative proof which works directly for all Lie groups G and uses ideas of Chen–Sun [23] given in the context of extremal Kähler metrics.

The Hilbert–Mumford weight associated to a connection $A \in \mathcal{A}(P)$ and an infinitesimal gauge action $\xi \in \Omega^0(\Sigma, \mathrm{ad}(P))$ is defined by

$$w_\tau(A,\xi) := \lim_{t\to\infty} \langle *F_{e^{\mathrm{i}t\xi}A} - \tau, \xi\rangle \in \mathbb{R}\cup\{\infty\} \tag{1.7}$$

where $\tau \in Z(\mathfrak{g})$ is given by (1.3) as before. These weights are closely related to parabolic reductions of the holomorphic bundle (P^c, J_A) and we show that positivity of these weights is equivalent to the algebraic notion of stability. We also give a new proof of the moment weight inequality following an approach outlined by Donaldson [39]. Together with the dominant weight theorem, this yields the following result.

Theorem 3 (**Atiyah-Bott** [4]). *For all $A \in \mathcal{A}(P)$ and $0 \neq \xi \in \Omega^0(\Sigma, ad(P))$ it holds*

$$-\frac{w_\tau(A,\xi)}{||\xi||} \leq \inf_{g \in \mathcal{G}^c(P)} ||*F_A - \tau||^2. \tag{1.8}$$

When the right-hand-side is positive, then there exists up to scaling a unique $0 \neq \xi_0 \in \Omega^0(\Sigma, ad(P))$ which yields equality. Moreover, it agrees up to scaling with $u(*F_{A_\infty} - \tau)u^{-1}$, where $u \in \mathcal{G}(P)$ and A_∞ is the limit of the Yang-Mills flow starting at A_0.

Proof. This is reformulated as Theorem B in the introduction of Chapter 2 and proved in Theorem 2.5.12 and Theorem 2.7.1. □

The proof of the general case is only sketched by Atiyah and Bott. They use some deep results from Lie theory to reduce the general case to the unitary case. We give a different argument for the general case which relies on Theorem 2. This relies heavily on the Yang–Mills flow whose analytic properties had not been established when Atiyah and Bott wrote their paper. In other words, we use analysis to avoid the algebraic difficulties in their argument.

Chapter 3: Convergence of the Yang–Mills–Higgs flow and applications

This chapter extends the discussion of the previous chapter to the symplectic vortex equations. In doing so we fill several gaps in the literature on the Yang–Mills–Higgs functional and the characterization of stability for gauged holomorphic maps. In particular, we extend Mundet's Kobayashi–Hitchin correspondence and the Kempf–Ness theorem to the semistable and unstable case, establish a sharp moment–weight inequality and prove the analogue of the Ness uniqueness theorem. The main analytic result in our work is a Łojasiewicz gradient inequality and uniform convergence of the Yang–Mills–Higgs flow under suitable technical assumptions. This chapter has been published in [115].

Let Σ be a Riemann surface, G a compact Lie group and $P \to \Sigma$ a principle G-bundle. We fix an area form and hence a Riemannian structure on Σ, and an invariant inner product on the Lie algebra $\mathfrak{g}$. The latter allows us to identify $\mathfrak{g}$ with its dual space. Let X be a Kähler manifold equipped with an Hamiltonian action of G which is generated by a moment map $\mu : X \to \mathfrak{g}$. Denote by $\mathcal{S}(P, X)$ the space of sections of the associated Kähler fibration $P \times_G X$. The symplectic vortex equations for a pair $(A, u) \in \mathcal{A}(P) \times \mathcal{S}(P, X)$ are given by

$$\bar{\partial}_A u = 0, \qquad *F_A + \mu(u) = 0. \tag{1.9}$$

We view the first equation as integrability condition, which formally defines the Kähler submanifold of holomorphic pairs

$$\mathcal{H}(P,X) := \{(A,u) \in \mathcal{A}(P) \times \mathcal{S}(P,X) \,|\, \bar{\partial}_A u = 0\}. \tag{1.10}$$

The second equation in (1.9) is a moment map for the action of the gauge group $\mathcal{G}(P)$ on $\mathcal{A}(P) \times \mathcal{S}(P,X)$. We assume in the following, that the action of G extends to a holomorphic action of its complexification G^c on X. This gives rise to a natural action of the complexified gauge group $\mathcal{G}(P)^c$ on $\mathcal{A}(P) \times \mathcal{S}(P,X)$ which preserves $\mathcal{H}(P,X)$. The moment map squared functional

$$\mathcal{F} : \mathcal{H}(P,X) \to \mathbb{R}, \qquad \mathcal{F}(A,u) := \frac{1}{2}\int_\Sigma || * F_A + \mu(u)||^2 \, dvol_\Sigma \tag{1.11}$$

agrees up to topologoical terms with the Yang–Mills–Higgs functional

$$\mathcal{YMH}(A,u) := \frac{1}{2}\int_\Sigma ||F_A||^2 + ||d_A u||^2 + ||\mu(u)||^2 \, dvol_\Sigma \tag{1.12}$$

on holomorphic pairs $(A,u) \in \mathcal{H}(A,u)$. Its negative gradient flow on $\mathcal{H}(P,X)$ is

$$\begin{gathered} A(0) = A_0, \qquad u(0) = u_0, \qquad \bar{\partial}_A(u) = 0 \\ \partial_t A = - * d_A(*F_A + \mu(u)), \qquad \partial_t u = JL_u(*F_A + \mu(u)) \end{gathered} \tag{1.13}$$

Longtime existence of this flow has been proved by Venugopalan [119] in the case of vector bundles. We slightly generalize her proof and show that assumption (C) below suffices to obtain longtime existence. The proof of convergence relies on a new Łojasiewicz gradient inequality for the Yang–Mills–Higgs functional. For this we need to make the following assumptions:

(A) The Kähler metric on X and the moment map $\mu : X \to \mathfrak{g}$ are both analytic.

(B) X is holomorphically aspherical.

(C) μ is proper and X is equivariantly convex at infinity, i.e. there exists a proper G-invariant function $f : X \to [0,\infty)$ and $c_0 > 0$ such that

$$f(x) \geq c_0 \quad \Longrightarrow \quad \begin{gathered} \langle \nabla_v \nabla f(x), v\rangle + \langle \nabla_{Jv} \nabla f(x), Jv\rangle \geq 0 \\ df(x) JL_x \mu(x) \geq 0 \end{gathered} \tag{1.14}$$

for every $x \in X$ and $v \in T_x X$.

Theorem 4 (**Convergence**). *Assume **(C)** and let $(A_0,u_0) \in \mathcal{H}(P,X)$ be given. Then there exists a unique solution $(A,u) : [0,\infty) \to \mathcal{H}(P,X)$ of (1.13) which exists for all times $t \geq 0$. If in addition **(A)** and **(B)** are satisfied, then there exists a critical point $(A_\infty, u_\infty) \in \mathcal{A}^{1,2}(P) \times \mathcal{S}^{2,2}(P,X)$ of Sobolev class $W^{1,2} \times W^{2,2}$ and $T, C, \epsilon > 0$ such that for all $t > T$ the pointwise distance between $u(t)$ and u_∞ is smaller then the injectivity radius of X along $u_\infty(P)$ and*

$$||A(t) - A_\infty||_{W^{1,2}} + ||\exp_{u_\infty}^{-1} u(t)||_{W^{2,2}} \leq Ct^{-\epsilon}.$$

Proof. This is reformulated as Theorem A in the introduction of Chapter 3 and proved in Theorem 3.4.3 and Theorem 3.4.8. □

This theorem is the key ingredient which allows us to extend many of the analytic arguments from Chapter 2 to the present setting. As a first application, we get the following analogue of the Ness–Uniqueness theorem.

Theorem 5 (**Uniqueness of critical points**). *Assume **(A)**, **(B)** and **(C)**. Let $(A_0, u_0) \in \mathcal{H}(P, X)$ and let (A_∞, u_∞) be the limit of the gradient flow (1.13) starting at (A_0, u_0). Then*

$$|| * F_{A_\infty} + \mu(u_\infty)||_{L^2} = \inf_{g\in\mathcal{G}^c(P)} || * F_{gA_0} + \mu(gu_0)||_{L^2} =: m.$$

Moreover, for every $(B, v) \in \overline{\mathcal{G}^c(A_0, u_0)}$ in the $W^{1,2}\times W^{2,2}$-closure of $\mathcal{G}^c(A_0, u_0)$ with $|| * F_B + \mu(v)||_{L^2} = m$, it holds $\mathcal{G}(B, v) = \mathcal{G}(A_\infty, u_\infty)$.

Proof. This is reformulated as Theorem B in the introduction of Chapter 3 and proved in Theorem 3.5.1. □

The Hilbert–Mumford weight associated to a pair $(A, u) \in \mathcal{H}(P, X)$ and an infinitesimal gauge action $\xi \in \Omega^0(\Sigma, \mathrm{ad}(P))$ is defined by

$$w((A, u), \xi) := \lim_{t\to\infty} \left\langle *F_{e^{\mathbf{i}t\xi}A} + \mu(e^{\mathbf{i}t\xi}u),\, \xi\right\rangle_{L^2} \in \mathbb{R}\cup\{\infty\}. \tag{1.15}$$

In order to prove the moment weight inequality in this context, we need to assume the following property for pairs $(A, u) \in \mathcal{H}(P, X)$:

$$\textbf{(H)} \qquad w((A, u), \xi) \le 0 \quad\Longrightarrow\quad \sup_{t>0} ||\mu(e^{\mathbf{i}t\xi}u)||_{L^2} < \infty$$

for all $\xi \in \Omega^0(\Sigma, \mathrm{ad}(P))$. When $w((A, u), \xi) < \infty$, then the corresponding property for the curvature term $\sup_{t>0} ||F_{e^{\mathbf{i}t\xi}A}||_{L^2} < \infty$ is automatically satisfied. This is the reason why such an assumption did not occur in our previous discussion of the Yang–Mills equations.

Theorem 6 (**Sharp moment-weight inequality**). *Suppose that $(A, u) \in \mathcal{H}(P, X)$ satisfies **(H)**. Then for all $\xi \in \Omega^0(\Sigma, ad(P))\backslash\{0\}$ it holds*

$$-\frac{w((A, u), \xi)}{||\xi||_{L^2}} \le \inf_{g\in\mathcal{G}^c(P)} || * F_{gA} + \mu(gu)||_{L^2}. \tag{1.16}$$

*If in addition **(A)**, **(B)**, **(C)** are satisfied and the right hand side is positive, then there exists a unique $\xi_0 \in \Omega^0(\Sigma, ad(P))$ with $||\xi_0||_{L^2} = 1$ which yields equality.*

Proof. This is reformulated as Theorem E in the introduction of Chapter 3 and proved in Theorem 3.6.3. □

This is a crucial ingredient in our extension of Mundet's Kobayashi–Hitchin correspondence [89] to the semistable and polystable case. Consider the following properties for a pair $(A, u) \in \mathcal{H}(P, X)$:

(SS) For all $\xi \in \Omega^0(\Sigma, \mathrm{ad}(P))$ it holds $w((A,u),\xi) \geq 0$.

(PS) For all $\xi \in \Omega^0(\Sigma, \mathrm{ad}(P))$ with $\exp(\xi) = \mathbb{1}$ and $w((A,u),\xi) = 0$ the limit $\lim_{t\to\infty} e^{\mathbf{i}t\xi}(A,u) \in (\mathcal{G}^c)^{2,2}(A,u)$ exists in $W^{1,2} \times W^{2,2}$ and remains in the Sobolev completion of the complex group orbit.

Theorem 7 (**Polystable and semistable correspondence**). *Assume **(A)**, **(B)**, **(C)** and suppose that $(A,u) \in \mathcal{H}(P,X)$ satisfies **(H)**.*

1. *(A,u) is polystable if and only if it satisfies **(SS)** and **(PS)**.*
2. *(A,u) is semistable if and only if it satisfies **(SS)**.*

Proof. This is reformulated as Theorem D in the introduction of Chapter 3 and proved in Theorem 3.6.5 and Theorem 3.6.4. □

Mundet's Kobayashi–Hitchin correspondence [89] establishes this correspondence for stable pairs in greater generality.

Chapter 4: Donaldson's moment map approach to Teichmüller theory

This chapter provides a self-contained exposition of a general moment map found by Donaldson [38] for the diffeomorphism group. The main applications considered in this chapter is the construction of a hyperkähler moduli space $\mathcal{M}$ associated to a closed oriented surface Σ with $\mathrm{genus}(\Sigma) \geq 2$. This embeds naturally into the cotangent bundle $T^*\mathcal{T}(\Sigma)$ and can be viewed as the Feix–Kaledin hyperkähler extension of the Weil–Petersson metric on Teichmüller space. Donaldson outlined various remarkable properties of this moduli space for which we provide complete proofs: The moduli space $\mathcal{M}$ parametrizes the class of almost-Fuchsian 3-manifolds. These are quasi-Fuchsian 3-manifolds which contain a unique minimal surface with principal curvatures in $(-1,1)$. The area of this minimal surface then provides a Kähler potential for the hyperkähler metric. Moreover, the moduli space $\mathcal{M}$ embeds naturally into the $\mathrm{SL}(2,\mathbb{C})$-representation variety of Σ and the hyperkähler structure on $\mathcal{M}$ extends the Goldman holomorphic symplectic structure on the representation variety. The various identifications are obtained using the work of Uhlenbecks [117] on germs of hyperbolic 3-manifolds, an explicit map from $\mathcal{M}$ to $\mathcal{T}(\Sigma) \times \overline{\mathcal{T}(\Sigma)}$ found by Hodge [61], the simultaneous uniformization theorem of Bers [8], and the theory of Higgs bundles introduced by Hitchin [58].
Another motivation for such a detailed account on Donaldson's framework is the fact that there are several interesting variants and extensions of the theory. We will explore some of these in the remaining three chapters of this thesis building upon the discussion in this chapter.

Let (M,ρ) be a closed manifold equipped with a volume form ρ and denote by $P \to M$ its $\mathrm{SL}(n,\mathbb{R})$-frame bundle. Let (X,ω) be a symplectic manifold with Hamiltonian $\mathrm{SL}(n,\mathbb{R})$ action generated by a moment map $\mu : X \to \mathfrak{sl}(n,\mathbb{R})^*$ and denote by $\mathcal{S}(P,X)$ the space of section of the associated bundle $P \times_{\mathrm{SL}(n,\mathbb{R})} X$. The group $\mathrm{Diff}(M,\rho)$ of volume preserving diffeomorphisms acts naturally on the frame bundle

P and thus also on the space of $\mathcal{S}(P,X)$. This action preserves the natural symplectic structure and the main result of Donaldson asserts that the subgroup of exact volume preserving diffeomorphisms $\mathrm{Diff}^{\mathrm{ex}}(M,\rho)$ acts in a Hamiltonian fashion. Its Lie algebra is the space of exact divergence free vector fields and its dual space can be identified with the space of exact 2-forms on M. The moment map $\underline{\mu}$ in the Theorem 8 below takes only values in the space of closed 2-forms. It is therefore not a moment map in the strict sense, but Theorem 8 asserts that it nevertheless satisfies the moment map equation. When $\dim(M)=2$, one can fix this by subtracting a suitable multiple of the area form from the moment map. In higher dimensions there is no easy way to fix this without destroying equivariance.

Theorem 8 (**Donaldson [38]**). *Fix a torsion free $SL(n,\mathbb{R})$ connection ∇ on M and define $\underline{\mu}:\mathcal{S}(P,X)\to\Omega^2(M)$ by*

$$\underline{\mu}(s) := \omega(\nabla s\wedge\nabla s) - \langle\mu_s, R\rangle - dc(\nabla\mu_s) \tag{1.17}$$

where $\mu_s\in\Omega^0(M,End_0(TM)^*)$ is obtained by composing the equivariant lift $\tilde{s}:P\to X$ of s with the moment map $\mu:X\to\mathfrak{sl}(n,\mathbb{R})^*$ and $c(\nabla\mu_s)\in\Omega^1(M)$ is defined as the contraction $(\mu_s)^i_{j;i}$ of $\nabla\mu_s$. Then the following holds:

1. *The map $\underline{\mu}$ is Diff(M,ρ)-equivariant and $\underline{\mu}(s)\in\Omega^2(M)$ is closed and independent of the connection ∇ used to define it.*

2. *Let $v\in Vect(M)$ be an exact divergence free vector field and choose a primitive $\alpha_v\in\Omega^{n-2}(M)$ with $d\alpha_v=\iota(v)\rho$. Then*

$$\partial_t\int_M \underline{\mu}(s(t))\wedge\alpha_v = \int_M \omega(\dot{s}(t),\mathcal{L}_v s(t))\rho \tag{1.18}$$

   for any smooth curve $s:\mathbb{R}\to\mathcal{S}(X,P)$, where $\mathcal{L}_v s$ denotes the infinitesimal action of v on s for the right action.

Proof. This is reformulated as Theorem A in the introduction of Chapter 4 and proved in Theorem 4.2.4. (Note that the formula for the moment map in [38] contains some obvious typos regarding the signs which we corrected in formula stated above.) □

Suppose Σ is a 2-dimensional surface with genus$(\Sigma)\geq 2$ and $X=\mathbb{H}^2$ the hyperbolic plane. In this case, X can be identified with the space of linear complex structures on $\mathbb{R}^2$, compatible with the standard orientation, and $\mathcal{S}(P,X)$ can be identified with the space of complex structures on Σ, compatible with the orientation determined by ρ. In this case, Theorem 8 yields that the Gaussian curvature form yields a moment map for the action of the Hamiltonian diffeomorphism group. After taking the action of the Flux group into account, this provides a construction of the Teichmüller space of Σ equipped with the Weil–Petersson metric.

Next, suppose $X\subset T^*\mathbb{H}^2$ is the unit disc bundle in the cotangent bundle of the hyperbolic plane. This carries a unique $S^1\times\mathrm{SL}(2,\mathbb{R})$-invariant hyperkähler metric, which extends the hyperbolic metric along the zero section and blows up when approaching the boundary of the disc bundle. The space of sections $\mathcal{S}(P,X)$ can be identified with the space $\mathcal{Q}_1(\Sigma)$ of pairs (J,σ) where $J\in\mathcal{J}(\Sigma)$ is a complex structure and $\sigma \in \Omega^0(\Sigma, S^2(T^*\Sigma \otimes_J \mathbb{C}))$ a quadratic differential satisfying the pointwise constraint $|\sigma|_J < 1$. Then, Theorem 8 yields a hyperkähler moment map for the action of the Hamiltonian diffeomorphism group.

Theorem 9 (**Donaldson [38]**).

1. *The action of Ham(Σ, ρ) on $\mathcal{Q}_1(\Sigma)$ admits a hyperkähler moment map given by*

$$\begin{aligned} \underline{\mu}_1(J,\sigma) &= \frac{|\bar{\partial}\sigma|^2 - |\partial\sigma|^2}{\sqrt{1-|\sigma|^2}}\rho + 2\sqrt{1-|\sigma|^2}K_J\rho + 2\boldsymbol{i}\bar{\partial}\partial\sqrt{1-|\sigma|^2} - 2c\rho \\ \underline{\mu}_2(J,\sigma) + \boldsymbol{i}\underline{\mu}_3(J,\sigma) &= -2\boldsymbol{i}\overline{\bar{\partial} r(\bar{\partial}\sigma)} \end{aligned} \tag{1.19}$$

   where $c := 2\pi(2 - 2genus(\Sigma))/vol(\Sigma, \rho)$ and $r : \Omega^{0,1}(\Sigma, S^2(T^*\Sigma \otimes_J \mathbb{C})) \to \Omega^{1,0}(\Sigma)$ is the contraction defined by the metric $\rho(\cdot, J\cdot)$.

2. *The action of Symp$_0(\Sigma, \rho)$ on $\mathcal{Q}_1(\Sigma)$ is Hamiltonian for the second and third symplectic form with moment maps*

$$\langle \underline{\tilde{\mu}}_2(J,\sigma), v\rangle + \boldsymbol{i}\langle \underline{\tilde{\mu}}_3(J,\sigma), v\rangle = -2\boldsymbol{i}\int_\Sigma \overline{\iota(v) r(\bar{\partial}_J\sigma)}\rho \tag{1.20}$$

   for any symplectic vector field $v \in$ Vect(Σ) satisfying $d\iota(v)\rho = 0$.

Proof. This is reformulated as Theorem C in the introduction of Chapter 4 and proved in Theorem 4.5.13. □

We give an alternative proof of the second statement, since we found it difficult to translate the conceptual arguments given by Donaldson into a rigorous proof. We proceed by generalizing the proof of Theorem A where we use that the canonical holomorphic symplectic form $\omega_2 + \mathbf{i}\omega_3$ on X is exact. The methods employed in the proof of this result are the starting point for our discussion of holomorphic differentials of arbitrary degree in Chapter 5.

After carefully taking the action of the flux group into account and suitable rescaling of the quadratic differential, Theorem 9 gives rise to a hyperkähler structure on the moduli space

$$\mathcal{M} := \left\{ (g,\sigma) \in \mathrm{Met}(\Sigma) \times Q(g) \;\middle|\; \begin{array}{c} \bar{\partial}\sigma = 0,\ |\sigma| < 1, \\ K_g - \frac{c}{2}|\sigma|^2 = \frac{c}{2} \end{array} \right\} \Big/ \mathrm{Diff}_0(\Sigma) \tag{1.21}$$

where $c := 2\pi(2 - 2\mathrm{genus}(\Sigma))/\mathrm{vol}(\Sigma, \rho)$ as above. This takes a particularly simple form when we scale the volume of Σ such that $c = -2$. Donaldson proposed the following three geometric interpretations:

1. $\mathcal{M}$ embeds into $T^*\mathcal{T}(\Sigma)$ and the hyperkähler metric on $\mathcal{M}$ yields the Feix–Kaledin extension of the Weil–Petersson metric on $\mathcal{T}(\Sigma)$.

2. $\mathcal{M}$ parametrizes the class of almost-Fuchsian hyperbolic 3-manifolds. These are quasi-Fuchsian 3-manifolds which possess an incompressible minimal surface with principal curvatures in $(-1, 1)$. This surface is then unique and its area provides a Kähler potential for the hyperkähler metric.

3. $\mathcal{M}$ embeds as an open subset into the smooth locus of the $\mathrm{SL}(2,\mathbb{C})$ representation variety $\mathcal{R}_{\mathrm{SL}(2,\mathbb{C})}(\Sigma) := \mathrm{Hom}\left(\pi_1(\Sigma), \mathrm{SL}(2,\mathbb{C})\right)/\mathrm{SL}(2,\mathbb{C})$. The hyperkähler structure on $\mathcal{M}$ is compatible with the natural holomorphic symplectic structure introduced by Goldman [52], where the natural complex structure coincides with the second complex structure on $\mathcal{M}$.

The class of almost-Fuchsian manifolds is strictly smaller the the class of quasi-Fuchsian manifold: There are examples of quasi-Fuchsian manifolds which admit more then one minimal surface (see [121, 63, 57]) and these cannot be almost-Fuchsian (see Lemma 4.6.5). The isomorphism between $\mathcal{M}$ and the space of almost-Fuchsian manifolds follows from Uhlenbeck's theory of minimal surfaces in hyperbolic 3-manifolds [117]. Her result gives rise to the following theorem in our context.

Theorem 10 (**Uhlenbeck** [117]). *Let $g \in Met(\Sigma)$ and $\sigma \in Q(g)$ satisfy the equations $K_g + |\sigma|^2 = -1$, $\bar{\partial}\sigma = 0$, and $|\sigma|_g < 1$. For every such pair we define an almost-Fuchsian metric on $Y := \Sigma \times \mathbb{R}$ by*

$$g^Y = g^Y_{g,\sigma} = \begin{pmatrix} g\left(\cosh(t)\mathbb{1} - \sinh(t) g^{-1} Re(\sigma)\right)^2 & 0 \\ 0 & 1 \end{pmatrix}. \tag{1.22}$$

This is the unique almost-Fuchsian metric which restricts to g along $\Sigma \times \{0\}$ and such that $Re(\sigma)$ is the second fundamental form of $\Sigma \times \{0\} \subset Y$.

Proof. This is reformulated as Theorem C in the introduction of Chapter 4 and proved in Theorem 4.6.4. □

Let $(Y := \Sigma \times \mathbb{R}, g^Y)$ be an almost Fuchsian manifold. Its boundary at infinity is the disjoint union of two disjoint unions of Σ, which are both equipped with an induced conformal structure. This gives rise to an embedding of the space of almost Fuchsian metrics into the product space $\mathcal{T}(\Sigma) \times \overline{\mathcal{T}(\Sigma)}$ and we show that the second complex structure on $\mathcal{M}$ corresponds to the complex structure $(\hat{J}_1, \hat{J}_2) \mapsto (-J_1\hat{J}_1, J_2\hat{J}_2)$ on $\mathcal{T}(\Sigma) \times \overline{\mathcal{T}(\Sigma)}$. With this understood, we then verify the following remarkable observation suggested by Donaldson.

Theorem 11. *Let $A : \mathcal{AF}(\Sigma) \to \mathbb{R}$ be the area functional, which assigns to an almost Fuchisan manifold Y the area of its unique minimal surface. Then*

$$2\boldsymbol{i}\bar{\partial}_{J_2}\partial_{J_2} A = \underline{\omega}_2. \tag{1.23}$$

Hence A provides a Kähler potential with respect to the natural complex structure on $\mathcal{AF}(\Sigma)$ which agrees (up to sign) with the second complex structure on $\mathcal{M}$.

Proof. This is reformulated as Theorem D in the introduction of Chapter 4 and proved in Theorem 4.6.9. □

By the Cartan–Ambrose–Higgs theorem, one can express every complete hyperbolic 3-manifold as quotient of hyperbolic space $\mathbb{H}^3$. This gives rise to a natural embedding of the almost Fuchsian moduli space into $\mathcal{R}_{\mathrm{PSL}(2,\mathbb{C})}(\Sigma)$. A classical result of Bers [9] asserts that the restriction of this complex structure to $\mathcal{M}$ corresponds to the standard complex structure on $\mathcal{T}(\Sigma) \times \overline{\mathcal{T}(\Sigma)}$ which differs by a sign from our conventions. In particular, the second complex structure on $\mathcal{M}$ corresponds to multiplication by $-\mathbf{i}$ on $\mathcal{R}_{\mathrm{PSL}(2,\mathbb{C})}(\Sigma)$.

The representation associated to an almost-Fuchisan manifold lifts to $\mathrm{SL}(2,\mathbb{C})$ and a corresponding embedding of $\mathcal{M}$ into $\mathcal{R}_{\mathrm{SL}(2,\mathbb{C})}(\Sigma)$ can be constructed directly using the theory of Higgs bundles [58]. This has been suggested by Donaldson [38] and goes as follows: Let $g \in \mathrm{Met}(\Sigma)$ and $\sigma \in Q(g)$ be given. Choose a holomorphic line bundle $L \to \Sigma$ with $L^2 = T\Sigma$ and define $E = L \oplus L^{-1}$. The Levi-Civita connection for g induces a unique $U(1)$-connection $a \in \mathcal{A}(L)$. Then consider the pair

$$A = \begin{pmatrix} a & \frac{\bar{\sigma}}{2} \\ -\frac{\sigma}{2} & -a \end{pmatrix} \in \mathcal{A}(E) \quad \text{and} \quad \phi = \frac{1}{2}\begin{pmatrix} 0 & \mathbf{1} \\ 0 & 0 \end{pmatrix} \in \Omega^{1,0}(\mathrm{End}(E)) \tag{1.24}$$

where $\sigma \in \Omega^{1,0}(L^{-2}) = \Omega^{1,0}(\mathrm{Hom}(L, L^{-1}))$ and $\mathbf{1} \in \Omega^0(\mathrm{End}(T\Sigma)) = \Omega^{1,0}(L^2) = \Omega^{1,0}(\mathrm{Hom}(L^{-1}, L))$.

Theorem 12. *Let $g \in Met(\Sigma)$ and $\sigma \in Q(g)$ satisfy the equations $K_g + |\sigma|^2 = -1$, $\bar{\partial}\sigma = 0$, and $|\sigma|_g < 1$. The corresponding pair (A, ϕ) defined by (1.24) satisfies the Hitchin equation*

$$\bar{\partial}_A \phi = 0, \qquad F_A + [\phi \wedge \phi^*] = 0.$$

and $B := A + \phi + \phi^* \in \mathcal{A}^c(E)$ is a flat $SL(2,\mathbb{C})$ connection. The holonomy representation $\rho_B : \pi_1(\Sigma) \to SL(2,\mathbb{C})$ agrees up to conjugation with the representation associated to the almost Fuchian metric $g^Y_{g,\sigma}$ defined in Theorem 10.

Proof. This is reformulated as Theorem D in the introduction of Chapter 4 and proved in Theorem 4.6.12). □

Finally, we show that the natural map of $\mathcal{M}$ into $T^*\mathcal{T}(\Sigma)$ is a well-defined embedding (see Theorem 4.6.14). This follows by a standard application of the continuation method and the proof is due to Uhlenbeck [117].

Chapter 5: Moduli spaces of holomorphic differentials over Riemann surfaces

We describe a generalization of Donaldson's construction of Teichmüller space and its Feix–Kaledin hyperkähler extension to moduli spaces of tuples of holomorphic differentials of mixed degree. These moduli spaces are closely related to Hitchin's higher Teichmüller components [59]. We hope that this might lead to a new construction of the Hitchin component using the diffeomorphism group instead of the gauge group.

Let (Σ, ρ) be a closed 2-dimensional surface with fixed area form $\rho \in \Omega^2(\Sigma)$. Assume $\mathrm{genus}(\Sigma) \geq 2$ and denote by $P \to \Sigma$ its $\mathrm{SL}(2,\mathbb{R})$-frame bundle. For $k \geq 2$, we define

$$X_k = \left\{(z, w) \in \mathbb{H} \times \mathbb{C} \;\middle|\; \mathrm{Im}(z)^k |w|^2 < 1\right\} \tag{1.25}$$

which we view as unit disc bundle in $(T^*\mathbb{H})^{k/2}$. This can naturally be identified with the space of pairs (J, γ) where $J \in \mathcal{J}(\mathbb{R}^2)$ is a linear complex structure on $\mathbb{R}^2$ and $\gamma : (\mathbb{R}^2, J)^k \to \mathbb{C}$ is a complex symmetric multilinear form with $|\gamma| < 1$. The space of sections of the associated bundle $P \times_{\mathrm{SL}(2,\mathbb{R})} X_k$ then admits a natural identification with

$$\mathcal{D}^1_k(\Sigma) := \left\{(J,\tau) \mid J \in \mathcal{J}(\Sigma),\, \tau \in \Omega^0(\Sigma, S^k(T^*\Sigma \otimes_J \mathbb{C})),\, |\tau|_J < 1\right\} \tag{1.26}$$

which parametrizes complex structures and complex differentials of order k. In order to obtain a symplectic structure on $\mathcal{D}^1_k(\Sigma)$ we need to define a $\mathrm{SL}(2,\mathbb{R})$-invariant symplectic structure on the total space of $X_k \subset (T^*\mathbb{H})^{k/2}$. In the case $k = 2$ there is a natural choice, namely the Feix–Kaledin hyperkähler extension of the hyperbolic metric on $\mathbb{H}$. For $k > 2$ we do not expect that there exists a hyperkähler setup. Instead we obtain a family of symplectic forms on X_k parametrized by a single functions $f : [0,1) \to [0,1)$ with $f(0) = 0$ and $f' > 0$: There exists a unique $\mathrm{SL}(2,\mathbb{R})$-invariant symplectic form $\omega_f \in \Omega^2(X_k)$ satisfying

$$\begin{aligned} \omega_f(\mathbf{i}, w) = &-\frac{\mathbf{i}}{2}\left(1 - f(|w|^2) + k|w|^2 f'(|w|^2)\right) d\bar{z} \wedge dz \\ &- \frac{2\mathbf{i}}{k} f'(|w|^2)\, d\bar{w} \wedge dw + f'(|w|^2)\left(\bar{w} d\bar{z} \wedge dw - w d\bar{w} \wedge dz\right). \end{aligned} \tag{1.27}$$

Here is a more geometric description of these forms: (1) The symplectic connections of ω_f yields the standard connection on $(T^*\mathbb{H})^{k/2}$ obtained from the Levi–Civita connection on the hyperbolic plane and (2) the S^1 action which rotates the fibres is Hamiltonian with $H(z,w) = -\frac{2}{k} f(\mathrm{Im}(z)^k |w|^2)$ and the Marsden-Weinstein quotient $H^{-1}(-\frac{2}{k} r^2)/S^1$ is symplectomorphic to the hyperbolic plane scaled by $(1 - f(r^2))$. None of these symplectic forms extends over the whole space $(T^*\mathbb{H})^{k/2}$ and the restriction to a disc bundle is necessary. After calculating the moment map for the $\mathrm{SL}(2,\mathbb{R})$-action on the fibre X_k and simplifying the resulting equations we deduce from Theorem 8 the following moment map.

Theorem 13 (**Moment map on $\mathcal{D}^1_k(\Sigma)$**)**.** *The action of $Ham(\Sigma,\rho)$ on the space $\{(J,\tau) \in \mathcal{D}^1_k(\Sigma) \mid \bar{\partial}_J \tau = 0\}$ is Hamiltonian with respect to ω_f and generated by the moment map*

$$\underline{\mu}_f(J,\tau) = \left(2K_J + \Delta F(|\tau|^2_J) - 2c\right)\rho \tag{1.28}$$

where $c := 2\pi(2 - genus(\Sigma))/vol(\Sigma,\rho)$, $F : [0,1) \to \mathbb{R}$ is defined by

$$F(t) := \int_0^t \frac{f(s)}{ks} + f'(s)\, ds$$

and $\Delta = d^*d$ is the positive Laplacian of the metric $\rho(\cdot, J\cdot)$.

Proof. This is reformulated as Theorem A in the introduction of Chapter 5 and proved in Theorem 5.2.9. □

It is natural to ask if there exists a preferred symplectic structure $\omega_f \in \Omega^2(X_k)$. In the case $k = 2$ this is answered by the Feix–Kaledin hyperkähler metric. For $k > 2$ we were unable to find a satisfactory answer. However, we show that there are functions f_k for which the resulting moduli spaces admits a particularly simple description, namely

$$\mathcal{M}_{f_k}(k) \cong \left\{ (g,\tau) \;\middle|\; \begin{array}{c} g \in \mathrm{Met}(\Sigma),\ (J_g,\tau) \in \mathcal{D}^1_k(\Sigma) \\ \bar\partial\tau = 0,\ K_g - \frac{c}{k}|\tau|_g^2 = c\frac{k-1}{k} \end{array} \right\} \Big/ \mathrm{Diff}_0(\Sigma) \tag{1.29}$$

where $c := 2\pi(2-2\mathrm{genus}(\Sigma))/\mathrm{vol}(\Sigma,\rho)$ and for $g \in \mathrm{Met}(\Sigma)$ we denote by $J_g \in \mathcal{J}(\Sigma)$ the unique complex structure compatible with g. In the case $k=2$ this corresponds to the Donaldson's hyperkähler extension of Teichmüller space.

The discussion so far extends naturally to tuples $(J,\tau_1,\dots,\tau_n)$ of complex differentials of mixed order. For $\mathbf{k} = (k_1,\dots,k_n) \in \mathbb{Z}^n_{\geq 2}$ define

$$\mathcal{D}^1_{\mathbf{k}}(\Sigma) := \{(J,\tau_1,\dots,\tau_n) \,|\, (J,\tau_i) \in \mathcal{D}_{k_i}(\Sigma) \text{ for } i = 1,\dots,n\}. \tag{1.30}$$

Then, the main observation is that $\mathcal{D}^1_{\mathbf{k}}(\Sigma)$ embeds naturally into the product manifold $\prod_{i=1}^n \mathcal{D}^1_{k_i}(\Sigma)$ as a symplectic submanifold.

Chapter 6: The Ricci form and Calabi–Yau Teichmüller space

This chapter summarizes joint work with Oscar Garcia–Prada and Dietmar A. Salamon [50]. We show that the Ricci form yields a moment map for the action of the group of exact volume preserving diffeomorphims on the space of almost complex structures. This gives rise to an extended Weil–Petersson symplectic form on the Calabi–Yau Teichmüller space of isotopy classes of complex structures with real first Chern class zero and nonempty Kähler cone. We also discuss variants of the theory for Kähler–Einstein pairs which have not been included into our joint paper. The presentation in this chapter is rather brief and we only sketch the arguments for the more technical results. Full detail and complete proofs can be found in our joint article [50].

Let (M,ρ) be a closed $2n$-dimensional manifold with fixed volume form ρ. We define the Ricci form $\mathrm{Ric}_{\rho,J} \in \Omega^2(M)$ associated to the volume form ρ and an almost complex structure $J \in \mathcal{J}(M)$ by

$$\mathrm{Ric}_{\rho,J}(u,v) := \frac{1}{4}\mathrm{tr}\left((\nabla_u J)J(\nabla_v J)\right) + \frac{1}{2}\mathrm{tr}\left(JR^\nabla(u,v)\right) + \frac{1}{2}d\lambda_J^\nabla \tag{1.31}$$

for $u,v \in \mathrm{Vect}(M)$, where ∇ is a torsion free ρ-connection on M and the 1-form λ_J^∇ is defined by $\lambda_J^\nabla(u) := \mathrm{tr}\left((\nabla J)u\right)$ for $u \in \mathrm{Vect}(M)$. The next theorem can be derived as a special case of Donaldson's moment map in Theorem 8. In [50] we give a direct and independent proof of this result.

Theorem 14 (**Ricci form**, [50]). *The Ricci form $Ric_{\rho,J} \in \Omega^2(M)$ does not depend on the choice of the connection ∇ used to define it, represents the cohomology class $2\pi c_1(TM,J)$ and agrees with the usual definition of the Ricci form on Kähler manifolds. The map $J \mapsto 2Ric_{\rho,J}$ satisfies the moment map equation for the action of the exact volume preserving diffeomorphism group on the space of almost complex structures.*

Proof. This is reformulated as Theorem A in the introduction of Chapter 6 and proved in Theorem 6.2.1. □

A useful generalization of the moment map equation involves the 1-form $\Lambda_\rho \in \Omega^1(\mathcal{J}(M), \Omega^1(M))$ defined by

$$\Lambda_\rho(J, \hat{J})(u) := \mathrm{tr}\left((\nabla \hat{J})u + \frac{1}{2}\hat{J}J\nabla_u J\right) \tag{1.32}$$

for $u \in \mathrm{Vect}(M)$, where ∇ is a torsion free ρ-connection on M. Then, the linearisation of $\mathrm{Ric}_{\rho,J}$ when varying J in direction $\hat{J}$ is given by $\frac{1}{2}d\Lambda_\rho(J,\hat{J})$ and

$$\int_M \Lambda_\rho(J, \hat{J}) \wedge \iota(v)\rho = \frac{1}{2}\int_M \mathrm{tr}\left(\hat{J}J\mathcal{L}_v J\right)\rho \tag{1.33}$$

for all $v \in \mathrm{Vect}(M)$. This setup leads to a new construction of the Weil–Petersson symplectic form on the Calabi–Yau Teichmüller space

$$\mathcal{T}_0(M) := \left\{ J \in \mathcal{J}_{\mathrm{int}}(M) \;\middle|\; \begin{array}{c} c_1(TM, J) = 0 \in H^2(M, \mathbb{R}) \\ \text{and } J \text{ admits a Kähler form} \end{array} \right\} \Big/ \mathrm{Diff}_0(M). \tag{1.34}$$

This moduli space has been studied extensively in the polarized cased [64, 90, 98] and for $K3$-surfaces, see [44] Chapter 16. The Bogomolov–Tian–Todorov Theorem [11, 111, 113] asserts that $\mathcal{T}_0(M)$ is a smooth manifold. However, it is not Hausdorff in general [54, 120]. The construction of the Weil–Petersson metric involves three main steps:

1. The natural inclusion of

$$\mathcal{T}_0(M, \rho) := \{J \in \mathcal{J}_{\mathrm{int},0}(M) \,|\, \mathrm{Ric}_{\rho,J} = 0\} /\mathrm{Diff}_0(M, \rho) \tag{1.35}$$

into Teichmüller space $\mathcal{T}_0(M)$ is a bijection.

2. The group $\mathrm{Diff}_0(M,\rho)/\mathrm{Diff}^{\mathrm{ex}}(M,\rho)$ acts trivially on

$$\mathcal{T}_0^{\mathrm{ex}}(M, \rho) := \mathcal{J}_{\mathrm{int},0}(M,\rho)/\mathrm{Diff}^{\mathrm{ex}}(M,\rho). \tag{1.36}$$

Hence, $\mathcal{T}_0(M,\rho) = \mathcal{T}_0^{\mathrm{ex}}(M,\rho)$ embeds into the Marsden–Weinstein quotient of $\mathcal{J}(M)$ and carries a natural closed 2-form.

3. The subspace of integrable structures $\mathcal{J}_{\mathrm{int}}(M) \subset \mathcal{J}(M)$ is not a symplectic submanifold and it is not obvious that the closed 2-form on $\mathcal{T}_0(M,\rho)$ is non-degenerated. We give a complete characterization of the kernel of the restriction of the symplectic form which then proves non-degeneracy of the Weil–Petersson symplectic form on the quotient.

The tangent spaces at the space of integrable complex structures are

$$T_J\mathcal{J}_{\mathrm{int}}(M) = \ker\left(\bar{\partial}_J : \Omega_J^{0,1}(M, TM) \to \Omega_J^{0,2}(M, TM)\right). \tag{1.37}$$

If $\mathrm{Ric}_{\rho,J} = 0$ and $\bar{\partial}_J \hat{J} = 0$ then there exist smooth functions $f, g : M \to \mathbb{R}$ such that

$$\Lambda_\rho(J, \hat{J}) = -df \circ J + dg \tag{1.38}$$

Moreover, for every $J \in \mathcal{J}_{\mathrm{int}}(M)$ with vanishing real first Chern class and non-empty Kähler cone, there exists a unique volume form ρ_J with $\mathrm{Ric}_{\rho_J,J} = 0$ and $\int_M \rho_J = V$.

Theorem 15 (**Weil–Petersson symplectic form**, [50]). *The Weil–Petersson symplectic form on $\mathcal{T}_0(M,\rho)$ is given by*

$$\Omega_J(\hat{J}_1,\hat{J}_2) = \int_M \left(\frac{1}{2} tr\left(\hat{J}_1 J \hat{J}_2\right) - f_1 g_2 + f_2 g_1\right)\rho_J \tag{1.39}$$

for $J \in \mathcal{J}_{int}(M)$ with vanishing real first Chern class and non-empty Kähler cone, $\hat{J}_i \in \Omega^{0,1}(M,TM)$ with $\bar{\partial}_J \hat{J}_i = 0$ and f_i, g_i defined by (1.38). This symplectic form is $\mathrm{Diff}_0(M)$ equivariant and thus the mapping class group acts on $\mathcal{T}_0(M)$ by symplectomorphism.

Proof. This is reformulated as Theorem B in the introduction of Chapter 6 and the proved in Theorem 6.3.5. □

The Weil–Petersson symplectic form gives rise to a symplectic connection on the bundle $\mathcal{E}_0(M)$ of isotopy classes of Ricci-flat Kähler structures over the space $\mathcal{B}_0(M)$ of symplectic forms with vanishing first Chern class.

Theorem 16 (**A symplectic connection**, [50]). *The projection $\mathcal{E}_0(M) \to \mathcal{B}_0(M)$ is a submersion and for every Ricci flat Kähler structure (ω, J) on M and for every closed 2-form $\hat{\omega}$, there exists a unique element $\hat{J} = \mathcal{A}_{\omega,J}(\hat{\omega}) \in \Omega^{0,1}_J(M,TM)$ satisfying*

$$\Omega_J(\hat{J},\hat{J}') = 0 \qquad \text{for all } \hat{J}' \in \Omega^{0,1}_J(M,TM) \text{ with } \bar{\partial}_J \hat{J}' = 0 \text{ and } \hat{J}' = (\hat{J}')^*$$

and the tangency conditions

$$\bar{\partial}_J \hat{J} = 0, \quad \Lambda_\rho(J,\hat{J}) = -d\langle\hat{\omega},\omega\rangle \circ J, \quad \hat{\omega}(\cdot,\cdot) - \hat{\omega}(J\cdot,J\cdot) = \langle(\hat{J}-\hat{J}^*)\cdot,\cdot\rangle.$$

This connection is $\mathrm{Diff}_0(M)$-equivariant and satisfies $\mathcal{A}_{\omega,J}(d\iota(v)\omega) = \mathcal{L}_v J$ for all $v \in \mathrm{Vect}(M)$ with $d\iota(Jv)\rho = 0$.

Proof. This is reformulated as Theorem C in the introduction of Chapter 6 and Theorem 6.3.5. □

The final section discusses variants of the theory for Kähler–Einstein manifolds which have not been included into our joint paper. Fix a volume form $\rho \in \Omega^{2n}(M)$ and cohomology classes $a, c \in H^2(M)$ such that $2\pi c = \kappa a$ for some $\kappa \in \mathbb{R}$. Denote by $\mathcal{S}_a(M,\rho) \subset \Omega^2(M)$ the space of symplectic forms on M with volume form $\omega^n/n! = \rho$ and denote by $\mathcal{J}_c(M)$ the space of almost complex structures with $c_1(TM,J) = c$. We call $a \in H^2(M,\mathbb{R})$ a Lefschetz class when $\cdot \cup a^{n-1} : H^1(M,\mathbb{R}) \to H^{2n-1}(M,\mathbb{R})$ is an isomorphism. Then $\mathcal{S}_a(M,\rho)$ is a symplectic manifold with the Lefschetz symplectic form

$$\Omega_\omega(\hat{\omega}_1,\hat{\omega}_2) := \int_M \lambda_1 \wedge \lambda_2 \wedge \frac{\omega^{n-1}}{(n-1)!} \tag{1.40}$$

for $\omega \in \mathcal{S}_a(M,\rho)$ and exact 2-forms $\hat{\omega} \in \Omega^2(M)$ with $\hat{\omega} \wedge \omega^{n-1} = 0$, where $\lambda_i \in \Omega^1(M)$ satisfy $d\lambda_i = \hat{\omega}$ and $\lambda_i \wedge \omega^{n-1}$ is exact. The motivation for this symplectic form comes from a moment map description of the equation $\omega^n/n! = \rho$.

Theorem 17 (Kähler–Einstein pairs). *The action of* $\mathrm{Diff}^{ex}(M,\rho)$ *on the product space* $\mathcal{J}_c(M)\times\mathcal{S}_a(M,\rho)$ *is Hamiltonian for the product symplectic form*

$$\Omega_{J,\omega}((\hat{J}_1,\hat{\omega}_1),(\hat{J}_2,\hat{\omega}_2)) := \int_M \frac{1}{2}tr\left(\hat{J}_1 J\hat{J}_2\right)\rho - 2\kappa\lambda_1\wedge\lambda_2\wedge\frac{\omega^{n-1}}{(n-1)!} \tag{1.41}$$

where $\lambda_i\in\Omega^1(M)$ *satisfy* $d\lambda_i=\hat{\omega}$ *and* $\lambda_i\wedge\omega^{n-1}$ *is exact. A moment map for this action is* $\mu:\mathcal{J}_c(M)\times\mathcal{S}_a(M,\rho)\to\Omega^2_{ex}(M)$ *defined by*

$$\mu(J,\omega)=2(Ric_{\rho,J}-\kappa\omega). \tag{1.42}$$

Proof. This is reformulated as Theorem D in the introduction of Chapter 6 and proved in Theorem 6.4.3. □

This leads to a Weil–Petersson metric on the Teichmüller space of Kähler–Einstein manifolds with a fixed symplectic form $\omega\in\mathcal{S}_a(M)$. Although this yields a new perspective on the subject, the symplectic form has been studied extensively, see [71, 98, 105] and the references therein.

Chapter 7: Universal Hitchin moduli spaces

This chapter contains joint work with Oscar Garcia–Prada, Luis Álvarez-Consul and Mario Garcia-Fernandez. We investigate variants of Hitchin's equations [58] on a Riemann surface Σ. In contrast to the classical theory, we do not fix the complex structure on the surface and investigate moment maps for the action of the extended gauge group. This yields various universal Hitchin moduli spaces which fibre naturally over Teichmüller space with fibre being the corresponding Hitchin moduli space. Most of the material is still work in progress and has not yet been explored in full detail.

Let (Σ,ρ) be a closed 2-dimensional surface equipped with an area form ρ. For a principal bundle $P\to\Sigma$ the extended gauge group $\tilde{\mathcal{G}}(P)$ of P consists of bundle isomorphisms covering Hamiltonian diffeomorphisms of Σ. Every connection $A\in\mathcal{A}(P)$ defines a splitting

$$\mathrm{Lie}(\tilde{\mathcal{G}}(P))\cong\Omega^0(\Sigma,\mathrm{ad}(P))\oplus\{v\in\mathrm{Vect}(\Sigma)\,|\,d\iota(v)\rho\text{ is exact}\} \tag{1.43}$$

which we denote by $\tilde{v}\mapsto(A(\tilde{v}),\pi_*\tilde{v})$. This is given by decomposing $\mathrm{Lie}(\tilde{\mathcal{G}}(P))\subset\mathrm{Vect}(P)$ into its A-horizonal and A-vertical component.

Real reductive groups. Suppose $G=(G,H,\theta,B)$ is a real reductive group, i.e. a quadruple consisting of a real Lie group G with reductive Lie algebra $\mathfrak{g}$, a maximal compact subgroup $H\subset G$, a Cartan involution $\theta:\mathfrak{g}\to\mathfrak{g}$ which defines a splitting $\mathfrak{g}=\mathfrak{h}\oplus\mathfrak{m}$ and a θ- and G-invariant bilinear form $B:\mathfrak{g}\times\mathfrak{g}\to\mathbb{R}$. The adjoint representation of G restricts to the so-called isotropy representation $\iota:H\to\mathrm{Aut}(\mathfrak{m})$. Now let $P\to\Sigma$ be a principal H bundle and denote by $P(\mathfrak{m}):=P\times_\iota\mathfrak{m}$ the associated $\mathfrak{m}$-bundle. In this setting, the holomorphicity condition of the Higgs field has no interpretation in terms of moment maps and we consider the configuration space

$$\mathcal{X}_1:=\left\{(J,A,\phi)\;\middle|\;\begin{array}{c}J\in\mathcal{J}(\Sigma),\,A\in\mathcal{A}(P)\\ \phi\in\Omega^{1,0}_J(\Sigma,P(\mathfrak{m})\otimes\mathbb{C}),\,\bar{\partial}_{A,J}\phi=0\end{array}\right\}. \tag{1.44}$$

This carries a natural symplectic structure obtained from B.

Theorem 18. *The natural $\tilde{\mathcal{G}}(P)$-action on $\mathcal{X}_1$ is Hamiltonian with moment map*

$$\langle \mu(J, A, \phi), \tilde{v}\rangle = -\int_\Sigma B\left(A(\tilde{v}),\, F_A - [\phi^* \wedge \phi]\right) + \int_\Sigma H\left(2K_J - 2c\right)\rho$$

where $\tilde{v} \in Lie(\tilde{\mathcal{G}}(P))$, with $\pi_*\tilde{v} \in Vect(\Sigma)$ and $A(\tilde{v}) \in \Omega^0(\Sigma, ad(P))$ defined by (1.43) and $\pi_*\tilde{v} = v_H$ is the Hamiltonian vector field for $H : \Sigma \to \mathbb{R}$, and $c := 2\pi(2g - 2)/vol(\Sigma, \rho)$.

Proof. This is reformulated as Theorem A in the introduction of Chapter 7 and proved in Proposition 7.3.2 and Theorem 7.3.1. □

Complex reductive groups. Assume that G is a compact group and $P \to \Sigma$ a principal G bundle. We consider the space

$$\mathcal{X}_2 := \mathcal{J}(\Sigma) \times \mathcal{A}(P) \times \Omega^1(\Sigma, \mathbf{i}\mathrm{ad}(P)). \tag{1.45}$$

For every fixed J, there is a natural isomorphism

$$\mathcal{A}(P) \times \Omega^1(\Sigma, \mathbf{i}\mathrm{ad}(P)) \cong \mathcal{A}(P) \times \Omega^{0,1}_J(\Sigma, \mathbf{i}\mathrm{ad}(P)) \tag{1.46}$$

and both spaces can be identified with $T^*\mathcal{A}(P)$. The later model carries a natural hyperkähler structure which gives rise to three symplectic forms on $\mathcal{X}_2$.

Theorem 19. *The action of $\tilde{\mathcal{G}}(P)$ on $\mathcal{X}_2$ is Hamiltonian for all three symplectic forms with moment maps*

$$\mu_1(J, A, \psi) = \left(\left(F_A + \frac{1}{2}[\psi \wedge \psi]\right),\, (2K_J - 2c)\rho + dtr\left(\psi\Lambda_\rho(d_A\psi)\right)\right)$$

$$\hat{\mu}_2(J, A, \psi) = \left(\boldsymbol{i}d_A * \psi,\, (2K_J - 2c)\rho + \boldsymbol{i}d * tr\left(\psi\Lambda_\rho(F_A)\right)\right)$$

$$\mu_3(J, A, \psi) = \left(\boldsymbol{i}d_A\psi,\, (2K_J - 2c)\rho + \boldsymbol{i}dtr\left(\psi\Lambda_\rho\left(F_A + \frac{1}{2}[\psi \wedge \psi]\right)\right)\right)$$

where $c := 2\pi(2genus(\Sigma) - 2)/vol(\Sigma, \rho)$ and

$$\Lambda_\rho : \Omega^2(\Sigma, ad(P) \otimes \mathbb{C}) \to \Omega^0(\Sigma, \boldsymbol{i}ad(P) \otimes \mathbb{C})$$

is the natural map induced by ρ. All three moment maps take values in the space $\Omega^2(\Sigma, ad(P)) \oplus \Omega^2_{ex}(\Sigma)$ which is the dual space of $Lie(\tilde{\mathcal{G}}(P))$ by (1.43).

Proof. This is reformulated as Theorem B in the introduction of Chapter 7 and proved in Theorem 7.4.5. □

Note that this theorem does not quite yield a hyperkähler moment map. Nevertheless, we have

$$(J, A, \psi) \in \mu_1^{-1}(0) \cap \hat{\mu}_2^{-1}(0) \cap \mu_3^{-1}(0) \qquad \Longleftrightarrow \qquad \left\{ \begin{array}{c} d_A\psi = 0,\, d_A^*\psi = 0 \\ F_A + \frac{1}{2}[\psi \wedge \psi] = 0 \\ 2K_J = c \end{array} \right.$$

After taking the action of the flux group into account, this yields a moduli space which fibres over Teichmüller space with the corresponding Hitchin moduli space as fibre.

Fibrations over Donaldson's moduli space. Consider the configuration space

$$\mathcal{X}_3 := \mathcal{Q}_1(\Sigma) \times \mathcal{A}(P) \times \Omega^1(\Sigma, \mathbf{i}\mathrm{ad}(P)). \tag{1.47}$$

where $\mathcal{Q}_1(\Sigma)$ denotes the space of pairs (J, σ) consisting of a complex structure $J \in \mathcal{J}(\Sigma)$ and a quadratic differential $\sigma \in \Omega^0(\Sigma, S^2(T^*\Sigma \otimes_J \mathbb{C}))$ with pointwise norm $|\sigma|_J < 1$. Donaldson [38] observed that the space $\mathcal{Q}_1(\Sigma)$ carries a hyperkähler structure whose hyperkähler quotient (after taking the action of the flux group into account) yields the Feix–Kaledin hyperkähler extension $\mathcal{M}$ of Teichmüller space. It is reasonable to expect that the space $\mathcal{X}_3$ carries three symplectic forms for which the action of the extended gauge group is Hamiltonian. This should give rise to a moduli space which fibres over $\mathcal{M}$ with fibres being the corresponding Hitchin moduli spaces.

This is still work in progress and has not yet been written up. There are two intriguing aspects which we would like to mention: First, there is a construction of Donaldson [38] which associates to every element in $\mathcal{M}$ a solution of the SU(2) Hitchin equations over Σ and thus $\mathcal{M}$ really parametrizes pairs of solutions to Hitchin's equation. Second, the resulting moduli space is naturally a hyperkähler fibration over the hyperkähler space $\mathcal{M}$. It is probably too optimistic to expect that they combine to a hyperkähler structure on the whole moduli space, but this is certainly something to be investigated more closely.

Chapter 2

The GIT picture for the Yang-Mills equations over Riemann surfaces

The content of this chapter has been published in [114]. We give a self-contained exposition of the Atiyah-Bott picture [4] for the Yang-Mills equation over Riemann surfaces with an emphasis on the analogy to finite dimensional geometric invariant theory. The main motivation is to provide a careful study of the semistable and unstable orbits: This includes the analogue of the Ness uniqueness theorem for Yang-Mills connections, the Kempf-Ness theorem, the Hilbert-Mumford criterion and a new proof of the moment-weight inequality following an approach outlined by Donaldson [39]. A central ingredient in our discussion is the Yang-Mills flow for which we assume longtime existence and convergence (see [97]).

2.1 Introduction

The aim of this chapter is threefold: The first goal is to provide a self-contained and essentially complete exposition of the geometric invariant theory for the Yang-Mills equation over Riemann surfaces from the differential geometric point of view. We follow closely the line of arguments of finite dimensional GIT (e.g. as it is explained in [51]) and emphasize this analogy throughout.

The second goal is to include a careful study of the semistable and unstable orbits. This is in contrast to most of the developments after the landmark paper [4] of Atiyah and Bott, which deal with the characterization of stable objects in more general moduli problems, i.e. the analogue of the Narasimhan-Seshadri theorem. In the unitary case Daskalopoulos [27] established the Morse theoretic picture of Atiyah and Bott. A direct corollary of this stratification is the analogue of the Ness uniqueness theorem and the moment limit theorem (see Theorem A below). We present an alternative proof of this result following the arguments discovered by Calabi-Chen [17] and Chen-Sun [23] in a different infinite dimensional setting. This argument does not depend on the Harder-Narasimhan filtration or on other aspects from the holomorphic point of view and works for general structure groups. Following an approach outlined by Donaldson [39], we also carry out a new proof of the moment-weight inequality which is essentially contained in the work of Atiyah and Bott.

The third goal is to provide a transparent exposition of the central ideas used in gauge theoretical moduli problems. While several results are known in greater generality, the key ideas are still immanent in our treatment. We hope that this enables non experts to explore the beauty of this subject without having to worry about the technical difficulties which come along with more general moduli problems.

We concentrates on the stability questions in Yang–Mills theory and do not discuss the topology of the resulting moduli space, which is one of the main topics in the work of Atiyah and Bott. There is no claim of originality (except to my knowledge Theorem A has only been proven in the case $G = U(n)$ in the existing literature). However, the various results and underlying ideas are spread over the literature and the present paper provides a unified exposition. The main technical ingredients in our discussion are long time existence and convergence of the Yang–Mills flow. The presented arguments allow for various generalizations to moduli problems in gauge theory, where the main obstacles are again long time existence and convergence of the relevant parabolic gradient flow. These obstructions can be overcome for the Yang–Mills–Higgs flow under suitable assumptions and the results of this article can be carried over to the symplectic vortex equation over Riemann surfaces, see Chapter 3 and [119]. For the extension of the theory to bundles over higher dimensional Kähler manifolds, the situation is more delicate and various known results are discussed at the end of the introduction.

There are two essentially different perspectives on GIT - the algebraic geometric and the symplectic point of view. The recent survey of Thomas [109] provides some background from both perspectives and explores several finite and infinite dimensional examples. Originally, Mumford [87] introduced GIT as a method to construct quotients and moduli spaces in algebraic geometry. The work of Atiyah-Bott [4] and the thesis of Kirwan [69] have shown that GIT is closely related to moment maps and symplectic reduction, where the link between both theories lies in the Morse-Bott stratification of the moment map squared functional. This leads to an entirely differential geometric version of GIT. Another important ingredient in this approach is the Kempf-Ness function: Let (X, J, ω, μ) be a closed Kähler manifold with Hamiltonian G-action and moment map μ. Here G denote a (real) compact Lie group with complexification G^c. For a given point $x \in X$ there exists a G-invariant function

$$\Phi_x : G^c/G \to \mathbb{R}$$

such that the gradient flow of Φ_x intertwines with the gradient flow of the moment map squared functional under the map $g \mapsto g^{-1}x$. The global analytic properties of Φ_x are related to the algebraic weights of x and to the solvability of the equation $\mu(gx) = 0$ by the Kempf-Ness theorem.

We follow throughout this survey the differential geometric approach. For a modern algebraic treatment we refer to [3] and the references therein. The new edition of [88] also contains a discussion of the GIT picture for the Yang-Mills equations.

Nevertheless, it leaves some refined question open: What are the appropriate analoguos versions of the Ness uniqueness theoerem, the Kempf-Ness theorem or the Hilbert-Mumford criterion? The analog of the Kempf-Ness functional has been used to provide analytic proofs for various generalizations of the Narasimhan-Seshadri theorem, but it has seen little discussion beyond these applications in the literature. The recent work of Calabi, Chen, Donaldson and Sun [17, 39, 18, 19, 23] has shown that the underlying geometric properties of the Kempf-Ness functional can be used to provide analytic proofs for the Ness uniqueness theorem and the Kempf-Ness theorem. We follow their ideas and obtain new proofs of the corresponding results in the Yang-Mills case. The exposition [51] provides a finite dimensional discussion of these arguments.

Main results

Let G be a compact connected Lie group and let Σ be a closed Riemann surface. Fix a volume form on Σ, compatible with the orientation, and let $P \to \Sigma$ be a principal G bundle. Atiyah and Bott [4] observed that the curvature

$$\mu(A) := *F_A \in \Omega^0(\Sigma, \mathrm{ad}(P))$$

defines a moment map for the action of the gauge group $\mathcal{G}(P)$ on the space of connections $\mathcal{A}(P)$. For any constant central section τ, the symplectic quotient

$$\mathcal{A}(P)//\mathcal{G}(P) := \mu^{-1}(\tau)/\mathcal{G}(P)$$

yields the moduli space of projectively flat connections on P with constant central curvature τ.

Let G^c be the complexification of G and $P^c := P \times_G G^c$ the associated principal G^c bundle. The complexification of the gauge group is $\mathcal{G}^c(P) := \mathcal{G}(P^c)$. The space $\mathcal{A}(P)$ can naturally be identified with the space $\mathcal{J}(P^c)$ of holomorphic structures on P^c (see Lemma 2.2.5) and the complexified gauge group $\mathcal{G}^c(P)$ acts naturally on this space. The corresponding GIT quotient

$$\mathcal{A}^{ss}(P)//\mathcal{G}^c(P)$$

of $\mathcal{A}(P)$ by $\mathcal{G}^c(P)$ is obtained in two steps. First, one defines a dense and open subset $\mathcal{A}^{ss}(P) \subset \mathcal{A}(P)$ of semistable connections or holomorphic structures on P and second, one identifies two semistable orbits in the quotient if they cannot be separated in $\mathcal{A}^{ss}(P)$. The restriction to semistable orbits is necessary to obtain a *good quotient* in the sense of algebraic geometry. There are two approaches to define semistable objects. In the symplectic approach, one chooses a moment map for the gauge action on $\mathcal{A}(P)$ to define semistable objects. In the algebraic geometric approach, one defines a notion of semistability $\mathcal{J}^{ss}(P^c) \subset \mathcal{J}(P^c)$ on the space of holomorphic structures on P^c. A classical result due to Narasimhan and Seshadri [91] in the case $G = U(n)$ and due to Ramanathan [95] for general G shows that both of these notions agree if one restricts to further open subsets of stable objects.

The Yang-Mills picture introduced by Atiyah and Bott [4] shed new light on this result and inspired Donaldson [30] to an analytic proof of the Narasimhan-Seshadri theorem. The Yang-Mills functional is given by the formula

$$\mathcal{YM} : \mathcal{A}(P) \to \mathbb{R}, \qquad \mathcal{YM}(A) := \frac{1}{2}\int_\Sigma ||F_A||^2 \, \mathrm{dvol}_\Sigma.$$

Standard arguments from Chern-Weil theory show that there exists a unique central element $\tau \in Z(\mathfrak{g})$ such that

$$\mathcal{YM}(A) = \inf_{B\in\mathcal{A}(P)} \mathcal{YM}(B) \qquad \Longleftrightarrow \qquad *F_A = \tau. \tag{2.1}$$

We shall consider in the following connections of Sobolev class $W^{1,2}$ and gauge transformations of Sobolev class $W^{2,2}$. Råde [97] showed in this setting that for every initial data $A_0 \in \mathcal{A}(P)$ the negative gradient flow of the Yang-Mills functional

$$\partial_t A(t) = -\nabla\mathcal{YM}(A(t)) = -d^*_{A(t)} F_{A(t)}, \qquad A(0) = A_0 \tag{2.2}$$

has a unique (weak) solution which exists for all time. Moreover, this solution remains in a single complexified $\mathcal{G}^c(P)$-orbit and converges in the $W^{1,2}$-topology to a Yang-Mills connection $A_\infty \in \overline{\mathcal{G}^c(A_0)}$. The following is the analogue of the Ness uniqueness theorem in finite dimensional GIT.

Theorem A (**Uniqueness of Yang-Mills connections**). *Let $A_0 \in \mathcal{A}(P)$ and let A_∞ be the limit of the Yang-Mills flow (2.2) starting at A_0. Then*

1. $\mathcal{YM}(A_\infty) = \inf_{g\in\mathcal{G}^c(P)} \mathcal{YM}(gA)$.
2. *If $B \in \overline{\mathcal{G}^c(A_0)}$ is contained in the $W^{1,2}$-closure of $\mathcal{G}^c(A_0)$ and*

$$\mathcal{YM}(B) = \inf_{g\in\mathcal{G}^c(P)} \mathcal{YM}(gA)$$

then $\mathcal{G}(B) = \mathcal{G}(A_\infty)$.

In the case $G = U(n)$ one can replace P by a hermitian vector bundle $E \to \Sigma$. Daskalopoulos [27] established in this case the convergence of the Yang-Mills flow over Riemann surfaces by different methods. He proves a suitable slice theorm near Yang-Mills connections and shows that the limiting Yang-Mills connection A_∞ is determined up to a unitary gauge transformation by the isomorphism class of the Harder-Narasimhan filtration of $(E, \bar\partial_{A_0})$. This proves Theorem A in the unitary case and it should be possible to deduce the general result from this using the methods in [10]. We present a different proof of Theorem A in Theorem 2.4.14 and Theorem 2.4.15 by following the line of arguments from finite dimensional GIT ([51], Chapter 6). These arguments were originally given by Calabi-Chen [17] and Chen-Sun [23] in the context of extremal Kähler metrics.

A connection $A \in \mathcal{A}(P)$ is called μ_τ-semistable resp. μ_τ-unstable if

$$\inf_{g\in\mathcal{G}^c(P)} || * F_{gA} - \tau||_{L^2} = 0 \qquad \text{resp.} \qquad \inf_{g\in\mathcal{G}^c(P)} || * F_{gA} - \tau||_{L^2} > 0$$

where τ is defined by (2.1). Moreover, A is called μ_τ-polystable if there exists $g \in \mathcal{G}^c(P)$ with $*F_{gA} = \tau$ and it is called μ_τ-stable if gA is in addition irreducible. Then Theorem A implies that the map which sends $A_0 \in \mathcal{A}(P)$ to the limit A_∞ of the Yang-Mills flow starting at A_0 yields the identifications

$$\mathcal{A}^{ss}(P)//\mathcal{G}^c(P) \cong \mathcal{A}^{ps}(P)/\mathcal{G}^c(P) \cong \mu^{-1}(\tau)/\mathcal{G}(P).$$

Conversely, the μ_τ-unstable orbits converge to higher critical points of the Yang-Mills functional. More details on this correspondence are given in Theorem 2.4.18.

The theory has greatly evolved since the paper [4] of Atiyah and Bott. The main goal in those developments has been the characterization of stable objects in more general moduli problems (e.g. [31], [32], [118], [58], [102], [12]). The characterization of unstable orbits is in general much more difficult as it refers to higher critical points of the Yang-Mills functional. Given a connection $A \in \mathcal{A}(P)$ and $\xi \in \Omega^0(\Sigma, \mathrm{ad}(P))$ the weight $w_\tau(A, \xi)$ is defined by

$$w_\tau(A, \xi) := \lim_{t \to \infty} \langle *F_{e^{\mathrm{i}t\xi}A} - \tau, \xi \rangle \in \mathbb{R} \cup \{\infty\}.$$

The first part of the following theorem is the analogue of the moment-weight inequality and the last two claims are the analogue of the Kempf existence and uniqueness theorem in finite dimensional GIT.

Theorem B (Atiyah-Bott). *Let $A \in \mathcal{A}(P)$ and let $\tau \in Z(\mathfrak{g})$ be defined by (2.1). Then*

1. *For all $0 \neq \xi \in \Omega^0(\Sigma, ad(P))$ it holds*

$$-\frac{w_\tau(A, \xi)}{||\xi||} \leq \inf_{g \in \mathcal{G}^c(P)} || * F_A - \tau||^2.$$

2. *If the right-hand-side is positive, then there exists up to scaling a unique $0 \neq \xi_0 \in \Omega^0(\Sigma, ad(P))$ such that*

$$-\frac{w_\tau(A, \xi_0)}{||\xi_0||} = \inf_{g \in \mathcal{G}^c(P)} || * F_A - \tau||^2.$$

   Moreover, ξ_0 is rational in the sense that it generates a closed one parameter subgroup of $\mathcal{G}(P)$.

3. *Let A_∞ be the limit of the Yang-Mills flow starting at A_0. Then there exists $u \in \mathcal{G}(P)$ such that ξ_0 agrees up to scaling with $u(*F_{A_\infty} - \tau)u^{-1}$.*

This is essentially contained in the work of Atiyah and Bott ([4], Prop. 8.13 and Prop. 10.13). A connection $A \in \mathcal{A}(P)$ induces a holomorphic structure on the complexified bundle $P^c := P \times_G G^c$ and its Lie algebra bundle $\mathrm{ad}(P^c)$. Atiyah and Bott explicitly determine the infimum of the Yang-Mills functional over $\mathcal{G}^c(A)$ in terms of the Harder-Narasimhan filtration of $\mathrm{ad}(P^c)$. The analogous result has been shown by Calabi, Chen, Donaldson and Sun [17, 39, 18, 19, 23] in the context of extremal Kähler metrics. Donaldson [39] compares the Atiyah-Bott picture in the vector bundle case $G = U(n)$ with their results on the Calabi functional and mentions that their methods should lead to a new proof of the moment-weight inequality in the Atiyah-Bott case. We carry out this proof in Theorem 2.5.12. We reformulate and prove the last two claims in Theorem 2.7.1. The case $G = U(n)$ follows along the line of arguments of Atiyah and Bott from the Harder-Narasimhan filtration and the Narasimhan-Seshadri theorem. The general case can be reduced to this by the use of Theorem A. For this, choose a faithful representation $G \hookrightarrow U(n)$. Then any G-connection A can be considered as $U(n)$-connection and Theorem A implies

$$\inf_{g \in \mathcal{G}^c(E)} \mathcal{YM}(gA) = \inf_{g \in \mathrm{GL}(E)} \mathcal{YM}(gA).$$

It now remains to compare the weights for the gauge action with respect to the two structure groups G and $U(n)$ to conclude the proof. We would also like to mention the work of Bruasse and Teleman [15, 14]. They prove for more general gauge theoretical moduli problems that whenever the supremum over the normalized weights is positive, then it is attained in a unique direction corresponding to the Harder-Narasimhan filtration.

There is a classical algebraic geometric notion of stability for holomorphic principal bundles (see Definition 2.3.2). In the vector bundle case $G = U(n)$ this corresponds to the notion of (slope-)stable holomorphic vector bundles, which are easier to define: A holomorphic vector bundle E is called stable (semistable) if

$$\frac{c_1(F)}{\mathrm{rk}(F)} < \frac{c_1(E)}{\mathrm{rk}(E)} \qquad \left(\frac{c_1(F)}{\mathrm{rk}(F)} \leq \frac{c_1(E)}{\mathrm{rk}(E)}\right)$$

holds for every proper holomorphic subbundle $0 \neq F \subset E$. Moreover, E is called polystable if it decomposes as the direct sum of stable vector bundles all having the same slope and E is called unstable if it is not semistable.

Theorem C (Generalized Narasimhan-Seshadri-Ramanathan theorem). *Let $A \in \mathcal{A}(P)$ and define τ by (2.1). Then A induces a holomorphic structure J_A on the complexified bundle $P^c := P \times_G G^c$ and the following holds true:*

1. *(P^c, J_A) is stable if and only if A is μ_τ-polystable and the kernel of the infinitesimal action $L_A : \Omega^0(\Sigma, ad(P^c)) \to \Omega^1(\Sigma, ad(P))$*

   $$L_A(\xi + \boldsymbol{i}\eta) := -d_A\xi - *d_A\eta$$

   contains only constant central sections.

2. *(P^c, J_A) is polystable if and only if A is μ_τ-polystable.*

3. *(P^c, J_A) is semistable if and only if A is μ_τ-semistable.*

4. *(P^c, J_A) is unstable if and only if A is μ_τ-unstable.*

Proposition 2.5.9 characterizes the stability of (P^c, J_A) in terms of the weights $w_\tau(A, \xi)$ and shows that this theorem is the appropriate analog of the Hilbert-Mumford criterion in finite dimensional GIT. The first claim is the Narasimhan-Seshadri-Ramanathan theorem. We present an analytic proof of this classical result in Theorem 2.6.5 which was originally given by Bradlow [12] and Mundet [89] for more general moduli problems. The main step in their proof is to establish the analogue of the Kempf-Ness theorem (see Theorem 2.6.2) in the stable case. The polystable case is deduced from the stable case by induction on the dimension of G. The unstable and semistable case follow directly from Theorem B by Proposition 2.5.9. We reformulate and prove Theorem C in Theorem 2.3.10.

Outline

In Section 2 we review the necessary preliminaries. The first part deals with the relevant background on gauge theory. Besides fixing notation, the main goals are to provide an explicit description of the complexified gauge action in both the vector bundle and principal bundle case and to describe the moment map picture of Atiyah and Bott. We show that this picture remains valid if one considers connections and gauge transformations in suitable Sobolev completions. The second part discusses parabolic subgroups of complex reductive Lie groups. These play a crucial role in the algebraic geometric definition of stability and the geometric description of the weights.

In Section 3 we discuss the algebraic and symplectic definitions of stability. The main result in this section is the generalized Narasimhan-Seshadri-Ramanathan theorem (Theorem 2.3.10) which states that these definitions are essentially equivalent. The proof of this theorem is based on the whole remainder of the exposition.

In Section 4 we review the analytical properties of the Yang-Mills flow which Råde [97] established in his thesis. We prove Theorem A in Theorem 2.4.14 and Theorem 2.4.15 and close this section with Theorem 2.4.18 which characterizes the μ_τ-stability of a connection $A \in \mathcal{A}(P)$ in terms of the limit A_∞ of the Yang-Mills flow starting at A.

In Section 5 we introduce the weights $w_\tau(A, \xi)$ and show that they are closely related to holomorphic parabolic reductions of the complexified bundle (P^c, J_A). Proposition 2.5.9 shows that the weights provide an alternative describes of the algebraic notion of stability. We close this section with the proof of the moment weight inequality (Theorem 2.5.12) following the approach outlined by Donaldson [39].

In Section 6 we describe a general procedure which associates to a given connection $A \in \mathcal{A}(P)$ a $\mathcal{G}(P)$-invariant functional $\Phi_A : \mathcal{G}^c(P) \to \mathbb{R}$. We call this the Kempf-Ness functional of A. The slope of this functional at infinity agrees with the weights discussed in Chapter 5 and hence relates to the algebraic notion of stability by Proposition 2.5.9. The analogue of the Kempf-Ness theorem (see Theorem 2.6.2) relates the global behavior of Φ_A to the symplectic μ_τ-stability of A. This provides a link between the algebraic and symplectic notions of stability and leads to an analytic proof of the Narasimhan-Seshadri-Ramanathan theorem in Theorem 2.6.5. These arguments are given by Bradlow [12] and Mundet [89] in more general settings.

In Section 7 we establish the analogue of the Kempf existence and uniqueness theorem (see Theorem 2.7.1). We include a self-contained account on the Harder-Narasimhan filtration for the convenience of the reader.

Higher dimensional base manifolds

We restrict our discussion to the case where Σ is a Riemann surface, although several results remain valid in greater generality. The main reason for this is to simplify the presentation. Let us indicate in the following to which degree the discussion could be generalized.

Replace Σ by a closed Kähler manifold (X, J, ω) and denote by

$$\Lambda : \Omega^{1,1}(X) \to \Omega^0(X)$$

the adjoint operator of $f \mapsto f\omega$. The Hermitian Einstein equation is given by

$$\Lambda F_A = \tau$$

for some constant central element $\tau \in \Omega^0(X, \mathrm{ad}(P))$. Denote by $\mathcal{A}^{1,1}$ the space of connections on P whose curvature F_A is of type $(1,1)$. This space can be given a Kähler structure and $\mu(A) = \Lambda F_A$ yields a moment map for the gauge action. In the vector bundle case, the Narasimhan-Seshadri theorem has been generalized to this setting by Donaldson [31, 32] in the algebraic framework and by Uhlenbeck and Yau [118] in the analytic framework over arbitrary Kähler manifolds. We would like to point out an observation by Anouche and Biswan [2]. They show that a holomorphic principal bundle P^c is polystable (resp. semistable), if and only if the associated holomorphic vector bundle $\mathrm{ad}(P^c)$ is polystable (resp. semistable). Further generalizations involving more complicated moduli problems have been studied by Hitchin [58], Simpson [102] and Bradlow [12]. In his thesis [89], Mundet generalizes this correspondence to a very general moduli problem.

Our discussion of the Yang-Mills flow in Chapter 4 relies heavily on the fact that Σ is a Riemann surface. In particular, the group of $W^{2,2}$ gauge transformations acts no more continuously on the space of $W^{1,2}$ connections for higher dimensional base manifolds. To avoid this issue, one could consider the flow directly on the space of smooth connections. Donaldson showed in [31] that the Yang-Mills flow starting at smooth $\mathcal{A}^{1,1}$ connections admits a smooth solution which exists for all time. In the stable case, Donaldson used this flow to prove his extension of the Narasimhan-Seshadri theorem. See [106] for a survey on this approach. The main issue is the complicated limiting behavior of solutions which yields profound technical difficulties. Bando and Siu ([6], Theorem 4) showed that the limit "breaks up" into Hermitian-Einstein sheaves in the unstable case and conjectured that the limit corresponds essentially to the Harder-Narasimhan filtration. This is very similar to our discussion in Chapter 7. The Bando-Siu conjecture has been confirmed by Daskalopoulos-Wentworth [28] in the case of Kähler surfaces and by Sibling [100] and Jacob [65, 66] for general Kähler manifolds. This yields the analogue of Theorem C for vector bundles over Kähler manifolds.

Our calculation of the weights in Chapter 5 remains valid over arbitrary Kähler manifold. However, the weakly holomorphic filtration yields in this case only a filtration by torsion-free subsheaves. The proof of the moment-weight inequality generalizes ad verbatim to this case. The proof which we present for the Narasimhan-Seshadri-Ramanathan theorem remains valid in this setting as well (see [89]).

The Harder-Narasimhan filtration is well defined for holomorphic vector bundles over Kähler manifolds, but consists of torsion-free subsheaves instead of holomorphic subbundles. It corresponds again to the supremum over the normalized weights. This is shown by Bruasse [14] and we present part of his argument in Chapter 7. It is a nontrivial result that the infimum of $||\Lambda F_{gA}||$ over the (smooth) complexified gauge orbit yields the same value and follows from the Bando-Siu conjecture. Bruasse gives an alternative and direct argument to prove that the supremum is in fact attained.

General assumptions

Let G be a compact connected (real) Lie group, Σ a closed Riemann surface and $P \to \Sigma$ a principal G bundle. We fix a volume form $dvol_\Sigma$ on Σ and assume for convenience that the volume form is scaled such that

$$\mathrm{vol}(\Sigma) = 1.$$

Note that the volume form also induces a fixed Riemannian metric on Σ.

Unless stated otherwise, all Lie groups are assumed to be connected. When G is a compact connected Lie group, then its complexification G^c, its parabolic subgroups $Q(\zeta)$ and their Levi subgroups $L(\zeta)$ are automatically connected (see Lemma 2.2.12).

As a general rule, we consider connections of Sobolev class $W^{1,2}$ and gauge transformations of Sobolev class $W^{2,2}$. The gauge action extends smoothly over these Sobolev spaces, since the base manifold is a Riemann surface. These regularity assumptions do not affect the overall picture and we shall discuss them in more detail in the preliminaries below.

2.2 Preliminaries

First, we review the underlying notions from gauge theory and set up our notation. The main goal is to describe the complexification of the gauge action and the moment map picture of Atiyah and Bott. We also discuss the regularity assumptions which are crucial for our further analytic discussion. In the second subsection, we describe parabolic subgroups of complex reductive Lie groups. We also include a brief discussion of the root space decomposition of semisimple Lie algebras for the sake of completeness.

2.2.1 Gauge theory

We consider throughout this section fiber bundles over a closed connected Riemann surface Σ.

Basic gauge theory

We start with the general framework of fiber bundles and specialize our discussion afterwards to the cases of vector bundles and principal bundles.

Fiber bundles. Let E, F and B be smooth manifolds. The manifold E together with a projection map $\pi : E \to B$ is called a fiber bundle over B with fiber F, if for every $x \in B$ there exists a neighborhood $x \in U \subset B$ and a diffeomorphisms

$$\psi : \pi^{-1}(U) \to U \times F$$

such that $\mathrm{pr}_1 \circ \psi = \pi|_U$. Here $\mathrm{pr}_1 : U \times F \to U$ denotes the projection onto the first factor. The map ψ is called a local trivialization of the fiber bundle E. Suppose ψ_α and ψ_β are local trivializations over U_α and U_β. Then there exists a unique map $g_{\beta\alpha} : U_\alpha \cap U_\beta \to \mathrm{Diff}(F)$ satisfying

$$\psi_{\beta\alpha}(x, u) := (\psi_\beta \circ \psi_\alpha^{-1})(x, u) = (x, g_{\beta\alpha}(x)u)$$

for all $x \in U_\alpha \cap U_\beta$ and $u \in F$. A reduction of the structure group of E to a subgroup $G \subset \mathrm{Diff}(F)$ consists of an open cover $\{U_\alpha\}$ of B together with local trivializations ψ_α such that all transition maps $g_{\beta\alpha}$ take values in G. The bundle E together with a fixed choice of such trivialization is called a fiber bundle with structure group G.

The tangent bundle TE contains a canonical vertical subbundle $V := \ker d\pi$. A connection on E is a splitting of the exact sequence

$$0 \to V \to TE \to TE/V \to 0$$

and corresponds to a horizontal distribution $H \subset TE$ satisfying $TE = H \oplus V$. Identifying H with the projection of TE onto V, we can describe a connection by a V-valued 1-form $A \in \Omega^1(E, V)$. The curvature of a connection is the 2-form $F_A \in \Omega^2(E, V)$ defined by

$$F_A(x; v, w) := A_x\left([v - A_x(v), w - A_x(w)]\right) = [v^{hor}, w^{hor}]^{vert}.$$

It measures the integrability of the horizontal distribution $H_A \subset TE$.

Affine connections and vector bundles. A vector bundle is a fiber bundle E whose fiber $F = V$ is a vector space and whose structure group $G \subset \mathrm{GL}(V)$ is linear. In this case every fiber $E_z := \pi^{-1}(z)$ has a canonical structure of a vector space and we have well-defined maps

$$\forall \lambda \in \mathbb{C} : \quad S_\lambda : E \to E, \qquad x \mapsto \lambda x$$

$$a : E \oplus E \to E, \qquad (x, y) \mapsto x + y.$$

A connection on E is a connection $A \in \Omega^1(E, V)$ of the underlying fiber bundle which is compatible with the linear structure on the fibers: Denote by $H_A \subset TE$ the horizontal distribution corresponding to A and by $\tilde{H}_A \subset T(E \oplus E)$ the induced horizontal distribution consisting of pairs $(v, w) \in H \oplus H$ satisfying $d\pi(v) = d\pi(w)$. Then one requires

$$dS_\lambda(H) \subset H \quad \forall \lambda \in \mathbb{C} \qquad \text{and} \qquad da(\tilde{H}) \subset H. \tag{2.3}$$

Alternatively, one can think of a connection as a covariant derivation

$$d_A : \Omega^0(\Sigma, E) \xrightarrow{d} \Omega^1(\Sigma, TE) \xrightarrow{A} \Omega^1(\Sigma, V) \cong \Omega^1(\Sigma, E)$$

where the last map comes from the canonical identification of the vertical bundle with the vector bundle itself. The linearity condition (2.3) says precisely that this defines an affine connection.

Definition 2.2.1. *Let $E \to \Sigma$ be a complex vector bundle. An **affine connection** on E is a linear operator $D : \Omega^0(\Sigma, E) \to \Omega^1(\Sigma, E)$ which satisfies the Leibniz rule*

$$D(fs) = df \otimes s + f \otimes Df$$

for all $f : \Sigma \to \mathbb{C}$ and $s \in \Omega^0(\Sigma, E)$.

We denote by $\mathcal{A}(E)$ the space of affine connections on E. Let $\psi_\alpha : E|_{U_\alpha} \to U_\alpha \times V$ be a local trivialization and denote for a local section $s : U_\alpha \to E$ with respect to this trivialization $s_\alpha := \mathrm{pr}_2 \circ \psi_\alpha$. Then an affine connection D has the shape

$$(Ds)_\alpha = ds_\alpha + A_\alpha s_\alpha$$

for some $A_\alpha \in \Omega^1(U_\alpha, \mathrm{End}(V))$. These A_α are called connection potentials for the affine connection D. If all connection potentials take values in the Lie algebra $\mathfrak{g} \subset \mathrm{End}(V)$ of the structure group $G \subset \mathrm{GL}(V)$, then the affine connection D is called a G-connection. We denote by $\mathcal{A}_G(E)$ the space of all G-connections on E.

An affine connection D induces higher covariant derivations by the formula

$$D : \Omega^k(\Sigma, E) \to \Omega^{k+1}(\Sigma, E), \qquad D(\tau \otimes s) = d\tau \otimes s + (-1)^k \tau \wedge Ds$$

for $\tau \in \Omega^k(\Sigma)$ and $s \in \Omega^0(\Sigma, E)$. The curvature $F_D \in \Omega^2(\Sigma, \mathrm{End}(E))$ is the unique tensor satisfying

$$(D \circ D)s = F_D \cdot s$$

for all $s \in \Omega^0(\Sigma, E)$. It is the obstruction to $D^2 = 0$ and not directly related to the curvature of the horizontal distribution defined by D. It rather corresponds to curvature of the induced horizontal distribution in the frame bundle of E as we shall see below.

Connections on principal bundles. Let G be a Lie group with Lie algebra $\mathfrak{g}$. A principal G bundle over Σ is a fiber bundle $\pi : P \to \Sigma$ together with a fiber preserving right action $P \times G \to P$ which is free and transitive on the fibers. In particular, the fibers are isomorphic to G and using the right action we can always construct equivariant local trivializations of P. For $p \in P$ and $\xi \in \mathfrak{g}$ the infinitesimal action of ξ is defined by

$$p\xi := \left.\frac{d}{dt}\right|_{t=0} p \exp(t\xi) \in T_pP.$$

The collection of these tangent vectors defines the vertical subbundle

$$V = \ker d\pi = \{p\xi \,|\, p \in P, \xi \in \mathfrak{g}\} \subset TP.$$

A connection on P is an equivariant connection of the underlying fiber bundle and corresponds to an equivariant horizontal distribution $H \subset TP$ satisfying $TP = V \oplus H$. Identifying H with the projection $\Pi : TP = V \oplus H \to V$, we can describe such a connection by a $\mathfrak{g}$-valued 1-form $A \in \Omega^1(P, \mathfrak{g})$ via the relation $\Pi_p(\hat{p}) = pA_p(\hat{p})$ for all $p \in P$ and $\hat{p} \in T_pP$. The connection 1-Form A satisfies the conditions

$$A_p(p\xi) = \xi \quad \text{and} \quad A_{pg}(\hat{p}g) = g^{-1} A_p(\hat{p}) g \tag{2.4}$$

for all $g \in G$, $\xi \in \mathfrak{g}$, $p \in P$ and $\hat{p} \in T_pP$. Conversely, the kernel of any $A \in \Omega^1(P, \mathfrak{g})$ satisfying (2.4) gives rise to an equivariant horizontal distribution $H \subset TP$. We define by

$$\mathcal{A}(P) := \{A \in \Omega^1(P, \mathfrak{g}) \,|\, A \text{ satisfies (2.4)}\}$$

the space of connections on P.

The curvature of a connection $A \in \mathcal{A}(P)$ is defined as

$$F_A := dA + \frac{1}{2}[A \wedge A] \in \Omega^2(P, \mathfrak{g})$$

where $[A \wedge A]$ is given by the usual formula for the exterior product with multiplication replaced by the Lie bracket. This curvature is linked to the curvature of the corresponding horizontal distribution by the relation

$$[X, Y]^{vert} = [X^{hor}, Y^{hor}]^{vert} = pF_A(p; X, Y)$$

for $p \in P$ and $X, Y \in T_pP$.

Associated bundles. Let $P \to \Sigma$ be a principal G bundle as above. A smooth manifold F together with a representation $\rho : G \to \mathrm{Diff}(F)$ gives rise to the associated fiber bundle $P \times_\rho F$ with fiber F which is defined by

$$P \times_\rho F := (P \times F)/G$$

where G acts diagonally by $g(p, x) = (pg, \rho(g)^{-1}x)$. We denote the orbit of $(p, x) \in P \times F$ under this action by $[p, x]$. A connection $A \in \mathcal{A}(P)$ induces a connection on the fiber bundle $P \times_\rho F$, which is given by the image of the horizontal distribution under $TP \subset TP \times TF \to T(P \times_\rho F)$.

Important examples arise from the action of G on itself by inner automorphism and from the adjoint action of G on its Lie algebra. We denote the associated bundles for these actions by

$$\mathrm{Ad}(P) := P \times_G G \qquad \text{and} \qquad \mathrm{ad}(P) := P \times_{ad} \mathfrak{g}.$$

Note that the bundle $\mathrm{Ad}(P)$ is a fiber bundle with fiber G but not a principal bundle. The fibers of $\mathrm{ad}(P)$ inherit from $\mathfrak{g}$ a well-defined Lie algebra structure.

The difference $a := A_1 - A_2$ of two connection 1-forms $A_1, A_2 \in \mathcal{A}(P)$ satisfies

$$a_p(p\xi) = 0 \qquad \text{and} \qquad a_{pg}(\hat{p}g) = g^{-1}a_p(\hat{p})g$$

for all $p \in P$, $\hat{p} \in T_pP$, $\xi \in \mathfrak{g}$ and $g \in G$. Hence a corresponds to a $\mathrm{ad}(P)$-valued 1-form $\bar{a}$ on Σ by the formula $\bar{a}(\pi(p); d\pi(p)\hat{p}) = [p, a(p; \hat{p})]$. This describes $\mathcal{A}(P)$ as an affine space with underlying linear space $\Omega^1(\Sigma, \mathrm{ad}(P))$ and with respect to any reference connection $A_0 \in \mathcal{A}(P)$ it holds

$$\mathcal{A}(P) = \{A_0 + a \,|\, a \in \Omega^1(\Sigma, \mathrm{ad}(P)\}.$$

Similarly, the curvature F_A of a connection A is an equivariant and horizontal 2-form on P and can thus be identified with an element $F_A \in \Omega^2(\Sigma, \mathrm{ad}(P))$.

Let H be a Lie group and let $\tilde{\rho} : G \to H$ be a homomorphism of Lie groups. Then left-multiplication $\rho(g) := L_{\tilde{\rho}(g)} \in \mathrm{Diff}(H)$ yields a representation of G and the associated bundle $P_H := P \times_\rho H$ is a principal H bundle. If $A \in \mathcal{A}(P)$, then A induces a connection $\rho(A) \in \mathcal{A}(P_H)$ by the formula

$$\rho(A)([p, h]; [\hat{p}, \hat{h}]) := h^{-1}\hat{h} + h^{-1}\dot{\rho}(A(p; \hat{p}))h$$

where $\dot{\rho} := d\rho(\mathbb{1}) : \mathfrak{g} \to \mathfrak{h}$ denotes the induced homomorphism of Lie algebras. The curvature of the induced connection is given by

$$F_{\rho(A)} = \dot{\rho}(F_A)$$

where $\dot{\rho}$ denotes the induced bundle map $\mathrm{ad}(P) \to \mathrm{ad}(P_H)$.

From principal bundles to vector bundles and back. Let V be a vector space and let $\rho : G \hookrightarrow \mathrm{GL}(V)$ be a faithful representation. The associated bundle $E := P \times_\rho V$ is then a vector bundle and the trivialization maps of P yield a natural reduction of the structure group of E to G. For a connection $A \in \mathcal{A}(P)$, the induced connection on E is compatible with the linear structure and defines an affine G-connection in $\mathcal{A}_G(E)$. The bundles $\mathrm{Aut}(E)$ and $\mathrm{End}(E)$ can be described as associated bundles

$$\mathrm{Aut}(E) = P \times_{\mathrm{Ad}(\rho)} \mathrm{GL}(V) \qquad \text{and} \qquad \mathrm{End}(E) = P \times_{\mathrm{Ad}(\rho)} \mathrm{End}(V)$$

where $\mathrm{Ad}(\rho) : G \to \mathrm{GL}(\mathrm{End}(V))$ is defined as the composition of ρ and the adjoint action of $\mathrm{GL}(V)$ on $\mathrm{End}(V)$. The induced map $\dot{\rho} : \mathfrak{g} \to \mathrm{End}(V)$ provides an inclusion $\mathrm{ad}(P) \to \mathrm{End}(E)$ and with respect to this map holds

$$F_{d_A} = \dot{\rho}(F_A)$$

for any connection $A \in \mathcal{A}(P)$.

Conversely, let $E \to \Sigma$ be a vector bundle with structure group $G \subset \mathrm{GL}(n)$. The frame bundle of E is defined by

$$\mathrm{Fr}(E) := \{(z, e) \,|\, z \in \Sigma,\, e : V \to E_z \text{ such that } \mathrm{pr}_2 \circ \psi_\alpha \circ e \in G\}$$

where $\psi_\alpha : E|_{U_\alpha} \to U_\alpha \times V$ is any trivialization of E with $z \in U_\alpha$. It follows directly from the definition that $\mathrm{Fr}(E)$ is a principal G bundle. An affine G-connection $D \in \mathcal{A}_G(E)$ induces a connection $A_D \in \mathcal{A}(\mathrm{Fr}(E))$ as follows: Let $\gamma : [0, 1] \to \Sigma$ be a curve. We call $e \in \Omega^0([0, 1], \gamma^*\mathrm{Fr}(E))$ a horizontal lift of γ if for every $v \in V$ the section

$$e_v \in \Omega^0([0, 1], \gamma^*E), \qquad e_v(t) := e(t)v \in E_{\gamma(t)}$$

satisfies $D_t(e_v) := D_{\dot{\gamma}(t)}(e_v(t)) = 0$. In a local trivialization this condition is equivalent to the ODE

$$\dot{e}_\alpha + A_\alpha(\gamma)e_\alpha = 0.$$

This shows that horizontal lifts exist when the connection potentials A_α take values in $\mathfrak{g}$. The tangent vector along horizontal lifts trace out an equivariant horizontal distribution in $\mathrm{Fr}(E)$ and hence determine a connection $A \in \mathcal{A}(\mathrm{Fr}(E))$.

The frame bundle construction is inverse to the construction of associated bundles in the sense that

$$\mathrm{Fr}(P \times_G V) \cong V \qquad \text{and} \qquad \mathrm{Fr}(E) \times_G V \cong E$$

whenever $G \subset \mathrm{GL}(V)$. This also provides a one-to-one correspondence between $\mathcal{A}(P)$ and $\mathcal{A}_G(E)$.

The Gauge group. The Gauge group of a principal G bundle P is defined as

$$\mathcal{G}(P) := \Omega^0(\Sigma, \mathrm{Ad}(P)).$$

This group is isomorphic to the group $\mathrm{Aut}(P)$ of fiber preserving equivariant automorphism of P under the map

$$\psi : \Omega^0(\Sigma, \mathrm{Ad}(P)) \cong \mathrm{Aut}(P), \qquad \psi_g(p) := pg(p).$$

It is useful think of $\mathcal{G}(P)$ as an infinite dimensional Lie group with Lie algebra

$$\mathrm{Lie}(\mathcal{G}(P)) = \Omega^0(\Sigma, \mathrm{ad}(P))$$

where all Lie theoretic operations are performed fiberwise. The Gauge group acts naturally on the space of connections via pull back

$$g(A) := \psi^*_{g^{-1}} A = -(dg)g^{-1} + gAg^{-1}.$$

The Gauge group of a vector bundle E with structure group G is the group

$$\mathcal{G}(E) := \Omega^0(\Sigma, G(E)) \subset \Omega^0(\Sigma, \mathrm{GL}(E))$$

which consists of all automorphisms of E taking values in G in any trivialization. We think of $\mathcal{G}(E)$ again as Lie group with Lie algebra $\Omega^0(\Sigma, \mathfrak{g}(E))$. The Gauge group acts naturally on the space of affine G-connection $\mathcal{A}_G(E)$ via pullback

$$(g^{-1})^* D = g \circ D \circ g^{-1}.$$

This action is more explicitly described in terms of the connection potential by

$$(gA)_\alpha = -dg_\alpha g_\alpha^{-1} + g_\alpha A_\alpha g_\alpha^{-1}$$

where $g_\alpha := (\mathrm{pr}_2 \circ \psi_\alpha)_* g : U_\alpha \to G$.

Suppose that $\rho : G \hookrightarrow \mathrm{GL}(V)$ is a faithful representation and $E := P \times_\rho V$ is an associated vector bundle. Then ρ induces an isomorphism $\mathrm{Ad}(P) \cong G(E)$ and hence $\mathcal{G}(P) \cong \mathcal{G}(E)$. The derivative $\dot\rho := d\rho(\mathbb{1}) : \mathfrak{g} \hookrightarrow \mathrm{End}(V)$ yields an isomorphism of $\mathrm{ad}(P) \cong \mathfrak{g}(E)$ and hence an identification of the Lie algebras of $\mathcal{G}(P)$ and $\mathcal{G}(E)$. From the naturality of the gauge action it is clear that the identification $\mathcal{A}(P) \cong \mathcal{A}_G(E)$ is equivariant with respect to the action of $\mathcal{G}(P)$ and $\mathcal{G}(E)$.

The moment map picture. Fix an invariant inner product $\langle\cdot,\cdot\rangle$ on $\mathfrak{g}$. This induces an inner product on the fibers of $\mathrm{ad}(P)$ and hence an invariant inner product on $\mathrm{Lie}(\mathcal{G}(P)) = \Omega^0(\Sigma, \mathrm{ad}(P))$ by the formula

$$\langle\xi,\eta\rangle := \int_\Sigma \langle\xi,\eta\rangle \, dvol_\Sigma.$$

This provides a natural hermitian structure on the space $\mathcal{A}(P)$ as follows. Since $\mathcal{A}(P)$ is an affine space, it suffices to define the hermitian structure on the underlying linear space $\Omega^1(\Sigma, \mathrm{ad}(P))$. For $a, b \in \Omega^1(\Sigma, \mathrm{ad}(P))$ we define

$$\omega_{\mathcal{A}}(a,b) := \int_\Sigma \langle a \wedge b\rangle, \qquad \langle a,b\rangle := \int_\Sigma \langle a \wedge *b\rangle, \qquad J_{\mathcal{A}} a := *a = -a \circ j_\Sigma.$$

The following observation is due to Atiyah and Bott [4].

Lemma 2.2.2. *The action of the Gauge group is Hamiltonian with moment map $\mu(A) := *F_A$. More explicitly, for every $\xi \in \Omega^0(\Sigma, ad(P))$ the infinitesimal action on $A \in \mathcal{A}(P)$ is given by*

$$L_A\xi := \left.\frac{d}{dt}\right|_{t=0} \exp(t\xi)(A) = -d_A\xi.$$

The function $\mathcal{A}(P) \to \mathbb{R}$, $A \mapsto \langle *F_A, \xi\rangle_{L^2}$, is differentiable and its differential is the 1*-form*

$$T_A\mathcal{A}(P) \to \mathbb{R}, \qquad a \mapsto \int_\Sigma \langle L_A\xi \wedge a\rangle = \omega_{\mathcal{A}}(L_A\xi, \cdot).$$

Proof. Let $\xi \in \Omega^0(\Sigma, \mathrm{ad}(P))$ be given and think of it as an equivariant map $\xi : P \to \mathfrak{g}$. We then compute

$$\begin{aligned}
\left.\frac{d}{dt}\right|_{t=0} \exp(t\xi)(A) &= \left.\frac{d}{dt}\right|_{t=0} d\exp(t\xi)^{-1}\exp(t\xi) + \exp(t\xi)A\exp(-t\xi) \\
&= \left.\frac{d}{dt}\right|_{t=0} \left(d\exp(t\xi)^{-1}\right) + [\xi, A] \\
&= -d\xi - [A, \xi]
\end{aligned}$$

The last expression agrees with $-d\xi$ along horizontal vectors in P and vanishes along vertical vectors. Hence it coincides with $-d_A\xi$ for the induced affine connection d_A on $\mathrm{ad}(P)$ and this proves the formula for the infinitesimal action.

From the formula

$$F_{A+a} = F_A + d_A a + \frac{1}{2}[a \wedge a]$$

we see that the variation of F_A in the direction $a \in \Omega^1(\Sigma, \mathrm{ad}(P))$ is given by $d_A a$. This yields

$$\langle d\mu(A)[a]; \xi\rangle = \int_\Sigma \langle d_A a, \xi\rangle = \int_\Sigma \langle a \wedge d_A\xi\rangle = \omega_{\mathcal{A}}(L_A\xi, a).$$

Here we used integration by part in the penultimate step and the formula

$$d\langle a, \xi\rangle = \langle d_A a, \xi\rangle - \langle a \wedge d_A\xi\rangle$$

which follows from the G-invariance of the inner product. □

The complexified gauge action

Let G be a compact connected Lie group and let $P \to \Sigma$ be a principal G bundle. We denote by G^c the complexification of G and call $P^c := P \times_G G^c$ the complexification of P. The complexified gauge group of P is defined as

$$\mathcal{G}^c(P) := \mathcal{G}(P^c).$$

One can think of elements in $\mathcal{G}^c(P)$ as G-equivariant maps from P to G^c. The Lie algebra bundle $\mathrm{ad}(P^c)$ is the complexification of the bundle $\mathrm{ad}(P)$ and since all Lie theoretic operations on the gauge group are defined fiberwise, it is reasonable to think of $\mathcal{G}^c(P)$ as the complexification of $\mathcal{G}(P)$. By the Peter-Weyl theorem, G admits a faithful representation $G \hookrightarrow U(n)$. Identifying G with its image in $U(n)$, we can describe its complexification $G^c \subset \mathrm{GL}(n)$ explicitly as the image of $G \times \mathfrak{g}$ under the diffeomorphism

$$U(n) \times \mathfrak{u}(n) \to \mathrm{GL}(n), \qquad (u, \eta) \mapsto u \exp(\mathbf{i}\eta).$$

In terms of the associated bundle $E := P \times_G \mathbb{C}^n$ the complexification of the gauge group is then given by

$$\mathcal{G}^c(E) = \Omega^0(\Sigma, G^c(E)).$$

The goal of this section is to explain how the $\mathcal{G}(P)$-action on $\mathcal{A}(P)$ extends naturally to a holomorphic action of $\mathcal{G}^c(P)$.

Proposition 2.2.3. *There exists a natural action of $\mathcal{G}^c(P)$ on $\mathcal{A}(P)$ whose infinitesimal action satisfies*

$$L_A(\xi + \boldsymbol{i}\eta) = L_A\xi + *L_A\eta = -d_A\xi - *d_A\eta \tag{2.5}$$

for all $\xi, \eta \in \Omega^0(\Sigma, ad(P))$ and $A \in \mathcal{A}(P)$.

Proof. See page 39. □

Holomorphic principal bundles. An almost complex structure J on a manifold M is an endomorphism $J \in \mathrm{End}(TM)$ satisfying $J^2 = -\mathbb{1}$. It is called an integrable or holomorphic structure if it endows M with the structure of a complex manifold. A holomorphic structure on the principal bundle $P^c = P \times_G G^c$ is an almost complex structure $J \in \mathrm{End}(TP^c)$ of the total space, which is G^c invariant and coincides with the canonical complex structure on the vertical subbundle, i.e. $J(p\zeta) = p(\mathbf{i}\zeta)$ for any $p \in P^c$ and $\zeta \in \mathfrak{g}^c$. We denote by $\mathcal{J}(P^c)$ the space of all holomorphic structures on P^c. The next lemma justifies this notation.

Lemma 2.2.4. *Every $J \in \mathcal{J}(P^c)$ is integrable.*

Proof. The Newlander-Nirenberg theorem states that an almost complex structure J on a manifold M is integrable if and only if the Nijenhuis-tensor $N_J : TM \otimes TM \to TM$ given by

$$N_J(v, w) := [v, w] + J[Jv, w] + J[v, Jw] - [Jv, Jw]$$

vanishes. We apply this to $M = P^c$. If $v, w \in T_p^{vert}(P^c)$ are both in the vertical bundle, we have $N_J(v,w) = 0$ as the fiber is a complex manifold. If $v \in T_p^{vert}(P^c)$ and $w \in T_p^{hor}(P^c)$ the Lie bracket $[v,w] = \mathcal{L}_v(w)$ vanishes, since the horizontal distribution is equivariant. In particular $N_J(v,w) = 0$ as all four terms vanish separately. Let finally $v, w \in T_p^{hor}(P^c)$ be horizontal vectors. We may assume that $p \in P$ and denote by $\bar{v} := d\pi(p)v$ and $\bar{w} := d\pi(p)w$ the projections onto $T\Sigma$. By definition of the curvature, we obtain the vertical component of the Nijenhuis tensor by

$$\begin{aligned} A_p(N_J(v,w)) &= F_A(\bar{v},\bar{w}) + \mathbf{i}F_A(j_\Sigma \bar{v},\bar{w}) + \mathbf{i}F_A(\bar{v}, j_\Sigma \bar{w}) - F_A(j_\Sigma \bar{v}, j_\Sigma \bar{w}) \\ &= 4F_A^{0,2}(\bar{v},\bar{w}) = 0. \end{aligned}$$

In the last step we use that Σ is a complex one-dimensional manifold and thus $\Omega^{0,2}(\Sigma) = 0$. The horizontal part of $N_J(v,w)$ gets identified under $d\pi(p)$ with $N_J(\bar{v},\bar{w})$ and vanishes as Σ is a complex manifold. This completes the proof of $N_J = 0$. $\square$

As a consequence, every holomorphic principal bundles admits holomorphic local trivializations with holomorphic transition maps. The next lemma is due to Singer [104].

Lemma 2.2.5. *There exists a one to one correspondence between connections $A \in \mathcal{A}(P)$ and holomorphic structures $J \in \mathcal{J}(P^c)$.*

Proof. A connection $A \in \mathcal{A}(P)$ induces a connection on P^c and thus determines for every $p \in P^c$ a splitting $T_p(P^c) = T_p^{hor}(P^c) \oplus T_p^{vert}(P^c)$. The vertical part is isomorphic to $\mathfrak{g}^c$ and has a canonical complex structure. The differential of the projection $\pi : P^c \to \Sigma$ restricts to an isomorphism $d\phi(p) : T_p^{hor}(P^c) \to T_{\pi(p)}\Sigma$ and induces a complex structure on $T_p^{hor}(P^c)$.

Conversely, let $J \in \mathcal{J}(P^c)$ be given and think of $P \subset P^c$ as a subbundle. For $p \in P$ we define $H_p := T_pP \cap J_p(T_pP)$ and claim that $T_pP + J_p(T_pP) = T_p(P^c)$. Indeed, since $T_p^{vert}P \cong \mathfrak{g}$, the sum clearly contains the vertical fiber $T_p^{vert}(P^c) \cong \mathfrak{g}^c$ and $d\pi(p)$ maps T_pP already onto $T_{\pi(p)}\Sigma$. It is immediate from the construction that H_p is invariant under J_p and defines a (real) two dimensional complement of $T_p^{vert}(P^c)$ in $T_p(P^c)$. As p varies over P we obtain an equivariant distribution along P and hence a connection $A \in \mathcal{A}(P)$. $\square$

Let $A \in \mathcal{A}(P)$, $g \in \mathcal{G}(P)$ and let $J_A \in \mathcal{J}(P^c)$ be the holomorphic structure induced by A. Then $g(A)$ induces the holomorphic structure $(\psi_{g^{-1}})^* J_A$, since the construction above is clearly functorial. The action of $\mathcal{G}(P)$ on $\mathcal{J}(P^c)$ has a natural extension to the complexified gauge group via

$$\mathcal{G}^c(P) \times \mathcal{J}(P^c) \to \mathcal{J}(P^c), \qquad g(J) := (\psi_{g^{-1}})^* J$$

where $\psi_{g^{-1}} \in \mathrm{Aut}(P^c)$ is the automorphism corresponding to g^{-1}. Using the identification of $\mathcal{J}(P^c)$ with $\mathcal{A}(P)$ this yields the desired action of $\mathcal{G}^c(P)$ on $\mathcal{A}(P)$ and the quotient $\mathcal{A}(P)/\mathcal{G}^c(P)$ parametrizes the isomorphism classes of holomorphic structures on P^c.

Holomorphic vector bundles. We consider the special case $G = U(n)$ and denote by $E := P \times_{U(n)} \mathbb{C}^n$ the associated vector bundle. A holomorphic structure on E is an almost complex structure $J \in \mathrm{End}(TE)$ of the total space which restricts to the linear complex structure on the fibers. Similarly as in the case of principal bundles, one shows that every such structure is indeed integrable and that every holomorphic vector bundle admits holomorphic trivializations. It is then easy to see that every holomorphic vector bundle E carries a natural operator

$$\bar{\partial}_E : \Omega^0(\Sigma, E) \to \Omega^{0,1}(\Sigma, E)$$

which in any holomorphic trivialization agrees with the usual $\bar{\partial}$ operator on $\mathbb{C}^n$. This operator is a particular Cauchy-Riemann operator on E.

Definition 2.2.6. *Let $E \to \Sigma$ be a complex vector bundle. A Cauchy Riemann operator on E is a linear operator*

$$D'' : \Omega^0(\Sigma, E) \to \Omega^{0,1}(\Sigma, E)$$

which satisfies the Leibniz rule

$$D''(fs) = \bar{\partial} f \otimes s + f \otimes D''s$$

for all $f : \Sigma \to \mathbb{C}$ and $s \in \Omega^0(\Sigma, E)$.

The converse is also true: Every Cauchy-Riemann operator determines a holomorphic structure on the complex bundle E, whose local holomorphic sections are solutions of the Cauchy-Riemann equation $D''s = 0$. This is another instance of the Newlander-Nirenberg theorem. In the case of Riemann surfaces a simpler proof of this result is given by Atiyah and Bott ([4], Section 5).

Note that the associated vector bundle E carries a canonical hermitian metric, which in any trivialization coincides with the standard hermitian metric on $\mathbb{C}^n$. We claim that there is a one to one correspondence between unitary connections on E and Cauchy-Riemann operators. For a unitary connection D we obtain a Cauchy-Riemann operator by the formula

$$D''s := (Ds)^{0,1} := \frac{1}{2}\left(Ds + \mathbf{i}(Ds) \circ j_\Sigma\right) = \frac{1}{2}\left(Ds - \mathbf{i} * (Ds)\right).$$

To show that this correspondence is bijective, it suffices to examine this correspondence locally. In a unitary trivialization $\psi : E|_U \to U \times \mathbb{C}^n$ the connection D can be described in terms of a 1-form $A \in \Omega^1(U, \mathfrak{u}(n))$ such that

$$Ds := ds + As, \qquad D''s := \bar{\partial}s + A^{0,1}s$$

holds for any section $s \in \Omega^0(U, \mathbb{C}^n)$ with $A^{0,1} := \frac{1}{2}(A + \mathbf{i}A \circ j_\Sigma)$. In particular, we recover A as twice the skew-hermitian part of $A^{0,1}$ and therefore it is uniquely determined by $A^{0,1}$. Conversely, any Cauchy Riemann operator D'' is given in this local trivialization by

$$D''s := \bar{\partial}s + Bs$$

for some $B \in \Omega^{0,1}(U, \mathfrak{gl}(n))$. Since B satisfies $B(j_\Sigma v) = -\mathbf{i}B(v)$ for any tangent vector $v \in T\Sigma|_U$, the skew-hermitian and hermitian part of B interchange if we compose B with j_Σ. This shows that B has the form $B = \frac{1}{2}(A + \mathbf{i}A \circ j_\Sigma)$ for some $A \in \Omega^1(U, \mathfrak{u}(n))$ and this proves the claim.

On the level of Cauchy-Riemann operators the complexified Gauge group $\mathcal{G}^c(E) = \Omega^0(\Sigma, \mathrm{GL}(E))$ acts naturally via

$$g(\bar\partial_A) := g \circ \bar\partial_A \circ g^{-1} = \bar\partial_A - \bar\partial_A(g)g^{-1}.$$

The next lemma summarizes the discussion above and provides an explicit formulas for this action on $\mathcal{A}(E)$.

Lemma 2.2.7. *Let $E \to \Sigma$ be a complex vector bundle.*

1. *For every holomorphic structure $\bar\partial_E$ and hermitian metric H exists a unique connection*

$$D := D(\bar\partial_E, H) =: D' + D'' \in \mathcal{A}^{1,0}(E) \oplus \mathcal{A}^{0,1}(E)$$

such that D is unitary with respect to H and $D'' = \bar\partial_E$.

2. *Let $g \in \Omega^0(\Sigma, GL(E))$ and denote $h := g^*g$ (with respect to H). Then*

$$D(g(\bar\partial_E), H) = g\left(D + h^{-1}D'(h)\right)g^{-1}$$

$$F(g(\bar\partial_E), H) = g\left(F + D''(h^{-1}D'(h))\right)g^{-1}.$$

Proof. For the first part, note that there is a one to one correspondence between hermitian metrics H and reductions of the structure group of E to $U(n)$: Using the Gram-Schmidt process we can always find local trivializations which identify H with the standard hermitian product on $\mathbb{C}^n$ and the transition map between such trivializations are clearly unitary. The second part follows from the formula

$$D(g(\bar\partial_E), H) = g(D) = g \circ D'' \circ g^{-1} + (g^{-1})^* \circ D' \circ g^*$$

and $F = D \circ D$. □

Remark 2.2.8. Consider the general case and assume that $G \subset U(n)$ is a compact connected subgroup. The structure group of E is then contained in G and the explicit formula in the lemma above shows that the subspace $\mathcal{A}_G(E)$ of G-connections is preserved by the action of $\mathcal{G}^c(E) = \Omega^0(\Sigma, G^c(E))$. Since holomorphic structures on E and its frame determine one another, it is clear that this action corresponds to the action described on holomorphic principal bundles above.

We may now deduce the formula for the infinitesimal action (2.5).

Proof of Proposition 2.2.3. As in Lemma 2.2.2 one calculates

$$\left.\frac{d}{dt}\right|_{t=0} \exp(t\zeta)(\bar\partial_A) = -\bar\partial_A\zeta$$

for $\zeta \in \Omega^0(\Sigma, \mathfrak{g}^c(E))$. Write $\zeta = \xi + \mathbf{i}\eta$ with $\xi, \eta \in \Omega^0(\Sigma, \mathfrak{g}(E))$ and use the formula $\bar\partial_A(\mathbf{i}\eta) = *\bar\partial_A\eta$ to deduce

$$\begin{aligned} L_A\zeta = -\bar\partial_A\zeta + (\bar\partial_A\zeta)^* &= -(\bar\partial_A\xi - (\bar\partial_A\xi)^*) - *(\bar\partial_A\eta - (\bar\partial_A\eta)^*) \\ &= -d_A\xi - *d_A\eta = L_A\xi + *L_A\eta. \end{aligned}$$

□

Regularity assumptions

Let G be a compact connected Lie group and let $P \to \Sigma$ be a principal G bundle. We shall always consider connections of Sobolev class $W^{1,2}$ and gauge transformations of Sobolev class $W^{2,2}$. More precisely, the space of $W^{1,2}$ connections on P is defined with respect to some smooth reference connection A_0 as

$$\mathcal{A}(P) := \{A_0 + a \,|\, a \in W^{1,2}(\Sigma, T^*\Sigma \otimes \mathrm{ad}(P))\}$$

and the $W^{2,2}$ completion of the gauge group and its complexification are

$$\mathcal{G}(P) := W^{2,2}(\Sigma, \mathrm{Ad}(P)), \qquad \mathcal{G}^c(P) := W^{2,2}(\Sigma, \mathrm{Ad}(P^c)).$$

We use the same notation as for the smooth groups, since all the results from the previous section carry over. In particular, the action of the gauge group and its complexification extend smoothly over these Sobolev completions, since $W^{2,2} \hookrightarrow C^0$ is in the good range of the Sobolev embedding. A connection still determines a holomorphic structure up to isomorphism due to the following regularity result.

Lemma 2.2.9. *For every $W^{1,2}$ connection $A \in \mathcal{A}(P)$ exists a complex $W^{2,2}$ gauge transformation $g \in \mathcal{G}^c(P)$ such that $g(A)$ is smooth.*

Proof. This is Lemma 14.8 in [4]. By Proposition 2.2.3, the infinitesimal action of the complex gauge group is given by

$$L_A : W^{2,2}(\Sigma, \mathrm{ad}(P^c)) \to W^{1,2}(\Sigma, T^*\Sigma \otimes \mathrm{ad}(P))$$
$$L_A(\xi + \mathbf{i}\eta) = -d_A\xi - *d_A\eta$$

For any smooth reference connection A_0, this is a compact perturbation of L_{A_0} which is a Fredholm operator. Hence L_A is also Fredholm and in particular its cokernel is finite dimensional.

It follows from the implicit function theorem in Banach spaces that we can choose a finite dimensional slice N orthogonal to the $\mathcal{G}^c$-orbit through A. Say $\dim(N) = r$ and fix $r+1$ connections $B_0, \ldots, B_r \in N$ which span an r-simplex containing A in its interior. A small perturbation of the vertices yields smooth connections $\tilde{B}_0, \ldots, \tilde{B}_r$ and the simplex spanned by these connections will still intersect the orbit $\mathcal{G}^c(A)$. This intersection point yields a smooth connection in the $\mathcal{G}^c$ orbit of A. □

2.2.2 Parabolic subgroups

Let G be a compact connected Lie group with Lie algebra $\mathfrak{g}$ and denote its complexification by G^c. Fix an invariant inner product on $\mathfrak{g}$. This induces a (real valued) inner product on $\mathfrak{g}^c = \mathfrak{g} \oplus \mathbf{i}\mathfrak{g}$ where we define both factors to be orthogonal. We define parabolic subgroups of G^c first by using toral generators of $\mathfrak{g}^c$. Then we recall briefly the root space decomposition of reductive Lie algebras and give an alternative intrinsic definition of parabolic subgroups. The first definition occures naturally in the geometric description of the weights in Chapter 5. The intrinsic version turns out to be useful in the proof of Proposition 2.5.9 which relates the algebraic notion of stability with the weights.

Toral generators

An element $\zeta \in \mathfrak{g}^c$ is called a **toral generator** if

$$T_\zeta := \overline{\{\exp(t\zeta)\,|t \in \mathbb{R}\}} \subset G^c$$

is a compact torus. We denote by $\mathcal{T}^c$ the set of toral generators. Certainly $\mathfrak{g} \subset \mathcal{T}^c$. Since any maximal compact subgroup of G^c is conjugated to G, for every $\zeta \in \mathcal{T}^c$ exists $g \in G^c$ such that $g^{-1}T_\zeta g^{-1} \subset G$. The relation $gT_\zeta g^{-1} = T_{g\zeta g^{-1}}$ then yields $g\zeta g^{-1} \in \mathfrak{g}$ and hence

$$\mathcal{T}^c = \mathrm{Ad}(G^c)(\mathfrak{g}) = \{g\xi g^{-1}\,|\, g \in G^c,\, \xi \in \mathfrak{g}\}.$$

Definition 2.2.10. *A **parabolic subgroup** of G^c is a subgroup of the form*

$$Q(\zeta) := \{g \in G^c \,|\, \textit{the limit } \lim_{t\to\infty} e^{\mathbf{i}t\zeta} g e^{-\mathbf{i}t\zeta} \textit{ exists in } G^c\}$$

*for some $\zeta \in \mathcal{T}^c$. The **Levi subgroup** of $Q(\zeta)$ is defined by*

$$L(\zeta) := \{g \in G^c \,|\, e^{\mathbf{i}\zeta} g e^{-\mathbf{i}\zeta} = g\}.$$

Remark 2.2.11. We consider $G^c = Q(0)$ as parabolic subgroup of itself.

Lemma 2.2.12. *Consider the setting described above and let $\zeta \in \mathcal{T}^c$.*

1. *$Q(\zeta)$ is a closed connected Lie subgroup of G^c with Lie algebra*

$$\mathfrak{q}(\zeta) := \{\rho \in \mathfrak{g}^c \,|\, \textit{the limit } \lim_{t\to\infty} e^{\mathbf{i}t\zeta}\rho e^{-\mathbf{i}t\zeta} \textit{ exists in } \mathfrak{g}^c\}.$$

2. *$L(\zeta)$ is a closed connected Lie subgroup of G^c with Lie algebra*

$$\mathfrak{l}(\zeta) := \{\rho \in \mathfrak{g}^c \,|\, e^{\mathbf{i}t\zeta}\rho e^{-\mathbf{i}t\zeta} = \rho\}$$

3. *$L(\zeta)$ is a maximal reductive subgroup of $Q(\zeta)$.*

4. *$Q(\zeta) = G^c$ if and only if ζ is contained in the center of $\mathfrak{g}^c$.*

Proof. Since $Q(g\zeta g^{-1}) = gQ(\zeta)g^{-1}$ and $L(g\zeta g^{-1}) = gL(\zeta)g^{-1}$, we may assume $\zeta = \xi \in \mathfrak{g}$. By the Peter-Weyl theorem, there exists a faithful representation $G \hookrightarrow U(n)$ and we may identify G with a closed subgroup of $U(n)$. Then $\mathbf{i}\xi$ yields a hermitian endomorphism of $\mathbb{C}^n$ which is diagonalizable with real eigenvalues $\lambda_1 < \cdots < \lambda_r$. Denote the eigenspace corresponding to λ_j by V_j. They yield an orthogonal decomposition

$$\mathbb{C}^n = V_1 \oplus \cdots \oplus V_r.$$

In this decomposition we can write $g \in G^c \subset \mathrm{GL}(n, \mathbb{C})$ as

$$g = \begin{pmatrix} g_{11} & g_{12} & \cdots & g_{1r} \\ g_{21} & g_{22} & \cdots & g_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ g_{r1} & g_{r2} & \cdots & g_{rr} \end{pmatrix}$$

with $g_{ij} \in \mathrm{Hom}(V_j, V_i)$. Then

$$e^{\mathbf{i}t\xi} g e^{-\mathbf{i}t\xi} = \begin{pmatrix} g_{11} & e^{(\lambda_1-\lambda_2)t} g_{12} & \cdots & e^{(\lambda_1-\lambda_r)t} g_{1r} \\ e^{(\lambda_2-\lambda_1)t} g_{21} & g_{22} & \cdots & e^{(\lambda_2-\lambda_r)t} g_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ e^{(\lambda_r-\lambda_1)t} g_{r1} & e^{(\lambda_r-\lambda_2)t} g_{r2} & \cdots & g_{rr} \end{pmatrix}.$$

Thus $g \in Q(\xi)$ if and only if g is upper triangular (i.e. $g_{ij} = 0$ for $i > j$) and $g \in L(\xi)$ if and only if g is block diagonal (i.e. $g_{ij} = 0$ for $i \neq j$). This shows that $L(\xi)$ and $Q(\xi)$ are closed subgroups of G^c and the formulas for $\mathfrak{l}(\xi)$ and $\mathfrak{q}(\xi)$ are immediate.

As the spaces V_j are pairwise orthogonal, the intersection $G \cap Q(\xi)$ consists of block diagonal matrices and hence agrees with the centralizer of the torus T_ξ in G. Since the centralizers of tori in compact groups are connected (see [70] Corollary 4.51) we conclude that $G \cap Q(\xi)$ is connected. Since $L(\xi)$ is the complexification of $G \cap Q(\xi)$ it is connected and reductive. Moreover $Q(\xi)/L(\xi)$ can be identified with the unipotent matrices in $Q(\xi)$ and hence $L(\xi)$ is a maximal reductive subgroup of $Q(\xi)$. We observe that

$$Q(\xi) \to L(\xi), \qquad g \mapsto \lim_{t\to\infty} e^{\mathbf{i}t\xi} g e^{-\mathbf{i}t\xi}$$

defines a continuous retraction of $Q(\xi)$ onto $L(\xi)$ and hence $Q(\xi)$ is connected.

Finally, since G^c is reductive, we have $G^c = Q(\xi)$ if and only if $G^c = L(\xi)$. The later is clearly equivalent to $\xi \in Z(\mathfrak{g})$. □

The root-space decomposition

We recall the necessary background on Lie theory briefly and refer to [70] for the proofs. Note that the discussion remains valid for any G-invariant inner product on $\mathfrak{g}$, which does not need to be the negative Killing form.

Reductive Lie groups. Using the invariant inner product on $\mathfrak{g}$, it is easy to show that the adjoint action of $\mathfrak{g}$ on itself is completely reducible. This yields an orthogonal decomposition

$$\mathfrak{g} = \mathfrak{z} \oplus [\mathfrak{g}, \mathfrak{g}]$$

where $\mathfrak{z}$ denotes the center of $\mathfrak{g}$ and the commutator $[\mathfrak{g}, \mathfrak{g}]$ is a direct sum of simple ideals and hence a semisimple Lie algebra. The same decomposition is valid for the complexification. To see this extend the inner product on $\mathfrak{g}$ to a non-degenerated $\mathbb{C}$-bilinear form $B : \mathfrak{g}^c \times \mathfrak{g}^c \to \mathbb{C}$ by

$$B(\xi_1 + \mathbf{i}\eta_1; \xi_2 + \mathbf{i}\eta_2) = \langle \xi_1, \xi_2 \rangle - \langle \eta_1, \eta_2 \rangle + \mathbf{i}(\langle \xi_1, \eta_2 \rangle + \langle \eta_1, \xi_2 \rangle).$$

This bilinear form is nondegenerate and G^c-invariant. Moreover, the B-orthogonal complement of a complex subspace $W \subset \mathfrak{g}^c$ is a G^c-invariant complement and the same argument as above yields the decomposition

$$\mathfrak{g}^c = \mathfrak{z}^c \oplus [\mathfrak{g}^c, \mathfrak{g}^c].$$

Root space decomposition. Fix a maximal torus $T \subset G$ with Lie algebra $\mathfrak{t}$ and decompose it orthogonally as $\mathfrak{t} = \mathfrak{z} \oplus \mathfrak{t}_0$. A nonzero imaginary valued real linear map

$$\alpha = \mathbf{i}a : \mathfrak{t}_0 \to \mathbf{i}\mathbb{R}, \qquad a \in \mathrm{Hom}(\mathfrak{t}_0, \mathbb{R})$$

is called a **root** of G with respect to T if there exists $e_\alpha \in [\mathfrak{g}^c, \mathfrak{g}^c]$ satisfying

$$[t, e_\alpha] = \alpha(t) e_\alpha \qquad \text{for all } t \in \mathfrak{t}_0.$$

The element e_α is uniquely determined by α up to scaling. We denote by $\mathfrak{g}_\alpha := \mathbb{C} \cdot e_\alpha$ the one dimensional root space corresponding to α and denote by R the set of all roots (relative to T). The **root space decomposition** of $\mathfrak{g}^c$ is the vector space decomposition

$$\mathfrak{g}^c = \mathfrak{z}^c \oplus \mathfrak{t}_0^c \oplus \bigoplus_{\alpha \in R} \mathfrak{g}_\alpha.$$

For a proof see [70] Chapters II.1-4 and IV.5.

Lemma 2.2.13. *Denote* $\mathfrak{g}_0 := \mathfrak{t}_0^c$.

1. *For* $\alpha, \beta \in R \cup \{0\}$ *the Lie bracket satisfies the relation*

$$[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subset \mathfrak{g}_{\alpha+\beta}$$

where the right-hand side is defined to be zero when $\alpha + \beta \notin R \cup \{0\}$.

2. *For* $\alpha, \beta \in R \cup \{0\}$ *with* $\alpha \neq -\beta$ *the subspaces* $\mathfrak{g}_\alpha$ *and* $\mathfrak{g}_\beta$ *are* B*-orthogonal.*

3. *If* $\alpha \in R$*, then* $-\alpha \in R$*. Moreover, if* $e_\alpha \in \mathfrak{g}_\alpha$ *then* $\bar{e}_\alpha \in \mathfrak{g}_{-\alpha}$ *and*

$$(\mathfrak{g}_\alpha \oplus \mathfrak{g}_{-\alpha}) \cap \mathfrak{g} = \mathbb{R}(e_\alpha + \bar{e}_\alpha) \oplus \mathbb{R}(\boldsymbol{i} e_\alpha - \boldsymbol{i} \bar{e}_\alpha)$$

Proof. The first and the last statement follow directly from the definitions. For the second statement consider first the case $\beta = 0$ and $\alpha \in R$. Then follows for all $s, t \in \mathfrak{t}_0^c$

$$B(\alpha(t)e_\alpha, s) = B([t, e_\alpha], s) = -B(e_\alpha, [t, s]) = 0$$

where we used in the second step that B is G^c-invariant. This shows that $\mathfrak{t}_0^c$ is B-orthogonal to $\mathfrak{g}_\alpha$. Now consider $\alpha, \beta \in R$ with $\alpha + \beta \neq 0$. A similar calculation shows for all $s, t \in \mathfrak{t}_0^c$

$$B(\alpha(t)e_\alpha, \beta(s)e_\beta) = B([t, e_\alpha], \beta(s)e_\beta) = -B(t, [e_\alpha, \beta(s)e_\beta]) = 0$$

where the last equality follows from the observation $[e_\alpha, \beta(s)e_\beta] \in \mathfrak{g}_{\alpha+\beta}$. □

The Weyl group. Using the inner product on $\mathfrak{g}$, we identify the roots $\alpha = \mathbf{i}a \in R$ with vectors $t_\alpha \in \mathfrak{t}_0$ by the relation

$$a(t) := \langle t_\alpha, t \rangle \qquad \text{for all } t \in \mathfrak{t}_0.$$

This yields a subset $\Phi_R = \{t_\alpha \,|\, \alpha \in R\} \subset \mathfrak{t}_0$ which satisfies the properties of an abstract root system:

1. Φ_R is a spanning set for $\mathfrak{t}_0$.
2. For every $t_\alpha \in \Phi_R$, the orthogonal reflection along $\ker\alpha$

$$s_\alpha : \mathfrak{t}_0 \to \mathfrak{t}_0, \qquad s_\alpha(t) := t - \frac{2\langle t, t_\alpha\rangle}{||t_\alpha||^2}$$

carries Φ_R to itself.

3. $\frac{2\langle t_\beta, t_\alpha\rangle}{||t_\alpha||^2}$ is an integer for all $t_\alpha, t_\beta \in \Phi_R$.

This is discussed in [70] Chapters II.5. The subgroup W generated by all the root reflection s_α inside the orthogonal group $O(\mathfrak{t}_0)$ is called the **Weyl group**. Since Φ_R is a spanning set of $\mathfrak{t}_0$, any orthogonal transformation which fixes Φ_R must be the identity and hence the Weyl group is always finite. After removing all hyperplanes $\ker(\alpha)$ the Weyl group acts transitively and freely on $\mathfrak{t}_0 \setminus \cup \{\ker(\alpha) \,|\, \alpha \in R\}$. The closure of a connected component of this space is called a **Weyl chamber** $\Omega_W \subset \mathfrak{t}_0$. In particular, Ω_W is the closure of a fundamental domain for the action of W. The Weyl group can alternatively be described as

$$W \cong N_G(T)/Z_G(T).$$

Here the normalizer $N_G(T)$ acts on the maximal torus T by conjugation. This action is trivial on the connected component of the center $Z_0(G) \subset T$ and its derivative induces an action on $\mathfrak{t}_0$. Since the inner product on $\mathfrak{g}$ is G-invariant, this identifies $N_G(T)/Z_G(T)$ with a subgroup of the orthogonal group $O(\mathfrak{t}_0)$ and it is easy to check that this group permutes the roots t_α. The equivalence of both descriptions of the Weyl-group is shown in [70] Chapters IV.6. Since any two maximal tori in G are conjugated, this shows that the conjugation classes in G are parametrized by T/W and in particular any element $\xi \in \mathfrak{g}$ is conjugated to an element in the Weyl chamber $\Omega_W \subset \mathfrak{t}_0$.

Simple roots. Consider a notion of positivity on the set R satisfying the properties

1. For every root $\alpha \in R$ exactly one of α and $-\alpha$ is positive.
2. If α and β are positive, then $\alpha + \beta$ is positive.

An easy way to define such a notion goes as follows. Choose a real linear functional $\phi : \mathfrak{t}_0 \to \mathbb{R}$ such that $\ker\phi \cap \Phi_R = \emptyset$ and define a root $\alpha \in R$ to be positive whenever $\phi(t_\alpha) > 0$. We write $\alpha > 0$ for a positive root α and denote by R^+ the collection of positive roots. This induces a partial ordering on the roots according to the rule

$$\alpha > \beta \quad \text{if and only if} \quad \alpha - \beta > 0.$$

A root $\alpha \in R^+$ is called simple if it cannot be decomposed as $\alpha = \beta + \gamma$ with $\beta, \gamma \in R^+$. In other words, a simple root is a minimal positive root. We denote by $R_0^+ = \{\alpha_1, \ldots, \alpha_r\}$ the set of simple roots. It is easy to deduce from the definitions that any root α can be written as

$$\alpha = \sum_{j=1}^{r} x_j \alpha_j \tag{2.6}$$

with coefficients $x_1, \dots, x_r \in \mathbb{Z}$ having all the same sign (or vanish). In particular $\Phi_{R_0^+}$ is a spanning set of $\mathfrak{t}_0$. A less obvious fact is that $\Phi_{R_0^+}$ is linear independent (see [70] II.5 Prop 2.49). Hence every root has a unique expression (2.6) and a root is positive if and only if all the coefficients are nonnegative. This observation shows that the collection of simple roots and the partial ordering determine one another.

Any collection of simple roots $R_0^+ = \{\alpha_1, \dots, \alpha_r\}$ determines a canonical Weyl chamber by the formula

$$\Omega_W = \{t \in \mathfrak{t}_0 \,|\, a_j(t) \geq 0 \text{ for all } j = 1, \dots, r\}$$

where we denote $\alpha_j = \mathbf{i}a_j$ as above. Conversely, given a Weyl chamber Ω_W we can recover the collection of positive roots by the rule

$$\alpha > 0 \text{ if and only if } \langle t, t_\alpha \rangle \geq 0 \text{ for all } t \in \Omega_W.$$

Hence the choice of a Weyl chamber and a partial ordering determine one another as well. Since any two Weyl chambers are conjugated by an element in G, this shows that all the choices in this section are canonical up to conjugation.

We denote the simple roots in $\Phi_{R_0^+}$ for convenience by $t_j := t_{\alpha_j}$. Since they define a basis of $\mathfrak{t}_0$, we can define a dual basis $\{\check{t}_1, \dots, \check{t}_r\}$ by

$$\left\langle \check{t}_i, \frac{2t_j}{||t_j||^2} \right\rangle = \delta_{ij} \tag{2.7}$$

for $i, j = 1, \dots, r$. They are clearly contained in the Weyl chamber determined by the simple roots and yield the characterization

$$t = \sum_{j=1}^{r} x_j \check{t}_j \in W_\Omega \qquad \Leftrightarrow \qquad x_j \geq 0 \text{ for } j = 1, \dots, r.$$

The dual elements

$$\lambda_j : \mathfrak{t}_0 \to \mathbf{i}\mathbb{R}, \qquad \lambda_j(t) := \mathbf{i}\langle \check{t}_j, t \rangle$$

are called the fundamental weights associated to the simple roots.

An intrinsic definition of parabolic subgroups

We provide an intrinsic definition of parabolic subgroups following the presentation [99] by Serre. Let $\xi \in [\mathfrak{g}, \mathfrak{g}]$ be given and choose a maximal torus $T \subset G$ such that $\xi \in \mathfrak{t}_0$. Moreover, let $R_0^+ := \{\alpha_1, \dots, \alpha_r\}$ be a choice of simple roots such that ξ is contained in the corresponding Weyl chamber. Denote

$$R(\xi) := \{\alpha \in R \,|\, \langle \xi, t_\alpha \rangle \geq 0\} \qquad \text{and} \qquad \tilde{R}(\xi) := \{\alpha \in R \,|\, \langle \xi, t_\alpha \rangle = 0\}. \tag{2.8}$$

Define the Lie subalgebras

$$\mathfrak{q}(\xi) := \mathfrak{z} \oplus \mathfrak{t}_0 \oplus \bigoplus_{\alpha \in R(\xi)} \mathfrak{g}_\alpha \tag{2.9}$$

and

$$\mathfrak{l}(\xi) := \mathfrak{z} \oplus \mathfrak{t}_0 \oplus \bigoplus_{\alpha \in \tilde{R}(\xi)} \mathfrak{g}_\alpha. \tag{2.10}$$

The next lemma shows that this notation is consistent with our definition in the section on toral generators.

Lemma 2.2.14. *Consider the setting from above and define $\mathfrak{q}(\xi)$ and $\mathfrak{l}(\xi)$ by (2.9) and (2.10) respectively. Then*

$$\mathfrak{q}(\xi) = \{\rho \in \mathfrak{g}^c \,|\, \textit{the limit} \lim_{t\to\infty} e^{\boldsymbol{i}t\xi}\rho e^{-\boldsymbol{i}t\xi} \textit{ exists in } \mathfrak{g}^c\}$$

and

$$\mathfrak{l}(\xi) = \{\rho \in \mathfrak{g}^c \,|\, e^{\boldsymbol{i}t\xi}\rho e^{-\boldsymbol{i}t\xi} = \rho\}.$$

Proof. Decompose $\rho \in \mathfrak{g}^c$ with respect to the root space decomposition as

$$\rho = \rho_0 + \sum_{\alpha \in R} \rho_\alpha$$

with $\rho_0 \in \mathfrak{t}$ and $\rho_\alpha \in \mathfrak{g}_\alpha$. By definition of the roots we have

$$[\mathbf{i}\xi, \rho_a] = -a(\xi)\rho(\xi) \cdot \rho_\alpha = -\langle t_\alpha, \xi\rangle \rho_\alpha$$

and hence

$$e^{\mathbf{i}t\xi}\rho e^{-\mathbf{i}t\xi} = \rho_0 + \sum_{\alpha\in R} e^{-\langle t_\alpha,\xi\rangle t}\rho_\alpha.$$

This converges for $t \to \infty$ if and only if $\rho_\alpha = 0$ for all $\alpha \notin R(\xi)$. Similarly, we have $\rho = e^{\mathbf{i}\xi}\rho e^{-\mathbf{i}\xi}$ if and only if $\rho_\alpha = 0$ for all $\alpha \notin \tilde{R}(\xi)$. □

We could now define the parabolic subgroup $Q(\xi)$ and its Levi subgroup $L(\xi)$ as those connected subgroups of G^c whose Lie algebras are given by $\mathfrak{q}(\xi)$ and $\mathfrak{l}(\xi)$ respectively. These are closed subgroups, since both agree with their normalizer in G^c.

Lemma 2.2.15. *Let $\check{t}_1, \ldots, \check{t}_r$ be defined by (2.7) and let*

$$\xi = x_1\check{t}_1 + \cdots + x_r\check{t}_r \in \Omega_W$$

with $x_j \geq 0$. Then $Q_j := Q(\check{t}_j)$ are maximal proper parabolic subgroups of G^c and $Q(\xi) \subset Q(\check{t}_j)$ if and only if $x_j > 0$. Moreover,

$$Q(\xi) = \bigcap_{\{j \,|\, x_j > 0\}} Q(\check{t}_j).$$

Proof. The proof is a simple matter of comparing $R(\check{t}_j)$ and $R(\xi)$. □

2.3 Algebraic and symplectic stability

Let G be a compact connected Lie group and let $P \to \Sigma$ be a principal G bundle over Σ. Denote by G^c the complexification of G and by $P^c := P \times_G G^c$ the complexified principal bundle.

The algebraic geometric construction of the moduli space of holomorphic structures on P^c, in the sense of Mumford's geometric invariant theory [88], depends on the notion of stable and semistable objects. For vector bundles this notion is due to Mumford [87] and it was later extended by Ramanathan [95] to principal bundles. We discuss these two definitions in the first subsection and denote the corresponding moduli space of holomorphic structures on P^c by

$$\mathcal{J}^{ss}(P^c)//\mathcal{G}(P^c).$$

As mentioned in the introduction, this space is obtained by identifying two orbits in $\mathcal{J}^{ss}(P)/\mathcal{G}(P^c)$ when they cannot be separated.

The $\mathcal{G}(P)$-action on $\mathcal{A}(P)$ is Hamiltonian with moment map $\mu(A) = *F_A$ by Lemma 2.2.2. For every central element $\tau \in Z(\mathfrak{g})$ one obtains the symplectic quotient

$$\mathcal{A}(P)//\mathcal{G}(P) := \mu^{-1}(\tau)/\mathcal{G}(P).$$

Note that the moment map is not uniquely determined by the gauge action and another moment map is given by $\mu_\tau(A) := *F_A - \tau$. In other words, different choices of τ correspond to different choices for the moment map. The symplectic version of GIT (see [51]) defines stable and semistable objects in $\mathcal{A}(P)$ in terms of the moment map. We show in the second subsection that there exists a natural choice for $\tau \in Z(\mathfrak{g})$ determined by the topological type of P and define the corresponding symplectic notion of stability. It will follow from Theorem 2.4.14 and Theorem 2.4.15 in the next section that this definitions leads to identifications

$$\mu_\tau^{-1}(0)/\mathcal{G}(P) \cong \mathcal{A}^{ss}(P)//\mathcal{G}^c(P).$$

The right hand side is again obtained by identifying orbits in $\mathcal{A}^{ss}(P)/\mathcal{G}^c(P)$ if they cannot be separated.

Recall from Lemma 2.2.5 that $\mathcal{J}(P^c)$ can naturally be identified with $\mathcal{A}(P)$. We prove in Theorem 2.3.10 that the different notions of stability on $\mathcal{A}(P)$ and $\mathcal{J}(P^c)$ are essentially equivalent under this identification. In particular, this yields isomorphism

$$\mathcal{J}^{ss}(P^c)//\mathcal{G}(P^c) \cong \mathcal{A}^{ss}(P)//\mathcal{G}^c(P) \cong \mu_\tau^{-1}(0)/\mathcal{G}(P)$$

for a suitable choice of $\tau \in Z(\mathfrak{g})$. The proof of this theorem will be based on the whole remainder of the chapter, namely on Proposition 2.5.9, the moment-weight inequality (Theorem 2.5.12), the Harder-Narasimhan-Ramanathan theorem (Theorem 2.6.5) and the dominant weight theorem (Theorem 2.7.1).

2.3.1 Algebraic stability

We discuss the algebraic notion of stability on the space $\mathcal{J}(P^c)$ of holomorphic structures on the principal G^c bundle P^c. This definition depends only on the complexified bundle P^c itself and not on the reduction $P \subset P^c$. Consider as a warmup the case $G^c = \mathrm{GL}(n)$. This allows us to identify P^c with a complex vector bundle. The slope or normalized Chern class of a vector bundle $E \to \Sigma$ is defined as

$$\mu(E) := \frac{c_1(E)}{\mathrm{rk}(E)}.$$

The following Definition is due to Mumford [87].

Definition 2.3.1. *Let $E \to \Sigma$ be a holomorphic vector bundle.*

1. *E is called **stable** if for every proper holomorphic subbundle $0 \neq F \subset E$ we have $\mu(F) < \mu(E)$.*
2. *E is called **polystable** if E is the direct sum of stable vector bundles all having the same slope.*
3. *E is called **semistable** if for every proper holomorphic subbundle $0 \neq F \subset E$ we have $\mu(F) \leq \mu(E)$.*
4. *E is called **unstable** if E is not semistable.*

The analogue of this definition for general Lie groups was formulated by Ramanathan [95]. Lemma 2.3.4 below shows that Definition 2.3.1 corresponds to the special case $G^c = \mathrm{GL}(n)$ in Definition 2.3.2.

Definition 2.3.2. *Let G^c be a connected reductive Lie group and $P^c \to \Sigma$ be a holomorphic principal G^c bundle.*

1. *P^c is called **stable** if for every holomorphic reduction $P_Q \subset P^c$ to a maximal proper parabolic subgroup $Q \subset G^c$ the subbundle $ad(P_Q) \subset ad(P^c)$ satisfies $c_1(ad(P_Q)) < 0$.*
2. *P^c is called **polystable** if there exists a parabolic subgroup $Q \subset G^c$ and a holomorphic reduction $P_L \subset P^c$ to a Levi subgroup of Q satisfying the following*
   (a) P_L is a stable principal L bundle.
   (b) For every character $\chi : L \to \mathbb{C}^*$, which is trivial on the center of G^c, the associated line bundle $\chi(P_L) := P_L \times_\chi \mathbb{C}$ satisfies $c_1(\chi(P_L)) = 0$.
3. *P^c is called **semistable** if for every holomorphic reduction $P_Q \subset P^c$ to a maximal proper parabolic subgroup $Q \subset G^c$ the subbundle $ad(P_Q) \subset ad(P^c)$ satisfies $c_1(ad(Q)) \leq 0$.*
4. *P^c is called **unstable** if $ad(P^c)$ is not semistable.*

Remark 2.3.3. Let $L_1, \cdots, L_r$ and G^c be complex connected reductive Lie groups such that the product $L_1 \times \cdots \times L_r \subset G^c$ embeds as a subgroup. Let P_j be stable principal L_j bundles for $j = 1, \ldots, r$. Then it is easy to see that $P_L := P_{L_1} \times \cdots \times P_{L_r}$ is a stable principal L bundle. However, the extension $P^c := P_L \times_L G^c$ is in general not a semistable G^c-bundle. The second condition in the definition of polystability in needed to guarantee the semistability of P^c. To see this let $P_{Q'} \subset P^c$ be the reduction to a maximal parabolic subgroup and consider the determinant of the adjoint action of $Q' \subset G^c$ on its Lie algebra. This character is clearly trivial on the center of G^c and either restricts to L or to a maximal parabolic subgroup $Q'' = Q' \cap L \subset L$. In the first case, it follows from the definition of polystability that $c_1(\mathrm{ad}(P_{Q'})) = 0$. In the other case observe that $P_{Q'}$ determines a maximal parabolic reduction $P_{Q''} \subset P_L$ and $c_1(\mathrm{ad}(P_{Q'})) = c_1(\mathrm{ad}(P_{Q''})) < 0$, since P_L is stable.

Lemma 2.3.4. *A holomorphic vector bundle E is stable, polystable, semistable or unstable if and only its $GL(n)$-frame bundle $P^c := Fr(E)$ is stable, polystable, semistable or unstable respectively.*

Proof. We discuss the stable (resp. semistable) case first. A maximal parabolic subgroup of $GL(n)$ is the stabilizer a subspace $0 \neq V \subset \mathbb{C}^n$ and the holomorphic reduction P_Q of the $\mathrm{GL}(n)$-frame bundle to a maximal parabolic subgroup is thus the stabilizer of a holomorphic subbundle $F \subset E$. Consider the orthogonal splitting $E = F \oplus G$ with respect to some fixed hermitian metric on E. Then $\mathrm{ad}(P_Q) \subset \mathrm{End}(E)$ is given by the space of upper block diagonal matrices. We choose unitary connections A_1 on E and A_2 of G and denote by A the induced connection of $E = F \oplus G$. This induces also a connection on $\mathrm{ad}(P_Q)$ and the curvature of this connection is given by the endomorphism

$$\xi \mapsto F_A\xi - \xi F_A$$

for $\xi \in \mathrm{ad}(P_Q)$. Since $F_A = \mathrm{diag}(F_{A_1}, F_{A_2})$ is block-diagonal, a short calculation shows that the trace of this map is given by $\mathrm{rk}(G)\mathrm{tr}(F_{A_1}) - \mathrm{rk}(F)\mathrm{tr}(F_{A_2})$ and Chern-Weyl theory yields

$$\begin{aligned} c_1(\mathrm{ad}(P_Q)) &= \mathrm{rk}(G)c_1(F) - \mathrm{rk}(F)c_1(G) \\ &= \mathrm{rk}(E/F)\mathrm{rk}(F)\left(\frac{c_1(F)}{\mathrm{rk}(F)} - \frac{c_1(E/F)}{\mathrm{rk}(E/F)}\right). \end{aligned}$$

This expression is nonpositive if and only if $c_1(F)/\mathrm{rk}(F) \leq c_1(E)/\mathrm{rk}(E)$ and negative whenever strict inequality holds. This proves the equivalence of both definitions in the stable and semistable case.

The unstable case is equivalent to the semistable case and it remains to discuss the polystable case. A general parabolic subgroup of $\mathrm{GL}(V)$ is the stabilizer of a filtration $V_1 \subset \cdots \subset V_r = V$ and a Levi subgroup in given as the stabilizer of a splitting $V = W_1 \oplus \cdots \oplus W_r$ with $V_j = W_1 \oplus \cdots \oplus W_j$. Hence, a holomorphic reduction $P_L \subset P^c$ to the Levi factor of a parabolic subgroup corresponds to the $L = \mathrm{GL}(n_1) \times \cdots \times \mathrm{GL}(n_r)$ frame bundle of a holomorphic splitting

$$E = E_1 \oplus \cdots \oplus E_r.$$

We claim that P_L is a stable principal L bundle if and only if all factors E_j are stable holomorphic vector bundles. Indeed, a maximal parabolic subgroup of L has the shape

$$Q = \mathrm{GL}(n_1) \oplus \cdots \oplus \mathrm{GL}(n_{j-1}) \oplus Q_j \oplus \mathrm{GL}(n_{j+1}) \cdots \oplus \mathrm{GL}(n_r)$$

where $Q_j \subset \mathrm{GL}(n_j)$ is a maximal parabolic subgroup. Then

$$\mathrm{ad}(P_Q) = \mathrm{End}(E_1) \oplus \cdots \oplus \mathrm{ad}(P_{Q_j}) \oplus \cdots \oplus \mathrm{End}(E_r)$$

and hence $c_1(\mathrm{ad}(P_Q)) = c_1(\mathrm{ad}(P_{Q_j}))$. The claim follows now from our discussion of the stable case.

It remains to verify that the slopes of all subbundles satisfy $\mu(E_j) = \mu(E)$ if and only if for every character $\chi : L \to \mathbb{C}^*$ which is trivial on the center of $\mathrm{GL}(n)$ the associated line bundle $\chi(P_L)$ has degree zero. Every character $\chi : L \to \mathbb{C}^*$ factors as $\chi = \chi_1 \cdots \chi_r$ with $\chi_j : \mathrm{GL}(n_i) \to \mathbb{C}^*$ and induces on the Lie algebra the representation

$$\dot\chi = \dot\chi_1 + \cdots + \dot\chi_r$$

with $\dot\chi_j := d\chi_j(\mathbb{1}) : \mathfrak{gl}(n_j) \to \mathbb{C}$. Since every traceless matrix in $\mathfrak{gl}(n_j)$ is a commutator, there exist $\lambda_j \in \mathbb{C}$ such that

$$\dot\chi_j(\rho_j) = \lambda_j \mathrm{tr}(\rho_j)$$

for all $\rho_j \in \mathfrak{gl}(n_j)$. We choose unitary connections $A_j \in \mathcal{A}(E_j)$ and denote by $A = A_1 \oplus \cdots \oplus A_r$ the induced unitary connection on E. Then follows from Chern-Weil theory

$$\begin{aligned} c_1(\chi(P_L)) &= \frac{\mathbf{i}}{2\pi} \int_\Sigma F_{\chi(A)} \, dvol_\Sigma = \frac{\mathbf{i}}{2\pi} \int_\Sigma \dot\chi(F_A) \, dvol_\Sigma \\ &= \lambda_1 c_1(E_1) + \cdots + \lambda_r c_1(E_r). \end{aligned}$$

Note that χ vanishes on the center of $\mathrm{GL}(n)$ if and only if $n_1\lambda_1 + \cdots + n_r\lambda_r = 0$ is satisfied. If in addition $\mu(E_j) = \mu(E)$ holds for all j, then

$$c_1(\chi(P_L)) = \sum_{j=1}^r n_j \lambda_j \mu(E_j) = 0.$$

For the converse consider the character $\chi : \mathrm{GL}(n_1) \times \cdots \times \mathrm{GL}(n_r) \to \mathbb{C}^*$

$$\chi(B_1, \ldots, B_r) := \det(B_j)^n \det(B)^{-n_j}.$$

This vanishes on the center of $\mathrm{GL}(n)$ and satisfies $\dot\chi(\xi) = n\mathrm{tr}(\xi_j) - n_i \mathrm{tr}(\xi)$. Hence

$$c_1(\chi(P_L)) = nc_1(E_j) - n_j c_1(E)$$

and this vanishes precisely when $\mu(E_j) = \mu(E)$ is satisfied. □

The next lemma shows that we can always reduce to the case where G^c has discrete center.

Lemma 2.3.5. *Let G^c be a complex connected reductive Lie group and $P^c \to \Sigma$ be a principal G^c bundle. Denote by $Z_0(G^c)$ the connected component of the center of G^c containing the identity. Let $H^c := G^c/Z_0(G^c)$ and denote by*

$$P_{H^c} := P^c/Z_0(G^c)$$

the associated H^c bundle. This carries a natural induced holomorphic structure and P^c is stable, polystable, semistable or unstable if and only if P_{H^c} is stable, polystable, semistable or unstable respectively.

Proof. The Lie algebra of G^c splits as $\mathfrak{g}^c = Z(\mathfrak{g}^c) \oplus [\mathfrak{g}^c, \mathfrak{g}^c]$ and $[\mathfrak{g}^c, \mathfrak{g}^c]$ can be identified with the semisimple Lie algebra of H^c. This splitting is preserved by the adjoint action of G^c and produces a splitting $\mathrm{ad}(P^c) = V \oplus \mathrm{ad}(P_{H^c})$ where $V = \Sigma \times Z(\mathfrak{g}^c)$ is a trivial bundle. Parabolic subgroups $Q \subset G^c$ correspond bijectively to parabolic subgroups $\bar{Q} := Q/Z_0(G^c) \subset H$ and parabolic reductions $P_Q \subset P^c$ correspond bijectively to parabolic reductions $P_{\bar{Q}} := P_Q/Z_0(G^c) \subset P_{H^c}$. Since $\mathrm{ad}(P_Q) = V \oplus \mathrm{ad}(P_{\bar{Q}})$, we have $c_1(\mathrm{ad}(P_Q)) = c_1(\mathrm{ad}(P_{\bar{Q}}))$ and this shows that P^c is stable (resp. semistable) if and only if P_{H^c} is stable (resp. semistable).

If L is a Levi subgroup of the parabolic subgroup $Q \subset G^c$, then $\bar{L} := L/Z_0$ is a Levi-subgroup of $\bar{Q} = Q/Z_0 \subset H^c$. Moreover, reductions $P_L \subset P^c$ to L correspond bijectively to reductions $P_{\bar{L}} = P_L/Z_0(G^c) \subset P_{H^c}$. We have already shown that P_L is stable if and only if $P_{\bar{L}}$ is stable. The characters $\chi : Q \to \mathbb{C}^*$ which are trivial on the center $Z(G^c)$ of G^c correspond bijectively to the characters $\bar{\chi} : \bar{Q} \to \mathbb{C}^*$ which are trivial on $Z(G^c)/Z_0(G^c)$ and

$$\chi(P_Q) \cong \bar{\chi}(P_{\bar{Q}}).$$

Thus P^c is polystable if and only if P_{H^c} is polystable. $\square$

2.3.2 Symplectic stability

Let G be a compact connected Lie group and $P \to \Sigma$ a principal G bundle. Let $\chi : G \to S^1$ be a character and denote by $\dot{\chi} = d\chi(\mathbb{1}) : \mathfrak{g} \to \mathbf{i}\mathbb{R}$ the induced character on the Lie algebra. Since $\mathfrak{g} = Z(\mathfrak{g}) \oplus [\mathfrak{g}, \mathfrak{g}]$, we may identify $-2\pi\mathbf{i}\dot{\chi}$ with an element in $Z(\mathfrak{g})^* = \mathrm{Hom}(Z(\mathfrak{g}), \mathbb{R})$. Denote by $\chi(P) := P \times_\chi \mathbb{C}$ the line bundle associated to P via χ. Then

$$c_1(\chi(P)) = \frac{\mathbf{i}}{2\pi} \int_\Sigma \dot{\chi}(F_A) \tag{2.11}$$

for any connection $A \in \mathcal{A}(P)$. The assignment $\dot{\chi} \mapsto -2\pi\mathbf{i}c_1(\chi(P))$ extends to a unique element in $Z(\mathfrak{g})^{**}$, since the lattice of all infinitesimal characters spans $Z(\mathfrak{g})^*$ as a vector space. This corresponds under the canonical isomorphism $Z(\mathfrak{g})^{**} \cong Z(\mathfrak{g})$ to an element $\tau \in Z(\mathfrak{g})$ which satisfies

$$\alpha(\tau) = \int_\Sigma \alpha(F_A) \qquad \text{for all } \alpha \in Z(\mathfrak{g})^* \text{ and } A \in \mathcal{A}(P). \tag{2.12}$$

Here we identify $Z(\mathfrak{g})^* \subset \mathfrak{g}^*$ with the subspace of linear functionals vanishing on $[\mathfrak{g}, \mathfrak{g}]$. We call τ the **central type** of P.

Remark 2.3.6. Recall our standing assumption $\mathrm{vol}(\Sigma) = 1$ and suppose that $A \in \mathcal{A}(P)$ satisfies $*F_A = \lambda$ for some $\lambda \in Z(\mathfrak{g})$. Then (2.12) yields

$$\alpha(\tau) = \int_\Sigma \alpha(F_A) = \int_\Sigma \alpha(\lambda)\, dvol_\Sigma = \alpha(\lambda)$$

for all $\alpha \in Z(\mathfrak{g})^*$ and hence $\lambda = \tau$.

Let $\tau \in Z(\mathfrak{g})$ be defined by (2.12). It follows from Lemma 2.2.7 that

$$\mu_\tau : \mathcal{A}(P) \to L^2(\Sigma, \mathrm{ad}(P)), \qquad \mu_\tau(A) := *F_A - \tau \tag{2.13}$$

is a moment map for the $\mathcal{G}(A)$-action on $\mathcal{A}(P)$. The following definition is the precise analogue of Definition 7.1 in [51] with respect to this moment map.

Definition 2.3.7. *Let G be a compact connected Lie group, let $P \to \Sigma$ be a principal G bundle with central type $\tau \in Z(\mathfrak{g})$ defined by (2.12), and define μ_τ by (2.13). For $A \in \mathcal{A}(P)$ denote by $\overline{\mathcal{G}^c(A)}$ the $W^{1,2}$-closure of the complex gauge orbit $\mathcal{G}^c(A)$.*

1. *A is called μ_τ-**stable** if and only if $\mathcal{G}^c(A) \cap \mu_\tau^{-1}(0) \cap \mathcal{A}^*(P) \neq \emptyset$ where $\mathcal{A}^*(P)$ denotes the irreducible connections on P.*
2. *A is called μ_τ-**polystable** if and only if $\mathcal{G}^c(A) \cap \mu_\tau^{-1}(0) \neq \emptyset$.*
3. *A is called μ_τ-**semistable** if and only if $\overline{\mathcal{G}^c(A)} \cap \mu_\tau^{-1}(0) \neq \emptyset$.*
4. *A is called μ_τ-**unstable** if and only if $\overline{\mathcal{G}^c(A)} \cap \mu_\tau^{-1}(0) = \emptyset$.*

Remark 2.3.8. We call $A \in \mathcal{A}(P)$ an irreducible connection if the map

$$d_A : W^{2,2}(\Sigma, \mathrm{ad}(P)) \to W^{1,2}(\Sigma, T^*\Sigma \otimes \mathrm{ad}(P))$$

is injective. In particular, μ_τ-stable connections can only exist when the center of G is discrete and $\tau = 0$. Since the infinitesimal action of the gauge group is given by $L_A\xi := -d_A\xi$, a connection A is irreducible if and only if the isotropy group of the orbit $\mathcal{G}(A)$ is discrete. Suppose that A is an irreducible connection satisfying $\mu_0(A) = *F_A = 0$. The infinitesimal action of the complexified gauge group

$$L_A(\xi + \mathbf{i}\eta) = -d_A\xi - *d_A\eta$$

is readily seen to be injective in this case: Assume $L_A(\xi + \mathbf{i}\eta) = 0$ and apply d_A to the equation. Then follows $d_A^* d_A\eta = 0$ and hence $d_A\eta = 0$. Since A is irreducible, we conclude $\eta = 0$ and then $\xi = 0$. This argument shows that the μ_τ-stable orbits are precisely the μ_τ-polystable orbits with discrete $\mathcal{G}^c(P)$ isotropy.

The next lemma relates the different notions of stability on P and on the quotient bundle $P_H := P/Z_0(G)$ with fiber $H := G/Z_0(G)$. Note that P_H has central type 0 since its center is discrete.

Lemma 2.3.9. *Let G be a compact connected Lie group, let $P \to \Sigma$ be a principal G bundle of central type $\tau \in \mathfrak{g}$ defined by (2.12) and let $P_H := P/Z_0(G)$ be the associated $H := G/Z_0(G)$ bundle. Let $A \in \mathcal{A}(P)$ and denote by $A_H \in \mathcal{A}(P_H)$ the induced connection.*

1. *A_H is μ_0-stable if and only if A is μ_τ-polystable and the kernel of the infinitesimal action*

$$L_A : W^{2,2}(\Sigma, ad(P^c)) \to W^{1,2}(\Sigma, T^*\Sigma \otimes ad(P))$$

$$L_A(\xi + \boldsymbol{i}\eta) = -d_A\xi - *d_A\eta$$

consists of constant central sections.

2. *A_H is μ_0-polystable if and only if A is μ_τ-polystable.*

3. *A_H is μ_0-semistable if and only if A is μ_τ-semistable.*

4. *A_H is μ_0-unstable if and only if A is μ_τ-unstable.*

Proof. We begin with the polystable case. Every constant central curvature connections on P clearly induces a flat connection on P_H. Conversely, assume that A_1 is a flat connection on P_H. As a general property of compact Lie groups, there exists an exact sequence

$$1 \to F \to Z_0(G) \times [G, G] \to G \to 1 \tag{2.14}$$

where $F = Z_0(G) \cap [G, G]$ is a finite group. From this follows the exact sequence

$$1 \to F \to G \to (G/Z_0(G)) \times (Z_0(G)/F) \to 1. \tag{2.15}$$

Consider the associated $(G/Z_0(G)) \times (Z_0(G)/F)$ bundle

$$\tilde{P} = P \times_G ((G/Z_0(G)) \times (Z_0(G)/F)) = P_H \times_\Sigma P_2 \tag{2.16}$$

where P_2 is a principal $Z_0(G)/F$-bundle over Σ. Since $Z_0(G)/F$ is connected and abelian, it is a torus and P_2 is isomorphic to the direct sum of S^1 bundles. It follows from Hodge theory that every line bundle admits a connection with constant central curvature and these yield a connection A_2 on P_2 with constant central curvature. Together with A_1 we obtain an induces a connection on $\tilde{P}$ which lifts to a connection on P with constant central curvature. It follows from Remark 2.3.6 that the curvature of this connection is given by τ.

For the proof of the stable case observe that $\mathrm{ad}(P) \cong V \oplus \mathrm{ad}(P_H)$ where $V = \Sigma \times Z(\mathfrak{g})$ denotes the trivial $Z(\mathfrak{g})$ bundle. The infinitesimal action

$$L_A : W^{2,2}(\Sigma, \mathrm{ad}(P)) \to W^{1,2}(\Sigma, T^*\Sigma \otimes \mathrm{ad}(P))$$

agrees with L_{A_H} on $\mathrm{ad}(P_H)$. Since d_A restricts to a flat connection on V, it follows that $\ker(L_A) \cong Z(\mathfrak{g}) \oplus \ker(L_{A_H})$ and this shows the claim.

It remains to discuss the semistable case. Assume first that A is μ_τ-semistable. Then exist connections $A^k \in \mathcal{G}^c(A)$ such that $A^k \to A^+$ for $k \to \infty$ and $\mu_\tau(A^+) = 0$. The induced connections $A_H^k \in \mathcal{A}(P_H)$ are clearly contained in $\mathcal{G}^c(A_H)$ and converge to the induced connection A_H^+. Since $\mu_\tau(A^+) = 0$, it follows that $\mu_0(A_H^+) = 0$ and hence A_H is μ_0-semistable.

For the converse, we consider the exact sequences (2.14) and (2.15) from above. Then (2.16) yields a finite covering

$$P \to \tilde{P} = P_H \times_\Sigma P_2$$

with covering group $F = Z_0(G) \cap [G, G]$. We have seen above that P_2 is a polystable $Z_0(G)/F$-bundle. Note that the natural identification $\mathcal{A}(\tilde{P}) = \mathcal{A}(P_H) \times \mathcal{A}(P_2)$ yields an inclusion

$$\mathcal{G}^c(P_H) \times \mathcal{G}^c(P_2) \subset \mathcal{G}^c(\tilde{P}). \tag{2.17}$$

Moreover, since $\mathrm{Ad}(P^c) \to \mathrm{Ad}(\tilde{P}^c)$ is a finite covering with covering group $F \subset Z_0(G^c)$, it is easy to see that every gauge transformation in $\mathcal{G}^c(\tilde{P})$ lifts to an element in $\mathcal{G}^c(P)$ and this lift commutes with the natural identification $\mathcal{A}(P) = \mathcal{A}(\tilde{P})$.

Now assume that $A \in \mathcal{A}(P)$ induces a μ_0-semistable connection $A_H \in \mathcal{A}(P_H)$. Since P_2 is polystable, it follows from (2.17) that there exists $g_0 \in \mathcal{G}^c(P)$ such that $g_0(A)$ induces $A_H \in \mathcal{A}(P_H)$ and a connection $A_2 \in \mathcal{A}(P_2)$ with constant central curvature. Since A_H is μ_0-semistable, using (2.16) again, there exists gauge transformations $g_k \in \mathcal{G}^c(P)$ such that $g_k(g_0(A))$ induce the same connection A_2 on P_2 and induce a sequence of connections A_H^k on P_H which converges to a flat connection A_H^+. Clearly, $g_k(g_0 A)$ converges to the connection A^+ which is induced by A_2 and A_H^+. Hence A^+ has constant central curvature and it follows from Remark 2.3.6 that $*F_{A^+} = \tau$. This completes the proof of the semistable case.

□

2.3.3 Equivalence of algebraic and symplectic stability

The following theorem shows that the algebraic notion of stability from Definition 2.3.2 and the symplectic notion of μ_τ-stability from Definition 2.3.2 are essentially equivalent.

Theorem 2.3.10 (Generalized Narasimhan-Seshadri-Ramanathan Theorem). *Let G be a compact connected Lie group and $P \to \Sigma$ a principal G bundle with central type $\tau \in Z(\mathfrak{g})$ defined by (2.12). Let $A \in \mathcal{A}(P)$ and consider the complexified bundle $P^c := P \times_G G^c$ with the induced holomorphic structure J_A.*

1. *(P^c, J_A) is stable if and only if A is μ_τ-polystable and the kernel of*

$$L_A : W^{2,2}(\Sigma, ad(P^c)) \to W^{1,2}(\Sigma, T^*\Sigma \otimes ad(P))$$

$$L_A(\xi + \boldsymbol{i}\eta) = -d_A\xi - *d_A\eta$$

   contains only constant central sections.

2. *(P^c, J_A) is polystable if and only if A is μ_τ-polystable.*

3. *(P^c, J_A) is semistable if and only if A is μ_τ-semistable.*

4. *(P^c, J_A) is unstable if and only if A is μ_τ-unstable.*

The stable case was first proven by Narasimhan and Seshadri [91] for $G = U(n)$ and later extended by Ramanathan [95] to arbitrary compact Lie groups. They establish these results using algebraic geometric methods.The first analytic proof was given by Donaldson [30] for the case $G = U(n)$. We present a different approach given by Bradlow [12] and Mundet [89] in Theorem 2.6.5. The equivalence of both definitions for semistability is essentially contained in the work of Atiyah and Bott [4].

Proof of Theorem 2.3.10. We assume the following results for the proof:

- the characterization of algebraic stability in Proposition 2.5.9
- the moment-weight inequality (Theorem 2.5.12)

- the Narasimhan-Seshadri-Ramanathan theorem (Theorem 2.6.5)
- the dominant weight theorem (Theorem 2.7.1)

We establish these results independently in the remainder of this chapter.

The stable case is equivalent to Theorem 2.6.5. By Lemma 2.3.5 and Lemma 2.3.9 we may assume in the following that $Z(G)$ is discrete and $\tau = 0$. We then deduce the polystable case from the stable case by an inductive argument: Assume first that P^c is polystable. Then there exists a reductive subgroup $L \subset G^c$ and a holomorphic reduction $P_L \subset P^c$ which is stable. We may assume that $L = K^c$ is the complexification of a compact subgroup $K \subset G$. Since $G^c/L \cong G/K$, we have an induced reduction $P_K \subset P$ and P_L agrees with the complexification of P_K. It follows from the construction in Lemma 2.2.5 that A restricts to a connection on P_K. Assuming the stable case (i.e. Theorem 2.6.5) we conclude that there exists a gauge transformation $g \in \mathcal{G}^c(P_K) \subset \mathcal{G}^c(P)$ such that $*F_{gA} = \tau_K \in Z(\mathfrak{k})$. It remains to show that $\tau_K \in Z(\mathfrak{g}) = 0$ vanishes. If $\tau_K \neq 0$ then exists a character $\chi : L \to \mathbb{C}^*$ with $\dot\chi(\tau_K) \neq 0$. Since $Z(G)$ is finite, we may replacing χ by a suitable power and assume that it is trivial on $Z(G^c)$. Using the definition of polystability then yields the contradiction

$$0 = c_1(\chi(P_L)) = \frac{\mathbf{i}}{2\pi}\int_\Sigma \dot\chi(F_A)\, dvol_\Sigma = \frac{\mathbf{i}}{2\pi}\dot\chi(\tau_K) \neq 0.$$

For the converse, assume that $A \in \mathcal{A}(P)$ is a flat connection. Let $H \subset G$ be the holonomy subgroup and $P_H \subset P$ be a reduction to the holonomy. Let $K := C_G(Z(H))$ be the centralizer of the center of the holonomy and denote the induced connection on $P_K = P_H \times_H K$ again by A. It is well-known that the isotropy subgroup of A consists of constant gauge transformations and is naturally isomorphic to the centralizer of its holonomy, i.e.

$$\mathcal{G}_A := \{g \in \mathcal{G}(P_K) \,|\, g(A) = A\} \cong C_K(H).$$

Comparing the Lie algebras of both sides, one checks that $C_K(H) = Z_0(K)$ is satisfied and $A \in \mathcal{A}(P_K)$ has only trivial isotropy. It follows now from the stable case (i.e. Theorem 2.6.5) that P_K^c is a stable principal $L = K^c$ bundle. Note that L is a Levi-subgroup of a parabolic subgroup of G^c, since K is the centralizer of a torus in G. Since $F_A = 0$, we have for any character $\chi : L \to \mathbb{C}^*$

$$c_1(\chi(P_L)) = \frac{\mathbf{i}}{2\pi}\int_\Sigma \dot\chi(F_A)\, dvol_\Sigma = 0$$

and hence P^c is polystable.

Assume that P^c is unstable. By Proposition 2.5.9 there exists $\xi \in \Omega^0(\Sigma, \mathrm{ad}(P))$ with $w_0(A,\xi) < 0$. The moment-weight inequality (Theorem 2.5.12) yields $\mu_0(gA) \geq -w_0(A,\xi)/||\xi|| > 0$ for all $g \in \mathcal{G}^c(P)$ and hence A is μ_0-unstable. Assume conversely that A is μ_0-unstable. The dominant-weight theorem (Theorem 2.7.1) shows that there exists $\xi \in \Omega^0(\Sigma, \mathrm{ad}(P))$ such that $w_0(A,\xi) < 0$ and hence P^c is unstable by Proposition 2.5.9. This completes the proof of the unstable case and the semistable case is equivalent to this case. □

2.4 The Yang-Mills flow and symplectic stability

Let G be a compact connected Lie group and let $P \to \Sigma$ be a principal G bundle of central type $\tau \in Z(\mathfrak{g})$ defined by (2.12). In the differential geometric approach towards GIT the moment map squared functional plays a crucial role. This is defined by

$$\mathcal{F}_\tau : \mathcal{A}(P) \to \mathbb{R}, \qquad \mathcal{F}_\tau(A) := \frac{1}{2}\int_\Sigma || * F_A - \tau||^2 \, dvol_\Sigma. \tag{2.18}$$

Note that (2.12) implies $\int_\Sigma \langle F_A, \tau\rangle = ||\tau||^2$ for every connection $A \in \mathcal{A}(P)$ and hence

$$\mathcal{F}_\tau(A) = \frac{1}{2}\int_\Sigma || * F_A - \tau||^2 dvol_\Sigma = \frac{1}{2}\left(\int_\Sigma ||F_A||^2 dvol_\Sigma - ||\tau||^2\right). \tag{2.19}$$

Thus $\mathcal{F}_\tau$ agrees up to a constant shift with the Yang-Mills functional

$$\mathcal{YM} : \mathcal{A}(P) \to \mathbb{R}, \qquad \mathcal{YM}(A) := \frac{1}{2}\int_\Sigma ||F_A||^2 \, dvol_\Sigma. \tag{2.20}$$

Råde showed in his thesis [97] that the negative gradient flow of the Yang-Mills functional is well-defined and converges if the base manifold has dimension 2 or 3. We summarize his results in the first subsection. Recall that we always consider the $W^{1,2}$-topology on $\mathcal{A}(P)$ when nothing else is specified.

A crucial observation is the following: Any solution of the Yang-Mills flow remains in a single complexified orbit and there exists a canonical lift of a solution $A(t)$ of the Yang-Mills flow under the projection $\mathcal{G}^c \to \mathcal{G}^c(A)$ to a curve in $\mathcal{G}^c(P)$. Since the Yang-Mills flow is $\mathcal{G}(P)$-invariant, the geometric importance lies within the projection of such curves in $\mathcal{G}^c(P)/\mathcal{G}(P)$. The fibers of this quotient coincide with the homogeneous space G^c/G which is a complete, connected, simply connected Riemannian manifold of nonpositive sectional curvature (see [51] Appendix A and B). This underlying geometry is crucial for the following application.

As a first application, we establish the moment limit theorem (Theorem 2.4.14) and the analogue of the Ness uniqueness theorem in Theorem 2.4.15 following the line of arguments in [51]. The first result says that the limit $A_\infty := \lim_{t\to\infty} A(t)$ of the Yang-Mills flow starting at $A_0 \in \mathcal{A}(P)$ minimizes the Yang-Mills functional over the complexified orbit $\mathcal{G}^c(A_0)$. The second result asserts that any connection in the $W^{1,2}$-closure of $\mathcal{G}(A_0)$ which minimizes the Yang-Mills functional over this orbit must be contained in $\mathcal{G}(A_\infty)$. In particular, every μ_τ-semistable orbit contains a unique μ_τ-polystable orbit in its closure. This yields the identification

$$\mathcal{A}^{ss}(P)//\mathcal{G}^c(P) \cong \mathcal{A}^{ps}/\mathcal{G}^c(P) \cong \mu_\tau^{-1}(0)/\mathcal{G}(P)$$

where two semistable orbits on the left hand side are identified if and only if they contain the same polystable orbit in their closure.

In the last section we extend this observation and characterize in Theorem 2.4.18 the μ_τ-stability of $A \in \mathcal{A}(P)$ in terms of the limit of the Yang-Mills flow starting at A. We observe in particular that $\mathcal{A}^{ss}(P)$ and $\mathcal{A}^s(P)$ are both open subsets of $\mathcal{A}(P)$ in the $W^{1,2}$-topology.

2.4.1 Analytical foundations

The Yang Mills flow on low dimensional manifolds

Recall for $A \in \mathcal{A}(P)$ and $a \in W^{1,2}(\Sigma, T^*\Sigma \otimes \mathrm{ad}(P))$ the formula

$$F_{A+a} = F_A + d_A a + \frac{1}{2}[a \wedge a].$$

From this follows directly that the L^2-gradient of the Yang-Mills functional (2.20) is given by

$$\nabla \mathcal{YM} : \mathcal{A}(P) \to W^{-1,2}(\Sigma, \mathrm{ad}(P)), \qquad \nabla \mathcal{YM}(A) := d_A^* F_A.$$

The critical points of the Yang-Mills functional (2.20) are called Yang-Mills connections and satisfy the equation

$$d_A^* F_A = 0.$$

It follows from the strong Uhlenbeck compactness result (see e.g. [122] Theorem E) and elliptic regularity that every Yang-Mills connection is gauge equivalent to a smooth Yang-Mills connection and the set $\Lambda := \{\mathcal{YM}(A) \,|\, d_A^* F_A = 0\}$ of critical values is discrete. The negative gradient flow of the Yang-Mills functional is given by the degenerated parabolic equation

$$\partial_t A(t) + d_{A(t)}^* F_{A(t)} = 0. \tag{2.21}$$

Definition 2.4.1 (**Weak solutions**). *Let $A_0 \in \mathcal{A}(P)$ be a connection of Sobolev class $W^{1,2}$. We call $A \in C^0([0,\infty), \mathcal{A}(P))$ a weak solution of the initial value problem*

$$\partial_t A(t) + d_{A(t)}^* F_{A(t)} = 0, \qquad A(0) = A_0 \tag{2.22}$$

if $A(0) = A_0$ and there exists a sequence $A_k : [0,\infty) \to \mathcal{A}(P)$ of smooth solutions of (2.21) which converges in $C^0_{\ell oc}([0,\infty), \mathcal{A}(P))$ to A, where $\mathcal{A}(P)$ is endowed with the $W^{1,2}$-topology.

The next two theorems state that the initial value problem (2.22) has a unique (weak) solution for every initial data $A_0 \in \mathcal{A}(P)$ existing for all time and that this solution converges to a Yang-Mills connection.

Theorem 2.4.2 (**Long time existence**). *Let G be a compact connected Lie group, $P \to \Sigma$ a principal G bundle and $A_0 \in \mathcal{A}(P)$.*

1. *There exists a unique weak solution $A(t) \in C^0_{\ell oc}([0,\infty), \mathcal{A}(P))$ for the initial value problem (2.22). The curvature has the additional regularity properties $F_{A(t)} \in C^0_{\ell oc}([0,\infty), L^2)$ and $F_{A(t)} \in L^2_{\ell oc}([0,\infty), W^{1,2})$.*

2. *The solution $A(t)$ and its curvature $F_{A(t)}$ depend smoothly on the initial data A_0 in these topologies.*

3. *If A_0 is smooth, then the solution $A(t)$ is smooth and satisfies (2.21).*

Proof. This is Theorem 1 in [97]. □

Theorem 2.4.3 (Convergence). *Assume the setting of Theorem 2.4.2 and let* $A(t) \in C^0_{\ell oc}([0,\infty), \mathcal{A}(P))$ *be a weak solution of (2.22). Then exists a Yang-Mills connection* $A_\infty \in \mathcal{A}(P)$ *and constants* $c, \beta > 0$ *such that*

$$||A(t) - A_\infty||_{W^{1,2}} \leq ct^{-\beta}$$

holds for all times $t > 0$.

Proof. This is Theorem 2 in [97]. □

The key ingredient in the proof of the convergence result is the appropriate analogue of the Łojasiewicz gradient inequality. This approach was introduced by Simon [101] for a general class of evolution equations.

Proposition 2.4.4 (Łojasiewicz gradient inequality). *Let* $A_\infty \in \mathcal{A}(P)$ *be a Yang-Mills connection. There exist constants* $\epsilon > 0$, $\gamma \in [\frac{1}{2}, 1)$ *and* $c > 0$ *such that for every* $A \in \mathcal{A}(P)$ *with* $||A - A_\infty||_{W^{1,2}} < \epsilon$ *the estimate*

$$||d_A^* F_A||_{W^{-1,2}} \geq c|\mathcal{YM}(A) - \mathcal{YM}(A_\infty)|^\gamma$$

is satisfied.

Proof. This is Proposition 7.2 and (9.1) in [97]. □

In finite dimensions the Łojasiewicz inequality always guarantees convergence by some standard arguments. We recall these arguments in the following and discuss additional technical difficulties arising in the infinite dimensional setting. Suppose that $A(t)$ satisfies (2.22). It follows from the weak Uhlenbeck compactness (see [97], Proposition 7.1) that there exists a $\mathcal{G}(P)$-orbit $\mathcal{G}(A_\infty)$ of Yang-Mills connections such that

$$\inf_{t>0} \mathcal{YM}(A(t)) = \mathcal{YM}(A_\infty)$$

and for every $\delta > 0$ exists $T > 0$ and $g \in \mathcal{G}(P)$ such that

$$||A(T) - g(A_\infty)||_{W^{1,2}} < \delta.$$

Since the Yang-Mills functional and the Łojasiewicz inequality are invariant under the action of $\mathcal{G}(P)$, the constant $\epsilon = \epsilon(g(A_\infty)) > 0$ from the Łojasiewicz inequality does not depend on g. Now choose $\delta < \epsilon$ and define

$$\overline{T} := \inf\{t > T \,|\, ||A(t) - g(A_\infty)||_{W^{1,2}} \geq \epsilon\}.$$

For any $s_1, s_2 \in (T, \overline{T})$ with $s_1 < s_2$ we obtain

$$\begin{aligned} ||A(s_1) - A(s_2)||_{L^2} &\leq \int_{s_1}^{s_2} ||d_A^* F_A||_{L_2}\, dt \\ &\leq \int_{s_1}^{s_2} \frac{||d_A^* F_A||_{L_2}^2}{c|\mathcal{YM}(A) - \mathcal{YM}(A_\infty)|^\gamma}\, dt \\ &\leq \frac{1}{c}\left(\mathcal{YM}(A(s_1)) - \mathcal{YM}(A(s_2))\right)^{1-\gamma}. \end{aligned}$$

To conclude the convergence result, one needs to show $\overline{T} = \infty$ and extend the estimate above to the $W^{1,2}$-norm. Both can be achieved by using the following lemma.

Lemma 2.4.5. *Let $A_\infty \in \mathcal{A}(P)$ be a Yang-Mills connection and $\epsilon = \epsilon(A_\infty) > 0$ as in Proposition 2.4.4. There exists a constant $c > 0$ with the following significance: For every solution $A(t)$ of the Yang-Mills flow (2.22) and real numbers $0 \le s_1 \le s_2 - 1$ such that $||A(t) - A_\infty||_{W^{1,2}} \le \epsilon$ for all $t \in [s_1, s_2]$ we have*

$$\int_{s_1+1}^{s_2} ||d_A^* F_A||_{W^{1,2}}\, dt \le c \int_{s_1}^{s_2} ||d_A^* F_A||_{L^2}\, dt.$$

Proof. This is Lemma 7.3 in [97]. □

Now the calculation above yields

$$||A(s_1+1) - A(s_2)||_{W^{1,2}} \le C \left(\mathcal{YM}(A(s_1)) - \mathcal{YM}(A_\infty)\right)^{1-\gamma} \tag{2.23}$$

for any $T \le s_1 < s_1 + 1 < \bar{T}$. Since the solutions of the Yang-Mills flow depend continuously on the initial condition in the $C^0_{\ell oc}([0,\infty), W^{1,2})$ topology, there exists a constant $c_1 > 0$ such that

$$||A(T+t) - g(A_\infty)||_{W^{1,2}} \le c_1 ||A(T) - g(A_\infty)||_{W^{1,2}}$$

holds for all $t \in [0,1]$. This follows as we may view $g(A_\infty)$ as constant flow line and the constant c_1 depends only on the orbit $\mathcal{G}(A_\infty)$. For sufficiently small δ, we have $\delta c_1 < \epsilon$ and hence $\bar{T} > 1$. Then (2.23) yields

$$||A(T+1) - A(t)||_{W^{1,2}} \le C \left(\mathcal{YM}(A(T)) - \mathcal{YM}(A_\infty)\right)^{1-\gamma} \le C\delta^{1-\gamma}$$

for any $T + 1 \le t \le \bar{T}$. For sufficiently small $\delta > 0$ the right hand side is smaller than ϵ and this yields $\bar{T} = \infty$. The calculation above then shows then that the integral $\int_0^\infty ||\partial_t A(t)||_{W^{1,2}}\, dt < \infty$ is finite and $A(t)$ converges uniformly to a Yang-Mills connection $\tilde{A}_\infty$.

Replacing A_∞ in the argument above by the limiting connection $\tilde{A}_\infty$ yields

$$||A(t) - \tilde{A}_\infty||_{W^{1,2}} \le C \left(\mathcal{YM}(A(t)) - \mathcal{YM}(\tilde{A}_\infty)\right)^{1-\gamma}.$$

Let $T > 0$ be such that for every $t > T$ the Łojasiewicz inequality in Lemma 2.4.5 for $A(t)$ with respect to the Yang-Mills connection $\tilde{A}_\infty$. Then

$$\begin{aligned} \partial_t \left(\mathcal{YM}(A(t)) - \mathcal{YM}(\tilde{A}_\infty)\right) &= -||\nabla\mathcal{YM}(A(t))||_{L^2} \\ &\le \left(\mathcal{YM}(A(t)) - \mathcal{YM}(\tilde{A}_\infty)\right)^{2\gamma} \end{aligned}$$

and hence $\left(\mathcal{YM}(A(t)) - \mathcal{YM}(\tilde{A}_\infty)\right)^{1-\gamma} \le C(t-T)^{\frac{1}{1-2\gamma}}$. This shows

$$||A(t) - \tilde{A}_\infty||_{W^{1,2}} \le C(t-T)^{\frac{1-\gamma}{1-2\gamma}}$$

for all $t > T$ and completes the proof of the convergence result. This argument also proves the following result:

Corollary 2.4.6. *Let $B \in \mathcal{A}(P)$ be a Yang-Mills connection and let $\epsilon > 0$. Then there exists $\delta > 0$ such that for every solution $A(t)$ of the Yang-Mills flow (2.22) with $||A(0) - B||_{W^{1,2}} < \delta$ we have either*

$$\sup_{t \ge 0} ||A(0) - A(t)||_{W^{1,2}} < \epsilon$$

for all $t \ge 0$ or there exists $T > 0$ with $\mathcal{YM}(A(T)) < \mathcal{YM}(B)$.

The Kempf-Ness flow

By Proposition 2.2.3 the infinitesimal action of the complexified Gauge action is given by

$$L_A(\xi + \mathbf{i}\eta) := \frac{d}{dt}\bigg|_{t=0} \exp(t\xi + \mathbf{i}t\eta)A = -d_A\xi - *d_A\eta$$

for $\xi, \eta \in W^{2,2}(\Sigma, \mathrm{ad}(P))$ and $A \in \mathcal{A}(P)$. With this formula we can express the gradient of the Yang-Mills functional as

$$\nabla\mathcal{YM}(A) = d_A^* F_A = - * d_A * F_A = L_A(\mathbf{i} * F_A).$$

This implies that any solution of the Yang-Mills flow (2.20) remains in a single complexified orbit.

Proposition 2.4.7. *Let $A_0 \in \mathcal{A}(P)$ and let $A(t)$ be the (weak) solution of the Yang-Mills flow (2.22) starting at A_0. Let $g : [0,\infty) \to \mathcal{G}^c(P)$ be the solution of the ODE*

$$g(t)^{-1}\dot{g}(t) = \boldsymbol{i}(*F_{A(t)}), \qquad g(0) = \mathbb{1}. \tag{2.24}$$

Then holds $g \in C^0_{\ell oc}([0,\infty), \mathcal{G}^c(P))$ and

$$A(t) = g(t)^{-1}A_0$$

for all $t \in [0,\infty)$. Moreover, g depends continuously on A_0.

Proof. Recall from Lemma 2.2.7 the formula

$$B(t) := g_t^{-1}(A_0) = A_0 + g_t^{-1}d_{A_0}g_t - g_t^{-1}(h_t^{-1}\partial_{A_0}h_t)g_t$$

with $h_t := (g_t^{-1})^* g_t^{-1}$. By Theorem 2.4.2 holds $F \in L^2_{\ell oc}([0,\infty), W^{1,2})$ and hence $g \in W^{1,2}_{\ell oc}([0,\infty), W^{1,2})$ and $B \in W^{1,2}_{\ell oc}([0,\infty), L^1)$. The same calculation as in the smooth case shows

$$\dot{B}(t) = L_{B(t)}(g_t^{-1}\dot{g}_t) = -d^*_{B(t)}F_{A(t)}.$$

Approximation of A_0 with smooth connections shows $A \in W^{1,2}_{\ell oc}([0,\infty), L^1)$ and

$$\dot{A}(t) = -d^*_{A(t)}F_{A(t)}.$$

Define $C(t) := A(t) - B(t)$ and $\Psi(t) := *F_{A(t)} \in L^2_{\ell oc}([0,\infty), L^1)$. The calculation above shows that C solves the linear ODE

$$\dot{C}(t) = *[C(t), \Psi(t)], \qquad C(0) = 0.$$

and hence $C = 0$. The Sobolev embedding $W^{1,2}([0,t_0], L^1) \hookrightarrow C^0([0,t_0], L^1)$ then yields $A(t) = B(t) = g_t^{-1}A_0$ for all t.

Since A maps continuously in $W^{1,2}$, it follows from the expression

$$A(t) = g_t^{-1}A_0 = A_0 + g_t^{-1}d_{A_0}g_t - g_t^{-1}(h_t^{-1}\partial_{A_0}h_t)g_t$$

that $A(t)^{0,1} = A_0^{0,1} + g_t^{-1}\bar{\partial}_{A_0}g_t$ and $g_t^{-1}\bar{\partial}_{A_0}g_t$ maps continuously into $W^{1,2}$. Let $\tilde{A}$ be a smooth reference metric and write $A_0 = \tilde{A} + a_0$. Then $g_t^{-1}\bar{\partial}_{\tilde{A}}g_t^{-1}$ maps continuously into L^p for any $p < \infty$ and by elliptic regularity, g maps continuously in $W^{1,p}$. Since $W^{1,p} \hookrightarrow C^0$, we can rerun the argument where $g_t^{-1}\bar{\partial}_{\tilde{A}} g_t$ now maps continuously in $W^{1,2}$ and conclude $g \in C^0_{\ell oc}([0,\infty), W^{2,2})$. Since A and F_A depend continuously on A_0, the solution g depends continuously on A_0 in $W^{1,2}_{\ell oc}([0,\infty), W^{1,2})$ and then by elliptic regularity also in $C^0_{\ell oc}([0,\infty), W^{2,2})$. □

Remark 2.4.8. Let $A_0 \in \mathcal{A}(P)$ and let $A(t)$ be as in Proposition 2.4.7. For $g_0 \in \mathcal{G}^c(P)$ consider the more general equation

$$g(t)^{-1}\dot{g}(t) = \mathbf{i} * F_{A(t)}, \qquad g(0) = g_0. \tag{2.25}$$

Then $\tilde{g}(t) = g_0^{-1} g(t)$ solves equation (2.24) with respect to $\tilde{A}_0 = g_0^{-1}(A_0)$. Hence (2.25) has a unique solution in $C^0_{\ell oc}([0,\infty), \mathcal{G}^c(P))$, which depends continuously on g_0 and A_0.

We shall consider the following variant of this equation.

Definition 2.4.9 (Kempf-Ness flow). *Let $A_0 \in \mathcal{A}(P)$ and $g_0 \in \mathcal{G}^c(P)$. We say that $g(t) \in C^0_{\ell oc}([0,\infty), \mathcal{G}^c(P))$ is a weak solution of the equation*

$$g^{-1}(t)\dot{g}(t) = \boldsymbol{i} * F_{g^{-1}(t)A_0}, \qquad g(0) = g_0 \tag{2.26}$$

if there exist a sequence of smooth initial data $(A_k, g_0^k) \in \mathcal{A}(P) \times \mathcal{G}^c(P)$ converging to (A_0, g_0) and smooth solutions $g_k(t)$ of the equation

$$g_k^{-1}(t)\dot{g}_k(t) = \boldsymbol{i} * F_{g_k^{-1}(t)A_k}, \qquad g_k(0) = g_0^k$$

such that $g_k(t)$ converges to $g(t)$ in $C^0_{\ell oc}([0,\infty), W^{2,2})$.

Remark 2.4.10. We call a solution $g \in C^0_{\ell oc}([0,\infty), \mathcal{G}^c(P))$ of (2.25) a solution of the Kempf-Ness flow starting at g_0 (with respect to A_0). We show in Section 6 that there exists a $\mathcal{G}(P)$-invariant functional

$$\Phi_{A_0} : \mathcal{G}^c(P) \to \mathbb{R}$$

whose negative gradient flow lines correspond to solution of (2.26).

Lemma 2.4.11. *For every initial data $(A_0, g_0) \in \mathcal{A}(P) \times \mathcal{G}^c(P)$ there exists a unique (weak) solution of (2.26) which depends continuously on the initial data.*

Proof. We use the notation introduced in Definition 2.4.9. Then $A_k(t) := g_k(t)^{-1} A_k$ satisfies

$$\partial_t A_k(t) = L_{A_k(t)}\left(g_k^{-1}(t)\dot{g}_k(t)\right) = -L_{A_k(t)}\left(\mathbf{i} * F_{A_k(t)}\right) = -d^*_{A_k(t)} F_{A_k(t)}$$

and thus $A_k(t)$ yields a smooth solution of the Yang-Mills flow. Conversely, the solution $A_k(t)$ is uniquely determined by the initial condition $(g_0^k)^{-1} A_k$ and we may recover $g_k(t)$ from this solution via Proposition 2.4.7 and Remark 2.4.8. Since solutions of the Yang-Mills flow and solutions of (2.25) depend continuously on the initial data, it follows that the weak solution $g(t)$ of (2.26) is uniquely determined by the weak solution $A(t)$ of the Yang-Mills flow starting at $g_0^{-1} A_0$. □

The next proposition shows that solutions of the Kempf-Ness flow (2.26) remain at bounded distance in the homogeneous space $\mathcal{G}^c/\mathcal{G}$.

Proposition 2.4.12. *Let $A_0 \in \mathcal{A}(P)$ and let $g, \tilde{g} \in C^0_{\ell oc}([0,\infty), \mathcal{G}^c(P))$ be (weak) solutions of (2.26) starting at $g_0, \tilde{g}_0 \in \mathcal{G}^c(P)$. Define $\eta(t) \in W^{2,2}(\Sigma, ad(P))$ and $u(t) \in \mathcal{G}(P)$ by the equation*

$$g(t)\exp(\boldsymbol{i}\eta(t))u(t) = \tilde{g}(t).$$

Then the following holds:

1. *$\rho(t) := ||\eta(t)||_{L^2}$ is non-increasing in t. More precisely, if $\eta(t) \neq 0$ then*

$$\dot{\rho}(t) = -\frac{1}{\rho(t)}\int_0^1 ||d_{A_{s,t}}\eta(t)||^2_{L^2}\, ds$$

   with $A_{s,t} := e^{-\boldsymbol{i}s\eta(t)}g_t^{-1}A_0$.

2. *The differential inequality*

$$(\partial_t + \Delta)||\eta||^2 \leq 0$$

   is satisfied. In particular, $||\eta(t)||_{L^\infty}$ is non-increasing by the maximum principle for the heat equation

3. *η is uniformly bounded in $W^{2,2}$.*

4. *u is uniformly bounded in $W^{2,2}$.*

Proof. We prove 1.) and 2.): By approximation, we can assume that A_0, g and $\tilde{g}$ are all smooth. Let $\pi : G^c \to G^c/G$ denote the projection and define $\gamma(s,t) := \pi(g(t)e^{\mathbf{i}s\eta(t)})$. Pointwise $\gamma(\cdot,t)$ is the unique geodesic of length $|\eta(t)|$ connecting $\pi(g)$ and $\pi(\tilde{g})$. The following calculation is pointwise valid:

$$\begin{aligned}
\partial_t||\eta||^2 &= \partial_t\int_0^1 \langle\partial_s\gamma, \partial_s\gamma\rangle\, ds = 2\int_0^1 \langle\nabla_t\partial_s\gamma, \partial_s\gamma\rangle\, ds \\
&= 2\int_0^1 \langle\nabla_s\partial_t\gamma, \partial_s\gamma\rangle\, ds = 2\int_0^1 \partial_s\langle\partial_t\gamma, \partial_s\gamma\rangle\, ds \\
&= 2\left(\langle\partial_t\gamma(1,t), \partial_s\gamma(1,t)\rangle - \langle\partial_t\gamma(0,t), \partial_s\gamma(0,t)\rangle\right) \\
&= 2\langle\tilde{g}^{-1}(t)\dot{\tilde{g}}(t) - g^{-1}(t)\dot{g}(t), \mathbf{i}\eta(t)\rangle \\
&= 2\langle *F_{\tilde{g}(t)^{-1}A_0} - *F_{g(t)^{-1}A_0}, \eta(t)\rangle
\end{aligned}$$

With $A_{s,t} := e^{-\mathbf{i}s\eta(t)}g^{-1}A_0$ this yields

$$\partial_t||\eta||^2 = 2\int_0^1 \langle\eta(t), *d_{A_{s,t}} * d_{A_{s,t}}\eta(t)\rangle = -\Delta||\eta||^2 - 2\int_0^1 ||d_{A_{s,t}}\eta(t)||^2\, ds.$$

This proves the second claim and the first one is obtained by integrating this inequality over Σ.

We prove 3.) and 4.): Recall that $\tilde{A}(t) := \tilde{g}_t^{-1}(A_0)$ and $A(t) := g_t^{-1}(A_0)$ are solutions of the Yang-Mills flow. Since they converge in $W^{1,2}$, they are both uniformly bounded in $W^{1,2}$. With $a(t) := e^{\mathbf{i}\eta_t}u_t$ we have $\tilde{A} = a^{-1}(A)$ and hence

$$\tilde{A}^{0,1} = A^{0,1} + a^{-1}\bar{\partial}_A a$$

This shows that $\bar{\partial}_A a$ is uniformly bounded in $W^{1,2}$ and hence a is uniformly bounded in $W^{2,2}$. From the formula $aa^* = e^{2\mathbf{i}\eta}$ we conclude that η is uniformly bounded in $W^{2,2}$ and then u is also uniformly bounded in $W^{2,2}$.

□

2.4.2 Uniqueness of Yang-Mills connections

We follow the arguments from ([51], Chapter 6) to prove the analog of the Ness uniqueness theorem and the moment limit theorem. These are originally due to Calabi-Chen [17] and Chen-Sun [23] in the context of extremal Kähler metrics.

Proposition 2.4.13. *Let $A_0, A_1 \in \mathcal{A}(P)$ be Yang-Mills connections with $\mathcal{G}^c(A_0) = \mathcal{G}^c(A_1)$. Then holds $\mathcal{G}(A_0) = \mathcal{G}(A_1)$.*

Proof. Choose $\tilde{g} \in \mathcal{G}^c(P)$ such that

$$A_1 = \tilde{g}^{-1}A_0.$$

holds. Since A_0 and A_1 are Yang-Mills connections, they generate constant flow lines $A_0(t) \equiv A_0$ and $A_1(t) \equiv A_1$. Let $g_0, g_1 \in C^0_{\ell oc}([0,\infty), \mathcal{G}^c)$ be the solutions of the equation

$$g_0^{-1}\dot{g}_0 = *F_{A_0}, \quad g_0(0) = \mathbb{1} \quad \text{and} \quad g_1^{-1}\dot{g}_1 = *F_{A_1}, \quad g_1(0) = \tilde{g}$$

from Proposition 2.4.7 and the following Remark. They satisfy

$$A_0 = g_0^{-1}(t)A_0 \quad \text{and} \quad A_1 = g_1^{-1}(t)A_0$$

and g_0 and g_1 are solutions of the Kempf-Ness flow (2.26) with respect to A_0. Define $\eta(t) \in W^{2,2}(\Sigma, \mathrm{ad}(P))$ and $u(t) \in \mathcal{G}$ by the equation

$$g_0(t)\exp(\mathbf{i}\eta(t))u(t) = g_1(t)$$

as in Proposition 2.4.12. Then there exist $\eta_\infty \in W^{2,2}(\Sigma, \mathrm{ad}(P))$, $u_\infty \in \mathcal{G}(P)$ and a sequence $t_i \to \infty$ such that

$$\lim_{i\to\infty} \dot{\rho}(t_i) = 0, \qquad \eta(t_i) \overset{H^2}{\rightharpoonup} \eta_\infty, \qquad u(t_i) \overset{H^2}{\rightharpoonup} u_\infty.$$

By taking a further subsequence if necessary, we may assume that

$$\lim_{i\to\infty} ||d_{A_{s,t_i}}\eta(t_i)||_{L^2} = 0$$

holds for almost every $s \in [0,1]$, where we defined

$$A_{s,t} = e^{-\mathbf{i}s\eta(t)}(g_0^{-1}(t)A_0) = e^{-\mathbf{i}s\eta(t)}A_0.$$

Moreover, by Rellich's theorem, $\eta(t_i)$ and $u(t_i)$ converge for every $p < \infty$ strongly in $W^{1,p}$ to η_∞ and u_∞. By continuity of the Gauge action $\mathcal{G}^{1,p} \times \mathcal{A}^p \to \mathcal{A}^p$ for $p > 2$, we conclude

$$A_{s,t_i} \overset{L^p}{\to} A_{s,\infty} := e^{-\mathbf{i}s\eta_\infty} A_0, \quad \text{and} \quad d_{A_{s,t_i}}\eta(t_i) \overset{L^p}{\to} d_{A_{s,\infty}}\eta_\infty.$$

This implies that for almost every $s \in [0,1]$, we must have $d_{A_{s,\infty}}\eta_\infty = 0$. For $s \to 0$ we conclude $d_{A_{0,\infty}}\eta_\infty = d_{A_0}\eta_\infty = 0$ and hence $e^{-\mathbf{i}\eta_\infty} A_0 = A_0$. It follows now

$$A_1 = g_1(t_i)^{-1} A_0 = u(t_i)^{-1} e^{-\mathbf{i}\eta(t_i)} A_0 \overset{L^p}{\longrightarrow} u_\infty^{-1} e^{-\mathbf{i}\eta_\infty} A_0 = u_\infty^{-1} A_0.$$

This shows $A_1 = u_\infty^{-1} A_0$ and thus A_0 and A_1 lie in the same $\mathcal{G}$-orbit. □

Theorem 2.4.14 (Moment Limit Theorem). *Let $A_0 \in \mathcal{A}(P)$ and $A : [0,\infty) \to \mathcal{A}(P)$ be the solution of the Yang-Mills flow starting at A_0. The limit $A_\infty := \lim_{t\to\infty} A(t)$ satisfies*

$$\mathcal{YM}(A_\infty) = \inf_{g \in \mathcal{G}^c(P)} \mathcal{YM}(gA_0).$$

Moreover, the $\mathcal{G}(P)$-orbit of A_∞ depends only on the complexified orbit $\mathcal{G}^c(A_0)$.

Proof. Let $g_0 \in \mathcal{G}^c(P)$ be given and define $g, \tilde{g} \in C^0_{\ell oc}([0,\infty), \mathcal{G}^c)$ by

$$g^{-1}\dot{g} = *F_A, \quad g(0) = \mathbb{1} \quad \text{and} \quad \tilde{g}^{-1}\dot{\tilde{g}} = *F_A, \quad g(0) = g_0$$

as in Proposition 2.4.7 and the following Remark. Let $A(t)$ and $\tilde{A}(t)$ be the solutions of the Yang-Mills flow starting at A_0 and $\tilde{A}_0 := g_0^{-1} A_0$. Then

$$A_0(t) = g_t^{-1}(A_0), \qquad \tilde{A}(t) = \tilde{g}_t^{-1}(A_0)$$

and $g, \tilde{g}$ are solutions of the Kempf-Ness flow (2.26) with respect to A_0. Define $\eta(t) \in W^{2,2}(\Sigma, \mathrm{ad}(P))$ and $u(t) \in \mathcal{G}(P)$ by the equation

$$g_0(t) \exp(\mathbf{i}\eta(t)) u(t) = g_1(t)$$

as in Proposition 2.4.12. It follows that there exist $\eta_\infty \in W^{2,2}(\Sigma, \mathrm{ad}(P))$ and $u_\infty \in \mathcal{G}(P)$ and a sequence $t_i \to \infty$ such that

$$\eta(t_i) \overset{W^{2,2}}{\rightharpoonup} \eta_\infty, \qquad u(t_i) \overset{W^{2,2}}{\rightharpoonup} u_\infty.$$

By Rellich's theorem we obtain strong convergence in $W^{1,p}$ and using the Sobolev embedding $W^{1,2} \hookrightarrow L^p$ for every $p < \infty$ we obtain:

$$\tilde{A}_\infty \overset{W^{1,2}}{\longleftarrow} \tilde{A}(t_i) = u(t_i)^{-1} e^{\mathbf{i}\eta(t_i)} A(t_i) \overset{L^p}{\longrightarrow} u_\infty^{-1} \eta_\infty^{-1} A_\infty.$$

Hence $\tilde{A}_\infty = u_\infty^{-1}\eta_\infty^{-1} A_\infty$. Thus $\tilde{A}_\infty$ and A_∞ are Yang-Mills connections lying in a common complexified orbit and Proposition 2.4.13 shows that in fact $\mathcal{G}(A_\infty) = \mathcal{G}(\tilde{A}_\infty)$. This shows $\mathcal{YM}(A_\infty) = \mathcal{YM}(\tilde{A}_\infty) \leq \mathcal{YM}(g_0^{-1} A_0)$ and completes the proof. □

The following theorem is the analog of the Ness uniqueness theorem in finite dimensional GIT.

Theorem 2.4.15 (Uniqueness of Yang-Mills connections). *Let $A_0 \in \mathcal{A}(P)$ and $A', A'' \in \overline{\mathcal{G}^c(A_0)}$ be in the $W^{1,2}$-closure of a single complexified orbit satisfying*

$$\mathcal{YM}(A') = \mathcal{YM}(A'') = \inf_{g \in \mathcal{G}^c} \mathcal{YM}(gA_0).$$

Then follows $\mathcal{G}(A') = \mathcal{G}(A'')$.

Corollary 2.4.16. *Let $P \to \Sigma$ be a principal G bundle of constant central type $\tau \in Z(\mathfrak{g})$ defined by* (2.12). *Suppose $A \in \mathcal{A}(P)$ is μ_τ-semistable. Then the $W^{1,2}$-closure $\overline{\mathcal{G}^c(A)}$ contains a unique μ_τ-polystable orbit.*

Proof. It follows from (2.19) that solutions of the equation $*F_A = \tau$ correspond to global minima of the Yang-Mills functional on $\mathcal{A}(P)$. □

Proof of Theorem 2.4.15. Let $A(t)$ be the solution of the Yang-Mills flow starting at A_0 and let $A_\infty := \lim_{t\to\infty} A(t)$. Then Theorem 2.4.14 implies

$$\mathcal{YM}(A_\infty) = \inf_{g \in \mathcal{G}^c} \mathcal{YM}(gA_0) =: m.$$

Since $A_\infty \in \overline{\mathcal{G}^c(A_0)}$, it suffices to show that any connection $B \in \overline{\mathcal{G}^c(A_0)}$ with $\mathcal{YM}(B) = m$ is contained in $\mathcal{G}(A_\infty)$. For this let $A_i \in \mathcal{G}^c(A_0)$ be a sequence which converges to B. Denote by $A_i(t)$ the corresponding solutions of the Yang-Mills flow and set $B_i := \lim_{t\to\infty} A_i(t)$. Note that B is necessarily a Yang-Mills connection, since

$$\mathcal{YM}(B(t)) = \lim_{i\to\infty} \mathcal{YM}(A_i(t)) \geq m = \mathcal{YM}(B(0))$$

where $B(t)$ denotes the solution of the Yang-Mills flow starting at B. Thus, we may apply Corollary 2.4.6 with respect to B and conclude that $||A_i - B_i||_{W^{1,2}}$ converges to zero and hence

$$\lim_{i\to\infty} B_i = B.$$

By Theorem 2.4.14 holds $\mathcal{G}(B_i) = \mathcal{G}(A_\infty)$ and hence there exists $u_i \in \mathcal{G}(P)$ such that $u_i^{-1}(A_\infty) = B_i$. Since the connections B_i are uniformly bounded in $W^{1,2}$, the gauge transformations u_i are uniformly bounded in $W^{2,2}$. Thus there exists $u_\infty \in \mathcal{G}(P)$ such that after passing to a subsequence u_i converges weakly in $W^{2,2}$ to u_∞ and strongly in $W^{1,p}$ for any $p < \infty$. Using the continuity of the Gauge action $\mathcal{G}^{1,p} \times \mathcal{A}^p \to \mathcal{A}^p$ we conclude

$$B_i = u_i^{-1}(A_\infty) \xrightarrow{L^p} u_\infty^{-1} A_\infty$$

and in particular $B = u_\infty^{-1} A_\infty \in \mathcal{G}(A_\infty)$

□

2.4.3 Yang-Mills characterization of μ_τ-stability

We characterize the μ_τ-stability of a connection $A \in \mathcal{A}(P)$ in terms of the the limit A_∞ of the Yang-Mills flow starting at A. This is Theorem 2.4.18 below. The proof relies on the following proposition.

Proposition 2.4.17. *Let $P \to \Sigma$ be a principal G bundle of central type $\tau \in Z(\mathfrak{g})$ defined by (2.12). The subsets of μ_τ-semistable connections*

$$\mathcal{A}^{ss}(P) := \{A \in \mathcal{A}(P) \,|\, A \text{ is } \mu_\tau\text{-semistable}\}$$

and μ_τ-stable connections

$$\mathcal{A}^{s}(P) := \{A \in \mathcal{A}(P) \,|\, A \text{ is } \mu_\tau\text{-stable}\}$$

are open subsets of $\mathcal{A}(P)$ with respect to the $W^{1,2}$-topology.

Proof. It follows from (2.19) that

$$\inf_{A\in\mathcal{A}(P)} \mathcal{YM}(A) \geq \frac{1}{2}||\tau||^2 =: m.$$

Moreover

$$\mathcal{A}^{ss}(P) := \left\{ A \in \mathcal{A}(P) \,\middle|\, \inf_{g\in\mathcal{G}^c(P)} \mathcal{YM}(gA) = m \right\} \tag{2.27}$$

and $\mathcal{YM}(A) = m$ is equivalent to $*F_A = \tau$.

Step 1: *$\mathcal{A}^{ss}(P)$ is open.*

Let $A_0 \in \mathcal{A}^{ss}(P)$ be given. Let $A(t)$ be the solution of the Yang-Mills flow starting at A_0 and $A_\infty := \lim_{t\to\infty} A(t)$. It follows from Theorem 2.4.14 and (2.27) that A_∞ is a Yang-Mills connection satisfying $\mathcal{YM}(A_\infty) = m$. By the Łojasiewicz inequality (Proposition 2.4.4) there exists $\epsilon > 0$, $c > 0$ and $\gamma \in [\frac{1}{2}, 1)$ such that for all $B \in \mathcal{A}(P)$ with $||B - A_\infty||_{W^{1,2}} < \epsilon$ the inequality

$$||d_B^* F_B||_{L^2} \geq c|\mathcal{YM}(B) - m|^\gamma \tag{2.28}$$

is satisfied. By Corollary 2.4.6 there exists $\delta > 0$ such that for every $B \in \mathcal{A}(P)$ with $||B - A_\infty||_{W^{1,2}} < \delta$ we have $||B_\infty - A_\infty||_{W^{1,2}} < \epsilon$. In particular, (2.28) applies to B_∞ and yields $\mathcal{YM}(B_\infty) = m$. This shows

$$U := \{B \in \mathcal{A}(P) \,|\, ||B - A_\infty||_{W^{1,2}} < \delta\} \subset \mathcal{A}^{ss}(P).$$

Now choose $T > 0$ such that $A(T) \in U$ and choose $g \in \mathcal{G}^c(P)$ with $A(T) = g^{-1}A_0$. By continuity of the gauge action there exists an open neighborhood V of A_0 with $g^{-1}V \subset U$ and hence $V \subset \mathcal{A}^{ss}(P)$.

Step 2: *Denote by $\mathcal{A}^*(P) \subset \mathcal{A}(P)$ the space of irreducible connections. This is an open subset and*

$$\mathcal{Z} := \{A \in \mathcal{A}^* \,|\, \mathcal{YM}(A) = m\}/\mathcal{G}$$

is a finite dimensional smooth submanifold of $\mathcal{A}^*/\mathcal{G}$.

We may assume that $Z(G)$ is discrete, $\tau = 0$ and $m = 0$, since otherwise $\mathcal{A}^*(P) = \emptyset$. Let $A_0 \in \mathcal{A}^*(P)$ be a smooth irreducible connection. The Laplacian $d_{A_0}^* d_{A_0}$ is then injective and by elliptic regularity there exists $c_0 > 0$ such that

$$||d_{A_0}^* d_{A_0}\xi||_{L^2} \geq c_0||\xi||_{W^{2,2}}$$

for all $\xi \in W^{2,2}(\Sigma, \mathrm{ad}(P))$. For $a \in W^{1,2}(\Sigma, T^*\Sigma \otimes \mathrm{ad}(P))$ expand

$$d^*_{A_0+a} d_{A_0+a}\xi = d^*_{A_0} d_{A_0}\xi + d^*_{A_0}[a,\xi] - *[a, *d_{A_0}\xi] - *[a, *[a,\xi]].$$

Since $\dim_{\mathbb{R}}(\Sigma) = 2$, we have the Sobolev estimate $||fg||_{L^2} \leq c||f||_{W^{1,2}}||g||_{W^{1,2}}$ and $||fg||_{W^{1,2}} \leq ||f||_{W^{1,2}}||g||_{W^{2,2}}$. This yields

$$||d^*_{A_0+a} d_{A_0+a}\xi||_{L^2} \geq c_0||\xi||_{W^{2,2}} - c||a||_{W^{1,2}}||\xi||_{W^{2,2}}$$

and $A_0 + a$ is irreducible if $||a||_{W^{1,2}}$ is sufficiently small. Hence $\mathcal{A}^*(P)$ is open.

Now fix an irreducible connection A_0 with $*F_{A_0} = 0$. We may assume without loss of generality that A_0 is smooth and work in a Coulomb gauge relative to A_0. This allows us to identify a neighborhood of $[A_0]$ in $\mathcal{A}^*(P)/\mathcal{G}(P)$ with $a \in W^{1,2}(\Sigma, T^*\Sigma \otimes \mathrm{ad}(P))$ satisfying $||a||_{W^{1,2}} < \epsilon$ and $d^*_{A_0}a = 0$ under the map $a \mapsto [A_0 + a]$. Consider

$$\phi : \{a \in W^{1,2}(\Sigma, T^*\Sigma \otimes \mathrm{ad}(P)) \,|\, d^*_{A_0}a = 0, ||a||_{W^{1,2}} < \epsilon\} \to L^2(\Sigma, \mathrm{ad}(P))$$

$$\phi(a) := *F_{A_0+a}$$

and define $Z_{A_0} := \phi^{-1}(0)$. We claim that 0 is a regular value for ϕ (after possibly shrinking ϵ). Once this is established, the claim follows from the implicit function theorem. The derivative of ϕ at a point a is given by

$$d\phi(a) : \{\hat{a} \in W^{1,2}(\Sigma, T^*\Sigma \otimes \mathrm{ad}(P)) \,|\, d^*_{A_0}\hat{a} = 0\} \to L^2(\Sigma, \mathrm{ad}(P))$$

$$d\phi(a)\hat{a} = *d_{A_0}\hat{a} + *[a \wedge \hat{a}].$$

Since $d\phi(a)$ is the restriction of a compact perturbation of the Fredholm operator $*(d_{A_0} \oplus d^*_{A_0})$, its kernel is finite dimensional. We denote by

$$K := \{\hat{a} \in W^{1,2}(\Sigma, T^*\Sigma \otimes \mathrm{ad}(P)) \,|\, d_{A_0}\hat{a} = 0, d^*_{A_0}\hat{a} = 0\}$$

the space of A_0-harmonic 1-forms with values in $\mathrm{ad}(P)$ and define V by the L^2-orthogonal decomposition

$$W^{1,2}(\Sigma, T^*\Sigma \otimes \mathrm{ad}(P)) = V \oplus K.$$

Then the restriction of the Fredholm-operator $d_{A_0} \oplus d^*_{A_0}$ to V defines an isomorphism

$$d_{A_0} \oplus d^*_{A_0} : V \to L^2(\Sigma, \mathrm{ad}(P)) \oplus L^2(\Sigma, \Lambda^2 T^*\Sigma \otimes \mathrm{ad}(P))$$

It is injective by definition of V and to prove surjectivity let $f \in L^2(\Sigma, \mathrm{ad}(P))$ and $\omega \in L^2(\Sigma, \Lambda^2 T^*\Sigma \otimes \mathrm{ad}(P))$ be given. Then by Hodge theory we can solve the equation

$$\Delta_{A_0}\hat{a} = d^*_{A_0}\omega + d_{A_0}f.$$

From this follows

$$d^*_{A_0}(d_{A_0}\hat{a} - \omega) = d_{A_0}(f - d^*_{A_0}\hat{a}).$$

Since $*F_{A_0} = 0$, both sides of the equation are orthogonal and hence must vanish. Since A_0 is irreducible, it follows $d_{A_0}\hat{a} = \omega$ and $d^*_{A_0}\hat{a} = f$. In particular, for any $s \in L^2(\Sigma, \mathrm{ad}(P))$ exists a solution $\hat{a} \in V$ of the equations

$$d_{A_0}\hat{a} + [a \wedge \hat{a}] = *s, \qquad d^*_{A_0}\hat{a} = 0 \tag{2.29}$$

for $a = 0$. Since the equation is linear in a, another application of the inverse function theorem shows that after possibly shrinking ϵ the equation (2.29) has a solution $\hat{a}(a) \in V$ for all a with $||a||_{W^{1,2}} < \epsilon$.

Step 3: $\mathcal{A}^s$ *is open.*

We may assume that $Z(G)$ is discrete, $\tau = 0$ and $m = 0$, since otherwise $\mathcal{A}^s(P) = \emptyset$. Let $A \in \mathcal{A}^s(P)$ be given. By definition there exists $g \in \mathcal{G}^c(P)$ such that $A_0 = g^{-1}A$ is smooth and satisfies $\mathcal{YM}(A_0) = 0$. Let Z_{A_0} be as in Step 2 and consider the map

$$\psi : Z_{A_0} \times W^{2,2}(\Sigma, \mathrm{ad}(P)) \times W^{2,2}(\Sigma, \mathrm{ad}(P)) \to \mathcal{A}$$

$$\psi(A, \xi, \eta) := e^{\mathbf{i}\eta} e^{\xi} A.$$

We have seen that Z_{A_0} is a smooth manifold with tangent space

$$T_{A_0} Z_{A_0} = \{\hat{a} \in W^{1,2}(\Sigma, \mathrm{ad}(P)) \,|\, d_{A_0}^* \hat{a} = 0,\, d_{A_0}\hat{a} = 0\}.$$

The differential of ψ at the point $(A_0, 0, 0)$ is given by

$$d\psi(A_0, 0, 0)[\hat{a}, \hat{\xi}, \hat{\eta}] := \hat{a} - d_{A_0}\hat{\xi} - *d_{A_0}\hat{\eta}.$$

Since $F_{A_0} = 0$, it follows as in Step 2 from Hodge theory that $d\psi(A_0, 0, 0)$ is an isomorphism. The implicit function theorem yields thus an open neighborhood U of A_0 with

$$A_0 \in U \subset \mathrm{Im}(\psi) \subset \mathcal{A}^s.$$

Finally, by continuity of the gauge action, there exists an open neighborhood V of A with $g^{-1}V \subset U$ and hence $\mathcal{A}^s(P)$ is open.

□

Theorem 2.4.18. *Let $P \to \Sigma$ be a principal G bundle of central type $\tau \in Z(\mathfrak{g})$ defined by (2.12) and denote $m := \frac{1}{2}||\tau||^2$. Let $A_0 \in \mathcal{A}(P)$ and denote by A_∞ the limit of the the Yang-Mills flow $A(t)$ starting at A_0.*

1. *A_0 is μ_τ-stable if and only if A_∞ is irreducible.*
2. *A_0 is μ_τ-polystable if and only if $\mathcal{YM}(A_\infty) = m$ and $A_\infty \in \mathcal{G}^c(A_0)$.*
3. *A_0 is μ_τ-semistable if and only if $\mathcal{YM}(A_\infty) = m$.*
4. *A_0 is μ_τ-unstable if and only if $\mathcal{YM}(A_\infty) > m$.*

Proof. It follows from (2.19) that m is a lower bound for the Yang-Mills functional on $\mathcal{A}(P)$ and $A \in \mathcal{A}(P)$ satisfies $\mathcal{YM}(A) = m$ if and only if $*F_A = \tau$. Thus the characterization for μ_τ-unstable and μ-semistable orbits follows from Theorem 2.4.14.

Suppose next that A_0 is μ_τ-polystable. Then exists $g_0 \in \mathcal{G}^c(P)$ such that $\tilde{A}_0 := g_0^{-1}(A_0)$ satisfies $\mathcal{YM}(\tilde{A}_0) = m$. The Yang-Mills flow line $\tilde{A}(t)$ starting at $\tilde{A}_0$ is constant and it follows from Theorem 2.4.14 and Theorem 2.4.15 that $A_\infty \in \mathcal{G}(\tilde{A}_0) \subset \mathcal{G}^c(A_0)$. The converse is immediate and this proves the criterion for μ_τ-polystable orbits.

Suppose now that A_0 is μ_τ-stable. Then the orbit $\mathcal{G}^c(A_0)$ has only discrete $\mathcal{G}^c(P)$ isotropy. Since A_0 is in particular μ_τ-polystable, we have $A_\infty \in \mathcal{G}^c(A_0)$. Hence the infinitesimal action $L_{A_\infty} : \xi \mapsto -d_{A_\infty}\xi$ is injective and A_∞ is irreducible. Suppose conversely that A_∞ is irreducible. Since A_∞ is a Yang-Mills connection, it satisfies $d_{A_\infty} * F_{A_\infty} = 0$ and hence $F_{A_\infty} = 0$. This shows that $\mathcal{G}^c(A_\infty)$ is stable. By Proposition 2.4.17, the subset $\mathcal{A}^s(P)$ of μ_τ-stable connections is open and hence $A(t) \in \mathcal{A}^s(P)$ for all sufficiently large t. Since the notion of μ_τ-stability is $\mathcal{G}^c(P)$-invariant, and since $A(t) \in \mathcal{G}^c(A_0)$, we conclude that A_0 is μ_τ-stable. □

2.5 Maximal weights

Let G be a compact connected Lie group, let $P \to \Sigma$ be a principal G bundle and let $\tau \in Z(\mathfrak{g})$ denote the central type of P defined by (2.12). It follows from Lemma 2.2.2 that $\mu_\tau(A) = *F_A - \tau$ defines a moment map for the action of $\mathcal{G}(P)$ on $\mathcal{A}(P)$. The **weights** associated to the gauge action with respect to this moment map are defined by

$$w_\tau(A,\xi) := \lim_{t\to\infty} \langle *F_{e^{\mathrm{i}t\xi}A} - \tau, \xi\rangle \tag{2.30}$$

for every $\xi \in W^{2,2}(\Sigma, \mathrm{ad}(P))$ and $A \in \mathcal{A}(P)$. Differentiating the right hand side in time yields

$$\frac{d}{dt}\langle *F_{e^{\mathrm{i}t\xi}A} - \tau, \xi\rangle = \langle - * d_{e^{\mathrm{i}t\xi}A} * d_{e^{\mathrm{i}t\xi}A}\xi, \xi\rangle = ||d_{e^{\mathrm{i}t\xi}A}\xi||^2_{L^2} \geq 0 \tag{2.31}$$

and therefore $w_\tau(A,\xi) \in \mathbb{R} \cup \{+\infty\}$ is well-defined.

Remark 2.5.1. The weights can be defined when ξ is only of Sobolev class $W^{1,2}$. The calculation above shows

$$w_\tau(A,\xi) = \langle *F_A - \tau, \xi\rangle + \int_0^\infty ||d_{e^{\mathrm{i}t\xi}A}\xi||^2_{L^2}\, dt \tag{2.32}$$

and the right hand side is well-defined for $\xi \in W^{1,2}(\Sigma, \mathrm{ad}(P))$.

We show in Proposition 2.5.2 and Lemma 2.5.7 that there exists a one to one correspondence between finite weights $w_\tau(A,\xi) < \infty$ and

$$\left\{ (P_Q, \xi_0) \,\middle|\, \begin{array}{c} \xi_0 \in \mathfrak{g},\ Q = Q(\xi_0) \\ P_Q \text{ is a principal } Q \text{ bundle} \\ P_Q \subset (P^c, J_A) \text{ is a holomorphic reduction} \end{array} \right\}.$$

For the definition of the parabolic subgroup $Q(\xi_0) \subset G^c$ see Definition 2.2.10. Using a deep regularity result of Uhlenbeck and Yau [118], we note that for every finite weight the section $\xi \in \Omega^0(\Sigma, \mathrm{ad}(P))$ is smooth provided A is a smooth connection.

Using this geometric description, we show in Proposition 2.5.9 that the algebraic stability of (P^c, J_A) is equivalent to the conditions on the weights $w_\tau(A,\xi)$ required in the Hilbert-Mumford criterion. In the last subsection we prove the moment weight inequality

$$-\frac{w_\tau(A,\xi)}{||\xi||_{L^2}} \leq \inf_{g\in\mathcal{G}^c(P)} ||\mu_\tau(gA)||_{L^2}.$$

This shows that A is μ_τ-unstable whenever there exists a negative weight. By Proposition 2.5.9 the later is true if and only if (P^c, J_A) is unstable.

2.5.1 Finite weights

It is more convenient to describe the weights in the language of vector bundles: We fix a faithfull representation $G \hookrightarrow U(n)$, identify G with a subgroup of $U(n)$ and denote by $E := P \times_G \mathbb{C}^n$ the associated vector bundle with structure group G. Consider the bundles

$$G(E),\ \mathfrak{g}(E),\ G^c(E),\ \mathfrak{g}^c(E)\ \subset\ \mathrm{End}(E)$$

which consist of endomorphisms that in any trivialization are contained in G, $\mathfrak{g}$, G^c and $\mathfrak{g}^c$ respectively. There are canonical identifications

$$\mathcal{G}(P) \cong \mathcal{G}(E) = \Omega^0(\Sigma, G(E)), \qquad \mathrm{ad}(P) \cong \mathfrak{g}(E) \subset \mathrm{End}(E)$$

and

$$\mathcal{G}(P^c) \cong \mathcal{G}^c(E) = \Omega^0(\Sigma, G^c(E)), \qquad \mathrm{ad}(P^c) \cong \mathfrak{g}^c(E) \subset \mathrm{End}(E).$$

We denote by $\mathcal{A}_G(E)$ the space of G-connections on E which is canonically isomorphic to $\mathcal{A}(P)$. Assume for convenience that the invariant inner product on $\mathfrak{g}$ is obtained by restriction of the standard inner product

$$\langle \xi, \eta \rangle := \mathrm{tr}(\xi\eta^*)$$

on $\mathfrak{u}(n)$.

Proposition 2.5.2. *Consider the setting described above. Let $A \in \mathcal{A}_G(E)$ be a smooth connection and let $\xi \in W^{1,2}(\Sigma, \mathfrak{g}(E))\backslash\{0\}$. If $w_\tau(A,\xi) < \infty$, then the following holds:*

1. *The endomorphism $\boldsymbol{i}\xi$ has constant eigenvalues $\lambda_1 < \cdots < \lambda_r$. The corresponding eigenspaces are unitary subbundles D_j and decompose E as orthogonal direct sum $E = D_1 \oplus \cdots \oplus D_r$.*

2. *Each partial sum $E_j := D_1 \oplus \cdots \oplus D_j$ is a holomorphic subbundle of E. This yields a holomorphic filtration*
$$0 < E_1 < E_2 < \cdots < E_r = E.$$

3. *The weight of ξ is given by the formula*
$$w_\tau(A,\xi) = 2\pi \sum_{j=1}^{r} \lambda_j c_1(D_j) - \langle \tau, \xi \rangle$$

This is Lemma 4.2 in [89]. Before giving the proof, we need to discuss the regularity of weakly holomorphic subbundles.

Definition 2.5.3. *Let E be a holomorphic hermitian vector bundle. A weakly holomorphic subbundle of E is a section $\pi \in W^{1,2}(\Sigma, End(E))$ satisfying $\pi = \pi^2 = \pi^*$ and $(\mathbb{1} - \pi)\bar\partial(\pi) = 0$.*

The following theorem is a special case of a more general result of Uhlenbeck and Yau [118]. They prove that weakly holomorphic subbundles of holomorphic hermitian vector bundles over arbitrary Kähler manifolds correspond to torsion-free coherent subsheaves. Since any torsion-free coherent sheaf over a Riemann surface is locally free, this reduces to the following:

Theorem 2.5.4 (Uhlenbeck and Yau [118])**.** *If $\pi \in W^{1,2}(\Sigma, End(E))$ is a weakly holomorphic subbundle, then π is the projection on a smooth holomorphic subbundle $E' \subset E$.*

Proof of Proposition 2.5.2. Let $0 \neq \xi \in W^{1,2}(\Sigma, \mathfrak{g}(E))$ be given and assume $w_\tau(A,\xi) < \infty$. Since $\mathfrak{g}^c = \mathfrak{g} \oplus \mathbf{i}\mathfrak{g}$ is per definitionem an orthogonal decomposition we have

$$\begin{aligned} ||d_{e^{\mathbf{i}t\xi}A}\xi||^2 &= \frac{1}{2}||\bar\partial_{e^{\mathbf{i}t\xi}A}\xi||^2 = \frac{1}{2}||\mathrm{Ad}\left(e^{\mathbf{i}t\xi}\right) \circ \bar\partial_A \circ \mathrm{Ad}\left(e^{-\mathbf{i}t\xi}\right)(\xi)||^2 \\ &= \frac{1}{2}||e^{\mathbf{i}t\xi}\bar\partial_A(\xi)e^{-\mathbf{i}t\xi}||^2. \end{aligned}$$

and from (2.32) follows

$$w_\tau(A,\xi) = \int_\Sigma \langle *F_A - \tau, \xi\rangle \, dvol_\Sigma + \frac{1}{2}\int_0^\infty ||e^{\mathbf{i}t\xi}\bar\partial_A(\xi)e^{-\mathbf{i}t\xi}||^2 \, dt. \tag{2.33}$$

Denote $A_t := e^{\mathbf{i}t\xi}(A)$ and let $k \geq 1$ be an integer. Then follows

$$\bar\partial \mathrm{tr}(\xi^k) = \mathrm{tr}(\bar\partial_{A_t}(\xi^k)) = k\mathrm{tr}(\xi^{k-1}\bar\partial_{A_t}(\xi))$$

and the Cauchy-Schwarz inequality $|\mathrm{tr}(AB)| \leq ||A|| \cdot ||B||$ yields

$$\begin{aligned} \int_\Sigma ||\bar\partial\, \mathrm{tr}(\xi^k)||\, dvol_\Sigma &\leq k\int_\Sigma ||\xi^{k-1}|| \cdot ||\bar\partial_{A_t}\xi||\, dvol_\Sigma \\ &= k||\xi^{k-1}||_{L^2} \cdot ||e^{\mathbf{i}\xi t}\bar\partial_A(\xi)e^{-\mathbf{i}\xi t}||_{L^2}. \end{aligned}$$

Since $w_\tau(A,\xi)$ is finite, it follows from (2.32) that there exists a sequence $t_j \to \infty$ such that

$$\lim_{j\to\infty} ||e^{\mathbf{i}\xi t_j}\bar\partial_A(\xi)e^{-\mathbf{i}\xi t_j}||_{L^2} = 0. \tag{2.34}$$

Hence $\bar\partial\mathrm{tr}(\xi^k) = 0$ and it follows from the maximum principle that $\mathrm{tr}(\xi^k)$ is constant. Denote the eigenvalues of $\mathbf{i}\xi$ with repetition according to their multiplicity by $\lambda_1' \leq \cdots \leq \lambda_n'$. Then

$$\mathrm{tr}(\xi^k) = (\lambda_1')^k + \ldots + (\lambda_n')^k$$

is constant for every $k \geq 1$. This is only possible if all the functions λ_j' are constant and hence $\mathbf{i}\xi$ has constant eigenvalues.

Let $\lambda_1 < \cdots < \lambda_r$ be the distinct eigenvalues of $\mathbf{i}\xi$. Since $\mathbf{i}\xi$ is a normal (hermitian) operator, the eigenspaces are pairwise orthogonal. Moreover, if Γ_j is a small loop around the eigenvalue λ_j in the complex plane, the orthogonal projection $\pi_j' : E \to D_j$ onto the eigenspace of λ_j is given by

$$\pi_j' := \frac{1}{2\pi\mathbf{i}}\int_{\Gamma_j} (z\mathbb{1} - \mathbf{i}\xi)^{-1}\, dz.$$

These projections have regularity $\pi'_j \in W^{1,2}(\Sigma, \mathrm{End}(E))$ and satisfy $\pi'_j = (\pi'_j)^2 = (\pi'_j)^*$.

We show next that the projections $\pi_j := \pi'_1 + \cdots + \pi'_j : E \to E_j$ define weakly holomorphic subbundles. By construction

$$\mathbf{i}\xi = m_1\pi_1 + \cdots + m_r\pi_r \tag{2.35}$$

for some $m_1, \ldots, m_r \in \mathbb{R}$. Write $\bar\partial_A(\xi) = [\hat\xi_{ij}]$ with respect to the splitting $E = D_1 \oplus \cdots \oplus D_r$. Then holds

$$\left[e^{\mathbf{i}t\xi}\hat\xi e^{-\mathbf{i}t\xi}\right]_{ij} = e^{(\lambda_j-\lambda_i)t}\hat\xi_{ij}$$

and (2.34) implies $\hat\xi_{ij} = 0$ for $i > j$. Thus $\bar\partial_A(\mathbf{i}\xi)$ is upper triangular and (2.35) yields

$$0 = (\mathbb{1} - \pi_j)(\bar\partial_A\xi)\pi_j = \sum_{k=1}^r m_k(\mathbb{1} - \pi_j)\bar\partial_A(\pi_k)\pi_j. \tag{2.36}$$

The Leibniz rule provides the formula

$$(\mathbb{1} - \pi_j)\bar\partial_A(\pi_k)\pi_j = \begin{cases} (\mathbb{1} - \pi_k)(\mathbb{1} - \pi_j)\bar\partial_A(\pi_j) & \text{for } k > j \\ (\mathbb{1} - \pi_j)\bar\partial_A(\pi_j) & \text{for } k = j \\ (\mathbb{1} - \pi_j)(\bar\partial_A(\pi_k) - \pi_k\bar\partial_A(\pi_j)) & \text{for } k < j \end{cases}.$$

This implies together with (2.36) the formula $(\mathbb{1} - \pi_j)\bar\partial_A(\pi_j) = 0$ by induction on j. Hence π_j defines a weakly holomorphic subbundle and E_j is smooth by Theorem 2.5.4. This proves the first two parts of the theorem.

Write $\bar\partial_A$ with respect to the splitting $E = D_1 \oplus \cdots \oplus D_r$ as

$$\bar\partial_A = \begin{pmatrix} \bar\partial_{A_1} & A_{12} & \ldots & A_{1r} \\ 0 & \bar\partial_{A_2} & \ldots & A_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \bar\partial_{A_r} \end{pmatrix}$$

where $A_{ij} \in \Omega^{0,1}(D_i \otimes D_j^*)$ and $\bar\partial_{A_j}$ is the Cauchy-Riemann operator corresponding to the induced unitary connection $A_j \in \mathcal{A}(D_j) \cong \mathcal{A}(E_j/E_{j-1})$. Decompose $\bar\partial_A = \bar\partial_{A_+} + A_0$ with

$$\bar\partial_{A_+} = \begin{pmatrix} \bar\partial_{A_1} & 0 & \ldots & 0 \\ 0 & \bar\partial_{A_2} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \bar\partial_{A_r} \end{pmatrix}, \qquad A_0 = \begin{pmatrix} 0 & A_{12} & \ldots & A_{1r} \\ 0 & 0 & \ldots & A_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 0 \end{pmatrix}$$

We claim that $e^{\mathbf{i}t\xi}(A)$ converges uniformly to $A_+ := A_1 \oplus \cdots \oplus A_r$ as $t \to \infty$. In fact

$$\bar\partial_{A_t} - \bar\partial_{A_+} = e^{\mathbf{i}t\xi}A_0e^{-\mathbf{i}t\xi}$$

and

$$[e^{\mathbf{i}t\xi}A_0e^{-\mathbf{i}t\xi}]_{ij} = -\mathbf{i}e^{t(\lambda_i-\lambda_j)}(\lambda_j - \lambda_i)A_{ij}$$

decays exponentially to zero, since A_0 is strictly upper triangular. This in turn implies that $e^{\mathbf{i}t\xi}A$ converges to A_+ and hence

$$\begin{aligned}
w_\tau(A,\xi) &= \lim_{t\to\infty}\langle *F_{e^{\mathbf{i}t\xi}A},\xi\rangle = \langle *F_{A_+}-\tau,\xi\rangle = \sum_{j=1}^r \langle *F_{A_j},\xi\rangle - \langle\tau,\xi\rangle \\
&= \sum_{j=1}^r \mathbf{i}\lambda_j\int_\Sigma \mathrm{tr}(F_{A_j})\,dvol_\Sigma - \langle\tau,\xi\rangle \\
&= 2\pi\sum_{j=1}^r \lambda_j c_1(D_j) - \langle\tau,\xi\rangle.
\end{aligned}$$

□

Corollary 2.5.5. *Suppose $\xi\in\Omega^0(\Sigma,\mathfrak{g}(E))$ yields a finite weight $w_\tau(A,\xi)$. Then the limit*

$$A_+ := \lim_{t\to\infty} e^{it\xi}A$$

exists in $\mathcal{A}_G(E)$. Moreover, the splitting $E=D_1\oplus\cdots\oplus D_r$ is holomorphic with respect to A_+ and on each factor the holomorphic structure agrees with the one induced by the isomorphism $D_j\cong E_j/E_{j-1}$

Proof. This follows directly from the proof of Proposition 2.5.2. □

Remark 2.5.6. The Corollary shows that $A_+\in\mathcal{G}^c(A)$ if and only if the holomorphic filtration determined by ξ splits holomorphically.

We reformulate the characterization of the finite weights in intrinsic terms. Let $A\in\mathcal{A}(P)\cong\mathcal{A}_G(E)$ and suppose that ξ is a smooth section of $\mathrm{ad}(P)\cong\mathfrak{g}(E)$ which yields a finite weight $w_\tau(A,\xi)$. By Proposition 2.5.2 this defines a holomorphic filtration

$$0<E_1<E_2<\cdots<E_r=E$$

and there exist unitary trivializations of this filtration such that $\xi=\xi_0$ where $\xi_0=-\mathbf{i}\mathrm{diag}(\lambda_1,\ldots,\lambda_r)$ is a block diagonal matrix with $\lambda_1<\lambda_2<\cdots<\lambda_r$. This trivialization yields a reduction $P_{K(\xi)}\subset P$ to $K(\xi_0):=C_G(\xi_0)$. Note that ξ_0 gives rise to a constant central section of $\mathrm{ad}(P_{K(\xi)})\subset\mathrm{ad}(P)$ and agrees with ξ in $\mathrm{ad}(P)$. We can rewrite the formula for the weight as

$$w_\tau(A,\xi) := \int_\Sigma\langle *F_{A_+},\xi\rangle\,dvol_\Sigma - \langle\tau,\xi\rangle$$

where $A_+\in\mathcal{A}(P_{K(\xi)})$ is a $K(\xi)$-connection. It follows from Chern-Weyl theory that the right hand side does not change when we replace A_+ by another $K(\xi)$-connection. The weight depends therefore only on the reduction $P_{K(\xi)}\subset P$ and ξ. The complexification yields a reduction $P^c_{K(\xi)}=P_{L(\xi)}\subset P^c$ to the Levi subgroup $L(\xi_0)\subset G^c$ (see Definition 2.2.10). The reduction $P_{L(\xi)}\subset P^c$ is holomorphic if and only if $\bar\partial_A$ takes values in $\mathfrak{l}(\xi_0)$ and this is the case if and only if the filtration determined by ξ splits holomorphically. In contrast, the extension $L(\xi_0)\subset Q(\xi_0)$

yields a reduction $P_{Q(\xi_0)} \subset P^c$ to the stabilizer of the filtration determined by ξ_0 within G^c. This reduction is always holomorphic, since $\bar{\partial}_A$ is upper block triangular.

Conversely, let $P_Q \subset P^c$ be a holomorphic reduction to a parabolic subgroup $Q = Q(\xi_0) \subset G^c$. This yields a canonical reduction $P_K \subset P$ to $K = C_G(\xi_0)$, since $G^c/Q(\xi_0) \cong G/C_G(\xi_0)$. Since ξ_0 is contained in the center of K, it gives rise to a constant section in $\mathrm{ad}(P_K)$ and its image under the embedding $\mathrm{ad}(P_K) \subset \mathrm{ad}(P)$ yields a section $\xi \in \Omega^0(\Sigma, \mathrm{ad}(P)$ which gives rise to a finite weight $w_\tau(A, \xi)$. We summarize our discussion in the following lemma.

Lemma 2.5.7. *Let $P \to \Sigma$ be a principal G bundle, let $A \in \mathcal{A}(P)$ be a smooth connection and let $P^c := P \times_G G^c$ denote the complexification of P endowed with the holomorphic structure determined by A. There exists a one-to-one correspondence between*

$$\{\xi \in \Omega^0(\Sigma, ad(P)) \,|\, w_\tau(A, \xi) < \infty\}$$

and

$$\left\{ (P_Q, \xi_0) \;\middle|\; \begin{array}{c} \xi_0 \in \mathfrak{g},\ Q = Q(\xi_0) \\ P_Q \textit{ is a principal } Q \textit{ bundle} \\ P_Q \subset P^c \textit{ is a holomorphic reduction} \end{array} \right\}$$

Every reduction $P_Q \subset P^c$ yields a canonical reduction $P_K \subset P$ to $K = C_G(\xi_0)$. The toral generator ξ_0 yields a constant section of $ad(P_K)$ and its image in $ad(P)$ yields ξ. Moreover, the weight is given by the formula

$$w_\tau(A, \xi) = \int_\Sigma \langle *F_B - \tau, \xi \rangle \, dvol_\Sigma$$

for any connection $B \in \mathcal{A}(P_K)$.

Proof. This follows directly from the preceding discussion □

The next lemma describes how the weights behave under an extension $G \hookrightarrow H$ of the structure group.

Lemma 2.5.8. *Let H be a compact connected Lie group and fix an invariant inner product on its Lie algebra $\mathfrak{h}$. Suppose that there exists a monomorphism $G \hookrightarrow H$ which identifies G with a subgroup of H and assume that the invariant inner product on $\mathfrak{g}$ is obtained by restriction of the one on $\mathfrak{h}$. Let $P \to \Sigma$ be a principal G bundle of central type $\tau \in Z(\mathfrak{g})$ defined by (2.12) and denote by $P_H := P \times_G H$ the associated H bundle.*

1. *The central type $\tau_H \in Z(\mathfrak{h})$ of P_H is the image of τ under the orthogonal projection*
$$Z(\mathfrak{g}) \hookrightarrow \mathfrak{h} \cong Z(\mathfrak{h}) \oplus [\mathfrak{h}, \mathfrak{h}] \to Z(\mathfrak{h}).$$

2. *Let $A \in \mathcal{A}(P)$, let $\xi \in \Omega^0(\Sigma, ad(P))$ and denote by $\xi_H \in \Omega^0(\Sigma, ad(P_H))$ the image of ξ under the embedding $ad(P) \subset ad(P_H)$. Then*

$$w_\tau(A, \xi) = w_{\tau_H}(A, \xi_H) + \int_\Sigma \langle \tau_H - \tau, \xi_0 \rangle \, dvol_\Sigma.$$

3. *Let $A \in \mathcal{A}(P)$, let $\xi_H \in \Omega^0(\Sigma, ad(P_H))$ be a section with $w_{\tau_H}(A, \xi) < \infty$ and denote by $\xi \in \Omega^0(\Sigma, ad(P))$ the image of ξ_H under the orthogonal projection $ad(P_H) \to ad(P)$. Then*

$$w_\tau(A, \xi) = w_{\tau_H}(A, \xi_H) + \int_\Sigma \langle \tau_H - \tau, \xi \rangle \, dvol_\Sigma.$$

Proof. For the first part, note that $\mathfrak{h} = Z(\mathfrak{h}) \oplus [\mathfrak{h}, \mathfrak{h}]$ yields an orthogonal decomposition with respect to any invariant inner product of $\mathfrak{h}$. The orthogonal projection of τ onto $Z(\mathfrak{h})$ does therefore depend only on the embedding of G into H and it is easy to verify that it satisfies (2.12) for P_H.

By Lemma 2.5.7 there exists $\xi_0 \in \mathfrak{g}$ and a reduction $P_K \subset P$ to a principal $K = C_G(\xi_0)$ bundle such that ξ is the image of the constant section ξ_0 under the embedding $\mathrm{ad}(P_K) \subset \mathrm{ad}(P)$. Moreover,

$$w_\tau(A, \xi) = \int_\Sigma \langle *F_B - \tau, \xi \rangle \, dvol_\Sigma$$

for any connection $B \in \mathcal{A}(P_K)$. Define $\tilde{K} = C_H(\xi_0)$ and $P_{\tilde{K}} := P_K \times_K \tilde{K} \subset P_H$. Then ξ_H agrees with the image of ξ_0 under the embedding $\mathrm{ad}(P_{\tilde{K}}) \subset \mathrm{ad}(P_H)$ and Lemma 2.5.7 yields

$$w_{\tau_H}(A, \xi) = \int_\Sigma \langle *F_B - \tau_H, \xi_0 \rangle$$

for any connection $B \in \mathcal{A}(P_{\tilde{K}})$. In particular, for $B \in \mathcal{A}(P_K) \subset \mathcal{A}(P_{\tilde{K}})$, we get

$$w_\tau(A, \xi) - w_{\tau_H}(A, \xi) = \int_\Sigma \langle \tau_H - \tau, \xi_0 \rangle \, dvol_\Sigma$$

and this proves the second part.

The third part follows by a similar argument. Note that the proof of Proposition 2.5.2 implies that there exists a connection $B = A_+ \in \mathcal{A}(P) \cap \mathcal{A}(P_{\tilde{K}})$ for the reduction $P_{\tilde{K}} \subset P_H$ associated to ξ_H. For such a connection holds $\langle \xi_H, F_B \rangle = \langle \xi, F_B \rangle$ and the claim follows as in the second part. □

2.5.2 Weights and algebraic stability

The following proposition characterizes the (algebraic) stability of the holomorphic principal bundle (P^c, J_A) in terms of the associated weights $w_\tau(A, \xi)$.

Proposition 2.5.9 (Characterization of Stability). *Let P be a principal G bundle of central type $\tau \in Z(\mathfrak{g})$ defined by (2.12). Let $A \in \mathcal{A}(P)$ be a smooth connection and let $P^c := P \times_G G^c$ be the complexified principal bundle endowed with the induced holomorphic structure J_A.*

1. *(P^c, J_A) is stable if and only if $w_\tau(A, \xi) > 0$ for all $\xi \in W^{1,2}(\Sigma, ad(P))$ which are not constant central sections.*

2. *(P^c, J_A) is polystable if and only if $w_\tau(A, \xi) \geq 0$ for all $\xi \in W^{1,2}(\Sigma, ad(P))$ and whenever $w_\tau(A, \xi) = 0$ the associated (smooth) reduction $P_{L(\xi)} \subset P_{Q(\xi)} \subset P^c$ is holomorphic.*

3. *(P^c, J_A) is semistable if and only if $w_\tau(A,\xi) \geq 0$ for all $\xi \in W^{1,2}(\Sigma, ad(P))$.*

4. *(P^c, J_A) is unstable if and only if there exists $\xi \in W^{1,2}(\Sigma, ad(P))$ with $w_\tau(A,\xi) < 0$.*

Proof. Using the geometric interpretation of the finite weights in Lemma 2.5.7 we can reduce the proof to a lemma of Ramanathan [95]. The proof will be given on page 78 below. □

Reduction argument

We reduce the theorem to the case where $Z(G)$ is discrete and $\tau = 0$. Recall that the invariant inner product on $\mathfrak{g}$ yields the decomposition $\mathfrak{g} = Z(\mathfrak{g}) \oplus [\mathfrak{g}, \mathfrak{g}]$ of the Lie algebra into its center and a semisimple subalgebra. The center yields a trivial $Z(\mathfrak{g})$ subbundle $V \subset \mathrm{ad}(P)$ and its orthogonal complement can be identified with $\mathrm{ad}(P/Z_0(G))$.

Lemma 2.5.10. *Assume the setting of Proposition 2.5.9. Let $\xi \in \Omega^0(\Sigma, ad(P))$ with $w_\tau(A,\xi) < \infty$ and decompose $\xi = \xi^z + \xi^{ss}$ with respect to the splitting $ad(P) = V \oplus ad(P/Z_0(G))$. Then*

$$w_\tau(A,\xi) = w_0(\bar{A}, \xi^{ss})$$

where $\bar{A} \in \mathcal{A}(P/Z_0(G))$ denotes the induced connection on $P/Z_0(G)$.

Proof. By Lemma 2.5.7 exists a reduction $P_K \subset P$ and an element $\xi_0 \in \mathfrak{g}$ which gives rise to a constant central section in $\mathrm{ad}(P_K)$ and such that ξ is the image of ξ_0 under the embedding $\mathrm{ad}(P_K) \subset \mathrm{ad}(P)$. Decompose $\xi_0 = \xi_0^z + \xi_0^{ss}$ with respect to $\mathfrak{g} = Z(\mathfrak{g}) \oplus [\mathfrak{g}, \mathfrak{g}]$. Then ξ_0^z yields ξ^z and ξ_0^{ss} yields ξ^{ss} under the embedding $\mathrm{ad}(P_K) \subset \mathrm{ad}(P)$. By Lemma 2.5.7 the weight is given by

$$w_\tau(A,\xi) = \int_\Sigma \langle *F_B - \tau, \xi^{ss} \rangle \, dvol_\Sigma + \int_\Sigma \langle *F_B - \tau, \xi^z \rangle \, dvol_\Sigma.$$

for any connection $B \in \mathcal{A}(P_K)$. The second integral vanishes by (2.12) and in the first integral yields

$$\int_\Sigma \langle *F_B - \tau, \xi^{ss} \rangle \, dvol_\Sigma = \int_\Sigma \langle *F_B, \xi^{ss} \rangle \, dvol_\Sigma = w_0(\bar{A}, \xi^{ss}).$$

since $\tau \in Z(\mathfrak{g})$ is orthogonal to $[\mathfrak{g}, \mathfrak{g}]$. This completes the proof. □

The main argument

The following result is a reformulation of Lemma 2.1 in [95].

Lemma 2.5.11. *Assume the setting of Proposition 2.5.9 and suppose in addition that $Z_0(G)$ is discrete and $\tau = 0$. (P^c, J_A) is stable (resp. semistable) with respect to Definition 2.3.2 if and only if $w_0(A,\xi) > 0$ (resp. $w_0(A,\xi) \geq 0$) for all $\xi \in W^{1,2}(\Sigma, ad(P))$.*

Proof. Let $\xi \in \Omega^0(\Sigma, \mathrm{ad}(P))$ with $w_0(A,\xi) < \infty$ be given. By Lemma 2.5.7 exists a reduction $P_K \subset P$ and an element $\xi_0 \in \mathfrak{g}$ such that $K = C_G(\xi_0)$ and ξ is the image of ξ_0 under the embedding $\mathrm{ad}(P_K) \subset \mathrm{ad}(P)$.

Let $T \subset G$ be a maximal torus whose Lie algebra contains ξ_0 and let $R_0^+ = \{\alpha_1, \dots, \alpha_r\}$ be a system of simple roots with respect to T whose Weyl-chamber contains ξ_0. Recall that $\alpha_j = \mathbf{i}a_j$ with $a_j \in \mathrm{Hom}(\mathfrak{t}, \mathbb{R})$ and define $t_j \in \mathfrak{t}$ by $a_j = \langle t_j, \cdot \rangle$. The elements $\check{t}_1, \dots, \check{t}_r \in \mathfrak{t}$ defined by (2.7) yield a basis of $\mathfrak{t}$ and ξ_0 has the shape

$$\xi_0 = \sum_{j=1}^r x_j \check{t}_j$$

with $x_j \geq 0$. Note that $\check{t}_j$ lies in the center of the Lie algebra of $K = C_G(\xi_0)$ when $x_j > 0$. Then $\check{t}_j$ gives rise to a constant central section of $\mathrm{ad}(P_K)$ and

$$w_0(A,\xi) = \sum_{j=1}^r x_j w_0(A, \check{t}_j). \tag{2.37}$$

Fix $1 \leq j \leq r$ with $x_j > 0$ and denote $Q_j := Q(\check{t}_j)$. This is a maximal parabolic subgroup of G^c which contains $Q(\xi_0)$ and the extension $P_{Q_j} := P_{Q(\xi)} \times_{Q(\xi)} Q_j \subset P^c$ yields a maximal parabolic reduction. Let $\chi : Q_j \to \mathbb{C}^*$ be the determinant of the action of Q_j on its Lie algebra and denote by $\dot{\chi} : \mathfrak{q}_j \to \mathbb{C}$ the induced map on the Lie algebra. Chern-Weyl theory yields the relation

$$c_1(\mathrm{ad}(P_{Q_j})) = \frac{\mathbf{i}}{2\pi} \int_\Sigma \dot{\chi}(F_B)$$

for a connection $B \in \mathcal{A}(P_K)$. For $\eta \in \mathfrak{q}_j$ the value of $\dot{\chi}(\eta)$ is given as the trace of $\mathrm{ad}(\eta) := [\eta, \cdot]$ acting on

$$\mathfrak{q}_j = \mathfrak{t} \oplus \bigoplus_{\alpha \in R(\check{t}_j)} \mathfrak{g}_\alpha. \tag{2.38}$$

where $R(\check{t}_j)$ is defined by (2.8). This decomposition is unitary and by definition of the roots we have $\mathrm{ad}(t)e_\alpha = \alpha(t)e_\alpha$ for $t \in \mathfrak{t}$. This shows

$$\dot{\chi}(\eta) = \sum_{\alpha \in R(\check{t}_j)} \alpha(\eta) \tag{2.39}$$

for all $\eta \in \mathfrak{t}$. Since $\dot{\chi}$ vanishes on $[\mathfrak{q}_j, \mathfrak{q}_j]$ it vanishes on all root space $\mathfrak{g}_\alpha$ with $\{\alpha, -\alpha\} \subset R(\check{t}_j)$. These are the roots in $\tilde{R}(\check{t}_j)$ which produce the Levi subgroup $L(\check{t}_j)$. The remaining root spaces $\mathfrak{g}_\alpha$ with $\alpha \in R(\check{t}_j) \backslash \tilde{R}(\check{t}_j)$ form a nilpotent subalgebra. This shows that (2.39) remains valid for all $\eta \in \mathfrak{q}_j$ if one extends the roots by complex linearity over $\mathfrak{t}^c$ and by zero over the root spaces.

Denote by R^+ the positive roots and by $R^-(\check{t}_j) = R(\check{t}_j) \backslash R^+$ the negative roots whose root spaces are contained in $\mathfrak{q}_j$. Then $\dot{\chi} = \gamma_1 + \gamma_2$ with

$$\gamma_1 = \sum_{\alpha \in R^+} \alpha, \qquad \gamma_2 := \sum_{\alpha \in R^-(\check{t}_j)} \alpha$$

and

$$\langle \alpha_i, \gamma_1 \rangle = \sum_{\alpha \in R^+} \langle t_i, t_\alpha \rangle = |t_i|^2 + \sum_{\alpha \in R^+ \setminus \{\alpha_j\}} \langle t_i, t_\alpha \rangle$$

holds for every simple root α_i. The root reflection

$$s_j : \mathfrak{t} \to \mathfrak{t}, \qquad s_j(t) := t - \frac{2\langle t, t_j \rangle}{|t_j|^2} t_j$$

restricts to a permutation of $R^+ \setminus \{\alpha_j\}$. Indeed, any root has a unique representation $t_\alpha = \sum_{k=1}^r c_k t_k$ and all coefficients happen to have the same sign. Applying the reflection s_j changes only the coefficient c_j and thus $s_j(\alpha)$ remains positive if $c_k > 0$ for some coefficient $k \neq j$. Using this symmetry we conclude

$$\langle \alpha_i, \gamma_1 \rangle = |t_i|^2. \tag{2.40}$$

A similar argument shows for $i \neq j$

$$\langle \alpha_i, \gamma_2 \rangle = \sum_{\alpha \in R^-(\check{t}_j)} \langle t_i, t_\alpha \rangle = -|t_i|^2 + \sum_{\alpha \in R^-(\check{t}_j) \setminus \{\alpha_j\}} \langle t_i, t_\alpha \rangle = -|t_i|^2. \tag{2.41}$$

This shows $\dot{\chi}(t_i) = 0$ for $i \neq j$. As a general property of root systems (see [70] Lemma 2.51) it holds $\langle t_j, t_i \rangle \leq 0$ for distinct simple roots α_i, α_j and thus

$$\dot{\chi}(t_j) = |t_j|^2 + \sum_{\alpha \in R^-(\check{t}_j)} \langle t_j, t_\alpha \rangle > 0 \tag{2.42}$$

Combining (2.40), (2.41) and (2.42) we conclude

$$\dot{\chi}(t) = \mathbf{i} m \langle \check{t}_j, t \rangle$$

for some $m > 0$. Hence

$$c_1(\mathrm{ad}(P_{Q_j})) = \frac{-m}{2\pi} \int_\Sigma \langle F_B, \check{t}_j \rangle = \frac{-m}{2\pi} w_0(A, \check{t}_j). \tag{2.43}$$

Suppose now that P^c is stable (resp. semistable). Then the left hand side in (2.43) is negative (resp. nonpositive) and (2.37) implies $w_0(A, \xi) > 0$ (resp. $w_0(A, \xi) \geq 0$). Conversely, Lemma 2.5.7 show that every holomorphic reduction $P_Q \subset P^c$ to a proper maximal parabolic subgroup $Q(\xi_0) \subset Q$ is induced by some $\xi \in \Omega^0(\Sigma, \mathrm{ad}(P))$ with $w_0(A, \xi) < \infty$. Lemma 2.2.15 shows that in (2.38) exactly one coefficient x_j does not vanishes. Hence (2.43) implies that $c_1(\mathrm{ad}(P_Q))$ is negative or vanishes if and only if $w_0(A, \xi)$ is positive or vanishes respectively. This establishes the converse direction and completes the proof of the lemma. □

Completion of the proof

Proof of Proposition 2.5.9. We may assume by Lemma 2.3.5 and Lemma 2.5.10 that $Z_0(G)$ is discrete and $\tau = 0$. The stable and semistable case follow then from Lemma 2.5.11 and the unstable case is equivalent to the semistable case.

Assume that P^c is polystable. Then there exists a holomorphic reduction $P_L \subset P^c$ to a Levi subgroup $L \subset G^c$ and P_L is a stable L bundle. Let $\xi \in \Omega^0(\Sigma, \mathrm{ad}(P))$ with $w_0(A,\xi) = 0$ be given. By Lemma 2.5.7 exists $\xi_0 \in \mathfrak{g}$ and a reduction $P_K \subset P$ to a principal $K = C_G(\xi_0)$ bundle such that ξ agrees with the image of ξ_0 under the embedding $\mathrm{ad}(P_K) \subset \mathrm{ad}(P)$. Using the notation from the proof of Lemma 2.5.11 above, write ξ_0 with respect to a system of simple roots as

$$\xi_0 = \sum_{j=1}^{r} x_j \check{t}_j$$

with $x_j \geq 0$. Since P^c is in particular semistable, the proof of Lemma 2.5.11 shows that $w_0(A,\xi) = 0$ if and only if

$$x_j > 0 \quad \Rightarrow \quad c_1(\mathrm{ad}(P_{Q_j})) = 0$$

where $Q_j := Q(\check{t}_j)$. We may assume (after conjugation) that $L = L(\eta_0)$ for some $\eta_0 \in \mathfrak{g}$ and η_0 is contained in the Weyl-chamber determined by our choice of simple roots. If L is not contained in Q_j, then $Q'_j := L \cap Q_j$ is a maximal parabolic subgroup of L and we have an induced reduction $P_{Q'_j} \subset P_L$. Since L and G^c are reductive, the Lie algebra bundles $\mathrm{ad}(P_L)$ and $\mathrm{ad}(P^c)$ carry a non degenerated symmetric $\mathbb{C}$-bilinear form. Hence they are both self-dual and have vanishing first Chern-class. This shows

$$\begin{aligned} c_1(\mathrm{ad}(P_{Q_j})) &= -c_1(\mathrm{ad}(P^c)/\mathrm{ad}(Q_j)) = -c_1(\mathrm{ad}(P_L)/\mathrm{ad}(Q'_j)) \\ &= c_1(\mathrm{ad}(P_{Q'_j})) < 0 \end{aligned}$$

where the last step follows from the stability of P_L. We have thus proven that $L \subset Q_j$ whenever $x_j > 0$ and this yields $L \subset L(\xi_0)$. Since the reduction to L is holomorphic, so is the reduction to $L(\xi_0)$.

Assume conversely, that all weights are nonnegative and if $\xi \in \Omega^0(\Sigma, P^c)$ is a section with $w_0(A,\xi) = 0$ then $P_{L(\xi)} \subset P^c$ is a holomorphic reduction (where $P_{L(\xi)} = P^c_{K(\xi)}$ and $P_{K(\xi)}$ is determined by Lemma 2.5.7). It follows from Lemma 2.5.11 that P^c is semistable. If P^c is in fact stable, then we are done. Otherwise there exists a vanishing weight $w_0(A,\xi) = 0$ and by assumption this yields a holomorphic reduction $P_{L(\xi)} \subset P^c$. In particular A restricts to a connection on $P_{K(\xi)} \subset P$ and $P_{K(\xi)}$ is again of central type 0. For the later claim let $\eta \in \mathfrak{g}$ be contained in the center of the Lie algebra of K and consider its image η' under the embedding $\mathrm{ad}(P_K) \subset \mathrm{ad}(P)$. Then follows

$$\int_\Sigma \langle *F_B, \eta \rangle \, dvol_\Sigma = w_0(A, \eta') \geq 0.$$

for any connection $B \in \mathcal{A}(P_K)$. Replacing η by $-\eta$ shows that this expression must vanish and hence $P_{K(\xi)}$ is of central type 0. Now Lemma 2.5.11 shows that $P_{L(\xi)}$ is again semistable. If $P_{L(\xi)}$ is not stable, then there exists $\tilde{\xi} \in \Omega^0(\Sigma, \mathrm{ad}(P_{K(\xi)})$ with $w_0(A,\tilde{\xi}) = 0$. We can consider $\tilde{\xi}$ as section ξ' of $\mathrm{ad}(P)$ which then satisfies $w_0(A,\xi') = 0$ and thus yields a strictly smaller holomorphic reduction $P_{L(\xi')} \subset P_{L(\xi)}$. If we replace ξ by ξ' and rerun the argument from above we obtain after finitely many iterations a section ξ which satisfies $w_0(A,\xi) = 0$ and yields a stable holomorphic reduction $P_{L(\xi)} \subset P^c$.

Let $\chi : L \to \mathbb{C}^*$ be a character. We need to show $c_1(\chi(P_{L(\xi)})) = 0$. Decompose $\xi_0 = \sum_{j=1}^r x_j \check{t}_j$ as above and denote

$$S := \{j \,|\, x_j > 0\}.$$

Since $\dot{\chi} : \mathfrak{l}(\xi) \to \mathbb{C}$ vanishes on $[\mathfrak{l}(\xi), \mathfrak{l}(\xi)]$, it vanishes on all the root spaces $\mathfrak{g}_\alpha$ belonging to $\mathfrak{l}(\xi)$ and the dual vectors $t_\alpha \in \mathfrak{t}$. In particular, $\dot{\chi}$ vanishes on the simple roots t_j with $j \notin S$ and has the shape

$$\dot{\chi}(\eta) = \sum_{j \in S} \mathbf{i} r_j \langle \eta, \check{t}_j \rangle$$

for some $r_j \in \mathbb{R}$. Chern-Weyl theory yields

$$c_1(\chi(P_{L(\xi)})) = \frac{\mathbf{i}}{2\pi} \int_\Sigma \dot{\chi}(F_B) = \frac{\mathbf{i}}{2\pi} \sum_{j \in S} \mathbf{i} r_j \int_\Sigma \langle *F_B, \check{t}_j \rangle$$

for some connection $B \in \mathcal{A}(P_{K(\xi)})$. We claim that each summand vanishes separately in the last expression. This follows from the assumption

$$0 = w_0(A, \xi) = \sum_{j=1}^r x_j \int_\Sigma \langle *F_B, \check{t}_j \rangle \, dvol_\Sigma = \sum_{j \in S} x_j \int_\Sigma \langle *F_B, \check{t}_j \rangle \, dvol_\Sigma$$

and

$$w_0(A, \check{t}_j) = \int_\Sigma \langle *F_B, \check{t}_j \rangle \, dvol_\Sigma \geq 0$$

since P^c is semistable. □

2.5.3 The moment weight inequality

The moment-weight inequality provides a lower bound for the norm of the moment-map $\mu_\tau(A) = *F_A - \tau$ on the complexified orbit $\mathcal{G}^c(A)$.

Theorem 2.5.12 (The moment-weight inequality). *Let $P \to \Sigma$ be a principal G bundle of central type $\tau \in Z(\mathfrak{g})$ defined by (2.12). Let $A \in \mathcal{A}(P)$ be a smooth connection and $\xi \in W^{1,2}(\Sigma, ad(P))$. Then*

$$-\frac{w_\tau(A, \xi)}{||\xi||_{L^2}} \leq \inf_{g \in \mathcal{G}^c} || * F_{g(A)} - \tau ||_{L^2}. \tag{2.44}$$

The moment weight-inequality is essentially proven by Atiyah and Bott ([4], Prop. 8.13 and Prop. 10.13). They explicitly determine the infimum of the Yang-Mills functional over $\mathcal{G}^c(A)$ in terms of the Harder-Narasimhan filtration of the holomorphic vector bundle ad(P^c). It follows from the proof of the dominant weight theorem (Theorem 2.7.1) in the next section that the same description yields the supremum over the left-hand side whenever it is positiv. We provide a different approach following the arguments in [51] for the finite dimensional case which are essentially due to Chen [19, 18] and Donaldson [39].

Proof. We reduce the proof to the case where $Z(G)$ is discrete and $\tau = 0$. Denote by $\bar{A} \in \mathcal{A}(P/Z_0(G))$ the induced connection on the quotient bundle and decompose $\xi = \xi^{ss} + \xi^z$ as in Lemma 2.5.7. Let $g \in \mathcal{G}^c(P)$ be given and decompose $F_{gA} = F^{ss} + F^z$ in the same way. Note that $F_{g\bar{A}} = F^{ss}$. Suppose that the moment-weight inequality is satisfied on $P/Z_0(G)$, i.e.

$$-\frac{w_0(\bar{A}, \xi^{ss})}{||\xi^{ss}||_{L^2}} \le || * F_{g\bar{A}}||_{L^2}.$$

We may assume $w_\tau(A, \xi) \le 0$. Then Lemma 2.5.7 implies

$$-\frac{w_\tau(A, \xi)}{||\xi||_{L^2}} \le -\frac{w_0(\bar{A}, \xi^{ss})}{||\xi^{ss}||_{L^2}} \le || * F_{g\bar{A}}||_{L^2} \le ||F_{gA} - \tau||_{L^2}$$

and this completes the reduction argument.

Now assume that $Z(G)$ is discrete and $\tau = 0$. Let $\xi \in W^{1,2}(\Sigma, \mathrm{ad}(P))$ with $w_0(A, \xi) < \infty$. Then ξ is smooth by Proposition 2.5.2 and the limit

$$\lim_{t\to\infty} e^{\mathbf{i}t\xi} A =: A_+ \tag{2.45}$$

exists by Corollary 2.5.5.

Let $g_0 = u_0 e^{\mathbf{i}\eta_0} \in \mathcal{G}^c(P)$ be given and define $\eta(t) \in W^{2,2}(\Sigma, \mathrm{ad}(P))$ and $u(t) \in \mathcal{G}$ by the equation

$$e^{\mathbf{i}\xi t} = e^{\mathbf{i}\eta(t)} u(t) g_0.$$

From this follows pointwise the estimate

$$||\eta(t) - t\xi|| \le ||\eta_0||. \tag{2.46}$$

To see this, denote by $\pi : G^c \to G^c/G$ the canonical projection and recall that G^c/G is a complete simply-connected Riemannian manifold with nonpositive sectional curvature. For a fixed time t and $z \in \Sigma$ define $p := \pi(e^{\mathbf{i}t\xi(z)})$ and $q := \pi(e^{\mathbf{i}\eta(t,z)})$. Then

$$\gamma : [0, 1] \to G^c/G, \qquad \gamma(s) := \pi(e^{\mathbf{i}t\xi(z)} e^{-\mathbf{i}s\eta_0(z)})$$

is the unique geodesic from p to q in G^c/G of length $||\eta_0(z)||$. Since the exponential map on a Riemannian manifold with nonpositive curvature is distance increasing, this yields

$$||\eta(t, z) - t\xi(z)|| \le \mathrm{dist}_{G^c/G}(p, q) = ||\eta_0(z)||$$

and hence (2.46). With this estimate we get

$$\begin{aligned}
\left|\left|\frac{\xi}{||\xi||_{L^2}} - \frac{\eta(t)}{||\eta(t)||_{L^2}}\right|\right|_{L^2} &\le \left|\left|\frac{t\xi - \eta(t)}{t||\xi||_{L^2}} + \frac{\eta(t)}{t||\xi||_{L^2}} - \frac{\eta(t)}{||\eta(t)||_{L^2}}\right|\right|_{L^2} \\
&\le \frac{||t\xi - \eta(t)||_{L^2}}{t||\xi||_{L^2}} + \left|\frac{||\eta(t)||_{L^2} - t||\xi||_{L^2}}{t||\xi||_{L^2}}\right| \\
&\le 2\frac{||\eta_0||_{L^2}}{t||\xi||_{L^2}}
\end{aligned}$$

and hence

$$\lim_{t\to\infty}\left|\left|\frac{\eta(t)}{||\eta(t)||_{L^2}}-\frac{\xi}{||\xi||_{L^2}}\right|\right|_{L^2}=0. \tag{2.47}$$

By (2.31) the map

$$s\mapsto\langle *F_{e^{\mathbf{i}su^{-1}\eta u}g_0A},u^{-1}\eta u\rangle$$

is nondecreasing in s. With the relation $e^{\mathbf{i}u^{-1}\eta u}g_0=u^{-1}e^{\mathbf{i}t\xi}$ follows

$$\begin{aligned}
-||*F_{g_0A}||_{L^2}&\le\frac{1}{||\eta||_{L^2}}\left\langle *F_{g_0A},u^{-1}\eta u\right\rangle\le\frac{1}{||\eta||_{L^2}}\left\langle *F_{e^{\mathbf{i}u^{-1}\eta u}g_0A},u^{-1}\eta u\right\rangle\\
&\le\frac{1}{||\eta||_{L^2}}\left\langle *F_{u^{-1}e^{\mathbf{i}t\xi}A},u^{-1}\eta u\right\rangle=\frac{1}{||\eta||_{L^2}}\langle *F_{e^{\mathbf{i}t\xi}A},\eta\rangle\\
&\le\frac{\langle *F_{e^{\mathbf{i}t\xi}A},\xi\rangle}{||\xi||_{L^2}}+\left\langle *F_{e^{\mathbf{i}t\xi}A},\frac{\eta}{||\eta||_{L^2}}-\frac{\xi}{||\xi||_{L^2}}\right\rangle
\end{aligned}$$

It follows from (2.45) and (2.46) that the right and side converges to $\frac{w_0(A,\xi)}{||\xi||}$ for $t\to\infty$ and this proves the theorem.

□

2.6 The Kempf-Ness functional

Let G be a compact connected Lie group and let $P\to\Sigma$ be a principal G bundle of central type $\tau\in Z(\mathfrak{g})$ defined by (2.12). Let $A\in\mathcal{A}(P)$ be a smooth connection. The **Kempf-Ness functional** associated to A is the $\mathcal{G}(P)$-invariant functional

$$\Phi_A:\mathcal{G}^c(P)\to\mathbb{R},\qquad \Phi_A(e^{\mathbf{i}\xi}u)=\int_0^1\langle *F_{e^{-\mathbf{i}t\xi}A}-\tau,-\xi\rangle\,dt. \tag{2.48}$$

We show in Lemma 2.6.1 below that the derivative of Φ_A is given by

$$\alpha_A(g;\hat g)=-\langle *F_{g^{-1}A}-\tau,\mathrm{Im}(g^{-1}\hat g)\rangle. \tag{2.49}$$

The asymptotic slope of Φ_A along the geodesic ray $t\mapsto e^{-\mathbf{i}t\xi}$ yields the weight $w_\tau(A,\xi)$. This is related to the stability of the associated holomorphic principal bundle (P^c,J_A) by Proposition 2.5.9. On the other hand, it follows directly from (2.49) that $g\in\mathcal{G}^c(P)$ is a critical point of Φ_A if and only if $*F_{g^{-1}A}=\tau$. The analog of the Kempf-Ness theorem in classical GIT is Theorem 2.6.2 below. It characterizes the different notions of μ_τ-stability in terms of the global behaviour of Φ_A and thus provides a link between the algebraic and the symplectic notions of stability. We can deduce from this the Narasimhan-Seshadri-Ramanathan theorem in the second subsection.

2.6.1 The generalized Kempf-Ness theorem

Lemma 2.6.1. *Let $P \to \Sigma$ be a principal G bundle and define $\Phi_A : \mathcal{G}^c(P) \to \mathbb{R}$ by (2.48).*

1. *The derivative of Φ_A is given by*
$$\alpha_A(g;\hat{g}) = -\langle *F_{g^{-1}A} - \tau, Im(g^{-1}\hat{g})\rangle.$$

2. *Let $g, h \in \mathcal{G}^c(P)$, then*
$$\Phi_{h^{-1}A}(h^{-1}g) = \Phi_A(g) - \Phi_A(h).$$

Proof. Let $g \in \mathcal{G}^c(P)$, $\hat{g} \in T_g\mathcal{G}^c(P)$ and let $u \in \mathcal{G}(P)$ be given. Then

$$\begin{aligned}\alpha_A(gu^{-1}, \hat{g}u^{-1}) &= \langle *F_{ug^{-1}A}, \mathrm{Im}(ug^{-1}\hat{g}u^{-1})\rangle \\ &= \langle u * F_{g^{-1}A}u^{-1}, u\mathrm{Im}(g^{-1}\hat{g})u^{-1}\rangle \\ &= \alpha_A(g, \hat{g})\end{aligned}$$

shows that α_A is invariant under the right-action of $\mathcal{G}(P)$ and hence descends to a 1-form on $\mathcal{G}^c(P)/\mathcal{G}(P)$.

We claim that α_A is closed. Denote by $\pi : G^c \to G^c/G$ the canoncial projection and let $\hat{g}_1 = d\pi(g)g\mathbf{i}\xi$ and $\hat{g}_2 := d\pi(g)g\mathbf{i}\eta$ be two tangent vectors in $T_{\pi(g)}\mathcal{G}^c(P)/\mathcal{G}(P)$. Then

$$\begin{aligned}d\alpha_A(g;\hat{g}_1,\hat{g}_2) &= d\alpha_A(g;\hat{g}_2)[\hat{g}_1] - d\alpha_A(g;\hat{g}_1)[\hat{g}_2] - \alpha_A(g;[\hat{g}_1,\hat{g}_2]) \\ &= d\langle F_{g^{-1}A} - \tau, \eta\rangle[g\mathbf{i}\xi] - d\langle F_{g^{-1}A} - \tau, \xi\rangle[g\mathbf{i}\eta] \\ &= \langle d^*_{g^{-1}A}d_{g^{-1}A}\xi, \eta\rangle - \langle d^*_{g^{-1}A}d_{g^{-1}A}\eta, \xi\rangle = 0.\end{aligned}$$

We used in the second step that $[\hat{g}_1, \hat{g}_2] \in T_g\mathcal{G}(P)$ is tangent to the real gauge orbit and thus lies in the kernel of $\alpha_A(g;\cdot)$.

Denote for $p, q \in \mathcal{G}^c(P)/\mathcal{G}(P)$ by $[p, q]$ the geodesic segment connecting p to q. Then (2.48) can be reformulated as

$$\Phi_A(g) = \int_{[\pi(\mathbb{1}),\pi(g)]} \alpha_A. \tag{2.50}$$

For $h \in \mathcal{G}^c(P)$ we have $\alpha_{h^{-1}A}(h^{-1}g, h^{-1}\hat{g}) = \alpha_A(g, \hat{g})$ and hence

$$\Phi_{h^{-1}A}(h^{-1}g) = \int_{[\pi(\mathbb{1}),\pi(h^{-1}g)]} \alpha_{h^{-1}A} = \int_{[\pi(h),\pi(g)]} \alpha_A.$$

Since α_A is closed we have

$$\int_{[\pi(h),\pi(g)]} \alpha_A = \int_{[\pi(\mathbb{1}),\pi(g)]} \alpha_A - \int_{[\pi(\mathbb{1}),\pi(h)]} \alpha_A = \Phi_A(g) - \Phi_A(h)$$

and this establishes the second part of the lemma.

Using the second part, we can can reduce the proof of the first part to the case $g = \mathbb{1}$ and in this case the claim follows directly from (2.48).

□

The difficult part of the following theorem is the stable case. The proof of this case is due to Bradlow [12] and Mundet [89] in the context of more general moduli problems.

Theorem 2.6.2 (**Generalized Kempf-Ness Theorem**). *Let G be a compact connected Lie group, let $P \to \Sigma$ be a principal G bundle with central type $\tau \in Z(\mathfrak{g})$ defined by (2.12) and let $A \in \mathcal{A}(P)$.*

1. *A is μ_τ-stable if and only if $\mathcal{G}^c(A)$ has discrete $\mathcal{G}^c(P)$ isotropy and for every $R > 0$ such that*

$$M_R := \{\xi \in W^{2,2}(\Sigma, ad(P) \,|\, ||*F_{e^{-i\xi}A} - \tau||_{L^2} \le R\}$$

   is nonempty, there exist constants $c_1, c_2 > 0$ such that

$$\Phi_A(e^{\mathbf{i}\xi}) \le c_1 ||\xi||_{L^\infty} + c_2 \qquad \text{for all } \xi \in M_R. \tag{2.51}$$

2. *A is μ_τ-polystable if and only if Φ_A has a critical point.*
3. *A is μ_τ-semistable if and only if Φ_A is bounded below.*
4. *A is μ_τ-unstable if and only if Φ_A is unbounded below.*

Proof. We consider both implications of the stable case in the following lemmas first. The proof will then be given on page 87 below. □

Lemma 2.6.3. *Assume the setting of Theorem 2.6.2. Suppose that the orbit $\mathcal{G}^c(A) \subset \mathcal{A}^*(P)$ contains only irreducible connections and that there exist $c_1, c_2, R > 0$ such that M_R is nonempty and (2.51) holds. Then exists $\xi_0 \in M_R$ such that*

$$\Phi_A(e^{\mathbf{i}\xi_0}) \le \Phi_A(e^{\mathbf{i}\xi}) \qquad \text{for all } \xi \in M_R \tag{2.52}$$

and $B := e^{-\mathbf{i}\xi_0}A$ satisfies $F_B = 0$.

Proof. Suppose first that $\xi_0 \in M_R$ satisfies (2.52). Let $B := e^{-\mathbf{i}\xi_0}A$ and let $\eta \in W^{2,2}(\Sigma, \mathrm{ad}(P))$ be a solution of the equation

$$\Delta_B \eta = d_B^* d_B \eta = *F_B$$

which exists since B is irreducible. Then follows

$$\left.\frac{d}{dt}\right|_{t=0} \Phi_A(e^{\mathbf{i}\xi_0} e^{\mathbf{i}\eta t}) = \alpha_A(e^{\mathbf{i}\xi_0}, e^{\mathbf{i}\xi_0}\mathbf{i}\eta) = -\langle *F_B, \eta\rangle = -||d_B\eta||^2_{L^2}$$

and

$$\begin{aligned}\left.\frac{d}{dt}\right|_{t=0} ||*F_{e^{-\mathbf{i}\eta t}e^{-\mathbf{i}\xi_0}A}||^2 &= 2\left\langle *F_B, *\left.\frac{d}{dt}\right|_{t=0} F_{e^{-\mathbf{i}\eta t}B}\right\rangle \\ &= 2\langle *F_B, *d_B * d_B\eta\rangle = -2\langle *F_B, \Delta_B\eta\rangle = -2||*F_B||^2_{L^2}\end{aligned}$$

Now decompose $e^{\mathbf{i}\xi_0}e^{\mathbf{i}\eta t} = e^{\mathbf{i}\xi_1}u$. Then the calculation shows that for sufficiently small t we have $\xi_1 \in M_B$ and $\Phi_A(e^{\mathbf{i}\xi_1}) \le \Phi_A(e^{\mathbf{i}\xi_0})$ with equality if and only if $F_B = 0$. Since (2.52) yields the converse inequality, we have indeed equality and hence $F_B = 0$.

It remains to prove the existence of a minimizer $\xi_0 \in M_R$. Let $\{\xi_k\} \subset M_R$ be a minimizing sequence satisfying

$$\lim_{k\to\infty} \Phi_A(e^{\mathbf{i}\xi_k}) = \inf_{\xi\in M_R} \Phi_A(e^{\mathbf{i}\xi}). \tag{2.53}$$

By definition of M_R, the curvature $F_{e^{\mathbf{i}\xi_k}A}$ is uniformly bounded in L^2. Hence the Uhlenbeck compactness theorem asserts that there exists $u_k \in \mathcal{G}(P)$ such that $A_k := u_k e^{\mathbf{i}\xi_k} A$ converges weakly in $W^{1,2}$. For $g_k := u_k e^{\mathbf{i}\xi_k}$ the expression

$$\bar\partial_{A_k} - \bar\partial_A = g_k^{-1}\bar\partial_A g_k$$

is thus uniformly bounded in $W^{1,2}$. Since ξ_k is uniformly bounded in L^∞ by (2.51) and (2.53), we conclude that g_k and ξ_k are uniformly bounded in $W^{2,2}$. Hence, after taking a subsequence, there exists $\xi_0 \in M_R$ such that ξ_k converges to ξ_0 weakly in $W^{2,2}$ and strongly in $W^{1,p}$ for $2 < p < \infty$. From this follows

$$\lim_{k\to\infty} \langle *F_{e^{-\mathbf{i}t\xi_k}A}, -\xi_k\rangle = \langle *F_{e^{-\mathbf{i}t\xi_0}}A, -\xi_0\rangle.$$

Hence $\lim_{k\to\infty} \Phi_A(e^{\mathbf{i}\xi_k}) = \Phi_A(e^{\mathbf{i}\xi_0})$ and ξ_0 satisfies (2.52). $\square$

Lemma 2.6.4. *Assume the setting of Theorem 2.6.2. Suppose that $Z_0(G)$ is discrete, $\tau = 0$ and $w_0(A,\xi) > 0$ for all nonzero $\xi \in W^{1,2}(\Sigma, ad(P))$. Let $R > 0$ be given such that M_R is nonempty. Then exist constants $c_1, c_2 > 0$ such that (2.51) is satisfied.*

Proof. The proof consists of several steps.

Step 1: *There exists $C > 0$ such that*

$$||\xi||_{C^0} \le C\,(||\xi||_{L^1} + 1) \qquad \textit{for all } \xi \in M_R.$$

We observe that

$$\begin{aligned} 2\langle *F_{e^{\mathbf{i}\xi}A} - *F_A, \xi\rangle &= 2\int_0^1 \langle \Delta_{e^{\mathbf{i}t\xi}A}\xi, \xi\rangle\, dt = \Delta||\xi||^2 + 2\int_0^1 ||d_{e^{\mathbf{i}t\xi}A}\xi||^2\, dt \\ &\ge \Delta||\xi||^2 \ge 2||\xi||\Delta||\xi|| \end{aligned}$$

and hence

$$\Delta||\xi|| \le ||*F_{e^{\mathbf{i}\xi}A} - *F_A||. \tag{2.54}$$

An argument due to Simpson ([102], Prop 2.1) shows that this implies the claim. For this denote

$$f : \Sigma \to \mathbb{R}, \qquad f(z) := ||\xi(z)||.$$

For $z_0 \in \Sigma$ choose a local coordinate which identifies z_0 with the origin in $\mathbb{C}$. Let $B_{r_0}(0)$ be a ball contained in the image of this local coordinate and let $r \in (0, r_0)$. Let w, h be solutions of

$$\Delta w = ||*F_{e^{\mathbf{i}\xi}A} - *F_A||,\ w|_{\partial B_r(0)} = 0 \qquad \Delta h = 0,\ h|_{\partial B_r(0)} = f|_{B_r(0)}.$$

Here we consider the Laplacian of Σ which agrees with the Laplacian on $\mathbb{C}$ up to a positive factor. Hence (2.54) and the maximum principle show that $f - w - h \leq 0$ and the mean value theorem yields

$$f(0) - w(0) \leq h(0) = \frac{1}{2\pi r} \int_{\partial B_r(0)} f.$$

Moreover, by definition of M_R and elliptic regularity follows

$$|w(0)| \leq C||w||_{W^{2,2}} \leq C||\Delta w||_{L^2} \leq C(|| * F_A||_{L^2} + R).$$

Hence

$$f(z_0) \leq C\left(r + \frac{1}{r}\int_{\partial B_r(0)} f\right).$$

Now choose $r \in (r_0/2, r_0)$ such that $\frac{r_0}{2}\int_{\partial B_r(0)} f \leq ||f||_{L^1}$ holds. Then follows

$$f(z_0) \leq C\left(r_0 + \frac{1}{r_0^2}||f||_{L^1}\right).$$

Since Σ is compact, we can perform this argument within finitely many charts and choose the final constant C to be independent of z_0.

Step 2: *There exist $c_1, c_2 > 0$ such that*

$$||\xi||_{L^1} \leq c_1 \Phi_A(e^{\mathbf{i}\xi}) + c_2 \qquad \text{for all } \xi \in M_R.$$

Suppose the claim is false. Then exists $C_k > 0$ and $\xi_k \in M_R$ such that

$$\lim_{k\to\infty} C_k = \infty, \quad \lim_{k\to\infty} ||\xi_k||_{L^1} = \infty \quad \text{and} \quad ||\xi_k||_{L^1} \geq C_k \Phi_A(e^{\mathbf{i}\xi_k}).$$

It follows from Step 2 that $\eta_k := -\xi_k/||\xi_k||_{L^1}$ is uniformly bounded in L^∞. Denote $\ell_k := ||\xi_k||_{L^2}$. Then

$$\frac{1}{C_k} \geq \frac{\Phi_A(e^{\mathbf{i}\xi_k})}{||\xi_k||_{L^1}} = \int_0^1 \langle *F_{e^{\mathbf{i}t\xi_k}A}, \eta_k\rangle\, dt = \frac{1}{\ell_k}\int_0^{\ell_k} \langle *F_{e^{\mathbf{i}t\eta_k}A}, \eta_k\rangle\, dt$$

The integrand is increasing by (2.31). Hence, for any fixed $t > 0$ follows

$$\frac{1}{C_k} \geq \frac{\ell_k - t}{\ell_k}\langle *F_{e^{\mathbf{i}t\eta_k}}, \eta_k\rangle + \frac{t}{\ell_k}\langle F_A, \eta_k\rangle. \tag{2.55}$$

It follows from (2.33) that

$$\langle *F_{e^{\mathbf{i}t\eta_k}A}, \eta_k\rangle = \langle *F_A, \eta_k\rangle + \frac{1}{2}\int_0^t ||e^{i\eta_k s}(\bar\partial_A \eta_k)e^{-\mathbf{i}\eta_k s}||_{L^2}^2\, ds$$

and, since η_k is uniformly bounded in C^0, we conclude that $\bar\partial_A\eta_k$ is uniformly bounded in L^2. Since A is irreducible and $||\bar\partial_A\eta_k||^2 = \frac{1}{2}||d_A\eta_k||^2$ this shows that η_k is uniformly bounded in $W^{1,2}$. Hence, after taking a subsequence, there exists $\eta \in W^{1,2} \cap L^\infty$ such that $\eta_k \to \eta$ converges weakly in $W^{1,2}$ and strongly in L^p for every $1 \le p < \infty$. In particular $||\eta||_{L^1} = 1$ shows that $\eta \neq 0$ and

$$\lim_{k\to\infty} \langle *F_{e^{\mathrm{i}t\eta_k}A}, \eta_k \rangle = \langle *F_{e^{\mathrm{i}t\eta}A}, \eta \rangle.$$

Now (2.55) implies $\langle *F_{e^{\mathrm{i}t\eta}A}, \eta\rangle \le 0$ and as $t \to \infty$ we obtain $w_0(A, \eta) \le 0$. This contradicts our assumptions and proves Step 2.

Step 3: *There exist $c_1, c_2 > 0$ such that*

$$||\xi||_{L^\infty} \le c_1 \Phi_A(e^{i\xi}) + c_2 \qquad \textit{for all } \xi \in M_R.$$

This follows directly from Step 1 and Step 2.

□

Proof of Theorem 2.6.2. Suppose A is μ_τ-stable. Then $Z_0(G)$ is discrete, $\tau = 0$ and $\mathcal{G}^c(A)$ has discrete $\mathcal{G}^c(P)$ isotropy. We claim that $w_0(A, \xi) > 0$ for all $\xi \in W^{1,2}(\Sigma, \mathrm{ad}(P))$. By Proposition 2.5.9 this condition is equivalent to the stability of the induced holomorphic structure J_A on $P^c := P \times_G G^c$. In particular, this condition is invariant under the action of $\mathcal{G}^c(P)$ and we may assume that $F_A = 0$. Then (2.32) shows

$$w_0(A, \xi) = \int_0^\infty ||d_{e^{\mathrm{i}t\xi}A}\xi||_{L^2}^2 \, dt > 0$$

since A is irreducible. Thus Lemma 2.6.4 applies and shows that the estimate (2.51) is satisfied. The converse direction follows from Lemma 2.6.3.

The characterization of the μ_τ-polystable case follows directly from (2.49).

In the following let $A(t)$ denote the solution of the Yang-Mills flow (2.22) starting at A and let $A_\infty := \lim_{t\to\infty} A(t)$. Suppose A is μ_τ-unstable. Theorem 2.4.14 and (2.19) show that

$$||F_{gA} - \tau||_{L^2} \ge ||F_{A_\infty} - \tau||_{L^2} = c > 0$$

for all $g \in \mathcal{G}^c(P)$. Now define $g(t)$ by (2.24). Then $A(t) = g(t)^{-1}(A)$ and

$$\begin{aligned} \frac{d}{dt}\Phi_A(g(t)) &= \alpha_A(g(t), \dot{g}(t)) = -\langle *F_{A(t)} - \tau, *F_{A(t)} \rangle \\ &= -|| *F_{A(t)} - \tau||_{L^2}^2 \le -c \end{aligned}$$

where the penultimate step follows from (2.12). This shows that Φ_A is unbounded below.

Suppose conversely that A is μ_τ-semistable. It follows from Theorem 2.4.14 and (2.19) that A_∞ is a global minimum for the Yang-Mills functional on $\mathcal{A}(P)$ and $*F_{A_\infty} = \tau$. It follows from the Łojasiewicz inequality (Lemma 2.4.5) that there exists $\gamma \in [\frac{1}{2}, 1)$ and $C, T > 0$ such that

$$\begin{aligned}
|| *F_{A(t)} - \tau||_{L^2}^2 &= 2|\mathcal{YM}(A(t)) - \mathcal{YM}(A_\infty)| \\
&\le C||d_{A(t)}^* F_{A(t)}||_{L^2}^{\frac{1}{\gamma}} \\
&\le C||d_{A(t)}^* F_{A(t)}||_{L^2}^2 (\mathcal{YM}(A(t)) - \mathcal{YM}(A_\infty))^{1-2\gamma} \\
&= \frac{d}{dt} C\left(\mathcal{YM}(A(t)) - \mathcal{YM}(A_\infty)\right)^{2-2\gamma}
\end{aligned}$$

for all $t > T$. Since the right hand side is integrable, the solution $g(t)$ of (2.24) satisfies

$$\lim_{t\to\infty} \Phi_A(g(t)) = -\int_0^\infty || *F_{A(t)} - \tau||_{L^2}^2 \, dt =: a > -\infty.$$

We claim that a is a global minimum for Φ_A. For this let $\tilde{g}_0 \in \mathcal{G}^c(P)$ and let $\tilde{g}(t)$ be the solution of (2.26) starting at $\tilde{g}_0$. This is a negative gradient flow line of Φ_A and satisfies

$$\frac{d}{dt}\Phi_A(\tilde{g}(t)) = -\alpha_A(\tilde{g}(t), \dot{\tilde{g}}(t)) = -|| *F_{\tilde{g}(t)^{-1}A} - \tau||_{L^2}^2 \le 0.$$

Define $\eta(t) \in W^{2,2}(\Sigma, \mathrm{ad}(P))$ and $u(t) \in \mathcal{G}(P)$ by the equation

$$g(t)\exp(\mathbf{i}\eta(t))u(t) = \tilde{g}(t)$$

and

$$\beta_t : [0, 1] \to \mathcal{G}^c(P), \qquad \beta_t(s) = g(t)e^{\mathbf{i}s\eta(t)}.$$

Then $(\Phi_A \circ \beta_t)$ satisfies

$$\begin{aligned}
\left.\frac{d}{ds}\right|_{s=0} (\Phi_A \circ \beta_t)(s) &= \alpha_A(g(t), \partial_s\beta_t(s)) = -\langle *F_{g(t)^{-1}A} - \tau, \eta\rangle \\
&\ge -|| *F_{g(t)^{-1}A} - \tau||_{L^2} \cdot ||\eta(t)||_{L^2}
\end{aligned}$$

and

$$\begin{aligned}
\frac{d^2}{ds^2}(\Phi_A \circ \beta_t)(s) &= -\frac{d}{ds}\langle *F_{e^{-\mathbf{i}\eta(t)s}g(t)^{-1}A} - \tau, \eta(t)\rangle \\
&= \langle d_{A_{s,t}}^* d_{A_{s,t}}\eta(t), \eta(t)\rangle = ||d_{A_{s,t}}\eta(t)||_{L^2}^2 \ge 0
\end{aligned}$$

where we abbreviated $A_{s,t} := e^{-\mathbf{i}\eta(t)s}g(t)^{-1}A$. In particular, $\Phi_A \circ \beta_t$ is convex and since $\eta(t)$ is uniformly bounded in L^∞ by Proposition 2.4.12 there exists a constant $C > 0$ such that

$$\Phi_A(\tilde{g}(t)) \ge \Phi_A(g(t)) - C||F_{g(t)^{-1}A} - \tau||_{L^2}.$$

Since $\Phi_A(\tilde{g}_0) \ge \Phi_A(\tilde{g}(t))$ for all t and the right hand side converges to a as $t \to \infty$ we conclude $\Phi_A(\tilde{g}_0) \ge a$. This establishes the claim and completes the proof of the theorem. $\square$

2.6.2 The Narasimhan-Seshadri-Ramanathan theorem

The Narasimhan-Seshadri-Ramanathan theorem relates the notion of stable objects in Definition 2.3.2 and Definition 2.3.7. This was first proven by Narasimhan-Seshadri [91] in the case $G = U(n)$ and later extended by Ramanathan [95] to general compact connected Lie groups. Both of these proofs are entirely of algebraic geometric nature.

In the case $G = U(n)$ Donaldson [30] gave an analytic proof of this result. His argument uses the moment weight inequality and an induction argument which is based on the Harder-Narasimhan filtration. We present a different proof which is due to Bradlow [12] and Mundet [89]. The main step in their proof consists of establishing the stable case in Theorem 2.6.2.

Theorem 2.6.5 (Narasimhan-Seshadri-Ramanathan). *Let G be a compact connected Lie group and $P \to \Sigma$ a principal G bundle with central type $\tau \in Z(\mathfrak{g})$ defined by (2.12). Let $A \in \mathcal{A}(P)$ and consider the complexified bundle $P^c := P \times_G G^c$ with the holomorphic structure induced by A. Then (P^c, J_A) is stable if and only if there exists a complex gauge transformation $g \in \mathcal{G}^c(P)$ such that $*F_{gA} = \tau$ and the kernel of*

$$L_A : W^{2,2}(\Sigma, ad(P^c)) \to W^{1,2}(\Sigma, T^*\Sigma \otimes ad(P))$$

$$L_A(\xi + \boldsymbol{i}\eta) = -d_A\xi + *d_A\eta$$

contains only constant central sections.

Proof. We may assume by Lemma 2.3.5 and Lemma 2.3.9 that $Z_0(G)$ is discrete and $\tau = 0$.

Suppose there exists $g \in \mathcal{G}^c(P)$ such that $*F_{gA} = 0$ and gA is irreducible. Then (2.32) shows

$$w_0(gA, \xi) = \int_0^\infty ||d_{e^{\mathrm{i}t\xi}gA}\xi||_{L^2}^2 \, dt > 0$$

for all $0 \neq \xi \in W^{1,2}(\Sigma, \mathrm{ad}(P))$ and by Proposition 2.5.9 (P^c, J_{gA}) is stable. Since the notion of stability is $\mathcal{G}^c(P)$ invariant, (P^c, J_A) is stable.

Assume conversely that (P^c, J_A) is stable. For every $g \in \mathcal{G}^c(P)$ then (P^c, J_{gA}) is stable as well and Proposition 2.5.9 implies $w_0(gA, \xi) > 0$ for every nonzero $\xi \in W^{1,2}(\Sigma, \mathrm{ad}(P))$. In particular, gA is irreducible and Lemma 2.6.4 is applicable and shows that A is μ_0-stable. □

2.7 The dominant weight theorem

The dominant weight theorem strengthens the moment weight inequality (Theorem 2.5.12). It shows that there exists (up to scaling) a unique section $\xi \in \Omega^0(\Sigma, \mathrm{ad}(P))$ which yields equality in the moment weight inequality, whenever the right hand side is positive. In particular, it relates the notion of unstable objects in Definition 2.3.2 and Definition 2.3.7. A key ingredient in its proof is the Harder-Narasimhan filtration associated to a holomorphic holomorphic vector bundle. We review this first before we proceed to the proof of the dominant weight theorem.

Theorem 2.7.1 (The dominant weight theorem). *Let G be a compact connected Lie group, let $P \to \Sigma$ be a principal G bundle of central type $\tau \in Z(\mathfrak{g})$ defined by (2.12) and let $A \in \mathcal{A}(P)$ be a smooth μ_τ-unstable connection.*

1. *There exists an element $\hat{\xi} \in \Omega^0(\Sigma, ad(P))$ such that*

$$\sup_{0 \neq \xi \in \Omega^0(\Sigma, ad(P))} -\frac{w_\tau(A,\xi)}{||\xi||_{L^2}} = -\frac{w_\tau(A,\hat{\xi})}{||\hat{\xi}||_{L^2}} = \inf_{g \in \mathcal{G}^c(P)} || * F_{gA} - \tau||_{L^2}. \tag{2.56}$$

2. *The normalized section $\hat{\xi}/||\hat{\xi}||_{L^2}$ is uniquely determined. Moreover, it is rational in the sense that it generates a closed $\mathbb{C}^*$ subgroup of $\mathcal{G}^c(P)$.*

3. *If A_∞ is the limit of the Yang-Mills flow (2.22) starting at A, then there exists $u \in \mathcal{G}(P)$ such that $\hat{\xi} = u(*F_{A_\infty} - \tau)u^{-1}$ satisfies (2.56).*

Proof. This result is essentially contained in the work of Atiyah and Bott. They determine in ([4], Prop. 8.13 and Prop. 10.13) the infimum of the Yang-Mills functional on the complexified orbit $\mathcal{G}^c(A)$ in terms of the Harder-Narasimhan filtration of $\mathrm{ad}(P^c)$.

Bruasse and Teleman [15, 14] show in a more general gauge theoretical setting that the supremum over the normalized weights is attained in a unique direction whenever it is positive. This corresponds to the case where (P, J_A) is unstable and they identify again the dominant weight with the Harder-Narasimhan filtration.

We follow these ideas in our proof below, but simplify the arguments considerably by using the moment weight inequality and the analytic properties of the Yang-Mills flow. The proof will be given on page 94. □

2.7.1 The Harder-Narasimhan filtration

Let F and G be holomorphic vector bundles over a Riemann surface Σ and let $\alpha : F \to G$ be a holomorphic bundle map. The kernel and cokernel of α are in general not well-defined as holomorphic vector bundles and one may think of them as vector bundles with singularities. These considerations lead naturally to the larger category of coherent analytic sheaves on Σ which is closed under taking kernels and cokernels. The next lemma, however, allows us to get away without considering sheaves.

Lemma 2.7.2. *Let F and G be holomorphic vector bundles over a Riemann surface Σ and let $\alpha : F \to G$ be a nonzero holomorphic bundle map. Then there exists a commutative diagram of holomorphic vector bundles and holomorphic bundle maps*

$$\begin{tikzcd}
0 \arrow[r] & F' \arrow[r] & F \arrow[r] \arrow[d, "\alpha"] & F'' \arrow[r] \arrow[d, "\beta"] & 0 \\
0 & G'' \arrow[l] & G \arrow[l] & G' \arrow[l] & 0 \arrow[l]
\end{tikzcd}$$

with exact rows and $rk(F'') = rk(G')$, $\det(\beta) \neq 0$ and $c_1(F'') \leq c_1(G')$.

Proof. This lemma is most easily understood in the language of analytic sheaves. Denote by $\mathcal{O}$ the sheaf of germs of holomorphic functions on Σ. There exists a one to one correspondence between holomorphic vector bundles and locally free $\mathcal{O}$-sheafs on Σ, which associates to a vector bundle its sheaf of holomorphic sections. The homomorphism α induces a homomorphism between the associated sheaves and the sheaf kernel and sheaf image are clearly torsion free subsheaves. Since the stalks of $\mathcal{O}$ are isomorphic to the principal ideal domain $\mathbb{C}[[z]]$, these sheaves are locally free and correspond to the the vector bundles F' and G'. □

Recall that we denote for a complex vector bundle $E \to \Sigma$ by

$$\mu(E) := \frac{c_1(E)}{\mathrm{rk}(E)}$$

its slope or normalized Chern-class.

Corollary 2.7.3. *Let F and G be holomorphic vector bundles over Σ.*

1. *Suppose F is semistable, G is stable and $\mu(F) = \mu(G)$. Then any nonzero holomorphic bundle map $\alpha : F \to G$ is surjective.*

2. *Suppose F and G are stable and $\mu(F) = \mu(G)$. Then any nonzero holomorphic bundle map $\alpha : F \to G$ is an isomorphism.*

3. *Suppose F and G are semistable and $\mu(F) > \mu(G)$. Then every holomorphic bundle map $\alpha : F \to G$ vanishes.*

Proof. We prove the first part. Suppose $\alpha : F \to G$ is neither zero nor surjective. Using the notation of Lemma 2.7.2 we see that G' is a proper subbundle and thus

$$\mu(G) > \mu(G') \geq \mu(F'') \geq \mu(F)$$

contradicting the assumption $\mu(G) = \mu(F)$. In other two parts follow from a similar argument. □

Lemma 2.7.4. *Let E be a holomorphic semistable vector bundle. Then there exists a filtration*

$$0 < E_1 < E_2 < \cdots < E_r = E$$

such that each quotient E_j/E_{j-1} is stable and $\mu(E_j/E_{j-1}) = \mu(E)$.

Proof. Let $F \subset E$ be a stable subbundle with $\mu(F) = \mu(E)$. Since $E \cong F \oplus (E/F)$ as C^∞-bundles, it follows $\mu(E/F) = \mu(E)$. Moreover, any holomorphic subbundle $G \subset E/F$ with $\mu(G) > \mu(E)$ would lift under the projection map $E \to E/F$ to a holomorphic subbundle $\tilde{G} \subset E$ with $\mu(\tilde{G}) > \mu(E)$ and this contradicts the semistability of E. Hence E/F is semistable and the lemma follows by induction. □

The Harder-Narasimhan filtration generalizes Lemma 2.7.4 to general holomorphic vector bundles.

Proposition 2.7.5 (Harder-Narasimhan filtration). *Let E be an holomorphic vector bundle. Then there exists a unique holomorphic filtration*

$$0 = E_0 < E_1 < \cdots < E_r = E$$

such that all quotients E_i/E_{i-1} are semistable and the slopes

$$\mu_j := \frac{c_1(E_j/E_{j-1})}{rk(E_j/E_{j-1})}$$

satisfy $\mu_1 > \mu_2 > \cdots > \mu_r$.

Proof. The degree of any holomorphic subbundles of E is uniformly bounded by Lemma 2.7.6 below. Let $E_1 \subset E$ be a semistable subbundle for which $\mu(E_1) =: \mu_1$ is maximal and such that E_1 has maximal rank among all such subbundles. We claim that every proper holomorphic subbundle $G' \subset E/E_1$ satisfies $\mu(G') < \mu_1$. Otherwise, the preimage of G' under the projection $E \to E/E_1$ would be a subbundle $\tilde{G} \subset E$ with $\mu(\tilde{G}) \geq \mu_1$ and of strictly greater rank then E_1. This proves the claim and the existence of the Harder-Narasimhan filtration follows by induction.

Let $0 = \tilde{E}_0 < \tilde{E}_1 < \cdots < \tilde{E}_\ell = E$ be another filtration of E such that all quotients $\tilde{E}_j/\tilde{E}_{j-1}$ are semistable and the slopes $\tilde{\mu}_j := \mu(\tilde{E}_j/\tilde{E}_{j-1})$ are strictly decreasing. In particular, $\tilde{E}_1$ is semistable and the construction above shows

$$\mu(E_1) \geq \mu(\tilde{E}_1) = \tilde{\mu}_1 > \tilde{\mu}_2 > \cdots > \tilde{\mu}_\ell.$$

The last part of Corollary 2.7.3 shows that the projection $E_1 \to E/\tilde{E}_{\ell-1}$ must be zero, since $\mu(E_1) > \tilde{\mu}_\ell$ and hence $E_1 \subset \tilde{E}_{\ell-1}$. Repeating the argument, it follows by induction that $E_1 \subset \tilde{E}_j$ for all $j \geq 1$. If $\mu(E_1) > \mu(\tilde{E}_1)$, we could go one step further and obtain the contradiction $E_1 \subset \tilde{E}_0 = 0$. This shows $\mu_1 = \tilde{\mu}_1$. Finally, consider the projection

$$\alpha : E_1 \to E \to E/\tilde{E}_1.$$

If it is nonzero, we can apply Lemma 2.7.2 with $F = E_1$ and $G = E/\tilde{E}_1$ to obtain the contradiction

$$\mu_1 = \mu(E_1) \leq \mu(F'') \leq \mu(G') \leq \tilde{\mu}_2 < \tilde{\mu}_1 = \mu_1.$$

This shows $E_1 \subset \tilde{E}_1$ and by maximality of $\mathrm{rk}(E_1)$ equality must hold. The uniqueness of the Harder-Narasimhan filtration follows now by induction. □

Lemma 2.7.6. *Let (E, h) be a hermitian holomorphic vector bundle over Σ and denote by $A \in \mathcal{A}(E)$ the associated unitary connection from Lemma 2.2.7. For a holomorphic subbundle $F \subset E$ the following holds:*

1. *Let $E = F \oplus G$ be an orthogonal decomposition and identify G with E/F. Denote by A_F and A_G the induced connections on F and G. Then A has the shape*

$$A = \begin{pmatrix} A_F & \eta \\ -\eta^* & A_G \end{pmatrix}$$

with $\eta \in \Omega^{0,1}(\Sigma, End(G, F))$. Moreover, the curvature has the shape

$$F_A = \begin{pmatrix} F_{A_F} - \eta \wedge \eta^* & d_A \eta \\ -d_A \eta^* & F_{A_G} - \eta^* \wedge \eta \end{pmatrix}.$$

2. *There exists a constant $C > 0$, which does not depend on F, such that*

$$c_1(F) \leq C(1 - ||\eta||_{L^2}^2).$$

Proof. We leave the first part as an exercise to the reader, see e.g. [53] Chapter 0.5. For the second part, we calculate

$$c_1(F) = \frac{\mathbf{i}}{2\pi}\int_\Sigma \mathrm{tr}(F_{A_F}) = \frac{\mathbf{i}}{2\pi}\int_\Sigma \mathrm{tr}\left(F_A|_F\right) + \mathrm{tr}\left(\eta \wedge \eta^*\right).$$

In local coordinates write $\eta = \tilde{\eta} d\bar{z}$ and hence $\eta \wedge \eta^* = 2\mathbf{i}\tilde{\eta}\tilde{\eta}^* dx \wedge dy$. This yields precisely the L^2-norm of η. Since F_A is uniformly bounded in L^∞, the estimate follows. □

We show next that the Harder-Narasimhan filtration is maximal among all holomorphic filtrations in a certain sense. For this we need to introduce some notation. Let

$$\mathcal{E} : 0 = E_0 < E_1 < \cdots < E_r = E$$

be a holomorphic filtration of E. Denote $n_j := \mathrm{rk}(E_j/E_{j-1})$, $k_j := c_1(E_j/E_{j-1})$ and define the characteristic vector of the filtration $\mathcal{E}$ to be

$$\vec{\mu}(\mathcal{E}) = \left(\frac{k_1}{n_1}, \ldots, \frac{k_1}{n_1}, \ldots, \frac{k_r}{n_r}, \ldots, \frac{k_r}{n_r}\right) \in \mathbb{R}^n \tag{2.57}$$

where we repeat each entry k_j/n_j exactly n_j-times. Moreover define

$$\ell_\mathcal{E} : \{0, \ldots, n\} \to \mathbb{R}, \qquad \ell_\mathcal{E}(m) = \sum_{j=1}^m \left[\vec{\mu}(\mathcal{E})\right]_j$$

where $\left[\vec{\mu}(\mathcal{E})\right]_j$ denotes the j-th entry of the vector $\vec{\mu}(\mathcal{E})$. The graph of $\ell_\mathcal{E}$ interpolates linearly between the points $(0,0)$, (n_1, k_1), $(n_1 + n_2, k_1 + k_2)$, …, (n, k). We consider the following ordering on the space of holomorphic filtrations:

$$\mathcal{E} \geq \mathcal{F} \quad \text{if and only if} \quad \ell_\mathcal{E} \geq \ell_\mathcal{F}.$$

We call a filtration $\mathcal{E}$ concave if the function $\ell_\mathcal{E}$ is concave, or equivalently, if the entries of $\vec{\mu}(\mathcal{E})$ are decreasing.

Proposition 2.7.7. *Let E be a holomorphic vector bundle over Σ. The Harder-Narasimhan filtration of E is the unqiue maximal concave filtration on E.*

Proof. Let

$$\mathcal{E}_{HN} : \qquad 0 < E_1 < E_2 < \cdots < E_r = E$$

be the Harder-Narasimhan filtration of E and let $F < E$ be a holomorphic subbundle. It suffices to prove that the point $p_F := (\mathrm{rk}(F), c_1(F))$ lies on or below the graph of $\ell_\mathcal{E}$. We prove this by induction on r.

Suppose $r = 1$. Then $\mathcal{E}$ is semistable and $\mu(E) \geq \mu(F)$. In particular, $\ell_\mathcal{E}$ is a straight line of slope $\mu(E)$ and p_F clearly lies below that line.

Suppose now $r > 1$. The Harder-Narasimhan filtration of E/E_1 is given by

$$\mathcal{E}'_{HN}: \qquad 0 < E_2/E_1 < E_3/E_1 < \cdots < E_r/E_1 = E/E_1$$

and the induction hypothesis applies to $\mathcal{E}'_{HN}$. Consider the commutative diagram from Lemma 2.7.2

$$\begin{tikzcd}
0 \arrow[r] & F' \arrow[r] & F \arrow[r] \arrow[d, "\alpha"] & F'' \arrow[r] \arrow[d, "\beta"] & 0 \\
0 & G'' \arrow[l] & E/E_1 \arrow[l] & G' \arrow[l] & 0 \arrow[l]
\end{tikzcd}$$

with $\alpha : F \to E \to E/E_1$. By the induction hypothesis, the point of $(\mathrm{rk}(G'), c_1(G'))$ lies below $\ell_{\mathcal{E}'}$. Since $\mathrm{rk}(F'') = \mathrm{rk}(G')$ and $c_1(F'') \leq c_1(G')$ the same holds with G' replaced by F''. This shows

$$c_1(E_1) + c_1(F'') \leq \ell_{\mathcal{E}}(\mathrm{rk}(E_1) + \mathrm{rk}(F'')). \tag{2.58}$$

Since F' gets maped to zero under α, we have $F' \subset E_1$ and $\mu(F') \leq \mu(E_1)$ by semistability of E_1. This shows $c_1(F') \leq \ell_{\mathcal{E}}(\mathrm{rk}(F'))$ and with (2.58) follows

$$c_1(F) = c_1(F') + c_1(F'') \leq \ell_{\mathcal{E}}(\mathrm{rk}(E_1) + \mathrm{rk}(F'')) + \ell_{\mathcal{E}}(\mathrm{rk}(F')) - \ell_{\mathcal{E}}(\mathrm{rk}(E_1)).$$

Since $\ell_{\mathcal{E}}$ is concave and $\mathrm{rk}(F') \leq \mathrm{rk}(E_1)$ we have

$$\ell_{\mathcal{E}}(\mathrm{rk}(E_1) + \mathrm{rk}(F'')) - \ell_{\mathcal{E}}(\mathrm{rk}(E_1)) \leq \ell_{\mathcal{E}}(\mathrm{rk}(F') + \mathrm{rk}(F'')) - \ell_{\mathcal{E}}(\mathrm{rk}(F'))$$

and thus

$$c_1(F) \leq \ell_{\mathcal{E}}(\mathrm{rk}(F') + \mathrm{rk}(F'')) = \ell_{\mathcal{E}}(\mathrm{rk}(F)).$$

This completes the proof. □

Corollary 2.7.8. *Let E be a holomorphic vector bundle over Σ. Let $\mathcal{E}$ be a concave filtration of E and $\mathcal{E}_{HN}$ the Harder-Narasimhan filtration of E. Then follows*

$$||\vec{\mu}(\mathcal{E})||_2 \leq ||\vec{\mu}(\mathcal{E}_{HN})||_2$$

where $||\cdot||_2$ denotes the standard euclidean norm on $\mathbb{R}^n$. Moreover, equality holds if and only if $\mathcal{E} = \mathcal{E}_{HN}$.

Proof. An easy calculation shows that for two concave filtrations with $\mathcal{E}_1 \leq \mathcal{E}_2$ the estimate $||\vec{\mu}(\mathcal{E}_1)||_2 \leq ||\vec{\mu}(\mathcal{E}_2)||_2$ is satisfied. Moreover, equality holds if and only if $\mathcal{E}_1 = \mathcal{E}_2$. □

2.7.2 Proof of the dominant weight theorem

We proceed now to the proof of Theorem 2.7.1. We consider first the case $G = U(n)$ and deduce the general case afterwards by choosing a faithful representation $G \hookrightarrow U(n)$.

μ_τ-unstable orbits in the unitary case.

Assume $G = U(n)$ and denote by $E := P \times_G \mathbb{C}^n$ the associated hermitian vector bundle. Note that the constant central type τ of P is related to the slope of E by the formula

$$\tau = -2\pi\mathbf{i}\mu(E) \cdot \mathbb{1}. \tag{2.59}$$

If $(E, \bar{\partial}_A)$ is unstable, then Proposition 2.5.9 implies that there exists a negative weight $w_\tau(A,\xi) < 0$ and the moment weight inequality (Theorem 2.5.12) shows that A is μ_τ-unstable. The following lemma proves the converse direction.

Lemma 2.7.9. *Let $A \in \mathcal{A}(E)$ be a unitary connection and suppose $(E, \bar{\partial}_A)$ is a semistable holomorphic vector bundle. Then the limit A_∞ of the Yang-Mills flow $A(t)$ starting at A satisfies*

$$*F_{A_\infty} = -2\pi\boldsymbol{i}\mu(E) \cdot \mathbb{1}.$$

Proof. We show first that the $W^{1,2}$-closure $\overline{\mathcal{G}^c(A)}$ contains a connection $\bar{A}$ with $F_{\bar{A}} = -2\pi\mathbf{i}\mu(E) \cdot \mathbb{1}$. For this, consider the refined Harder-Narasimhan filtration from Lemma 2.7.4

$$0 < E_1 < E_2 < \cdots < E_r = E$$

with stable quotients E_j/E_{j-1} all having the same slope as E. Choose an orthogonal splitting $E = D_1 \oplus \cdots \oplus D_r$ such that $E_j = D_1 \oplus \cdots \oplus D_j$. With respect to this splitting $\bar{\partial}_A$ has the shape

$$\bar{\partial}_A = \begin{pmatrix} \bar{\partial}_{A_1} & A_{12} & \dots & A_{1r} \\ 0 & \bar{\partial}_{A_2} & \dots & A_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \bar{\partial}_{A_r} \end{pmatrix}.$$

Define $g_t := \operatorname{diag}(t^{-1}, t^{-2}, \dots, t^{-r})$. Then

$$\bar{\partial}_{g_t(A)} = \begin{pmatrix} \bar{\partial}_{A_1} & tA_{12} & \dots & t^{r-1}A_{1r} \\ 0 & \bar{\partial}_{A_2} & \dots & t^{r-2}A_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \bar{\partial}_{A_r} \end{pmatrix} \to \begin{pmatrix} \bar{\partial}_{A_1} & 0 & \dots & 0 \\ 0 & \bar{\partial}_{A_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \bar{\partial}_{A_r} \end{pmatrix}$$

as $t \to 0$. Since $E_j/E_{j-1} \cong (D_j, \bar{\partial}_{A_j})$ are stable holomorphic vector bundles, Theorem 2.6.5 shows that there exist complex gauge transformations $g_j \in \mathcal{G}^c(D_j)$ such that $\bar{A}_j = g_j(A_j)$ satisfies $*F_{\bar{A}_j} = -2\pi\mathbf{i}\mu(D_j)$. Since $\mu(D_j) = \mu(E)$, we conclude that the induced connection $\bar{A} = \bar{A}_1 \oplus \cdots \oplus \bar{A}_r$ has curvature $F_{\bar{A}} = -2\pi\mathbf{i}\mu(E) \cdot \mathbb{1}$.

It follows from (2.19) that $\bar{A}$ minimizes the Yang-Mills functional over $\mathcal{A}_{U(n)}(E)$. The lemma follows thus from Theorem 2.4.14 and Theorem 2.4.15. □

Proof of Theorem 2.7.1 for $G = U(n)$.

Let ξ be a section of skew-hermitian endomorphism in $\mathfrak{u}(E) \subset \mathrm{End}(E)$ satisfying $||\xi|| = 1$ and

$$-w_\tau(A,\xi) = \sup_{0 \neq \eta \in \Omega^0(\Sigma,\mathfrak{u}(E))} -\frac{w_\tau(A,\eta)}{||\eta||}. \tag{2.60}$$

Proposition 2.5.2 shows that ξ determines a holomorphic filtration and orthogonal splitting

$$\mathcal{E}: \qquad E_1 < E_2 < \cdots < E, \qquad E_j = D_1 \oplus \cdots \oplus D_j$$

of $(E, \bar{\partial}_A)$. With respect to this orthogonal splitting ξ has the shape

$$\mathbf{i}\xi = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_r)$$

with $\lambda_1 < \lambda_2 < \cdots < \lambda_r$ and the weight is given by

$$w_\tau(A,\lambda) = 2\pi \sum_{j=1}^r \lambda_j \left(c_1(D_j) - \mathrm{rk}(D_j)\mu(E)\right).$$

By maximality of the weight $-w_\tau(A,\xi)$ we conclude that $\lambda = (\lambda_1, \ldots, \lambda_r)$ is a global minimum of the function

$$f(x_1, \ldots, x_r) = \sum_{j=1}^r x_j \left(c_1(D_j) - \mathrm{rk}(D_j)\mu(E)\right)$$

on the ellipsoid $\{\sum_{j=1}^r x_j^2 \mathrm{rk}(D_j) = 1\}$ under the open condition

$$x_1 < x_2 < \cdots < x_r.$$

Since $(E, \bar{\partial}_A)$ is unstable, Proposition 2.5.9 implies that this minimum is negative and f does not vanish identically. Thus ∇f vanishes nowhere and λ must lie on the ellipsoid. It satisfies there the Lagrange condition

$$\left(c_1(D_j) - \mathrm{rk}(D_j)\right)\mu(E) = c\lambda_j \mathrm{rk}(D_j)$$

for $j = 1, \ldots, r$ and some constant $c \neq 0$. Since $f(\lambda) < 0$ we must have $c < 0$. Since the λ_j are increasing this yields

$$\mu(D_1) > \mu(D_2) > \cdots > \mu(D_r)$$

and $\mathcal{E}$ is a concave filtration of E. Solving the Lagrange problem we get

$$\lambda_j = \frac{\mu(E) - \mu(D_j)}{\sqrt{\sum_{j=1}^r \mathrm{rk}(D_j)(\mu(D_j) - \mu(E))^2}} = \frac{\mu(E) - \mu(D_j)}{\sqrt{||\vec{\mu}(\mathcal{E})||_2^2 - \mathrm{rk}(E)\mu(E)^2}}$$

and

$$-w_\tau(A,\xi) = 2\pi\sqrt{||\vec{\mu}(\mathcal{E})||_2^2 - \mathrm{rk}(E)\mu(E)^2}. \tag{2.61}$$

Now Corollary 2.7.8 shows that $\mathcal{E} = \mathcal{E}_{HN}$ must agree with the Harder-Narasimhan filtration of E and ξ is uniquely determined.

Conversely, we can use the Harder-Narasimhan filtration to define ξ and the argument from above shows that it satisfies (2.60). It remains to show it also yields equality in the moment-weight inequality. It follows from the proof of Proposition 2.5.2 that the limit

$$A_+ := \lim_{t\to\infty} e^{\mathbf{i}t\xi} A$$

exists and splits as $A_+ = A_1 \oplus \cdots \oplus A_r$ with $A_j \in \mathcal{A}(D_j) \cong \mathcal{A}(E_j/E_{j-1})$. The Yang-Mills flow $A_+(t)$ starting at A_+ is the product of the Yang-Mills flow on each factor and clearly remains in the closure $\overline{\mathcal{G}^c(A)}$. It follows from Lemma 2.7.9 that the limit $A_\infty := \lim_{t\to\infty} A_+(t)$ of this flow satisfies

$$F_{A_\infty} = -2\pi\mathbf{i}\begin{pmatrix} \mu(D_1) & & & \\ & \mu(D_2) & & \\ & & \ddots & \\ & & & \mu(D_r)\end{pmatrix}.$$

Now (2.59) and (2.61) yield

$$\inf_{g\in\mathcal{G}^c} ||F_{gA} - \tau|| \le ||F_{A_\infty} - \tau|| = 2\pi\sqrt{\sum_{j=1}^r \mathrm{rk}(D_j)(\mu(E) - \mu(D_j))^2} = -w(A,\xi).$$

The converse inequality follows from the moment-weight inequality (Theorem 2.5.12) and this completes the proof in the unitary case.

Proof of Theorem 2.7.1 for general compact connected Lie groups G.

Let G be a compact connected Lie group. We show first that one restrict the argument to the case where $Z_0(G)$ is discrete. Recall that the Lie algebra of G decomposes as $\mathfrak{g} = Z(\mathfrak{g}) \oplus [\mathfrak{g},\mathfrak{g}]$. The center yields a trivial $Z(\mathfrak{g})$ subbundle $V \subset \mathrm{ad}(P)$ and its orthogonal complement has fiber $[\mathfrak{g},\mathfrak{g}]$ and is canonically isomorphic to $\mathrm{ad}(P/Z_0(G))$. This yields the orthogonal decomposition

$$\mathrm{ad}(P) \cong V \oplus \mathrm{ad}(P/Z_0(G)). \tag{2.62}$$

Let $A \in \mathcal{A}(P)$ and denote by $\bar{A} \in \mathcal{A}(P/Z_0(G))$ the induced connection. Decompose $\xi \in \Omega^0(\Sigma, \mathrm{ad}(P))$ as $\xi = \xi^z + \xi^{ss}$ with respect to the decomposition (2.62). Then (2.12) and Lemma 2.5.10 yields

$$w_\tau(A,\xi) = w_\tau(A,\xi^{ss}) = w_0(\bar{A},\xi^{ss}).$$

Decompose similarly $F_{gA} = F^z + F^{ss}$ and note that $F^{ss} = F_{g\bar{A}}$. This yields

$$|| * F_{gA} - \tau||^2 = || * F^{ss}||^2 + || * F^z - \tau||^2 \ge || * F_{g\bar{A}}||^2.$$

As in Lemma 2.3.9 one shows that g can be modified to a gauge transformation $\tilde{g}$ such that $g\bar{A} = \tilde{g}\bar{A}$ and $*F^z = \tau$. Hence

$$\inf_{g\in\mathcal{G}^c(P)} || * F_{gA} - \tau|| = \inf_{g\in\mathcal{G}^c(P/Z_0(G))} || * F_{g\bar{A}}||.$$

This completes the reduction argument.

Now assume that $Z_0(G)$ is discrete and $\tau = 0$. Choose a faithful representation $G \hookrightarrow U(n)$ and identify G with its image in $U(n)$. It follows from Lemma 2.5.8 and (2.59) that the associated vector bundle $E = P \times_G \mathbb{C}^n$ satisfies $\mu(E) = 0$. For $A \in \mathcal{A}(P)$ Theorem 2.4.14 yields

$$\inf_{g \in \mathcal{G}^c(P)} ||*F_{gA}|| = ||*F_{A_\infty}|| = \inf_{g \in \mathrm{GL}(n)} ||*F_{gA}||$$

where we consider A as G-connection for the left equality and as $U(n)$-connection for the right equality. It follows from the unitary case that there exists (up to scaling) a unique section $\xi \in \Omega^0(\Sigma, \mathfrak{u}(E))$ satisfying

$$-\frac{w_0(A,\xi)}{||\xi||} = \inf_{g \in \mathcal{G}^c} ||F_{gA}||.$$

Let $\tilde{\xi}$ be the orthogonal projection of ξ onto $\mathfrak{g}(E) \subset \mathfrak{u}(E)$. Then Lemma 2.5.8 shows $w_0(A,\xi) = w_0(A,\tilde{\xi})$ and hence

$$\inf_{g \in \mathcal{G}^c} ||F_{gA}|| = -\frac{w_0(A,\xi)}{||\xi||} \leq -\frac{w_0(A,\tilde{\xi})}{||\tilde{\xi}||}$$

with equality if and only if $\xi = \tilde{\xi}$. The moment weight inequality (Theorem 2.5.12) yields the converse inequality and this completes the proof.

Chapter 3

Convergence of the Yang–Mills–Higgs flow and applications

The content of this chapter has been published in [115]. The symplectic vortex equations [25, 26, 89] are an equivariant version of the J-holomorphic curves equation in symplectic geometry. These equations also generalize the Yang–Mills equations [4], the notion of Bradlow pairs [12] and are closely related to Hitchin's selfduality equations [58] and Higgs-bundles. The symplectic vortex equations admit a variational description as global minimum of the Yang–Mills–Higgs functional. We study its negative gradient flow on holomorphic pairs (A, u) where A is a connection on a principal G-bundle P over a closed Riemann surface Σ and $u : P \to X$ is an equivariant map into a Kähler Hamiltonian G-manifold. The connection A induces a holomorphic structure on the Kähler fibration $P \times_G X$ and we require that u descends to a holomorphic section of this fibration. We prove a Łojasiewicz type gradient inequality and show uniform convergence of the negative gradient flow in the $W^{1,2} \times W^{2,2}$-topology when X is equivariantly convex at infinity with proper moment map, X is holomorphically aspherical and its Kähler metric is analytic.

As applications we establish several results inspired by finite dimensional GIT: First, we prove a certain uniqueness property for the critical points of the Yang–Mills–Higgs functional which is the analogue of the Ness uniqueness theorem. Second, we extend Mundet's Kobayashi–Hitchin correspondence to the polystable and semistable case. The arguments for the polystable case lead to a new proof in the stable case. Third, in proving the semistable correspondence, we establish the moment-weight inequality for the vortex equation and prove the analogue of the Kempf existence and uniqueness theorem. Our proofs are inspired by the work of Calabi, Chen, Donaldson, Sun [17, 23, 18, 19, 39] on extremal Kähler metrics; see [51] for a finite dimensional discussion.

3.1 Introduction

Throughout this chapter we assume the following. G is a compact (real) Lie group with Lie algebra $\mathfrak{g}$ together with a fixed choice of an invariant inner product on $\mathfrak{g}$, Σ is a closed Riemann surface with fixed volume form $dvol_\Sigma$ and induced Riemannian metric, $P \to \Sigma$ is a principal G bundle and (X, J, ω) is a Kähler manifold equipped with a Hamiltonian G action induced by an equivariant moment map $\mu : X \to \mathfrak{g}$.

Geometric invariant theory for the vortex equation

Atiyah–Bott [4] observed that the curvature $F_A \in \Omega^2(\Sigma, \mathrm{ad}(P))$ defines a moment map for the action of the gauge group $\mathcal{G}(P)$ on the space of connections $\mathcal{A}(P)$ (see Lemma 2.2.2). The vortex equations are obtained as an extension of this picture. Consider the associated Kähler fibration

$$P(X) := P \times_G X := (P \times X)/G.$$

and denote by $\mathcal{S}(P, X)$ its space of sections. The symplectic vortex equations on pairs $(A, u) \in \mathcal{A}(P) \times \mathcal{S}(P, X)$ are given by

$$\bar{\partial}_A u = 0, \qquad *F_A + \mu(u) = 0. \tag{3.1}$$

The connection $A \in \mathcal{A}(P)$ induces a holomorphic structure on the total space of the Kähler fibration $P(X)$ and the equation $\bar{\partial}_A u = 0$ requires u to be a holomorphic section. The subspace

$$\mathcal{H}(P, X) := \{(A, u) \in \mathcal{A}(P) \times \mathcal{S}(P, X) \,|\, \bar{\partial}_A u = 0\}$$

is formally a Kähler submanifold of $\mathcal{A}(P) \times \mathcal{S}(P, X)$. It is well known that

$$\Phi : \mathcal{A}(P) \times \mathcal{S}(P, X) \to \Omega^0(\Sigma, \mathrm{ad}(P)), \qquad \Phi(A, u) := *F_A + \mu(u) \tag{3.2}$$

provides a moment map for the $\mathcal{G}(P)$-action on $\mathcal{H}(P, X)$ (see Lemma 3.2.1) and solutions of (3.1) give rise to the symplectic moduli space

$$\mathcal{M}_{symp}(P, X) := \{(A, u) \in \mathcal{H}(P, X) \,|\, *F_A + \mu(u) = 0\} / \mathcal{G}(P).$$

This moduli space admits an alternative description as complex GIT quotient of $\mathcal{H}(P, X)$. For this let G^c be the complexification of G, let $P^c := P \times_G G^c$ be the complexification of P and define the complexified gauge group as $\mathcal{G}^c(P) := \mathcal{G}(P^c)$. There exists a one to one correspondence between smooth connections on P and holomorphic structures on P^c (see [104] or Lemma 2.2.5). This yields a natural action of $\mathcal{G}^c(P)$ on $\mathcal{A}(P)$ which extends the gauge action. Assume that the G-action on (X, J, ω) extends to a holomorphic G^c-action on (X, J) such that $\mathcal{G}^c(P)$ acts naturally on $\mathcal{S}(P, X)$.

Definition 3.1.1. *Let $(A, u) \in \mathcal{H}(P, X)$ and denote by $\overline{\mathcal{G}^c(A, u)}$ the $W^{1,2} \times W^{2,2}$-closure of its complexified orbit[1]. Denote by $\Phi(A, u) := *F_A + \mu(u)$ the moment map (3.2).*

[1] Here it suffices to consider the closure within the space $\mathcal{H}(P, X)$ of smooth holomorphic pairs. In the main part of the paper we will consider pairs (A, u) of Sobolev class $W^{1,2} \times W^{2,2}$ and gauge transformations of Sobolev class $W^{2,2}$. This does not affect the overall picture since (a) every complex orbit contains a dense set of smooth representatives and (b) every $W^{1,2} \times W^{2,2}$ solution to the vortex equation is gauge equivalent to a smooth solution. See Section 3.2.7 and Lemma 3.2.10.

1. *(A,u) is called **stable**, if $\Phi^{-1}(0)\cap\mathcal{G}^c(A,u)\neq\emptyset$ and the isotropy subgroup $\mathcal{G}_{(A,u)}:=\{k\in\mathcal{G}(P)\,|\,k(A,u)=(A,u)\}$ is discrete.*

2. *(A,u) is called **polystable**, if $\Phi^{-1}(0)\cap\mathcal{G}^c(A,u)\neq\emptyset$.*

3. *(A,u) is called **semistable**, if $\Phi^{-1}(0)\cap\overline{\mathcal{G}^c(A,u)}\neq\emptyset$.*

4. *(A,u) is called **unstable**, if $\Phi^{-1}(0)\cap\overline{\mathcal{G}^c(A,u)}=\emptyset$.*

Denote by $\mathcal{H}^s\subset\mathcal{H}^{ps}\subset\mathcal{H}^{ss}$ and $\mathcal{H}^{us}$ the corresponding $\mathcal{G}^c(P)$-invariant subspaces.

The GIT quotient of $\mathcal{H}(P,X)$ by $\mathcal{G}^c(P)$ is defined as the quotient space

$$\mathcal{M}_{GIT}(P,X):=\mathcal{H}^{ss}(P,X)//\mathcal{G}^c(P):=(\mathcal{H}^{ss}(P,X)/\mathcal{G}^c(P))/\sim$$

under the orbit closure relation $\mathcal{G}^c(A,u)\sim\mathcal{G}^c(B,v)$ if and only if $\overline{\mathcal{G}^c(A,u)}\cap\overline{\mathcal{G}^c(B,v)}\cap\mathcal{H}^{ss}(P,X)\neq\emptyset$. It follows from our main results that each equivalence class in this quotient contains a unique $\mathcal{G}(P)$-orbit of solutions to the symplectic vortex equations and $\mathcal{M}_{GIT}(P,X)\cong\mathcal{M}_{symp}(P,X)$ (see Corollary 3.1.7).

The main theorem

The moment map squared functional plays a crucial role in the differential geometric version of GIT. It is defined by

$$\mathcal{F}:\mathcal{H}(P,X)\to\mathbb{R},\qquad \mathcal{F}(A,u):=\frac{1}{2}\int_\Sigma ||*F_A+\mu(u)||^2\,dvol_\Sigma \tag{3.3}$$

and closely related to the Yang–Mills–Higgs functional

$$\mathcal{YMH}(A,u):=\frac{1}{2}\int_\Sigma ||F_A||^2+||d_Au||^2+||\mu(u)||^2\,dvol_\Sigma \tag{3.4}$$

by the energy identity in Proposition 3.2.2. In particular, for $(A,u)\in\mathcal{H}(P,X)$ it holds $\nabla\mathcal{YMH}(A,u)=\nabla\mathcal{F}(A,u)$, albeit the gradients look quite different at first glance. The negative gradient flow on $\mathcal{H}(P,X)$ has the following form

$$\begin{gathered} A(0)=A_0,\qquad u(0)=u_0,\qquad \bar\partial_A(u)=0\\ \partial_tA=*d_A(*F_A+\mu(u)),\qquad \partial_tu=-JL_u(*F_A+\mu(u)) \end{gathered} \tag{3.5}$$

Our main result says that solutions exist for all time and converge under the following hypothesis:

(A) The Kähler metric on X and the moment map $\mu:X\to\mathfrak{g}$ are both analytic.

(B) X is holomorphically aspherical.

(C) μ is proper and X is equivariantly convex at infinity, i.e. there exists a proper G-invariant function $f:X\to[0,\infty)$ and $c_0>0$ such that

$$f(x)\geq c_0\quad\Longrightarrow\quad \begin{gathered}\langle\nabla_v\nabla f(x),v\rangle+\langle\nabla_{Jv}\nabla f(x),Jv\rangle\geq 0\\ df(x)JL_x\mu(x)\geq 0\end{gathered} \tag{3.6}$$

for every $x\in X$ and $v\in T_xX$.

Theorem A (Convergence). *Assume **(C)** and let $(A_0, u_0) \in \mathcal{H}(P, X)$ be given. Then there exists a unique solution*

$$(A, u) : [0, \infty) \to \mathcal{H}(P, X)$$

*of (3.5) which exists for all times $t \geq 0$. If in addition **(A)**, **(B)** are satisfied, then there exists a critical point $(A_\infty, u_\infty) \in \mathcal{A}^{1,2}(P) \times \mathcal{S}^{2,2}(P, X)$ of Sobolev class $W^{1,2} \times W^{2,2}$ and $T, C, \epsilon > 0$ such that for all $t > T$ the pointwise distance between $u(t)$ and u_∞ is smaller then the injectivity radius of X along $u_\infty(P)$ and*

$$||A(t) - A_\infty||_{W^{1,2}} + ||\exp_{u_\infty}^{-1} u(t)||_{W^{2,2}} \leq Ct^{-\epsilon}.$$

Proof. Long time existence of the flow is established in Theorem 3.4.3 together with certain continuity and regularity assertions on the flow. The convergence part is proved in Theorem 3.4.8. □

Remark 3.1.2 (Regularity of the Limit.). Starting at a smooth initial condition $(A_0, u_0) \in \mathcal{H}(P, X)$, the solution $(A(t), u(t))$ of (3.5) remains smooth for all times $t > 0$. However, it is an open question if the limit (A_∞, u_∞) is smooth.

Lin [77] and Venugopalan [119] discussed the flow (3.5) independently and they proved under certain hypotheses that solutions exist for all times. Lin [77] considered in fact a generalization of (3.5), where Σ is replaced by a compact Kähler manifold, and showed that smooth solutions exist for all times when X is compact. His proof follows ideas of Donaldson [31] and he translates (3.5) into a heat flow on the space of complex gauge connections. Venugopalan [119] extended the arguments given by Råde [97] for the Yang–Mills flow and proved short time existence together with an uniform lower bound of the existence interval. For this argument she needed to assume that the flow remains in a compact region of X. We verify in Lemma 3.2.5 that this property follows from **(C)** and the maximum principle.

The main ingredient in our proof of the convergence of solutions to (3.5) is a Łojasiewicz gradient inequality for the Yang–Mills–Higgs functional (Theorem 3.4.4). This approach was introduced by Simon [101] and in its implementation we follow the arguments given by Råde [97] for the Yang–Mills flow.

Remark 3.1.3 (On assumption (A)). The proof depends on a suitable version of the Łojasiewicz gradient inequality and requires an analytic setup. In the finite dimensional case, it follows from the Marle and Guillemin-Sternberg normal form that the moment map squared functional is locally analytic (see Lerman [76]). If an analogous result is valid in our infinite dimensional setting, one might hope to remove this assumption.

Remark 3.1.4 (On assumption (B)).

1. Holomorphically aspherical means that every holomorphic map $\mathbb{C}P^1 \to X$ is constant.

2. This assumption prevents bubbling of holomorphic spheres within the fiber and is needed to establish sequential compactness along the flow lines (see Proposition 3.4.9).

3. X is necessarily noncompact under this assumption. Suppose otherwise that X is compact and there exists $\xi \in \mathfrak{g}\backslash\{0\}$ and $x_0 \in X$ such that $\exp(\xi) = \mathbb{1}$ and the infinitesimal action $L_{x_0}\xi \neq 0$ is nontrivial. Let $x : \mathbb{R} \to X$ be the solution of $\dot{x} = -JL_x\xi = -\nabla H_\xi(x)$ with $H_\xi := \langle \mu, \xi \rangle$ starting at $x(0) = x_0$. Since H_ξ is a Morse–Bott function, $x(t)$ converges exponentially to critical points $x^\pm$ as $t \to \pm\infty$ satisfying $L_{x^\pm}\xi = 0$. Using the S^1 action obtained from integrating the infinitesimal action of ξ, one can rotate this flow line within X and construct a nontrivial holomorphic sphere.

4. When X has nonpositive curvature, the distance function is plurisubharmonic and every holomorphic sphere $\mathbb{C}P^1 \to X$ is constant.

Remark 3.1.5. Important examples in which our assumptions are satisfied arise when X is a complex vector space (see [7, 13]).

Remark 3.1.6 (Higgs bundles). Let $X = \mathfrak{g}^c$ and consider the adjoint action of G on $\mathfrak{g}^c$. This action is Hamiltonian with moment map $\mu(\zeta) = \frac{i}{2}[\zeta, \zeta^*]$ where $\zeta^* := -\mathrm{Re}(\zeta) + \mathbf{i}\mathrm{Im}(\zeta)$. Then $P(X) = \mathrm{ad}(P^c)$ is a holomorphic vector bundle and our assumptions are satisfied. Higgs bundles are obtained as a slight variant of this setup where one considers holomorphic sections of the twisted bundle $P(X) \otimes K = \Omega^{1,0}(\Sigma, \mathrm{ad}(P^c))$. While this is not covered by our general discussion, the proof generalizes ad verbatim to this case.

Consequences of the main theorem

The infinitesimal action of $\xi \in \Omega^0(\Sigma, \mathrm{ad}(P))$ on $\mathcal{A}(P) \times \mathcal{S}(X, P)$ is given by

$$\mathcal{L}_{(A,u)}\xi := \left.\frac{d}{dt}\right|_{t=0} \exp(t\xi)(A, u) = (-d_A\xi, L_u\xi)$$

where $L_x : \mathfrak{g} \to T_xX$ denotes the infinitesimal action of $\mathfrak{g}$ on X. Denote by $\mathcal{L}^c_{(A,u)}$ the infinitesimal action of $\mathcal{G}^c(P)$ which agrees with the complexification of $\mathcal{L}_{(A,u)}$. Then

$$\nabla\mathcal{F}(A, u) = \mathcal{L}^c_{(A,u)}\mathbf{i}(*F_A + \mu(u))$$

implies that solutions of (3.5) remain in a single complexified orbit. The following result is the analogue of the Ness uniqueness theorem in finite dimensional GIT.

Theorem B (Uniqueness of critical points). *Assume **(A)**, **(B)** and **(C)**.*

1. *Let $(A_0, u_0) \in \mathcal{H}(P, X)$ and let (A_∞, u_∞) be the limit of the gradient flow (3.5) starting at (A_0, u_0). Then*

$$|| * F_{A_\infty} + \mu(u_\infty)||_{L^2} = \inf_{g \in \mathcal{G}^c(P)} || * F_{gA_0} + \mu(gu_0)||_{L^2} =: m.$$

2. *Suppose $(B_0, v_0), (B_1, v_1) \in \overline{\mathcal{G}^c(A_0, u_0)}$ (the $W^{1,2} \times W^{2,2}$-closure) and*

$$|| * F_{B_0} + \mu(v_0)||_{L^2} = m = || * F_{B_1} + \mu(v_1)||_{L^2}.$$

Then there exists $k \in \mathcal{G}(P)$ such that $(B_1, v_1) = k(B_0, v_0)$.

Proof. This is reformulated and proven in Theorem 3.5.1. The equivalence of both formulations follows from Proposition 3.2.2. □

Corollary 3.1.7. *Assume **(A)**, **(B)** and **(C)**. Every semistable orbit contains a unique polystable orbit in its $W^{1,2} \times W^{2,2}$-closure and every polystable orbit contains a unique $\mathcal{G}(P)$-orbit of solutions to the symplectic vortex equations.*

The corollary shows $\mathcal{M}_{symp}(P, X) \cong \mathcal{M}_{GIT}(P, X)$. More explicitly, this isomorphism is obtained by the map which sends $(A_0, u_0) \in \mathcal{H}(P, X)$ to its limit (A_∞, u_∞) under (3.5). Theorem 3.5.2 gives a complete characterization for the different stability conditions in Definition 3.1.1 in terms of the limit (A_∞, u_∞).

Next, we need to recall the general construction behind the Kempf–Ness theorem. Given $(A, u) \in \mathcal{H}(P, X)$ there exists a $\mathcal{G}(P)$-invariant functional

$$\Psi_{(A,u)} : \mathcal{G}^c(P) \to \mathbb{R}$$

whose gradient flow intertwines with (3.5) under the map $g \mapsto g^{-1}(A, u)$. The Kempf–Ness theorem characterizes the stability conditions of (A, u) in Definition 3.1.1 in terms of the global properties of $\Psi_{(A,u)}$. The stable case is the main step in Mundet's proof of the Kobayashi–Hitchin correspondence [89] and relates the stability of (A, u) to a certain properness of $\Psi_{(A,u)}$. The remaining cases are the content of the next theorem, whose proof is a relatively easy consequence of Theorem A and Theorem B.

Theorem C (Kempf–Ness Theorem). *Assume **(A)**, **(B)**, **(C)** and let $(A, u) \in \mathcal{H}(P, X)$.*

1. *(A, u) is polystable if and only if $\Psi_{(A,u)}$ has a critical point.*
2. *(A, u) is semistable if and only if $\Psi_{(A,u)}$ is bounded below.*
3. *(A, u) is unstable if and only if $\Psi_{(A,u)}$ is unbounded below.*

Proof. This is established in Theorem 3.5.5. □

The weights for the $\mathcal{G}^c(P)$-action are defined as the asymptotic slopes of $\Psi_{(A,u)}$ along geodesics rays in $\mathcal{G}^c(P)/\mathcal{G}(P)$. For $(A, u) \in \mathcal{H}(P, X)$ and $\xi \in \Omega^0(\Sigma, \mathrm{ad}(P))$ one has the explicit description

$$w((A, u), \xi) := \lim_{t\to\infty} \left\langle *F_{e^{\mathbf{i}t\xi}A} + \mu(e^{\mathbf{i}t\xi}u),\, \xi \right\rangle_{L^2} \in \mathbb{R} \cup \{\infty\}.$$

Mundet's Kobayashi–Hitchin correspondence asserts that (A, u) is stable if and only if $w((A, u), \xi) > 0$ for all $\xi \neq 0$. We extend this correspondence to the polystable and semistable case under the following technical assumption on a pair $(A, u) \in \mathcal{H}(P, X)$:

(H) For all $\xi \in \Omega^0(\Sigma, \mathrm{ad}(P))$ it holds:

$$w((A, u), \xi) \leq 0 \quad \Longrightarrow \quad \sup_{t>0} ||\mu(e^{\mathbf{i}t\xi}u)||_{L^2} < \infty.$$

Remark 3.1.8 (**On assumption (H)**).

1. **(H)** is trivially satisfied for stable pairs (A, u) and, by Proposition 3.6.6, it is always satisfied for polystable pairs.

2. By Proposition 3.6.2, $w((A, u), \xi) < \infty$ implies that $A_+ := \lim_{t\to\infty} e^{\mathbf{i}t\xi}A$ exists in C^∞.

3. Proposition 3.6.2 provides a strong tool to verify **(H)**. When X is a unitary vector space with linear $G \subset U(n)$ action, one can show that
$$w((A, u), \xi) < \infty \qquad \Longrightarrow \qquad \lim_{t\to\infty} e^{\mathbf{i}t\xi}(A, u) =: (A_+, u_+)$$
where the limit exists in C^∞ and **(H)** is satisfied in this case. Similarly, using Proposition 3.6.2, one verifies **(H)** for Higgs bundles.

4. **(H)** admits the following geometric description: For $(A, u) \in \mathcal{H}(P, X)$ denote by $\Psi_{(A,u)} : \mathcal{G}^c(P) \to \mathbb{R}$ its Kempf–Ness functional.

   (H') For all $\xi \in \Omega^0(\Sigma, \mathrm{ad}(P))$ it holds:
   $$\sup_{t>0} \Psi_{(A,u)}\left(e^{-\mathbf{i}t\xi}\right) < \infty \quad \Rightarrow \quad \sup_{t>0} \left|\left|\nabla\Psi_{(A,u)}\left(e^{-\mathbf{i}t\xi}\right)\right|\right|_{L^2} < \infty.$$

   Unraveling the definitions shows **(H)** $\Leftrightarrow$ **(H')**. This property is reasonable to expect, since $\Psi_{(A,u)}$ is convex along geodesics. However, one can construct examples which show that convexity of $\Psi_{(A,u)}$ alone does not guarantee **(H')**.

5. Unfortunately, we know little about the validity of **(H)** in general: We could neither prove that it is always satisfied, nor construct an explicit counterexample. This question is already meaningful (and open) in the finite dimensional case where $\Sigma = \{pt\}$.

Consider the following properties for a pair $(A, u) \in \mathcal{H}(P, X)$:

(SS) For all $\xi \in \Omega^0(\Sigma, \mathrm{ad}(P))$ it holds $w((A, u), \xi) \geq 0$.

(PS) For all $\xi \in \Omega^0(\Sigma, \mathrm{ad}(P))$ with $\exp(\xi) = \mathbb{1}$ and $w((A, u), \xi) = 0$ the limit $\lim_{t\to\infty} e^{\mathbf{i}t\xi}(A, u) \in (\mathcal{G}^c)^{2,2}(A, u)$ exists in $W^{1,2} \times W^{2,2}$ and remains in the Sobolev completion of the complex group orbit.

Theorem D (**Polystable and semistable correspondence**). *Assume **(A)**, **(B)**, **(C)** and suppose that $(A, u) \in \mathcal{H}(P, X)$ satisfies **(H)**.*

1. *(A, u) is polystable if and only if it satisfies **(SS)** and **(PS)**.*

2. *(A, u) is semistable if and only if it satisfies **(SS)**.*

Proof. This is established in Theorem 3.6.5 and Theorem 3.6.4. □

The polystable case has been established for twisted Higgs-bundles over Riemann surface by García-Prada, Gothen and Mundet [49] by different methods. They construct a Jordan-Hölder reduction and then deduce the polystable case from the stable case. For our proof the opposite is true and the stable case can be recovered as a special case of the polystable case. The proof is based on arguments of Chen–Sun [23]. The semistable correspondence follows from a sharp version of the moment weight inequality stated next.

Theorem E (**Sharp moment-weight inequality**). *Suppose that $(A,u)\in\mathcal{H}(P,X)$ satisfies **(H)**. Then for all $\xi\in\Omega^0(\Sigma,ad(P))\backslash\{0\}$ it holds*

$$-\frac{w((A,u),\xi)}{||\xi||_{L^2}}\le\inf_{g\in\mathcal{G}^c(P)}||*F_{gA}+\mu(gu)||_{L^2}. \tag{3.7}$$

*If in addition **(A)**, **(B)**, **(C)** are satisfied and the right hand side is positive, then there exists a unique $\xi_0\in\Omega^0(\Sigma,ad(P))$ with $||\xi_0||_{L^2}=1$ which yields equality.*

Proof. This is established in Theorem 3.6.3. □

For finite dimensional projective spaces the estimate (3.7) is due to Mumford [88] and Ness [92, Lemma 3.1], and the existence of a dominant weight is due to Kempf [68]. Around the same time Atiyah–Bott [4] established this result for the Yang–Mills equations over Riemann surfaces. Its generalization to the hermitian Yang–Mills equations over higher dimensional base manifolds is essentially equivalent to the Bando-Siu conjecture [6], established by Daskalopoulos–Wentworth [28], Sibling [100] and Jacob [65, 66]. In the context of K-stability and extremal Kähler metrics moment-weight inequalities are due to Tian [112], Donaldson [36, 37, 39] and Chen [18, 19]. In this context Chen–Sun [23] found an analytic proof of the Kempf existence theorem on finite dimensional spaces and we extend their argument to our infinite dimensional setting to prove existence of the dominant weight. The survey [51] by Georgoulas–Robbin–Salamon provides an overview on the different proofs of the moment weight inequality for Hamiltoninan actions on closed Kähler manifolds and its importance for geometric invariant theory.

3.2 Preliminaries

3.2.1 The moment map picture

We recall the natural Kähler structures on $\mathcal{A}(P)$ and $\mathcal{S}(P,X)$. Since $\mathcal{A}(P)$ is an affine space over the linear space $\Omega^1(\Sigma,\mathrm{ad}(P))$, it suffices to specify the Kähler structure on the later one. For $a,b\in\Omega^1(\Sigma,\mathrm{ad}(P))$ this is defined as

$$\omega_{\mathcal{A}}(a,b):=\int_\Sigma\langle a\wedge b\rangle,\qquad J_{\mathcal{A}}a=*a,\qquad \langle a,b\rangle_{\mathcal{A}}:=\int_\Sigma\langle a\wedge *b\rangle.$$

For $u\in\mathcal{S}(P,X)$ let $\tilde{u}:P\to X$ be the equivariant map determined by $u(z)=[p,\tilde{u}(p)]$ for $z\in\Sigma$ and $p\in P_z$. The tangent space $T_u\mathcal{S}(P,X)$ is represented by G-equivariant sections of the vector bundle $\tilde{u}^*TX\to P$ or equivalently by sections of the quotient bundle $\tilde{u}^*TX/G\to P/G=\Sigma$. The quotient bundle is again a vector bundle over Σ and we denote it in the following by u^*TX/G for simplicity. For $\hat{u}_1,\hat{u}_2\in T_u\mathcal{S}(P,X)=\Omega^0(\Sigma,u^*TX/G)$ one defines

$$\omega_{\mathcal{S}}(\hat{u}_1,\hat{u}_2):=\int_\Sigma\omega(\hat{u}_1,\hat{u}_2)\,dvol_\Sigma,\quad J_{\mathcal{S}}\hat{u}_1=J\hat{u}_1,\quad \langle\hat{u}_1,\hat{u}_1\rangle_{\mathcal{S}}:=\int_\Sigma\langle\hat{u}_1,\hat{u}_2\rangle\,dvol_\Sigma.$$

On $\mathcal{A}(P)\times\mathcal{S}(P,X)$ denote the product Kähler structure by $(\omega_{\mathcal{A}\times\mathcal{S}},J_{\mathcal{A}\times\mathcal{S}},\langle\cdot,\cdot\rangle_{\mathcal{A}\times\mathcal{S}})$.

Lemma 3.2.1. *The diagonal $\mathcal{G}(P)$-action on $\mathcal{A}(P) \times \mathcal{S}(P,X)$ is Hamiltonian with moment map*

$$\Phi : \mathcal{A}(P) \times \mathcal{S}(P,X) \to \Omega^0(\Sigma, ad(P)), \qquad \Phi(A,u) := *F_A + \mu(u). \tag{3.8}$$

Proof. For $(A,u) \in \mathcal{A}(P) \times \mathcal{S}(P,X)$ and $\xi \in \Omega^0(\Sigma, \mathrm{ad}(P))$ the infinitesimal action is given by

$$\mathcal{L}_{(A,u)}\xi := \left.\frac{d}{dt}\right|_{t=0} \exp(t\xi)(A,u) = (-d_A\xi, L_u\xi) \tag{3.9}$$

where $L_x : \mathfrak{g} \to T_xX$ denotes the infinitesimal action of $\mathfrak{g}$ on X. The verification of (3.9) is straightforward and left to the reader. The differential of the function

$$\mathcal{A}(P) \times \mathcal{S}(P,X) \to \mathbb{R}, \qquad (A,u) \mapsto \int_\Sigma \langle *F_A + \mu(u), \xi\rangle \, dvol_\Sigma$$

is the 1-form

$$T_A\mathcal{A}(P) \times T_u\mathcal{S}(P,X) \to \mathbb{R}, \qquad (a,\hat{u}) \mapsto \int_\Sigma \langle -d_A\xi \wedge a\rangle + \int_\Sigma \omega(L_u\xi, \hat{u}).$$

This is precisely $\omega_{\mathcal{A}\times\mathcal{S}}(\mathcal{L}_{(A,u)}\xi, \cdot)$ and (3.8) satisfies the moment map equation. □

3.2.2 Connections on $P(X)$ and the space $\mathcal{S}(P,X)$

For $x \in X$ the infinitesimal action of $\mathfrak{g}$ defines a map $L_x : \mathfrak{g} \to T_xX$. A smooth connection $A \in \mathcal{A}(P)$ induces on the Kähler fibration $P(X)$ the covariant derivative

$$d_A : \Omega^0(\Sigma, P(X)) \to \Omega^1(\Sigma, u^*TX/G), \qquad d_Au := du + L_uA$$

with values in the the vertical tangent bundle along u which is isomorphic to u^*TX/G. Moreover, A and the Levi-Civita connection induce a covariant derivative

$$\nabla_A : \Omega^0(\Sigma, u^*TX/G) \to \Omega^1(\Sigma, u^*TX/G), \qquad \nabla_A\xi := \nabla\xi + \nabla_\xi(L_uA).$$

All these covariant derivatives extend to first order elliptic operators between suitable Sobolev spaces.

3.2.3 The Yang–Mills–Higgs functional

The moment map squared functional (3.3) and the Yang–Mills–Higgs functional (3.4) are related by the following energy identity.

Proposition 3.2.2. *Let $(A,u) \in \mathcal{A}(P) \times \mathcal{S}(P,X)$, then*

$$\mathcal{YMH}(A,u) = \mathcal{F}(A,u) + \int_\Sigma ||\bar{\partial}_A(u)||^2 \, dvol_\Sigma + \langle \omega - \mu, u\rangle \tag{3.10}$$

where $\langle \omega - \mu, u\rangle = \int_\Sigma u^*\omega - d\langle \mu(u), A\rangle$.

Proof. This is Proposition 3.1 in [25]. □

Remark 3.2.3. The term $\langle \omega - \mu, u \rangle$ describes the pairing between the equivariant homology class $[u] \in H_2^G(X, \mathbb{Z})$ determined by u and the equivariant cohomology class $[\omega - \mu] \in H_G^2(X, \mathbb{R})$, see [25] for more details. In particular, this term is constant on the homotopy class of (A, u) and solutions of the symplectic vortex equation (3.1) minimize the Yang–Mills–Higgs functional in their homotopy class.

Lemma 3.2.4.

1. *The L^2-gradient of $\mathcal{F}$ is given by*

$$\nabla \mathcal{F}(A, u) = \begin{pmatrix} - * d_A(*F_A + \mu(u)) \\ JL_u(*F_A + \mu(u)) \end{pmatrix}.$$

2. *The L^2-gradient of $\mathcal{YMH}$ is given by*

$$\nabla \mathcal{YMH}(A, u) = \begin{pmatrix} d_A^* F_A + L_u^* d_A u \\ \nabla_A^* d_A u + d\mu(u)^* \mu(u) \end{pmatrix}.$$

3. *If $(A, u) \in \mathcal{H}(P, X)$, then both gradients are tangential to $\mathcal{H}(P, X)$ and agree. That is*

$$\begin{aligned} - * d_A(*F_A + \mu(u)) &= d_A^* F_A + L_u^* d_A u \\ J(u) L_u(*F_A + \mu(u)) &= \nabla_A^* d_A u + d\mu(u)^* \mu(u) \end{aligned}$$

holds for all $(A, u) \in \mathcal{H}(P, X)$.

Proof. We leave the first two parts to the reader (or refer to [119] and [77] for full details). For the last claim, note that

$$\nabla \mathcal{F}(A, u) = \begin{pmatrix} - * d_A(*F_A + \mu(u)) \\ J(u) L_u(*F_A + \mu(u)) \end{pmatrix} = \mathcal{L}^c_{(A,u)} \mathbf{i}(*F_A + \mu(u))$$

is tangential to the complexified orbit $\mathcal{G}^c(A, u) \subset \mathcal{H}(P, X)$ and hence tangential to $\mathcal{H}(P, X)$. Since $\mathcal{H}(P, X)$ minimizes the functional

$$(A, u) \mapsto \int_\Sigma ||\bar{\partial}_A u||^2 \, dvol_\Sigma$$

its gradient vanishes for $(A, u) \in \mathcal{H}(P, X)$ and the claim follows from the energy identity (3.10). □

3.2.4 Equivariant convexity at infinity

The next lemma shows that under assumption **(C)** solutions of (3.5) remain in a compact region of X.

Lemma 3.2.5. *Suppose X is equivariantly convex at infinity, let $f : X \to [0, \infty)$ and $c_0 > 0$ be as in (3.6). Let $T \in (0, \infty]$ and suppose $(A, u) : [0, T) \to \mathcal{H}(P, X)$ is a smooth map satisfying*

$$\partial_t u = -JL_u(*F_A + \mu(u)).$$

Then, for $c > c_0$ and $S_c := f^{-1}[0, c]$, it holds

$$u_0(P) \subset S_c \qquad \Longrightarrow \qquad u_t(P) \subset S_c$$

for every $t \in [0, T]$.

Proof. The proof is similar to the calculation in [24], Lemma 2.7.

In local trivializing coordinates $z = x + \mathbf{i}y$ define

$$v_x := \partial_x^A u := \partial_x u + L_u A(\partial_x), \qquad v_y := \partial_y^A u := \partial_y u + L_u A(\partial_y)$$

Denote by $\tilde{\Delta} := \partial_x^2 + \partial_y^2$ the standard Laplacian. Then

$$\begin{aligned}\tilde{\Delta} f(u) &= \partial_x \langle \nabla f(u), v_x\rangle + \partial_y \langle \nabla f(u), v_y\rangle \\ &= \langle \nabla_x^A \nabla f(u), v_x\rangle + \langle \nabla_y^A \nabla f(u), v_y\rangle + \langle \nabla f(u), \nabla_x^A v_x + \nabla_y^A v_y\rangle\end{aligned}$$

and since f is G-invariant, we obtain

$$\tilde{\Delta} f(u) = \langle \nabla_{v_x} \nabla f(u), v_x\rangle + \langle \nabla_{v_y} \nabla f(u), v_y\rangle + \langle \nabla f(u), \nabla_x^A v_x + \nabla_y^A v_y\rangle \tag{3.11}$$

Using the characteristic equation for the curvature

$$\nabla_x^A v_y - \nabla_y^A v_x = L_u F_A(\partial_x, \partial_y)$$

and the assumption $(A, u) \in \mathcal{H}(P, X)$, which is equivalent to $v_x + J v_y = 0$, we obtain

$$\nabla_x^A v_x + \nabla_y^A v_y = -J\left(\nabla_x^A v_x - \nabla_y^A v_x\right) = -J L_u F_A(\partial_x, \partial_y).$$

Inserting this in (3.11) yields

$$\tilde{\Delta} f(u) = \langle \nabla_{v_x} \nabla f(u), v_x\rangle + \langle \nabla_{v_y} \nabla f(u), v_y\rangle - \langle \nabla f(u), J L_u F_A(\partial_x, \partial_y)\rangle \tag{3.12}$$

If $f(u) \geq c_0$, then the convexity assumption implies that the first two terms in (3.12) are positive and thus

$$f(u) \geq c_0 \qquad \Longrightarrow \qquad \Delta f(u) \leq \langle \nabla f(u), J L_u * F_A\rangle$$

where $\Delta = d^* d$ denotes the positive Laplacian (which corresponds to $-\lambda\tilde{\Delta}$ in local coordinates for some function $\lambda > 0$). This yields

$$f(u) \geq c_0 \qquad \Longrightarrow \qquad (\partial_t + \Delta) f(u) \leq \langle \nabla f(u), -J L_u d\mu(u)\rangle \leq 0 \tag{3.13}$$

where we used the second equation in the convexity assumption.

We deduce the claim from (3.13) by contradiction. Suppose there exists $M > c$ such that

$$t_1 := \inf\{t \in [0, T) \,|\, f(u_t(z)) \geq M \text{ for some } z \in \Sigma\}$$

satisfies $0 < t_1 < T$ (i.e. $\inf \emptyset = \infty$ is excluded). Let $D \subset \Sigma$ be a small disc and let $t_0 \in (0, t_1)$ be such that

$$f(u_t(z)) > c_0 \qquad \forall (t, z) \in [t_0, t_1] \times D$$

and $f(u_{t_1}(z_0)) = M$ for some interior point $z_0 \in D$. It follows from (3.13) that in local coordinates $f(u_t(z))$ is a subsolution to a parabolic equation on $[t_0, t_1] \times D$ and by construction it attains its maximum on $\{t_1\} \times D$. By the strong maximum principle for parabolic equations (see [47] Chapter 2, Theorem 1), it follows that $f(u_t(x)) \equiv M$ is constant on $[t_0, t_1] \times D$. This contradicts the definition of t_1 and completes the proof of the lemma.

□

3.2.5 Sobolev spaces

We discuss mixed Sobolev spaces of time dependent sections of vector bundles. Following Råde [97] and Venugopalan [119] we use a norm on $H^r([0,t_0],H^s(\Sigma,V))$ which depends on the length t_0 of the time interval. For convenience, we use the abbreviation $H^s := W^{s,2}$ for L^2-Sobolev spaces.

Fractional Sobolev spaces on bounded domains

The refer to [1] for the general theory of Sobolev spaces. The definition of fractional Sobolev spaces (also called Bessel potential spaces) uses deep results from harmonic analysis (see [107] Chapter V.3 or [56] Chapter 2.1-3). For $s \in \mathbb{R}$ and $p \in (1,\infty)$ one defines

$$W^{s,p}(\mathbb{R}^n) := (1-\Delta)^{-s/2}\left(L^p(\mathbb{R}^n)\right), \qquad ||f||_{W^{s,p}} := ||(1-\Delta)^{s/2}f||_{L^p}. \tag{3.14}$$

For a smoothly bounded domain $\Omega \subset \mathbb{R}^n$ and $f \in C^\infty(\Omega)$ one defines

$$||f||_{W^{s,p}(\Omega)} = \inf_{f=F|_\Omega} ||F||_{W^{s,p}(\mathbb{R}^n)}$$

where the infimum ranges over all $F \in C_0^\infty(\mathbb{R}^n)$ which restrict to f. The space $W^{s,p}(\Omega)$ (resp. $W_0^{s,p}(\Omega)$) is the closure of $C^\infty(\Omega)$ (resp. $C_0^\infty(\Omega)$) under this norm. The extension theorem shows that $W^{s,p}(\Omega)$ is the set of restriction to Ω of functions in $W^{s,p}(\mathbb{R}^n)$. In the special case $p=2$ one obtains the Hilbert spaces $H^s(\Omega) = W^{s,2}(\Omega)$ (see [78]).

Interpolation. The spaces $W^{s,p}(\Omega)$ form a family of interpolation spaces in both parameters: the degree s of differentiability and the degree p of summability. For $1 < p_0, p_1 < \infty$, $s_0, s_1 \in \mathbb{R}$ and $0 < \theta < 1$ it holds

$$W^{s_\theta,p_\theta}(\Omega) \cong [W^{s_0,p_0}(\Omega), W^{s_1,p_1}(\Omega)]_\theta \tag{3.15}$$

with $s_\theta = (1-\theta)s_0 + \theta s_1$, $p_\theta = (1-\theta)p_0 + \theta p_1$ and $[\cdot,\cdot]_\theta$ refers to the holomorphic interpolation method. The same remains valid for the spaces $W_0^{s,p}(\Omega)$. (See [56] Chapter 2.4-5, [116] Chapters 1.9, 2.4 and 4.3).

Duality. For $1 < p, q < \infty$, $\frac{1}{p} + \frac{1}{q} = 1$ and $s \in \mathbb{R}_0^+$ there exists a natural identification

$$W^{-s,p}(\Omega) \cong W_0^{s,q}(\Omega)^*. \tag{3.16}$$

which is obtained by extending the L^2-product.

Products. Let $s, t, u \in \mathbb{R}$ with $u \leq \min\{s,t\}$ and $s + t \geq 0$. Let $1 < p, q, r < \infty$ with $s \neq n/p$, $t \neq n/q$, $u \neq -n/r$, and

$$\max\left\{\left(\frac{1}{p}-\frac{s}{n}\right), \left(\frac{1}{q}-\frac{t}{n}\right), \left(\frac{1}{p}-\frac{s}{n}\right)+\left(\frac{1}{q}-\frac{t}{n}\right)\right\} \leq \left(\frac{1}{r}-\frac{u}{n}\right) \tag{3.17}$$

Then, if $f \in W^{s,p}(\Omega)$ and $g \in W^{t,q}(\Omega)$, the product fg is contained in $W^{u,r}(\Omega)$ and satisfies an estimate

$$||fg||_{W^{u,r}(\Omega)} \leq C||f||_{W^{s,p}(\Omega)}||g||_{W^{t,q}(\Omega)}. \tag{3.18}$$

This follows for $s,t,u \in \mathbb{Z}_0^+$ from the Sobolev embedding theorem and Hölder's inequality. The general case is obtained from this by interpolation (3.15) and duality (3.16). (See [93] Theorem 9.6 for the details)

Sobolev spaces of sections $H^s(\Sigma, V)$

Let $V \to \Sigma$ be a Riemannian vector bundle over Σ. One can describe $H^s(\Sigma, V)$ in local coordinates as follows: Let $\{U_\alpha\}$ be an open trivializing cover of Σ by charts and choose unitary trivializations $V|_{U_\alpha} \cong U_\alpha \times \mathbb{R}^n$. A partition of unity subordinate to the cover $\{U_\alpha\}$ divides a section $\sigma \in \Omega^0(\Sigma, V)$ into a collection of functions $\sigma_\alpha^j \in C_0^\infty(U_\alpha)$. Using the charts we identify U_α with open bounded subsets $\Omega_\alpha \subset \mathbb{R}^2$ and define

$$||\sigma||_{H^s} := \sum_{\alpha,j} ||\sigma_\alpha^j||_{H^s(\Omega_\alpha)}. \tag{3.19}$$

The space $H^s(\Sigma, V)$ is the completion of $\Omega^0(\Sigma, V)$ in this norm.

Remark 3.2.6. Let ∇ be a smooth Riemannian connection on V. For $s = k \in \mathbb{Z}_0^+$ the norm

$$||\sigma|| := \sum_{j=0}^{k} ||\nabla^j \sigma||_{L^2} \tag{3.20}$$

is equivalent to the H^k-norm defined in (3.19). This leads to an alternative construction of $H^s(\Sigma, V)$ starting with (3.20) for positive integers and then using interpolation and duality.

The product formula (3.18) takes under the assumptions $p = q = r = 2$ and $n = 2$ the following simpler form. Let $s, t, u \in \mathbb{R}$ with $s, t \neq +1$, $u \neq -1$, $s + t \geq 0$, and

$$u \leq \min\{s, t, s + t - 1\}.$$

Then, if $f \in H^s(\Sigma)$ and $g \in H^t(\Sigma)$, the product fg is contained in $H^u(\Sigma)$ and satisfies an estimate

$$||fg||_{H^u} \leq C||f||_{H^s}||g||_{H^t}.$$

Time dependent Sobolev spaces $H^r([0, t_0], H^s(\Sigma, V))$

Let K be a separable Hilbert space, let $t_0 > 0$ and let $f : [0, t_0] \to K$ be a smooth function. The following is a slight variant of (3.14). For $r \in \mathbb{R}$ we define

$$||f||_{H^r([0,t_0],K)} := \inf_{F|_{[0,t_0]}=f} \left(\int_{-\infty}^{\infty} (\tau^2 + t_0^{-2})^r ||\hat{F}(\tau)||_K^2 \, d\tau \right)^{\frac{1}{2}} \tag{3.21}$$

where the infimum is taken over all $F \in C_0^\infty(\mathbb{R}, K)$ which restrict to f on $[0, t_0]$ and $\hat{F}$ denotes the Fourier transform.

Remark 3.2.7. If $r = k \in \mathbb{Z}_0^+$, then (3.21) is equivalent to the norm

$$||f|| := \sum_{j=0}^{k} \left|\left| t_0^{-(k-j)} \frac{d^j}{dt^j} f \right|\right|^2_{L^2([0,t_0],K)}. \tag{3.22}$$

As before, one could construct the spaces $H^r([0,t_0],K)$ using (3.22) for positive integers and then use interpolation and duality.

The dependence of the norms on t_0 has the advantage that for $r_1 \geq r_2$ the inclusion

$$H^{r_1}([0,t_0],K) \hookrightarrow H^{r_2}([0,t_0],K)$$

has norm $\leq Ct_0^{r_1-r_2}$. In particular, this norm can be controlled by t_0.

For $K = H^s(\Sigma, V)$ one obtains the spaces $H^r([0,t_0], H^s(\Sigma, V))$. These form again a family of interpolation spaces ([119] Lemma 6.36): For $s_0, s_1, r_0, r_1 \in \mathbb{R}$ and $\theta \in (0,1)$ we have

$$[H^{r_0}([0,t_0], H^{s_0}(\Sigma,V)), H^{r_1}([0,t_0], H^{s_1}(\Sigma,V))]_\theta \cong H^{r_\theta}([0,t_0], H^{s_\theta}(\Sigma,V))$$

with $r_\theta = (1-\theta)r_0 + \theta r_1$, $s_\theta = (1-\theta)s_0 + \theta s_1$ and $[\cdot,\cdot]_\theta$ denotes the holomorphic interpolation method.

3.2.6 The heat equation

Let $V \to \Sigma$ be a Riemannian vector bundle and let ∇ be a Riemannian connection on V.

Lemma 3.2.8. *For every $\sigma_0 \in \Omega^0(\Sigma, V)$ and $t_0 > 0$, there exists a unique smooth solution $\sigma : [0,t_0] \to \Omega^0(\Sigma, V)$ solving the initial value problem*

$$\partial_t \sigma + \nabla^* \nabla \sigma = 0, \qquad \sigma(0,\cdot) = \sigma_0. \tag{3.23}$$

Moreover, there exists a constant $C > 0$ such that the following estimate holds

$$||\sigma||_{L^2([0,t_0],H^1(\Sigma,V))} \leq Ct_0^{\frac{1}{2}} ||\sigma_0||_{L^2}.$$

Proof. This is a special case of Lemma 6.33 in [119]. □

From this we deduce the following estimates.

Lemma 3.2.9. *Let $f : [0,t_0] \to \Omega^0(\Sigma,V)$ be smooth. There exists a unique smooth solution ψ of the equation*

$$\partial_t \psi + \nabla^* \nabla \psi = f, \qquad \psi(0,\cdot) = 0. \tag{3.24}$$

Moreover, the solution satisfies the estimates

$$||\psi||_{L^2([0,t_0],H^1(\Sigma,V))} \leq Ct_0^{\frac{1}{2}} ||\psi||_{L^1([0,t_0],L^2(\Sigma,V))} \tag{3.25}$$

and

$$||\psi||_{L^2([0,t_0],H^1(\Sigma,V))} \leq Ct_0^{\frac{1}{4}} ||\psi||_{L^2([0,t_0],H^{-\frac{1}{2}}(\Sigma,V))}. \tag{3.26}$$

Proof. Let P_t denote the solution operator of (3.23), i.e. $P_0 = \mathbb{1}$ and $P_t\sigma_0(\cdot) = \sigma(t,\cdot)$ satisfies (3.23). The solution of (3.24) is then given by

$$\psi(t,\cdot) = \int_0^t P_{t-s} f(s,\cdot)\, ds.$$

The Minkowski inequality and Lemma 3.2.8 yield

$$\begin{aligned} ||\psi||_{L^2(H^1)} &\le \left(\int_0^{t_0} \left(\int_0^t ||P_{t-s} f(s,\cdot)||_{H^1}\, ds \right)^2 dt \right)^{\frac{1}{2}} \\ &\le \int_0^{t_0} \left(\int_s^{t_0} ||P_{t-s} f(s,\cdot)||_{H^1}^2\, dt \right)^{\frac{1}{2}} ds \\ &\le C t_0^{\frac{1}{2}} \int_0^{t_0} ||f(s,\cdot)||_{L^2}\, ds \end{aligned}$$

and this proves (3.25). Abbreviate

$$H^{r,s} := H^r([0,t_0], H^s(\Sigma, V)).$$

Parabolic regularity (see [119] Lemma 6.35) yields

$$||\psi||_{H^{\frac{3}{4},-\frac{1}{2}}} \le C||f||_{H^{-\frac{1}{4},-\frac{1}{2}}}, \qquad ||\psi||_{H^{-\frac{1}{4},\frac{3}{2}}} \le C||f||_{H^{-\frac{1}{4},-\frac{1}{2}}}$$

and, since $H^{0,1}$ is an interpolation space between $H^{\frac{3}{4},-\frac{1}{2}}$ and $H^{-\frac{1}{4},\frac{3}{2}}$, it follows

$$||\psi||_{H^{0,1}} \le C||f||_{H^{-\frac{1}{4},-\frac{1}{2}}} \le C t_0^{\frac{1}{4}} ||f||_{H^{0,-\frac{1}{2}}}.$$

This establishes (3.26) and completes the proof. □

3.2.7 Sobolev completions and regularity assumptions

For the main part of the article, we need to consider suitable Sobolev completions of the various spaces defined in the introduction. The space

$$\mathcal{S}^{2,2}(P,X) := W^{2,2}(\Sigma, P(X))$$

contains all continuous sections $u : \Sigma \to P(X)$ which in any trivialization of $P(X)$ and local coordinates on Σ and X are of Sobolev class $W^{2,2}_{\ell oc}$. It carries a natural topology, since $W^{2,2}_{\ell oc}(\mathbb{R}^2) \hookrightarrow C^0(\mathbb{R}^2)$ is in the good range of the Sobolev embedding: For $u \in \mathcal{S}(P,X)$ let $\epsilon > 0$ be smaller then the injectivity radius of X along the image of u. Then

$$\{\hat{u} \in W^{2,2}(\Sigma, u^*TX/G) \,|\, ||\hat{u}||_{W^{2,2}} < \epsilon\} \to \mathcal{S}^{2,2}(P,X), \qquad \hat{u} \mapsto \exp_u \hat{u}$$

defines a homeomorphism onto its image.

With respect to a smooth reference connection $A_0 \in \mathcal{A}(P)$, we define

$$\mathcal{A}^{1,2}(P) := \{A_0 + a \,|\, a \in W^{1,2}(\Sigma, T^*\Sigma \otimes \mathrm{ad}(P))\}$$

and denote

$$\mathcal{H}^{1,2}(P,X) := \{(A,u) \in \mathcal{A}^{1,2}(P) \times \mathcal{S}^{2,2}(P,X) \,|\, \bar\partial_A u = 0\}.$$

The $W^{2,2}$ completion of the gauge groups

$$\mathcal{G}^{2,2}(P) := W^{2,2}(\Sigma, \mathrm{Ad}(P)), \qquad (\mathcal{G}^c)^{2,2}(P) := W^{2,2}(\Sigma, \mathrm{Ad}(P \times_G G^c))$$

are similar defined as $\mathcal{S}^{2,2}(P,X)$ by requiring their sections to be of Sobolev class $W^{2,2}$ in any local trivialization. These groups act continuously on $\mathcal{S}^{2,2}(P,X)$, $\mathcal{A}^{1,2}(P)$ and $\mathcal{H}^{1,2}(P,X)$ as one readily checks.

Lemma 3.2.10. *Let $(A,u) \in \mathcal{H}^{1,2}(P,X)$.*

1. *There exists $g \in (\mathcal{G}^c)^{2,2}(P)$ such that $g(A,u)$ is smooth.*

2. *If (A,u) is a critical point of $\mathcal{YMH}$ satisfying*

$$d_A(*F_A + \mu) = 0, \qquad L_u(*F_A + \mu(u)) = 0 \tag{3.27}$$

then there exits $k \in \mathcal{G}^{2,2}(P)$ such that $k(A,u)$ is smooth.

Proof. This lemma is proven as in the Yang-Mills case. First, there exists $g \in (\mathcal{G}^c)^{2,2}$ such that gA is smooth (see [4], Lemma 14.8). Then $\bar\partial_{gA}(gu) = 0$ and elliptic regularity yields that gu is smooth. This proves the first part of the lemma.

For the second part we pass to a Coulomb gauge and choose a smooth reference connection $A_0 \in \mathcal{A}(P)$ and $k \in \mathcal{G}^{2,2}(P)$ such that $d^*_{A_0}(kA - A_0) = 0$. By (3.27), $a := kA - A_0$ satisfies

$$\Delta_{A_0} a = d_{A_0} * F_{A_0} + \frac{1}{2}[a \wedge a] + d_{A_0}(\mu(ku)) + [a \wedge (*F_{kA} + \mu(ku))]. \tag{3.28}$$

Suppose first that $a \in H^1$ and $u \in H^2$. Using the multiplication theorem $H^1 \otimes L^2 \to H^{-\frac{1}{2}}$, one sees that the right hand side of (3.28) is in $H^{-\frac{1}{2}}$ and hence $a \in H^{\frac{3}{2}}$. With this improved regularity, the right hand side of (3.28) is now contained in H^0 and hence $a \in H^2$. The holomorphicity condition

$$0 = \bar\partial_{kA}(ku) = \bar\partial_{A_0}(ku) + (L_{ku}a)^{0,1}$$

then yields $u \in H^3$. Repeating this argument, one shows $k(A,u) \in H^\ell \times H^{\ell+1}$ for every $\ell \geq 2$ and this completes the bootstrapping argument. □

3.3 The Łojasiewicz inequality for Gelfand triples

We establish an infinite dimensional version of the Łojasiewicz gradient inequality following closely the arguments of Råde [97] and Simon [101]. This result is an important ingredient in proving convergence of the Yang–Mills–Higgs gradient flow.

Let H be Hilbert space and let $V \subset H$ be a dense subset. Suppose V is a Hilbert space in its own right with respect to an inner product $\langle \cdot, \cdot \rangle_V$ and assume that the inclusion $V \subset H$ is compact. Identifying H with its dual, we obtain the Gelfand triple $V \subset H = H^* \subset V^*$. Let $F : V \to \mathbb{R}$ be a real analytic function and denote its differential by

$$M := dF : V \to V^*.$$

Assume F vanishes to the first order at the origin, i.e. $F(0) = 0$ and $M(0) = 0$. The linearization of M at the origin is given by

$$L = dM(0) : V \to V^*$$

and we call this map the Hessian of F at the origin.

Theorem 3.3.1. *Assume the setting described above and suppose there are constants $\delta, c > 0$ such that*

$$||Lx||_{V^*} \geq \delta ||x||_V - c||x||_H \tag{3.29}$$

is satisfied for all $x \in V$. Then there exist $\epsilon, C > 0$ and $\gamma \in [\frac{1}{2}, 1)$ such that for all $x \in V$ with $||x||_V \leq \epsilon$ it holds

$$||dF(x)||_{V^*} \geq C|F(x)|^\gamma.$$

Proof. The proof consists of six steps.

Step 1: *L has finite dimensional kernel and closed range.*

The proof is left as an exercise and uses the assumption that $V \subset H$ is compact. The result follows as in [86] Lemma A.1.1.

Step 2: *Construction of the finite dimensional approximation.*

Let $K := \ker(L)$ and denote its orthogonal complement by W'. The image $W'' := \mathrm{Im}(L) \subset V^*$ agrees with the annihilator of K. Identifying $K^* \subset V^*$ with the annihilator of W' yields decompositions

$$V = K \oplus W', \qquad V^* = K^* \oplus W''$$

and L restricts to an isomorphism $L : W' \to W''$. It follows from the implicit function theorem that there exists $\epsilon > 0$ and $\delta > 0$ such that for every $x \in K$ with $||x||_V < \epsilon$ there exists a unique $\phi(x) \in W'$ with $||\phi(x)||_V < \delta$ solving the equation $M(x + \phi(x)) \in K^*$. Moreover, the function

$$\phi : B_\epsilon(0; K) \to B_\delta(0; W')$$

is analytic. Define

$$f : B_\epsilon(0; K) \to \mathbb{R}, \qquad f(x) := F(x + \phi(x)).$$

This is a real analytic function on a finite dimensional domain.

Step 3: *For $x \in B_\epsilon(0; K)$ it holds $df(x) = M(x + \phi(x)) \in K^*$.*

For $x, y \in K$ the chain rule yields

$$\langle df(x), y\rangle_{V^* \times V} = \langle M(x+\phi(x)), y + d\phi(x)y\rangle_{V^*\times V}.$$

and this proves the claim, since $M(x+\phi(x)) \in K^*$ annihilates $d\phi(x)y \in W'$.

Step 4: *Decompose $x \in V$ with $||x||_V < \epsilon$ as*

$$x = x_0 + \phi(x_0) + x' \tag{3.30}$$

with $x_0 \in K$ and $x' \in W'$. For sufficiently small $\epsilon > 0$ there exists $C > 0$ such that

$$||M(x)||_{V^*} \geq C\left(||df(x_0)||_{V^*} + ||x'||_V\right) \tag{3.31}$$

holds for all $x \in V$ with $||x||_V < \epsilon$.

The terms in the decomposition (3.30) satisfy the estimates

$$||x_0||_V \leq C||x||_V, \quad ||\phi(x_0)||_V \leq C||x||_V, \quad ||x'||_V \leq C||x||_V. \tag{3.32}$$

Using Step 3 we obtain

$$\begin{aligned} M(x) &= M(x_0 + \phi(x_0) + x') \\ &= df(x_0) + \int_0^1 dM(x_0 + \phi(x_0) + tx')x'\,dt \\ &= df(x_0) + Lx' + \int_0^1 \left(dM(x_0+\phi(x_0)+tx') - dM(0)\right)x'\,dt. \end{aligned}$$

Since dM is continuously differentiable, it follows from (3.32) that there exists an estimate

$$\sup_{t\in[0,1]} ||dM(x_0+\phi(x_0)+tx') - dM(0)||_{\mathrm{Hom}(V,V^*)} \leq C||x||_V \leq C\epsilon.$$

Since $df(x_0) \in K^*$ and $Lx' \in W''$ we have

$$||df(x_0) + Lx'||_{V^*} \geq C\left(||df(x_0)||_{V^*} + ||Lx'||_{V^*}\right) \geq C(||df(x_0)||_{V^*} + ||x'||_V).$$

Combining these estimates yields

$$||M(x)||_{V^*} \geq C_1(||df(x_0)||_{V^*} + ||x'||_V) - C_2\epsilon||x'||_V$$

and this proves (3.31) after possibly shrinking $\epsilon > 0$.

Step 5: *For sufficiently small $\epsilon > 0$ there exists $C > 0$ such that*

$$|F(x)| \leq f(x_0) + C||x'||_V^2 \tag{3.33}$$

for all $x \in V$ with $||x||_V < \epsilon$.

The Taylor expansion of F yields:

$$
\begin{aligned}
F(x) &= F(x_0 + \phi(x_0) + x') \\
&= f(x_0) + \int_0^1 \langle M(x_0 + \phi(x_0) + tx'), x' \rangle_{V^* \times V} \, ds \\
&= f(x_0) + \langle M(x_0 + \phi(x_0)), x' \rangle_{V^* \times V} \\
&\qquad + \int_0^1 \int_0^1 \langle dM(x_0 + \phi(x_0) + stx') sx', x' \rangle_{V^* \times V} \, dsdt \\
&= f(x_0) + \langle df(x_0), x' \rangle_{V^* \times V} + \frac{1}{2} \langle Lx', x' \rangle_{V^* \times V} + \langle L_2 x', x' \rangle_{V^* \times V}
\end{aligned}
$$

where

$$L_2 x' := \int_0^1 \int_0^1 s \left(dM(x_0 + \phi(x_0) + stx') - dM(0) \right) x' \, dsdt.$$

As in Step 4 one shows that this term satisfies an estimate

$$\langle L_2 x', x' \rangle_{V^* \times V} \leq C ||x||_V ||x'||_V^2 \leq C\epsilon ||x'||^2.$$

The open mapping theorem yields the estimate $\langle Lx', x' \rangle_{V^* \times V} \geq C ||x'||_V^2$. Combining these estimates yields

$$|F(x)| \leq f(x_0) + C_1 ||x'||_V^2 - C_2 \epsilon ||x'||_V^2$$

and this proves (3.33) for sufficiently small $\epsilon > 0$.

Step 6: *For suffiently small $\epsilon > 0$, there exists $C > 0$ and $\gamma \in [\frac{1}{2}, 1)$ such that*

$$||M(x)||_{V^*} \geq C |F(x)|^\gamma \tag{3.34}$$

for all $x \in V$ with $||x||_V < \epsilon$.

The gradient inequality of Łojasiewicz [80] shows that for sufficently small $\epsilon > 0$ there exists $C > 0$ and $\gamma \in [\frac{1}{2}, 1)$ such that

$$||df(x)||_V \geq |f(x)|^\gamma$$

for all $x \in K$ with $||x||_V < \epsilon$. Since K is finite dimensional, there exists a constant such that $C ||df(x_0)||_{V^*} \geq ||df(x_0)||_V$. Now the estimates (3.31) and (3.33) show

$$
\begin{aligned}
||M(x)||_{V^*} &\geq C \left(||df(x_0)||_{V^*} + ||x'||_V \right) \\
&\geq C_1 \left| |F(x)| - C_2 ||x'||_V^2 \right|^\gamma + C_3 ||x'||_V.
\end{aligned}
$$

We may assume that $|F(x)| < 1$ for all $x \in B_\epsilon(0, V)$ and then follows (3.34) with $C := \min\{C_1 2^{-\gamma}, C_3/\sqrt{2C_2}\}$. $\square$

3.4 Convergence of the Yang–Mills–Higgs flow

In the first section weak solutions of the gradient flow (3.5) are defined and the existence and regularity of solutions are discussed. The second section contains a proof of the Łojasiewicz gradient inequality for the Yang–Mills–Higgs functional. Combining this inequality with an interior regularity result in the third section, we can then prove that solutions convergence under the additional assumptions **(A)**, **(B)**. This approach is very similar to the one developed by Råde [97] in the Yang–Mills case.

3.4.1 The gradient flow equations

Definition 3.4.1 (**Negative gradient flow of $\mathcal{YMH}$**). *A (weak) solution of*

$$\begin{aligned} \partial_t A &= -d_A^* F_A - L_u^* d_A u \\ \partial_t u &= -\nabla_A^* d_A u - d\mu(u)^* \mu(u) \end{aligned} \tag{3.35}$$

is a continuous map $(A, u) : [0, \infty) \to \mathcal{H}^{1,2}(P, X)$, such that there exists a sequence of smooth solutions of (3.35) converging to (A, u) in $C^0([0, \infty), H^1 \times H^2)$.

Definition 3.4.2 (**Negative gradient flow of $\mathcal{F}$**). *A (weak) solution of*

$$\begin{aligned} \partial_t A &= *d_A(*F_A + \mu(u)) \\ \partial_t u &= -JL_u(*F_A + \mu(u)) \end{aligned} \tag{3.36}$$

is a continuous map $(A, u) : [0, \infty) \to \mathcal{H}^{1,2}(P, X)$, such that there exists a sequence of smooth solutions of (3.36) converging to (A, u) in $C^0([0, \infty), H^1 \times H^2)$.

By Lemma 3.2.4 both of these flows agree:

$$(A, u) \text{ is a weak solution of (3.35)} \quad \Longleftrightarrow \quad (A, u) \text{ is a weak solution of (3.36).}$$

The following theorem is a slight extension of a result of Venugopalan [119] (she works in the $H^1 \times C^0$ topology and needs to assume that the flow remains in a compact region of X).

Theorem 3.4.3. *Assume **(C)** and let $(A_0, u_0) \in \mathcal{H}^{1,2}(P, X)$.*

1. *There exists a unique solution $(A, u) \in C^0([0, \infty), \mathcal{H}^{1,2}(P, X))$ of (3.36) with $A(0) = A_0$ and $u(0, \cdot) = u_0$.*
2. *The map $\Phi : [0, \infty) \to L^2(\Sigma, T^*\Sigma \otimes ad(P))$*
$$\Phi(t) := *F_{A(t)} + \mu(u(t)) \tag{3.37}$$
is contained in the spaces $C^0([0, \infty), L^2)$ and $L^2_{\ell oc}([0, \infty), H^1)$.
3. *The solution $g : [0, \infty) \to (\mathcal{G}^c)^{2,2}(P)$ of the ODE*
$$g^{-1}(t)\dot{g}(t) = \boldsymbol{i}(*F_{A(t)} + \mu(u(t))), \qquad g(0) = \mathbb{1}. \tag{3.38}$$
is continuous with values in H^2 and satisfies $(A(t), u(t)) = g(t)^{-1}(A_0, u_0)$.

4. *The solution $(A(t), u(t))$ of (3.36), the map $\Phi(t)$ in (3.37) and the solution $g(t)$ of (3.38) depend continuously on the initial condition $(A_0, u_0) \in \mathcal{H}^{1,2}(P, X)$ in the respective topologies stated above.*

Proof. Let $f : X \to [0, \infty)$ and $c_0 \in \mathbb{R}$ be as in (3.6). By Lemma 3.2.5 the compact sets $S_c := f^{-1}[0, c]$ with $c > c_0$ have the following property: If $(A(t), u(t))$ is a gradient flow line starting at (A_0, u_0), then

$$u_0(P) \subset S_c \qquad \Longrightarrow \qquad u_t(P) \subset S_c \quad \forall t \geq 0. \tag{3.39}$$

Venugopalan proves long time existence by establishing short time existence together with an uniform lower bound on the existence intervall. When we restrict to the set S_c her analysis yields uniform lower bounds for the existence interval for any solution with $u_0(P) \subset S_c$. Now Theorem 1.1 in [119] shows that for any $A_0 \in H^1$ and $u_0 \in C^0$ there exists a unique (weak) solution $(A, u) \in C^0([0, \infty), H^1 \times C^0)$. Moreover, the proof shows the solution (A, u) depends continuously on the initial condition (A_0, u_0), the moment map term $\Phi(t) := *F_{A(t)} + \mu(u(t))$ is contained in the space $C^0([0, \infty), L^2) \cap L^2_{\ell oc}([0, \infty), H^1)$ and depends continously on the initial condition (A_0, u_0) in these topologies. The additional regularity $\Phi \in L^2_{\ell oc}([0, \infty), H^1)$ is somewhat hidden in her proof and follows from the consideration of the space $\tilde{U}_P(t_0)$ at the end of the proof of Proposition 3.3. There she shows $\Phi \in H^{\frac{1}{2}+\epsilon, -2\epsilon} \cap H^{-\frac{1}{2}, 2}$ and this embedds into $L^2(H^1)$ by interpolation.

By the Sobolev embedding $H^1 \times H^2 \hookrightarrow H^1 \times C^0$, we obtain for any initial condition $(A_0, u_0) \in H^1 \times H^2$ a solution $(A, u) \in C^0([0, \infty), H^1 \times C^0)$. We claim that there exists a continuous path of complex gauge transformations $g : [0, \infty) \to (\mathcal{G}^c)^{2,2}(P)$, depending continuously on the initial condition (A_0, u_0), such that $(A(t), u(t)) = g(t)^{-1}(A_0, u_0)$. By continuity of the gauge action, this readily implies that $(A, u) \in C^0([0, \infty), H^1 \times H^2)$ and it depends continuously on the initial condition.

By (3.38) it holds $g \in H^1_{\ell oc}([0, \infty), H^1)$, since $\Phi \in L^2_{\ell oc}([0, \infty), H^1)$. And by (3.36) it holds $\partial_t A \in L^2_{\ell oc}([0, \infty), L^2)$. Hence $B(t) := A(t) - g(t)^{-1} A_0 \in H^1_{\ell oc}([0, \infty), L^1)$ and

$$\dot{B}(t) = [*B(t), \Phi(t)], \qquad B(0) = 0.$$

If B is smooth, this implies $B = 0$. In general, one can approximate weak solutions by smooth solutions and deduce then $B = 0$. This shows $A(t) = g(t)^{-1} A_0$ for all $t \geq 0$. Since $g \in H^1_{\ell oc}([0, \infty), H^1)$ and $A \in C^0([0, \infty), H^1)$ depend continuously on the initial condition in $H^1 \times H^2$, it follows from the equation

$$A(t)^{0,1} = (g(t)^{-1} A_0)^{0,1} = A_0 + g(t)^{-1} \bar{\partial}_{A_0} g(t)$$

and standard elliptic bootstrapping arguments that $g \in C^0([0, \infty), H^2)$ depends continuously on the initial condition. One readily checks $u(t) = g(t)^{-1} u_0$ and this completes the proof. □

3.4.2 The Łojasiewicz gradient inequality

The Łojasiewicz gradient inequality is the key ingredient in proving uniform convergence of the Yang–Mills–Higgs flow. This approach is due to Simon [101] and we follow quite closely the arguments of Råde [97] in the Yang–Mills case.

Theorem 3.4.4 (Łojasiewicz gradient inequality). *Assume **(A)** and let* $(A_\infty, u_\infty) \in \mathcal{H}^{1,2}(P,X)$ *be a critical point of* $\mathcal{YMH}$*. Then there exist* $\epsilon, C > 0$ *and* $\gamma \in [\frac{1}{2}, 1)$ *such that for all* $a \in H^1(\Sigma, T^*\Sigma \otimes ad(P))$ *and* $\hat{u} \in H^2(\Sigma, u^*TX/G)$ *with* $||a||_{H^1} + ||\hat{u}||_{H^2} < \epsilon$ *it holds*

$$\begin{aligned} &||\nabla\mathcal{YMH}(A_\infty + a, \exp_{u_\infty}\hat{u})||_{H^{-1}\times L^2} \\ &\qquad \geq c|\mathcal{YMH}(A_\infty + a, \exp_{u_\infty}\hat{u}) - \mathcal{YMH}(A_\infty, u_\infty)|^\gamma. \end{aligned} \tag{3.40}$$

Proof. See page 121. □

By Lemma 3.2.10 every critical point $(A_0, u_0) \in \mathcal{H}^{1,2}(P,X)$ of the Yang–Mills–Higgs functional is gauge equivalent to a smooth pair. Since the estimate (3.40) is $\mathcal{G}^{2,2}(P)$ invariant, we may assume in the following that $(A_\infty, u_\infty) \in \mathcal{H}(P,X)$ is smooth.

The infinitesimal gauge action induces for $s = \pm 1$ the L^2-orthogonal splittings

$$H^s(\Sigma, T^*\Sigma \otimes \mathrm{ad}(P)) \oplus H^{s+1}(\Sigma, u_\infty^* TX/G) = I^{s+1} \oplus V^{s,s+1} \tag{3.41}$$

with

$$\begin{aligned} I^{s+1} &:= \{(-d_{A_\infty}\xi, L_{u_\infty}\xi) \,|\, \xi \in H^{s+1}(\Sigma, \mathrm{ad}(P))\} \\ V^{s,s+1} &:= \{(a, \hat{u}) \in H^s \times H^{s+1} \,|\, d_{A_\infty}^* a + L_{u_\infty}^* \hat{u} = 0\}. \end{aligned}$$

Define

$$E : V^{1,2} \to \mathbb{R}, \qquad E(a,\hat{u}) := \mathcal{YMH}(A_\infty + a, \exp_{u_\infty}\hat{u}).$$

When $||\hat{u}||_{L^\infty}$ is smaller than the injectivity radius of X along $u_\infty(P)$, the L^2-gradient of E is given by

$$\nabla E(a,\hat{u}) = \Pi_V \circ T_{(a,\hat{u})} \nabla\mathcal{YMH}(A_\infty + a, \exp_{u_\infty}\hat{u})$$

where

$$\begin{aligned} T_{(a,\hat{u})} : T_{(A_\infty + a, \exp_{u_\infty}\hat{u})}(\mathcal{A} \times \mathcal{S}(P,X)) &\to T_{(A_\infty, u_\infty)}(\mathcal{A} \times \mathcal{S}(P,X)) \\ T_{(a,\hat{u})}(b,v) &:= (b, d\exp_{u_0}^{-1} v) \end{aligned}$$

and Π_V denotes the orthogonal projection onto $V^{-1,0}$ in (3.41). The next theorem establishes the Łojasiewicz inequality for E and we show below that this is equivalent to Theorem 3.4.4.

Theorem 3.4.5. *In the setting described above, there exist* $\epsilon, C > 0$ *and* $\gamma \in [\frac{1}{2}, 1)$ *such that for all* $(a,\hat{u}) \in V^{1,2}$ *with* $||a||_{H^1} + ||\hat{u}||_{H^2} < \epsilon$ *it holds*

$$||\nabla E(a,\hat{u})||_{H^{-1}\times L^2} \geq C|E(a,\hat{u}) - E(0,0)|^\gamma. \tag{3.42}$$

Proof. E is an analytic functional by assumption **(A)** and we claim that its Hessian

$$\nabla^2 E(0,0) : V^{1,2} \to V^{-1,0}$$

satisfies the elliptic estimate

$$||(b,v)||_{H^1 \times H^2} \leq C\left(||\nabla^2 E(0,0)(b,v)||_{H^{-1}\times L^2} + ||(b,v)||_{L^2 \times H^1}\right) \tag{3.43}$$

for all $(b,v) \in V^{1,2}$. Then Theorem 3.4.5 follows directly from Theorem 3.3.1.

The Hessian Q of the Yang–Mills–Higgs functional at (A_∞, u_∞) has the shape

$$Q(b,v) = (d^*_{A_\infty} d_{A_\infty} b, \nabla^*_{A_\infty} \nabla_{A_\infty} v) + R(b,v)$$

with some compact operator R. As a Hessian, this operator is symmetric and it follows from the gauge-invariance of the Yang–Mills–Higgs functional, that it restricts to the Hessian of E, i.e.

$$\nabla^2 E(0,0) = Q|_{V^{1,2}} : V^{1,2} \to V^{-1,0}$$

takes indeed values within $V^{-1,0}$. Now consider the operator

$$\Lambda := Q + \mathcal{L}_{(A_\infty,u_\infty)} \mathcal{L}^*_{(A_\infty,u_\infty)} : V^{1,2} \oplus I^2 \to V^{-1,0} \oplus I^0$$

This has the shape

$$\Lambda(b,v) = (d^*_{A_\infty} d_{A_\infty} b + d_{A_\infty} d^*_{A_\infty} b, \nabla^*_{A_\infty} \nabla_{A_\infty} v) + \tilde{R}(b,v)$$

with some compact operator $\tilde{R}$. In particular, Λ is Fredholm and satisfies the elliptic estimate

$$||(b,v)||_{H^1 \times H^2} \leq C \left(||\Lambda(b,v)||_{H^{-1} \times L^2} + ||(b,v)||_{L^2 \times H^1} \right). \tag{3.44}$$

Since $\mathcal{L}_{(A_\infty,u_\infty)} \mathcal{L}^*_{(A_\infty,u_\infty)}$ vanishes on $V^{1,2}$, we have $\Lambda|_{V^{1,2}} = \nabla^2 E$ and (3.43) follows from (3.44). □

Proof of Theorem 3.4.4. Since the Yang–Mills–Higgs functional is gauge-invariant, it follows from the implicit function theorem that we may assume $(a, \hat{u}) \in V^{1,2}$ with respect to the splitting (3.41). Let $\epsilon > 0$ be sufficiently small, let $(a, \hat{u}) \in V^{1,2}$ with $||a||_{H^1} + ||\hat{u}||_{H^2} < \epsilon$ and denote $(A, u) = (A_\infty + a, \exp_{u_\infty} \hat{u})$. Then

$$\begin{aligned} ||\nabla E(a,\hat{u})||_{H^{-1} \times L^2} &\leq C ||T_{(a,\hat{u})} \nabla \mathcal{YMH}(A,u)||_{H^{-1} \times L^2} \\ &\leq C ||\nabla \mathcal{YMH}(A,u)||_{H^{-1} \times L^2} \end{aligned}$$

and Theorem 3.4.4 follows from Theorem 3.4.5. □

Remark 3.4.6. Theorem 3.4.4 is in fact equivalent to Theorem 3.4.5. To see this denote by Π_I the projection onto I^0 in (3.41). With the same notation as above follows

$$\begin{aligned} &||\Pi_I \circ T_{(a,\hat{u})} \nabla \mathcal{YMH}(A,u)||_{H^{-1} \times L^2} \\ &\quad \leq C ||\mathcal{L}^*_{(A_\infty,u_\infty)} T_{(a,\hat{u})} \nabla \mathcal{YMH}(A,u)||_{H^{-2} \times H^{-1}} \\ &\quad = C ||(\mathcal{L}^*_{(A_\infty,u_\infty)} T_{(a,\hat{u})} - \mathcal{L}^*_{(A,u)}) \nabla \mathcal{YMH}(A,u)||_{H^{-2} \times H^{-1}} \end{aligned}$$

where the second equation uses $\mathcal{L}^*_{(A,u)} \nabla \mathcal{YMH}(A,u) = 0$. One verifies that the operator norm of $\mathcal{L}^*_{(A_\infty,u_\infty)} T_{(a,\hat{u})} - \mathcal{L}^*_{(A,u)}$ tends to zero as $||a||_{H^1} + ||\hat{u}||_{H^2}$ tends to zero. Hence

$$||\Pi_I \circ T_{(a,\hat{u})} \nabla \mathcal{YMH}(A,u)||_{H^{-1} \times L^2} \leq C(\epsilon) ||\nabla \mathcal{YMH}(A,u)||_{L^2}$$

where $C(\epsilon) > 0$ tends to zero as $\epsilon \to 0$ and therefore

$$||\nabla E(a,\hat{u})||_{H^{-1} \times L^2} \geq C ||\nabla \mathcal{YMH}(A,u)||_{H^{-1} \times L^2}$$

for sufficiently small $\epsilon > 0$.

3.4.3 Interior regularity

Theorem 3.4.7. *Let $(A_\infty, u_\infty) \in \mathcal{H}^{1,2}(P,X)$ be a critical point of $\mathcal{YMH}$. There exists $\epsilon_0 > 0$ such that for every $\epsilon \in (0, \epsilon_0)$ there exists $C > 0$ with the following significance: let $T > 1$ and let*

$$(a, \hat{u}) : [0,T] \to H^1(\Sigma, T^*\Sigma \otimes ad(P)) \times H^2(\Sigma, u_\infty^* TX/G)$$

be a continuous map such that $(A(t), u(t)) = (A_\infty + a(t), \exp_{u_\infty} \hat{u}(t)) \subset \mathcal{H}^{1,2}(P,X)$ is a solution of (3.35) and $||a(t)||_{H^1} + ||\hat{u}(t)||_{H^2} < \epsilon$ for all $t \in [0,T]$. Then

$$\int_1^T ||\partial_t a||_{H^1} + ||\partial_t \hat{u}||_{H^2}\, dt \leq C \int_0^T ||\nabla \mathcal{YMH}(A(t), u(t))||_{L^2}\, dt.$$

Proof. By Lemma 3.2.10, we may assume that $(A_\infty, u_\infty) \in \mathcal{H}(P,X)$ is smooth, after applying a suitable gauge transformation. The idea of the proof is then to show that the derivatives

$$\begin{aligned} b &:= \partial_t a = *d_A(*F_A + \mu(u)) \\ v &:= \partial_t \hat{u} = d\exp_{u_\infty}(\hat{u})^{-1}\left(-\nabla_A^* d_A u - JL_u\mu(u)\right) \end{aligned}$$

are solutions of the heat equation up to some perturbation which can be controlled. Note that the later expression is well-defined for sufficiently small $\epsilon > 0$, i.e. when $||\hat{u}||_{L^\infty} \leq C||\hat{u}||_{H^2} \leq C\epsilon$ is smaller than the injectivity radius of X along the image of u_∞. Using standard estimates for the heat equation, we can then deduce the interior regularity estimate. In the following fix a small time $0 < t_0 < \min\{1, T/2\}$.

Differentiating b in time gives

$$\partial_t b = \partial_t\left(*d_A(*F_A + \mu(u))\right) = -d_A^* d_A b + [*b, *F_A + \mu(u)] + *d_A(\partial_t \mu(u))$$

From the gauge-invariance follows $d_A^* b + L_u(\partial_t u) = 0$ and hence

$$\partial_t b + \Delta_A b = [*b, *F_A + \mu(u)] + d_A L_u(\partial_t u) + *d_A(\partial_t \mu(u)).$$

Finally the Bochner-Weizenböck formula yields the relation

$$\Delta_A b = \nabla_\infty^* \nabla_\infty b + F_A \times b + R_\Sigma \times b$$

where $\times$ denotes some bilinear expression and R_Σ is the Riemann curvature tensor of Σ. Then follows

$$\begin{aligned} (\partial_t + \nabla_\infty^*\nabla_\infty)b = {} & \nabla_\infty a \times b + a \times \nabla_\infty b + a \times a \times b + F_\infty \times b + R_\Sigma \times b + b \times \mu(u) \\ & + a \times L_u(\partial_t u) + a \times (\partial_t \mu(u)) + \nabla_\infty(L_u(\partial_t u)) + \nabla_\infty(\partial_t \mu(u)) \end{aligned}$$

Now choose a smooth cut-off function $\eta(t)$ such that $\eta(t) = 0$ for $\eta \in [0, t_0/2]$ and $\eta(t) = 1$ for $t \in [t_0, 2t_0]$. Then ηb satisfies

$$(\partial_t + \nabla_\infty^*\nabla_\infty)(\eta b) = \eta(\partial_t + \nabla_\infty^*\nabla_\infty)b + \eta'(t)b.$$

and vanishes at $t = 0$. Lemma 3.2.9 shows

$$
\begin{aligned}
||\eta b||_{L^2([0,2t_0],H^1)} \leq Ct_0^{\frac{1}{4}} &||\eta(\nabla_\infty a \times b + \dots + b \times \mu(u))||_{L^2([0,2t_0],H^{-\frac{1}{2}})} \\
&+ Ct_0^{\frac{1}{4}} ||\eta(a \times L_u(\partial_t u) + \dots + \nabla_\infty(\partial_t \mu(u)))||_{L^2([0,2t_0],H^{-\frac{1}{2}})} \\
&+ Ct_0^{\frac{1}{2}} ||\eta_t' b||_{L^1([0,2t_0],L^2)}
\end{aligned}
$$

Using the assumption $||a(t)||_{H^1} + ||\hat{u}(t)||_{H^2} < \epsilon$ and the multiplication theorem $H^1 \otimes L^2 \to H^{-\frac{1}{2}}$ it follows:

$$
||\eta(\nabla_\infty a \times b + \dots + b \times \mu(u))||_{L^2([0,2t_0],H^{-\frac{1}{2}})} \leq C\epsilon ||\eta b||_{L^2([0,2t_0],H^1)}
$$

$$
||\eta(a \times L_u(\partial_t u) + \dots + \nabla_\infty(\partial_t \mu(u)))||_{L^2([0,2t_0],H^{-\frac{1}{2}})} \leq C\epsilon ||\eta v||_{L^2([0,2t_0],H^2)}
$$

By choosing t_0 sufficiently small we thus obtain

$$
||\eta b||_{L^2([0,2t_0],H^1)} \leq c||\eta' b||_{L^1([0,2t_0],L^2)} + Ct_0^{\frac{1}{4}} ||\eta v||_{L^2([0,2t_0],H^2)} \tag{3.45}
$$

Next, we need to obtain a similar estimate for v. Define

$$
\begin{gathered}
\Psi : H^1(\Sigma, T^*\Sigma \otimes \mathrm{ad}(P)) \times H^2(\Sigma, u_\infty^* TX/G) \to H^{-\frac{1}{2}}(\Sigma, u_\infty^* TX/G) \\
\Psi(a, \hat{u}) := d\exp_{u_\infty}^{-1}(-\nabla_A^* d_A u - JL_u \mu(u))
\end{gathered}
$$

with $A = A_\infty + a$ and $u = \exp_{u_\infty} \hat{u}$. This is continuously differentiable and satisfies $\Psi(0,0) = 0$. In particular,

$$
v = \Psi(a, \hat{u}) = d\Psi(0,0)(a, \hat{u}) + q(a, \hat{u})
$$

where q vanishes to the first order. Differentiating this equation with respect to t yields

$$
\begin{aligned}
\partial_t v &= d\Psi(0,0)(b, v) + \partial_t q(a, \hat{u}) \\
&= -\nabla_{A_\infty}^* \nabla_{A_\infty} v - \nabla_v(JL_{u_\infty} \mu(u_\infty)) - \nabla_{A_\infty}^* L_{u_\infty} b + \nabla b \times d_{A_\infty} u_\infty \\
&\quad + dq(a_t, \hat{u}_t)[b, v]
\end{aligned}
$$

Let η be the same cut-off function as above. Then ηv vanishes at $t = 0$ and solves the equation

$$
\begin{aligned}
(\partial_t + \nabla_{A_\infty}^* \nabla_{A_\infty})(\eta v) = &-\nabla_{\eta v}(JL_{u_\infty} \mu(u_\infty) - \nabla_{A_\infty}^* L_{u_\infty}(\eta b) + \nabla(\eta b) \times d_\infty u_\infty \\
&+ \eta \partial_t q(a_t, \hat{u}_t)[b, v] + \eta' v
\end{aligned}
$$

It follows from Lemma 3.2.9 that

$$
\begin{aligned}
||\eta v||_{L^2([0,2t_0],H^1)} \leq Ct_0^{\frac{1}{4}} &|| - \nabla_{\eta v}(JL_{u_\infty} \mu(u_\infty) - \ldots + \nabla(\eta b) \times d_\infty u_\infty||_{L^2([0,2t_0],H^{-\frac{1}{2}})} \\
&+ Ct_0^{\frac{1}{4}} ||\eta dq(a_t, \hat{u}_t)[b, v]||_{L^2([0,2t_0],H^{-\frac{1}{2}})} \\
&+ Ct_0^{\frac{1}{2}} ||\eta_t' v||_{L^1([0,2t_0],L^2)}
\end{aligned}
$$

The first term satisfies the estimate

$$\begin{aligned}&||-\nabla_{\eta v}(JL_{u_\infty}\mu(u_\infty)-\cdots+\nabla(\eta b)\times d_\infty u_\infty||_{L^2([0,2t_0],H^{-\frac{1}{2}})}\\&\qquad\le C||\eta v||_{L^2([0,2t_0],H^1)}+C||\eta b||_{L^2([0,2t_0],H^1)}\end{aligned}$$

and it follows from the definition of q that

$$||\eta dq(a_t,\hat{u}_t)[b,v]||_{L^2([0,2t_0],H^{-\frac{1}{2}})}\le C\epsilon\left(||\eta b||_{L^2([0,2t_0],H^1)}+||\eta v||_{L^2([0,2t_0],H^2)}\right)$$

For sufficiently small $t_0>0$ we thus get

$$\begin{aligned}||\eta v||_{L^2([0,2t_0],H^1)}&\le C||\eta_t'v||_{L^1([0,2t_0],L^2)}\\&\qquad+Ct_0^{\frac{1}{4}}(||\eta v||_{L^2([0,2t_0],H^2)}+||\eta b||_{L^2([0,2t_0],H^1)})\end{aligned}\tag{3.46}$$

Finally, differentiating the holomorphicity condition $\bar{\partial}_A u=\bar{\partial}u+L_uA^{0,1}=0$, we obtain by elliptic regularity the estimate

$$||v_t||_{H^2}\le C(||v_t||_{H^1}+||b_t||_{H^1}).\tag{3.47}$$

Combining (3.45, (3.46) and (3.47) yields for sufficiently small $t_0>0$

$$||\eta b||_{L^2([0,2t_0],H^1)}+||\eta v||_{L^2([0,2t_0],H^2)}\le C\left(||\eta' b||_{L^1([0,2t_0],L^2)}+||\eta' v||_{L^1([0,2t_0],L^2)}\right).$$

In particular

$$\begin{aligned}&||b||_{L^1([t_0,2t_0],H^1)}+||\eta v||_{L^1([t_0,2t_0],H^2)}\\&\qquad\le t_0^{\frac{1}{2}}\left(||\eta b||_{L^2([0,2t_0],H^1)}+||\eta v||_{L^2([0,2t_0],H^2)}\right)\\&\qquad\le Ct_0^{\frac{1}{2}}\left(||\eta' b||_{L^1([0,2t_0],L^2)}+||\eta' v||_{L^1([0,2t_0],L^2)}\right)\\&\qquad\le Ct_0^{-\frac{1}{2}}\left(||b||_{L^1([0,2t_0],L^2)}+||v||_{L^1([0,2t_0],L^2)}\right).\end{aligned}$$

The proof follows now by subdividing the interval $[0,T]$ into smaller intervals of length t_0 and applying the estimate above to each pair of successive subintervals.

□

3.4.4 The convergence theorem

The next theorem establishes uniform convergence of the Yang–Mills–Higgs flow.

Theorem 3.4.8. *Assume **(A)**, **(B)** and **(C)**. Let $(A_0,u_0)\in\mathcal{H}^{1,2}(P,X)$ and let $(A(t),u(t))$ be the solution of (3.36). There exist $C,\beta>0$ such that for all $T>0$*

$$\int_T^\infty ||\partial_t A(t)||_{H^1}+||\partial_t u(t)||_{H^2}\,dt\le CT^{-\beta}.$$

In particular, $(A(t),u(t))$ converges uniformly in $H^1\times H^2$ to a critical point (A_∞,u_∞) of $\mathcal{YMH}$.

The following compactness result arises from a combination of Gromov compactness for holomorphic curves and Uhlenbeck compactness.

Proposition 3.4.9. *Assume* **(B)**, **(C)** *and let* $(A(t), u(t)) \subset \mathcal{H}^{1,2}(P, X)$ *be a solution of (3.36). Then there exists sequences of times* $t_j \to \infty$ *and of gauge transformations* $k_j \in \mathcal{G}^{2,2}(P)$ *and a critical point* (A_∞, u_∞) *of* $\mathcal{YMH}$ *such that* $k_j(A(t_j), u(t_j))$ *converges to* (A_∞, u_∞) *in* $H^1 \times H^2$.

Proof. This is a special case of Theorem 1.2 in [119]. □

Remark 3.4.10. In general, we expect the convergence of $u(t_j)$ only modulo bubbling in finitely many fibers, as stated in [119] Theorem 1.2. Assumption **(B)** rules the formation of bubbles out and is crucial for the result stated in Proposition 3.4.9.

Proof of Theorem 3.4.8. Let $(A(t), u(t)) \subset \mathcal{H}^{1,2}(P, X)$ be a solution of (3.36). Let $t_j \to \infty$, $k_j \in \mathcal{G}^{2,2}(P)$ and (A_∞, u_∞) be as in Proposition 3.4.9 above. Choose $\epsilon > 0$, such that the Łojasiewicz gradient inequality in Theorem 3.4.4 is satisfied with respect to (A_∞, u_∞). Let $\delta \in (0, \epsilon)$ and choose $j \geq 1$ such that

$$||A_\infty - k_j A(t_j)||_{H^1} + ||\exp_{u_\infty}^{-1}(k_j u(t_j))||_{H^2} < \delta.$$

Since the gradient flow and the Łojasiewicz inequality are $\mathcal{G}^{2,2}(P)$-equivariant, we may assume $k_j = \mathbb{1}$ and $t_j = 0$.

The gradient flow depends continuously in the $C^0(H^1 \times H^2)$ topology on the initial conditions by Theorem 3.4.3. Since the flow is constant at the critical point (A_∞, u_∞) this yields

$$||A(1) - A(0)||_{H^1} + ||\exp_{u_\infty}^{-1} u_1 - \exp_{u_\infty}^{-1} u_0||_{H^2} \leq \rho(\delta)$$

where $\rho(\delta) \to 0$ as $\delta \to 0$. Define

$$\overline{T} := \inf\{t > 0 \,|\, ||A(t) - A_\infty||_{H^1} + ||\exp_{u_\infty}^{-1} u_t||_{H^2} \geq \epsilon\}.$$

By choosing $\delta > 0$ sufficiently small, we can guarantee $\overline{T} > 1$. For $1 < s < \overline{T}$ define $\hat{u}(s) := \exp_{u_\infty}^{-1} u(s)$. The interior regularity estimate in Theorem 3.4.7 and the Łojasiewicz gradient inequality in Theorem 3.4.4 yield

$$\begin{aligned}
||A(s)-A_\infty||_{H_1} &+ ||\hat{u}(s)||_{H^2} \\
&\leq \rho(\delta) + \int_1^s ||\partial_t A(t)||_{H^1} + ||\partial_t \hat{u}(t)||_{H^2}\, dt \\
&\leq \rho(\delta) + C\int_0^s ||\partial_t A(t)||_{L^2} + ||\partial_t u(t)||_{L^2}\, dt \\
&\leq \rho(\delta) + C\int_0^s \frac{||\nabla \mathcal{YMH}(A,u)||^2_{L^2 \times L^2}}{(\mathcal{YMH}(A,u) - \mathcal{YMH}(A_\infty, u_\infty))^\gamma}\, dt \\
&\leq \rho(\delta) + C\left(\mathcal{YMH}(A(0), u(0)) - \mathcal{YMH}(A_\infty, u_\infty)\right)^{1-\gamma}.
\end{aligned}$$

For $\delta > 0$ sufficiently small, this shows $\overline{T} = \infty$ and the integral $\int_1^\infty ||\partial_t A(t)||_{H^1} + ||\partial_t u(t)||_{H^2}\, dt < \infty$ is finite. This proves that $(A(t), u(t))$ converges uniformly in $H^1 \times H^2$ to a critical point $(\tilde{A}_\infty, \tilde{u}_\infty)$ of the Yang–Mills–Higgs functional.

Repeating the argument from above, with respect to the critical point $(\tilde{A}_\infty, \tilde{u}_\infty)$ we obtain for all sufficently large T

$$\int_T^\infty ||\partial_t A(t)||_{H^1} + ||\partial_t \hat{u}(t)||_{H^2}\, dt \leq f(T-1)^{1-\gamma}$$

with $f(t) := (\mathcal{YMH}(A(t), u(t)) - \mathcal{YMH}(\tilde{A}_\infty, \tilde{u}_\infty))$. Since

$$f'(t) = -||\nabla \mathcal{YMH}(A(t), u(t))||_{L^2}^2 \leq -Cf(t)^{2\gamma}$$

it follows $f(t) \leq Ct^{\frac{1}{1-2\gamma}}$ and hence

$$\int_T^\infty ||\partial_t A(t)||_{H^1} + ||\partial_t \hat{u}(t)||_{H^2}\, dt \leq C(T-1)^{\frac{1-\gamma}{1-2\gamma}}$$

for all sufficiently large T. This is equivalent to the estimate in the Theorem and completes the proof. □

We state some consequences of the proof for later reference.

Corollary 3.4.11. *Assume **(A)**, **(B)**, **(C)** and let $(B, v) \in \mathcal{H}(P, X)$ be a critical point of the Yang–Mills–Higgs functional. There exist $C, \epsilon_0 > 0$ and $\gamma \in [\frac{1}{2}, 1)$ with the following significance: let $(A, u) : [0, \infty) \to \mathcal{H}(P, X)$ be a solution of (3.36) satisfying $||A(0) - B||_{H^1} + ||\exp_v^{-1} u(0)||_{H^2} < \epsilon_0$ and $\mathcal{YMH}(A(t), u(t)) \geq \mathcal{YMH}(B, v)$ for all $t > 0$. Then*

1. *The limit satisfies $\mathcal{YMH}(A_\infty, u_\infty) = \mathcal{YMH}(B, v)$.*
2. *For every $\epsilon > 0$ exists $\delta \in (0, \epsilon_0)$ such that*

$$\int_0^\infty ||\partial_t A(t)||_{H^1} + ||\partial_t u(t)||_{H^2}\, dt < \epsilon$$

whenever $||A(0) - B||_{H^1} + ||\exp_v^{-1} u(0)||_{H^2} < \delta$.

3.5 Uniqueness and the Kempf–Ness theorem

3.5.1 Uniqueness of critical points

The next result is a reformulation of Theorem B in the introduction and the analogue of the Ness uniqueness theorem in finite dimensional GIT. The proof is based on arguments of Chen–Sun [23] in the finite dimensional differentiable setting.

Theorem 3.5.1 (**Uniqueness of critical points**)**.** *Assume **(A)**, **(B)** and **(C)**. Let $(A_0, u_0) \in \mathcal{H}^{1,2}(P, X)$ and (A_∞, u_∞) be the limit of the Yang–Mills–Higgs flow (3.36) starting at (A_0, u_0). Then $(A_\infty, u_\infty) \in \overline{(\mathcal{G}^c)^{2,2}(A_0, u_0)}$ (the $H^1 \times H^2$ closure) and*

$$\mathcal{YMH}(A_\infty, u_\infty) = \inf_{g \in (\mathcal{G}^c)^{2,2}(P)} \mathcal{YMH}(gA_0, gu_0).$$

Moreover, if $(B, v) \in \overline{(\mathcal{G}^c)^{2,2}(A_0, u_0)}$ and $\mathcal{YMH}(B, v) = \mathcal{YMH}(A_\infty, u_\infty)$, then $(B, v) \in \mathcal{G}^{2,2}(A_\infty, u_\infty)$.

Proof. The proof consists of four steps.

Step 1: *Let $(B, v) \in \mathcal{H}^{1,2}(P, X)$ and let $g_j : [0, \infty) \to (\mathcal{G}^c)^{2,2}(P)$ satisfy*

$$g_j(t)^{-1}\dot{g}_j(t) = \boldsymbol{i}\left(F_{g_j(t)^{-1}B} + \mu(g_j(t)^{-1}v)\right)$$

for $j \in \{0,1\}$. Using the Cartan decomposition, write

$$g_1(t) = g_0(t)e^{\mathbf{i}\eta(t)}k(t)$$

with $\eta(t) \in H^2(\Sigma, ad(P))$ and $k(t) \in \mathcal{G}^{2,2}(P)$. Then $\eta(t)$ and $k(t)$ are uniformly bounded in H^2.

Denote by $\pi : G^c \to G^c/G$ the canonical projection. The homogeneous space G^c/G is a complete Riemannian manifold with nonpositive curvature and for $t > 0$ the curve $\gamma(s,t) := \pi(g_0(t)e^{\mathbf{i}s\eta(t)})$ is pointwise the unique geodesic of length $||\eta(t)||$ connecting $\pi(g_0(t))$ and $\pi(g_1(t))$. This yields

$$\begin{aligned}
\partial_t||\eta(t)||^2 &= 2\int_0^1 \langle \nabla_t \partial_s\gamma, \partial_s\gamma\rangle \, ds = 2\int_0^1 \partial_s\langle \partial_t\gamma, \partial_s\gamma\rangle \, ds \\
&= 2\langle g_1(t)^{-1}\dot{g}_1(t), \mathbf{i}\eta(t)\rangle - \langle g_0(t)^{-1}\dot{g}_0(t), \mathbf{i}\eta(t)\rangle \\
&= 2\langle *F_{g_1(t)^{-1}B} - *F_{g_0(t)^{-1}B}, \eta(t)\rangle + 2\langle \mu(g_1(t)^{-1}v) - \mu(g_0(t)^{-1}v), \eta(t)\rangle.
\end{aligned}$$

Abbreviate $(B_{s,t}, v_{s,t}) := e^{-\mathbf{i}s\eta(t)}g_0(t)^{-1}(B,v)$. Then

$$\begin{aligned}
\partial_t||\eta(t)||^2 &= 2\int_0^1 \partial_s\langle *F_{B_{s,t}} + \mu(v_{s,t}), \eta(t)\rangle \, ds \\
&= -2\langle \Delta_{B_{s,t}}\eta + L^*_{v_{s,t}}L_{v_{s,t}}\eta(t), \eta(t)\rangle \\
&= -\Delta(||\eta(t)||^2) - 2\int_0^1 \left(||L_{u_{s,t}}\eta(t)||^2 + ||d_{A_{s,t}}\eta(t)||^2\right) \, ds
\end{aligned}$$

Thus $||\eta||^2$ satisfies the differential inequality $(\partial_t + \Delta)||\eta||^2 \leq 0$ and by the maximum principle for the heat equation $\eta(t)$ is uniformly bounded in L^∞. Since $(B_j(t), v_j(t)) := (g_j(t)^{-1}B, g_j(t)^{-1}v)$ satisfies (3.36), it converge uniformly in $H^1 \times H^2$ by Theorem 3.4.8. Hence it follows from the equation

$$B_1(t) = g_1(t)^{-1}B = k(t)^{-1}e^{-\mathbf{i}\eta(t)}g_0(t)^{-1}B = k(t)^{-1}e^{-\mathbf{i}\eta(t)}B_0(t)$$

and elliptic bootstrapping that $\eta(t)$ and $k(t)$ are uniformly bounded in H^2.

Step 2: *Let $(B_0, v_0), (B_1, v_1) \in \mathcal{H}^{1,2}(P, X)$ be critical points of the Yang–Mills–Higgs functional. If $(B_1, v_1) \in (\mathcal{G}^c)^{2,2}(B_0, v_0)$, then $(B_1, v_1) \in \mathcal{G}^{2,2}(B_0, v_0)$.*

Choose $\tilde{g} \in (\mathcal{G}^c)^{2,2}(P)$ such that $\tilde{g}^{-1}(B_1, v_1) = (B_0, v_0)$. By Theorem 3.4.3 there exist $g_0, g_1 : [0,\infty) \to (\mathcal{G}^c)^{2,2}(P)$ solving

$$\begin{aligned}
g_0^{-1}\dot{g}_0 &= *F_{B_0} + \mu(v_0), \qquad g_0(0) = \mathbb{1}, \qquad g_0^{-1}(t)(B_0, v_0) = (B_0, v_0) \\
g_1^{-1}\dot{g}_1 &= *F_{B_1} + \mu(v_1), \qquad g_1(0) = \tilde{g}, \qquad g_1^{-1}(t)(B_0, v_0) = (B_1, v_1).
\end{aligned}$$

Both of these curves satisfy the conditions of Step 1 with $(B, v) = (B_0, v_0)$. Using the same notation as in Step 1, write $g_1(t) = g_0(t)e^{\mathbf{i}\eta(t)}k(t)$, and conclude that there exists a sequence $t_j \to \infty$ such that $k(t_j) \to k_\infty$ and $\eta(t_j) \to \eta_\infty$ converge weakly in H^2 and strongly in $W^{1,p}$. Then

$$B_{s,t_j} \xrightarrow{L^p} B_{s,\infty} := e^{-\mathbf{i}s\eta_\infty}B_0, \qquad d_{B_{s,t_j}}\eta(t_j) \xrightarrow{L^p} d_{B_{s,\infty}}\eta_\infty$$

where $(B_{s,t}, v_{s,t}) := e^{-\mathbf{i}s\eta(t)}(B_0, v_0)$ as in Step 1. It follows from the calculation in Step 1 that

$$\partial_t ||\eta||^2_{L^2} = -2 \int_0^1 ||d_{B_{s,t}}\eta||^2_{L^2} + ||L_{v_{s,t}}\eta||^2_{L^2}\, ds$$

and we may assume in addition

$$\lim_{j\to\infty} ||d_{B_{s,t_j}}\eta(t_j)||_{L^2} + ||L_{v_{s,t_j}}\eta(t_j)||_{L^2} = 0.$$

At $s = 0$ we obtain $\mathcal{L}_{(B_0,v_0)}\eta_\infty = 0$ and hence

$$(B_1, v_1) = g_1(t_j)^{-1}(B_1, v_1) \xrightarrow{L^p} k_\infty^{-1} e^{-\mathbf{i}\eta_\infty}(B_0, v_0) = k_\infty^{-1}(B_0, v_0).$$

This shows $(B_1, v_1) \in \mathcal{G}^{2,2}(B_0, v_0)$ and completes the proof of Step 2.

Step 3: *Let $(A_0, u_0), (B_0, v_0) \in \mathcal{H}^{1,2}(P, X)$ and denote by $(A_\infty, u_\infty), (B_\infty, v_\infty)$ the limits of the Yang–Mills–Higgs flow (3.36) starting at (A_0, u_0), (B_0, v_0) respectively. If $(B_0, v_0) \in (\mathcal{G}^c)^{2,2}(A_0, u_0)$, then $(B_\infty, v_\infty) \in \mathcal{G}^{2,2}(A_\infty, u_\infty)$.*

Denote by $(A(t), u(t))$ and $(B(t), v(t))$ the solutions of (3.36) starting at (A_0, u_0) and (B_0, v_0) respectively. Choose $\tilde{g} \in (\mathcal{G}^c)^{2,2}(P)$ such that $(B_0, v_0) = \tilde{g}^{-1}(A_0, u_0)$. Then, by Theorem 3.4.3, there exist $g_0, g_1 : [0, \infty) \to (\mathcal{G}^c)^{2,2}(P)$ solving

$$g_0^{-1}\dot{g}_0(t) = \Phi(A(t), u(t)), \qquad g_0(0) = \mathbb{1}, \qquad g_0^{-1}(t)(A_0, u_0) = (A(t), u(t))$$

$$g_1^{-1}\dot{g}_1(t) = \Phi(B(t), u(t)), \qquad g_1(0) = \tilde{g}, \qquad g_1^{-1}(t)(A_0, u_0) = (B(t), v(t)).$$

Both of these curves satisfy the conditions of Step 1 with $(B, v) = (A_0, u_0)$. Using the same notation as in Step 1, write $g_1(t) = g_0(t)e^{\mathbf{i}\eta(t)}k(t)$. Then there exists a sequence $t_j \to \infty$ such that $\eta(t_j) \to \eta_\infty$ and $k(t_j) \to k_\infty$ converge weakly in H^2 and strongly in $W^{1,p}$. As j tends to infinity in the equation

$$(B(t_j), v(t_j)) = k(t_j)^{-1}e^{-\mathbf{i}\eta(t_j)}(A(t_j), u(t_j))$$

both sides converge in $L^p \times W^{1,p}$ and this yields $(B_\infty, v_\infty) = k_\infty^{-1}e^{-\mathbf{i}\eta_\infty}(A_\infty, u_\infty)$. This proves $(\mathcal{G}^c)^{2,2}(A_\infty, u_\infty) = (\mathcal{G}^c)^{2,2}(B_\infty, v_\infty)$ and Step 3 follows from Step 2.

Step 4: *If $(B, v) \in \overline{(\mathcal{G}^c)^{2,2}(A_0, u_0)}$ and $\mathcal{YMH}(B, v) = \mathcal{YMH}(A_\infty, u_\infty)$, then $(B, v) \in \mathcal{G}^{2,2}(A_\infty, u_\infty)$.*

It follows from Step 3 that

$$\mathcal{YMH}(A_\infty, u_\infty) = \inf_{g\in(\mathcal{G}^c)^{2,2}(P)} \mathcal{YMH}(gA_0, gu_0) =: m.$$

Note that the solution $(B(t), v(t))$ of the Yang–Mills–Higgs flow (3.36) starting at (B, v) remains in the closure $\overline{(\mathcal{G}^c)^{2,2}(A_0, u_0)}$. Hence $\mathcal{YMH}(B(t), v(t)) = m$ is constant, $(B(t), v(t))$ a constant flow line and (B, v) a critical point.

Choose $(A^{(j)}, u^{(j)}) \in (\mathcal{G}^c)^{2,2}(A_0, u_0)$ converging in $H^1 \times H^2$ to (B, v) and denote the limit of the Yang–Mills–Higgs flow starting at $(A^{(j)}, u^{(j)})$ by $(B^{(j)}, v^{(j)})$. Corollary

3.4.11 shows that $(B^{(j)}, v^{(j)})$ converges to (B, v) in the $H^1 \times H^2$ topology. By Step 3, there exists $k_j \in \mathcal{G}^{2,2}(P)$ such that $(B^{(j)}, v^{(j)}) = (k_j A_\infty, k_j u_\infty)$. Since the connections $B^{(j)}$ are uniformly bounded in H^1, the gauge transformations k_j are uniformly bounded in H^2 and after passing to a subsequence, we may assume that $k_j \to k_\infty$ converges weakly in H^2 and strongly in $W^{1,p}$. It follows $(B, v) = (k_\infty A_\infty, k_\infty u_\infty)$ and this completes the proof.

□

Theorem 3.5.2. *Assume **(A)**, **(B)** and **(C)**. Let $(A, u) \in \mathcal{H}^{1,2}(P, X)$ and denote by (A_∞, u_∞) the limit of the flow (3.5) starting at (A, u). Then*

1. *(A, u) is stable if and only if $\mathcal{L}_{(A_\infty, u_\infty)}$ is injective.*
2. *(A, u) is polystable if and only if $(A_\infty, u_\infty) \in \mathcal{G}^c(A, u)$ remains in the complex orbit and $*F_{A_\infty} + \mu(u_\infty) = 0$.*
3. *(A, u) is semistable if and only if $*F_{A_\infty} + \mu(u_\infty) = 0$.*
4. *(A, u) is unstable if and only if $*F_{A_\infty} + \mu(u_\infty) \neq 0$.*

Remark 3.5.3. We call an element $(A, u) \in \mathcal{H}^{1,2}(P, X)$ stable, polystable, semistable or unstable, if every smooth element of $(\mathcal{G}^c)^{2,2}(A, u)$ is stable in the sense of Definition 3.1.1. Note that for every stable pair (A, u) the extension of the infinitesimal action

$$\mathcal{L}^c_{(A,u)} : H^2(\Sigma, \mathrm{ad}(P)) \to H^1(\Sigma, T^*\Sigma \otimes \mathrm{ad}(P)) \oplus H^2(\Sigma, u^*TX/G)$$

remains injective. This follows from Lemma 3.2.10 and elliptic regularity.

Proof. The unstable, semistable and polystabe characterization follow directly from Theorem 3.5.1.

For the stable case, note that every stable orbit has discrete $\mathcal{G}^c$-isotropy and this proves one direction. Conversely, the limit satisfies the critical point equation $\mathcal{L}_{(A_\infty, u_\infty)}(*F_{A_\infty} + \mu(u_\infty)) = 0$. Hence, when $\mathcal{L}_{(A_\infty, u_\infty)}$ is injective, (A_∞, u_∞) is stable. Since the subset of stable pairs $\mathcal{H}^{1,2}_s(P, X)$ is open by Proposition 3.5.4 below, this implies $(A(t), u(t)) \in \mathcal{H}^{1,2}_s(P, X)$ for sufficiently large t and (A, u) is stable. □

Proposition 3.5.4. *Assume **(A)**, **(B)** and **(C)**. The subsets of stable and semistable pairs $\mathcal{H}^{1,2}_{ss}(P, X) \subset \mathcal{H}^{1,2}_s(P, X) \subset \mathcal{H}^{1,2}(P, X)$ are open subsets in the $H^1 \times H^2$-topology.*

Proof. The semistable case follows from Corollary 3.4.11. The stable case follows from a suitable application of the implicit function theorem: Suppose $(A, u) \in \mathcal{H}^{1,2}(P, X)$ solves the vortex equation $*F_A + \mu(u) = 0$ and $\mathcal{L}_{(A,u)}$ is injective. Then $\mathcal{L}^c_{(A,u)}$ is also injective and, since

$$\langle \mathcal{L}^c_{(A,u)} \mathbf{i}\xi, (a, \hat{u}) \rangle_{L^2 \times L^2} = \langle \xi, *d_A a + d\mu(u)\hat{u} \rangle_{L^2}$$

for all $\xi \in H^2(\Sigma, \mathrm{ad}(P))$ and $(a, \hat{u}) \in T_{(A,u)}(\mathcal{A}(P) \times \mathcal{S}(P, X))$, (A, u) is a regular point for the moment map $\Phi(A, u) = *F_A + \mu(u)$. It follows that

$$\mathcal{Z} := \Phi^{-1}(0) \subset \mathcal{A}^{1,2}(P) \times \mathcal{S}^{2,2}(P, X)$$

is a submanifold locally around (A,u) and the orthogonal complement of $T_{(A,u)}\mathcal{Z}$ coincides with the image of $H^2(\Sigma, \mathbf{i}\,\mathrm{ad}(P))$ under $\mathcal{L}^c_{(A,u)}$. Hence

$$H^2(\Sigma, T^*\Sigma \otimes \mathrm{ad}(P^c)) \times \mathcal{Z} \to \mathcal{A}^{1,2}(P) \times \mathcal{S}^{2,2}(P,X), \qquad (\xi, z) \mapsto \exp(\mathbf{i}\xi)z$$

restricts to a diffeomorphism between neighborhoods of $(0;(A,u))$ and (A,u). In particular, (A,u) is an interior point of $\mathcal{H}^{1,2}_s(P,X)$.

□

3.5.2 The Kempf–Ness theorem

Let $(A,u) \in \mathcal{H}^{1,2}(P,X)$ and define the 1-form $\alpha_{(A,u)} : T(\mathcal{G}^c)^{2,2}(P) \to \mathbb{R}$ by

$$\alpha_{(A,u)}(g,\hat{g}) := -\int_\Sigma \left\langle *F_{g^{-1}A} + \mu(g^{-1}u); \mathrm{Im}(g^{-1}\hat{g}) \right\rangle \, dvol_\Sigma. \tag{3.48}$$

It is straight forward to check that $\alpha_{(A,u)}$ is exact, $\mathcal{G}^{2,2}(P)$-invariant and integrates to a unique $\mathcal{G}^{2,2}(P)$-invariant functional

$$\Psi_{(A,u)} : (\mathcal{G}^c)^{2,2}(P) \to \mathbb{R} \tag{3.49}$$

satisfying $\Psi_{(A,u)}(\mathbb{1}) = 0$ (see e.g. [89]). We call $\Psi_{(A,u)}$ the Kempf–Ness functional associated to (A,u).

Theorem 3.4.3 shows that for every $g_0 \in (\mathcal{G}^c)^{2,2}(P)$ the negative gradient flow

$$g^{-1}(t)\dot{g}(t) = -g^{-1}(t)\nabla\Psi_{(A,u)}(g(t)) = -\mathbf{i}(*F_{g(t)^{-1}A} + \mu(g(t)^{-1}u)) \tag{3.50}$$

has a unique solution $g \in C^0([0,\infty), (\mathcal{G}^c)^{2,2}(P))$ satisfying $g(0) = g_0$. This flow intertwines with the Yang–Mills–Higgs flow in the following sense

$$g(t) \text{ solves } (3.50) \quad \Longrightarrow \quad (A(t), u(t)) := (g(t)^{-1}A, g(t)^{-1}u) \text{ solves } (3.36).$$

We will repetitively make use of the fact that $\Psi_{(A,u)}$ is convex along geodesics in $(\mathcal{G}^c)^{2,2}(P)/\mathcal{G}^{2,2}(P)$. This amounts to the formula

$$\frac{d^2}{dt^2}\Psi_{(A,u)}(ge^{\mathbf{i}t\xi}) = \left|\left|\mathcal{L}_{e^{-\mathbf{i}t\xi}g^{-1}(A,u)}\xi\right|\right|^2_{L^2} \geq 0 \tag{3.51}$$

for $g \in (\mathcal{G}^c)^{2,2}(P)$ and $\xi \in H^2(\Sigma, \mathrm{ad}(P))$.

The Kempf–Ness theorem relates the stability of the pair (A,u) to global properties of the functional $\Psi_{(A,u)}$. The stable case is due to Mundet [89], see Remark 3.5.6 below. The remaining cases are the content of the next theorem which is a reformulation of Theorem C in the introduction.

Theorem 3.5.5. *Assume* ***(A)***, ***(B)*** *and* ***(C)*** *and let* $(A,u) \in \mathcal{H}^{1,2}(P,X)$.

1. (A,u) *is polystable if and only if* $\Psi_{(A,u)}$ *has a critical point.*

2. (A,u) *is semistable if and only if* $\Psi_{(A,u)}$ *is bounded below.*

3. (A,u) *is unstable if and only if* $\Psi_{(A,u)}$ *is unbounded below.*

Proof. The polystable case follows from (3.48). For the other two cases let $g_0 \in (\mathcal{G}^c)^{2,2}(P)$ and $g : [0,\infty) \to (\mathcal{G}^c)^{2,2}(P)$ be the solution of (3.50) starting at g_0. Then

$$\frac{d}{dt}\Psi_{(A,u)}(g(t)) = \alpha_{(A,u)}(g(t),\dot g(t)) = -||*F_{g(t)^{-1}A} + \mu(g(t)^{-1}u)||_{L^2}^2.$$

If (A,u) is unstable, Theorem 3.5.1 shows that the right hand side is bounded above by a strictly negative constant and hence $\Psi_{(A,u)}$ is unbounded below. Conversely, assume that (A,u) is semistable. Then $(A(t),u(t)) := (g(t)^{-1}A, g(t)^{-1}u)$ satisfies (3.36) and its limit (A_∞,u_∞) solves $*F_{A_\infty} + \mu(u_\infty)$ by Theorem 3.5.2. By Proposition 3.2.2 and Theorem 3.4.4 there exist $\gamma \in [\frac{1}{2},1)$ and $C,T>0$ such that for all $t>T$

$$\begin{aligned}
||*F_{g(t)^{-1}A} + \mu(g(t)^{-1}u)||_{L^2}^2 &= 2\left(\mathcal{F}(A(t),u(t)) - \mathcal{F}(A_\infty,u_\infty)\right)\\
&= 2\left(\mathcal{YMH}(A(t),u(t)) - \mathcal{YMH}(A_\infty,u_\infty)\right)\\
&\le 2\left(\mathcal{YMH}(A(t),u(t)) - \mathcal{YMH}(A_\infty,u_\infty)\right)^\gamma\\
&\le C||\nabla\mathcal{YMH}(A(t),u(t))||_{L^2}\\
&= C||\partial_t(A(t),u(t))||_{L^2}.
\end{aligned}$$

Theorem 3.4.8 shows that the right-hand-side is integrable and hence

$$m := \lim_{t\to\infty}\Psi_{(A,u)}(g(t)) > -\infty.$$

We claim $m = \inf\Psi_{(A,u)}$. For this let $\tilde g_0 \in (\mathcal{G}^c)^{2,2}(P)$ and denote by $\tilde g(t)$ the solution of (3.50) starting at $\tilde g_0$. It follows from Step 1 of the proof of Theorem 3.5.1, that the pointwise geodesic distance between $g(t)$ and $\tilde g(t)$ in G^c/G remains uniformly bounded. Since $\Psi_{(A,u)}$ is convex along geodesics in $(\mathcal{G}^c)^{2,2}(P)/\mathcal{G}^{2,2}(P)$ by (3.51) and its gradient converges to zero along $g(t)$ and $\tilde g(t)$, it follows that $|\Psi_{(A,u)}(g(t)) - \Psi_{(A,u)}(\tilde g(t))|$ converges to zero. This proves the claim and $\Psi_{(A,u)}$ is bounded below m. □

Remark 3.5.6 (The stable case). In finite dimensions the Kempf–Ness functional of a point is proper if and only if this point is stable. Mundet [89] established the following analogous result for the vortex in equations in great generality: (A,u) is stable if and only if the complexified orbit $\mathcal{G}^c(A,u)$ has discrete $\mathcal{G}^c$-isotropy and for every $R>0$ there exist $c_1,c_2>0$ such that

$$||*F_{e^{-\mathbf{i}\xi}A} + \mu(e^{-\mathbf{i}\xi}A)||_{L^2} < R \quad\Longrightarrow\quad ||\xi||_{L^\infty} \le c_1\Psi_{(A,u)}(e^{\mathbf{i}\xi}) + c_2. \tag{3.52}$$

3.6 Polystability and the moment-weight inequality

3.6.1 The Kobayashi–Hitchin correspondence

Finite weights

The weights of $(A,u) \in \mathcal{H}(P,X)$ are defined as the asymptotic slopes of $\Psi_{(A,u)}$ along the geodesic rays $[\exp(-\mathbf{i}t\xi)]$ in $\mathcal{G}^c/\mathcal{G}$. Here $\xi \in \mathrm{Lie}(\mathcal{G})$ is a section of $T^*\Sigma \otimes \mathrm{ad}(P)$ and one may hope to replace the conditions on $\Psi_{(A,u)}$ in Theorem 3.5.5 by conditions on these weights. In general, one needs to consider sections ξ of very low regularity, namely of Sobolev class H^1. For bundles over a Riemann surface and smooth pairs (A,u) every finite weight is obtained from a smooth section by Proposition 3.6.2 below.

Definition 3.6.1. *For $(A,u) \in \mathcal{H}(P,X)$ and $\xi \in H^1(\Sigma, ad(P))$ define*

$$w((A,u),\xi) := \lim_{t\to\infty} \langle *F_{e^{it\xi}A} + \mu(e^{it\xi}u), \xi\rangle_{L^2} \in \mathbb{R}\cup\{+\infty\}.$$

By (3.51), the right-hand-side is monotone increasing in t and the limit exists.

Similarly, define by

$$w(A,\xi) := \lim_{t\to\infty} \langle *F_{e^{\mathrm{i}t\xi}A}, \xi\rangle, \qquad w(u,\xi) := \lim_{t\to\infty} \langle \mu(e^{\mathbf{i}t\xi}u), \xi\rangle$$

the weights for the $\mathcal{G}(P)$-action on $\mathcal{A}(P)$ and $\mathcal{S}(P,X)$ respectively. They are well-defined in $\mathbb{R}\cup\{+\infty\}$ and satisfy $w((A,u),\xi) = w(A,\xi) + w(u,\xi)$.

Proposition 3.6.2. *Let $A \in \mathcal{A}(P)$ be smooth and let $\xi \in H^1(\Sigma, ad(P))\backslash\{0\}$ with $w(A,\xi) < \infty$.*

1. *Endow $P^c := P \times_G G^c$ with the holomorphic structure induced by A. Then there exists $\xi_0 \in \mathfrak{g}\backslash\{0\}$ and a holomorphic reduction $P_Q \subset P^c$ to the parabolic subgroup*

$$Q = Q(\xi_0) := \left\{ q \in G^c \;\middle|\; \text{the limit } \lim_{t\to\infty} e^{it\xi_0} q e^{-it\xi_0} =: q_+ \text{ exists} \right\}.$$

   The reduction $P_Q \subset P^c$ induces a smooth reduction $P_K \subset P$ to the centralizer $K = C_G(\xi_0)$ and ξ is the image of ξ_0 under the following map

$$Z(Lie(K)) \to \Omega^0(\Sigma, ad(P_K)) \to \Omega^0(\Sigma, ad(P))$$

   where the first arrow identifies central elements with constant sections and the second map is obtained from the inclusion $P_K \subset P$.

2. *The limit $A_+ := \lim_{t\to\infty} e^{it\xi} A$ exists in H^1 and A_+ restricts to a smooth connection on P_K.*

Proof. This is an intrinsic version of [89] Lemma 4.2 and makes use of a deep reularity result of Uhlenbeck and Yau [118] on weakly holomorphic subbundles, see Lemma 2.5.7. The reduction $P_K \subset P$ is induced by the isomorphism $G^c/Q(\xi_0) \cong G/C_G(\xi_0)$. □

Stable Kobayashi–Hitchin correspondence

The Kobayashi–Hitchin correspondence says that $(A,u) \in \mathcal{H}(P,X)$ is stable if and only if $w((A,u),\xi) > 0$ for all $\xi \in \Omega^0(\Sigma, \mathrm{ad}(P))\backslash\{0\}$. This was established by Mundet [89] in greater generality and we briefly recall his argument. Suppose (A,u) is stable and satisfies the vortex equation. Then

$$w((A,u),\xi) = \langle *F_A + \mu(u), \xi\rangle_{L^2} + \int_0^\infty ||\mathcal{L}_{(e^{\mathrm{i}t\xi}A, e^{\mathrm{i}t\xi}u)}\xi||^2_{L^2}\, dt \tag{3.53}$$

is positive. It is a less obvious fact that this condition is $\mathcal{G}^c(P)$-invariant and hence $w(g(A,u),\xi) > 0$ for every $g \in \mathcal{G}^c(P)$. The converse direction depends on the Kempf–Ness theorem. Mundet shows by contradiction when no estimate (3.52) holds, then there exists a destabilizing direction ξ with $w((A,u),\xi) \leq 0$. Once the estimate (3.52) is established, one obtains a solution to the vortex equation by direct methods of the calculus of variations.

Our proof of the polystable case in Theorem 3.6.5 below yields an alternative proof of the stable case under more restrictive assumptions.

Semistable Kobayashi–Hitchin correspondence

We need to assume the following technical property on the pair $(A,u) \in \mathcal{H}(P,X)$:

(H) For all $\xi \in \Omega^0(\Sigma, \mathrm{ad}(P))$ it holds:

$$w((A,u),\xi) \leq 0 \qquad \Longrightarrow \qquad \sup_{t>0} ||\mu(e^{\mathbf{i}t\xi}u)||_{L^2} < \infty.$$

We refer to Remark 3.1.8 for a discussion of this assumption. Following the ideas of Chen [19, 18], Chen–Sun [23] and Donaldson [39] we prove the following version of the moment weight inequality which is Theorem E in the introduction.

Theorem 3.6.3 (**Sharp moment weight inequality**). *Suppose $(A,u) \in \mathcal{H}(P,X)$ satisfies **(H)**. Then for all $\xi \in \Omega^0(\Sigma, ad(P))\backslash\{0\}$ it holds*

$$-\frac{w((A,u),\xi)}{||\xi||_{L^2}} \leq \inf_{g \in \mathcal{G}^c(P)} || * F_{gA} + \mu(gu)||_{L^2}. \tag{3.54}$$

*If in addition **(A)**, **(B)**, **(C)** are satisfied and the right hand side is positive, then there exists a unique $\xi_0 \in \Omega^0(\Sigma, ad(P))$ with $||\xi||_{L^2} = 1$ which yields equality.*

Proof. The proof is given in the next subsection on page 134. □

Theorem 3.6.4 (**Semistable correspondence**). *Assume **(A)**, **(B)**, **(C)** and suppose that $(A,u) \in \mathcal{H}(P,X)$ satisfies **(H)**. Then the following are equivalent:*

1. *(A,u) is semistable in the sense of Definition 3.1.1.*
2. *$\inf_{g \in \mathcal{G}^c(P)} || * F_{gA} + \mu(gu)||_{L^2} = 0$.*
3. *$w((A,u),\xi) \geq 0$ for all $\xi \in \Omega^0(\Sigma, ad(P))$.*

Proof. This is a direct consequence of Theorem 3.5.1 and Theorem 3.6.3. □

Polystable Kobayashi–Hitchin correspondence

Consider for $(A,u) \in \mathcal{H}(P,X)$ the following properties

(SS) For all $\xi \in \Omega^0(\Sigma, \mathrm{ad}(P))$ it holds $w((A,u),\xi) \geq 0$.

(PS1) For all $\xi \in \Omega^0(\Sigma, \mathrm{ad}(P))$ with $\exp(\xi) = \mathbb{1}$ and $(w(A,u), \xi) = 0$ the limit

$$\lim_{t\to\infty} e^{\mathbf{i}t\xi}(A,u) \in (\mathcal{G}^c)^{2,2}(A,u)$$

exists in $H^1 \times H^2$ and remains in the (Sobolev completion of the) complexified group orbit $(\mathcal{G}^c)^{2,2}(A,u)$.

(PS2) For all $\xi \in \Omega^0(\Sigma, \mathrm{ad}(P))$ with $(w(A,u), \xi) = 0$ the limit

$$\lim_{t\to\infty} e^{\mathbf{i}t\xi}(A,u) \in (\mathcal{G}^c)^{2,2}(A,u)$$

exists in $H^1 \times H^2$ and remains in the (Sobolev completion of the) complexified group orbit $\mathcal{G}^c(A,u)$.

Theorem 3.6.5 (Polystable correspondence). *Assume **(A)**, **(B)**, **(C)** and suppose that $(A,u) \in \mathcal{H}(P,X)$ satisfies **(H)**. Then the following are equivalent*

1. *(A,u) is polystable, i.e. there exits $g \in \mathcal{G}^c(P)$ such that $*F_{gA} + \mu(gu) = 0$.*
2. *(A,u) satisfies **(SS)** and **(PS1)**.*
3. *(A,u) satisfies **(SS)** and **(PS2)**.*

Proof. See page 140. □

Assumption **(H)** is only needed for the application of Theorem 3.6.4. For twisted Higgs-bundles over Riemann surfaces a polystable Kobayashi–Hitchin correspondence was established by García-Prada, Gothen and Mundet [49] by different methods. We present a more general proof following the ideas of Chen–Sun [23].

3.6.2 Proof of the moment-weight inequality

The purpose of this section is to prove Theorem 3.6.3. Section 3.6.2 contains the proof of the inequality (3.54). The proof is essentially due to Chen [19, 18] and Donaldson [39]. Section 3.6.2 contains a proof of the equality in the unstable case. This is the analog of the Kempf existence theorem in finite dimension. The proof is based on arguments given by Chen–Sun [23] in the finite dimensional differentiable case. Section 3.6.2 contains a proof of the uniqueness claim. This is the analogue of the Kempf uniqueness theorem. The proof is the one given in [51], Theorem 11.3, for the finite dimensional setting and extends almost ad verbum to our setting.

Proof of the inequality

Let $(A,u) \in \mathcal{H}(P,X)$, $g_0 \in \mathcal{G}^c(P)$ and $\xi \in \Omega^0(\Sigma, \mathrm{ad}(P))\backslash\{0\}$ be given and assume $w((A,u),\xi) \leq 0$. Define $\eta(t) \in \Omega^1(\Sigma, \mathrm{ad}(P))$ and $u(t) \in \mathcal{G}(P)$ by

$$g_0^{-1} = e^{-\mathbf{i}\xi t} e^{-\mathbf{i}\eta(t)} u(t). \tag{3.55}$$

Let $\pi : G^c \to G^c/G$ denote the canonical projection. Since the left-invariant metric on G^c/G has nonpositive curvature, the exponential map is distance increasing and it holds pointwise

$$||\xi t - \eta(t)|| \leq \mathrm{dist}_{G^c/G}(\pi(e^{\mathbf{i}\xi t}), \pi(e^{\mathbf{i}\eta(t)})) \leq \mathrm{dist}_{G^c/G}(\pi(\mathbb{1}), \pi(g_0^{-1})).$$

In particular, there exists $C > 0$ such that $||\xi t - \eta(t)||_{L^2} \leq C$ and this implies

$$\left|\left|\frac{\xi}{||\xi||_{L^2}} - \frac{\eta(t)}{||\eta(t)||_{L^2}}\right|\right|_{L^2} \leq \frac{C}{2t||\xi||_{L^2}} \tag{3.56}$$

Define

$$g : [0,1] \to \mathcal{G}^c(P), \qquad g(s) := g_0^{-1} \exp\left(\mathbf{i} s u(t)\eta(t)u^{-1}(t)\right).$$

Then $\gamma := \pi \circ g$ is the unique geodesic connecting $\pi(g_0^{-1})$ to $\pi(e^{-\mathbf{i}t\xi})$. It follows from (3.48) and the fact that $\Psi_{(A,u)}$ is convex along geodesics (3.51) that

$$\begin{aligned} -|| * F_{g_0 A} + \mu(g_0 u)||_{L^2} &\leq \frac{1}{||\eta||_{L^2}} \alpha_{(A,u)}\left(\gamma(0), \dot{\gamma}(0)\right) \\ &\leq \frac{1}{||\eta||_{L^2}} \alpha_{(A,u)}\left(\gamma(1), \dot{\gamma}(1)\right) \\ &= \left\langle *F_{e^{\mathbf{i}\xi t}A} + \mu(e^{\mathbf{i}\xi t}u), \frac{\eta(t)}{||\eta(t)||_{L^2}} \right\rangle_{L^2}. \end{aligned}$$

Assumption **(H)**, Proposition 3.6.2 and (3.56) show that the right-hand side converges to $\frac{w((A,u),\xi)}{||\xi||_{L^2}}$ and this completes the proof.

Existence of the dominant weight

Suppose that $(A_0, u_0) \in \mathcal{H}(P, X)$ is unstable. We prove in this section that there exists $\xi \in \Omega^0(\Sigma, \mathrm{ad}(P))$ such that

$$-\frac{w((A_0,u_0),\xi)}{||\xi||_{L^2}} = \inf_{g \in \mathcal{G}^c} || * F_{gA_0} + \mu(gu_0)||_{L^2}. \tag{3.57}$$

Let $(A, u) : [0, \infty) \to \mathcal{H}(P, X)$ be the solution of (3.36) starting at (A_0, u_0), let $g : [0, \infty) \to \mathcal{G}^c(P)$ be the solution of (3.38). Define $\xi(t) \in \Omega^0(\Sigma, \mathrm{ad}(P))$ and $k(t) \in \mathcal{G}(P)$ by

$$g(t) = e^{-\mathbf{i}\xi(t)} k(t). \tag{3.58}$$

The strategy of the proof is to show that the limit

$$\lim_{t\to\infty} \frac{\xi(t)}{t} =: \xi_\infty \tag{3.59}$$

exists in $W^{1,p}$ and satisfies (3.57).

Step 1: *The limit (3.59) exists in L^2.*

Denote by $\pi : G^c \to G^c/G$ the canonical projection and let $\gamma := \pi \circ g$. Since $g^{-1}\dot{g} = \mathbf{i}(*F_A + \mu(u))$ takes values in $\mathbf{i}\mathfrak{g}$, it holds $\nabla_t \dot{\gamma} = d\pi(g)\mathbf{i}(\partial_t(g^{-1}\dot{g}))$ and Theorem 3.4.8 yields the estimate

$$\int_T^\infty ||\nabla_t \dot{\gamma}(t)||_{L^2}\, dt \leq C \int_T^\infty ||\partial_t(A,u)||_{H^1}\, dt \leq C T^{-\epsilon}. \tag{3.60}$$

Define $\gamma_t : [0,t] \to \mathcal{G}^c(P)/\mathcal{G}(P)$ by $\gamma_t(s) := \pi\left(e^{-\mathbf{i}\frac{\xi(t)}{t}s}\right)$. Pointwise this is the geodesic segment connection $\pi(\mathbb{1})$ to $\gamma(t)$ and we define

$$\rho_t(s) : \Sigma \to \mathbb{R}, \qquad \rho_t(s) = \mathrm{dist}_{G^c/G}(\gamma(s), \gamma_t(s)).$$

Since G^c/G has nonpositve sectional curvature, there holds pointwise the estimate $\ddot{\rho}_t(s) \geq -||\nabla\dot{\gamma}(s)||$ (see [51] Appendix A). Hence (3.60) yields

$$||\dot{\rho}_t(s)||_{L^2} \leq \int_s^\infty ||\nabla_t\dot{\gamma}(t)||_{L^2} \leq Cs^{-\epsilon} \tag{3.61}$$

and integrating this estimate shows

$$||\rho_t(s)||_{L^2} \leq \int_0^s ||\dot{\rho}_t(r)||_{L^2}\, dr \leq Cs^{1-\epsilon}. \tag{3.62}$$

Since the exponential map on G^c/G is distance increasing, it follows pointwise for $0 < t_1 < t_2$

$$\left|\left|\frac{\xi(t_1)}{t_1} - \frac{\xi(t_2)}{t_2}\right|\right| \leq ||\dot{\gamma}_{t_1}(0) - \dot{\gamma}_{t_2}(0)|| \leq \frac{\rho_{t_2}(t_1)}{t_1} \tag{3.63}$$

Now (3.62) and (3.63) show that $\frac{\xi(t)}{t}$ is a L^2-Cauchy sequence and the limit (3.59) exists in L^2.

Step 2: *The limit (3.59) exists in $W^{1,p}$ for every $p \in (2,\infty)$.*

Let $\xi(t)$ be as in (3.58) and define

$$R(t) := *F_{e^{\mathbf{i}\xi(t)}A_0} - *F_{A_0} + \mu(e^{\mathbf{i}\xi(t)}u_0) - \mu(u_0).$$

A similar calculation as in the proof of Theorem 3.5.1 shows

$$\begin{aligned} 2\langle R(t), \xi(t)\rangle &= \Delta||\xi(t)||^2 + 2\int_0^1 \left(||d_{e^{\mathbf{i}s\xi(t)}A_0}\xi(t)||^2 + ||L_{e^{\mathbf{i}s\xi(t)}u_0}\xi(t)||^2\right)\, ds \\ &\geq 2||\xi(t)||\Delta||\xi(t)||. \end{aligned}$$

Thus $||\xi(t)|| : \Sigma \to [0,\infty)$ are positive functions satisfying $\Delta\xi(t) \leq ||R(t)||$ at points where $\xi(t) \neq 0$. An argument of Donaldson [32] (see [103] Prop 2.1) using the mean-value property of harmonic functions shows that this implies an estimate

$$||\xi(t)||_{C^0} \leq C\left(1 + ||R(t)||_{L^2} + ||\xi(t)||_{L^1}\right). \tag{3.64}$$

Since $(A(t), u(t))$ satisfies (3.36) and

$$||*F_{e^{\mathbf{i}\xi(t)}A_0} + \mu(e^{\mathbf{i}\xi(t)}u_0)||_{L^2} = ||*F_{A(t)} + \mu(u(t))||_{L^2}$$

the term $||R(t)||_{L^2}$ is uniformly bounded and (3.64) simplifies to

$$||\xi(t)||_{C^0} \leq C(1 + ||\xi(t)||_{L^1}). \tag{3.65}$$

In particular, $\frac{\xi(t)}{t}$ is uniformly bounded in C^0. Since $g(t)^{-1}\bar{\partial}_{A_0}g(t) = (A_0 - A(t))^{0,1}$ is uniformly bounded in H^1 and pointwise

$$||g^{-1}(t)\bar{\partial}_{A_0}g(t)||^2 = ||e^{\mathbf{i}\xi(t)}\bar{\partial}_{A_0}e^{-\mathbf{i}\xi(t)}||^2 + ||(\bar{\partial}_{A_0}k(t))k(t)^{-1}||^2$$

it follows that $e^{\mathbf{i}\xi(t)}\bar{\partial}_{A_0}e^{-\mathbf{i}\xi(t)}$ is uniformly bounded in L^p for every $p \in (1,\infty)$. Now

$$e^{\mathbf{i}\xi(t)}\bar{\partial}_{A_0}e^{-\mathbf{i}\xi(t)} = te^{\mathbf{i}\xi(t)/t}\bar{\partial}_{A_0}e^{-\mathbf{i}\xi(t)/t}$$

implies that $e^{\mathbf{i}\xi(t)/t}\bar{\partial}_{A_0}e^{-\mathbf{i}\xi(t)/t}$ converges to zero in L^p and by elliptic regularity the limit (3.59) exists in $W^{1,p}$.

Step 3: *The limit ξ_∞ defined by (3.59) yields equality in (3.54).*

The Kempf–Ness functional (3.49) satisfies $\Psi_{(A_0,u_0)}(\mathbb{1}) = 0$, decreases along $\gamma(t)$ and is convex along geodesics. Hence $\Psi_{(A_0,u_0)}(e^{\mathbf{i}s\xi(t)}) \leq 0$ for $0 < s < t$ and by continuity with respect to the $W^{1,p}$-topology, it takes nonpositive values along the geodesic ray $\gamma_\infty(t) := \pi\left(e^{\mathbf{i}\xi_\infty t}\right)$. This implies $w((A_0,u_0),\xi_\infty) \leq 0$ and ξ_∞ is smooth by Proposition 3.6.2. Using again that $\Psi_{(A_0,u_0)}$ is convex along geodesics it follows

$$\left|\Psi_{(A_0,u_0)}(\gamma(t)) - \Psi_{(A_0,u_0)}(\gamma_\infty(t))\right| \leq M \cdot \mathrm{dist}_{\mathcal{G}^c/\mathcal{G}}(\gamma(t),\gamma_\infty(t)) \tag{3.66}$$

where $\mathrm{dist}_{\mathcal{G}^c/\mathcal{G}}$ denotes the L^2-geodesic distance and

$$M := \sup_{t>0}\max\left\{||*F_{g(t)^{-1}A_0} + \mu(g(t)^{-1}u_0)||_{L^2},\ ||*F_{e^{\mathbf{i}\xi_\infty t}A_0} + \mu(e^{\mathbf{i}t\xi}u_0)||_{L^2}\right\}.$$

which is finite by (**H**) and Proposition 3.6.2. As $t \to \infty$ in (3.62) one obtains $\mathrm{dist}_{\mathcal{G}^c/\mathcal{G}}(\gamma(t),\gamma_\infty(t)) \leq Ct^{1-\epsilon}$ and hence

$$\left|\Psi_{(A_0,u_0)}(\gamma(t)) - \Psi_{(A_0,u_0)}\gamma_\infty(t)\right| \leq Ct^{1-\epsilon}. \tag{3.67}$$

Then

$$\begin{aligned}
-w((A_0,u_0),\xi_\infty) &= \lim_{t\to\infty}\frac{1}{t}\int_0^t \langle e^{-\mathbf{i}\xi_\infty s}(A_0,u_0),\xi_\infty\rangle\,ds \\
&= \lim_{t\to\infty}\frac{\Psi_{(A_0,u_0)}(\gamma_\infty(t))}{t} \\
&= \lim_{t\to\infty}\frac{\Psi_{(A_0,u_0)}(\gamma(t))}{t} \\
&= \lim_{t\to\infty}\frac{1}{t}\int_0^t ||*F_{A(s)} + \mu(u(s))||_{L^2}^2\,ds \\
&= ||*F_{A_\infty} + \mu(u_\infty)||_{L^2}^2
\end{aligned}$$

By Theorem 3.5.2

$$||*F_{A_\infty} + \mu(u_\infty)||_{L^2} = \inf_{g\in\mathcal{G}^c}||*F_{gA_0} + \mu(gu_0)|| =: m$$

and thus $-w((A_0,u_0),\xi_\infty) = m^2$. Now

$$||\xi_\infty||_{L^2} = \lim_{t\to\infty}\left|\left|\frac{\xi(t)}{t}\right|\right|_{L^2} \leq \lim_{t\to\infty}\frac{1}{t}\int_0^t ||\dot{\gamma}(s)||_{L^2}\,ds = ||*F_{A_\infty} + \mu(u_\infty)||_{L^2} = m$$

shows $-w(A,\xi_\infty)/||\xi_\infty||_{L^2} \geq m$ and the converse inequality follows from (3.54).

Uniqueness of the dominant weight

Suppose that $(A,u) \in \mathcal{H}(P,X)$ is unstable and $\xi_0, \xi_1 \in \Omega^0(\Sigma, \mathrm{ad}(P))$ satisfy $||\xi_1||_{L^2} = ||\xi_2||_{L^2} = 1$ and

$$-w((A,u),\xi_1) = -w((A,u),\xi_2) = \inf_{g \in \mathcal{G}^c(P)} ||*F_{gA} + \mu(gu)||_{L^2} =: m > 0.$$

We prove in the following that this implies $\xi_1 = \xi_2$.

Define $\eta(t) \in \Omega^0(\Sigma, \mathrm{ad}(P))$ and $k(t) \in \mathcal{G}(P)$ by

$$e^{-\mathbf{i}t\xi_0} e^{\mathbf{i}\eta(t)} = e^{-\mathbf{i}t\xi_1} k(t). \tag{3.68}$$

Let $\pi : G^c \to G^c/G$ denote the canonical projection and let $p(t) := \pi(e^{-\mathbf{i}t\xi_0} e^{\mathbf{i}\eta(t)/2})$ denote the midpoint between the geodesic rays spanned by ξ_1 and ξ_2. Since G^c/G has nonpositive curvature, the exponential map (based at $p(t)$) is distance increasing and this yields

$$\begin{aligned} d(1,p(t))^2 &\le \frac{d(\pi(\mathbb{1}),\pi(e^{\mathbf{i}\xi_1 t}))^2 + d(\pi(\mathbb{1}),\pi(e^{\mathbf{i}\xi_2 t}))^2}{2} - \frac{d(\pi(e^{\mathbf{i}\xi_1 t}),\pi(e^{\mathbf{i}\xi_2 t}))^2}{4} \\ &\le t^2 \left(1 - \frac{||\xi_1 - \xi_2||_{L^2}}{4}\right) \end{aligned}$$

where $d(\cdot,\cdot) = \mathrm{dist}_{\mathcal{G}^c/\mathcal{G}}(\cdot,\cdot)$ denotes the L^2-geodesic distance. Since $\Psi_{(A,u)}$ is convex along geodesics, it follows $\Psi_{(A,u)}(p(t)) \le -tm$ and hence

$$\frac{\Psi_{(A,u)}(p(t))}{d(1,p(t))} \le \frac{-m}{\sqrt{1 - ||\xi_0 - \xi_1||_{L^2}/4}}. \tag{3.69}$$

Denote for $r > 0$

$$S_r := \left\{ \pi(e^{i\xi}) \;\middle|\; \xi \in \Omega^0(\Sigma, \mathrm{ad}(P)),\ ||\xi||_{L^2} = r \right\} \subset \mathcal{G}^c(P)/\mathcal{G}(P).$$

We claim

$$\lim_{r\to\infty} \frac{1}{r} \inf_{S_r} \Psi_{(A,u)} = -m. \tag{3.70}$$

As $t \to \infty$ in (3.69) the claim implies $\xi_1 = \xi_2$. The inequality "$\le$" in (3.70) follows by considering the values along the geodesics ray $\pi(e^{\mathbf{i}\xi_1 t})$. For the other direction let $h \in \mathcal{G}^c(P)$ be given and using (3.51) one estimates

$$\Psi_{h^{-1}(A,u)}(g) \ge -||*F_{h^{-1}A} + \mu(h^{-1}u)||_{L^2} \cdot d(\pi(\mathbb{1}),\pi(g))$$

Suppose h is chosen such that $||*F_{h^{-1}A} + \mu(h^{-1}u)||_{L^2} \le m + \epsilon$. Then

$$\begin{aligned} \Psi_{(A,u)}(g) &= \Psi_{h^{-1}(A,u)}(h^{-1}g) + \Psi_{(A_0,u_0)}(h) \\ &\ge (-m-\epsilon) d(\pi(\mathbb{1}),\pi(h^{-1}g)) + \Psi_{(A_0,u_0)}(h) \\ &\ge (-m-\epsilon) d(\pi(\mathbb{1}),\pi(g)) + (-m-\epsilon) d(\pi(\mathbb{1}),\pi(h^{-1})) + \Psi_{(A_0,u_0)}(h) \end{aligned}$$

and as $d(\pi(\mathbb{1}),\pi(g)) \to \infty$ and $\epsilon \to 0$ this proves (3.70).

3.6.3 Proof of the polystable correspondence

The purpose of this section is to prove Theorem 3.6.5.

Proposition 3.6.6. *Let $(A, u) \in \mathcal{H}(P, X)$ be polystable, then (A, u) satisfies **(SS)** and **(PS2)**.*

Proof of Proposition 3.6.6. Choose $g \in \mathcal{G}^c(P)$ such that $*F_{gA} + \mu(gu) = 0$. Then

$$w(g(A,u),\xi) = \langle *F_{gA} + \mu(gu), \xi\rangle_{L^2} + \int_0^\infty ||\mathcal{L}_{(e^{\mathbf{i}t\xi}gA, e^{\mathbf{i}t\xi}gu)}\xi||_{L^2}^2 \, dt \tag{3.71}$$

shows $w(g(A,u),\xi) \geq 0$. Equality holds if and only if $\mathcal{L}_{g(A,u)}\xi = 0$ and $e^{\mathbf{i}t\xi}g(A,u) = g(A,u)$ is constant. In particular, $g(A,u)$ satisfies **(SS)** and **(PS2)**. The proposition follows now from Lemma 3.6.7 below. □

Lemma 3.6.7. *Let $(A, u) \in \mathcal{H}(P, X)$ and let $(B, v) \in \mathcal{G}^c(A, u)$.*

1. *If (A, u) satisfies **(SS)** and **(PS1)** then (B, v) satisfies **(SS)** and **(PS1)**.*
2. *If (A, u) satisfies **(SS)** and **(PS2)** then (B, v) satisfies **(SS)** and **(PS2)**.*

Proof. We prove the second part first. Choose $g \in \mathcal{G}^c(P)$ such that $(B,v) = g(A,u)$ and let $\xi \in \Omega^0(\Sigma, \mathrm{ad}(P))$ be such that $w(g(A,u),\xi) \leq 0$. Let $\xi_0 \in \mathfrak{g}$ and $P_Q \subset P^c$ be the $Q(\xi_0)$-bundle determined by ξ as asserted in Proposition 3.6.2. It is possible to decompose $g = qk$ with $q \in \mathcal{G}(P_Q)$ and $k \in \mathcal{G}(P)$ (e.g. by using the identity $G^c/B = G/Z(G)$ for any Borel subgroup $B \subset Q(\xi_0)$). By definition of $Q(\xi_0)$

$$q_+ := \lim_{t\to\infty} e^{\mathbf{i}t\xi} q e^{-\mathbf{i}t\xi} \tag{3.72}$$

Using the assumption $w(qk(A,u),\xi) \leq 0$ it follows for $t > 0$

$$0 \geq \Psi_{qk(A,u)}(e^{-\mathbf{i}t\xi}) = \Psi_{k(A,u)}(q^{-1}e^{-\mathbf{i}t\xi}) - \Psi_{k(A,u)}(q^{-1}). \tag{3.73}$$

Let $\pi : G^c \to G^c/G$ deonte the canoncial projection. Then

$$\mathrm{dist}_{\mathcal{G}^c/\mathcal{G}}\left(\pi\left(e^{-\mathbf{i}t\xi}\right), \pi\left(q^{-1}e^{-\mathbf{i}t\xi}\right)\right) = \mathrm{dist}_{\mathcal{G}^c/\mathcal{G}}\left(\pi(\mathbb{1}), \pi\left(e^{\mathbf{i}t\xi}q^{-1}e^{-\mathbf{i}t\xi}\right)\right) \leq C \tag{3.74}$$

which is bounded by (3.72). For $t > s > 0$ define $\eta_{s,t} \in \Omega^0(\Sigma, \mathrm{ad}(P))$ and $k_{s,t} \in \mathcal{G}(P)$ by

$$e^{-\mathbf{i}s\xi}e^{\mathbf{i}\eta_{s,t}} = q^{-1}e^{-\mathbf{i}t\xi}k_{s,t}.$$

Since the exponential map in G^c/G is distance increasing, it follows from (3.74)

$$\lim_{t\to\infty} \left|\left|\frac{\eta_{s,t}}{t-s} - \xi\right|\right|_{L^2} = 0. \tag{3.75}$$

If $\sup\{\Psi_{k(A,u)}(e^{-\mathbf{i}t\xi}) \,|\, t > 0\} < \infty$, then clearly $w(k(A,u),\xi) \leq 0$. Otherwise, (3.73) shows that for all sufficently large $s > 0$ and every $t > s$ we have

$$\Psi_{k(A,u)}(e^{-\mathbf{i}s\xi}) > \Psi_{k(A,u)}(q^{-1}e^{-\mathbf{i}t\xi}).$$

Since $\Psi_{k(A,u)}$ is convex along the geodesic segement $r \mapsto e^{-\mathbf{i}s\xi} \cdot e^{\mathbf{i}\eta_{s,t}r}$, it follows

$$d\Psi_{k(A,u)}\left(\pi(e^{-\mathbf{i}s\xi}); d\pi(e^{-\mathbf{i}s\xi})\mathbf{i}\frac{\eta_{s,t}}{t-s}\right) = \left\langle *F_{e^{\mathbf{i}s\xi}kA} + \mu(e^{-\mathbf{i}s\xi}ku), \frac{\eta_{s,t}}{t-s}\right\rangle < 0.$$

Now (3.75) implies $\left\langle F_{e^{\mathbf{i}s\xi}A} + \mu(e^{\mathbf{i}s\xi}u); \xi\right\rangle \leq 0$ for all sufficiently large s and hence $w(k(A,u),\xi) \leq 0$. Since (A,u) satisfies **(SS)** by assumption, it follows $w((A,u),k^{-1}\xi k) = 0$ and **(PS2)** implies that the limit

$$(A_+,u_+) := \lim_{t\to\infty} e^{\mathbf{i}tk^{-1}\xi k}(A,u)$$

exists in $H^1 \times H^2$ and $(A_+,u_+) \in (\mathcal{G}^c)^{2,2}(A,u)$. Hence

$$(B_+,v_+) := \lim_{t\to\infty} e^{\mathbf{i}t\xi}(B,v) = \lim_{t\to\infty} e^{\mathbf{i}t\xi}qe^{-\mathbf{i}t\xi}ke^{\mathbf{i}tk^{-1}\xi k}(A,u) = q_+k(A_+,u_+)$$

exists and $(B_+,v_+) \in (\mathcal{G}^c)^{2,2}(B,v)$. Moreover,

$$w((B,v),\xi) = \langle *F_{q_+kA_+} + \mu(q_+ku_+),\xi\rangle$$

and it remains to verify that this vanishes. By Proposition 3.6.2, there exists a reduction $P_K \subset P$ to the centralizer $K = C_G(\xi_0)$ and q_+ restricts to an element in $\mathcal{G}^c(P_K)$. Let $h : [0,1] \to \mathcal{G}^c(P_K)$ be a smooth path connecting $\mathbb{1}$ to q_+ with $h^{-1}(t)\dot{h}(t) = \alpha(t) + \mathbf{i}\beta(t)$. A short calculation shows

$$\begin{aligned}
&\partial_t\langle *F_{h(t)kA_+} + \mu(h(t)ku_+),\xi\rangle \\
&\quad = \left\langle -\left[*F_{h(t)kA_+} + \mu(h(t)ku_+),\alpha(t)\right],\xi\right\rangle + \left\langle \mathcal{L}_{h(t)k(A_+,u_+)}\beta(t), \mathcal{L}_{h(t)k(A_+,u_+)}\xi\right\rangle \\
&\quad = 0
\end{aligned}$$

where the last step uses $[\alpha(t),\xi] = 0$ and $\mathcal{L}_{h(t)k(A_+,u_+)}\xi = h(t)\mathcal{L}_{k(A,u)}\xi = 0$. Hence

$$\begin{aligned}
w((B,v),\xi) &= \langle *F_{kA_+} + \mu(ku_+),\xi\rangle = \langle *F_{A_+}, \mu(u_+), k^{-1}\xi k\rangle \\
&= w((A,u),k^{-1}\xi k) = 0
\end{aligned}$$

and this completes the proof of the second part.

The first part follows from the same argument, since $\exp(\xi) = \mathbb{1}$ implies $\exp(k^{-1}\xi k) = \mathbb{1}$. □

Proof of Theorem 3.6.5. If (A,u) is polystable then it satisfies **(SS)** and **(PS2)** by Proposition 3.6.6. For the converse direction let $(A_0,u_0) \in \mathcal{H}(P,X)$ be given and assume that it satisfies **(SS)** and **(PS1)**. Denote by $(A,u) : [0,\infty) \to \mathcal{H}(P,X)$ the solution of (3.36) starting at (A_0,u_0) with limit (A_∞,u_∞). Theorem 3.6.4 shows that (A_0,u_0) is semistable and hence by Theorem 3.5.2 the limit solves $*F_{A_\infty} + \mu(u_\infty) = 0$. Denote the isotropy groups at the limit and their Lie algebras by

$$H := \{h \in \mathcal{G}^{2,2}(P) \,|\, (hA_\infty, hu_\infty) = (A_\infty,u_\infty)\}, \qquad \mathfrak{h} = \ker(\mathcal{L}_{A_\infty,u_\infty})$$

$$H^c := \{h \in (\mathcal{G}^c)^{2,2}(P) \,|\, (hA_\infty, hu_\infty) = (A_\infty,u_\infty)\}, \qquad \mathfrak{h}^c = \ker(\mathcal{L}^c_{(A_\infty,u_\infty)}).$$

By Theorem 3.5.2 we may assume that these groups are not discrete. There are two important properties to note: (1) By Lemma 3.2.10 A_∞ is gauge equivalent to a smooth connection. In particular, as a subgroup of the isotropy group of A_∞, one can identify H with a closed and hence compact subgroup of G. (2) Using the equation $*F_{A_\infty} + \mu(u_\infty) = 0$, a short calculation shows that H^c is indeed the complexification of H.

Step 1: *There exists an H-invariant holomorphic coordinate chart*

$$\psi : (T_{(A_\infty,u_\infty)}\mathcal{A}^{1,2}(P) \times \mathcal{S}^{2,2}(P,X), 0) \to \left(\mathcal{A}^{1,2}(P) \times \mathcal{S}^{2,2}(P,X), (A_\infty, u_\infty)\right)$$

defined on a neighborhood of the origin satisfying $d\psi(0,0) = id$.

Let $\{g_p\}_{p\in P}$ be a smooth G-invariant family of Riemannian metrics on X, compatible with the holomorphic structure, such that g_p is flat in a neighborhood of $u_\infty(p)$. Then $\exp_{g_p} : T_{u_\infty(p)}X \to X$ is holomorphic in a neighbourhood of the origin and $\phi_{\mathcal{S}}(\hat{u})(p) := \exp_{g_p}(\hat{u}(p))$ provides a holomorphic chart for $\mathcal{S}(P,X)$. Define

$$\psi(a,\hat{u}) := \int_H h^{-1}(A_\infty + hah^{-1})\, d\mu_H(h) + \phi_{\mathcal{S}} \int_H \phi_{\mathcal{S}}^{-1}\left(h^{-1}\phi_{\mathcal{S}}(h\hat{u})\right)\right)\, d\mu_H(h)$$

where μ_H denotes the Haar-measure on H with $\mu_H(H) = 1$. This is well-defined for $||\hat{u}||_{L^\infty} \leq c||\hat{u}||_{H^2}$ sufficiently small and satisfies the desired properties.

Step 2: *The linearization of the holomorphicity condition $\bar{\partial}_A u = 0$ is the operator $D : H^1(\Sigma, ad(P)) \oplus H^2(\Sigma, u_\infty^* TX) \to H^1(\Sigma, \Lambda^{0,1} \otimes u_\infty^* TX)$*

$$D(a,\hat{u}) = \left(\nabla_{A_\infty}\hat{u} + L_{u_\infty} a\right)^{0,1}.$$

There exists an H-invariant holomorphic coordinate chart ψ as in Step 1 with the additional property that

$$\psi(a,\hat{u}) \in \mathcal{H}^{1,2}(P,X) \qquad \Longrightarrow \qquad D(a,\hat{u}) = 0$$

for every pair $(a,\hat{u})$ in the domain of ψ.

Since ∇_{A_∞} is a Fredholm operator with closed range and finite dimensional cokernel, it follows that the image of D is closed with finite codimension. Now any choice of complements for the kernel and image of D yield a pseudoinverse

$$T : H^1(\Sigma, \Lambda^{0,1} \otimes u_\infty^* TX) \to H^1(\Sigma, \mathrm{ad}(P)) \oplus H^2(\Sigma, u_\infty^* TX)$$

which is a bounded linear operator satisfying $DTD = D$ and $TDT = T$. Since D is complex linear we can choose complex complements to obtain a complex linear pseudoinverse T. Moreover, D is H-equivariant and for every $h \in H$ the operator $T_h := hTh^{-1}$ yields another complex linear pseudoinverse for D. The average

$$\int_H \int_H T_{h_1} D T_{h_2} d\mu_H(h_1) d\mu_H(h_2)$$

with respect to the Haar measure μ_H provides a H-equivariant pseudoinverse.

Let ψ be defined as in Step 1 and let T be a complex linear H-invariant pseudoinverse of D. Consider on the domain of ψ the map $\tilde{f}(a,\hat{u}) := f(\psi(a,\hat{u}))$ where $f(A,u) := \bar{\partial}_A u$. The map

$$\theta(a,\hat{u}) := (a,\hat{u}) + T(\tilde{f}(a,\hat{u}) - D(a,\hat{u}))$$

satisfies $\theta(0) = 0$ and $d\theta(0) = \mathbb{1}$. Hence there exists a local holomorphic inverse θ^{-1} around the origin by the inverse function theorem. It follows from the construction that $\tilde{f}_0 := (\mathbb{1} - DT) \circ \tilde{f} \circ \theta^{-1}$ takes values in $\ker(T)$ and

$$\tilde{f} \circ \theta^{-1} = \tilde{f}_0 + D.$$

Since $\ker(T)$ is a complement of $\mathrm{Im}(D)$ this implies

$$\tilde{f} \circ \theta^{-1}(a,\hat{u}) = 0 \qquad \Longleftrightarrow \qquad D(a,\hat{u}) = 0, \quad \tilde{f}_0(a,\hat{u}) = 0.$$

Step 2 follows from this discussion after replacing ψ by $\psi \circ \theta^{-1}$.

Step 3: *Denote by $\mathfrak{h}^\perp$ the L^2-orthogonal complement of $\mathfrak{h}$ in $H^2(\Sigma, ad(P))$ and by V the L^2-orthogonal complement of the image of $\mathcal{L}^c_{A_\infty,u_\infty}$. Then there exists $t_0 > 0$ and maps*

$$(a,\hat{u}) : [t_0,\infty) \to \mathit{ker}(D) \cap V, \qquad \xi,\eta : [t_0,\infty) \to \mathfrak{h}^\perp$$

such that $(a(t),\hat{u}(t))$ is in the domain of the chart ψ constructed in Step 2 and

$$(A(t),u(t)) = e^{\boldsymbol{i}\eta(t)} e^{\xi(t)} \psi(a(t),\hat{u}(t)) \tag{3.76}$$

for all $t > t_0$.

The map $\mathfrak{h}^\perp \times \mathfrak{h}^\perp \times V \to \mathcal{A}^{1,2}(P) \times \mathcal{S}^{2,2}(P,X)$ defined by $(\xi,\eta,(a,\hat{u})) \mapsto e^{\mathbf{i}\eta} e^{\xi} \psi(a,\hat{u})$ is smooth near the origin with invertible derivative. Step 3 follows now from the implicit function theorem and Step 2.

Step 4: *Let $g : [0,\infty) \to \mathcal{G}^c(P)$ be the solution of the equation $g^{-1}\dot{g} = \boldsymbol{i}(*F_{A(t)} + \mu(u(t)))$ with $g(0) = \mathbb{1}$ obtained in Theorem 3.4.3. There exists $t_1 \geq t_0$ with the following significance:*

$$h : [t_1,\infty) \to (\mathcal{G}^c)^{2,2}(P), \qquad h(t) := e^{-\boldsymbol{i}\xi(t)} e^{-\boldsymbol{i}\eta(t)} g^{-1}(t) g(t_1) e^{\boldsymbol{i}\eta(t_1)} e^{\boldsymbol{i}\xi(t_1)}$$

satisfies $h(t) \in H^c$ and $(a(t),\hat{u}(t)) = h(t)^{-1}(a(t_1),\hat{u}(t_1))$ for every $t \geq t_1$.

Let $t_1 \geq t_0$ be fixed. Rewrite the identity $(A(t),u(t)) = g(t)^{-1}(A_0,u_0)$ as $\psi(a(t),\hat{u}(t)) = h(t)\psi(a(t_1),\hat{u}(t_1))$ and differentiate this to obtain

$$d\psi(a(t),\hat{u}(t))[\partial_t(a(t),\hat{u}(t))] = \partial_t(\psi(a(t),\hat{u}(t))) = \mathcal{L}^c_{\psi(a(t),\hat{u}(t))} \dot{h}(t) h^{-1}(t). \tag{3.77}$$

For $(a,\hat{u}) \in V$ consider the operator

$$N_{(a,\hat{u})} : H^2(\Sigma, \mathrm{ad}(P)^c) \times V \to T_{(A_\infty,u_\infty)}(\mathcal{A}^{1,2}(P) \times \mathcal{S}^{2,2}(P,X))$$

$$N_{(a,\hat{u})}(\zeta,(b,\hat{v})) := \mathcal{L}^c_{\psi(a,\hat{u})}\zeta - d\psi(a,\hat{u})[b,\hat{v}].$$

Then (3.77) can be reformulated as

$$(h(t)^{-1}\dot{h}(t), \partial_t(a(t),\hat{u}(t))) \in \ker(N_{(a(t),\hat{u}(t))}). \tag{3.78}$$

Since $N_{(0,0)}$ is surjective with kernel $\mathfrak{h}^c$, it follows that $N_{(a,\hat{u})}$ is a surjective Fredholm operator with index $\dim(\mathfrak{h}^c)$ for $||a||_{H^1} + ||\hat{u}||_{H^2}$ sufficiently small. For $\xi \in \mathfrak{h}$ it holds

$$\left.\frac{d}{ds}\right|_{s=0} \psi(e^{\xi s}(a,\hat{u})) = \left.\frac{d}{ds}\right|_{s=0} e^{\xi s}\psi(a,\hat{u}) = \mathcal{L}_{\psi(a,\hat{u})}\xi.$$

Since V is H-invariant this shows $\mathcal{L}_{\psi(a,\hat{u})}\mathfrak{h} \subset d\psi(a,\hat{u})V$. Moreover, since ψ is holomorphic and V a complex subspaces, it follows $\mathcal{L}^c_{\psi(a,\hat{u})}\mathfrak{h}^c \subset d\psi(a,\hat{u})V$. This implies that the kernel of $N_{(a,\hat{u})}$ projects onto $\mathfrak{h}^c$. For sufficiently large t_1 the same is true for all operators $N_{(a(t),\hat{u}(t))}$ with $t \geq t_1$. Then (3.78) shows $h^{-1}(t)\dot{h}(t) \in \mathfrak{h}^c$ for all $t \geq t_1$ and hence $h(t) \in H^c$. Since ψ is holomorphic and H-equivariant this completes the proof of Step 4.

Step 5: *There exists* $\xi_0 \in \mathfrak{h}$ *with* $\exp\xi_0 = \mathbb{1}$ *such that* $w(h(t_1)^{-1}(A(t_1),u(t_1)),\xi_0) = 0$ *and*

$$\lim_{t\to\infty} e^{\mathbf{i}t\xi_0}h(t_1)^{-1}(A(t_1),u(t_1)) = (A_\infty,u_\infty). \tag{3.79}$$

In particular, $(A_\infty,u_\infty) \in (\mathcal{G}^c)^{2,2}(A_0,u_0)$ *and* (A_0,u_0) *is polystable.*

The group H acts on the finite dimensional vector space $X_0 := V \cap \ker(D)$ by unitary automorphism. Step 4 shows that the origin is contained in the closure of the H^c-orbit of $(a(t_1),\hat{u}(t_1))$. The classical Hilbert-Mumford criterion (see [51] Theorem 14.2) shows that there exists $\xi_0 \in \mathfrak{h}$ with $\exp\xi_0 = \mathbb{1}$ such that

$$\lim_{t\to\infty} e^{\mathbf{i}t\xi_0}(a(t_1),\hat{u}(t_1)) = 0.$$

Since $\psi(e^{\mathbf{i}t\xi_0}(a(t_1),\hat{u}(t_1)) = e^{\mathbf{i}t\xi_0}\psi(a(t_1),\hat{u}(t_1)) = e^{\mathbf{i}t\xi_0}h(t_1)^{-1}(A(t_1),u(t_1))$ for all $t \geq 0$, it follows

$$\lim_{t\to\infty} e^{\mathbf{i}t\xi_0}h(t_1)^{-1}(A(t_1),u(t_1)) = \psi(0) = (A_\infty,u_\infty)$$

and $w(h(t_1)^{-1}(A(t_1),u(t_1)),\xi_0) = \langle *F_{A_\infty} + \mu(u_\infty),\xi_0\rangle = 0$.

By Lemma 3.2.10 there exists $k \in \mathcal{G}^{2,2}(P)$ such that $k(A_\infty,u_\infty)$ is smooth. Then $kh(t_1)^{-1} \in \ker(\mathcal{L}^c_{k(A_\infty,u_\infty)}$ is smooth and

$$w(kh(t_1)^{-1}(A(t_1),u(t_1)),k\xi_0k^{-1}) = w(h(t_1)^{-1}(A(t_1),u(t_1)),\xi_0) = 0$$

By Lemma 3.6.7 $kh(t_1)^{-1}(A(t_1),u(t_1))$ satisfies **(PS1)** and together with (3.79) this yields

$$\begin{aligned}(A_\infty,u_\infty) &= \lim_{t\to\infty} e^{\mathbf{i}t\xi_0}h(t_1)^{-1}(A(t_1),u(t_1)) \\ &= k^{-1}\lim_{t\to\infty} e^{k\xi_0k^{-1}}kh(t_1)^{-1}(A(t_1),u(t_1)) \in (\mathcal{G}^c)^{2,2}(A_0,u_0)\end{aligned}$$

Hence (A_0,u_0) is polystable by Theorem 3.5.2 and this completes the proof.

□

Chapter 4

Donaldson's moment map approach to Teichmüller theory

This chapter provides a self-contained exposition of a general moment map framework for the diffeomorphism group introduced by Donaldson [38].
The main applications considered in this chapter is the construction of a hyperkähler moduli space $\mathcal{M}$ associated to a closed oriented surface Σ with genus$(\Sigma) \geq 2$. This embeds naturally into the cotangent bundle $T^*\mathcal{T}(\Sigma)$ and can be viewed as the Feix–Kaledin hyperkähler extension of the Weil–Petersson metric on Teichmüller space. Donaldson outlined various remarkable properties of this moduli space for which we provide complete proofs: The moduli space $\mathcal{M}$ parametrizes the class of almost-Fuchsian 3-manifolds. These are quasi-Fuchsian 3-manifolds which contain a unique minimal surface with principal curvatures in $(-1,1)$. The area of this minimal surface then provides a Kähler potential for the hyperkähler metric. Moreover, the moduli space $\mathcal{M}$ embeds naturally into the $\mathrm{SL}(2,\mathbb{C})$-representation variety of Σ and the hyperkähler structure on $\mathcal{M}$ extends the Goldman holomorphic symplectic structure on the representation variety. The various identifications are obtained using the work of Uhlenbecks [117] on germs of hyperbolic 3-manifolds, an explicit map from $\mathcal{M}$ to $\mathcal{T}(\Sigma) \times \overline{\mathcal{T}(\Sigma)}$ found by Hodge [61], the simultaneous uniformization theorem of Bers [8], and the theory of Higgs bundles introduced by Hitchin [58].
Another motivation for such a detailed account on Donaldson's framework is the fact that there are several interesting variants and extensions of the theory. We will explore some of these in the remaining three chapters of this thesis building upon the discussion in this chapter.

4.1 Introduction

Donaldson's moment map framework

Let (M, ρ) be a closed manifold equipped with a volume form ρ, denote by $P \to M$ its $\mathrm{SL}(n, \mathbb{R})$ frame bundle and let (X, ω) be a symplectic manifold with Hamiltonian $\mathrm{SL}(n, \mathbb{R})$ action generated by a moment map $\mu : X \to \mathfrak{sl}(n, \mathbb{R})^*$. Denote by $\mathcal{S}(P, X)$ the space of section of the associated symplectic fibration $P(X) := P \times_{\mathrm{SL}(n,\mathbb{R})} X$. This carries a natural symplectic form defined by

$$\underline{\omega}_s(\hat{s}_1, \hat{s}_2) := \int_M \omega_s(\hat{s}_1, \hat{s}_2)\rho \tag{4.1}$$

for vertical vector fields $\hat{s}_1, \hat{s}_2 \in \Omega^0(M, s^* T^{vert} P(X))$.

The group $\mathrm{Diff}(M, \rho)$ of volume preserving diffeomorphisms acts symplectically on $\mathcal{S}(P, X)$ by pullback. It is useful to view this group formally as an infinite dimensional Lie group with Lie algebra

$$\mathrm{Lie}\left(\mathrm{Diff}^{\mathrm{ex}}(M, \rho)\right) = \{v \in \mathrm{Vect}(M) \mid d\iota(v)\rho = 0\}. \tag{4.2}$$

The subgroup $\mathrm{Diff}^{\mathrm{ex}}(M, \rho) \subset \mathrm{Diff}_0(M, \rho)$ of exact volume preserving diffeomorphisms is the subgroup obtained by integrating exact divergence free vector fields. In other words, it is the subgroup corresponding to the Lie subalgebra

$$\mathrm{Lie}\left(\mathrm{Diff}^{\mathrm{ex}}(M, \rho)\right) = \{v \in \mathrm{Vect}(M) \mid \iota(v)\rho \text{ is exact}\}. \tag{4.3}$$

This space is isomorphic to $\Omega^{n-2}(M)/\ker(d)$ and thus its dual space can formally be identified with the space of exact 2-forms on M

$$\mathrm{Lie}\left(\mathrm{Diff}^{\mathrm{ex}}(M, \rho)\right)^* = d\Omega^1(M). \tag{4.4}$$

In this setup Donaldson proved the following theorem.

Theorem A (**Donaldson** [38])**.** *Fix a torsion free $SL(n, \mathbb{R})$ connection ∇ on M and define $\underline{\mu} : \mathcal{S}(P, X) \to \Omega^2(M)$ by*

$$\underline{\mu}(s) := \omega(\nabla s \wedge \nabla s) - \langle \mu_s, R^\nabla \rangle - dc(\nabla \mu_s) \tag{4.5}$$

where $\mu_s \in \Omega^0(M, End_0(TM)^*)$ is obtained by composing the equivariant lift $\tilde{s} : P \to X$ of s with the moment map $\mu : X \to \mathfrak{sl}(n, \mathbb{R})^*$ and $c(\nabla \mu_s) \in \Omega^1(M)$ is defined as the contraction $(\mu_s)^i_{j;i}$ of $\nabla \mu_s$. Then the following holds.

1. *The map $\underline{\mu}$ is Diff(M, ρ)-equivariant and $\underline{\mu}(s) \in \Omega^2(M)$ is closed and independent of the connection ∇ used to define it.*

2. *Let $v \in Vect(M)$ be an exact divergence free vector field and choose a primitive $\alpha_v \in \Omega^{n-2}(M)$ with $d\alpha_v = \iota(v)\rho$. Then*

$$\partial_t \int_M \underline{\mu}(s(t)) \wedge \alpha_v = \int_M \omega(\dot{s}(t), \mathcal{L}_v s(t))\rho \tag{4.6}$$

for any smooth curve $s : \mathbb{R} \to \mathcal{S}(P, X)$, where $\mathcal{L}_v s$ denotes the infinitesimal action of v on s for the right action.

Proof. This is Theorem 9 in [38] and we include the proof in Theorem 4.2.4. (Note that the formula for the moment map in [38] contains some obvious typos regarding the signs which we corrected in formula stated above.) □

Remark 4.1.1. The map $\underline{\mu}$ is not a moment map in the strict sense, since it takes only values in the space of closed 2-forms and not in the space of exact 2-forms. Nevertheless, Theorem A asserts that $\underline{\mu}$ satisfies the moment map equation. When $\dim(M) = 2$, one can fix this by subtracting a suitable multiple of the area form. In higher dimensions, one could still subtract a suitable closed 2-form $\tau \in \Omega^2(M)$ such that $\underline{\mu}(s) - \tau$ is exact for every $s \in \mathcal{S}(P, X)$. However, the resulting moment map $\underline{\mu}(s) - \tau$ is in general not equivariant.

Construction of the hyperkähler moduli space $\mathcal{M}$

In the following let (Σ, ρ) be a closed oriented 2-dimensional surface with fixed area form $\rho \in \Omega^2(\Sigma)$ and assume $\text{genus}(\Sigma) \geq 2$.

The fundamental example considered by Donaldson arises form taking the hyperbolic plane $\mathbb{H}$ as fibre. This admits a canonical identification with the space $\mathcal{J}(\mathbb{R}^2)$ of linear complex structures on $\mathbb{R}^2$ and the space of sections $\mathcal{S}(P, X) \cong \mathcal{J}(\Sigma)$ gets then identified with the space of complex structures on Σ. Theorem A asserts that the action of $\text{Ham}(\Sigma, \rho)$ on $\mathcal{J}(\Sigma)$ is Hamiltonian and generated by the moment map

$$\underline{\mu}(J) = 2(K_J - c)\rho \qquad \text{with} \quad c := \frac{2\pi(2\text{genus}(\Sigma) - 2)}{\text{vol}(\Sigma, \rho)}, \tag{4.7}$$

where K_J denotes the Gaussian curvature of $\rho(\cdot, J\cdot)$ and c is determined by the Gauss–Bonnet theorem. After taking the action of $\text{Symp}_0(\Sigma, \rho)/\text{Ham}(\Sigma, \rho)$ on the Marsden–Weinstein into account, this yields the Teichmüller space

$$\mathcal{T}(\Sigma) := \mathcal{J}(\Sigma)/\text{Diff}_0(\Sigma) \cong \{J \in \mathcal{J}(\Sigma) \mid K_J = c\} / \text{Symp}_0(\Sigma, \rho) \tag{4.8}$$

equipped with the Weil–Petersson symplectic form (up to scaling).

A remarkable generalization of this construction is obtained by taking as fibre the unit disc bundle $X \subset T^*\mathbb{H}$. This carries a unique $S^1 \times \text{SL}(2, \mathbb{R})$-invariant hyperkähler metric, which extends the hyperbolic metric along the zero section and blows up when approaching the boundary of the disc bundle (see Theorem 4.5.1). A section $s \in \mathcal{S}(P, X)$ corresponds now to a pair (J, σ) consisting of a complex structure J and a quadratic differential σ with $|\sigma|_J < 1$, i.e. $\mathcal{S}(P, X)$ corresponds to

$$\mathcal{Q}_1(\Sigma) := \big\{(J, \sigma) \mid J \in \mathcal{J}(\Sigma),\, \sigma \in \Omega^0(\Sigma, S^2(T^*\Sigma \otimes_J \mathbb{C})),\, |\sigma|_J < 1\big\} \tag{4.9}$$

The hyperkähler structure on X then yields a hyperkähler structure on $\mathcal{Q}_1(\Sigma)$ and Theorem A asserts that there exists a hyperkähler moment map for the action of the Hamiltonian diffeomorphism group.

Theorem B (**Donaldson** [38])**.**

1. *The action of Ham(Σ, ρ) on $\mathcal{Q}_1(\Sigma)$ admits a hyperkähler moment map given by*

$$\begin{gathered} \underline{\mu}_1(J, \sigma) = \frac{|\bar{\partial}\sigma|^2 - |\partial\sigma|^2}{\sqrt{1 - |\sigma|^2}}\rho + 2\sqrt{1 - |\sigma|^2}K_J\rho + 2\boldsymbol{i}\bar{\partial}\partial\sqrt{1 - |\sigma|^2} - 2c\rho \\ \underline{\mu}_2(J, \sigma) + \boldsymbol{i}\underline{\mu}_3(J, \sigma) = -2\boldsymbol{i}\overline{\bar{\partial} r(\bar{\partial}\sigma)} \end{gathered} \tag{4.10}$$

where $c := 2\pi(2 - 2genus(\Sigma))/vol(\Sigma, \rho)$ and $r : \Omega^{0,1}(\Sigma, S^2(T^*\Sigma \otimes_J \mathbb{C})) \to \Omega^{1,0}(\Sigma)$ is the contraction defined by the metric $\rho(\cdot, J\cdot)$.

2. *The action of $Symp_0(\Sigma,\rho)$ on $\mathcal{Q}_1(\Sigma)$ is Hamiltonian for the second and third symplectic form with moment maps*

$$\langle \underline{\tilde{\mu}}_2(J,\sigma),v\rangle + \boldsymbol{i}\langle \underline{\tilde{\mu}}_3(J,\sigma),v\rangle = -2\boldsymbol{i}\int_\Sigma \overline{\iota(v)r(\bar{\partial}_J\sigma)}\rho \tag{4.11}$$

for any symplectic vector field $v \in Vect(\Sigma)$ satisfying $d\iota(v)\rho = 0$.

Proof. This is Proposition 17 in [38] and we present a proof in Theorem 4.5.13. We give an alternative proof of the second statement, since we found it difficult to translate the conceptual arguments given by Donaldson into a rigorous proof. We proceed by generalizing the proof of Theorem A where we use that the canonical holomorphic symplectic form $\omega_2 + \mathbf{i}\omega_3$ on X is exact. □

The construction of the hyperkähler quotient $\mathcal{M}$ based upon Theorem B requires some additional work. This was indicated very briefly in [38] and we include a careful exposition of this. First, we show that the quotient

$$\left(\underline{\mu}_1^{-1}(0)\cap\underline{\tilde{\mu}}_2^{-1}(0)\cap\underline{\tilde{\mu}}_2^{-1}(0)\right)\Big/\mathrm{Symp}_0(\Sigma,\rho)$$

inherits a canonical hyperkähler structure form $\mathcal{Q}_1(\Sigma)$. Then we show that by Moser isotopy and a suitable rescaling of the quadratic differential this moduli space is isomorphic to

$$\mathcal{M} := \left\{(g,\sigma)\in\mathrm{Met}(\Sigma)\times Q(g)\;\middle|\;\begin{array}{c}\bar{\partial}\sigma=0,\ |\sigma|<1,\\ K_g-\frac{c}{2}|\sigma|^2=\frac{c}{2}\end{array}\right\}\Big/\mathrm{Diff}_0(\Sigma) \tag{4.12}$$

where $c := \frac{2\pi(2-2\mathrm{genus}(\Sigma))}{\mathrm{vol}(\Sigma,\rho)}$ as above.

Geometric interpretations of the hyperkähler quotient

The moduli space $\mathcal{M}$ takes a particularly simple form when we scale the volume of Σ such that $c=-2$. Donaldson proposed under this assumption the following three geometric interpretations:

1. $\mathcal{M}$ can be embedded in $T^*\mathcal{T}(\Sigma)$ and the hyperkähler metric on $\mathcal{M}$ yields the Feix–Kaledin extension of the Weil–Petersson metric on $\mathcal{T}(\Sigma)$.

2. $\mathcal{M}$ parametrizes the class of almost-Fuchsian hyperbolic 3-manifolds. These are quasi-Fuchsian 3-manifolds which possess an incompressible minimal surface with principal curvatures in $(-1,1)$. This surface is then unique and its area provides a Kähler potential for the hyperkähler metric.

3. $\mathcal{M}$ embeds as an open subset into the smooth locus of the $\mathrm{SL}(2,\mathbb{C})$ representation variety $\mathcal{R}_{\mathrm{SL}(2,\mathbb{C})}(\Sigma) := \mathrm{Hom}\left(\pi_1(\Sigma),\mathrm{SL}(2,\mathbb{C})\right)/\mathrm{SL}(2,\mathbb{C})$. The hyperkähler structure on $\mathcal{M}$ is compatible with the natural holomorphic symplectic structure introduced by Goldman [52], where the natural complex structure coincides with the second complex structure on $\mathcal{M}$.

Remark 4.1.2. The class of almost-Fuchsian manifolds is strictly smaller the the class of quasi-Fuchsian manifold: There are examples of quasi-Fuchsian manifolds which admit more then one minimal surface (see [121, 63, 57]) and these cannot be almost-Fuchsian (see Lemma 4.6.5).

The isomorphism between $\mathcal{M}$ and the space of almost-Fuchsian manifolds follows from Uhlenbeck's theory of minimal surfaces in hyperbolic 3-manifolds [117]. Her result gives rise to the following theorem in our context.

Theorem C (Uhlenbeck [117]**).** *Let $g \in Met(\Sigma)$ and $\sigma \in Q(g)$ satisfy the equations $K_g + |\sigma|^2 = -1$, $\bar{\partial}\sigma = 0$, and $|\sigma|_g < 1$. For every such pair we define an almost-Fuchsian metric on $Y := \Sigma \times \mathbb{R}$ by*

$$g^Y = g^Y_{g,\sigma} = \begin{pmatrix} g\left(\cosh(t)\mathbb{1} - \sinh(t)g^{-1}Re(\sigma)\right)^2 & 0 \\ 0 & 1 \end{pmatrix}. \tag{4.13}$$

This is the unique almost-Fuchsian metric which restricts to g along $\Sigma \times \{0\}$ and such that $Re(\sigma)$ is the second fundamental form of $\Sigma \times \{0\} \subset Y$.

Proof. See Theorem 4.6.4. □

Let $(Y := \Sigma \times \mathbb{R}, g^Y)$ be an almost Fuchsian manifold. Its boundary at infinity is the disjoint union of two disjoint unions of Σ, which are both equipped with an induced conformal structure. This gives rise to an embedding of the space of almost Fuchsian metrics into the product space $\mathcal{T}(\Sigma) \times \overline{\mathcal{T}(\Sigma)}$. An alternative construction of this map was introduced by Hodge [61]. This is based on the $\mathrm{SL}(2,\mathbb{R})$-equivariant diffeomorphism $\alpha : X \to \mathbb{H} \times \overline{\mathbb{H}}$ defined by

$$\alpha(x + \mathbf{i}y, u + \mathbf{i}v) := \left(x - \frac{y^2 v}{1 - yu} + \mathbf{i}\frac{y\gamma}{1 - yu}, x + \frac{y^2 v}{1 + yu} + \mathbf{i}\frac{y\gamma}{1 + yu}\right) \tag{4.14}$$

where $\gamma := \sqrt{1 - y^2(u^2 + v^2)}$. The second complex structure on $X \subset T^*\mathbb{H}$ corresponds under this map to $(\mathbf{i}, -\mathbf{i})$ on $\mathbb{H} \times \overline{\mathbb{H}}$. On the space of sections, this gives rise to an embedding of $\mathcal{Q}_1(\Sigma)$ into $\mathcal{J}(\Sigma) \times \overline{\mathcal{J}(\Sigma)}$ which descends to the moduli spaces and yields the same embedding of $\mathcal{M}$ into $\mathcal{T}(\Sigma) \times \overline{\mathcal{T}(\Sigma)}$ as before. This embedding intertwines the second complex structure on $\mathcal{M}$ with the complex structure $(\hat{J}_1, \hat{J}_2) \mapsto (-J_1\hat{J}_1, J_2\hat{J}_2)$ on $\mathcal{T}(\Sigma) \times \overline{\mathcal{T}(\Sigma)}$ (see Proposition 4.6.8 and Proposition 4.6.7). We then verify the following remarkable observation of Donaldson.

Theorem D. *Let $A : \mathcal{AF}(\Sigma) \to \mathbb{R}$ be the area functional, which assigns to an almost Fuchisan manifold Y the area of its unique minimal surface. Then*

$$2\boldsymbol{i}\bar{\partial}_{J_2}\partial_{J_2}A = \underline{\omega}_2. \tag{4.15}$$

Hence A provides a Kähler potential with respect to the natural complex structure on $\mathcal{AF}(\Sigma)$ which agrees (up to sign) with the second complex structure on $\mathcal{M}$.

Proof. See Theorem 4.6.9. □

By the Cartan–Ambrose–Higgs theorem, one can express every complete hyperbolic 3-manifold as quotient of hyperbolic space $\mathbb{H}^3$. This gives rise to a natural embedding of the almost Fuchsian moduli space into $\mathcal{R}_{\mathrm{PSL}(2,\mathbb{C})}(\Sigma)$. A classical result of Bers [9] asserts that the restriction of this complex structure to $\mathcal{M}$ corresponds to the standard complex structure on $\mathcal{T}(\Sigma) \times \overline{\mathcal{T}(\Sigma)}$ which differs by a sign from our conventions. In particular, the second complex structure on $\mathcal{M}$ corresponds to multiplication by $-\mathbf{i}$ on $\mathcal{R}_{\mathrm{PSL}(2,\mathbb{C})}(\Sigma)$.

The representation associated to an almost-Fuchisan manifold lifts to $\mathrm{SL}(2,\mathbb{C})$ and a corresponding embedding of $\mathcal{M}$ into $\mathcal{R}_{\mathrm{SL}(2,\mathbb{C})}(\Sigma)$ can be constructed directly using the theory of Higgs bundles [58]. This has been suggested by Donaldson [38] and goes as follows: Let $g \in \mathrm{Met}(\Sigma)$ and $\sigma \in Q(g)$ be given. Choose a holomorphic line bundle $L \to \Sigma$ with $L^2 = T\Sigma$ and define $E = L \oplus L^{-1}$. The Levi-Civita connection for g induces a unique $U(1)$-connection $a \in \mathcal{A}(L)$. Then consider the pair

$$A = \begin{pmatrix} a & \frac{\bar{\sigma}}{2} \\ -\frac{\sigma}{2} & -a \end{pmatrix} \in \mathcal{A}(E) \quad \text{and} \quad \phi = \frac{1}{2}\begin{pmatrix} 0 & \mathbf{1} \\ 0 & 0 \end{pmatrix} \in \Omega^{1,0}(\mathrm{End}(E)) \tag{4.16}$$

where $\sigma \in \Omega^{1,0}(L^{-2}) = \Omega^{1,0}(\mathrm{Hom}(L, L^{-1}))$ and $\mathbf{1} \in \Omega^0(\mathrm{End}(T\Sigma)) = \Omega^{1,0}(L^2) = \Omega^{1,0}(\mathrm{Hom}(L^{-1}, L))$.

Theorem E. *Let $g \in Met(\Sigma)$ and $\sigma \in Q(g)$ satisfy the equations $K_g + |\sigma|^2 = -1$, $\bar{\partial}\sigma = 0$, and $|\sigma|_g < 1$. The corresponding pair (A, ϕ) defined by (4.16) satisfies the Hitchin equation*

$$\bar{\partial}_A \phi = 0, \qquad F_A + [\phi \wedge \phi^*] = 0$$

and $B := A + \phi + \phi^* \in \mathcal{A}^c(E)$ is a flat $SL(2,\mathbb{C})$ connection. The holonomy representation $\rho_B : \pi_1(\Sigma) \to SL(2,\mathbb{C})$ agrees up to conjugation with the representation associated to the almost Fuchian metric $g^Y_{g,\sigma}$ defined in Theorem C.

Proof. See Theorem 4.6.12). □

Finally, we show that the natural map of $\mathcal{M}$ into $T^*\mathcal{T}(\Sigma)$ is a well-defined embedding (see Theorem 4.6.14). This follows by a standard application of the continuation method and the proof is due to Uhlenbeck [117].

We should also mention the work of Taubes [108], which is closely related to our setup. He investigates the larger moduli space which one obtains by omitting the constraint $|\sigma|_g < 1$ in the definition of $\mathcal{M}$.

Overview

Section 2 discusses the relevant background on symplectic fibrations and contains a proof of Theorem A.

Section 3 summarizes basic properties of the hyperbolic plane $\mathbb{H}$. We provide an explicit formula for the identification $\mathbb{H} \cong \mathcal{J}(\Sigma)$ and show that the cotangent bundle $T^*\mathbb{H}$ can be identified with the space of pairs (J, q) where $J \in \mathcal{J}(\mathbb{R}^2)$ and $q \in Q(J)$ is a complex quadratic form on $(\mathbb{R}^2, J)$.

Section 4 contains Donaldson's construction of Teichmüller space. We include a detailed discussion on how this leads to the Weil–Petersson symplectic form on Teichmüller space. We also include an exposition of the identification of Teichmüller space with the space of Fuchsian representations.

Section 5 contains the construction of the moduli space $\mathcal{M}$: We present Donaldson's derivation of the hyperkähler metric on $X \subset T^*\mathbb{H}$ by elementary methods and then calculate the hyperkähler moment map for the $\mathrm{SL}(2,\mathbb{R})$ action on X. After this preparatory work, we proceed to the proof of Theorem B. This is the main ingredient for the construction of $\mathcal{M}$. Finally, we show that the hyperkähler structure on $\mathcal{Q}_1(\Sigma)$ induces a hyperkähler structure on $\mathcal{M}$.

Section 6 investigates the three geometric models for the hyperkähler moduli space $\mathcal{M}$. We construct various isomorphism between the different models and establish Theorem C, Theorem D and Theorem E. We also include a brief discussion of complete hyperbolic 3-manifolds, quasi-Fuchsian groups and the simultaneous uniformization theorem of Bers.

4.2 Donaldson's moment map

Let (M,ρ) be a closed oriented n-dimensional manifold with fixed volume form ρ and let $P \to M$ be its $\mathrm{SL}(n,\mathbb{R})$ frame bundle which is defined by

$$P := \{(z,\theta)\,|\, z \in M,\ \theta \in \mathrm{Hom}(\mathbb{R}^n, T_zM),\ \theta^*\rho_z = \mathrm{dvol}_{\mathbb{R}^n}\}.$$

Let (X,ω) be a symplectic manifold with Hamiltonian $\mathrm{SL}(n,\mathbb{R})$-action induced by an equivariant moment map $\mu : X \to \mathfrak{sl}^*(n,\mathbb{R})$ and consider the associated bundle

$$P(X) := P \times_{\mathrm{SL}(n,\mathbb{R})} X := (P \times X)/\mathrm{SL}(n,\mathbb{R})$$

where $\mathrm{SL}(n,\mathbb{R})$ acts diagonally. Denote by $\mathcal{S}(P,X)$ its space of sections. We establish the necessary background on symplectic fibrations and the action of the diffeomorphism group in the first two subsection. We then state the main result of this section is Theorem 4.2.4. This establishes Donaldson's moment map for the action of the subgroup $\mathrm{Diff}_{ex}(M,\rho)$ of exact volume preserving diffeomorphism on $\mathcal{S}(P,X)$. The action of the full group of volume preserving diffeomorphism is symplectic but in general not Hamiltonian.

4.2.1 Symplectic fibrations

The symplectic structure on $\mathcal{S}(P,X)$

The space $\mathcal{S}(P,X)$ is formally an infinite dimensional symplectic manifold. The tangent space at $s \in \mathcal{S}(P,X)$ is the space of vertical vector fields along s

$$T_s\mathcal{S}(P,X) = \Omega^0(M, s^*T^{vert}P(X)).$$

The symplectic form on X induces a symplectic structure on the vertical tangent bundle $T^{vert}P(X)$ and

$$\underline{\omega}_s : T_s\mathcal{S}(P,X) \times T_s\mathcal{S}(P,X) \to \mathbb{R}, \qquad \underline{\omega}_s(\hat{s}_1,\hat{s}_2) := \int_M \omega(\hat{s}_1,\hat{s}_2)\rho.$$

defines a symplectic form on $\mathcal{S}(P,X)$.

Symplectic connections and covariant differentiation

A connection $A \in \mathcal{A}(P)$ induces a covariant derivative on $\mathcal{S}(P, X)$ via

$$\begin{aligned} \nabla : \mathcal{S}(P,X) &\to \Omega^1(M, s^*T^{vert}P(X)) \\ \nabla_{\hat{p}} s(p) = ds(p)\hat{p} + L_{s(p)}A_p(\hat{p}) &= ds(p)\hat{p}^{hor}. \end{aligned} \tag{4.17}$$

In this formula $s : P \to X$ is an equivariant map, $L_x : \mathfrak{sl}(n,\mathbb{R}) \to T_xX$ denotes the infinitesimal action, and $\hat{p}^{hor} := \hat{p} - p \cdot A_p(\hat{p})$ is the horizontal component of a tangent vector $\hat{p}$ of P. The next lemma shows that there exists a closed 2-form $\Omega \in \Omega^2(P(X))$ defined on the total space which agrees with ω along the fibres and such that the horizontal and vertical subspaces are Ω-orthogonal. Conversely, any such 2-form gives rise to a symplectic connection on $P(X)$, where one recovers the horizontal distribution $T^{hor}P(X)$ as the Ω-orthogonal complement of the fibre.

Lemma 4.2.1 (A closed 2-form on the total space). *Let $A \in \mathcal{A}(P)$ be given and define $\Omega \in \Omega^2(P \times X)$ by*

$$\begin{aligned} &\Omega_{(p,x)}((\hat{p}_1, \hat{x}_1), (\hat{p}_2, \hat{x}_2)) \\ &\qquad := \omega_x(\hat{x}_1 + L_xA_p(\hat{p}_1), \hat{x}_2 + L_xA_p(\hat{p}_2)) - \langle \mu(x), F_A(\hat{p}_1, \hat{p}_2)\rangle. \end{aligned} \tag{4.18}$$

This is a closed, equivariant and horizontal 2-form on $P \times X$. In particular, it descends to a closed 2-form on $P(X)$ which restricts to ω along the fibres.

Proof. Define $\alpha \in \Omega^1(P \times X)$ by $\alpha_{(p,x)}(\hat{p}, \hat{x}) := \langle \mu(x), A_p(\hat{p})\rangle$. Then

$$\begin{aligned} &(d\alpha)_{(p,x)}((\hat{p}_1, \hat{x}_1), (\hat{p}_2, \hat{x}_2)) \\ &\qquad = \langle d\mu(x)\hat{x}_1, A_p(\hat{p}_2)\rangle - \langle d\mu(x)\hat{x}_2, A_p(\hat{p}_2)\rangle - \langle \mu(x), (dA)_p(\hat{p}_1, \hat{p}_2)\rangle \\ &\qquad = \omega_x\left(L_xA_p(\hat{p}_2), \hat{x}_1\right) - \omega_x\left(L_xA_p(\hat{p}_1), \hat{x}_2\right) - \langle \mu(x), (dA)_p(\hat{p}_1, \hat{p}_2)\rangle, \end{aligned}$$

where the second equation follows from the characteristic equation for the moment map. Equivariance of the moment map yields the identity

$$\omega_x(L_xA_p(\hat{p}_1), L_xA_p(\hat{p}_2)) = \langle \mu(x), [A_p(\hat{p}_1), A_p(\hat{p}_2)]\rangle.$$

Denote by $\mathrm{pr}_X : P \times X \to X$ the projection onto the second factor. Putting everything together, we have shown that $\Omega = \mathrm{pr}_X^*\omega - d\alpha$ and therefore Ω is closed.

The tangent space of the $\mathrm{SL}(n,\mathbb{R})$-orbit through a point $(p, x) \in P \times X$ is given by $\{(p\xi, -L_x\xi) \in T_pP \times T_xX \mid \xi \in \mathfrak{sl}(n,\mathbb{R})\}$. It now follows from (4.18) and Cartan's formula that Ω is $\mathrm{SL}(n,\mathbb{R})$-invariant. It follows directly from (4.18) that Ω is horizontal and thus descends to a closed 2-form on $P(X)$. $\square$

There exists a natural isomorphism $\mathrm{ad}(P) \cong \mathrm{End}_0(TM)$ and we denote by $R \in \Omega^2(M, \mathrm{End}_0(TM))$ the curvature of the connection $A \in \mathcal{A}(P)$ under this identification. Let $s \in \mathcal{S}(P, X)$ be given. The composition $\mu \circ s : P \to \mathrm{sl}(n,\mathbb{R})^*$ is equivariant and thus descends to a section $\mu_s \in \Omega^0(M, \mathrm{End}_0(TM)^*)$. Denote the dual pairing of these two sections by

$$\langle \mu_s, R\rangle \in \Omega^2(M). \tag{4.19}$$

Define $\omega(\nabla s \wedge \nabla s) \in \Omega^2(M)$ by coupling the exterior product on M with the symplectic from on $T^{vert}P(X)$:

$$\omega(\nabla s \wedge \nabla s) : TM \times TM \to \mathbb{R}, \qquad (u,v) \mapsto \omega(\nabla_u s, \nabla_v s) \tag{4.20}$$

Lemma 4.2.2. *Fix a connection $A \in \mathcal{A}(P)$. Let $s \in \mathcal{S}(P,X)$ and define $\Omega \in \Omega^2(P(X))$ by (4.18). Then*

$$s^*\Omega = \omega(\nabla s \wedge \nabla s) - \langle \mu_s, R \rangle. \tag{4.21}$$

where the two terms on the right hand side are define by (4.19) and (4.20).

Proof. Lift $s \in \mathcal{S}(P,X)$ to an equivariant map $s : P \to X$ and define $\tilde{s} : P \to P \times X$ by $\tilde{s}(p) := (p, s(p))$. It follows from (4.17) that

$$\tilde{s}^*\Omega_p(\hat{p}_1, \hat{p}_2) = \omega_{s(p)}(\nabla_{\hat{p}_1} s(p), \nabla_{\hat{p}_2} s(p)) - \langle \mu(s(p)), F_A(\hat{p}_1, \hat{p}_2) \rangle$$

The curvature form $F_A \in \Omega^2(P, \mathfrak{sl}(n,\mathbb{R}))$ descends to $R \in \Omega^2(M, \mathrm{End}_0(TM))$ and hence $\tilde{s}^*\Omega$ descends to the 2-form $\omega(\nabla s \wedge \nabla s) - \langle \mu_s, R \rangle$ on M. $\square$

4.2.2 Action of the diffeomorphism group

The group $\mathrm{Diff}(M,\rho)$ of volume preserving diffeomorphisms can be viewed as infinite dimensional Lie group with Lie algebra

$$\mathrm{Lie}\left(\mathrm{Diff}(M,\rho)\right) = \left\{ v \in \mathrm{Vect}(M) \,|\, d\iota(v)\rho = 0 \right\}.$$

Every $\phi \in \mathrm{Diff}(M,\rho)$ lifts naturally to an equivariant diffeomorphism of P defined by

$$\tilde{\phi} : P \to P, \qquad \tilde{\phi}(z,\theta) := (\phi(z), d\phi(z) \circ \theta)$$

for $z \in M$ and $\theta \in \mathrm{Hom}(\mathbb{R}^n, T_zM)$. This induces a natural action

$$\mathrm{Diff}(P,\rho) \times \mathcal{S}(P,X) \to \mathcal{S}(P,X), \qquad \phi^*s := s \circ \tilde{\phi}$$

where we view elements of $\mathcal{S}(P,X)$ as equivariant maps $s : P \to X$.

Infinitesimal action

There is a one-to-one correspondence between connections $A \in \mathcal{A}(P)$ and $\mathrm{SL}(n,\mathbb{R})$ connections ∇ on TM. For the calculation of the infinitesimal action it is useful to adopt the later point of view and to choose a torsion free $\mathrm{SL}(n,\mathbb{R})$ connections on TM as auxiliary data.

Lemma 4.2.3. *Choose a torsion-free $SL(n,\mathbb{R})$ connection ∇ on TM and denote by $A \in \mathcal{A}(P)$ the corresponding connection 1-form on P. Let $v \in Vect(M)$ with $d\iota(v)\rho = 0$ and denotes is flow by $\phi_v^t \in Diff(M,\rho)$.*

1. *The infinitesimal action of v on P is defined as*

$$\mathcal{L}_v(z,\theta) := \left.\frac{d}{dt}\right|_{t=0} (\phi_v^t(z), d\phi_v^t(z) \circ \theta) \in T_{(z,\theta)}P$$

and satisfies for all $(z,\theta) \in P$

$$d\pi(z,\theta)\mathcal{L}_v(z,\theta) = v(z), \qquad A_{(z,\theta)}(\mathcal{L}_v(z,\theta)) = \theta^{-1}(\nabla_{\theta(\cdot)}v)(z). \tag{4.22}$$

where $\pi : P \to M$ *denotes the projection map.*

2. *Denote by* $\nabla v : P \to \mathfrak{sl}(n,\mathbb{R})$ *the map* $(z,\theta) \mapsto \theta^{-1}(\nabla_{\theta(\cdot)}V)(z)$ *. Then*

$$\mathcal{L}_v s := \left.\frac{d}{dt}\right|_{t=0} (\phi_v^t)^* s = \nabla_v s - L_s(\nabla v). \tag{4.23}$$

Proof. The first part of (4.22) follows from

$$d\pi(z,\theta)\mathcal{L}_v(x,\theta) = \left.\frac{d}{dt}\right|_{t=0} \pi(\phi_v^t(z), d\phi_V^t(z) \circ \theta) = \left.\frac{d}{dt}\right|_{t=0} \phi_v^t(z) = v(z).$$

The following calculation uses that ∇ is a torsion free connection corresponding to $A \in \mathcal{A}(P)$. For every $\xi \in \mathbb{R}^n$ it holds

$$A_{(z,\theta)}(\mathcal{L}_v(z,\theta))\xi = \theta^{-1}\, \nabla_t d\phi_v^t(z)\theta(\xi)\big|_{t=0} = \theta^{-1}\nabla_{\theta(\xi)}\, \partial_t\phi_v^t(z)\big|_{t=0} = \theta^{-1}\nabla_{\theta(\xi)}v(z).$$

This completes the proof of (4.22). Next, let $s : P \to X$ be an equivariant map. Then, by the chain rule and (4.17), it follows

$$(\mathcal{L}_v s)(p) = ds(p)[\mathcal{L}_v(p)] = \nabla s(p)[\mathcal{L}_v(p)] - L_{(s(p))}A_p(\mathcal{L}_v(p))$$

for every $p \in P$. Equation (4.23) follows from this and (4.22). □

Exact volume preserving diffeomorphism

A diffeomorphism $\phi \in \mathrm{Diff}(M,\rho)$ is called exact, if there exists an isotopy $\phi : [0,1] \to \mathrm{Diff}(M,\rho)$ with $\phi_0 = \mathbb{1}$ and $\phi_1 = \phi$ and there exists a smooth map $v : [0,1] \to \mathrm{Vect}(M)$ such that $\partial_t\phi_t = v_t \circ \phi_t$ and $\iota(v_t)\rho$ is exact for all $t \in [0,1]$. This is the subgroup of $\mathrm{Diff}(M,\rho)$ corresponding to the Lie subalgebra

$$\mathrm{Lie}\left(\mathrm{Diff}_{ex}(M,\rho)\right) = \left\{v \in \mathrm{Vect}(M) \,|\, \iota(v)\rho \in d\Omega^{n-2}(M)\right\}.$$

The dual space of the Lie algebra can be identified with the space of exact 2-forms on M using the pairing

$$\Omega^2_{ex}(M) \times \mathrm{Lie}(\mathrm{Diff}_{ex}(M,\rho)) \to \mathbb{R}, \qquad \langle\tau, v\rangle := \int_M \tau \wedge \alpha_v$$

where $\alpha_v \in \Omega^{n-2}(M)$ satisfies $d\alpha_v = \iota(v)\rho$. By Stokes theorem, the pairing does not depend on the choice of this primitive, since τ is exact.

4.2.3 Donaldson's moment map

For $s \in \mathcal{S}(P,X)$ the moment map $\mu : X \to \mathfrak{sl}(n,\mathbb{R})^*$ gives rise to a section $\mu_s := \mu\circ s \in \Omega^0(M, \mathrm{End}_0(TM)^*)$. Its covariant derivative is a tensor $\nabla\mu_s \in \Omega^1(M, \mathrm{End}_0(TM)^*)$

and we denote by $c(\nabla\mu_s) \in \Omega^1(M)$ its contraction of the first and third index. This is given by

$$c(\nabla\mu_s) \in \Omega^1(M), \qquad \hat{m} \mapsto \sum_{i=1}^{n} \left\langle \nabla_{e_i}\mu_s, (\hat{m} \otimes e^i)_0 \right\rangle \tag{4.24}$$

where $(e_1, \ldots, e_n)$ is a local frame for TM with dual frame $(e^1, \ldots, e^n)$ and the contraction does not depend on the choice of this basis. We are now ready to state the main result of this section.

Theorem 4.2.4 (**Donaldson's moment map**). *Let ∇ be a torsion-free $SL(n, \mathbb{R})$ on TM and define*

$$\underline{\mu} : \mathcal{S}(P, X) \to \Omega^2(M), \qquad \underline{\mu}(s) := \omega(\nabla s \wedge \nabla s) - \langle \mu_s, R \rangle - dc(\nabla\mu_s) \tag{4.25}$$

where the expression on the right hand side are defined in (4.19), (4.20) and (4.24).

1. *$\underline{\mu}(s)$ is closed for every $s \in \mathcal{S}(P, X)$.*
2. *$\underline{\mu}$ is equivariant for the action of Diff(M, ρ).*
3. *Let $v \in$ Vect(M) be an exact divergence free vector field and choose a primitive $\alpha_v \in \Omega^{n-2}(M)$ with $d\alpha_v = \iota(v)\rho$. The derivative of the map*

$$\mathcal{S}(P, X) \to \mathbb{R}, \qquad s \mapsto \int_M \underline{\mu}(s) \wedge \alpha_v \tag{4.26}$$

   is the map $T_s\mathcal{S}(P, X) \to \mathbb{R}$ defined by

$$\hat{s} \mapsto \underline{\omega}(\hat{s}, \mathcal{L}_v s) = \int_M \omega(-\nabla_v s + L_s \nabla v, \hat{s})\rho \tag{4.27}$$

4. *$\underline{\mu}$ is independent of the choice of the torsion free $SL(n, \mathbb{R})$ connection ∇.*

Remark 4.2.5. The map $\underline{\mu}$ is not a moment map in the strict sense, since it takes values in the space of closed 2-forms. Let $v \in \mathrm{Vect}(M)$ such that $\iota(v)\rho$ is exact and choose $\alpha_v \in \Omega^{n-2}(M)$ with $d\alpha_v = \iota(v)\rho$. Then

$$\langle \underline{\mu}(s), v \rangle = \int_M \underline{\mu}(s) \wedge \alpha_v$$

depends on the choice of the primitive α_v. Different choices for α_v change the pairing only by a constant, and so its derivative is well-defined and independent of any choices. The equations (4.27) and (4.26) show that $\underline{\mu}$ satisfies the moment map equation.

Alternatively, it follows from Lemma 4.2.2 that the values of $\underline{\mu}(s)$ are contained in a single cohomology class in $H^2(M)$. Let τ be any representative of this class and consider $\underline{\mu}(s) - \tau$ as a moment map. This is a moment map in the strict sense, but it is not equivariant unless M is a surface.

The proof of Theorem 4.2.4 takes up the rest of this section.

Proof of the first two assertions in Theorem 4.2.4

It follows from Lemma 4.2.2 that

$$\underline{\mu}(s) := s^*\Omega - dc(\nabla\mu_s). \tag{4.28}$$

Since $\Omega \in \Omega^2(P(X))$ is closed, this implies that $\underline{\mu}(s)$ is closed.

Let $\phi \in \mathrm{Diff}(M,\rho)$ be given and define $\tilde{\nabla} := \phi^*\nabla$. This is again a torsion free $\mathrm{SL}(n,\mathbb{R})$ connection on TM. Since the pullback of connections is functorial with respect to the various induced connections, we get

$$c(\nabla\mu_{\phi^*s}) = c(\nabla\phi^*\mu_s) = c(\phi^*\tilde{\nabla}\mu_s) = \phi^*c(\tilde{\nabla}\mu_s).$$

and similarly

$$\omega(\nabla\phi^*s, \nabla\phi^*s) - \langle\mu_{\phi^*s}, R^\nabla\rangle = \phi^*\omega(\tilde{\nabla}s, \tilde{\nabla}s) - \phi^*\langle\mu_s, R^{\tilde{\nabla}}\rangle.$$

Hence equivariance of $\underline{\mu}$ will follow from the fourth assertion that $\underline{\mu}$ is independent of the choice of the torsion free $\mathrm{SL}(n,\mathbb{R})$ connection ∇.

Proof of the moment map equation

We prove the third assertion in Theorem 4.2.4. Let $v \in \mathrm{Vect}(M,\rho)$ be an exact divergence free vector field and choose $\alpha_v \in \Omega^{n-2}(M)$ with $d\alpha_v = \iota(v)\rho$. We claim that for any smooth curve $s : \mathbb{R} \to \mathcal{S}(P,X)$ it holds

$$\partial_t \int_M s(t)^*\Omega \wedge \alpha_v = \underline{\omega}(-\nabla_v s(t), \dot{s}(t)) \tag{4.29}$$

$$-\partial_t \int_M dc(\nabla\mu_{s(t)}) \wedge \alpha_v = \underline{\omega}(L_{s(t)}\nabla v, \dot{s}(t)). \tag{4.30}$$

These two equations together with (4.28) then yield the third assertion.

We prove (4.29): Since Ω is closed it follows from Cartan's formula:

$$\begin{aligned}
\partial_t \int_M s(t)^*\Omega \wedge \alpha_v &= \int_M ds(t)^*(\iota(\dot{s}(t))\Omega) \wedge \alpha_v \\
&= \int_M s(t)^*(\iota(\dot{s}(t))\Omega) \wedge \iota(v)\rho \\
&= \int_M \iota(v)s(t)^*(\iota(\dot{s}(t))\Omega)\rho \\
&= \int_M \Omega(\dot{s}(t), ds(t)v)\rho \\
&= \int_M \omega(\dot{s}(t), \nabla_v s(t))\rho = \underline{\omega}(-\nabla_v s(t), \dot{s}(t))
\end{aligned}$$

where we used in the penultimate equation that the horizontal and vertical tangent spaces of $P(X)$ are Ω-orthogonal.

We prove (4.30): The moment map equation on the fibre X yields

$$\underline{\omega}(L_{s(t)}\nabla v, \dot{s}(t)) = \int_M \langle d\mu(s(t))\dot{s}(t), \nabla v\rangle\rho = \partial_t \int_M \langle \mu_{s(t)}, \nabla v\rangle\rho.$$

On the other hand, integration by parts yields

$$\int_M dc(\nabla\mu_s) \wedge \alpha_v = \int_M c(\nabla\mu_s) \wedge \iota(v)\rho = \int_M \iota(v)c(\nabla\mu_s)\rho.$$

Thus (4.30) will follow from the observation

$$\int_M (\iota(v)c(\nabla\mu_s) + \langle\mu_s, \nabla v\rangle)\rho = 0. \tag{4.31}$$

In local coordinates, write $v = v^i e_i$ and $\mu_s = \mu_i^j e^i \otimes e_j$. Then $\iota(v)c(\nabla\mu_s) + \langle\mu_s, \nabla v\rangle$ is given by

$$v^i\mu_{i;k}^k + v_{;k}^i\mu_i^k = \mathrm{div}(v^i\mu_i^k e_k).$$

Hence the integrand in (4.31) is a divergence term and its integral vanishes.

Independence of the connection

We prove the fourth assertion in Theorem 4.2.4. Let ∇ and ∇' be two torsion free $\mathrm{SL}(n,\mathbb{R})$-connections on TM and define

$$\gamma := \nabla' - \nabla \in \Omega^1(M, \mathrm{End}_0(TM)).$$

This satisfies the additional symmetry $\gamma_{ij}^k = \gamma_{ji}^k$. We show in the following that the moment map (4.25) defined with respect to ∇ and ∇' agree.

Step 1: The formula $R' = R + d^\nabla\gamma + [\gamma,\gamma]$ yields

$$\langle\mu_s, R' - R\rangle = \langle\mu_s, d^\nabla\gamma + [\gamma,\gamma]\rangle \tag{4.32}$$

Step 2: We show

$$\begin{aligned} &dc((\nabla' - \nabla)\mu_s)(e_i, e_j) \\ &\quad = -\langle\mu_s, (d^\nabla\gamma)(e_i, e_j)\rangle + \langle\nabla_{e_j}\mu_s, \gamma(e_i)\rangle - \langle\nabla_{e_i}\mu_s, \gamma(e_j)\rangle. \end{aligned} \tag{4.33}$$

For $\Phi \in \Omega^0(M, \mathrm{End}_0(TM))$ it holds

$$\langle(\nabla' - \nabla)\mu_s, \Phi\rangle = \langle\mu_s, (\nabla - \nabla')\Phi\rangle = \langle\mu_s, [\Phi, \gamma]\rangle.$$

This yields the formula

$$c(\nabla'\mu_s - \nabla\mu_s)e_j = \left\langle \mu_s, \sum_{i=1}^n \left[e_j \otimes e^i, \gamma(e_i)\right] \right\rangle$$

Using the symmetry condition $\gamma_{ij}^k = \gamma_{ji}^k$ and $\mathrm{tr}(\gamma(e_k)) = \sum_{i=0}^n \gamma_{ik}^i = 0$, we obtain

$$\sum_{i=1}^n \left[e_j \otimes e^i, \gamma(e_i)\right] = \sum_{i,k=1}^n \gamma_{ik}^i e_j \otimes e^k - \gamma_{ij}^k e_k \otimes e^i = -\gamma(e_j)$$

and hence

$$c(\nabla'\mu_s - \nabla\mu_s) = -\sum_{j=1}^{n}\langle\mu_s, \gamma(e_j)\rangle e^j.$$

Differentiating this equations yields

$$\begin{aligned} dc((\nabla' - \nabla)\mu_s)(e_i, e_j) &= \mathcal{L}_{e_j}\langle\mu_s, \gamma(e_i)\rangle - \mathcal{L}_{e_i}\langle\mu_s, \gamma(e_j)\rangle - \langle\mu_s, \gamma[e_i, e_j]\rangle \\ &= \langle\mu_s, \nabla_{e_j}\gamma(e_i) - \nabla_{e_i}\gamma(e_j) - \gamma[e_i, e_j]\rangle \\ &\quad + \langle\nabla_{e_j}\mu_s, \gamma(e_i)\rangle - \langle\nabla_{e_i}\mu_s, \gamma(e_j)\rangle \\ &= -\langle\mu_s, (d^\nabla\gamma)(e_i, e_j)\rangle + \langle\nabla_{e_j}\mu_s, \gamma(e_i)\rangle - \langle\nabla_{e_i}\mu_s, \gamma(e_j)\rangle. \end{aligned}$$

and this proves (4.33).

Step 3: We show that

$$\begin{aligned} &\omega(\nabla' s, \nabla' s)(e_i, e_j) - \omega(\nabla s, \nabla s)(e_i, e_j) \\ &\quad = \langle\mu_s, [\gamma(e_i), \gamma(e_j)]\rangle + \langle\nabla_{e_j}\mu_s, \gamma(e_i)\rangle - \langle\nabla_{e_i}\mu_s, \gamma(e_j)\rangle \end{aligned} \tag{4.34}$$

Denote by $A' := A^{\nabla'}$ and $A := A^{\nabla}$ the corresponding connection 1-froms on P. Then $(A' - A)(z, \theta) = \theta^{-1}\gamma(z)\theta$ and (4.17) yields

$$(\nabla' s - \nabla s)(z, \theta) = L_s(\theta^{-1}\gamma(z)\theta). \tag{4.35}$$

Then follows

$$\begin{aligned} &\omega(\nabla'_{e_i} s, \nabla'_{e_j} s) - \omega(\nabla_{e_i} s, \nabla_{e_j} s) \\ &\quad = \omega(\nabla'_{e_i} s - \nabla_{e_i} s, \nabla_{e_j} s) + \omega(\nabla_{e_i} s, \nabla'_{e_j} s - \nabla_{e_j} s) \\ &\qquad + \omega(\nabla'_{e_i} s - \nabla_{e_i} s, \nabla'_{e_j} s - \nabla_{e_j} s) \\ &\quad = \omega(L_s(\theta^{-1}\gamma(e_i)\theta), \nabla_{e_j} s) - \omega(L_s(\theta^{-1}\gamma(e_j)\theta), \nabla_{e_i} s) \\ &\qquad + \omega(L_s(\theta^{-1}\gamma(e_i)\theta), L_s(\theta^{-1}\gamma(e_j)\theta)) \\ &\quad = \langle d\mu(s)[\nabla_{e_j} s], \theta^{-1}\gamma(e_i)\theta\rangle - \langle d\mu(s)[\nabla_{e_i} s], \theta^{-1}\gamma(e_j)\theta\rangle \\ &\qquad + \langle\mu(s), \theta^{-1}[\gamma(e_i), \gamma(e_j)]\theta\rangle \\ &\quad = \langle\nabla_{e_j}\mu_s, \gamma(e_i)\rangle - \langle\nabla_{e_i}\mu_s, \gamma(e_j) + \langle\mu_s, [\gamma(e_i), \gamma(e_j)]\rangle \end{aligned}$$

The last equation uses $\nabla_{\hat{p}}(\mu \circ s)(p) = d\mu(s(p))[\nabla_{\hat{p}} s(p)]$ which follows from (4.17).

Step 4: The moment map $\underline{\mu}$ is independent of the chosen connection.

This follows directly from (4.32), (4.33), and (4.34) and completes the proof of Theorem 4.2.4.

4.3 Complex structures on $\mathbb{R}^2$ and quadratic forms.

The main applications of Theorem 4.2.4 to Teichmüller theory arises when one takes as fibre the hyperbolic plane or its cotangent bundle. We recall in this section fundamental properties of these spaces and establish our notation. In particular, we show that the hyperbolic plane can be identified with the space of linear complex structures on $\mathbb{R}^2$. Its cotangent bundle can be identified with pairs (J, q) consisting of a complex structure and a quadratic form.

4.3.1 The hyperbolic plane

The upper half-plane model. Let $\mathbb{H} := \{z \in \mathbb{C} \,|\, \mathrm{Im}(z) > 0\}$ denote the upper half plane. It has a canoncial complex structure and we endow it with the hyperbolic metric and volume form

$$g_{\mathbb{H}}(x, y) = \frac{dx^2 + dy^2}{y^2}, \qquad \omega_{\mathbb{H}}(x, y) = \frac{dx \wedge dy}{y^2}.$$

The group $\mathrm{SL}(2, \mathbb{R})$ acts on $\mathbb{H}$ by Möbius transformations

$$\mathrm{SL}(2, \mathbb{R}) \times \mathbb{H} \mapsto \mathbb{H}, \qquad \begin{pmatrix} a & b \\ c & d \end{pmatrix} z := \frac{az + b}{cz + d}. \tag{4.36}$$

Every Möbiustransformation is a Kähler isometry of $\mathbb{H}$. This action is transitive with stabilizer $\mathrm{SO}(2)$ at $\mathbf{i}$ and therefore $\mathbb{H} \cong \mathrm{SL}(2, \mathbb{R})/\mathrm{SO}(2)$.

The disc model. Let $\mathbb{D} := \{z \in \mathbb{C} \,|\, |z| < 1\}$ denote the open unit disc in $\mathbb{C}$. It carries a canonical complex structure and we equip it with the hyperbolic metric and volume form

$$g_{\mathbb{D}} := \frac{4(dx^2 + dy^2)}{(1 - x^2 - y^2)^2}, \qquad \omega_{\mathbb{D}} = \frac{4dx \wedge dy}{(1 - x^2 - y^2)^2}.$$

The group

$$\begin{aligned} \mathrm{SU}(1, 1) :&= \left\{ A \in \mathrm{SL}(2, \mathbb{C}) \,\middle|\, \bar{A}^t \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right\} \\ &= \left\{ \begin{pmatrix} \alpha & \beta \\ \bar{\beta} & \bar{\alpha} \end{pmatrix} \,\middle|\, \alpha, \beta \in \mathbb{C},\ |\alpha|^2 - |\beta|^2 = 1 \right\} \end{aligned}$$

acts by fractional linear transformations on $\mathbb{D}$ via

$$\mathrm{SU}(1, 1) \times \mathbb{D} \mapsto \mathbb{D}, \qquad \begin{pmatrix} \alpha & \beta \\ \bar{\beta} & \bar{\alpha} \end{pmatrix} z := \frac{\alpha z + \beta}{\bar{\beta} z + \bar{\alpha}}.$$

Lemma 4.3.1 (Cayley transform).

1. *The map $f : \mathbb{H} \to \mathbb{D}$ defined by*

$$f(z) := \frac{z - \boldsymbol{i}}{z + \boldsymbol{i}} \tag{4.37}$$

is a Kähler isometry with inverse given by $f^{-1}(z) := \boldsymbol{i}\frac{1+z}{1-z}$.

2. *There exists a unique isomorphism $SL(2, \mathbb{R}) \cong SU(1, 1)$ such that (4.37) is equivariant. The isomorphism $SL(2, \mathbb{R}) \to SU(1, 1)$ is given by*

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \frac{1}{2} \begin{pmatrix} (a + d) + (b - c)\boldsymbol{i} & (a - d) - (b + c)\boldsymbol{i} \\ (a - d) + (b + c)\boldsymbol{i} & (a + d) - (b - c)\boldsymbol{i} \end{pmatrix}$$

with inverse $SU(1,1) \to SL(2,\mathbb{R})$ given by

$$\begin{pmatrix} \alpha & \beta \\ \bar{\beta} & \bar{\alpha} \end{pmatrix} \mapsto \begin{pmatrix} Re(\alpha)+Re(\beta) & Im(\beta)-Im(\alpha) \\ Im(\alpha)+Im(\beta) & Re(\alpha)-Re(\beta) \end{pmatrix}.$$

Proof. The first part is left to the reader. For the second part, let

$$\Psi = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2,\mathbb{R})$$

be given. The automorphism $\psi(z) := f\left(\Psi_* f^{-1}(z)\right)$ is then given by

$$\psi(z) := \frac{(a\mathbf{i} - b + c + d\mathbf{i})z + (a\mathbf{i} + b + c - d\mathbf{i})}{(a\mathbf{i} - b - c - d\mathbf{i})z + (a\mathbf{i} + b - c + d\mathbf{i})}$$

and after normalization this yields the desired isomorphism $\mathrm{SL}(2,\mathbb{R}) \cong \mathrm{SU}(1,1)$. □

4.3.2 The space of complex structures on the plane

Denote the space of linear complex structures on $\mathbb{R}^2$, compatible with the standard orientation, by

$$\mathcal{J}(\mathbb{R}^2) := \{J \in \mathrm{End}(\mathbb{R}^2) \mid J^2 = -\mathbb{1},\, \det(\cdot, J\cdot) > 0\}$$

This is a smooth manifold and the tangent spaces at $J \in \mathcal{J}(\mathbb{R}^2)$ consists of all J-antilinear endomorphism:

$$T_J\mathcal{J}(\mathbb{R}^2) = \{\hat{J} \in \mathrm{End}(\mathbb{R}^2) \mid J\hat{J} + \hat{J}J = 0\}.$$

The space $\mathcal{J}(\mathbb{R}^2)$ is a Kähler manifold, where the complex structure on $T_J\mathcal{J}(\mathbb{R}^2)$ is given by $\hat{J} \mapsto -J\hat{J}$ and the metric and symplectic form are

$$\omega_{\mathcal{J}}(\hat{J}_1, \hat{J}_2) = \frac{1}{2}\mathrm{tr}(\hat{J}_1 J \hat{J}_2), \qquad g_{\mathcal{J}}(\hat{J}_1, \hat{J}_2) = \frac{1}{2}\mathrm{tr}(\hat{J}_1\hat{J}_2) \tag{4.38}$$

The group $\mathrm{SL}(2,\mathbb{R})$ acts on $\mathcal{J}(\Sigma)$ by conjugation

$$\mathrm{SL}(2,\mathbb{R}) \times \mathcal{J}(\mathbb{R}^2) \to \mathcal{J}(\mathbb{R}^2), \qquad \Psi_* J = \Psi J \Psi^{-1}.$$

This action preserves the Kähler structure on $\mathcal{J}(\mathbb{R}^2)$. Moreover, it is transitive with stabilizer $\mathrm{SO}(2)$ at the standard complex structure

$$J_0 := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

Therefore, $\mathcal{J}(\mathbb{R}^2) \cong \mathrm{SL}(2,\mathbb{R})/SO(2)$ and the next lemma gives an explicit formula for the composition $\mathbb{H} \cong \mathrm{SL}(2,\mathbb{R})/\mathrm{SO}(2) \cong \mathcal{J}(\mathbb{R}^2)$.

Lemma 4.3.2. *There exists a unique Kähler isometry $j : \mathbb{H} \to \mathcal{J}(\mathbb{R}^2)$ which is $SL(2,\mathbb{R})$-equivariant and satisfies $j(\boldsymbol{i}) = J_0$. It is given by the formula*

$$j : \mathbb{H} \to \mathcal{J}(\mathbb{R}^2), \qquad j(x + \boldsymbol{i}y) := \begin{pmatrix} \frac{x}{y} & -\frac{x^2+y^2}{y} \\ \frac{1}{y} & -\frac{x}{y} \end{pmatrix} \tag{4.39}$$

Proof. Let $z = x + \mathbf{i}y \in \mathbb{H}$ and define

$$\Psi := \frac{1}{\sqrt{y}} \begin{pmatrix} y & x \\ 0 & 1 \end{pmatrix} \in \mathrm{SL}(2, \mathbb{R}).$$

Then $z = \Psi_* \mathbf{i}$ and $\mathrm{SL}(2,\mathbb{R})$-equivariance implies

$$j(z) = \Psi J_0 \Psi^{-1} = \frac{1}{y} \begin{pmatrix} y & x \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & -x \\ 0 & y \end{pmatrix} = \begin{pmatrix} \frac{x}{y} & -\frac{x^2+y^2}{y} \\ \frac{1}{y} & -\frac{x}{y} \end{pmatrix}$$

The derivative $dj(\mathbf{i}) : T_{\mathbf{i}}\mathbb{H} \to T_{J_0}\mathcal{J}(\mathbb{R}^2)$ is given by

$$dj(\mathbf{i})\hat{z} = \begin{pmatrix} \hat{x} & -\hat{y} \\ -\hat{y} & -\hat{x} \end{pmatrix}$$

for $\hat{z} = \hat{x} + \mathbf{i}\hat{y} \in \mathbb{C}$. This is complex linear, since $dj(\mathbf{i})[\mathbf{i}\hat{z}] = J_0 dj(\mathbf{i})\hat{z}$. Moroever,

$$\omega_{\mathcal{J}}\left(\partial_x j(\mathbf{i}), \partial_y j(\mathbf{i})\right) = \frac{1}{2}\mathrm{tr}\left(\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}\right) = 1$$

shows that $dj(\mathbf{i})$ the symplectic structures. Then, by compatibility, it also intertwines the metrics. It follows now from the $\mathrm{SL}(2,\mathbb{R})$-invariance of the Kähler structures on $\mathcal{J}(\mathbb{R}^2)$ and $\mathbb{H}$ that j is a Kähler isometry. □

Denote by $\omega_0 = dx \wedge dy$ the standard area form on $\mathbb{R}^2$. Every $J \in \mathcal{J}(\mathbb{R}^2)$ defines a hermitian form on $(\mathbb{R}^2, J)$ defined by

$$h_J : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{C}, \qquad h_J(\cdot,\cdot) := \omega_0(\cdot, J\cdot) + \mathbf{i}\omega_0(\cdot,\cdot) \tag{4.40}$$

This is complex anti-linear in the first coordinate and complex linear in the second coordinate with respect to J. A direct computation shows that $h_{j(z)}$ has the matrix representation

$$h_{j(z)}(v, w) = v^t \, \frac{1}{\mathrm{Im}(z)} \begin{pmatrix} 1 & -\bar{z} \\ -z & |z|^2 \end{pmatrix} w. \tag{4.41}$$

for $v, w \in \mathbb{R}^2$.

4.3.3 Complex quadratic forms

Let $J \in \mathcal{J}(\mathbb{R}^2)$ and define the associated hermitian form h_J by (4.40). We denote the space of complex quadratic forms on $(\mathbb{R}^2, J)$ by

$$Q(J) := \{q : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{C} \,|\, (J, \mathbf{i})\text{-complex bilinear and symmetric}\}.$$

This carries the complex structure $q \mapsto \mathbf{i}q$ and the hermitian structure

$$\begin{aligned} g_Q(q_1, q_2) &:= \mathrm{Re}\left(\frac{\overline{q_1(v,v)}q_2(v,v)}{h_J(v,v)^2}\right), \\ \omega_Q(q_1, q_2) &:= \mathrm{Im}\left(\frac{\overline{q_1(v,v)}q_2(v,v)}{h_J(v,v)^2}\right), \end{aligned} \tag{4.42}$$

where both expressions do not depend on the choice of the vector $v \in \mathbb{R}^2 \backslash \{0\}$.

Identification with tangent vectors. The map

$$T_J\mathcal{J}(\mathbb{R}^2) \xrightarrow{\cong} Q(J), \qquad \hat{J} \mapsto q_{(J,\hat{J})} := h_J(\hat{J}\cdot,\cdot) \tag{4.43}$$

is a complex linear unitary isomorphism identifying tangent vectors with quadratic forms. The next lemma summarizes some properties of this map.

Remark 4.3.3. The identification $T_J\mathcal{J}(\mathbb{R}^2) \cong Q(J)$ is a consequence of our choice for the complex structure on $\mathcal{J}(\mathbb{R}^2)$ being multiplication by $-J$. If one considers the opposite complex structure, i.e. the one given on $T_J\mathcal{J}(\mathbb{R}^2)$ by multiplication with J, we would end up with an identification $Q(J) \cong T_J^*\mathcal{J}(\mathbb{R}^2)$. However, with this choice of complex structure, the map $j : \mathbb{H} \to \mathcal{J}(\mathbb{R}^2)$ defined by (4.39) would be complex antilinear.

Lemma 4.3.4 (Quadratic forms and tangent vectors).

1. *For $J \in \mathcal{J}(\mathbb{R}^2)$ and $\hat{J} \in T_J\mathcal{J}(\mathbb{R}^2)$ it holds*
$$h_J(\hat{J}v, w) = h_J(\hat{J}w, v) \tag{4.44}$$
for all $v, w \in \mathbb{R}^2$. In particular, $q_{(J,\hat{J})} := h_J(\hat{J}\cdot,\cdot) \in Q(J)$.

2. *For every $J \in \mathcal{J}(\mathbb{R}^2)$ the map (4.43) is a unitary isomorphism with respect to the structures defined in (4.38) and (4.42).*

3. *The collection of maps (4.43) is $SL(2,\mathbb{R})$-equivariant in the following sense: Let $J \in \mathcal{J}(\mathbb{R}^2)$, $\hat{J} \in T_J\mathcal{J}(\mathbb{R}^2)$, and $\Psi \in SL(2,\mathbb{R})$, then*
$$q_{\Psi_*(J,\hat{J})}(v,w) = q_{(J,\hat{J})}(\Psi^{-1}v, \Psi^{-1}w)$$
for all $v, w \in \mathbb{R}^2$.

Proof. Differentiating the equation $\omega_0(Jv, Jw) = \omega_0(v,w)$ it follows

$$\omega_0(\hat{J}v, Jw) + \omega_0(Jv, \hat{J}w) = 0.$$

Hence $\omega_0(\hat{J}v, Jw) = \omega_0(v, J\hat{J}w)$ shows that $\hat{J}$ is self-adjoint with respect to the inner product $\omega_0(\cdot, J\cdot)$. Moroever, $\omega_0(\hat{J}v, w) = -\omega_0(v, \hat{J}w)$ and then follows

$$h_J(\hat{J}v, w) = \omega_0(\hat{J}v, Jw) + \mathbf{i}\omega_0(\hat{J}v, w) = \omega_0(v, J\hat{J}w) + \mathbf{i}\omega_0(\hat{J}w, v) = h_J(\hat{J}w, v)$$

This completes the proof of (4.44).

For the second part, it follows from (4.44) that

$$||\hat{J}||^2 = \frac{1}{2}\left(\frac{h_J(\hat{J}^2v, v)}{h_J(v,v)} + \frac{h_J(\hat{J}^2Jv, Jv)}{h_J(v,v)}\right) = \frac{h_J(\hat{J}v, \hat{J}v)}{h_J(v,v)} = \frac{|h_J(v, \hat{J}v)|^2}{h_J(v,v)^2} = ||q_{\hat{J}}||^2$$

where we used in the penultimate equation that $(\mathbb{R}^2, J)$ is complex one-dimensional and hence $|h_J(v, \hat{J}v)|^2 = h_J(v,v)h_J(\hat{J}v, \hat{J}v)$. Hence (4.43) is an isometry. It is clearly complex linear and by compatibility it also intertwines the symplectic structures.

Finally, let $\Psi \in \mathrm{SL}(2,\mathbb{R})$ be given and compute

$$\begin{aligned} q_{\Psi_*(J,\hat{J})} &= h_{\Psi J\Psi^{-1}}(\Psi\hat{J}\Psi^{-1}\cdot,\cdot) = \omega_0(\Psi\hat{J}\Psi^{-1}\cdot, \Psi J\Psi^{-1}\cdot) + \omega_0(\Psi\hat{J}\Psi^{-1}\cdot,\cdot) \\ &= \omega_0(\hat{J}\Psi^{-1}\cdot, J\Psi^{-1}\cdot) + \omega_0(\hat{J}\Psi^{-1}\cdot, \Psi^{-1}\cdot) = q_{(J,\hat{J})}(\Psi^{-1}\cdot, \Psi^{-1}\cdot). \end{aligned}$$

This proves equivariance and the lemma. $\square$

Identification with covectors. The Riemannian metric on $\mathcal{J}(\mathbb{R}^2)$ defines a complex anti-linear isomorphism of the tangent bundle and cotangent bundle of $\mathcal{J}(\mathbb{R}^2)$. This is given by

$$T_J\mathcal{J}(\mathbb{R}^2) \to T_J^*\mathcal{J}(\mathbb{R}^2), \qquad \hat{J} \mapsto \left(\hat{J}' \mapsto \frac{1}{2}\mathrm{tr}\left(\hat{J}\hat{J}'\right)\right). \tag{4.45}$$

Combining this map with the isomorphism $j : \mathbb{H} \to \mathcal{J}(\mathbb{R}^2)$ defined by (4.39) yields an identification between the $T^*\mathbb{H}$ and the bundle of quadratic forms over $\mathcal{J}(\Sigma)$, which is complex anti-linear along the fibres.

Lemma 4.3.5 (Quadratic forms and covectors). *Define the map*

$$(j,q) : T^*\mathbb{H} \to \mathcal{J}(\mathbb{R}^2) \times Hom(\mathbb{R}^2 \otimes \mathbb{R}^2, \mathbb{C})$$
$$j(z,w) := j(z) := \begin{pmatrix} \frac{x}{y} & -\frac{x^2+y^2}{y} \\ \frac{1}{y} & -\frac{x}{y} \end{pmatrix}, \qquad q(z,w) := \begin{pmatrix} \bar{w} & -\bar{z}\bar{w} \\ -\bar{z}\bar{w} & \bar{z}^2\bar{w} \end{pmatrix} \tag{4.46}$$

where $z = x + \boldsymbol{i}y \in \mathbb{H}$ and $w \in \mathbb{C}$. Then the following holds:

1. *(4.46) is $SL(2,\mathbb{R})$-equivariant in the sense that*
$$q(\Psi(z,w))(\cdot,\cdot) = q(z,w)(\Psi^{-1}\cdot,\Psi^{-1}\cdot) \in Q(\Psi j(z)\Psi^{-1})$$
for every $\Psi \in SL(2,\mathbb{R})$ and $(z,w) \in T^*\mathbb{H}$.
2. *For every $z \in \mathbb{H}$ the fibre map $q(z,\cdot) : T_z^*\mathbb{H} \to Q(j(z))$ is a complex anti-linear isometry satisfying*
   (a) $g_Q(q(z,w_1), q(z,w_2)) = Im(z)^2 Re(w_1\bar{w}_2)$.
   (b) $\omega_Q(q(z,w_1), q(z,w_2)) = Im(z)^2 Im(w_1\bar{w}_2)$.
   (c) $q(z,\boldsymbol{i}w) = -\boldsymbol{i}q(z,w)$.

Proof. We leave it to the reader to check that the map (q,j) is indeed constructed by combining (4.43), (4.45), and (4.39). Since all these maps are $\mathrm{SL}(2,\mathbb{R})$-equivariant isometries, with (4.43) and j being complex linear and (4.45) being complex anti-linear, it follows then that q is a $\mathrm{SL}(2,\mathbb{R})$-equivariant and complex anti-linear isometry. □

Duality. In our discussion so far, we viewed a covector $J^* \in T_J^*\mathcal{J}(\mathbb{R}^2)$ as real-linear map $J^* : T_J\mathcal{J}(\mathbb{R}^2) \to \mathbb{R}$. This extends uniquely to a complex linear map from $T_J\mathcal{J}(\mathbb{R}^2) \to \mathbb{C}$ and the resulting complex linear dual pairing is given by

$$T_J^*\mathcal{J}(\mathbb{R}^2) \times T_J\mathcal{J}(\mathbb{R}^2) \to \mathbb{C}, \qquad \langle J^*, \hat{J}\rangle = J^*(\hat{J}) + \mathbf{i}J^*(J\hat{J}) \tag{4.47}$$

Identify $T_J^*\mathcal{J}(\mathbb{R}^2)$ with $Q(J)$ using (4.43) and (4.45). This identification is complex anti-linear and the dual pairing (4.47) takes the form

$$Q(J) \times T_J\mathcal{J}(\mathbb{R}^2) \to \mathbb{C}, \qquad \langle q, \hat{J}\rangle_{Q\times T\mathcal{J}} = \overline{\frac{q(\hat{J}v,v)}{h_J(v,v)}} \tag{4.48}$$

where the right hand side does not depend on $v \in \mathbb{R}^2\backslash\{0\}$. This pairing is complex anti-linear in the first coordinate and complex linear in the second one.

Lemma 4.3.6. *Define* $(j, q) : T^*\mathbb{H} \to Hom(\mathbb{R}^2 \times \mathbb{R}^2, \mathbb{C})$ *by (4.46). Then*

$$w\hat{z} = \langle q(z, w), dj(z)\hat{z}\rangle_{Q\times T\mathcal{J}} \tag{4.49}$$

for all $z \in \mathbb{H}$ *and* $\hat{z}, w \in \mathbb{C}$. *Here we think of* $\hat{z} \in T_z\mathbb{H}$, $w \in T_z^*\mathbb{H}$, *and define the right hand side by (4.48).*

Proof. This follows directly from the construction of the dual pairing. Alternatively, one may use Lemma 4.3.5 to verify the formula at $z = \mathbf{i}$ and then use $\mathrm{SL}(2, \mathbb{R})$ equivariance of both sides in (4.49) to complete the proof. □

4.3.4 A moment map for Möbiustransformations

Lemma 4.3.7. *The action of* $SL(2, \mathbb{R})$ *on* $\mathcal{J}(\mathbb{R}^2)$ *and* $\mathbb{H}$ *is Hamiltonian and generated by the equivariant moment maps*

$$\mu_{\mathcal{J}} : \mathcal{J}(\mathbb{R}^2) \to \mathfrak{sl}^*(2, \mathbb{R}), \qquad \langle\mu_{\mathcal{J}}(J), \xi\rangle := -tr(J\xi),$$
$$\mu_{\mathbb{H}} : \mathbb{H} \to \mathfrak{sl}^*(2, \mathbb{R}), \qquad \langle\mu_{\mathbb{H}}(z), \xi\rangle := -tr(j(z)\xi)$$

for $\xi \in \mathfrak{sl}(2, \mathbb{R})$. *Here* $j : \mathbb{H} \to \mathcal{J}(\mathbb{R}^2)$ *is the isomorphism defined by (4.39).*

Proof. Let $J \in \mathcal{J}(\mathbb{R}^2)$, $\hat{J} \in T_J\mathcal{J}(\mathbb{R}^2)$ and $\xi \in \mathfrak{sl}(2, \mathbb{R})$. The infinitesimal action of ξ at J is given by $L_J\xi = [\xi, J]$ and therefore

$$\omega_{\mathcal{J}}\left(L_J\xi, \hat{J}\right) = \frac{1}{2}\mathrm{tr}\left([\xi, J]J\hat{J}\right) = \frac{1}{2}\left(-\mathrm{tr}\left(\xi\hat{J}\right) - \mathrm{tr}\left(J\xi J\hat{J}\right)\right) = -\mathrm{tr}(\hat{J}\xi)$$

This proves the first part of the lemma and the second part follows from this by equivariance of j. □

4.4 Teichmüller space as symplectic quotient

Throughout this section let (Σ, ρ) denote a closed 2-dimensional surface with fixed area form ρ and $\mathrm{genus}(\Sigma) \geq 2$. Let $X = \mathbb{H}$ be the upper half plane and denote by $P \to \Sigma$ the $\mathrm{SL}(2, \mathbb{R})$ frame bundle. Then

$$P(X) \cong \mathcal{J}(\Sigma) = \left\{J \in \Omega^0(\Sigma, \mathrm{End}(T\Sigma)) \,|\, J^2 = -\mathbb{1},\ \rho(\cdot, J\cdot) > 0\right\}$$

is the space of complex structure on Σ, compatible with the orientation determined by ρ. It follows from Theorem 4.2.4 that the natural action of $\mathrm{Ham}(\Sigma, \rho)$ on $\mathcal{J}(\Sigma)$ is Hamiltonian with moment being $2(K_J - c)\rho$, where K_J denotes the Gaussian curvature of the metric $\rho(\cdot, J\cdot)$ and $c = \frac{2\pi(2\mathrm{genus}(\Sigma)-2)}{\mathrm{vol}(\Sigma,\rho)}$. The details of this are given in Theorem 4.4.2 below. The resulting Marsden-Weinstein quotient is

$$\widetilde{\mathcal{T}}(\Sigma, \rho) := \mathcal{J}(\Sigma)//\mathrm{Ham}(\Sigma, \rho) = \{J \in \mathcal{J}(\Sigma) \,|\, 2K_J = c\}/\mathrm{Ham}(\Sigma, \rho)$$

and it follows from general principles that this carries a symplectic structure. This moduli space fibres over the Teichmüller space $\mathcal{T}(\Sigma)$, which has the description

$$\mathcal{T}(\Sigma) := \mathcal{J}(\Sigma)/\mathrm{Diff}_0(\Sigma) \cong \{J \in \mathcal{J}(\Sigma) \,|\, K_J = c\}/\mathrm{Symp}(\Sigma, \rho).$$

We show that the fibres of $\widetilde{\mathcal{T}}(\Sigma,\rho)$ over $\mathcal{T}(\Sigma)$ are symplectic submanifolds and that the symplectic form on $\widetilde{\mathcal{T}}(\Sigma,\rho)$ descends in a canonical way to $\mathcal{T}(\Sigma)$. The resulting symplectic form is the Weil–Petersson symplectic form on Teichmüller space and does not depend on the fixed area form ρ (apart from scaling). The final subsection recalls the well-known fact that Teichmüller space $\mathcal{T}(\Sigma)$ can be embedded into the representation variety $\mathrm{Ham}(\pi_1(\Sigma),\mathrm{SL}(2,\mathbb{R}))/\mathrm{SL}(2,\mathbb{R})$. The image consists of the so-called Fuchsian representations and this subsection can be viewed as model case for our discussion of the quasi-Fuchsian moduli space in the next section.

4.4.1 The Hamiltonian quotient

Let (Σ,ρ) be a closed, oriented 2-dimensional surface with fixed area form ρ and with $\mathrm{genus}(\Sigma)\geq 2$. Denote by P its $\mathrm{SL}(2,\mathbb{R})$ frame bundle and consider the associated bundle $P(\mathbb{H}):=P\times_{\mathrm{SL}(2,\mathbb{R})}\mathbb{H}$. Define $j:\mathbb{H}\to\mathcal{J}(\mathbb{R}^2)$ by (4.39). Then

$$P(\mathbb{H})\hookrightarrow\mathrm{End}(T\Sigma),\qquad [(z,\theta),\zeta]\mapsto\theta^{-1}j(\zeta)\theta \tag{4.50}$$

is a well-defined embedding which yields the identification

$$\mathcal{S}(P,\mathbb{H})=\mathcal{J}(\Sigma):=\left\{J\in\Omega^0(\Sigma,\mathrm{End}(T\Sigma))\,|\,J^2=-\mathbb{1},\ \rho(\cdot,J\cdot)>0\right\}.$$

The next lemma summarizes same important properties of this setup.

Lemma 4.4.1.

1. *Any torsion-free $SL(2,\mathbb{R})$ connection on $T\Sigma$ induces connections on $P(\mathbb{H})$ and End$(T\Sigma)$ which are compatible with respect to the embedding (4.50).*
2. *The inclusion (4.50) is Symp(Σ,ρ)-equivariant.*

Proof. The proof is a matter of unravelling the definition. We leave this as an exercise to the reader. □

Theorem 4.4.2. *The action of Ham(Σ,ρ) on $\mathcal{J}(\Sigma)$ is Hamiltonian and generated by the moment map*

$$\underline{\mu}:\mathcal{J}(\Sigma)\to\Omega^2(\Sigma),\qquad \underline{\mu}(J)=2(K_J-c)\rho$$

where K_J denotes the Gaussian curvature of the metric $\rho(\cdot,J\cdot)$ and $c=\frac{2\pi(2genus(\Sigma)-2)}{vol(\Sigma,\rho)}$.

Remark 4.4.3. This moment map was first established by Quillen and it has been generalized by Fujiki [48] in the integrable case and Donaldson [34] to the space of compatible almost complex structures on a symplectic manifold.

Proof. Let ∇ be the Levi-Civita connections of the metric $\rho(\cdot,J\cdot)$. By the Newlander-Nierenberg theorem, every complex structure $J\in\mathcal{J}(\Sigma)$ is integrable and in particular $\nabla J=0$. We calculate the moment map $\underline{\mu}$ using Theorem 4.2.4. Since $\nabla J=0$, the two terms $\omega(\nabla J,\nabla J)$ and $c(\nabla\mu_J)$ both vanish. The only remaining term is

$$\langle\mu_J,R^\nabla\rangle=-\mathrm{tr}(JR^\nabla)=2K_J\rho.$$

We used in the last equation that the Gaussian curvature K and Riemann curvature tensor R^∇ are related by

$$K_J := \frac{\langle R^\nabla(u,v)v, u\rangle}{|u|^2 \cdot |v|^2 - \langle u,v\rangle^2}, \qquad R^\nabla(u,v) = -K_J \rho(u,v) J$$

for all $u, v \in \mathrm{Vect}(\Sigma)$. The constant $c = 2\pi(2\mathrm{genus}(\Sigma) - 2)/\mathrm{vol}(\Sigma,\rho)$ is determined by the Gauss–Bonnet theorem and guarantees that $\underline{\mu}(J)$ is an exact 2-form. □

The Marsden-Weinstein quotient of $\mathcal{J}(\Sigma)$ by $\mathrm{Ham}(\Sigma,\rho)$ is defined as the quotient of $\underline{\mu}^{-1}(0)$ by $\mathrm{Ham}(\Sigma,\rho)$. This yields the quotient space

$$\widetilde{\mathcal{T}}(\Sigma,\rho) := \left\{ J \in \mathcal{J}(\Sigma) \,\middle|\, K_J = \frac{2\pi(2\mathrm{genus}(\Sigma) - 2)}{\mathrm{vol}(\Sigma,\rho)} \right\} \Big/ \mathrm{Ham}(\Sigma,\rho). \tag{4.51}$$

It follows from the formal properties of a moment map that the symplectic form on $\mathcal{J}(\Sigma)$ descends to a symplectic form on $\widetilde{\mathcal{T}}(\Sigma,\rho)$.

4.4.2 Donaldson's description of Teichmüller space

The Teichmüller space of Σ is defined as

$$\mathcal{T}(\Sigma) := \mathcal{J}(\Sigma)/\mathrm{Diff}_0(\Sigma).$$

Early and Eells [41] showed that $\mathrm{Diff}_0(\Sigma)$ acts freely on $\mathcal{J}(\Sigma)$ and $\mathcal{J}(\Sigma) \to \mathcal{T}(\Sigma)$ is a fibre bundle with fibre $\mathrm{Diff}_0(\Sigma)$. It is a classical theorem of Teichmüller that $\mathcal{T}(\Sigma)$ is homeomorphic to $\mathbb{R}^{6g-6}$ where $g = \mathrm{genus}(\Sigma)$. Fischer and Tromba [46] were able to prove Teichmüller's theorem directly from the fibre bundle description of Early and Eells. A corollary of this is the well-known fact that $\mathrm{Diff}_0(\Sigma,\rho)$ is contractible. This can also be proven directly and we will make use of this fact in the following.

Two descriptions of Teichmüller space

Lemma 4.4.4. *Assume genus*$(\Sigma) \geq 2$. *Then the inclusion Symp*$_0(\Sigma,\rho) \subset$ *Diff*$_0(\Sigma)$ *is a homotopy equivalence.*

Proof. This follows from a parametrized version of Moser isotopy (which I learned from [96]). Denote

$$\mathcal{V}(\Sigma) := \left\{ \omega \in \Omega^2(\Sigma) \,\middle|\, \omega \text{ is an area form and } \int_\Sigma \omega = \int_\Sigma \rho \right\}.$$

Standard Moser isotopy arguments show that there exists a continuous map

$$\mathcal{V} \to \mathrm{Diff}_0(\Sigma), \qquad \omega \mapsto \psi_\omega$$

satisfying $\psi_\omega^*\omega = \rho$. Then the map

$$\mathrm{Diff}_0(\Sigma) \to \mathcal{V}(\Sigma) \times \mathrm{Symp}_0(\Sigma,\rho), \qquad \psi \mapsto (\psi^*\rho, \psi \circ \psi_{\psi^*\rho})$$

is a homeomorphism with inverse $(\omega,\phi) \mapsto \phi \circ \psi_\omega^{-1}$. Since $\mathcal{V}(\Sigma)$ is convex, it follows that $\mathrm{Diff}_0(\Sigma)$ and $\mathrm{Symp}_0(\Sigma,\rho)$ are homotopy equivalent. □

Proposition 4.4.5. *The inclusion $\underline{\mu}^{-1}(0) \hookrightarrow \mathcal{J}(\Sigma)$ yields an isomorphism*

$$\left\{ J \in \mathcal{J}(\Sigma) \,\middle|\, K_J = \frac{2\pi(2genus(\Sigma)-2)}{vol(\Sigma,\rho)} \right\} \Big/ Symp_0(\Sigma,\rho) \cong \mathcal{T}(\Sigma).$$

Proof. Denote $c := 2\pi(2\mathrm{genus}(\Sigma)-2)/\mathrm{vol}(\Sigma,\rho)$. By Theorem 4.4.2, the moment map is given by $\underline{\mu}(J) = 2K_J\rho - 2c$ and this is clearly $\mathrm{Symp}_0(\Sigma,\rho)$-equivariant. In particular, $\underline{\mu}^{-1}(0)$ is preserved by the action of $\mathrm{Symp}_0(\Sigma,\rho)$ and the quotient $\underline{\mu}^{-1}(0)/\mathrm{Symp}_0(\Sigma,\rho)$ is well-defined.

The uniformization theorem shows that for every $J \in \mathcal{J}(\Sigma)$ there exists a unique metric g_J which is compatible with J and has constant curvature $K_{g_J} \equiv c$. Let ρ_J denote the area form of g_J. The triple (J, g_J, ρ_J) is satisfies the naturality condition

$$\phi^* g_J = g_{\phi^* J}, \qquad \phi^*\rho_J = \rho_{\phi^* J} \tag{4.52}$$

for every diffeomorphism $\phi : \Sigma \to \Sigma$, by of uniqueness of the metric g_J.

We show surjectivity: For $J \in \mathcal{J}(\Sigma)$ it follows from Gauss-Bonnet that $\mathrm{vol}(\Sigma,\rho_J) = \mathrm{vol}(\Sigma,\rho)$. Hence, by Moser isotopy, there exists $\phi \in \mathrm{Diff}_0(\Sigma)$ such that $\phi^*\rho_J = \rho$ and hence $\phi^* J \in \underline{\mu}^{-1}(0)$.

We show injectivity: Let $J_1, J_2 \in \mathcal{J}(\Sigma)$ with $\rho_{J_1} = \rho = \rho_{J_2}$ be given and suppose that $J_1 = \phi^* J_2$ for some $\phi \in \mathrm{Diff}_0(\Sigma)$. Then $\phi^*\rho_{J_2} = \rho_{J_1}$ implies $\phi \in \mathrm{Symp}(\Sigma,\rho)$. By Lemma 4.4.4, it now follows $\phi \in \mathrm{Diff}_0(\Sigma) \cap \mathrm{Symp}(\Sigma,\rho) = \mathrm{Symp}_0(\Sigma,\rho)$ and this completes the proof. □

Complex structure on Teichmüller space

The next lemma calculates the infinitesimal action of $\mathrm{Diff}_0(\Sigma)$ on $\mathcal{J}(\Sigma)$.

Lemma 4.4.6. *Let $J \in \mathcal{J}(\Sigma)$, $v \in Vect(\Sigma)$. The Lie derivative of J is given by*

$$\mathcal{L}_v J = 2J\bar{\partial}_J v$$

where $\bar{\partial}$ is the canonical Cauchy–Riemann operator on $(T\Sigma, J)$ determined by J.

Proof. Denote by ∇ the Levi-Civita connection for the metric $g := \rho(\cdot, J\cdot)$ and let $v, w \in \mathrm{Vect}(\Sigma)$. Then

$$(\mathcal{L}_v J)w = \mathcal{L}_v(Jw) - J(\mathcal{L}_v w) = [Jw, v] - J[w,v] = J\nabla_w v - \nabla_{Jw} v = 2J(\bar{\partial}_J v)w$$

and this completes the proof. □

It follows from the Riemann–Roch theorem, that Σ admits no holomorphic vector fields. The infinitesimal action is hence injective and the orbits $\mathrm{Diff}_0(\Sigma)\cdot J$ are complex submanifolds of $\mathcal{J}(\Sigma)$. The complex structure on $\mathcal{J}(\Sigma)$ thus descends canonically to the Teichmüller space $\mathcal{T}(\Sigma) = \mathcal{J}(\Sigma)/\mathrm{Diff}_0(\Sigma)$.

Symplectic structure on Teichmüller space

Denote by $\widetilde{\mathcal{T}}(\Sigma,\rho)$ the Marsden-Weinstein quotient (4.51) and define

$$H := \mathrm{Symp}_0(\Sigma,\rho)/\mathrm{Ham}(\Sigma,\rho).$$

Recall that $\mathrm{Ham}(\Sigma,\rho) < \mathrm{Symp}_0(\Sigma,\rho)$ is a normal subgroup (see [85], Proposition 10.2) and therefore H is a well-defined quotient group. It follows form Proposition 4.4.5 that Teichmüller space has the description $\mathcal{T}(\Sigma) \cong \widetilde{\mathcal{T}}(\Sigma,\rho)/H$.

Remark 4.4.7. The flux homomorpism associates to every path $[0,1] \to \mathrm{Symp}_0(\Sigma,\rho)$, $t \mapsto \psi_t$, a cohomology class in $H^1(\Sigma,\mathbb{R})$ defined by

$$\mathrm{Flux}(\{\psi_t\}) := \int_0^1 [\iota(\partial_t\psi_t)\omega]\, dt \in H^1(\Sigma,\mathbb{R}).$$

Since $\pi_1(\mathrm{Symp}_0(\Sigma))$ is trivial by Lemma 4.4.4, it follows that the the flux homomorphism descends to an isomorphism

$$\mathrm{Flux}:\ H := \mathrm{Symp}_0(\Sigma,\rho)/\mathrm{Ham}(\Sigma,\rho) \xrightarrow{\cong} H^1(\Sigma,\mathbb{R}).$$

See [85], Proposition 10.18 for more details.

Lemma 4.4.8. *The H-orbits in $\widetilde{\mathcal{T}}(\Sigma,\rho)$ are symplectic submanifolds.*

Proof. Let $[J] \in \widetilde{\mathcal{T}}(\Sigma,\rho)$. Then, by Lemma 4.4.6, we have

$$T_{[J]}\left(H\cdot[J]\right) = \frac{\{J\bar\partial_J v \,|\, v \in \mathrm{Vect}(\Sigma) \text{ with } d\iota(v)\rho = 0\}}{\{J\bar\partial_J v \,|\, v \in \mathrm{Vect}(\Sigma) \text{ with } \iota(v)\rho \text{ exact}\}}.$$

Next, consider the subspace $S_J \subset T_J\mathcal{J}(\Sigma)$ defined by

$$S_J := \{\bar\partial_J v \,|\, v \in \mathrm{Vect}(\Sigma) \text{ with } d\iota(v)\rho = 0 = d\iota(Jv)\rho\}.$$

This is a complex and hence symplectic subspace. The natural projection from $T_J\mathcal{J}(\Sigma)$ to $T_{[J]}\widetilde{\mathcal{T}}(\Sigma,\rho)$ restricts to a symplectic isomorphism from S_J to $T_{[J]}\left(H\cdot[J]\right)$ by the Hodge decomposition theorem. Hence $T_{[J]}\left(H\cdot[J]\right)$ is also symplectic and this proves the lemma. □

Lemma 4.4.9. *Let (Q,ω) be a symplectic manifold and let G be a Lie group acting symplectically, properly and freely on Q. Suppose that all G orbits are symplectic submanifold of Q. Then Q/G carries a natural symplectic structure which is obtained by declaring that $(T_q(G\cdot q))^\omega \to T_{[q]}Q/G$ is a symplectic isomorphism for every $q \in Q$.*

Proof. The tangent space $T_q(G\cdot q)$ of the G-orbit through q is by assumption symplectic and so its symplectic complement $(T_q(G\cdot q))^\omega$ is also a symplectc. The induced symplectic form on $T_{[q]}Q/G$ does not on the representative q, because G acts symplectically on Q. It follows that Q/G carries a well-defined non-degenerated 2-form $\omega_{Q/G} \in \Omega^2(Q/G)$. It remains to show that $\omega_{Q/G} \in \Omega^2(Q/G)$ is closed. We have

$$\begin{aligned} d\omega_{Q/G}(v_1,v_2,v_3) &= \omega_{Q/G}([v_1,v_2],v_3) + \omega_{Q/G}([v_2,v_3],v_1) + \omega_{Q/G}([v_3,v_1],v_2) \\ &\quad - \mathcal{L}_{v_3}(\omega_{Q/G}(v_1,v_2)) - \mathcal{L}_{v_1}(\omega_{Q/G}(v_2,v_3)) - \mathcal{L}_{v_2}(\omega_{Q/G}(v_3,v_1)) \end{aligned}$$

for $v_1, v_2, v_3 \in \mathrm{Vect}(Q/G)$. Let $\tilde{v}_j \in \mathrm{Vect}(Q)$ be the unique lift of v_j with $\tilde{v}_j(q) \in (T_q(G \cdot q))^\omega$ for all $q \in Q$. Then follows

$$\omega_{Q/G}([v_1, v_2], v_3) = \omega([\tilde{v}_1, \tilde{v}_2], \tilde{v}_3)$$

since $[\tilde{v}_1, \tilde{v}_2]$ projects to $[v_1, v_2]$. Moreover, $\mathcal{L}_{v_3}(\omega_{Q/G}(v_1, v_2)) = \mathcal{L}_{\tilde{v}_3}(\omega(\tilde{v}_1, \tilde{v}_2))$. Similar equations hold for the other terms in $d\tilde{\omega}$ and hence

$$d\omega_{Q/G}(v_1, v_2, v_3) = d\omega(\tilde{v}_1, \tilde{v}_2, \tilde{v}_3) = 0.$$

This completes the proof of the lemma. □

Weil–Petersson metric on Teichmüller space

It follows from Lemma 4.4.9 and Lemma 4.4.8 that $\mathcal{T}(\Sigma)$ carries a natural symplectic structure arising from the description

$$\mathcal{T}(\Sigma) = \widetilde{\mathcal{T}}(\Sigma, \rho)/H = \big(\underline{\mu}^{-1}(0)/\mathrm{Ham}(\Sigma, \rho)\big)/H.$$

The next proposition shows that this symplectic structure agrees up to scaling with the Weil–Petersson symplectic form, where the scaling factor is determined by the total volume $V = \mathrm{vol}(\Sigma, \rho)$. In particular, the induced symplectic structure on $\mathcal{T}(\Sigma)$ is independent of ρ apart from scaling.

Proposition 4.4.10 (**Weil–Petersson symplectic form**)**.** *For $J \in \mathcal{J}(\Sigma)$ denote by g_J the unique metric which is compatible with J and satisfies $K_{g_J} \equiv c$, where $c := 2\pi(2genus(\Sigma) - 2)/vol(\Sigma, \rho)$. Let h_J be the hermitian form*

$$h_J : T\Sigma \otimes T\Sigma \to \mathbb{C}, \qquad h_J(u, v) = g_J(u, v) + \boldsymbol{i} g_J(Ju, v).$$

By Lemma 4.4.6, the tangent space of $[J] \in \mathcal{T}(\Sigma)$ is given by

$$T_{[J]}\mathcal{T}(\Sigma) = coker(\bar{\partial}_J : \Omega^0(\Sigma, T\Sigma) \to \Omega^{0,1}(\Sigma, T\Sigma)) \cong \mathcal{H}^{0,1}_J(T\Sigma)$$

where $\mathcal{H}^{0,1}_J = \{\hat{J} \in \Omega^{0,1}_J(\Sigma, T\Sigma) \,|\, \bar{\partial}^*_J \hat{J} = 0\}$ denotes the harmonic representatives. The symplectic form on Teichmüller space is given by the formula

$$\omega_{\mathcal{T}} : \mathcal{H}^{0,1}_J(T\Sigma) \times \mathcal{H}^{0,1}_J(T\Sigma) \to \mathbb{R}, \qquad \omega_{\mathcal{T}}(\hat{J}_1, \hat{J}_2) = 2Re \int_\Sigma h_J(\hat{J}_1 \wedge \hat{J}_2) \tag{4.53}$$

and the corresponding Kähler metric

$$g_{\mathcal{T}} : \mathcal{H}^{0,1}_J(T\Sigma) \times \mathcal{H}^{0,1}_J(T\Sigma) \to \mathbb{R}, \qquad g_{\mathcal{T}}(\hat{J}_1, \hat{J}_2) = -2Im \int_\Sigma h_J(\hat{J}_1 \wedge \hat{J}_2) \tag{4.54}$$

is the Weil–Petersson metric.

Proof. The formulas (4.53) and (4.54) are $\mathrm{Diff}(\Sigma)$-invariant and descend to well-defined pairings on Teichmüller space. It follows from the Gauss–Bonnet theorem that $\mathrm{vol}(\Sigma, g_J) = \mathrm{vol}(\Sigma, \rho)$. Hence, by Moser isotopy and naturally (4.52), there exists $\phi \in \mathrm{Diff}_0(\Sigma)$ such that $\phi^* J \in \underline{\mu}^{-1}(0)$. We may therefore assume in the following that ρ is the volume form of g_J.

There are two point of views to understand $\hat{J}_1, \hat{J}_2 \in T_J\mathcal{J}(\Sigma)$, namely as sections in $\Omega^0(\Sigma, \mathrm{End}(T\Sigma))$ or as 1-forms in $\Omega^{0,1}_J(\Sigma, T\Sigma)$. These two perspectives are related by the formula

$$\frac{1}{2}\mathrm{tr}\left(\hat{J}_1\hat{J}_2\right)\rho + \frac{\mathbf{i}}{2}\mathrm{tr}\left(\hat{J}_1 J\hat{J}_2\right)\rho = -2\mathbf{i}h_J(\hat{J}_1 \wedge \hat{J}_2) \tag{4.55}$$

The left-hand side of this expression integrates to the Kähler structure on $\mathcal{J}(\Sigma)$. The right-hand side yields the Kähler structure (4.54) and (4.53). Therefore, it only remains to verify that every $\hat{J} \in \mathcal{H}^{0,1}_J(T\Sigma)$ is tangent to $\underline{\mu}^{-1}(0)$ and in the symplectic complement of the orbit $T_J(\mathrm{Diff}_0(\Sigma)\cdot J)$. Let $F:\Sigma \to \mathbb{R}$ be a Hamiltonian. Then Theorem 4.4.2 yields

$$\partial_{\hat{J}}\int_\Sigma 2F(K_J - c)\rho = \int_\Sigma \mathrm{tr}\left(-J(\bar{\partial}_J v_F)J\hat{J}\right)\rho = -\langle\bar{\partial}v_F, \hat{J}\rangle_{L^2} = -\langle v_F, \bar{\partial}^*\hat{J}\rangle_{L^2} = 0.$$

This shows that $\hat{J}$ is indeed tangential to $\underline{\mu}^{-1}(0)$. Moreover, for $v \in \mathrm{Vect}(\Sigma)$ we have

$$\int_\Sigma \mathrm{tr}\left(\hat{J}J(\bar{\partial}_J v)\right)\rho = \langle\hat{J}, J\bar{\partial}_J v\rangle_{L^2} = \langle\bar{\partial}^*_J\hat{J}, Jv\rangle_{L^2} = 0.$$

Hence, by Lemma 4.4.6, $\hat{J}$ is in the symplectic complement of $T_J(\mathrm{Diff}_0(\Sigma)\cdot J)$ and this completes the proof. □

4.4.3 The space of Fuchsian representations

The moduli space of Fuchsian representations of Σ is the space

$$\mathcal{F}(\Sigma) := \frac{\{\rho \in \mathrm{Hom}(\pi_1(\Sigma), \mathrm{PSL}(2,\mathbb{R}))\,|\,\text{discrete and properly discontinuous}\}}{\text{conjugation}}.$$

This is an open subset of the $\mathrm{PSL}(2,\mathbb{R})$ representation variety of Σ.

Remark 4.4.11. We suppressed the dependency of $\mathcal{F}(\Sigma)$ on the choice of a basepoint in Σ, since the moduli spaces for different choices are canonically isomorphic: Any path $\eta:[0,1]\to\Sigma$ with $\eta(0)=z_0$ and $\eta(1)=z_1$ induces an isomorphism $\pi_1(\Sigma,z_0)\to\pi_1(\Sigma,z_1)$, $[\gamma]\mapsto[\bar{\eta}\gamma\eta]$, defined by concatenation. This yields an identification of the corresponding Fuchsian moduli spaces which is independent of the choice of η.

The identification of the Fuchsian moduli space with Teichmüller spaces depends on the Dehn–Nielson–Baer theorem. Denote by

$$\mathrm{Out}(\pi_1(\Sigma)) := \mathrm{Aut}(\pi_1(\Sigma))/\mathrm{Inn}(\pi_1(\Sigma))$$

the space of outer automorphism of $\pi_1(\Sigma)$. Let $f \in \mathrm{Diff}(\Sigma)$ and fix a basepoint $z_0 \in \Sigma$. Let $\eta:[0,1]\to\Sigma$ be a path with $\eta(0)=z_0$ and $\eta(1)=f(z_0)$ and define $f_* \in \mathrm{Out}(\pi_1(\Sigma,z_0))$ by

$$f_*[\gamma] := [\eta f(\gamma)\bar{\eta}].$$

In the quotient space of outer isomorphism, f_* does not depend on the choice of the path η.

Theorem 4.4.12 (Dehn–Nielsen–Baer Theorem). *The map*

$$\mathit{Diff}(\Sigma) \to \mathit{Out}(\pi_1(\Sigma)), \qquad f \mapsto f_*$$

is surjective with kernel $\mathit{Diff}_0(\Sigma)$*. Therefore,* $\mathit{Diff}(\Sigma)/\mathit{Diff}_0(\Sigma) \cong \mathit{Out}(\pi_1(\Sigma))$*.*

Proof. See [42], Theorem 8.1 and Theorem 1.13. □

Theorem 4.4.13 (Fuchsian model for Teichmüller space). *Uniformization associates to every* $J \in \mathcal{J}(\Sigma) \cup \mathcal{J}(\bar{\Sigma})$ *a biholomorphic map* $\phi : \tilde{\Sigma} \to \mathbb{H}$ *between the universal cover of* (Σ, J) *and the upper half plane. The push-forward of the deck-transformation action of* $\pi_1(\Sigma)$ *on* $\tilde{\Sigma}$ *yields a representation* $\pi_1(\Sigma) \to PSL(2, \mathbb{R})$*. This construction is descends to a well-defined isomorphism*

$$\mathcal{T}(\Sigma) \cup \mathcal{T}(\bar{\Sigma}) \xrightarrow{\cong} \mathcal{F}(\Sigma)$$

where the right hand side is the disjoint union of the Teichmüller spaces obtained from both orientations of Σ*.*

Proof. We describe both directions of the isomorphism in the following:

Step 1: *Construction of the map* $\mathcal{T}(\Sigma) \cup \mathcal{T}(\bar{\Sigma}) \to \mathcal{F}(\Sigma)$*.*

Let J be a complex structure on Σ. Denote by $\tilde{\Sigma}$ the universal cover of Σ and by $\tilde{J}$ the lifted complex structure. By uniformization, there exists a biholomorphic map

$$\phi : \tilde{\Sigma} \to \mathbb{H}$$

which is unique up to postcomposing with an element of $\mathrm{Aut}(\mathbb{H}) \cong \mathrm{PSL}(2, \mathbb{R})$. The fundamental group $\pi_1(\Sigma)$ acts by deck-transformations on $\tilde{\Sigma}$. The pushforward of this action under ϕ yields a holomorphic action of $\pi_1(\Sigma)$ on $\mathbb{H}$ and thus a representation $\rho : \pi_1(\Sigma) \to \mathrm{PSL}(2, \mathbb{R})$. Postcomposing ϕ with an element of $\mathrm{PSL}(2, \mathbb{R})$ corresponds to conjugation of the representation and we obtain a well-defined map

$$\mathcal{J}(\Sigma) \to \mathcal{F}(\Sigma)$$

We show in the following that this map descends to $\mathcal{T}(\Sigma) \cup \mathcal{T}(\bar{\Sigma})$. To be precise, fix a base point $z_0 \in \Sigma$ and identify the universal covering of Σ with

$$\tilde{\Sigma} = \{\beta : [0, 1] \to \Sigma \,|\, \beta \text{ is a smooth path with } \beta(0) = z_0\}/\sim$$

where two paths are identified if they are homotopic with fixed endpoints. Let $f : [0, 1] \to \mathrm{Diff}(\Sigma)$ be an isotopy with $f_0 = \mathrm{id}$ and define

$$\tilde{f} : \tilde{\Sigma} \to \tilde{\Sigma}, \qquad \tilde{f}[\beta] := [t \mapsto f_t(\beta(t))] \,.$$

One readily checks that $\tilde{f}$ is an $\pi_1(\Sigma, z_0)$-equivariant lift of f_1. In particular $\tilde{f}^* \tilde{J}$ is a lift of $f_1^* J$ and $\phi \circ \tilde{f}$ yields the same representation as ϕ.

Step 2: *Construction of the map* $\mathcal{F}(\Sigma) \to \mathcal{T}(\Sigma) \cup \mathcal{T}(\bar{\Sigma})$.

First, we choose some reference data: Let $J_0 \in \mathcal{J}(\Sigma)$. By uniformization there exists a discrete and properly discontinuous representation $\rho_0 : \pi_1(\Sigma) \to \mathrm{PSL}(2,\mathbb{R})$ and a biholomorphic map

$$\phi_0 : (\Sigma, J_0) \to \mathbb{H}/\Gamma \quad \text{with} \quad \Gamma := \rho_0(\pi_1(\Sigma)).$$

Next, let $\rho : \pi_1(\Sigma) \to \mathrm{PSL}(2,\mathbb{R})$ be a discrete and properly discontinuous representation. Then $\mathbb{H}/\rho(\pi_1(\Sigma)) \cong \Sigma \cong \mathbb{H}/\Gamma$ as differentiable manifolds and there exists a diffeomorphism

$$f : \mathbb{H}/\Gamma \to \mathbb{H}/\rho(\pi_1(\Sigma)).$$

Without specifying the basepoints, f induces an isomorphism

$$f_* : \Gamma \cong \pi_1(\mathbb{H}/\Gamma) \to \pi_1(\mathbb{H}/\rho(\pi_1(\Sigma))) \cong \rho(\pi_1(\Sigma))$$

which is well-defined up to conjugation. It follows from the Dehn–Nielson–Baer theorem that the diffeomorphism f can be chosen in such a way that $f_* = \rho \circ \rho_0^{-1}$ holds up to conjugation and this determines f up to isotopy. Define $J_\rho \in \mathcal{J}(\Sigma)$ as the pullback of the complex structure on $\mathbb{H}/\rho(\pi_1(\Sigma))$ under the map $(f \circ \phi_0)$. This induces a map $\mathcal{F}(\Sigma) \to \mathcal{T}(\Sigma) \cup \mathcal{T}(\bar{\Sigma})$.

The construction of this map does not depend on the reference data (J_0, ρ_0, ϕ_0). Let (J_1, ρ_1, ϕ_1) be a difference choice for the reference data and let $\rho : \pi_1(\Sigma) \to \mathrm{PSL}(2,\mathbb{R})$ be a given representation. These yield complex structures J_ρ^0 and J_ρ^1 respectively where J_ρ^1 agrees with the pullback of J_ρ^0 under $h := (f_0 \circ \phi_0)^{-1} \circ (f_1 \circ \phi_1)$. It follows from the construction that $h_* \in \mathrm{Out}(\pi_1(\Sigma))$ is the identity. Hence $h \in \mathrm{Diff}_0(\Sigma)$ by the Dehn–Nielson–Baer theorem and both complex structures descend to the same element in Teichmüller space.

Both constructions are inverse to each other and this completes the proof of the theorem. □

4.5 Hyperkähler thickening of Teichmüller space

In this section, we consider the unit disc bundle in $T^*\mathbb{H}$ as fiber:

$$X := D^*\mathbb{H} := \left\{ (z,w) \in \mathbb{H} \times \mathbb{C} \,\middle|\, |w| < \frac{1}{\mathrm{Im}(z)} \right\}$$

This carries a natural $\mathrm{SL}(2,\mathbb{R})$-action and a holomorphic symplectic form. We show in Theorem 4.5.1 that X carries a (essentially unique) $\mathrm{SL}(2,\mathbb{R})$-invariant hyperkähler structure compatible with the holomorphic symplectic form. The restriction to the disc bundle is necessary for this discussion, because the hyperkähler structure ceases to exist on the total space of $T^*\mathbb{H}$.

As before let (Σ, ρ) be a closed 2-dimensional manifold equipped with an area form ρ and denote by $P \to \Sigma$ its $\mathrm{SL}(2,\mathbb{R})$ frame bundle. Using Lemma 4.3.5, one can identify the space of sections $\mathcal{S}(P,X)$ of the associated bundle $P(X) := P \times_{\mathrm{SL}(2,\mathbb{R})} X$ with

$$\mathcal{Q}_1(\Sigma) := \{(J,\sigma) \,|\, J \in \mathcal{J}(\Sigma),\, \sigma \in Q(J),\, |\sigma|_J < 1\}$$

where $Q(J)$ is the space of quadratic differentials $\sigma \in \Omega^0(\Sigma, S^2(T^*\Sigma \otimes_J \mathbb{C}))$. The hyperkähler structure of X induces a hyperkähler structure on the bundle $\mathcal{Q}_1(\Sigma)$. The general theory developed in Theorem 4.2.4 then yields a hyperkähler moment map for the action of $\mathrm{Ham}(\Sigma, \rho)$ on $\mathcal{Q}_1(\Sigma)$. This is the content of Theorem 4.5.13.

As in the construction of Teichmüller space, one would like to construct a symplectic quotient for the action of $\mathrm{Symp}_0(\Sigma, \rho)$ instead of $\mathrm{Ham}(\Sigma, \rho)$. This requires some additional work, as the $\mathrm{Symp}_0(\Sigma, \rho)$-action does not admit a hyperkähler moment map. A key ingrendient in constructing this moduli step is a slight extension of the arguments in the proof of Theorem 4.5.13, which yields two moment maps for the $\mathrm{Symp}_0(\Sigma, \rho)$-action. After suitable rescaling and applying standard Moser isotopy arguments, this yields the moduli space

$$\mathcal{M} := \left\{ (g, \sigma) \in \mathrm{Met}(\Sigma) \times Q(J_g) \;\middle|\; \begin{array}{c} \bar{\partial}_{J_g}\sigma = 0,\ |\sigma|_g < 1, \\ K_g - \frac{c}{2}|\sigma|_g^2 = \frac{c}{2} \end{array} \right\} \Big/ \mathrm{Diff}_0(\Sigma) \tag{4.56}$$

where $c = 2\pi(2-2\mathrm{genus}(\Sigma))/\mathrm{vol}(\Sigma, \rho)$ and $J_g \in \mathcal{J}(\Sigma)$ is the unique complex structure compatible with g. This moduli space comes equipped with a natural hyperkähler structure and an embedding $\mathcal{T}(\Sigma) \hookrightarrow \mathcal{M}$ of Teichmüller space. We investigate the rich geometry of this moduli space in the following section.

4.5.1 Hyperkähler extension of the hyperbolic plane

We identify the unit disc bundle inside the cotangent bundle of the hyperbolic plane with the set

$$X = \left\{ (z, w) \in \mathbb{H} \times \mathbb{C} \;\middle|\; |w| < \frac{1}{\mathrm{Im}(z)} \right\}.$$

This carries a natural $\mathrm{SL}(2, \mathbb{R})$ action induced by the action on $\mathbb{H}$

$$\mathrm{SL}(2, \mathbb{R}) \times X \to X, \qquad \begin{pmatrix} a & b \\ c & d \end{pmatrix} (z, w) = \left(\frac{az + b}{cz + d},\, (cz + d)^2 w \right) \tag{4.57}$$

and a S^1 action which is given by rotation of the fibre

$$S^1 \times X \to X, \qquad \left(e^{\mathbf{i}t}, (z, w)\right) \mapsto \left(z,\, e^{\mathbf{i}t} w\right) \tag{4.58}$$

As a complex cotangent bundle, X carries a natural holomorphic symplectic structure. This consists of the complex structure $J_1(\hat{z}, \hat{w}) := (\mathbf{i}\hat{z}, \mathbf{i}\hat{w})$ and the complex nondegenerate closed 2-form $dz \wedge dw \in \Omega^{2,0}_{J_1}(X)$. Denote the real and imaginary part of this form by

$$\omega_2 := dx \wedge du - dy \wedge dv, \qquad \omega_3 := dx \wedge dv + dy \wedge du$$

where $z = x + \mathbf{i}y$ and $w = u + \mathbf{i}v$. A hyperkähler metric on X, which is compatible with this holomorphic symplectic structure, is a Riemannian metric g satisfying the following: The relation

$$\omega_i(\cdot, J_i \cdot) = g(\cdot, \cdot) \qquad \text{for } i = 1, 2, 3$$

defines a Kähler form ω_1 and integrable complex structures J_2, J_3 which satisfy the quaternionic relations together with J_1.

Feix [43] and Kaledin [67] proved the following general existence result which applies to our situation: For any real analytic Kähler manifold, there exists a unique S^1-invariant hyperkähler metric defined on some neighbourhood of the zero section in the total space of the cotangent bundle, which is compatible with the canonical holomorphic symplectic structure. The following theorem calculates this metric in the case of the hyperbolic plane and shows that it exists on all of X.

Theorem 4.5.1 (Hyperkähler metric on X.). *Define the Riemannian metric g on X by*

$$g = \frac{d\bar{z}dz}{2Im(z)^2\sqrt{1-r^2}} + \frac{Im(z)^2}{2\sqrt{1-r^2}}d\bar{w}dw + \frac{\boldsymbol{i}Im(z)\bar{w}}{2\sqrt{1-r^2}}d\bar{z}dw - \frac{\boldsymbol{i}Im(z)w}{2\sqrt{1-r^2}}d\bar{w}dz$$

where $r := |w|Im(z)$. Then g is a $SL(2,\mathbb{R}) \times S^1$-invariant hyperkähler metric on X. It is compatible with the holomorphic symplectic structure and restricts to the hyperbolic metric along $\mathbb{H} \times \{0\}$ with curvature -1.

Proof. This is Lemma 16 in [38]. The proof takes up the rest of this subsection will be completed on page 178 below. □

Remark 4.5.2. It is easy to verify $\det g_{\bar{\alpha}\beta} \equiv \frac{1}{4}$ and then to deduce that g is a hyperkähler metric. The purpose of the following discussion to explain the derivation of the formula. This will be important for our discussion in Chapter 5.

Remark 4.5.3. The hyperkähler metric on X is not complete.

Remark 4.5.4 (The second complex structure). Hodge [61] showed that there exists a $\mathrm{SL}(2,\mathbb{R})$-equivariant diffeomorphism $\alpha : X \to \mathbb{H} \times \overline{\mathbb{H}}$ which identifies the second complex structure J_2 on X with $(\mathbf{i}, -\mathbf{i})$ on $\mathbb{H} \times \overline{\mathbb{H}}$. It is given by the formula

$$\alpha(z,w) = \left(\exp_z\left(\mathbf{i}f_z(w)\right), \exp_z\left(-\mathbf{i}f_z(w)\right)\right) \tag{4.59}$$

where $f_z : T_z^*\mathbb{H} \to T_z\mathbb{H}$ is given by

$$f_z(w) := \operatorname{arctanh}\left(-\mathrm{Im}(z)|w|\right)\frac{\mathrm{Im}(z)^2 w}{\mathrm{Im}(z)|w|}.$$

For $z = x + \mathbf{i}y$, $w = u + \mathbf{i}v$ and $\gamma = \sqrt{1-r^2} = \sqrt{1 - y^2(u^2+v^2)}$ it holds

$$\alpha(x+\mathbf{i}y, u+\mathbf{i}v) = \left(x - \frac{y^2v}{1-yu} + \mathbf{i}\frac{y\gamma}{1-yu}, x + \frac{y^2v}{1+yu} + \mathbf{i}\frac{y\gamma}{1+yu}\right). \tag{4.60}$$

By Remark 4.5.3, α is not an isometry.

We now proceed to the proof of Theorem 4.5.1. This consists of the following four steps: First, we describe X in the Poincaré disc model of hyperbolic space and introduce a suitable double cover of that space. We then determine successively the hyperkähler metric on the double cover, in the disc model and finally on X.

Step 1: The disc model and its double cover

In the disc model of the hyperbolic plane, X takes the form

$$X_{\mathbb{D}} := D^*\mathbb{D} := \left\{ (z,w) \in \mathbb{D} \times \mathbb{C} \,\middle|\, |w| < \frac{2}{(1-|z|^2)} \right\}.$$

This comes equipped with a natural complex and holomorphic symplectic structure. The $\mathrm{SU}(1,1)$-action on $\mathbb{D}$ induces the action

$$\mathrm{SU}(1,1) \times X_{\mathbb{D}} \to X_{\mathbb{D}}, \qquad \begin{pmatrix} \alpha & \beta \\ \bar{\beta} & \bar{\alpha} \end{pmatrix} (z,w) := \left(\frac{\alpha z + \beta}{\bar{\beta} z + \bar{\alpha}}, \, (\bar{\beta} z + \bar{\alpha})^2 w \right).$$

Moreover, the Cayley transformation (see Lemma 4.3.1) defines a diffeomorphism

$$X \xrightarrow{\cong} X_{\mathbb{D}}, \qquad (\zeta, \eta) \mapsto \left(\frac{\zeta - \mathbf{i}}{\zeta + \mathbf{i}}, \frac{(\zeta + \mathbf{i})^2 \eta}{2\mathbf{i}} \right). \tag{4.61}$$

This intertwines the various structures on X and $X_{\mathbb{D}}$ and it is equivariant with respect to the isomorphism $\mathrm{SL}(2,\mathbb{R}) \cong \mathrm{SU}(1,1)$ defined in Lemma 4.3.1.

Consider the following open subset of $\mathbb{C}^2$

$$\tilde{X} := \{ (z_1, z_2) \in \mathbb{C}^2 \,|\, 0 < |z_1|^2 - |z_2|^2 < 2 \}.$$

The group $\mathrm{U}(1,1)$ acts on $\tilde{X}$ by linear transformations and the next lemma shows that $\tilde{X}$ can be regarded as twofold cover of the punctured disc bundle $X_{\mathbb{D}} \backslash (\mathbb{D} \times \{0\})$.

Lemma 4.5.5. *The map $\pi : \tilde{X} \to X_{\mathbb{D}} \backslash (\mathbb{D} \times \{0\})$ defined by*

$$\pi(z_1, z_2) := \left(\frac{z_2}{z_1}, z_1^2 \right) \tag{4.62}$$

is a holomorphic double cover. Moreover, for $(z_1,z_2) \in \tilde{X}$ with $(z,w) := \pi(z_1,z_2) \in X_{\mathbb{D}} \backslash (\mathbb{D} \times \{0\})$ it holds

$$\pi(A(z_1,z_2)) = \overline{A}_*(z,w), \qquad \pi(e^{\mathbf{i}t}(z_1,z_2)) = (z, e^{2\mathbf{i}t} w) \tag{4.63}$$

for all $A \in SU(1,1)$ and $e^{\mathbf{i}t} \in S^1$.

Proof. The map π is well-defined, since $0 < |z_1|^2 - |z_2|^2 < 2$ is equivalent to the constraint $0 < |z_1^2| < 2/(1 - |z_2/z_1|^2)$. Moreover, $\pi(z_1,z_2) = (z,w)$ if and only if z_1 is a square root of w and $z_2 = z z_1$. Hence, π is indeed a double cover. For

$$A = \begin{pmatrix} \alpha & \beta \\ \bar{\beta} & \bar{\alpha} \end{pmatrix} \in \mathrm{SU}(1,1)$$

and $(z_1,z_2) \in \tilde{X}$ it holds

$$\pi(A(z_1,z_2)) = \left(\frac{\bar{\beta} + \bar{\alpha}(z_2/z_1)}{\alpha + \beta(z_2/z_1)}, \left(\alpha + \beta \frac{z_2}{z_1} \right)^2 z_1^2 \right) = \bar{A} \left(\frac{z_1}{z_2}, z_1^2 \right).$$

This proves the first equation in (4.63) and the second equation is obvious. $\square$

Step 2: The hyperkähler metric on the double cover

The next lemma constructions a U(1, 1)-invariant hyperkähler metric on $\tilde{X}$.

Lemma 4.5.6. *Define the Riemannian metric $\tilde{g}$ on $\tilde{X}$ by the formula*

$$\tilde{g} := \frac{r^3+4|z_2|^2}{r^2\sqrt{4-r^2}}d\bar{z}_1dz_1 + \frac{-r^3+4|z_1|^2}{r^2\sqrt{4-r^2}}d\bar{z}_2dz_2 - \frac{4z_2\bar{z}_1\,d\bar{z}_2dz_1+4z_1\bar{z}_2\,d\bar{z}_1dz_2}{r^2\sqrt{4-r^2}}.$$

where $r := |z_1|^2 - |z_2|^2$. This is a $U(1,1)$-invariant Kähler metric with $\det(\tilde{g}_{\bar{\alpha}\beta}) \equiv 1$.

Proof. Let $\tilde{g}$ be a Kähler metric on $\tilde{X}$ with Kähler form $\tilde{\omega}$. Its volume form is given by $\tilde{\omega}^2/2 = -\det(\tilde{g}_{\bar{\alpha}\beta})d\bar{z}_1 \wedge dz_1 \wedge d\bar{z}_2 \wedge dz_2$ and hence $\det(\tilde{g}_{\bar{\alpha}\beta}) \equiv 1$ is equivalent to

$$-\frac{\tilde{\omega}^2}{2} \equiv d\bar{z}_1 \wedge dz_1 \wedge d\bar{z}_2 \wedge dz_2. \tag{4.64}$$

Let $f : \tilde{X} \to \mathbb{R}$ be a smooth function of the form $f(z_1, z_2) = F(|z_1|^2 - |z_2|^2)$ for some function F. We make the ansatz

$$\tilde{\omega} = \mathbf{i}\bar{\partial}\partial f. \tag{4.65}$$

This is clearly U(1, 1)-invariant and we calculate

$$\begin{aligned}\bar{\partial}\partial f &= \bar{\partial}\left(F'\bar{z}_1dz_1 - F'\bar{z}_2dz_2\right)\\ &= \left(F''|z_1|^2 + F'\right)d\bar{z}_1dz_1 + \left(F''|z_2|^2 - F'\right)d\bar{z}_2dz_2\\ &\quad - F''z_2\bar{z}_1d\bar{z}_2dz_1 - F''z_1\bar{z}_2d\bar{z}_1dz_2\end{aligned} \tag{4.66}$$

Hence (4.64) is equivalent to

$$1 = F'F''(|z_2|^2 - |z_1|^2) - (F')^2.$$

Thus $G := F'$ satisfies $-1 = G(r)G'(r)r + G^2(r)$ and the solution of this equation is given by

$$G(r) = \sqrt{\frac{C_0}{r^2} - 1}$$

for some constant C_0. It follows from (4.65) and (4.66) that $\tilde{\omega}$ is completely determined by G. For $C_0 = 4$ the solution is defined on all of $\tilde{X}$ and blows up as $r = |z_1|^2 - |z_2|^2$ approaches 2. A short calculation shows that this yields the Kähler form

$$-\mathbf{i}\tilde{\omega} = \frac{-r^3-4|z_2|^2}{r^2\sqrt{4-r^2}}d\bar{z}_1dz_2 + \frac{r^3-4|z_1|^2}{r^2\sqrt{4-r^2}}d\bar{z}_2dz_2 + \frac{4z_2\bar{z}_1\,d\bar{z}_2dz_1+4z_1\bar{z}_2\,d\bar{z}_1dz_2}{r^2\sqrt{4-r^2}}.$$

The formula for the Kähler metric follows from this, since $\tilde{g}_{\bar{\alpha}\beta} = \mathbf{i}\tilde{\omega}_{\bar{\alpha}\beta}$. □

Step 3: The hyperkähler metric in the disc model

Lemma 4.5.7. *The metric $\tilde{g}$ determined in Lemma 4.5.6 descends under the covering map (4.62) to a metric on $X_{\mathbb{D}}\backslash(\mathbb{D} \times \{0\})$ which is given by the formula*

$$g = \frac{-r^3|w|+4|w|^2}{r^2\sqrt{4-r^2}}d\bar{z}dz + \frac{(1-|z|^2)^2}{4\sqrt{4-r^2}}d\bar{w}dw - \frac{(1-|z|^2)}{2\sqrt{4-r^2}}\left(\bar{z}wd\bar{w}dz + \bar{w}zd\bar{z}dw\right).$$

where $r = |w|(1 - |z|^2)$. Moreover, g extends smoothly over the whole space $X_{\mathbb{D}}$ and restricts to the hyperbolic metric along $\mathbb{D} \times \{0\}$.

Proof. Let $(z_1, z_2) \in \tilde{X}$ and $(z, w) = \pi(z_1, z_2) = (z_2/z_1, z_1^2)$ be the corresponding coordinates on $X_{\mathbb{D}}$. Then

$$dw = 2z_1 dz_1, \quad dz = \frac{dz_2}{z_1} - \frac{z_2}{z_1^2} dz_1 \quad \Longleftrightarrow \quad dz_1 = \frac{dw}{2z_1}, \quad dz_2 = z_1 dz + \frac{z_2}{2z_1^2} dw$$

and it follows

$$\begin{aligned}
d\bar{z}_1 dz_1 &= \frac{1}{4|w|} d\bar{w} dw \\
d\bar{z}_2 dz_2 &= |w| d\bar{z} dz + \frac{|z|^2}{4|w|} d\bar{w} dw + \frac{z\bar{w}}{2|w|} d\bar{z} dw + \frac{\bar{z} w}{2|w|} d\bar{w} dz \\
d\bar{z}_1 dz_2 &= \frac{w}{2|w|} d\bar{w} dz + \frac{z}{4|w|} d\bar{w} dw \\
d\bar{z}_2 dz_1 &= \frac{\bar{w}}{2|w|} d\bar{z} dw + \frac{\bar{z}}{4|w|} d\bar{w} dw
\end{aligned}$$

The coefficients of $\tilde{g}$ can be expressed as

$$\begin{aligned}
\tilde{g}_{\bar{1}1} &= \frac{r^3 + 4|w| \cdot |z|^2}{r^2\sqrt{4 - r^2}} \\
\tilde{g}_{\bar{2}2} &= \frac{-r^3 + 4|w|}{r^2\sqrt{4 - r^2}} \\
\tilde{g}_{\bar{2}1} &= \frac{-4z|w|}{r^2\sqrt{4 - r^2}} \\
\tilde{g}_{\bar{1}2} &= \frac{-4\bar{z}|w|}{r^2\sqrt{4 - r^2}}
\end{aligned}$$

with $r = |w|(1 - |z|^2)$. Combining the equations we obtain

$$\begin{aligned}
g_{\bar{w}w} &= \frac{(1 - |z|^2)^2}{4\sqrt{4 - r^2}} \\
g_{\bar{z}z} &= \frac{-r^3|w| + 4|w|^2}{r^2\sqrt{4 - r^2}} = \frac{4 - |w|^2(1 - |z|^2)^3}{(1 - |z|^2)^2\sqrt{4 - r^2}} \\
g_{\bar{w}z} &= -\frac{(1 - |z|^2)\bar{z}w}{2\sqrt{4 - r^2}} \\
g_{\bar{z}w} &= -\frac{(1 - |z|^2)z\bar{w}}{2\sqrt{4 - r^2}}
\end{aligned}$$

This calculation is valid for all $(z, w) \in X_{\mathbb{D}}$ with $w \neq 0$. The metric extends smoothly over $\mathbb{D} \times \{0\}$ taking the values

$$g_{\bar{w}w}(z, 0) = \frac{(1 - |z|^2)^2}{8}, \qquad g_{\bar{z}z}(z, 0) = \frac{2}{(1 - |z|^2)^2}$$

and $g_{\bar{w}z}(z, 0) = g_{\bar{z}w}(z, 0) = 0$. This completes the proof of the lemma. $\square$

The next result is the analogue of Theorem 4.5.1 for the disc model $X_{\mathbb{D}}$.

Theorem 4.5.8. *Define g as in Lemma 4.5.7. This is a $SU(1,1)\times S^1$-invariant hyperkähler metric on $X_{\mathbb{D}}$, where the action of S^1 is given by rotation of the fibres. This metric is compatible with the natural holomorphic symplectic structure and restricts to the hyperbolic metric along $\mathbb{D}\times\{0\}$ with constant curvature -1.*

Proof. Let g be the metric on $X_{\mathbb{D}}$ defined in Lemma 4.5.7. It follows from Lemma 4.5.5 and Lemma 4.5.6 that g is $\mathrm{SU}(1,1)\times S^1$-invariant and by Lemma 4.5.7 it restricts to the hyperbolic metric along $\mathbb{D}\times\{0\}$.

Since $\det(\tilde{g}_{\bar{\alpha}\beta})\equiv 1$, it follows

$$\mathrm{vol}_{\tilde{g}} = -d\bar{z}_1\wedge dz_1\wedge d\bar{z}_2\wedge dz_2.$$

Using $2dz_1\wedge dz_2 = dz\wedge dw$, we then get

$$\mathrm{vol}_g = -\frac{1}{4}d\bar{z}\wedge dz\wedge d\bar{w}\wedge dw.$$

and thus $\det(g_{\bar{\alpha}\beta})\equiv\frac{1}{4}$. Since $\det(g_{\bar{\alpha}\beta})$ is the metric tensor for the induced metric on the anti-canonical line bundle, it follows that the metric on $\Lambda^{2,0}T^*X_{\mathbb{D}}$ is constant. The Levi-Civita connection of g induces the unique connection on this bundle which is compatible with the metric and holomorphic structures. It follows that this is the trivial connection and

$$\nabla(dz\wedge dw) = 0.$$

In particular, $\omega_2 = \mathrm{Re}(dz\wedge dw)$ and $\omega_3 = \mathrm{Im}(dz\wedge dw)$ are parallel and it follows that $(X_{\mathbb{D}}, J_2, \omega_2)$ and $(X_{\mathbb{D}}, J_3, \omega_3)$ are Kähler manifolds.

It remains to verify that the complex structures J_1, J_2, J_3 satisfy the algebraic relations of the quaternions. Since $dz\wedge dw\in\Omega^{2,0}_{J_1}(X_{\mathbb{D}})$, we get

$$\omega_2(J_1\cdot,\cdot) = \omega_2(\cdot, J_1\cdot), \qquad \omega_3(J_1\cdot,\cdot) = \omega_3(\cdot, J_1\cdot), \qquad \omega_2(\cdot,\cdot) = \omega_3(J_1\cdot,\cdot).$$

Hence

$$\omega_2(\cdot, J_2J_1\cdot) = g(\cdot, J_1\cdot) = -g(J_1\cdot,\cdot) = -\omega_2(J_1\cdot, J_2\cdot) = -\omega_2(\cdot, J_1J_2\cdot)$$

implies $J_2J_1 = -J_1J_2$. Moreover,

$$g(J_2\cdot,\cdot) = \omega_2(\cdot,\cdot) = \omega_3(J_1\cdot,\cdot) = g(J_3J_1\cdot,\cdot)$$

yields $J_2 = J_3J_1$ and hence $J_3 = -J_2J_1 = J_1J_2$. □

Step 4: The hyperkähler metric in the upper half-plane model

Proof of Theorem 4.5.1. The Cayley transform (4.61) is a diffeomorphism $X\cong X_{\mathbb{D}}$ which intertwines the various structures on X and $X_{\mathbb{D}}$. Hence it suffices to verify that the metric on X defined in Theorem 4.5.1 is the pullback of the metric on $X_{\mathbb{D}}$ calculated in Theorem 4.5.8. Then Theorem 4.5.1 follows from Theorem 4.5.8.

Let (z,w) be the coordinates on $X_{\mathbb{D}}$ and (ζ,η) be the coordinates on $X_{\mathbb{H}}$. The Cayley transform (4.61) is given by

$$(z,w) = \left(\frac{\zeta-\mathbf{i}}{\zeta+\mathbf{i}}, \frac{(\zeta+\mathbf{i})^2\eta}{2\mathbf{i}}\right).$$

It follows

$$dz = \frac{2\mathbf{i}}{(\zeta+\mathbf{i})^2}d\zeta, \qquad dw = -\mathbf{i}(\zeta+\mathbf{i})\eta d\zeta + \frac{(\zeta+\mathbf{i})^2}{2\mathbf{i}}d\eta$$

and

$$d\bar{z}dz = \frac{4}{|\zeta+\mathbf{i}|^4}d\bar{\zeta}d\zeta$$

$$d\bar{w}dw = |\eta|^2\cdot|\zeta+\mathbf{i}|^2 d\bar{\zeta}d\zeta + \frac{|\zeta+\mathbf{i}|^2}{4}d\bar{\eta}d\eta + \frac{(\zeta+\mathbf{i})|\zeta+\mathbf{i}|^2\bar{\eta}}{2}d\bar{\zeta}d\eta + \frac{\overline{(\zeta+\mathbf{i})}|\zeta+\mathbf{i}|^2\eta}{2}d\bar{\eta}d\zeta$$

$$d\bar{z}dw = \frac{-2(\zeta+\mathbf{i})^3\eta}{|\zeta+\mathbf{i}|^4}d\bar{\zeta}d\zeta - \frac{(\zeta+\mathbf{i})^4}{|\zeta+\mathbf{i}|^4}d\bar{\zeta}d\eta$$

$$d\bar{w}dz = \frac{-2\overline{(\zeta+\mathbf{i})}^3\bar{\eta}}{|\zeta+\mathbf{i}|^4}d\bar{\zeta}d\zeta - \frac{\overline{(\zeta+\mathbf{i})}^4}{|\zeta+\mathbf{i}|^4}d\bar{\eta}d\zeta$$

With $r = |w|(1-|z|^2) = 2|\eta|\mathrm{Im}(\zeta)$, the components of the metric transform as follows:

$$g_{\bar{z}z} = \frac{-r^3|w|+4|w|^2}{r^2\sqrt{4-r^2}} = \frac{|\zeta+\mathbf{i}|^4}{4\mathrm{Im}(\zeta)^2\sqrt{4-r^2}} - \frac{r|\zeta+\mathbf{i}|^2|\eta|}{2\sqrt{4-r^2}}$$

$$g_{\bar{w}w} = \frac{(1-|z|^2)^2}{4\sqrt{4-r^2}} = \frac{4\mathrm{Im}(\zeta)^2}{|\zeta+\mathbf{i}|^4\sqrt{4-r^2}}$$

$$g_{\bar{z}w} = -\frac{(1-|z|^2)z\bar{w}}{2\sqrt{4-r^2}} = -\frac{\mathbf{i}\mathrm{Im}(\zeta)(\zeta-\mathbf{i})\overline{(\zeta+\mathbf{i})}^3\bar{\eta}}{|\zeta+\mathbf{i}|^4\sqrt{4-r^2}}$$

$$g_{\bar{w}z} = -\frac{(1-|z|^2)\bar{z}w}{2\sqrt{4-r^2}} = \frac{\mathbf{i}\mathrm{Im}(\zeta)\overline{(\zeta-\mathbf{i})}(\zeta+\mathbf{i})^3\eta}{|\zeta+\mathbf{i}|^4\sqrt{4-r^2}}$$

A lengthy computation combining these terms yield the desired expression:

$$g_{\bar{\zeta}\zeta} = \frac{1}{\mathrm{Im}(\zeta)^2\sqrt{4-r^2}}$$

$$g_{\bar{\eta}\eta} = \frac{\mathrm{Im}(\zeta)^2}{\sqrt{4-r^2}}$$

$$g_{\bar{\zeta}\eta} = \frac{\mathbf{i}\mathrm{Im}(\zeta)\bar{\eta}}{\sqrt{4-r^2}}$$

$$g_{\bar{\eta}\zeta} = \frac{-\mathbf{i}\mathrm{Im}(\zeta)\eta}{\sqrt{4-r^2}}$$

Note that in Theorem 4.5.1 we defined $r = |\eta|\mathrm{Im}(\zeta)$ in contrast to our convention of $r = 2|\eta|\mathrm{Im}(\zeta)$ in the calculation above. Hence we need to replace r by $2r$ to match both formulae. □

4.5.2 Moment maps on the fibre

The S^1-action on X defined by (4.58) is Hamiltonian for ω_1 and rotates the symplectic forms ω_2 and ω_3. The moment map for this action with respect to ω_1 yields a Kähler potential for the hyperkähler metric with respect to the second and third complex structure. This is a general feature for hyperkähler manifolds equipped with such an S^1-action, which has been observed in [60]. We recall the argument in Lemma 4.5.9 below. The $\mathrm{SL}(2,\mathbb{R})$-action on X defined by (4.57) preserves all three symplectic forms and admits a hyperkähler moment map. We calculate the first moment map in Proposition 4.5.10. The second and third moment map follow from a more general calculation in Proposition 4.5.11.

Lemma 4.5.9 (Rotation of the fibres). *Equip X with the hyperkähler structure obtained in Theorem 4.5.1 and consider the S^1-action on X defined by (4.58)*

1. *This action is Hamiltonian with respect to ω_1 and generated by the Hamiltonian function*

$$H : X \to \mathbb{R}, \qquad H(z,w) := \sqrt{1 - Im(z)^2|w|^2}. \tag{4.67}$$

2. *H is a Kähler potential for the hyperkähler structure with respect to the second and third complex structure, i.e.*

$$2\boldsymbol{i}\bar{\partial}_{J_2}\partial_{J_2}H = \omega_2, \qquad 2\boldsymbol{i}\bar{\partial}_{J_3}\partial_{J_3}H = \omega_3 \tag{4.68}$$

Proof. For $(z,w) \in X$ write $z = x + \mathbf{i}y$ and $w = u + \mathbf{i}v$. Then

$$\begin{aligned}
\omega_1((0,\mathbf{i}w),(\hat{z},\hat{w})) &= -g((0,w),(\hat{z},\hat{w})) \\
&= -2\mathrm{Re}\,(\bar{w}g_{\bar{w}z}\hat{z} + \bar{w}g_{\bar{w}w}\hat{w}) \\
&= -\frac{|w|^2y\hat{y} + (u\hat{u} + v\hat{v})y^2}{\sqrt{1-|w|^2y^2}} \\
&= dH(z,w)[\hat{z},\hat{w}].
\end{aligned}$$

This shows that the Hamiltonian vector field generated by H is given by $v_H(z,w) = (0,\mathbf{i}w)$ and hence H generates the S^1-action on X.

Denote by $\phi_t(z,w) := (z, e^{\mathbf{i}t}w)$ the rotation by $e^{\mathbf{i}t}$. Then

$$\mathcal{L}_{v_H}(\omega_2 + \mathbf{i}\omega_3) = \left.\frac{d}{dt}\right|_{t=0} \phi_t^*(\omega_2 + \mathbf{i}\omega_3) = \left.\frac{d}{dt}\right|_{t=0} e^{\mathbf{i}t}dz \wedge dw = \mathbf{i}\omega_2 - \omega_3$$

and therefore $\mathcal{L}_{v_H}\omega_2 = -\omega_3$ and $\mathcal{L}_{v_H}\omega_3 = \omega_2$. The identity

$$dH(J_2u) = \omega_1(v_H, J_2u) = g(J_1v_H, J_2u) = g(J_3v_H, u) = \omega_3(v_H, u)$$

then yields

$$2\mathbf{i}\bar{\partial}_{J_2}\partial_{J_2}H = d(dH \circ J_2) = d\iota(v_H)\omega_3 = \mathcal{L}_{v_H}\omega_3 = \omega_2.$$

This proves the first equation in (4.68). The second follows by a similar calculation and this proves the lemma. □

Proposition 4.5.10 (**Moment map for** ω_1.). *Let $\omega_1 \in \Omega^2(X)$ be the symplectic form obtained in Theorem 4.5.1 and let $j : \mathbb{H} \to \mathcal{J}(\mathbb{R}^2)$ be the isomorphism (4.39). Then $\mu_1 : X \to \mathfrak{sl}^*(2,\mathbb{R})$ defined by*

$$\langle \mu_1(z,w), \xi\rangle := -\sqrt{1 - Im(z)^2|w|^2)}\, tr(j(z)\xi), \qquad \text{for } \xi \in \mathfrak{sl}(2,\mathbb{R})$$

is an equivariant moment map for the $SL(2,\mathbb{R})$ action on X with respect to ω_1.

Proof. The proof consists of three steps.

Step 1: *For $0 < r < 1$ define $X_r := \{(z,w) \in X \mid |w| Im(z) = r\}$. Then*

$$\omega_1((\hat{z}_1, \hat{w}_1), (\hat{z}_2, \hat{w}_2)) = \sqrt{1 - r^2}\, \omega_{\mathbb{H}}(\hat{z}_1, \hat{z}_2). \tag{4.69}$$

for all $(z,w) \in X_r$ and $(\hat{z}_1, \hat{w}_1), (\hat{z}_2, \hat{w}_2) \in T_{(z,w)}X_r$.

It follows from Lemma 4.5.9 that $X_r = H^{-1}(\sqrt{1 - r^2})$. Hence X_r/S^1 is a Marsden–Weinstein quotient and ω_1 induces a well-defined $\mathrm{SL}(2,\mathbb{R})$-invariant symplectic form on X_r/S^1. Since $\mathrm{SL}(2,\mathbb{R})$ acts transitively on X_r, such a form is unique up to scaling and there exists $f(r) \in \mathbb{R}$ such that

$$\omega_1((\hat{z}_1, \hat{w}_1), (\hat{z}_2, \hat{w}_2)) = f(r)\, \omega_{\mathbb{H}}(\hat{z}_1, \hat{z}_2).$$

for all $(z,w) \in X_r$ and $(\hat{z}_1, \hat{w}_1), (\hat{z}_2, \hat{w}_2) \in T_{(z,w)}X_r$. We calculate $f(r)$ by evaluating ω_1 at $(\mathbf{i}, r) \in X_r$ on the tangent vectors $(1,0), (\mathbf{i}, -r) \in T_{(\mathbf{i},r)}X_r$.

$$f(r) = (\omega_1)_{(\mathbf{i},r)}((1,0), (\mathbf{i}, -r)) = 2\mathrm{Im}\,(g_{\bar{z}z}\mathbf{i} - g_{\bar{z}w} r) = \sqrt{1 - r^2}.$$

This establishes (4.69).

Step 2: *Let $(z,w) \in X$ with $w \neq 0$ and define by*

$$V_r(z,w) := \left(0, \frac{w}{Im(z)|w|}\right), \qquad V_\phi(z,w) := (0, \boldsymbol{i}w)$$

the radial and the angular vector fields along the fibre. Decompose $\xi \in \mathfrak{sl}(2,\mathbb{R})$ as $\xi = \xi_0 + \xi_1$ such that ξ_0 commutes with $j(z)$ and ξ_1 anti-commutes with $j(z)$. Then

$$L_{(z,w)}\xi_0 = -tr(j(z)\xi)V_\phi$$

$$\omega_1(L_{(z,w)}\xi_1, V_r) = 0 = \omega_1(L_{(z,w)}\xi_1, V_\phi).$$

Both identities are preserved by the $\mathrm{SL}(2,\mathbb{R})$-action and we may assume without loss of generality $z = \mathbf{i}$ and $w > 0$. Then $j(z) = J_0$ and ξ decomposes as

$$\xi = \begin{pmatrix} a & b \\ c & -a \end{pmatrix} \quad \Longrightarrow \quad \xi_0 = \begin{pmatrix} 0 & \frac{b-c}{2} \\ \frac{c-b}{2} & 0 \end{pmatrix}, \quad \xi_1 = \begin{pmatrix} a & \frac{b+c}{2} \\ \frac{b+c}{2} & -a \end{pmatrix}.$$

The infinitesimal action of ξ is given by

$$L_{(\mathbf{i},w)}\xi = (\mathbf{i}a + b - \mathbf{i}(\mathbf{i}c - a), 2(c\mathbf{i} - a)w) = (2\mathbf{i}a + (b+c), 2(c\mathbf{i} - a)w)\,.$$

This yields for ξ_0

$$L_{(\mathbf{i},w)}\xi_0 = (0, \mathbf{i}(c-b)w) = (c-b)V_\phi = -\mathrm{tr}(J_0\xi)V_\phi$$

which proves the first identity. For ξ_1 it follows

$$L_{(\mathbf{i},w)}\xi_1 = (2\mathbf{i}a + (b+c), (-2a + \mathbf{i}(b+c))w)\,.$$

and hence

$$\begin{aligned}\omega_1(L_{(\mathbf{i},w)}\xi_0, (0,\hat{w})) &= 2\mathrm{Im}\left(((b+c) - 2\mathbf{i}a)g_{\bar{z}w}\hat{w} - (2a + \mathbf{i}(b+c))\bar{w}g_{\bar{w}w}\hat{w}\right) \\ &= \frac{1}{\sqrt{1-w^2}}\mathrm{Im}\left(((b+c) - 2\mathbf{i}a)\mathbf{i}\bar{w}\hat{w} - (2a + \mathbf{i}(b+c))\bar{w}\hat{w}\right) \\ &= 0.\end{aligned}$$

for every $\hat{w} \in \mathbb{C}$. This proves the second identity.

Step 3: μ_1 *satisfies the moment map equation*

$$d\langle \mu_1(z,w), \xi\rangle[\hat{z},\hat{w}] = \omega_1(L_{(z,w)}\xi, (\hat{z},\hat{w})) \tag{4.70}$$

for every $(z,w) \in X$ *and* $(\hat{z},\hat{w}) \in T_{(z,w)}X$.

Suppose first $w = 0$. For tangent vectors $(\hat{z}, 0)$ along the base, the claim follows from Lemma 4.3.7. For tangent vectors $(0,\hat{w})$ along the fibre, the derivative of $\langle \mu_1, \xi\rangle$ in the direction of $(0,\hat{w})$ vanishes. Since $\omega_1(L_{(z,0)}\xi, (0,\hat{w})) = 0$, it follows that (4.70) is satisfied in the case $w = 0$.

Suppose next $r := |w|\mathrm{Im}(z) > 0$ and consider the case where $(\hat{z},\hat{w})$ is tangential to X_r. Since $L_{(z,w)}\xi$ is also tangential, it follows from (4.69) and Lemma 4.3.7

$$\begin{aligned}d\langle \mu_1(z,w), \xi\rangle[\hat{z},\hat{w}] &= -\sqrt{1-r^2}\, d\mathrm{tr}(j(z)\xi)[\hat{z}] \\ &= \sqrt{1-r^2}\,\omega_{\mathbb{H}}(L_z\xi, \hat{z}) \\ &= \omega_1(L_{(z,w)}\xi, (\hat{z},\hat{w}))\end{aligned}$$

Finally consider the case $r := |w|\mathrm{Im}(z) > 0$ and $(\hat{z},\hat{w}) = V_r(z,w)$. The vector fields V_r and V_ϕ defined in Step 3 satisfy

$$\omega_1(V_r(z,w), V_\phi(z,w)) = 2\mathrm{Im}\left(\frac{\bar{w}}{\mathrm{Im}(z)|w|}g_{\bar{w}w}\mathbf{i}w\right) = \frac{r}{\sqrt{1-r^2}}$$

Hence, it follows from Step 2 that

$$d\langle \mu_1(z,w), \xi\rangle[V_r] = \frac{r}{\sqrt{1-r^2}}\mathrm{tr}(j(z)\xi) = \omega_1\left(-\mathrm{tr}(j(z)\xi)V_\phi, V_r\right) = \omega_1(L_{(z,w)}\xi, V_r).$$

This completes the proof of the moment map equation (4.70). $\square$

Proposition 4.5.11 (Complex cotangent bundles). *Let G be a Lie group acting on a smooth complex manifold Y. Denote by $\pi : T^*Y \to Y$ the canonical projection and recall that the tautological 1-form $\lambda \in \Omega^1(T^*Y, \mathbb{C})$ is defined by*

$$\lambda_{(y,\alpha)} := \alpha \circ d\pi(y, \alpha) : T_{(y,\alpha)}(T^*Y) \to \mathbb{C}.$$

The holomorphic symplectic form on T^*Y is defined by

$$\omega_2 + \boldsymbol{i}\omega_3 = -d\lambda \in \Omega^2(T^*Y, \mathbb{C}).$$

The G-action on Y induces a natural action on T^*Y. This action is Hamiltonian with respect to ω_2 and ω_3 and admits the moment maps

$$\langle \mu_2(y, \alpha), \xi\rangle + \boldsymbol{i}\langle \mu_3(y, \alpha), \xi\rangle := \lambda_{(y,\alpha)}(L_{(y,\alpha)}\xi) = \alpha(L_y\xi), \qquad \text{for } \xi \in \mathfrak{g}.$$

Here $L_y : \mathfrak{g} \to T_yY$ and $L_{(y,\alpha)} : \mathfrak{g} \to T_{(y,\alpha)}T^*Y$ denote the infinitesimal action on Y and T^*Y respectively.

Proof. Let $g \in G$, $(y, \alpha) \in T^*Y$ and denote by $m_g : T_yY \to T_{gy}Y$ the derivative of the action by g. Then $g(y, \alpha) = (gy, \alpha \circ m_g^{-1})$ and $g^*\lambda = \lambda$. Hence the Lie derivative of λ in the direction $v_\xi(y, \alpha) := L_{(y,\alpha)}\xi$ vanishes. Then, by Cartan's formula, we get

$$0 = \mathcal{L}_{v_\xi}\lambda = d\iota(v_\xi)\lambda + \iota(v_\xi)d\lambda.$$

This yields $\omega_2(v_\xi, \cdot) + \mathbf{i}\omega_3(v_\xi, \cdot) = d\lambda(v_\xi)$ and proves the moment map equation. $\square$

4.5.3 The hyperkähler moment map on the space of sections

The main result of this subsection is Theorem 4.5.13, which calculates a hyperkähler moment map for the action of $\mathrm{Ham}(\Sigma, \rho)$ on $\mathcal{Q}_1(\Sigma)$ and two moment maps for the action of $\mathrm{Symp}_0(M, \rho)$. We begin our discussion with a careful look at the isomorphism $\mathcal{S}(P, X) \cong \mathcal{Q}_1(\Sigma)$.

Geometric description of the sections

Denote by $P \to (\Sigma, \rho)$ the $\mathrm{SL}(2, \mathbb{R})$ frame bundle, let $X \subset T^*\mathbb{H}$ be the unit disc-bundle inside the cotangent bundle of the hyperbolic plane, and consider the associated fibration $P(X) := P \times_{\mathrm{SL}(2,\mathbb{R})} X$. Denote by

$$(j, q) : X \to \mathcal{J}(\mathbb{R}^2) \times \mathrm{Hom}(\mathbb{R}^2 \otimes \mathbb{R}^2, \mathbb{C})$$

the map (4.46) defined in Lemma 4.3.5. We remind the reader that the fibre maps $q(\zeta, \cdot) : T^*_\zeta\mathbb{H} \to Q(j(\zeta))$ are complex anti-linear isometries for the canonical structures. This yields an embedding

$$\begin{aligned} P(X) &\hookrightarrow \mathrm{End}(T\Sigma) \times S^2(T^*\Sigma \otimes \mathbb{C}), \\ [(z, \theta), (\zeta, \eta)] &\mapsto \big(\theta j(\zeta)\theta^{-1}, \theta^* q(\zeta, \eta)\big) \end{aligned} \tag{4.71}$$

where $z \in \Sigma$, $\theta : \mathbb{R}^2 \to T_z\Sigma$ is a volume preserving frame and $(\zeta, \eta) \in X$. On the space of section this yields the identification

$$\mathcal{S}(P, X) \cong \mathcal{Q}_1(\Sigma) = \{(J, \sigma) \,|\, J \in \mathcal{J}(\Sigma),\ \sigma \in Q(J),\ |\sigma|_J < 1\}$$

where $Q(J) = \Omega^1(\Sigma, S^2(T^*\Sigma \otimes_J \mathbb{C}))$ denotes the space of J-quadratic differentials.

Lemma 4.5.12.

1. *Any torsion free* $SL(2,\mathbb{R})$ *connection on* $T\Sigma$ *induces connections on* $P(X)$ *and* $End(T\Sigma) \times S^2(T^*\Sigma \otimes \mathbb{C})$ *which are compatible with respect to (4.71).*

2. *The inclusion (4.71) is* $Symp(\Sigma,\rho)$*-equivariant.*

3. *The hyperkähler structure on* X *(see Theorem 4.5.1) induces a hyperkähler structure on* $P(X)$*.*

   (a) *The natural complex structure on* X *corresponds to*

   $$(\hat{J},\hat{\sigma}) \mapsto (-J\hat{J}, -\boldsymbol{i}\hat{\sigma})$$

   for $(J,\sigma) \in \mathcal{Q}_1(\Sigma)$ *and* $(\hat{J},\hat{\sigma}) \in T_{(J,q)}\mathcal{Q}_1(\Sigma)$*. The corresponding holomorphic symplectic form satisfies (pointwise) the equation*

   $$(\omega_2 + \boldsymbol{i}\omega_3)_{(J,\sigma)}\Big((\hat{J}_1,\hat{\sigma}_1),(\hat{J}_2,\hat{\sigma}_2)\Big) = \frac{\overline{\hat{\sigma}_2(\hat{J}_1 v, v)} - \overline{\hat{\sigma}_1(\hat{J}_2 v, v)}}{|v|_J^2}$$

   for $(J,\sigma) \in \mathcal{Q}_1(\Sigma)$ *and* $(\hat{J}_i,\hat{\sigma}_i) \in T_{(J,q)}\mathcal{Q}_1(\Sigma)$*.*

   (b) *The first symplectic form satisfies*

   $$(\omega_1)_{(J,q)}\left((0,\hat{\sigma}_1),(0,\hat{\sigma}_2)\right) = \frac{-\omega_Q(\hat{\sigma}_1,\hat{\sigma}_2)}{\sqrt{1-|\sigma|^2}}$$

   for $(J,\sigma) \in \mathcal{Q}_1(\Sigma)$ *and* $\hat{\sigma}_i \in Q(J)$*. Here we denote by* ω_Q *the pointwise symplectic structure on* $S^2(T^*\Sigma \otimes_J \mathbb{C})$ *determined by* J *and* ρ*.*

Proof. The first two claims are a matter of unravelling the definitions and left to the reader. The formula for the holomorphic symplectic structure follow from Lemma 4.3.5 and Lemma 4.3.6. The final property of the hyperkähler metric follows from Theorem 4.5.1. □

Calculation of the hyperkähler moment map

The symplectic forms on $P(X)$ integrates to symplectic forms on $\mathcal{Q}_1(\Sigma)$ defined by

$$\underline{\omega}_i((\hat{J}_1,\hat{\sigma}_1),(\hat{J}_2,\hat{\sigma}_2)) := \int_\Sigma \omega_i((\hat{J}_1,\hat{\sigma}_1),(\hat{J}_2,\hat{\sigma}_2))\rho$$

The next theorem calculates moment maps for these symplectic forms.

Theorem 4.5.13. *The action of* $Ham(\Sigma,\rho)$ *on* $\mathcal{Q}_1(\Sigma)$ *is Hamiltonian for all three symplectic structures* $\underline{\omega}_i$*. Moreover, the action of* $Symp_0(\Sigma,\rho)$ *on* $\mathcal{Q}_1(\Sigma)$ *is Hamiltonian for* $\underline{\omega}_2$ *and* $\underline{\omega}_3$*:*

1. *An equivariant moment map for the* $Ham(\Sigma,\rho)$*-action on* $\mathcal{Q}_1(\Sigma)$ *for* $\underline{\omega}_1$ *is*

   $$\underline{\mu}_1(J,\sigma) = \frac{|\bar{\partial}_J\sigma|_J^2 - |\partial_J\sigma|_J^2}{\sqrt{1-|\sigma|_J^2}}\rho + 2\sqrt{1-|\sigma|_J^2}K_J\rho + 2\boldsymbol{i}\bar{\partial}_J\partial_J\sqrt{1-|\sigma|_J^2} - 2c\rho \tag{4.72}$$

   where $c := 2\pi(2-2genus(\Sigma))/vol(\Sigma,\rho)$*,* K_J *denotes the Gaussian curvature of* $\rho(\cdot,J\cdot)$ *and all norms* $|\cdot|_J$ *are calculated with respect to this metric.*

2. *Define the contraction* $r : \Omega_J^{0,1}(\Sigma, S^2(T^*\Sigma \otimes_J \mathbb{C})) \to \Omega_J^{1,0}(\Sigma)$ *by* $r(\gamma) := \frac{\gamma(v)(v,\cdot)}{|v|_J^2}$ *which is independent of* $0 \neq v \in$ *Vect*(Σ). *An equivariant moment map for the Ham*(Σ, ρ)*-action on* $\mathcal{Q}_1(\Sigma)$ *for* $\underline{\omega}_2$ *and* $\underline{\omega}_3$ *is given by*

$$\underline{\mu}_2(J,\sigma) + \boldsymbol{i}\underline{\mu}_3(J,\sigma) = -2\boldsymbol{i}\overline{\bar{\partial}_J r(\bar{\partial}_J\sigma)} \tag{4.73}$$

3. *An equivariant moment map for the Symp*$_0(\Sigma, \rho)$*-action on* $\mathcal{Q}_1(\Sigma)$ *with respect to* $\underline{\omega}_2$ *and* $\underline{\omega}_3$ *is given by*

$$\left\langle \underline{\tilde{\mu}}_2(J,\sigma) + \boldsymbol{i}\underline{\tilde{\mu}}_3(J,\sigma), v \right\rangle = -2\boldsymbol{i}\int_\Sigma \overline{\iota(v) r(\bar{\partial}_J\sigma)}\rho. \tag{4.74}$$

for any symplectic vector field $v \in$ *Vect*(Σ) *satisfying* $d\iota(v)\rho = 0$.

Proof. Denote by ∇ the Levi-Civita connection of $\rho(\cdot, J\cdot)$. We deduce (4.72) in Step 1 from Theorem 4.2.4. For the proof of (4.74) we need to extend the arguments used in the derivation of Theorem 4.2.4. This is done in Step 2 and the derivation of (4.73) and (4.74) is completed in Step 3 and Step 4.

Step 1.1: $\underline{\omega}_1(\nabla(J,\sigma), \nabla(J,\sigma)) = \frac{1}{\sqrt{1-|\sigma|^2}}(|\bar{\partial}\sigma|^2 - |\partial\sigma|^2)\rho.$

For the Levi-Civita connection we have $\nabla J = 0$ and then Lemma 4.5.12 yields:

$$\underline{\omega}_1(\nabla(J,\sigma)\wedge\nabla(J,\sigma)) = \underline{\omega}_1((0,\nabla\sigma)\wedge(0,\nabla\sigma)) = -\frac{\omega_Q(\nabla\sigma,\nabla\sigma)}{\sqrt{1-|\sigma|^2}}. \tag{4.75}$$

For $u \in \mathrm{Vect}(\Sigma)$, we have

$$|\partial_u\sigma|^2 = \frac{1}{4}\left|\nabla_u\sigma - \mathbf{i}\nabla_{Ju}\sigma\right|^2 = \frac{1}{4}\left(|\nabla_u\sigma|^2 + |\nabla_{Ju}\sigma|^2\right) + \frac{1}{2}\omega_Q(\nabla_u\sigma, \nabla_{Ju}\sigma)$$

and

$$|\bar{\partial}_u\sigma|_J^2 = \frac{1}{4}\left|\nabla_u\sigma + \mathbf{i}\nabla_{Ju}\sigma\right|^2 = \frac{1}{4}\left(|\nabla_u\sigma|^2 + |\nabla_{Ju}\sigma|^2\right) - \frac{1}{2}\omega_Q(\nabla_u\sigma, \nabla_{Ju}\sigma)$$

which yields $|\partial_u\sigma|^2 - |\bar{\partial}_u\sigma|^2 = \omega_Q(\nabla_u\sigma, \nabla_{Ju}\sigma)$ and therefore

$$\left(|\partial\sigma|^2 - |\bar{\partial}\sigma|^2\right)\rho = \omega_Q(\nabla\sigma,\nabla\sigma) \tag{4.76}$$

Step 1.1 follows from (4.75) and (4.76).

Step 1.2: $\langle \mu_{(J,\sigma)}, R^\nabla\rangle = -2K_J\sqrt{1-|\sigma|^2}\rho$ *where* R^∇ *and* K_J *denotes the Riemannian and Gaussian curvature the Levi–Civita connection for* $\rho(\cdot, J\cdot)$.

The Riemann curvature tensor R^∇ and Gaussian curvature K_J are related by the formula $R^\nabla = -K_J J \otimes \rho$. By Proposition 4.5.10 if follows

$$\langle \mu_{(J,\sigma)}, R^\nabla\rangle = \sqrt{1-|\sigma|^2}K_J\mathrm{tr}(J^2)\rho = -2K_J\sqrt{1-|\sigma|^2}\rho$$

and this proves Step 1.2.

Step 1.3: $dc(\nabla\mu_{(J,\sigma)}) = -2\boldsymbol{i}\bar\partial\partial\sqrt{1-|\sigma|^2}$.

Using Proposition 4.5.10 and $\nabla J = 0$, we obtain

$$\begin{aligned}\nabla_u\mu_{(J,\sigma)}(\Psi) &= \mathcal{L}_u\left(-\sqrt{1-|\sigma|^2}\mathrm{tr}(J\Psi)\right) + \sqrt{1-|\sigma|^2})\mathrm{tr}(J\nabla_u\Psi) \\ &= -\mathcal{L}_u\left(\sqrt{1-|\sigma|^2}\right)\mathrm{tr}(J\Psi)\end{aligned}$$

for all $\Psi \in \Omega^0(\Sigma, \mathrm{End}(T\Sigma))$ and $u \in \mathrm{Vect}(\Sigma)$. Let e_1, $e_2 = Je_1$ be a local orthonormal frame for $T\Sigma$ and write $v \in \mathrm{Vect}(\Sigma)$ as $v = v_1e_1 + v_2e_2$. Then

$$\begin{aligned}c(\nabla\mu_{(J,\sigma)})(v) &= \nabla_{e_1}\mu_{(J,\sigma)}(e_1^*\otimes v) + \nabla_{e_2}\mu_{(J,\sigma)}(e_2^*\otimes v) \\ &= -\mathcal{L}_{e_1}\sqrt{1-|\sigma|^2}\mathrm{tr}(Je_1^*\otimes v) - \mathcal{L}_{e_2}\sqrt{1-|\sigma|^2}\mathrm{tr}(Je_2^*\otimes v) \\ &= \mathcal{L}_{e_1}\sqrt{1-|\sigma|^2}v_2 - \mathcal{L}_{e_2}\sqrt{1-|\sigma|^2}v_1 \\ &= -\mathcal{L}_{Jv}\left(\sqrt{1-|\sigma|^2}\right).\end{aligned}$$

Step 1.3. follows from this and the relation $d(df\circ J) = 2\mathbf{i}\bar\partial\partial f$ which holds for every smooth function $f : \Sigma \to \mathbb{C}$.

Step 1.4: *(4.72) defines an equivariant moment map for the action of* $Ham(\Sigma,\rho)$ *with respect to* $\underline{\omega}_1$.

We have identified in Steps 1.1-3 all components of the moment map in Theorem 4.2.4. It follows from the general theory (see Lemma 4.2.2) that the cohomology class of

$$\frac{|\bar\partial\sigma|^2 - |\partial\sigma|^2}{\sqrt{1-|\sigma|^2}}\rho + 2\sqrt{1-|\sigma|^2}K_J\rho + 2\mathbf{i}\bar\partial\partial\sqrt{1-|\sigma|^2}$$

does not depend on $(J,\sigma) \in \mathcal{Q}_1(\Sigma)$. For $\sigma = 0$, it follows from the Gauss–Bonnet theorem that the cohomology class is represented by $2c\rho$. Therefore $\underline{\mu}_1$ takes values in the space of exact 2-forms and the moment map equation follows from Theorem 4.2.4.

Step 2.1: *Let* $\lambda \in \Omega^1(X,\mathbb{C})$ *be the tautological* 1*-form on* X *which is defined by* $\lambda_{(\zeta,\eta)}(\hat\zeta,\hat\eta) = \eta\hat\zeta$ *for* $(\zeta,\eta) \in X$ *and* $(\hat\zeta,\hat\eta) \in T_{(\zeta,\eta)}X$. *Define* $\Lambda \in \Omega^1(P\times X,\mathbb{C})$ *by*

$$\Lambda_{(p,x)}(\hat p,\hat x) = \lambda_x(\hat x + L_xA_p(\hat p)). \tag{4.77}$$

This descends to a 1*-form on* $P(X) = P\times_{SL(2,\mathbb{R})} X$ *and satisfies*

$$\partial_{\hat s}\int_\Sigma s^*\Lambda\wedge\iota(v)\rho = \int_\Sigma (d\Lambda)_s(\hat s,\nabla_v s)\rho. \tag{4.78}$$

for every symplectic vector field $v \in Vect(\Sigma)$ *with* $d\iota(v)\rho = 0$.

Since λ is $\mathrm{SL}(2,\mathbb{R})$-equivariant, one readily verifies that $\Lambda \in \Omega^1(P\times X,\mathbb{C})$ is also equivariant and descends to a well-define 1-form $\Lambda \in \Omega^1(P(X),\mathbb{C})$. It now follows from Cartan's formula

$$\begin{aligned}\partial_{\hat{s}} \int_\Sigma s^*\Lambda \wedge \iota(v)\rho &= \int_\Sigma ds^*\iota(\hat{s})\Lambda \wedge \iota(v)\rho + \int_\Sigma s^*(\iota(\hat{s})d\Lambda) \wedge \iota(v)\rho \\ &= \int_\Sigma \iota(v)s^*(\iota(\hat{s})d\Lambda)\rho \\ &= \int_\Sigma (d\Lambda)_s(\hat{s}, \nabla_v s)\rho\end{aligned}$$

where the second equation uses integration by parts and $d\iota(v)\rho = 0$.

Step 2.2: *Let $v \in Vect(\Sigma)$ with $d\iota(v)\rho = 0$. By Lemma 4.2.3, the Lie derivative of s along v is given by $\mathcal{L}_v s = \nabla_v s - L_s\nabla v$. This satisfies*

$$\int_\Sigma (d\Lambda_s)(\hat{s}, \mathcal{L}_v s)\rho = \partial_{\hat{s}} \int_\Sigma s^*\Lambda \wedge \iota(v)\rho - \partial_{\hat{s}} \int_\Sigma \Lambda_s(L_s\nabla v)\rho$$

where $L_x : \mathfrak{sl}(2,\mathbb{R}) \to T_xX$ denotes the infinitesimal action on the fibre.

It follows from the moment map equation in Proposition 4.5.11 that

$$\int_\Sigma (d\Lambda)_s\,(L_s\nabla v, \hat{s})\,\rho = -\partial_{\hat{s}} \int_\Sigma \Lambda_s(L_s\nabla v)\rho.$$

The formula follows thus from Step 2.1 and $\mathcal{L}_v s = \nabla_v s - L_s\nabla v$.

Step 3: *(4.74) defines an equivariant moment maps for the action of $Symp_0(\Sigma,\rho)$ with respect to $\underline{\omega}_2$ and $\underline{\omega}_3$.*

For $s = (J,\sigma)$ we have

$$s^*\Lambda = (J,\sigma)^*\Lambda = \langle \sigma, \nabla J\rangle_{Q(J)\times T\mathcal{J}} = 0$$

$$\Lambda_s(L_s\nabla v) = \langle \sigma, L_J(\nabla v)\rangle_{Q(J)\times T\mathcal{J}} = \langle \sigma, -2J\bar{\partial}_J v\rangle_{Q(J)\times T\mathcal{J}}$$

where the pairing is defined by (4.48). In the second equation, we used that $L_J\xi = [\xi, J] = -2J\xi^{0,1}$ for $\xi \in \mathrm{End}_0(\mathbb{R}^n)$ and $J \in \mathcal{J}(\mathbb{R}^2)$. Along the vertical tangent bundle of $\mathcal{Q}_1(\Sigma)$, it holds $\underline{\omega}_2 + \mathbf{i}\underline{\omega}_3 = -d\Lambda$. It thus follows from Step 2 that

$$\left\langle \underline{\mu}_2\,(J,\sigma) + \mathbf{i}\underline{\mu}_3\,(J,\sigma)\,, v\right\rangle = \int_\Sigma \langle \sigma, -2J\bar{\partial}_J v)\rangle_{Q(J)\times T\mathcal{J}} \cdot \rho \tag{4.79}$$

where the pairing is defined by (4.48). In a holomorphic chart $U \subset \Sigma$ write

$$J(z) = J_0, \qquad \sigma(z) = f(z)dz^2, \qquad v = v(z), \qquad \rho = \lambda dx \wedge dy$$

for smooth functions $f, v : U \to \mathbb{C}$ and $\lambda : U \to \mathbb{R}^+$. Then (4.48) yields

$$\begin{aligned}\langle \sigma, -2J\bar{\partial}_J v\rangle_{Q\times T\mathcal{J}} \cdot \rho &= 2\mathbf{i}\overline{f(z)\frac{\partial}{\partial \bar{z}} v(z)}dx \wedge dy \\ &= \overline{\frac{\partial}{\partial \bar{z}}(f(z)v(z))}\, d\bar{z} \wedge dz - \overline{\frac{\partial f(z)}{\partial \bar{z}}} v(z)\, d\bar{z} \wedge dz \\ &= -\partial_J\overline{\iota(v)\sigma} - 2\mathbf{i}\iota(v)r(\overline{\bar{\partial}_J\sigma})\rho.\end{aligned}$$

Here we used that $r(\bar{\partial}\sigma) = \lambda^{-1}\frac{\partial f}{\partial \bar{z}} dz$ in local coordinates. Thus

$$\int_\Sigma \langle \sigma, -2J\bar{\partial}_J v\rangle_{Q(J)\times T\mathcal{J}} \cdot \rho = -2\mathbf{i}\int_\Sigma \overline{\iota(v) r(\bar{\partial}_J \sigma)}\rho. \tag{4.80}$$

Step 3 follows from (4.79), and (4.80).

Step 4: *(4.73) defines an equivariant moment maps for the action of Ham(Σ,ρ) with respect to $\underline{\omega}_2$ and $\underline{\omega}_3$.*

Let $H : \Sigma \to \mathbb{R}$ be a Hamiltonian and define $v_H \in \mathrm{Vect}(M)$ by $\iota(v_H)\rho = dH$. Then

$$-2\mathbf{i}\int_\Sigma \overline{\iota(v_H) r(\bar{\partial}_J \sigma)}\rho = -2\mathbf{i}\int_\Sigma \overline{r(\bar{\partial}_J\sigma)} \wedge dH = -2\mathbf{i}\int_\Sigma H\overline{\bar{\partial}_J r(\bar{\partial}_J \sigma)}$$

where we used integration by parts and that $r(\bar{\partial}_J\sigma)$ is a $(1,0)$ form. Equation (4.73) follows now from Step 3. □

4.5.4 Construction of the moduli space

The Hamiltonian quotient space

The hyperkähler quotient of $\mathcal{Q}_1(\Sigma)$ by $\mathrm{Ham}(\Sigma,\rho)$ is defined by

$$\begin{aligned}\mathcal{M}_0 &:= \underline{\mu}_1^{-1}(0)\cap \underline{\mu}_2^{-1}(0)\cap \underline{\mu}_3^{-1}(0)/\mathrm{Ham}(\Sigma,\rho) \\ &= \left\{(J,\sigma)\in \mathcal{Q}_1(\Sigma)\;\middle|\; \underline{\mu}_1(J,\sigma)=0,\ \bar{\partial}_J r(\bar{\partial}_J\sigma)=0\right\}\Big/\mathrm{Ham}(\Sigma,\rho)\end{aligned} \tag{4.81}$$

where $\underline{\mu}_1, \underline{\mu}_2, \underline{\mu}_3$ are the moment maps calculated in Theorem 4.5.13 for the $\mathrm{Ham}(\Sigma,\rho)$-action on $\mathcal{Q}_1(\Sigma)$. It follows from general principles that $\mathcal{M}_0$ is a hyperkähler manifold.

The next lemma is formulated in a finite dimensional setting, but extends formally to our case. It indicates that transversality for the hyperkähler moment map is an automatic consequence of our setup.

Lemma 4.5.14. *Let (M, g, I_1, I_2, I_3) be a hyperkähler manifold and let G be a Lie group. Suppose G acts freely on M by hyperkähler isometries and admits a hyperkähler moment map*

$$\mu = (\mu_1,\mu_2,\mu_3) : M \to \mathbb{R}^3 \otimes \mathfrak{g}^*.$$

Then 0 is a regular value of μ.

Proof. Let $x \in M$ with $\mu(x) = 0$ be given. Denote by $L_x : \mathfrak{g} \to T_xM$ its infinitesimal action and decompose $T_xM = W_0 \oplus W_1$ with

$$W_1 := \mathrm{Im}(L_x), \qquad W_0 = \mathrm{Im}(L_x)^\perp.$$

Equivariance of the moment map yields for all $\xi,\eta\in\mathfrak{g}$ the identity

$$\langle I_i L_x\xi, L_x\eta\rangle = \langle d\mu_i(x)L_x\eta, \xi\rangle = \langle \mu_i(x), [\eta,\xi]\rangle = 0.$$

This shows that the three complex structures map W_1 into W_0.

Let $\eta_1, \eta_2, \eta_3 \in \mathfrak{g}^*$ be given. Since G acts freely, L_x is injective, and the dual map $L_x^* : T_xM \to \mathfrak{g}^*$ is surjective with kernel W_0. Hence there exist $u_i \in W_1$ with $\eta_i = L_x^*(\alpha_i)$ for $i = 1, 2, 3$. For $v := -(I_1u_1 + I_2u_2 + I_3u_3)$ we then obtain

$$\langle d\mu_i(x)v, \xi\rangle = \omega_i(L_x\xi, v) = g(L_x\xi, I_iv) = g(L_x\xi, u_i) = \langle \eta_i, \xi\rangle.$$

This proves surjectivity of $d\mu : T_xM \to \mathbb{R}^3 \otimes \mathfrak{g}^*$ and the lemma. □

Construction of the symplectic quotient space

Denote by $\mathcal{M}_0 := \mathcal{Q}_1//\mathrm{Ham}(\Sigma, \rho)$ the Hamiltonian quotient (4.81). Recall from Remark 4.4.7 that the quotient

$$H := \mathrm{Symp}^0(\Sigma, \rho)/\mathrm{Ham}(\Sigma, \rho) \cong H^1(\Sigma, \mathbb{R})$$

is a well-defined group where the identification with $H^1(\Sigma, \mathbb{R})$ is obtained using the flux homomorphism. The Lie algebra of H is the quotient

$$\mathrm{Lie}(H) := \frac{\{v \in \mathrm{Vect}(\Sigma) \,|\, d\iota(v)\rho = 0\}}{\{v \in \mathrm{Vect}(\Sigma) \,|\, \iota(Jv)\rho \text{ exact}\}}$$

The $\mathrm{Symp}_0(\Sigma, \rho)$ action on $\mathcal{Q}_1(\Sigma)$ induces an action of H on $\mathcal{M}_0$ preserving the hyperkähler structure.

Proposition 4.5.15.

1. *The H-action is Hamiltonian with respect to $\underline{\omega}_2$ and $\underline{\omega}_3$: For $[v] \in Lie(H)$ and $[J, \sigma] \in \mathcal{M}_0$ the following maps*

$$\left\langle \tilde{\mu}_2^H([J, \sigma]) + \boldsymbol{i}\tilde{\mu}_3^H([J, \sigma])[v]\right\rangle = -2\boldsymbol{i}\int_\Sigma \overline{\iota(v)r(\bar{\partial}_J\sigma)}\rho \tag{4.82}$$

are well-defined equivariant moment maps for $\underline{\omega}_2$ and $\underline{\omega}_3$ respectively.

2. *On $\mathcal{M}_0$ the equation $\tilde{\mu}_2^H([J, \sigma]) = 0$ is equivalent to $\tilde{\mu}_3^H([J, \sigma]) = 0$ and*

$$(\tilde{\mu}_2^H)^{-1}(0) = (\tilde{\mu}_3^H)^{-1}(0) = \left\{[J, \sigma] \in \mathcal{M}_0 \,|\, \bar{\partial}_J\sigma = 0\right\} \tag{4.83}$$

3. *$(\tilde{\mu}_2^H)^{-1}(0) = (\tilde{\mu}_3^H)^{-1}(0)$ is a J_1–complex submanifold*

4. *The H-orbits in $(\tilde{\mu}_2^H)^{-1}(0) = (\tilde{\mu}_3^H)^{-1}(0)$ are J_1–complex submanifolds.*

Proof. Theorem 4.5.13 show that the action of $\mathrm{Symp}_0(\Sigma, \rho)$ on $\mathcal{Q}_1(\Sigma)$ is Hamiltonian with respect to $\underline{\omega}_2$ and $\underline{\omega}_3$. This directly implies that the action of H is Hamiltonian for the symplectic forms induced by $\underline{\omega}_2$ and $\underline{\omega}_3$. The formula for the moment maps (4.82) follows from (4.74).

For fixed J, one can identify the Lie algebra of H with

$$\mathfrak{h}_J := \{v \in \mathrm{Vect}(\Sigma) \,|\, d\iota(v)\rho = 0 = d\iota(Jv)\rho\}$$

by Hodge theory. This is a J-invariant subspace and it holds $\tilde{\mu}_2^H([J,\sigma],Jv) = \tilde{\mu}_3^H([J,\sigma],v)$ for all $v \in \mathfrak{h}_J$. Hence $(\tilde{\mu}_2^H)^{-1}(0) = (\tilde{\mu}_3^H)^{-1}(0)$. We prove (4.83) next. Let $[J,\sigma] \in (\tilde{\mu}_2^H)^{-1}(0) = (\tilde{\mu}_3^H)^{-1}(0)$ be given. Then

$$0 = \int_\Sigma \iota(v) r(\bar{\partial}_J\sigma)\rho = \int_\Sigma r(\bar{\partial}_J\sigma) \wedge \iota(v)\rho$$

for all $v \in \mathfrak{h}_J$. The defining equations of $\mathcal{M}_0$ show that $r(\bar{\partial}_J\sigma) \in \Omega_J^{1,0}(\Sigma,\mathbb{C})$ is closed. Since $\{\iota(v)\rho \,|\, v \in \mathfrak{h}_J\}$ parametrizes the space of (real) harmonic 1-forms, it follows from Poincaré duality that $[r(\bar{\partial}_J\sigma)] = 0 \in H_J^{1,0}(\Sigma)$. Hence $r(\bar{\partial}_J\sigma) = \partial_J f$ with $\bar{\partial}_J\partial_J f = 0$. Then f is constant and therefore $\bar{\partial}_J\sigma = 0$.

Let $[J,\sigma] \in (\tilde{\mu}_2^H)^{-1}(0) = (\tilde{\mu}_3^H)^{-1}(0)$. It follows from the moment map equations that multiplication with J_2 and J_3 yields isomorphism

$$J_2 : T_{[J,\sigma]}\left(H \cdot [J,\sigma]\right) \to \left(T_{[J,\sigma]}\left(\tilde{\mu}_2^H\right)^{-1}(0)\right)^\perp$$

$$J_3 : T_{[J,\sigma]}\left(H \cdot [J,\sigma]\right) \to \left(T_{[J,\sigma]}\left(\tilde{\mu}_3^H\right)^{-1}(0)\right)^\perp$$

Hence $J_1 = J_2J_3$ maps $T_{[J,\sigma]}\left(H \cdot [J,\sigma]\right)$ and $T_{[J,\sigma]}\left(\tilde{\mu}_2^H\right)^{-1}(0)$ onto themselves. Therefore $\left(\tilde{\mu}_2^H\right)^{-1}(0) = \left(\tilde{\mu}_3^H\right)^{-1}(0)$ is a J_1-complex submanifold of $\mathcal{M}_0$ and the H-orbits are complex submanifolds. □

Consider the moduli space

$$\mathcal{M}_s := \left\{(J,\sigma) \in \mathcal{Q}(\Sigma) \,\middle|\, \underline{\mu}_1(J,\sigma) = 0,\, \bar{\partial}_J\sigma = 0,\, |\sigma| < 1\right\} \Big/ \mathrm{Symp}_0(\Sigma,\rho) \tag{4.84}$$

It follows from Proposition 4.5.15 above that $\mathcal{M}_s$ carries a natural hyperkähler structure: Since $\mathcal{M}_s$ is a Marsden-Weinstein quotient of $\mathcal{M}_0$ for the symplectic structures induced by $\underline{\omega}_2$ and $\underline{\omega}_3$, it follows that they descend to symplectic structures on $\mathcal{M}_s$. Using Lemma 4.4.9, $\underline{\omega}_1$ also provides a natural symplectic structure on $\mathcal{M}_s$. This yields three algebraically compatible symplectic forms on $\mathcal{M}_s$ and a lemma of Hitchin ([58], Lemma 6.8) shows that this defines a hyperkähler structure.

Proposition 4.5.16. *Let $(J,\sigma) \in \mathcal{Q}_1(\Sigma)$ and assume $\bar{\partial}_J\sigma = 0$. Then*

$$\underline{\mu}_1(J,\sigma) = \left(2K_J + \Delta\log(1+\sqrt{1-|\sigma|^2})\right)\rho - 2c\rho \tag{4.85}$$

were $\Delta = d^*d$ is the positive Laplaction for the metric $\rho(\cdot,J\cdot)$. In particular,

$$\mathcal{M}_s = \left\{(J,\sigma) \in \mathcal{Q}(\Sigma) \,\middle|\, \begin{array}{c} \bar{\partial}_J\sigma = 0, \quad |\sigma| < 1 \\ K_J + \frac{1}{2}\Delta\log(1+\sqrt{1-|\sigma|^2}) = c \end{array}\right\} \Big/ Symp_0(\Sigma,\rho). \tag{4.86}$$

where $c := 2\pi(2 - 2genus(\Sigma))/vol(\Sigma,\rho)$.

Proof. Consider the function $h := |\sigma|^2$. It suffices to prove the lemma around a point where $h \neq 0$. We also simplify notation and abbreviate $\bar{\partial} := \bar{\partial}_J$ and $\partial := \partial_J$.

Step 1: *Suppose $\bar{\partial}\sigma = 0$, then*

$$\underline{\mu}_1(J,\sigma) = \boldsymbol{i}\frac{\bar{\partial}h \wedge \partial h}{2h\sqrt{1-h}} - \boldsymbol{i}\sqrt{1-h}\bar{\partial}\partial\log(h) + 2\boldsymbol{i}\bar{\partial}\partial\sqrt{1-h} - 2c\rho \tag{4.87}$$

Denote by h_Q the hermitian form on the bundle of quadratic differentials induced by J and ρ. Since $\bar{\partial}\sigma = 0$, it follows $\partial h = \partial h_Q(\sigma,\sigma) = h_Q(\sigma, \partial\sigma)$ and then

$$|\partial h|^2 = |h_Q(\sigma,\partial\sigma)|^2 = |\sigma|^2|\partial\sigma|^2 = h|\partial\sigma|^2. \tag{4.88}$$

Using $|\partial h|^2\rho = -\frac{\mathbf{i}}{2}\,\bar{\partial}h \wedge \partial h$, we then obtain

$$|\partial\sigma|^2\rho = -\frac{\mathbf{i}}{2h}\bar{\partial}h \wedge \partial h \tag{4.89}$$

Next choose holomorphic coordinates and write

$$\rho = \lambda dx \wedge dy, \qquad \sigma(z) = f(z)dz^2$$

for some positive function λ and a holomorphic function f. The Gaussian curvature K_J can be computed in these coordinates via

$$K_J = -\frac{1}{2}\lambda^{-1}(\partial_x^2\log(\lambda) + \partial_y^2\log(\lambda)).$$

Since $f(z)$ is holomorphic, $\log(|f(z)|^2)$ is harmonic and we compute

$$\begin{aligned}\bar{\partial}\partial\log(h) &= \frac{1}{4}(\partial_x^2 + \partial_y^2)\log(|f(z)|^2\lambda^{-2})2\mathbf{i}\,dx \wedge dy \\ &= -\mathbf{i}(\partial_x^2 + \partial_y^2)\log(\lambda)\,dx \wedge dy \\ &= 2\mathbf{i}K_J\rho.\end{aligned}$$

This shows

$$K_J\rho = -\frac{\mathbf{i}}{2}\bar{\partial}\partial\log(h). \tag{4.90}$$

Plugging (4.89) and (4.90) into (4.72) and using $\bar{\partial}\sigma = 0$ yields (4.87).

Step 2: *Suppose $\bar{\partial}\sigma = 0$, then μ_1 is given by (4.85)*

Form (4.87) and integration by parts, it follows

$$\begin{aligned}\underline{\mu}_1(J,\sigma) &= \mathbf{i}\frac{\bar{\partial}h \wedge \partial h}{2h\sqrt{1-h}} - \mathbf{i}\sqrt{1-h}\bar{\partial}\partial\log(h) + 2\mathbf{i}\bar{\partial}\partial\sqrt{1-h} - 2c\rho \\ &= \mathbf{i}\left[-\bar{\partial}\sqrt{1-h} \wedge \partial\log(h) - \sqrt{1-h}\bar{\partial}\partial\log(h) + 2\bar{\partial}\partial\sqrt{1-h}\right] - 2c\rho \\ &= \mathbf{i}\bar{\partial}\left[-\sqrt{1-h}\partial\log(h) + 2\partial\sqrt{1-h}\right] - 2c\rho\end{aligned}$$

A primitive for the inner expression is obtained by the following calculation:

$$\begin{aligned}
-\sqrt{1-h}\partial\log(h)+2\partial\sqrt{1-h} &= \frac{-1}{h\sqrt{1-h}}\partial h \\
&= \frac{-1+\sqrt{1-h}}{h\sqrt{1-h}}\partial h-\partial\log(h) \\
&= \frac{-1}{\sqrt{1-h}(1+\sqrt{1-h})}\partial h-\partial\log(h) \\
&= 2\partial\log(1+\sqrt{1-h})-\partial\log(h)
\end{aligned}$$

Using (4.90) this yields

$$\begin{aligned}
\underline{\mu}_1(J,\sigma) &= -\mathbf{i}\bar\partial\partial\log(h)+2\mathbf{i}\bar\partial\partial\log(1+\sqrt{1-h})-2c\rho \\
&= 2K_J\rho+\Delta(\log(1+\sqrt{1-h}))\rho-2c\rho
\end{aligned}$$

where $\Delta=d^*d$ is the positive Laplacian which satisfies $(\Delta h)\rho=2\mathbf{i}\bar\partial\partial h$. □

Metric description of the moduli space

Denote by $\mathrm{Met}(\Sigma)$ the space of Riemannian metrics g on Σ. For every $g\in\mathrm{Met}(\Sigma)$ there exists a unique complex structure $J=J_g\in\mathcal{J}(\Sigma)$ which is compatible with g and ρ. In the following, we always refer to this complex structure, when discussing holomorphic objects on (Σ,g). Define

$$\mathcal{M}_d := \left\{(g,\sigma)\;\middle|\; \begin{array}{c} g\in\mathrm{Met}(\Sigma),\ \sigma\in Q(g),\ |\sigma|<1 \\ \bar\partial\sigma=0,\ K_g+\frac{1}{2}\Delta\log(1+\sqrt{1-|\sigma|^2})=c \end{array}\right\}\Big/\mathrm{Diff}_0(\Sigma). \tag{4.91}$$

Standard Moser isotopy arguments (as in Proposition 4.4.5) show that the canonical inclusion $\mathcal{M}_s\to\mathcal{M}_d$ is an isomorphism, where $\mathcal{M}_s$ is defined by (4.86). The next proposition provides a simpler description of this moduli space.

Proposition 4.5.17. *Consider on the space of pairs (g,σ) with $g\in Met(\Sigma)$ and $\sigma\in Q(g)$ with $|\sigma|<1$ the self-map defined by*

$$(g,\sigma)\mapsto\left(\left(1+\sqrt{1-|\sigma|_g^2}\right)\cdot g,\ \sigma\right).$$

This induces a well-defined isomorphism between $\mathcal{M}_d$ and

$$\mathcal{M} := \left\{(g,\sigma)\;\middle|\; \begin{array}{c} g\in Met(\Sigma),\ \sigma\in Q(g),\ |\sigma|<1 \\ \bar\partial\sigma=0,\ K_g-\frac{c}{2}|\sigma|^2=\frac{c}{2} \end{array}\right\}\Big/Diff_0(\Sigma) \tag{4.92}$$

where $c:=2\pi(2-2genus(\Sigma))/vol(\Sigma,\rho)$.

Proof. Let $g_0\in\mathrm{Met}(\Sigma)$, let $\sigma\in Q(g_0)$ with $|\sigma|_{g_0}<1$ and define

$$f:=1+\sqrt{1-|\sigma|_{g_0}^2},\qquad g:=fg_0.$$

Then $|\sigma|_g = |\sigma|_{g_0}/f < 1$. For the converse direction, use the relation $(f-1)^2 = 1 - |\sigma|_{g_0}^2$ to obtain

$$\frac{1}{f} = \frac{1 + |\sigma|_{g_0}^2/f^2}{2} = \frac{1 + |\sigma|_g^2}{2}. \tag{4.93}$$

It follows that one can recover g_0 from (g, σ) via $g_0 = 2|\sigma|_g/(1+|\sigma|_g^2)\cdot g$. In particular

$$|\sigma|_{g_0} = \frac{2|\sigma|_g}{1 + |\sigma|_g}.$$

and this shows that $|\sigma|_{g_0} < 1$ if and only if $|\sigma|_g < 1$. The Gaussian curvature changes under the conformal change as follows

$$K_g = \frac{1}{f}\left(K_{g_0} + \frac{1}{2}\Delta_{g_0}\log(f)\right).$$

and (4.93) then yields

$$K_g = \frac{1 + |\sigma|_g^2}{2}\left(K_{g_0} + \frac{1}{2}\Delta_{g_0}\log(f)\right).$$

This proves the identification of $\mathcal{M}$ with $\mathcal{M}_d$ and the proposition. □

4.6 Three geometric models for the moduli space

We assume throughout this section that $\mathrm{genus}(\Sigma) \geq 2$ and that $V := \mathrm{vol}(\Sigma, \rho) = \pi(2\mathrm{genus}(\Sigma) - 2)$. The moduli space (4.92) constructed in the previous section is then given by

$$\mathcal{M} := \left\{ (g, \sigma) \;\middle|\; \begin{array}{c} g \in \mathrm{Met}(\Sigma),\, \sigma \in Q(g),\, |\sigma| < 1 \\ \bar{\partial}\sigma = 0,\, K_g + |\sigma|_g^2 = -1 \end{array} \right\} \Big/ \mathrm{Diff}_0(\Sigma) \tag{4.94}$$

It follows from the general construction that $\mathcal{M}$ carries a natural hyperkähler structure which extends the Weil–Petersson metric on Teichmüller space. The purpose of this section is to establish the following three geometric description of this moduli space proposed by Donaldson [38].

1. $\mathcal{M}$ embeds as an open neighbourhood of the zero section into the cotangent bundle of Teichmüller space $\mathcal{T}(\Sigma)$. The hyperkähler metric on $\mathcal{M}$ can then be viewed as the Feix–Kaledin hyperkähler extension of the Weil–Petersson metric on $\mathcal{T}(\Sigma)$.

2. $\mathcal{M}$ parametrizes the class of almost-Fuchsian hyperbolic 3-manifolds. These are quasi-Fuchsian 3-manifolds which possess an incompressible minimal surface with principal curvatures in $(-1, 1)$. This surface is then unique and its area provides a Kähler potential for the hyperkähler metric with respect to the second complex structure.

3. $\mathcal{M}$ embeds as an open subset into the smooth locus of the $\mathrm{SL}(2,\mathbb{C})$ representation variety $\mathcal{R}_{\mathrm{SL}(2,\mathbb{C})}(\Sigma) := \mathrm{Hom}\left(\pi_1(\Sigma), \mathrm{SL}(2,\mathbb{C})\right)/\mathrm{SL}(2,\mathbb{C})$. The natural complex structure in this picture corresponds to minus the second complex structure in the first picture and the Goldman holomorphic symplectic form on $\mathcal{R}_{\mathrm{SL}(2,\mathbb{C})}(\Sigma)$ (see [52]) restricts to $-\underline{\omega}_1 + \mathbf{i}\underline{\omega}_3$ along the moduli space $\mathcal{M}$.

First, we recall a construction of Uhlenbeck [117] which associate to every pair $[g,\sigma] \in \mathcal{M}$ a complete hyperbolic 3-manifold. The isomorphism between $\mathcal{M}$ and the almost-Fuchsian moduli space is then given in Theorem 4.6.4. Next, following Hodge [61] we describe an explicit embedding of $\mathcal{M}$ into $\mathcal{T}(\Sigma)\times\overline{\mathcal{T}(\Sigma)}$ in Theorem 4.6.4. This map is not surjective and both maps are related by the simultaneous uniformization theorem of Bers [8], stated in Theorem 4.6.6. By the Cartan–Ambrose–Higgs theorem, one can express every complete hyperbolic 3-manifold as quotient of hyperbolic space $\mathbb{H}^3$. This gives rise to a natural embedding of the almost Fuchsian moduli space into $\mathcal{R}_{\mathrm{PSL}(2,\mathbb{C})}(\Sigma)$. Theorem 4.6.12 was outlined by Donaldson [38] and describes a lift of this embedding from $\mathcal{M}$ into $\mathcal{R}_{\mathrm{SL}(2,\mathbb{C})}(\Sigma)$ using the theory of Higgs bundles [58]. Finally, we recall in Theorem 4.6.14 a well-known result of Uhlenbeck [117] which states that the natural map of $\mathcal{M}$ into $T^*\mathcal{T}(\Sigma)$ is a well-defined embedding.

4.6.1 Germs of hyperbolic 3-manifolds and almost-Fuchsian metrics

The equation

$$K_g + |\sigma|^2 = -1, \qquad \bar{\partial}\sigma = 0 \tag{4.95}$$

for pairs $g \in \mathrm{Met}(\Sigma)$ and $\sigma \in Q(g)$ appeared in Uhlenbeck [117] and Taubes [108] for minimal surfaces in hyperbolic 3-manifolds. It follows from Theorem 4.6.1 that solutions of (4.95) correspond to *minimal hyperbolic germs over* Σ. When the additional pointwise constraint $|\sigma|_g \leq 1$ is satisfied, then every solution of (4.95) embeds as a minimal surfaces into a complete hyperbolic 3-manifold $Y \cong \Sigma \times \mathbb{R}$. A hyperbolic 3-manifold is a connected, oriented, Riemannian 3-manifold (Y, g^Y) with constant sectional curvature -1. In particular, we do not assume completeness of (Y, g^Y). This section contains a proof of the following theorem of Uhlenbeck [117].

Theorem 4.6.1 (Germs of hyperbolic 3-manifolds).

1. *Suppose $g \in Met(\Sigma)$ and $\sigma \in Q(g)$ satisfy (4.95). Then there exists a hyperbolic 3-manifold (Y, g^Y) and a minimal isometric embedding*

$$\iota : (\Sigma, g) \hookrightarrow (Y, g^Y)$$

with second fundamental form $h = Re(\sigma)$. Moreover, the hyperbolic metric g^Y is uniquely determined by (g,σ) in a tubular neighbourhood of $\iota(\Sigma) \subset Y$.

2. *Let (Y, g^Y) be a hyperbolic 3-manifold, let $\iota : \Sigma \hookrightarrow (Y, g^Y)$ be a minimal embedding with second fundamental form h and denote $g := \iota^* g^Y \in Met(\Sigma)$. Then there exists $\sigma \in Q(g)$ with $h = Re(\sigma)$ such that (g,σ) satisfies (4.95).*

Proof. The proof consists of four steps.

Step 1: *Assume $\iota : (\Sigma, g) \hookrightarrow (Y, g^Y)$ is an isometric embedding. Then there exists $\epsilon > 0$, an open neighbourhood $\iota(\Sigma) \subset W \subset Y$ and a diffeomorphism $f : \Sigma \times (-\epsilon, \epsilon) \to W$ such that $\iota(z) = f(z, 0)$ for all $z \in \Sigma$ and the pullback metric f^*g^Y on $\Sigma \times (-\epsilon, \epsilon)$ has the form*

$$(f^*g^Y)(z,t) = \begin{pmatrix} g_t(z) & 0 \\ 0 & 1 \end{pmatrix} \tag{4.96}$$

for a smooth family of metrics $t \mapsto g_t \in Met(\Sigma)$ with $g_0 = g$.

Since Σ and Y are both orientable, there exists unit normal vector field $\nu \in \Omega^0(\Sigma, \iota^*(T^\perp \iota(\Sigma)))$. For sufficiently small $\epsilon > 0$ the map

$$f : \Sigma \times (-\epsilon, \epsilon) \to Y, \qquad f(z,t) := \exp_z(t\nu(z))$$

is a diffeomorphism onto its image which has all the desired properties.

Step 2: *Assume $Y = \Sigma \times (-\epsilon, \epsilon)$ with Riemannian metric (4.96). The second fundamental form of $\Sigma \times \{0\} \subset Y$, is given by*

$$h = -\frac{1}{2} \frac{d}{dt}\bigg|_{t=0} g_t. \tag{4.97}$$

Moreover, for $t \in (-\epsilon, \epsilon)$ define $k_t \in \Omega^0(\Sigma, T^*\Sigma \otimes T^*\Sigma)$ by

$$k_t(z; u, v) := g^Y \left(R^Y_{(z,t)}((u,0),(0,1))(0,1),(v,0) \right) \tag{4.98}$$

for all $z \in \Sigma$ and $u, v \in T_z\Sigma$. This satisfies the equation

$$k_t = -\frac{1}{2}\ddot{g}_t + \frac{1}{4}\dot{g}_t g_t^{-1} \dot{g}_t \tag{4.99}$$

for all $t \in (-\epsilon, \epsilon)$.

Choose local coordinates (x^1, x^2, x^3) on $Y = \Sigma \times (-\epsilon, \epsilon)$ such that (x^1, x^2) are coordinates for Σ and x^3 parametrizes $(-\epsilon, \epsilon)$ in unit speed. Then

$$\Gamma^3_{ij} = -\frac{1}{2}\partial_3 g_{ij}, \qquad \text{for } i, j \in \{1, 2\}$$

and this yields (4.97) in local coordinates. Moreover,

$$\Gamma^k_{i3} = \frac{1}{2}\sum_{\ell=1}^{2} g^{k\ell}\partial_3 g_{i\ell}, \qquad \Gamma^3_{i3} = \Gamma^3_{33} = \Gamma^k_{33} = 0$$

for $i, k \in \{1, 2\}$. Therefore

$$\begin{aligned}
R^3_{3jk} &= \partial_3 \Gamma^3_{jk} - \partial_j \Gamma^3_{3k} + \sum_{\nu=1}^{3} \left(\Gamma^\ell_{3\nu} \Gamma^\nu_{jk} - \Gamma^3_{j\nu} \Gamma^\nu_{3k} \right) \\
&= -\frac{1}{2} \partial_3^2 g_{jk} - \Gamma^3_{j1} \Gamma^1_{3k} - \Gamma^3_{j2} \Gamma^2_{3k} \\
&= -\frac{1}{2} \partial_3^2 g_{jk} + \frac{1}{4} \sum_{i,\ell=1}^{2} (\partial_3 g_{ij}) g^{i\ell} (\partial_3 g_{k\ell})
\end{aligned}$$

and this establishes (4.99) is local coordinates.

Step 3: *Suppose $g \in Met(\Sigma)$ and $\sigma \in Q(g)$ satisfies (4.95). Let $\epsilon > 0$ and let $(-\epsilon, \epsilon) \to Met(\Sigma)$, $t \mapsto g_t$, be a solution of the second order ODE*

$$g_0 = g, \qquad \dot{g}_0 = -2Re(\sigma), \qquad \frac{1}{2} \ddot{g}_t - \frac{1}{4} \dot{g}_t g_t^{-1} \dot{g}_t = g_t. \tag{4.100}$$

Then g^Y defined by (4.96) is a hyperbolic metric on $Y = \Sigma \times (-\epsilon, \epsilon)$.

The metric g^Y has constant sectional curvature -1 if and only if

$$\langle R^Y(v_1, v_2) v_3, v_4 \rangle_{g^Y} = (-1) \left(\langle v_1, v_4 \rangle_{g^Y} \langle v_2, v_3 \rangle_{g^Y} - \langle v_1, v_3 \rangle_{g^Y} \langle v_2, v_4 \rangle_{g^Y} \right)$$

for all $q \in Y$ and $v_1, v_2, v_3, v_4 \in T_q Y$. By symmetry of the curvature tensor, this is equivalent to the following three instances of the curvature equation:

$$\langle R^Y((u,0),(0,1))(0,1),(v,0) \rangle_{g^Y} = -g_t(u,v) \tag{4.101}$$

$$\langle R^Y((u,0),(v,0))(0,1),(w,0) \rangle_{g^Y} = 0 \tag{4.102}$$

$$\langle R^Y((u,0),(v,0))(v,0),(u,0) \rangle = -g_t(u,u) g_t(v,v) + g_t(u,v)^2. \tag{4.103}$$

for every point $(z,t) \in Y$ and $u, v, w \in T_z \Sigma$.

Equation (4.101) follows from (4.98), (4.99) and (4.100). By Lemma 4.6.2 below, (4.102), (4.103) are satisfied for $t = 0$. By differentiating these two equations, we will show that they remain valid for all $t \in (-\epsilon, \epsilon)$.

Extend $(u,0), (v,0), (w,0) \in T_{(z,t)} Y$ to vector fields which are parallel in t direction. The second Bianchi identity yields at the point (z,t):

$$\begin{aligned}
&\partial_t \langle R^Y((u,0),(v,0))(0,1),(w,0) \rangle_{g^Y} \\
&= -\langle \nabla_{(u,0)} R^Y((v,0),(0,1)(0,1),(w,0) \rangle_{g^Y} \\
&\qquad + \langle \nabla_{(v,0)} R^Y((u,0),(0,1))(0,1),(w,0) \rangle_{g^Y} \\
&= -\mathcal{L}_u k_t(v,w) + \mathcal{L}_v k_t(u,w) - k_t([u,v],w) + k_t(v, \nabla^t_u w) - k_t(u, \nabla^t_v w) \\
&= \mathcal{L}_u g_t(v,w) - \mathcal{L}_v g_t(u,w) + g_t([u,v],w) - g_t(v, \nabla^t_u w) + g_t(u, \nabla^t_v w) \\
&= 0
\end{aligned}$$

where ∇^t denotes the Levi-Civita connection on Σ for g_t. The penultimate equation uses (4.101), i.e $g_t = -k_t$, and this completes the proof of (4.102).

We deduce from (4.101) and (4.102) the following auxiliary result:

$$\nabla_t R((u,0),(0,1)) = 0 \tag{4.104}$$

for all $u \in \mathrm{Vect}(\Sigma)$. We get from (4.101) that

$$\begin{aligned}
&\langle \nabla_t R^Y((u,0),(0,1))(0,1),(0,v)\rangle_{g^Y} \\
&\quad = \partial_t \langle R^Y((u,0),(0,1))(0,1),(0,v)\rangle_{g^Y} - \langle R^Y(\nabla_t(u,0),(0,1))(0,1),(0,v)\rangle_{g^Y} \\
&\qquad - \langle R^Y((u,0),(0,1))(0,1),\nabla_t(0,v)\rangle_{g^Y} \\
&\quad = \partial_t \langle (u,0),(v,0)\rangle_{g^Y} - \langle \nabla_t(u,0),(v,0)\rangle_{g^Y} - \langle (u,0),\nabla_t(v,0)v\rangle_{g^Y} \\
&\quad = 0
\end{aligned}$$

for all $u, v \in \mathrm{Vect}(\Sigma)$. On the other hand it holds

$$\begin{aligned}
-\langle \nabla_t R^Y((u,0),(0,1))(0,v),(0,w)\rangle_{g^Y} &= \langle R^Y(\nabla_t(u,0),(0,1))(0,v),(0,w)\rangle_{g^Y} \\
&\quad + \langle R^Y((u,0),(0,1))\nabla_t(0,v),(0,w)\rangle_{g^Y} \\
&\quad + \langle R^Y((u,0),(0,1))(0,v),\nabla_t(0,w)\rangle_{g^Y} \\
&= 0
\end{aligned}$$

for all $u, v, w \in \mathrm{Vect}(\Sigma)$. Here we used that $\langle \nabla_t(v,0),(0,1)\rangle_{g^Y} = 0$ for any vector field $v \in \mathrm{Vect}(\Sigma)$, the equation (4.102) and symmetries of the curvature tensor to conclude that all three terms on the right hand side vanish separately. This establishes (4.104)

We proceed to the proof of (4.103): For two linearly independet vectors $u, v \in T_z\Sigma$ define linear maps $\pi_u, \pi_v : T_z\Sigma \to \mathbb{R}$ such that $w = \pi_u(w)u + \pi_v(w)w$ for all $w \in T_z\Sigma$. They are given by the formulas

$$\pi_u(w) = \frac{g_t(v,v)g_t(w,u) - g_t(w,v)g_t(u,v)}{g_t(u,u)g_t(v,v) - g_t(u,v)^2} \tag{4.105}$$

$$\pi_v(w) = \frac{g_t(u,u)g_t(w,v) - g_t(u,w)g_t(u,v)}{g_t(u,u)g_t(v,v) - g_t(u,v)^2}. \tag{4.106}$$

It follows from (4.104) that

$$\begin{aligned}
&\partial_t \langle R^Y((u,0),(v,0))(v,0),(u,0)\rangle_{g^Y} \\
&\quad = 2\langle R^Y(\nabla_t(u,0),(0,v))(0,v),(0,u)\rangle_{g^Y} + 2\langle R^Y((u,0),\nabla_t(0,v))(0,v),(0,u)\rangle_{g^Y} \\
&\quad = 2\left(\pi_u(\nabla_t(u,0)) + \pi_v(\nabla_t(0,v)\right)\langle R^Y((u,0),(v,0))(v,0),(u,0)\rangle_{g^Y}
\end{aligned}$$

For the last step, note that $\langle \nabla_t(u,0),(0,1)\rangle_{g^Y} = 0$ and so $\nabla_t(u,0)$ (and similarly $\nabla_t(v,0)$) can be identified with a tangent vector in $T_z\Sigma$. Using the relations

$$\langle \nabla_t(u,0),(u,0)\rangle_{g^Y} = \frac{1}{2}\dot{g}_t(u,u), \qquad \langle \nabla_t(v,0),(v,0)\rangle_{g^Y} = \frac{1}{2}\dot{g}_t(v,v)$$

$$\langle \nabla_t(u,0),(v,0)\rangle_{g^Y} + \langle (u,0),\nabla_t(v,0)\rangle_{g^Y} = \dot{g}_t(u,v)$$

and the equations (4.105), (4.106) it follows

$$
\begin{aligned}
&2\left(\pi_u(\nabla_t(u,0)) + \pi_v(\nabla_t(0,v)\right) \\
&\quad = \left(\frac{\dot g_t(u,u)g_t(v,v) + \dot g_t(v,v)g_t(u,u) - 2\dot g_t(u,v)g_t(u,v)}{g_t(u,u)g_t(v,v) - g_t(u,v)^2}\right).
\end{aligned}
$$

Hence $R(t) := \langle R^Y_{(z,t)}((u,0),(v,0))(v,0),(u,0)\rangle_{g^Y}$ satisfies the ODE

$$
\dot R(t) = \left(\frac{\dot g_t(u,u)g_t(u,u) + \dot g_t(v,v)g_t(v,v) - \dot g_t(u,v)}{g_t(u,u)g_t(v,v) - g_t(u,v)^2}\right) R(t)
$$

with initial condition $R(0) = -g(u,u)g(v,v) + g(u,v)^2$. The unique solution of this ODE is $R(t) = -g_t(u,u)g_t(v,v) + g_t(u,v)^2$ and this proves 4.103).

Step 4: *Completion of the proof.*

For the first part, choose $\epsilon > 0$ such that (4.100) has a positive definite solution g_t for $t \in (-\epsilon,\epsilon)$. Then $Y = \Sigma \times (-\epsilon,\epsilon)$ with the Riemannian metric g^Y defined by (4.96) is a hyperbolic manifold by Step 3 and $\iota : \Sigma \to Y$, $\iota(z) := (z,0)$, is an isometric embedding with second fundamental form $h = \mathrm{Re}(\sigma)$ by Step 1. Moreover, $\sigma(Ju,Ju) = -\sigma(u,u)$ for any $u \in \mathrm{Vect}(\Sigma)$ yields $\mathrm{tr}(h) = 0$ and thus $\Sigma \times \{0\}$ is minimal. Uniqueness follows from uniqueness of the solution of (4.100). Conversely, the second part follows from Step 1 and Lemma 4.6.2 below. □

Lemma 4.6.2. *Let (Y,g^Y) be a Riemannian 3-manifold and let $(\Sigma,g) \subset (Y,g^Y)$ be an isometrically embedded minimal surface with second fundamental form h. Then there exists exists a unique quadratic differential $\sigma \in Q(g)$ with $h = Re(\sigma)$ and the following is satisfied:*

1. *σ is holomorphic if and only if*

$$
R^Y_z(u,v)w \in T_z\Sigma \qquad \text{for all } z \in \Sigma \text{ and } u,v,w \in T_z\Sigma
$$

where R^Y denotes the curvature tensor of the ambient manifold Y.

2. *The intrinsic and extrinsic curvature along Σ are related by*

$$
\frac{\langle R^Y(u,v)v,u\rangle_{g^Y}}{|u|_g^2|v|_g^2 - \langle u,v\rangle_g^2} = K_g + |\sigma|_g^2 \qquad \text{for all } u,v \in Vect(\Sigma)
$$

where K_g denotes the Gaussian curvature of (Σ,g).

Proof. Let $z \in \Sigma$ and choose conformal coordinates (x,y) in a neighborhood of z. In these coordinates h can be written as

$$
h(x,y) = h_{11}(x,y)dx^2 + h_{22}(x,y)dy^2 + 2h_{12}(x,y)dxdy.
$$

Since $\Sigma \subset Y$ is minimal, its mean curvature vanishes and thus $h_{11} = -h_{22}$. Define a quadratic differential by

$$
\sigma(x,y) = (h_{11}(x,y) - \mathbf{i}h_{12}(x,y))\,dz^2. \tag{4.107}
$$

This satisfies $h = \mathrm{Re}(\sigma)$ and, since the expression is conformally invariant, it defines a quadratic differential on Σ.

We prove 1. For $z \in \Sigma$ denote by $\Pi(z) : T_zY \to T_z\Sigma$ the orthogonal projection determined by g^Y. The Mainardi-Codazzi equation yields for $u, v, w \in \mathrm{Vect}(\Sigma)$:

$$\begin{aligned}
\big(R^Y_z(u,v)w\big)^{\perp} :&= (\mathbb{1} - \Pi(z))R^Y_z(u,v)w \\
&= (\nabla^\Sigma_u h)_z(v,w) - (\nabla^\Sigma_v h)_z(u,w) \\
&= \mathcal{L}_u h(v,w) - \mathcal{L}_v h(u,w) + h([u,v],w) + h(u, \nabla^\Sigma_v w) - h(v, \nabla^\Sigma_u w) \\
&= \mathcal{L}_u h(v,w) - \mathcal{L}_v h(u,w) + (\mathbb{1} - \Pi(z))\, R^\Sigma_z(u,v)w \\
&= \mathcal{L}_u h(v,w) - \mathcal{L}_v h(u,w).
\end{aligned}$$

In a conformal chart around z the equation $\big(R^Y(u,v)w\big)^{\perp} = 0$ is thus equivalent to

$$\partial_1 h_{22}(x,y) = \partial_2 h_{12}(x,y), \qquad \partial_2 h_{11}(x,y) = \partial_1 h_{12}(x,y).$$

Since $h_{11} = -h_{22}$, these are the Cauchy–Riemann equations for the function $h_{11}(x,y) - \mathbf{i} h_{12}(x,y)$ and hence equivalent to holomorphicity of σ.

We prove 2. The Gauss-Codazzi equation yields for $u, v \in \mathrm{Vect}(\Sigma)$:

$$\langle R^Y(u,v)v, u\rangle_{g^Y} = \langle R^\Sigma(u,v)v, u\rangle_g - h(u,u)h(v,v) - h(u,v)^2$$

For a unit vector field $u \in \mathrm{Vect}(\Sigma)$ with $|u|_g = 1$, it follows

$$|\sigma|_g^2 = \mathrm{Re}(\sigma(u))^2 + \mathrm{Im}(\sigma(u))^2 = -h(u,u)h(Ju,Ju) - h(u,Ju)^2$$

and $\langle R^\Sigma(u,Ju)Ju, u\rangle_g = K_g$. This yields

$$\langle R^Y(u,Ju)Ju, u\rangle = K_g + |\sigma|_g^2.$$

Hence $K_g + |\sigma|_g^2$ agrees with the sectional curvature of $T_z\Sigma \subset T_zY$ and this proves the lemma. □

Definition 4.6.3 (Almost-Fuchsian metrics). *We call a complete hyperbolic metric g^Y on $Y := \Sigma \times \mathbb{R}$ almost-Fuchsian when it has the product shape*

$$g^Y(z,t) = \begin{pmatrix} g_t(z) & 0 \\ 0 & 1 \end{pmatrix}$$

with $g_t \in Met(\Sigma)$ and such that $\Sigma \times \{0\} \subset Y$ is a minimal surface with principal curvatures in $(-1,1)$. Denote by $\mathcal{AF}(\Sigma)$ the space of all almost-Fuchsian metrics.

An almost-Fuchsian manifold is a hyperbolic 3-manifolds Y which is isometric to $\Sigma \times \mathbb{R}$ equipped with an almost-Fuchsian metric. The work of Uhlenbeck [117] proves that a complete hyperbolic 3-manifold Y is almost-Fuchsian if and only if it admits a minimal and incompressible embedding $\iota : \Sigma \hookrightarrow Y$ with principal curvatures in $(-1,1)$. This follows by similar arguments as in Theorem 4.6.1 and Theorem 4.6.4 below. The later provides an explicit isomorphism between the moduli space $\mathcal{M}$ and $\mathcal{AF}(\Sigma)/\mathrm{Diff}_0(\Sigma)$.

Theorem 4.6.4. *Let $g \in Met(\Sigma)$ and $\sigma \in Q(g)$ satisfy the equations*

$$K_g + |\sigma|^2 = -1, \qquad \bar{\partial}\sigma = 0, \qquad |\sigma|_g < 1. \tag{4.108}$$

For every such pair we define an almost-Fuchsian metric by

$$g^Y = g^Y_{g,\sigma} = \begin{pmatrix} g\left(\cosh(t)\mathbb{1} - \sinh(t)g^{-1}Re(\sigma)\right)^2 & 0 \\ 0 & 1 \end{pmatrix}. \tag{4.109}$$

This is the unique almost-Fuchsian metric which restricts to g along $\Sigma \times \{0\}$ and such that $Re(\sigma)$ is the second fundamental form of $\Sigma \times \{0\} \subset Y$. In particular,

$$\mathcal{M} \xrightarrow{\cong} \mathcal{AF}(\Sigma)/Diff_0(\Sigma), \qquad [g,\sigma] \mapsto [g^Y_{g,\sigma}] \tag{4.110}$$

defines an isomorphism of the two moduli spaces.

Proof. We verify first that g^Y defined by (4.109) is indeed quasi-Fuchsian. It follows from Step 2 in the proof of Theorem 4.6.1 that $\mathrm{Re}(\sigma)$ is the second fundamental form of $\Sigma \times \{0\}$ in Y. Moreover, $\Sigma \times \{0\}$ is a minimal surface, since $g^{-1}\mathrm{Re}(\sigma) \in \Omega^0(\Sigma, \mathrm{End}(T\Sigma))$ has trace zero. Since $g^{-1}\mathrm{Re}(\sigma)$ is a traceless symmetric endomorphism, it is diagonalizable with eigenvalues $\pm\lambda := \pm\sqrt{-\det(g^{-1}h)}$. These are the principal curvatures of $\Sigma \times \{0\}$ and they satisfy

$$|\pm\lambda|^2 = |\det(g^{-1}\mathrm{Re}(\sigma))| = |\sigma|_g^2 < 1$$

by (4.107). This also implies that g^Y is positive definite. A direct calculation shows that g^Y satisfies (4.100). It then follows from Step 2 and Step 3 in the proof of Theorem 4.6.1 that g^Y is a hyperbolic metric and therefore almost-Fuchsian. Moreover, uniqueness of solutions to (4.100), shows that the metric g^Y is uniquely determined by the initial data $(g, \mathrm{Re}(\sigma))$.

Conversely, for every almost-Fuchsian metric g^Y, we can recover (g,σ) from the restriction of g^Y to $\Sigma \times \{0\}$ and its second-fundamental form. This proves bijectivity of (4.110) and concludes the proof of the theorem. □

Lemma 4.6.5. *Every almost-Fuchsian manifold $Y = (\Sigma \times \mathbb{R}, g^Y)$ contains a unique closed incompressible minimal surface, which is $\Sigma \times \{0\}$.*

Proof. By Theorem 4.6.4, we may assume that g^Y is given by (4.109). A direct calculation shows that the mean curvature along $\Sigma \times \{t\}$ is

$$H(z,t) = \frac{2\cosh(t)\sinh(t)(1 + |\sigma(z)|_g^2)}{\cosh(t)^2 + \sinh(t)^2|\sigma(z)|_g^2}.$$

As a vector, this points in positive t direction for $t > 0$ and in negative t direction for $t < 0$. Hence, by the maximum principle, there exists no bounded minimal surface in Y except $\Sigma \times \{0\}$. □

4.6.2 A Kähler potential and quasi-Fuchsian manifolds

This section begins with a brief recollection of well-known properties of hyperbolic space $\mathbb{H}^3$, quasi-Fuchisan groups and the simultaneous uniformization theorem of Bers. Classical references for this material are [110, 9].

Next, we describe work of Hodge [61] which gives rise to an explicit embedding $\mathcal{M} \hookrightarrow \mathcal{T}(\Sigma) \times \overline{\mathcal{T}(\Sigma)}$ which is equivariant with respect to the natural action of the mapping class group and intertwines the second complex structure on $\mathcal{M}$ with the canonical complex structure on $\mathcal{T}(\Sigma) \times \overline{\mathcal{T}(\Sigma)}$. This map is not surjective and its image can be identified with the space of almost-Fuchsian manifolds.

Finally, we describe a Kähler potential for the hyperkähler metric on $\mathcal{M}$: The functional, which assigns to every almost-Fuchsian manifold the area of its unique minimal surface, is a Kähler potential with respect to the standard complex structure obtained from $\mathcal{T}(\Sigma) \times \overline{\mathcal{T}(\Sigma)}$.

Hyperbolic space and Kleinian groups

The upper half plane model. The upper half plane model of hyperbolic space is $\mathbb{H}^3 := \mathbb{C} \times \mathbb{R}_{>0}$ endowed with the hyperbolic metric

$$g^{\mathbb{H}^3}_{(z,y)}\left((\hat{z}_1, \hat{y}_1), (\hat{z}_2, \hat{y}_2)\right) = \frac{\mathrm{Re}(\hat{z}_1)\mathrm{Re}(\hat{z}_2) + \mathrm{Im}(\hat{z}_1)\mathrm{Im}(\hat{z}_2) + \hat{y}_1\hat{y}_2}{y^2}.$$

Identify $(z, y) \in \mathbb{H}^3$ with the quaternion $z_1 + \mathbf{i}z_2 + \mathbf{j}y + \mathbf{k} \cdot 0$ and define

$$\mathrm{SL}(2,\mathbb{C}) \times \mathbb{H}^3 \to \mathbb{H}^3, \qquad \begin{pmatrix} a & b \\ c & d \end{pmatrix}(z, y) := (a(z + \mathbf{j}y) + b)(c(z + \mathbf{j}y) + d)^{-1}.$$

One readily checks that this action is well-defined, preserves the hyperbolic metric, acts transitively on the unit disc bundle, and identifies the isometry group of $\mathbb{H}^3$ with $\mathrm{PSL}(2,\mathbb{C})$. The boundary at infinity $\partial_\infty\mathbb{H}^3$ can be identified with $(\mathbb{C} \times \{0\}) \cup \{\infty\} \cong S^2$. It follows from the explicit formula above that isometries on $\mathbb{H}^3$ correspond to conformal automorphism of the boundary. The induced action of $\mathrm{SL}(2,\mathbb{C})$ on the boundary is the standard action given by Möbius transformations.

Kleinian groups. A Kleinian group is a discrete subgroup $\Gamma < \mathrm{PSL}(2,\mathbb{C})$. The limit set $L_\Gamma \subset \partial_\infty\mathbb{H}^3$ of a Kleinian group Γ is defined as follows: Choose $p \in \mathbb{H}^3$ and denote its orbit by $\Gamma(p) \subset \mathbb{H}^3$. Then $L_\Gamma \subset \partial_\infty\mathbb{H}^3$ is the set of points which can be approximated in the euclidean topology of the closed ball $\mathbb{H}^3 \cup \partial_\infty\mathbb{H}^3$ by sequences contained in the orbit $\Gamma(p)$. One readily checks that this definition does not depend on the choice of p. The complement $\Omega_\Gamma := \partial_\infty\mathbb{H}^3 \backslash L_\Gamma$ is called the region of discontinuity. This is the largest open subset of the boundary on which Γ acts properly and discontinuously. The Ahlfors finiteness theorem asserts that for a finitely generated Kleinian group the quotient Ω_Γ/Γ is the disjoint union of finitely many Riemann surfaces with finitely many points removed. The hyperbolic manifold $Y := \mathbb{H}^3/\Gamma$ can thus be viewed as hyperbolic cobordism between these surfaces. We describe a simplest instance of this picture in the following.

Fuchsian and quasi-Fuchsian groups. A quasi-Fuchsian group is a Kleinian group $\Gamma < \mathrm{PSL}(2,\mathbb{C})$ whose limit set L_Γ is a Jordan curve and such that both components of its region of discontinuity $\Omega_\Gamma =: D_+ \cup D_-$ are preserved by Γ. For these groups Marsden [81] proved that $\mathbb{H}^3/\Gamma$ is diffeomorphic to $(D_+/\Gamma)\times\mathbb{R}$ and $(\mathbb{H}^3\cup\Omega)/\Gamma$ is diffeomorphic to $(D_+/\Gamma)\times[0,1]$. A quasi-Fuchsian manifold is a complete hyperbolic 3-manifold Y which is isometric to $\mathbb{H}^3/\Gamma$ for some quasi-Fuchsian group Γ. A Fuchsian group is a quasi-Fuchsian group Γ whose limit set L_Γ is a circle.

Every Fuchsian group is conjugated to a discrete subgroup of $\mathrm{PSL}(2,\mathbb{R})$ and thus determines a hyperbolic surface $(\Sigma, g) := \mathbb{H}^2/\Gamma$. A direct calculation shows that the Fuchsian hyperbolic 3-manifold $Y := \mathbb{H}^3/\Gamma$ is isometric to $\Sigma\times\mathbb{R}$ equipped with the metric

$$g^Y(z,t) = \begin{pmatrix} \cosh(t)^2 g(z) & 0 \\ 0 & 1 \end{pmatrix} \tag{4.111}$$

where $\Sigma := \mathbb{H}^2/\Gamma$ and $g \in \mathrm{Met}(\Sigma)$ is the induced hyperbolic metric.

It follows from Definition 4.6.3 that every Fuchsian manifold is almost-Fuchsian, and conversely, that every almost-Fuchsian manifold is quasi-isometric to a Fuchsian manifold. In particular, every almost-Fuchsian manifold is quasi-Fuchsian, since every quasi-isometry of $\mathbb{H}^3$ induces a continuous map on its boundary at infinity. The converse is not true: There are examples of quasi-Fuchsian manifolds which admit more then one minimal surface (see [121, 63, 57]) and these cannot be almost-Fuchsian by Lemma 4.6.5.

Simultaneous uniformization

An *odd coupled pair* is a triple $(\Sigma_-, [f], \Sigma_+)$ consisting of two closed Riemann surfaces $\Sigma_\pm$ and the homotopy class of an orientation reversing diffeomorphism $f : \Sigma_- \to \Sigma_+$. Two odd coupled pairs $(\Sigma_-, [f], \Sigma^+)$ and $(\tilde{\Sigma}_-, [\tilde{f}], \tilde{\Sigma}_+)$ are called equivalent if there exist biholomorphic maps $h_- : \tilde{\Sigma}_- \to \Sigma_-$ and $h_+ : \tilde{\Sigma}_+ \to \Sigma_+$ such that f is homotopic to $h_+ \circ \tilde{f} \circ h_-^{-1}$.

Now fix a closed oriented Riemann surface Σ. It is not hard to see that every odd coupled pair of Riemann surfaces of the same genus as Σ is isomorphic to a couple of the form

$$(\Sigma_-, [f], \Sigma^+) \sim ((\bar{\Sigma}, J_-), [\mathrm{id}], (\Sigma, J_+))$$

for some complex structures $J_- \in \mathcal{J}(\bar{\Sigma})$ and $J_+ \in \mathcal{J}(\Sigma)$. More precisely, this gives rise to an identification of the space of odd coupled pairs with the quotient $\mathcal{T}(\bar{\Sigma})\times\mathcal{T}(\Sigma)/\mathrm{MCG}(\Sigma)$ where $\mathrm{MCG}(\Sigma) = \mathrm{Diff}_+(\Sigma)/\mathrm{Diff}_0(\Sigma)$ denotes the mapping class group.

Every quasi-Fuchsian group $\Gamma < \mathrm{PSL}(2,\mathbb{C})$ gives rise to an odd coupled pair: The Riemann surfaces $\Sigma_\pm$ are the two connected components of Ω_Γ/Γ, where Ω_Γ denotes the region of discontinuity on the boundary sphere. Moreover, the fundamental groups $\pi_1(\Sigma_\pm)$ are canonically isomorphic to Γ and hence give rise to an isomorphism $\pi_1(\Sigma_-) \to \pi_1(\Sigma_+)$. This determines a unique homotopy class $[f]$ by the Dehn–Nielsen–Baer Theorem 4.4.12. The simultaneous uniformization theorem of Bers asserts that this constructions provides a bijection between the moduli space of quasi-Fuchsian groups and odd coupled pairs.

Theorem 4.6.6 (**Simultaneous uniformization**, Bers [8]). *Let $(\Sigma_-, [f], \Sigma_+)$ be an odd coupled pair with closed Riemann surfaces with genus$(\Sigma_\pm) \geq 2$. Then this pair is equivalent to one which can be represented by a quasi-Fuchsian group $\Gamma < PSL(2, \mathbb{C})$, which is uniquely determined up to conjugation.*

Denote by $\mathcal{QF}(\Sigma)$ the space of quasi-Fuchsian groups Γ which are isomorphic to $\pi(\Sigma)$. Then the theorem above asserts that

$$\frac{\mathcal{QF}(\Sigma)}{\text{conjugation}} \cong \frac{\mathcal{T}(\bar{\Sigma}) \times \mathcal{T}(\Sigma)}{\text{MCG}(\Sigma)}$$

where $\text{MCG}(\Sigma) = \text{Diff}_+(\Sigma)/\text{Diff}_0(\Sigma)$ denotes the mapping class group of Σ.

Embedding of the moduli space $\mathcal{M}$ into the quasi-Fuchsian moduli space

We present two maps from $\mathcal{M}$ into $\mathcal{T}(\Sigma)\times\overline{\mathcal{T}(\Sigma)}$. The following proposition is a rather direct consequence of Definition 4.6.3 and makes no claim about holomorphicity.

Proposition 4.6.7. *For an almost-Fuchsian metric*

$$g^Y = g^Y_{g,\sigma} = \begin{pmatrix} g\left(\cosh(t)\mathbb{1} - \sinh(t) g^{-1} Re(\sigma)\right)^2 & 0 \\ 0 & 1 \end{pmatrix} \in \mathcal{AF}(\Sigma)$$

define $g^\infty_\pm := g(1 + |\sigma|^2_g) \mp 2Re(\sigma)$ and let $J_\pm(g^Y) := J_{g^\infty_\pm} \in \mathcal{J}(\Sigma)$ be the unique complex structures compatible with $g^\infty_\pm$. Then

1. *$(\Sigma \times \mathbb{R}, g^Y)$ is isomorphic to the quasi-Fuchsian manifold which corresponds to the odd coupled pair $((\Sigma, J_+(g_Y)), (\Sigma, -J_-(g^Y)), [id_\Sigma])$.*
2. *The map $\mathcal{M} \cong \mathcal{AF}(\Sigma)/Diff_0(\Sigma) \hookrightarrow \mathcal{T}(\Sigma) \times \mathcal{T}(\Sigma)$ defined by*
$$[g, \sigma] \mapsto \left[J_+(g^Y_{g,\sigma}), J_-(g^Y_{g,\sigma})\right] \tag{4.112}$$
is a mapping class group equivariant embedding.

Proof. The metric $g_t := g\left(\cosh(t)\mathbb{1} - \sinh(t) g^{-1}\text{Re}(\sigma)\right)^2$ is conformally equivalent to $g\left(\mathbb{1} - \tanh(t) g^{-1}\text{Re}(\sigma)\right)^2$. For $t \to \pm\infty$, this tends to

$$g^\infty_\pm := g\left(\mathbb{1} + g^{-1}\text{Re}(\sigma)\right)^2 = g(1 + |\sigma|^2_g) \mp 2\text{Re}(\sigma)$$

where we used the relation $(g^{-1}\text{Re}(\sigma))^2 = |\sigma|^2_g \mathbb{1}$. This establishes the given formula and the proposition. □

One can understand the map (4.112) more explicitly on the level of sections. Denote by $X \subset T^*\mathbb{H}$ the unit disc bundle equipped with the hyperkähler structure from Theorem 4.5.1. Hodge [61] showed that there exists an $\text{SL}(2, \mathbb{R})$-equivariant diffeomorphism $\alpha : X \to \mathbb{H} \times \bar{\mathbb{H}}$ which intertwines the second complex structure on X with $(\mathbf{i}, -\mathbf{i})$ on $\mathbb{H} \times \bar{\mathbb{H}}$. It is explicitly given by the formula

$$\alpha(x + \mathbf{i}y, u + \mathbf{i}v) = \left(x - \frac{y^2 v}{1 - yu} + \mathbf{i}\frac{y\gamma}{1 - yu}, x + \frac{y^2 v}{1 + yu} + \mathbf{i}\frac{y\gamma}{1 + yu}\right). \tag{4.113}$$

where $\gamma := \sqrt{1 - y^2(u^2+v^2)}$ (see Remark 4.5.4). It gives rise to a $\mathrm{Diff}(\Sigma)$-equivariant bundle map $\alpha : \mathcal{Q}_1(\Sigma) \to \mathcal{J}(\Sigma) \times \overline{\mathcal{J}(\Sigma)}$ which descends to a mapping class group equivariant map

$$\mathcal{M} \cong \mathcal{M}_s \to \mathcal{T}(\Sigma) \times \overline{\mathcal{T}(\Sigma)}. \tag{4.114}$$

where $\mathcal{M}_s$ denotes the moduli space (4.86).

Proposition 4.6.8 (Hodge [61]). *The two maps (4.112) and (4.114) agree. Moreover, the second complex structure on $\mathcal{M}$ corresponds to $(\hat{J}_1, \hat{J}_2) \mapsto (-J_1\hat{J}_1, J_2\hat{J}_2)$ on $\mathcal{T}(\Sigma) \times \overline{\mathcal{T}(\Sigma)}$.*

Proof. Let $(J,\sigma) \in \mathcal{Q}_1(\Sigma)$ and denote by $g := \rho(\cdot, J\cdot)$ the induced Riemannian metric. Choose a holomorphic chart $\phi : U \to \Sigma$ and

$$\phi^*J = J_0, \qquad \phi^*\rho = \lambda^2 dx \wedge dy, \qquad \phi^*g = \lambda^2(dx^2+dy^2), \qquad \phi^*\sigma = \lambda(u - \mathbf{i}v)dz^2$$

for a smooth functions $u, v : U \to \mathbb{R}$ and $\lambda : U \subset \mathbb{R}^2 \to \mathbb{R}_+$. This chart defines a canonical trivialization of the $\mathrm{SL}(2,\mathbb{R})$-frame bundle define by the frames $\theta_z := \lambda^{-1}d\phi(z)$. With respect to this trivialization corresponds the pair $(\phi^*J, \phi^*\sigma)$ under the isomorphism (4.46) to the section $s^{\ell oc} := (\mathbf{i}, u + \mathbf{i}v) : U \to X$. By (4.113) we then have

$$\alpha(s^{\ell oc}) := \left(\frac{-v}{1-u} + \mathbf{i}\frac{\gamma}{1-u}, \frac{v}{1+u} + \mathbf{i}\frac{\gamma}{1+u}\right).$$

This corresponds to the two complex structures

$$J_+^{\ell oc} := j\left(\frac{-v}{1-u} + \mathbf{i}\frac{\gamma}{1-u}\right) = \frac{1}{\gamma}\begin{pmatrix} v & -(1+u) \\ 1-u & -v \end{pmatrix} \in \mathcal{J}(\mathbb{R}^2)$$

$$J_-^{\ell oc} := j\left(\frac{v}{1+u} + \mathbf{i}\frac{\gamma}{1+u}\right) = \frac{1}{\gamma}\begin{pmatrix} -v & 1-u \\ 1+u & v \end{pmatrix} \in \mathcal{J}(\mathbb{R}^2)$$

where $j : \mathbb{H} \to \mathcal{J}(\mathbb{R}^2)$ is defined (4.39). These are compatible with the metrics

$$g_+ := \begin{pmatrix} 0 & 2\lambda^2\gamma \\ -2\lambda^2\gamma & 0 \end{pmatrix} J_+^{\ell oc} = 2\lambda^2\begin{pmatrix} 1-u & -v \\ -v & 1+u \end{pmatrix} = 2(\phi^*g - \phi^*\mathrm{Re}(\sigma)).$$

$$g_- := \begin{pmatrix} 0 & 2\lambda^2\gamma \\ -2\lambda^2\gamma & 0 \end{pmatrix} J_-^{\ell oc} = 2\lambda^2\begin{pmatrix} 1+u & v \\ v & 1-u \end{pmatrix} = 2(\phi^*g + \phi^*\mathrm{Re}(\sigma)).$$

This shows that the complex structures (J_+, J_-) associated to (J,σ) under the maps α are determined by $g_\pm := 2(g \mp \mathrm{Re}(\sigma))$. Finally, define $\tilde{g} := (1 + \sqrt{1 - |\sigma|_g^2})$. A short calculation shows

$$g_\pm = 2(g \mp \mathrm{Re}(\sigma)) = \tilde{g}(1 + |\sigma|_{\tilde{g}}^2) \pm 2\mathrm{Re}(\sigma)$$

and hence $J_\pm$ agree with the complex structures defined in Proposition 4.6.7. $\square$

A Kähler potential for the hyperkähler metric

Consider the area functional on $\mathcal{M} \cong \mathcal{AF}(\Sigma)$ which assigns to every almost-Fuchsian manifold the area of its unique closed minimal surface.

$$A : \mathcal{M} \cong \mathcal{AF}(\Sigma)/\mathrm{Diff}_0 \to \mathbb{R}, \qquad A([g,\sigma]) := \mathrm{vol}(\Sigma, g) \tag{4.115}$$

where $\mathcal{M}$ is the moduli space (4.94). The second complex structure on $\mathcal{M}$ corresponds by Proposition 4.6.8 to the standard complex structure on $\mathcal{AF}(\Sigma)$ obtained from the embedding into $\mathcal{T}(\Sigma) \times \overline{\mathcal{T}(\Sigma)}$. The next theorem verifies a remark of Donaldson which asserts that the area functional (4.115) is a Kähler potential for the hyperkähler metric on $\mathcal{M}$ with respect to this complex structure. This has been confirmed by direct arguments along $\mathcal{T}(\Sigma) \subset \mathcal{M}$ in [55].

Theorem 4.6.9. *The area functional (4.115) provides a Kähler potential for the hyperbolic metric. More precisely*

$$2\boldsymbol{i}\bar{\partial}_{J_2}\partial_{J_2} A = \underline{\omega}_2. \tag{4.116}$$

Proof. On the moduli space $\mathcal{M}_d$, defined by (4.91), the area functional has the shape

$$A_d : \mathcal{M}_d \to \mathbb{R}, \qquad A([g,\sigma]) := \int_\Sigma \left(1 + \sqrt{1 - |\sigma|_g^2}\right) \mathrm{dvol}_g. \tag{4.117}$$

This follows from the identification $\mathcal{M} \cong \mathcal{M}_d$ in Proposition 4.5.17. In particular, on the original moduli space $\mathcal{M}_s$, defined by (4.84), one has

$$A_s : \mathcal{M}_s \to \mathbb{R}, \qquad A([J,\sigma]) := \int_\Sigma \left(1 + \sqrt{1 - |\sigma|_J^2}\right) \rho \tag{4.118}$$

where the norm $|\cdot|_J$ is defined using the metric $\rho(\cdot, J\cdot)$. Consider the S^1-action

$$S^1 \times \mathcal{M}_s \to \mathcal{M}_s, \qquad e^{\mathbf{i}t}[g,\sigma] = [g, e^{-\mathbf{i}t}\sigma].$$

It follows from Lemma 4.5.9 that A_s is a Hamiltonian function on $(\mathcal{M}_s, \underline{\omega}_1)$ which generates this S^1-action. Denote by $V_A \in \mathrm{Vect}(\mathcal{M}_s)$ the Hamiltonian vector field generated by A_s. Moreover, the S^1-action rotates $\underline{\omega}_2$, $\underline{\omega}_3$ and satisfies $\mathcal{L}_{V_A}\underline{\omega}_2 = -\underline{\omega}_3$ and $\mathcal{L}_{V_A}\underline{\omega}_3 = \underline{\omega}_2$. Hence the same formal calculation as in Lemma 4.5.9 yields

$$dA_s(J_2 W) = \underline{\omega}_1(V_A, J_2 W) = \langle J_1 V, J_2 W\rangle = \langle J_3 V, W\rangle = \underline{\omega}_3(V, W)$$

for all $W \in \mathrm{Vect}(\mathcal{M}_s)$ and therefore

$$2\mathbf{i}\bar{\partial}_{J_2}\partial_{J_2} H = d(dH \circ J_2) = d\iota(V_A)\underline{\omega}_3 = \mathcal{L}_{V_A}\underline{\omega}_3 = \underline{\omega}_2.$$

This proves (4.116) and the theorem. □

4.6.3 Embedding into the $\mathrm{PSL}(2,\mathbb{C})$ representation variety

Let g^Y by a hyperbolic metric on $Y := \Sigma \times \mathbb{R}$. The universal cover $\tilde{Y}$ of Y is isometric to hyperbolic space by the Cartan–Ambrose–Higgs theorem and there exists an isometry $\phi : \tilde{Y} \to \mathbb{H}^3$. The push-forward of the desk-transformation action of $\pi_1(\Sigma)$ on $\tilde{Y}$ yields then a representation $\rho : \pi_1(\Sigma) \to \mathrm{PSL}(2, \mathbb{C})$. Different choices of the isometry ϕ differ by an element of $\mathrm{PSL}(2, \mathbb{C})$ and lead to conjugated representations. We thus obtain a well-defined embedding

$$\mathcal{M} \cong \mathcal{AF}(\Sigma)/\mathrm{Diff}_0(\Sigma) \to \mathcal{R}_{\mathrm{PSL}(2,\mathbb{C})}(\Sigma) := \frac{\mathrm{Ham}(\pi(\Sigma), \mathrm{PSL}(2\mathbb{C}))}{\text{conjugation}}. \tag{4.119}$$

The image is an open subset in the smooth locus of the the representation variety $\mathcal{R}_{\mathrm{PSL}(2,\mathbb{C})}(\Sigma)$ which carries a natural holomorphic symplectic structure, see Goldman [52]. A classical result of Bers [9] asserts that the restriction of this complex structure to $\mathcal{M}$ corresponds to the standard complex structure on $\mathcal{T}(\Sigma) \times \overline{\mathcal{T}(\Sigma)}$ which differs by a sign from our conventions. In particular, it follows from Proposition 4.6.8 that the second complex structure on $\mathcal{M}$ corresponds to multiplication by $-\mathbf{i}$ on $\mathcal{R}_{\mathrm{PSL}(2,\mathbb{C})}(\Sigma)$.

Remark 4.6.10 (Holomorphic symplectic structure). The quasi-Fuchsian moduli space carries a natural holomorphic symplectic structure which can be expressed in complex Fenchel–Nielson coordinates and corresponds to the Goldman holomorphic symplectic structure on $\mathcal{R}_{\mathrm{PSL}(2,\mathbb{C})}(\Sigma)$, see [94, 52]. We have seen above that the underlying complex structure agrees with minus the second complex structure on $\mathcal{M}$. Moreover, the holomorphic symplectic form corresponds to $-\underline{\omega}_1 + \mathbf{i}\underline{\omega}_3$ on $\mathcal{M}$. This can be seen by noting that both symplectic forms agree (up to sign) with the Weil–Petersson symplectic form along Teichmüller space, which we embed diagonally into the quasi-Fuchisan moduli space using α. In then follows from holomorphicity that both forms agree on all of $\mathcal{M}$. See Hodge [61] for more details on this.

The Hitchin equation

We present in the following a construction of Donaldson which associates to every pair $[g, \sigma] \in \mathcal{M}$ a solution of Hitchin's equation. By classical results of Hitchin [58] and Donaldson [33], such solutions determines a flat $\mathrm{SL}(2, \mathbb{C})$ connection together with a harmonic map of the universal cover $\tilde{\Sigma}$ into $\mathbb{H}^3 = \mathrm{SL}(2, \mathbb{C})/\mathrm{SU}(2)$. This gives rise to an alternative description of the embedding of $\mathcal{M}$ into $\mathcal{R}_{\mathrm{PSL}(2,\mathbb{C})}(\Sigma)$.

Let $g \in \mathrm{Met}(\Sigma)$ and $\sigma \in Q(g)$ be given. Choose a holomorphic line bundle $L \to \Sigma$ with $L^2 = T\Sigma$ and define $E = L \oplus L^{-1}$. The Levi-Civita connection for g induces a unique $U(1)$-connection $a \in \mathcal{A}(L)$. Then consider the pair

$$A = \begin{pmatrix} a & \frac{\bar{\sigma}}{2} \\ -\frac{\sigma}{2} & -a \end{pmatrix} \in \mathcal{A}(E) \quad \text{and} \quad \phi = \frac{1}{2} \begin{pmatrix} 0 & \mathbf{1} \\ 0 & 0 \end{pmatrix} \in \Omega^{1,0}(\mathrm{End}(E)) \tag{4.120}$$

where

$$\begin{aligned} \sigma &\in Q(J) = \Omega^{1,0}(L^{-2}) = \Omega^{1,0}(\mathrm{Hom}(L, L^{-1})) \\ \bar{\sigma} &\in \overline{Q(J)} = \Omega^{0,1}(L^2) = \Omega^{0,1}(\mathrm{Hom}(L^{-1}, L)) \\ \mathbf{1} &\in \Omega^0(\mathrm{End}(T\Sigma)) = \Omega^{1,0}(L^2) = \Omega^{1,0}(\mathrm{Hom}(L^{-1}, L)). \end{aligned}$$

The adjoint section ϕ^* is given by

$$\phi^* = \frac{1}{2} \begin{pmatrix} 0 & 0 \\ \mathbf{1}^* & 0 \end{pmatrix} \in \Omega^{0,1}(\mathrm{End}(E))$$

where $\mathbf{1}^* = 2\mathbf{i}\mathrm{dvol}_g \in \Omega^2(\Sigma,\mathbb{C}) = \Omega^{0,1}(\Sigma, T^*\Sigma) = \Omega^{0,1}(\Sigma, \mathrm{Hom}(L, L^{-1}))$ and we used the sign convention $\Lambda^{1,1}(T^*\Sigma) \cong \Lambda^{0,1}(T^*\Sigma) \otimes \Lambda^{1,0}(T^*\Sigma)$. The next lemma asserts that (A,ϕ) satisfies the Hitchin equations and yields a flat $\mathrm{SL}(2,\mathbb{C})$ connection.

Lemma 4.6.11. *Consider the setup described above. The pair (g,σ) satisfies (4.95) if and only if (A,ϕ) satisfies the Hitchin equations*

$$F_A + [\phi \wedge \phi^*] = 0, \qquad \bar{\partial}_A \phi = 0. \tag{4.121}$$

Moreover, if these conditions are satisfied, then $B := A + \phi + \phi^*$ is a flat $SL(2,\mathbb{C})$ connection.

Proof. The induced curvature form on L and L^{-1} are $\frac{\mathbf{i}}{2}K_g\mathrm{vol}_g$ and $-\frac{\mathbf{i}}{2}K_g\mathrm{vol}_g$ where K_g denote the curvature form. Moreover, σ yields a covariant constant section of $\Omega^{1,0}(L^{-2})$ since $\bar{\partial}\sigma = 0$ as quadratic differential and both connections are induced by the Levi-Civita connection. By the formula in Lemma 2.7.6, it follows

$$F_A = \begin{pmatrix} K + |\sigma|_g^2 & 0 \\ 0 & -K - |\sigma|_g^2 \end{pmatrix} \frac{\mathrm{dvol}_g}{2\mathbf{i}}$$

Moreover,

$$\left[\begin{pmatrix} 0 & \mathbf{1} \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ \mathbf{1}^* & 0 \end{pmatrix}\right] = \begin{pmatrix} \mathbf{1}\wedge\mathbf{1}^* & 0 \\ 0 & -\mathbf{1}^*\wedge\mathbf{1} \end{pmatrix} = \begin{pmatrix} -2\mathbf{i} & 0 \\ 0 & 2\mathbf{i} \end{pmatrix} \mathrm{dvol}_g$$

and hence

$$F_A + [\phi\wedge\phi^*] = \begin{pmatrix} K_g + 1 + |\sigma|_g^2 & 0 \\ 0 & -(K_g + 1 + |\sigma|^2) \end{pmatrix} \frac{1}{2\mathbf{i}} \mathrm{dovl}_g.$$

This proves the first part of the lemma. Moreover,

$$F_B = F_A + d_A(\phi + \phi^*) + \frac{1}{2}[(\phi+\phi^*)\wedge(\phi+\phi^*)] = F_A + [\phi\wedge\phi^*]$$

shows that $B = A + \phi + \phi^*$ is a flat $\mathrm{SL}(2,\mathbb{C})$ connection when $F_A + [\phi\wedge\phi^*] = 0$. □

The holonomy representation $\rho_{A,\phi} : \pi_1(\Sigma) \to \mathrm{SL}(2,\mathbb{C})$ of the flat connection $B := A + \phi + \phi^*$ is well-defined up to conjugation and therefore Lemma 4.6.11 yields again an embedding of $\mathcal{M}$ into $\mathcal{R}_{\mathrm{PSL}(2,\mathbb{C})}(\Sigma)$.The connection between hyperbolic 3-manifolds and Hitchin's equation was observed by Donaldson [33]. For this consider the following model of hyperbolic space

$$\mathbb{H}^3 = \mathrm{SL}(2,\mathbb{C})/\mathrm{SU}(2), \qquad \langle d\pi(g)g\mathbf{i}\xi, d\pi(g)g\mathbf{i}\eta\rangle := -2\mathrm{tr}(\xi\eta)$$

for $g \in \mathrm{SL}(2,\mathbb{C})$ and $\xi,\eta \in \mathrm{su}(2)$ where $\pi : \mathrm{SL}(2,\mathbb{C}) \to \mathrm{SL}(2,\mathbb{C})/\mathrm{SU}(2)$ denotes the canonical projection. Let P^c and P be the $\mathrm{SL}(2,\mathbb{C})$ and $\mathrm{SU}(2)$ frame bundle of $E = L \oplus L^{-1}$. Then B induces a flat connection on the $\mathbb{H}^3$-bundle

$$P(\mathbb{H}^3) := P^c \times_{\mathrm{SL}(2,\mathbb{C})} (\mathrm{SL}(2,\mathbb{C})/\mathrm{SU}(2)) = P^c/\mathrm{SU}(2)$$

and the reduction $P \subset P^c$ gives rise to a section $s_{A,\phi} \in \Omega^0(\Sigma, P(\mathbb{H}^3))$.

Theorem 4.6.12. *Suppose (g,σ) satisfies (4.95) and let (A,ϕ) be the corresponding solution of Hitchin's equation (see Lemma 4.6.11). Let $s_{A,\phi} \in \Omega^0(\Sigma, P(\mathbb{H}^3)$ be the corresponding section of the associated $\mathbb{H}^3$-bundle as described above. Finally, denote by $(\tilde{\Sigma},\tilde{g},\tilde{\sigma})$ be the universal cover of Σ equipped with the lifted Riemannian metric $\tilde{g}$ and quadratic differential $\tilde{\sigma}$. Then the following holds.*

1. *$s_{A,\phi}$ lifts to a $\pi_1(\Sigma)$-equivariant isometric immersion $\tilde{s}_{A,\phi} : (\tilde{\Sigma},\tilde{g}) \to \mathbb{H}^3$ and the second fundamental form of $\tilde{s}_{A,\phi}$ is given by Re$(\tilde{\sigma})$.*

2. *The holonomy representation $\rho_B : \pi_1(\Sigma) \to SL(2,\mathbb{C})$ of the flat connection $B := A+\phi+\phi^*$ agrees up to conjugation with the image of $[g,\sigma]$ under (4.119). In particular, $Y := \mathbb{H}^3/\rho_B$ is a smooth almost-Fuchsian manifold and $s_{A,\phi}$ defines a minimal isometric embedding $(\Sigma, g) \hookrightarrow Y$ with second fundamental form Re(σ).*

Proof. We recall some of the key observations of Donaldson [33]: First, the canonical isomorphism

$$\mathbf{i}\mathrm{ad}(P) \cong s^*_{A,\phi}(T^{vert}(P(\mathbb{H}^3))) \tag{4.122}$$

intertwines the connection induced by A on $\mathbf{i}\mathrm{ad}(P)$ and the connection on $(T^{vert}(P(\mathbb{H}^3))$ induced by the flat connection $B := A+\phi+\phi^*$ and the Levi-Civita connection of the hyperbolic metric on $\mathbb{H}^3$. Second, the associated section $s_{A,\phi}$ satisfies

$$\nabla s = (\phi+\phi^*) \in \Omega^0(\Sigma, \mathbf{i}\mathrm{ad}(P)) \subset \Omega^0(\Sigma, \mathrm{End}(E))$$

where we identify $\mathbf{i}\mathrm{ad}(P)$ with the space of self-adjoint endomorphism of E. By (4.121), it follows $d_A^*(\phi+\phi^*) = 0$, and this is equivalent to $\nabla^*\nabla s = 0$. Solutions to the later equation are called twisted harmonic sections – they are represented in any flat trivialization by harmonic maps into $\mathbb{H}^3$.

After this preliminary discussion, we can proceed to the proof of the theorem. It suffices to verify the first part locally. Let $U \subset \Sigma$ be a contractible holomorphic coordinate chart and suppose $g = \lambda^2(dx^2+dy^2)$ in these coordinates. This chart provides a trivialization in $T\Sigma = L^2$ along U and we choose compatible trivializations of L and L^{-1}. These trivializations are not unitary, and the bundle metric is given by $\lambda \oplus \lambda^{-1}$. In this trivialization, the section $s_{A,\phi}$ is represented by a map $s : U \to \mathbb{H}^3$. Moreover,

$$ds(v) = \frac{1}{2}\begin{pmatrix} 0 & v \\ \lambda^2\bar{v} & 0 \end{pmatrix} \in \mathrm{End}(\mathbb{C}^2)$$

when $ds(v)$ is viewed as section of $\mathbf{i}\mathrm{ad}(P) \subset \mathrm{End}(E)$ and

$$ds(v) = L_s\left[\frac{1}{2}\begin{pmatrix} 0 & \lambda v \\ \lambda\bar{v} & 0 \end{pmatrix}\right] \in T_s\mathbb{H}^3$$

when $ds(v)$ as section of $s^*T\mathbb{H}^3$, where $L_p : \mathfrak{sl}(2,\mathbb{C}) \to T_p\mathbb{H}^3$ is defined by $L_p\xi := \partial_t|_{t=0}pe^{t\xi}$. In particular, $|ds(v)|^2 = \lambda^2|v|^2$ shows that s is an isometric immersion. We calculate in the same chart

$$\nabla_u(ds(v)) = [A(u), ds(v)] = \frac{1}{2}\begin{pmatrix} \frac{1}{2}(\bar{\sigma}(u,v)+\sigma(u,v)) & a(u)v+va(u) \\ -\lambda^2(a(u)\bar{v}+\bar{v}a(u)) & -\frac{1}{2}(\bar{\sigma}(u,v)+\sigma(u,v)) \end{pmatrix}$$

for vector fields $u, v : U \to \mathbb{C}$. It follows from the formula for $ds(v)$ above that

$$\nu(s) := \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{pmatrix}.$$

corresponds to the unit normal vector field along the image of s. Hence its second fundamental form is given by $\mathrm{Re}(\sigma)$ and this completes the proof of the first part.

By Theorem 4.6.4 there exists a unique quasi-Fuchisan metric $g^Y \in \mathcal{AF}(\Sigma)$ on $Y := \Sigma \times \mathbb{R}$ for which $\Sigma \times \{0\}$ is a minimal surface with induced metric g and second fundamental form $\mathrm{Re}(\Sigma)$. This lifts a hyperbolic metric on $\tilde{Y} := \tilde{\Sigma} \times \mathbb{R}$. It follows from Step 2 and Step 3 in the proof of Theorem 4.6.1 that

$$\tilde{Y} \to \mathbb{H}^3, \qquad (z,t) \mapsto \exp_{\tilde{s}_{A,\phi}(z)}(t\nu(\tilde{s}_{A,\phi}(z)))$$

is a $\pi_1(\Sigma)$-equivariant isometry. This proves the second part and the theorem. □

4.6.4 The cotangent bundle of Teichmüller space

We identify the cotangent bundle of Teichmüller space with

$$T^*\mathcal{T}(\Sigma) := \{(J,\sigma) \,|\, J \in \mathcal{J}(\Sigma),\, \sigma \in Q(J),\, \bar{\partial}_J \sigma = 0\}/\mathrm{Diff}_0(\Sigma).$$

Remark 4.6.13. Recall that we chose the complex structure on Teichmüller space to be $\hat{J} \mapsto -J\hat{J}$ and with this complex structure it would be more natural to identify $\mathcal{Q}(\Sigma)/\mathrm{Diff}_0(\Sigma)$ with the tangent space of Teichmüller space, see Remark 4.3.3. To obtain nevertheless an identification with the cotangent bundle, we need to define the complex structure on $\mathcal{Q}(\Sigma)$ by $(\hat{J}, \hat{\sigma}) \mapsto (-J\hat{J}, -\mathbf{i}\hat{\sigma})$ which is consistent with the first complex structure on $\mathcal{M}$.

The next theorem is a special case of a result due to Uhlenbeck ([117], Theorem 4.4). It shows that $\mathcal{M}$ admits a natural embedding into $T^*\mathcal{T}(\Sigma)$.

Theorem 4.6.14. *Let $\mathcal{M}$ be the moduli space (4.94). For $g \in Met(\Sigma)$ denote by $J_g \in \mathcal{J}(\Sigma)$ the unique complex structure compatible with g and the orientation of Σ. Then*

$$\mathcal{M} \to T^*\mathcal{T}(\Sigma), \qquad [g,\sigma] \mapsto [J_g, \sigma] \tag{4.123}$$

is a smooth embedding.

Remark 4.6.15. The theorem does not hold without the restriction $|\sigma|_g < 1$, see [117, 62].

Remark 4.6.16. The hyperkähler structure of $\mathcal{M}$ along the image can be viewed as the Feix–Kaledin hyperkähler extension [43, 67] of the Weil–Petersson metric along Teichmüller space.

Proof. Let $J \in \mathcal{J}(\Sigma)$ and $\sigma \in Q(J)$. We need to show that there exists a unique metric g in the conformal class determined by J with $|\sigma|_g < 1$ and

$$K_g + |\sigma|_g^2 = -1 \tag{4.124}$$

By uniformization, there exists a unique hyperbolic metric $g_0 \in \mathrm{Met}(\Sigma)$ which is compatible with J. Every other metric in the conformal class of g_0 has the shape $g = e^{2u}g_0$ for some smooth function $u : \Sigma \to \mathbb{R}$.

Step 1: *$g := e^{2u}g_0 \in Met(\Sigma)$ solves (4.124) if and only if u solves*

$$\Delta_{g_0} u - 1 + e^{2u} + |\sigma|^2_{g_0} e^{-2u} = 0. \tag{4.125}$$

where $\Delta_{g_0} = d^*d$ denotes the positive Laplacian.

The Gaussian curvature changes as $K_g = e^{-2u}(\Delta_{g_0} u - 1)$ and the norm of σ changes by $|\sigma|^2_g = |\sigma|^2_{g_0} e^{-4u}$. Hence

$$K_g + |\sigma|^2_g + 1 = e^{-2u}\left(\Delta_{g_0} u - 1 + e^{2u} + |\sigma|^2_{g_0} e^{-2u}\right)$$

and this proves Step 1.

Step 2: *Fix $k \geq 2$ and define $F : W^{k,2}(\Sigma, \mathbb{R}) \to W^{k-2,2}(\Sigma, \mathbb{R})$ by*

$$F(u) := \Delta_{g_0} u - 1 + e^{2u} + |\sigma|^2_{g_0} e^{-2u} \tag{4.126}$$

Suppose $|\sigma|_g < 1$ pointwise, then $L_u := dF(u) : W^{k,2}(\Sigma, \mathbb{R}) \to W^{k-2,2}(\Sigma, \mathbb{R})$ is given by

$$L_u\xi := \Delta_{g_0}\xi + 2e^{2u}\xi - 2|\sigma|^2_{g_0} e^{-2u}\xi. \tag{4.127}$$

and this is a positive self-adjoint isomorphism.

The formula for the derivative is immediate. We then calculate

$$\begin{aligned}
\langle L_u\xi, \xi\rangle_{L^2} &= \int_\Sigma \left(|d\xi|^2_{g_0} + 2e^{2u}\xi^2 - 2|\sigma|^2_{g_0} e^{-2u}\xi^2\right) \mathrm{dvol}_{g_0} \\
&= \int_\Sigma \left(|d\xi|^2_g + 2\xi^2 - 2|\sigma|^2_g\xi^2\right) \mathrm{dvol}_g \\
&= \int_\Sigma \left(|d\xi|^2_g + 2(1 - |\sigma|^2_g)\xi^2\right) \mathrm{dvol}_g
\end{aligned}$$

This is strictly positive for $\xi \neq 0$ and hence L_u is injective. Since L_u is a lower order pertubation of the Laplacian Δ_{g_0}, it is a Fredholm operator of index 0, and therefore also surjective.

Step 3: *Let $g \in Met_V(\Sigma)$, $\sigma \in Q(g)$ with $|\sigma|_g < 1$ and suppose (g, σ) satisfies (4.124). Then there exists a unique smooth path $u : [0,1] \to W^{k,2}(\Sigma, \mathbb{R})$, $t \mapsto u_t$, such that*

$$\Delta_{g_0} u_t - 1 + e^{2u_t} + |t\sigma|^2_{g_0} e^{-2u_t} = 0 \tag{4.128}$$

for all $t \in [0,1]$ and $g = g_0 e^{2u_1}$.

First, let $0 \leq t_0 < 1$ and suppose that $u_t \in W^{k,2}(\Sigma, \mathbb{R})$ is a smooth family of functions satisfying (4.128) for $t \in (t_0, 1]$. We claim that

$$\partial_t |\sigma|^2_{g_t} \geq 0 \qquad \text{for all } t \in (t_0, 1]. \tag{4.129}$$

Indeed, differentiating the equation yields

$$L_{u_t} \dot{u}_t + 2t|\sigma|^2_{g_0} e^{-2u_t} = 0$$

where L_{u_t} is a positive elliptic operator by Step 2, provided that $|t\sigma|^2_{g_t} < 1$. In this case, it follows from the maximums principle that $\dot{u}_t < 0$ and then $\partial_t |\sigma|^2_{g_t} = \partial_t \left(|\sigma|^2_{g_0} e^{-4u_t} \right) > 0$. Therefore the set of times $t \in (t_0, 1]$ for which (4.129) holds is open, closed and contains 1. It follows that (4.129) is satisfied for all $t \in (t_0, 1]$

Next, consider the joint function $G : W^{k,2}(\Sigma, \mathbb{R}) \times \mathbb{R} \to W^{k-2,2}(\Sigma, \mathbb{R})$ defined by

$$G(u, t) = \Delta_{g_0} u - 1 + e^{2u} + |t\sigma|^2_{g_0} e^{-2u}.$$

We need to show that there exists a unique family u_t satisfying $G(u_t, t) = 0$ for all $t \in [0, 1]$ and $g = g_0 e^{2u_1}$. By Step 2, we can apply the inverse function theorem at $G(u_t, t)$ if $|t\sigma|^2_{g_t} < 1$ for $g_t := g_0 e^{2u_t}$. For $t = 1$ this is satisfied by assumption, and the solution exists on some interval $(t_0, 1]$. Moreover, it follows from (4.129) that the condition $|t\sigma|^2_{g_t} < 1$ remains satisfied for all $t \in (t_0, 1]$. This yields uniqueness of the solution and openness of the maximal existence interval. It remains to show that u_t converges as $t \to t_0$. The estimate in Step 2, shows that the family of operators $L_{u_t} : W^{2,2}(\Sigma, \mathbb{R}) \to L^2(\Sigma, \mathbb{R})$ is uniformly bounded and hence

$$\dot{u}_t = L^{-1}_{u_t} \left(2t|\sigma|^2_{g_0} e^{-2u_t} \right), \qquad t \in (t_0, 1]$$

is uniformly bounded in $W^{2,2}(\Sigma, \mathbb{R})$. Then, by elliptic regularity, $\dot{u}_t$ is also uniformly bounded in $W^{k,p}(\Sigma, \mathbb{R})$ and therefore u_t converges as $t \to t_0$.

Step 4: *The inclusion (4.123) is an embedding.*

Let $g \in \mathrm{Met}_V(\Sigma)$, $\sigma \in Q(g)$ with $|\sigma|_g < 1$ be given and suppose (g, σ) satisfies (4.124). By Step 3 there exists a unique path $u : [0, 1] \to W^{k,2}(\Sigma, \mathbb{R})$ satisfying

$$K_{g_t} + |t\sigma|^2_{g_t} = -1, \qquad g_t := g_0 e^{-2u_t}.$$

For $t = 0$, the maximum principle yields that $u_0 \equiv 0$. We may thus recover the metric $g = g_1$ by following the path of solutions defined $G(u_t, t) = 0$. This shows uniqueness of solutions within the conformal class under the constraint $|\sigma|_g < 1$ and this proves the theorem. $\square$

Chapter 5

Moduli spaces of holomorphic differentials over Riemann surfaces

Donaldson introduced in [38] a general moment map framework for the action of the diffeomorphism group on the space of sections for certain symplectic fibrations. He then applied this framework to construct Teichmüller space and a hyperkähler extension parametrizing complex structures together with holomorphic quadratic differentials. We generalize his construction in this chapter to holomorphic differentials of arbitrary degree and tuples of holomorphic differentials of mixed degree. These moduli spaces are closely related to Hitchin's higher Teichmüller components [59]. We hope that this might lead to a new construction of the Hitchin component using the diffeomorphism group instead of the gauge group.

5.1 Introduction

Let (Σ, ρ) be a closed 2-manifold with fixed area form $\rho \in \Omega^2(\Sigma)$ and assume $\mathrm{genus}(\Sigma) \geq 2$. Let $P \to \Sigma$ denote its $\mathrm{SL}(2,\mathbb{R})$ frame bundle. The unit disc bundle $X_k \subset (T^*\mathbb{H})^{k/2}$ can be identified with pairs (J, γ) consisting of a linear complex structure $J \in \mathcal{J}(\mathbb{R}^2)$ and a symmetric J-complex multilinear form $\gamma : (\mathbb{R}^2, J)^k \to \mathbb{C}$ with $|\gamma| < 1$. The space of sections of the associated bundle $P \times_{\mathrm{SL}(2,\mathbb{R})} X_k$ then admits a natural identification with the space

$$\mathcal{D}^1_k(\Sigma) := \left\{ (J, \tau) \mid J \in \mathcal{J}(\Sigma),\, \tau \in S^k(T^*\Sigma \otimes_J \mathbb{C}),\, |\tau| < 1 \right\} \tag{5.1}$$

which parametrizes complex structures and complex differentials of order k.

The fibre $X_k \subset (T^*\mathbb{H})^{k/2}$ carries in the case $k = 2$ a natural symplectic form coming from the Feix–Kaledin hyperkähler extension of the hyperbolic metric on $\mathbb{H}$. For $k > 2$, we do not expect that there exists a hyperkähler setup and obtain instead a family of symplectic forms on X_k parametrized by a single functions $f : [0, 1) \to [0, 1)$

with $f(0)=0$ and $f'>0$: There exists a unique $\mathrm{SL}(2,\mathbb{R})$-invariant symplectic form $\omega_f \in \Omega^2(X_k)$ satisfying

$$\begin{aligned}\omega_f(\mathbf{i},w) = &-\frac{\mathbf{i}}{2}\left(1-f(|w|^2)+k|w|^2f'(|w|^2)\right)\,d\bar{z}\wedge dz \\ &-\frac{2\mathbf{i}}{k}f'(|w|^2)\,d\bar{w}\wedge dw + f'(|w|^2)\left(\bar{w}d\bar{z}dw - wd\bar{w}dz\right).\end{aligned} \tag{5.2}$$

Here is a more geometric description of these forms: (1) The symplectic connections of ω_f yields the standard connection on $(T^*\mathbb{H})^{k/2}$ obtained from the Levi–Civita connection on the hyperbolic plane and (2) the S^1 action which rotates the fibres is Hamiltonian with $H(z,w) = -\frac{2}{k}f\left(\mathrm{Im}(z)^k|w|^2\right)$ and the Marsden-Weinstein quotient $H^{-1}\left(-\frac{2}{k}r^2\right)/S^1$ is symplectomorphic to the hyperbolic plane scaled by $(1-f(r^2))$. None of these symplectic forms extends over the whole space $(T^*\mathbb{H})^{k/2}$ and the restriction to a disc bundle is necessary. The general framework introduced by Donaldson then yields the following:

Theorem A. *The action of Ham(Σ,ρ) on the space $\{(J,\tau)\in\mathcal{D}^1_k(\Sigma)\,|\,\bar{\partial}_J\tau=0\}$ is Hamiltonian with respect to $\underline{\omega}_f$ and with moment map*

$$\underline{\mu}_f(J,\tau) = \left(2K_J+\Delta F(|\tau|^2)-2c\right)\rho \tag{5.3}$$

where $c := 2\pi(2-\text{genus}(\Sigma))/\text{vol}(\Sigma,\rho)$, $F:[0,1)\to\mathbb{R}$ is defined by

$$F(t) := \int_0^t \frac{f(s)}{ks} - f'(s)\,ds$$

and $\Delta = d^*d$ is the positive Laplacian of the metric $\rho(\cdot,J\cdot)$.

Proof. See Theorem 5.2.9. □

This theorem is the key step in showing that the moduli space

$$\mathcal{M}_f(k) := \left\{(J,\gamma)\in\mathcal{D}^1_k(\Sigma)\;\middle|\;\bar{\partial}\gamma=0,\quad K_J+\frac{1}{2}\Delta F(|\gamma|^2)=c\right\}\Big/\mathrm{Symp}_0(\Sigma,\rho)$$

carries a natural symplectic form. It is natural to ask the following:

1. Does there exists a preferred symplectic structure $\omega_f\in\Omega^2(X_k)$?
2. Are there symplectic structure ω_f for which $\mathcal{M}_f(k)$ admits a more concrete or geometric description?

In the case $k=2$ both questions are answered by the Feix–Kaledin hyperkähler metric. For $k>2$ we were unable to find a satisfactory answer to the first question. However, we found the following answer to the second question:

Theorem B. *There exists a unique monotone increasing function $F_k:[0,1]\to\mathbb{R}$ which satisfies $F_k(0)=\log((k-1)/k)$, $F_k(1)=0$ and $te^{-kF_k(t)}-ke^{-F_k(t)}+(k-1)=0$. Define $f_k:[0,1)\to[0,1)$ by*

$$f_k(t) := -\left(\int_0^t u_k'(s)s^{-1/k}\,ds\right)t^{1/k}. \tag{5.4}$$

Then $(J,\tau) \mapsto (e^{F_k(|\tau|_J^2)}\rho(\cdot, J\cdot), \tau)$ *induces an isomorphism*

$$\mathcal{M}_{f_k}(k) \cong \left\{ (g,\tau) \;\middle|\; \begin{array}{c} g \in Met(\Sigma),\ (J_g,\tau) \in \mathcal{D}^1_k(\Sigma) \\ \bar\partial\tau = 0, K_g - \frac{c}{k}|\tau|_g^2 = c\frac{k-1}{k} \end{array} \right\} \Big/ Diff_0(\Sigma) \tag{5.5}$$

where $c := 2\pi(2 - 2genus(\Sigma))/vol(\Sigma,\rho)$.

Proof. See Theorem 5.2.10. □

In the case $k = 2$ this yields $f_2(r) = 1 - \sqrt{1-r}$ which corresponds again to the Feix–Kaledin hyperkähler metric.

The discussion so far extends rather directly to tuples of complex differentials of mixed order: Let $\mathbf{k} = (k_1, \ldots, k_n) \in \mathbb{Z}$ with $2 \le k_1 \le \cdots \le k_n$ and define

$$\mathcal{D}^1_{\mathbf{k}}(\Sigma) := \left\{ (J, \tau_1, \ldots, \tau_n) \;\middle|\; J \in \mathcal{J}(\Sigma),\ \tau_i \in S^{k_i}(T^*\Sigma \otimes_J \mathbb{C}),\ |\tau_i| < 1 \right\} \tag{5.6}$$

A tuple $\mathbf{f} = (f_1, \ldots, f_n)$ of functions $f_i : [0,1) \to [0,1)$ with $f_i(0) = 0$ and $f_i' > 0$ and a weight vector $\alpha = (\alpha_1, \ldots, \alpha_n) \in (0,1)^n$ with $\sum_{i=1}^n \alpha_i = 1$ defines a symplectic form $\underline{\omega}^\alpha_{\mathbf{f}}$ on $\mathcal{D}^1_{\mathbf{k}}(\Sigma)$, which is obtained by combining the symplectic forms $\underline{\omega}_{f_i}$ on $\mathcal{D}_{k_i}(\Sigma)$ weighted with α. We then have the following

Theorem C. *The action of* $Ham(\Sigma,\rho)$ *on* $\{(J, \tau_1, \ldots, \tau_n) \in \mathcal{D}^1_k(\Sigma) \,|\, \bar\partial_J \tau_i = 0\}$ *is Hamiltonian with respect to* $\underline{\omega}^\alpha_{f}$ *and with moment map*

$$\underline{\mu}_{f}(J, \tau_1, \ldots, \tau_n) = \left(2K_J + \Delta\left(\sum_{i=1}^n \alpha_i F_i\left(|\tau_i|^2\right) \right) - 2c \right)\rho \tag{5.7}$$

where $c := 2\pi(2 - genus(\Sigma))/vol(\Sigma,\rho)$, $F_i : [0,1) \to \mathbb{R}$ *are defined by*

$$F_i(t) := \int_0^t \frac{f_i(t)}{kt} - f_i'(t)\, dt$$

and $\Delta = d^*d$ *is the positive Laplacian of the metric* $\rho(\cdot, J\cdot)$.

Proof. See Theorem 5.3.3. □

After taking the action of $\mathrm{Symp}_0(\Sigma,\rho)/\mathrm{Ham}(\Sigma,\rho)$, this construction gives rise to a symplectic form on the moduli space

$$\mathcal{M}_{\mathbf{f}}(\mathbf{k}) := \left\{ (J,\gamma) \in \mathcal{D}^1_k(\Sigma) \;\middle|\; \bar\partial\gamma = 0, \quad K_J + \frac{1}{2}\Delta\left(\sum_{i=1}^n \alpha_i F_i\left(|\tau_i|^2\right) \right) = c \right\} \Big/ \mathrm{Symp}_0(\Sigma,\rho)$$

The final section discusses the relation between these moduli spaces and Hitchin's higher Teichmüller components (see [59]). We also indicate some open questions and directions for further investigations.

5.2 Holomorphic differentials of order k

The first section introduces the space $\mathcal{D}_k(\mathbb{R}^2)$ of pairs (J, γ) where $J \in \mathcal{J}(\mathbb{R}^2)$ is a linear complex structure and $\gamma : (\mathbb{R}^2, J)^k \to \mathbb{C}$ is a multilinear form of order k. This is a line bundle over the space $\mathcal{J}(\mathbb{R}^2)$ and we investigate its total space both in the disc and upper half plane model of $\mathcal{J}(\Sigma)$.

The second section discusses the class of symplectic structures ω_f on the unit disc bundle $\mathcal{D}_k^1(\mathbb{R}^2) \subset \mathcal{D}_k(\mathbb{R}^2)$ and calculate the moment map for the natural $\mathrm{SL}(2, \mathbb{R})$ action.

In the third section, we then apply Donaldson's framework to this particular situation. This yields a moment map for the Hamiltonian action on the space of holomorphic differentials of order k and allows us to construct the moduli space $\mathcal{M}_f(k)$ which fibres over Teichmüller space.

5.2.1 Complex symmetric multilinear forms

Let $k \geq 2$ be a positive integer. For $J \in \mathcal{J}(\mathbb{R}^2)$ denote the space of J-complex symmetric multilinear forms of degree k by

$$\begin{aligned} D_k(J) : &= \left\{ \gamma : (\mathbb{R}^2)^k \to \mathbb{C} \;\middle|\; \begin{array}{c} \gamma \text{ is symmetric and} \\ (J, \mathbf{i})\text{-complex multilinear} \end{array} \right\} \\ &\cong \left\{ \gamma : \mathbb{R}^2 \to \mathbb{C} \;\middle|\; \begin{array}{c} \text{for all } \alpha, \beta \in \mathbb{R} \text{ and } v \in \mathbb{R}^2 \text{ it holds:} \\ \gamma(\alpha v + \beta J v) = (\alpha + \mathbf{i}\beta)^k \gamma(v) \end{array} \right\} \end{aligned}$$

Let h_J be the hermitian form on $(\mathbb{R}^2, J)$ determined by the standard area form on $\mathbb{R}^2$ and J, see (4.40). This induces on $D_k(J)$ the following hermitian structure

$$g_{D_k}(\gamma_1, \gamma_2) := \mathrm{Re}\left(\frac{\overline{\gamma}_1(v)\gamma_2(v)}{h_J(v,v)^k}\right), \qquad \omega_{D^k}(\gamma_1, \gamma_2) := \mathrm{Im}\left(\frac{\overline{\gamma}_1(v)\gamma_2(v)}{h_J(v,v)^k}\right)$$
$$J_{D_k}(\gamma)(v) = \mathbf{i}\gamma(v)$$

where none of these expressions depend on the choice of $v \in \mathbb{R}^2 \backslash \{0\}$. The spaces $D_k(J)$ form a hermitian line bundle $\mathcal{D}_k(\mathbb{R}^2) \to \mathcal{J}(\mathbb{R}^2)$ defined by

$$\mathcal{D}_k(\mathbb{R}^2) := \{(J, \gamma) \,|\, J \in \mathcal{J}(\mathbb{R}^2),\, \gamma \in D^k(J)\}$$

and the natural $\mathrm{SL}(2, \mathbb{R})$-action on $\mathcal{J}(\mathbb{R}^2)$ lifts to

$$\mathrm{SL}(2, \mathbb{R}) \times \mathcal{D}_k(\mathbb{R}^2) \to \mathcal{D}_k(\mathbb{R}^2), \qquad \Psi_*(J, \gamma) := (\Psi J \Psi^{-1}, (\Psi^{-1})^* \gamma).$$

The upper half-plane model

Denote by Y_k the space $\mathbb{H} \times \mathbb{C}$ equipped with the following $\mathrm{SL}(2, \mathbb{R})$ action

$$\mathrm{SL}(2, \mathbb{R}) \times Y_k \to Y_k, \qquad \begin{pmatrix} a & b \\ c & d \end{pmatrix}(z, w) := \left(\frac{az+b}{cz+d}, (cz+d)^k w\right).$$

We think of Y_k as $(T^*\mathbb{H})^{k/2}$ with the canonical complex structure and induced $\mathrm{SL}(2, \mathbb{R})$ action. The induced metric on the fibre is then given by

$$|w|_z^2 = \mathrm{Im}(z)^k |w|^2, \qquad \text{for } (z, w) \in Y_k.$$

Lemma 5.2.1. *Define* $\gamma_k : Y_k \to Hom((\mathbb{R}^2)^{\otimes k}, \mathbb{C})$ *by*

$$\gamma_k(z,w) : \mathbb{R}^2 \to \mathbb{C}, \qquad \gamma_k(v) \mapsto \bar{w}(v_1 - \bar{z}v_2)^k \tag{5.8}$$

and define $j : \mathbb{H} \to \mathcal{J}(\mathbb{R}^2)$ *by (4.39). Then the following holds:*

1. $\gamma_k(z,w) \in D_k(j(z))$ *for every* $(z,w) \in Y_k$.

2. *The fibre map* $\gamma_k(z,\cdot) : \mathbb{C} \to D_k(j(z))$ *is a complex anti-linear isometry for every* $z \in \mathbb{H}$.

3. *The bundle map* $(j,\gamma_k) : Y_k \to \mathcal{D}_k(\mathbb{R}^2)$ *is a* $SL(2,\mathbb{R})$*-equivariant bijection.*

Proof. For the first part, it suffices to check that

$$A(z) : \mathbb{R}^2 \to \mathbb{C}, \qquad A(z)v := v_1 - \bar{z}v_2$$

is $j(z)$-holomorphic. With $z = x + \mathbf{i}y$ this amounts to the calculation

$$A(z)j(z) = \frac{1}{y}(1, -\bar{z}) \begin{pmatrix} x & -x^2 - y^2 \\ 1 & -x \end{pmatrix} = \left(\frac{x - \bar{z}}{y}, \frac{-x^2 - y^2 + x\bar{z}}{y} \right) = \mathbf{i}A(z).$$

For the second part, let $(z,w) \in Y_k$ and denote $e_1 = (1,0) \in \mathbb{R}^2$. By (4.41) it follows $|h_{j(z)}(e_1,e_1)| = \mathrm{Im}(z)^{-1}$ and hence

$$|\gamma_k(z,w)|^2 = |\mathrm{Im}(z)|^k |\gamma_k(z,w)e_1|^2 = |\mathrm{Im}(z)|^k |w|^2 = |w|_z^2$$

This shows that the fibre map $\gamma_k(z,\cdot)$ is an isometry. For the third part let

$$\Psi = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2,\mathbb{R})$$

and $(z,w) \in Y_k$ be given. Then

$$\begin{aligned} \gamma_k(z,w)(\Psi^{-1}v) &= \overline{w}\left((dv_1 - bv_2) - \bar{z}(-cv_1 + av_2)\right)^k \\ &= \overline{w}\left((c + \bar{z}d)v_1 - (a\bar{z} + b)v_2\right)^k \\ &= (c\bar{z} + d)^k \bar{w} \left(v_1 - \frac{a\bar{z} + b}{c\bar{z} + d} v_2 \right)^k \\ &= \gamma_k(\Psi_*(z,w))(v). \end{aligned}$$

This proves equivariance of the bundle map (j,γ_k) and completes the proof of the lemma. □

The disc model

Denote by $Y_k^{\mathbb{D}}$ the space $\mathbb{D} \times \mathbb{C}$ equipped with the following SU(1,1)-action

$$\mathrm{SU}(1,1) \times Y_k^{\mathbb{D}} \to Y_k^{\mathbb{D}}, \qquad \begin{pmatrix} \alpha & \beta \\ \bar{\beta} & \bar{\alpha} \end{pmatrix} (z,w) := \left(\frac{\alpha z + \beta}{\bar{\beta} z + \bar{\alpha}}, (\bar{\beta} z + \bar{\alpha})^k w \right).$$

As in the upper half-plane model we think of $Y_k^{\mathbb{D}}$ as $(T^*\mathbb{D})^{k/2}$ with the canonical complex structure and induced $\mathrm{SL}(2,\mathbb{R})$ action. The induced metric on the fibre is then given by

$$|w|_z^2 = \left(\frac{1-|z|^2}{2}\right)^k |w|^2$$

for $(z,w) \in Y_k^{\mathbb{D}}$. Recall from Lemma 4.3.1 that the map $\mathrm{SL}(2,\mathbb{R}) \to \mathrm{SU}(1,1)$ defined by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \frac{1}{2}\begin{pmatrix} (a+d)+(b-c)\mathbf{i} & (a-d)-(b+c)\mathbf{i} \\ (a-d)+(b+c)\mathbf{i} & (a+d)-(b-c)\mathbf{i} \end{pmatrix} \tag{5.9}$$

is an isomorphism.

Lemma 5.2.2. *Let $\kappa := 2^{-k/2} \cdot e^{-k\pi \boldsymbol{i}/4}$ and define $F_k : Y_k \to Y_k^{\mathbb{D}}$ by*

$$F_k(z,w) := \left(\frac{z-\boldsymbol{i}}{z+\boldsymbol{i}},\ \kappa(z+\boldsymbol{i})^k w\right). \tag{5.10}$$

This map satisfies the following properties:

1. *F_k is biholomorph with inverse $G_k : Y_k^{\mathbb{D}} \to Y_k$ defined by*

$$G_k(\zeta,\eta) = \left(\boldsymbol{i}\frac{1+\zeta}{1-\zeta},\ \kappa\eta(1-\zeta)^k\right)$$

2. *F_k is $(SL(2,\mathbb{R}), SU(1,1))$-equivariant with respect to (5.9) and intertwines the Riemannian metrics along the fibres.*

Proof. A direct calculation, which we leave to the reader, shows that F_k and G_k are indeed inverse maps.

Next let $(z,w) \in Y_k$ and $(\zeta,\eta) := F_k(z,w) \in Y_k^{\mathbb{D}}$. Then

$$|\eta|_\zeta^2 = 2^{-2k}\left(1-\left|\frac{z-\mathbf{i}}{z+\mathbf{i}}\right|^2\right)^k |z+\mathbf{i}|^k |w|^2 = \mathrm{Im}(z)^k |w|^2 = |w|_z$$

shows that F_k intertwines the Riemannian structures along the fibres.

Finally, let $(z,w) \in Y_k$ and

$$\Psi = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2,\mathbb{R})$$

be given. Denote the image of Ψ under (5.9) by

$$\tilde{\Psi} := \begin{pmatrix} \alpha & \beta \\ \bar{\beta} & \bar{\alpha} \end{pmatrix} := \frac{1}{2}\begin{pmatrix} (a+d)+(b-c)\mathbf{i} & (a-d)-(b+c)\mathbf{i} \\ (a-d)+(b+c)\mathbf{i} & (a+d)-(b-c)\mathbf{i} \end{pmatrix} \in \mathrm{SU}(1,1).$$

We then compute

$$\begin{aligned} F_k(\Psi(z,w)) &= F_k\left(\frac{az+b}{cz+d}, (cz+d)^k w\right) \\ &= \left(\frac{(a-\mathbf{i}c)z+(b-\mathbf{i}d)}{(a+\mathbf{i}c)z+(b+\mathbf{i}d)}, \kappa\left((a+\mathbf{i}c)z+(b+\mathbf{i}d)\right)^k w\right) \end{aligned}$$

and

$$\begin{aligned}\tilde{\Psi}F_k(z,w) &= \tilde{\Psi}\left(\frac{z-\mathbf{i}}{z+\mathbf{i}}, \kappa(z+\mathbf{i})^k w\right)\\ &= \left(\frac{(\alpha+\beta)z+\mathbf{i}(\beta-\alpha)}{(\bar{\alpha}+\bar{\beta})z+\mathbf{i}(\bar{\alpha}-\bar{\beta})}, \kappa\left((\bar{\alpha}+\bar{\beta})z+\mathbf{i}(\bar{\alpha}-\bar{\beta})\right)^k w\right)\end{aligned}$$

Using $\bar{\alpha}+\bar{\beta}=a+\mathbf{i}c$ and $\mathbf{i}(\bar{\alpha}-\bar{\beta})=b+\mathbf{i}d$ it follows $F_k(\Psi(z,w))=\tilde{\Psi}F_k(z,w)$ and this proves equivariance. □

5.2.2 Symplectic structures and moment maps on the fibre

We consider the unit disc bundle in $(T^*\mathbb{H})^{k/2}$ as fibre. More precisely, define

$$X_k := \left\{(z,w)\in Y_k \;\middle|\; \mathrm{Im}(z)^k|w|^2<1\right\}. \tag{5.11}$$

Using Lemma 5.2.2 and the map (5.10), we can identify X_k with

$$X_k^{\mathbb{D}} := \left\{(z,w)\in Y_k^{\mathbb{D}} \;\middle|\; \left(\frac{1-|z|^2}{2}\right)^k |w|^2<1\right\} \tag{5.12}$$

which is the unit disc bundle in $(T^*\mathbb{D})^{k/2}$.

Symplectic structures

Extending Donaldson's construction of the Feix–Kaledin hyperkähler metric on the unit disc bundle in $T^*\mathbb{H}$, we obtain a class of $\mathrm{SL}(2,\mathbb{R})$, resp. $\mathrm{SU}(1,1)$, invariant symplectic forms on X_k and $X_k^{\mathbb{D}}$. As in the hypekähler case, these symplectic forms do not extend over the total space.

Lemma 5.2.3. *Let $f:[0,1)\to[0,1)$ be a smooth function with $f(0)=0$ and $f'(r)>0$ for all $r\in[0,1)$. Then there exists a unique $SU(1,1)$-invariant symplectic form $\omega_f^{\mathbb{D}}\in\Omega^2(X_k^{\mathbb{D}})$ satisfying*

$$\omega_f^{\mathbb{D}}(0,w) = -2\boldsymbol{i}(1-f(2^{-k}|w|^2))\,d\bar{z}\wedge dz - \frac{2\boldsymbol{i}}{2^k k}f'(2^{-k}|w|^2)\,d\bar{w}\wedge dw. \tag{5.13}$$

Moreover, f restricts to the hyperbolic area form along $\mathbb{D}\times\{0\}$.

Proof. Uniqueness of such a metric is follow from the fact that every $\mathrm{SU}(1,1)$-orbit contains an element of the form $(0,w)$. Moreover, if $\omega_f\in\Omega^2(X_k^{\mathbb{D}})$ is an $\mathrm{SU}(1,1)$-invariant 2-form satisfying (5.13), then the assumptions on f imply that ω_f is non-degenerate and restricts to the hyperbolic area form along $\mathbb{D}\times\{0\}$.

It remains to show the existence of a closed $\mathrm{SU}(1,1)$ invariant $\omega_f^{\mathbb{D}}\in\Omega^2(X_k^{\mathbb{D}})$ which satisfies (5.13). For this consider the open subset

$$\tilde{X}=\{(z_1,z_2)\in\mathbb{C}^2\,|\,0<|z_1|^2-|z_2|^2<2\}.$$

This is a k-fold covering of $X_k^{\mathbb{D}}\backslash(\mathbb{D}\times\{0\})$ with covering map

$$\pi:\tilde{X}\to X_k^{\mathbb{D}},\qquad \pi(z_1,z_2):=\left(\frac{z_2}{z_1}, z_1^k\right). \tag{5.14}$$

Now let $F : (0,2) \to \mathbb{R}$ be a smooth function and consider

$$\tilde{\omega} = \mathbf{i}\bar{\partial}\partial F(|z_1|^2 - |z_2|^2)$$

This is clearly SU$(1,1)$-invariant and a direct calculation shows it induces on $X_k^{\mathbb{D}} \backslash (\mathbb{D} \times \{0\})$ the following SU$(1,1)$-invariant metric

$$\begin{aligned}\omega_F^{\mathbb{D}}(z,w) = \mathbf{i}|w|^{2/k}\left(F''(\delta)|z|^2|w|^{2/k} - F'(\delta)\right) d\bar{z}\wedge dz \\ + \mathbf{i}\frac{\delta(F''(\delta)\delta + F'(\delta))}{k^2|w|^2}\, d\bar{w}\wedge dw \\ - \mathbf{i}\frac{|w|^{2/k}(F''(\delta)\delta + F'(\delta))}{k|w|^2}\,(z\bar{w}\, d\bar{z}dw + \bar{z}w\, d\bar{z}dw)\end{aligned}$$

where $\delta := (1-|z|^2)|w|^{2/k}$. On the fibre over $z=0$ this gives

$$\begin{aligned}\omega_F^{\mathbb{D}}(0,w) = -\mathbf{i}|w|^{2/k}F'(|w|^{2/k})\, d\bar{z}\wedge dz \\ + \mathbf{i}\frac{|w|^{2/k}(F''(|w|^{2/k})|w|^{2/k} + F'(|w|^{2/k}))}{k^2|w|^2}\, d\bar{w}\wedge dw\end{aligned}$$

Now, let $f : [0,1) \to [0,1)$ be a smooth function with $f(0) = 0$ and $f'(r) > 0$ for all $r \in [0,1)$ and choose $F : (0,2) \to \mathbb{R}$ such that

$$f(r^2) = 2 - \frac{1}{2}F'(r^{2/k})r^{2/k}.$$

The formula above then transforms to

$$\omega_F^{\mathbb{D}}(0,w) = -2\mathbf{i}(1 - f(2^{-k}|w|^2))\, d\bar{z}\wedge dz - \frac{2\mathbf{i}}{2^k k} f'(2^{-k}|w|^2)\, d\bar{w}\wedge dw$$

and hence we can take $\omega_f^{\mathbb{D}} := \omega_F^{\mathbb{D}} \in \Omega^2(\mathrm{X}_k^{\mathbb{D}})$ and this completes the proof. □

Lemma 5.2.4. *Let $f : [0,1) \to [0,1)$ be a smooth function with $f(0) = 0$ and $f'(r) > 0$ for all $r \in [0,1)$. Then there exists a unique $SL(2,\mathbb{R})$-invariant symplectic form $\omega_f \in \Omega^2(X_k)$ satisfying*

$$\begin{aligned}\omega_f(\boldsymbol{i}, w) = -\frac{\boldsymbol{i}}{2}\left(1 - f(|w|^2) + k|w|^2 f'(|w|^2)\right) d\bar{z}\wedge dz \\ - \frac{2\boldsymbol{i}}{k} f'(|w|^2)\, d\bar{w}\wedge dw + f'(|w|^2)\,(\bar{w}d\bar{z}dw - wd\bar{w}dz)\,.\end{aligned} \tag{5.15}$$

Moreover, f restricts to the hyperbolic area form along $\mathbb{H} \times \{0\}$.

Proof. This follows from Lemma 5.2.3 and Lemma 5.2.2 by direct calculation. □

Remark 5.2.5. For $k = 2$ and $f(r) = 1 - \sqrt{1-r}$ we recover the hyperkähler metric on the unit disc bundle in $T^*\mathbb{H}$ and $T^*\mathbb{D}$ respectively, see Theorem 4.5.1.

Moment maps for the action on the fibre

Fix $\omega_f \in \Omega^2(X_k)$ as in Lemma 5.2.4. We calculate in the following a moment map for the $\mathrm{SL}(2,\mathbb{R})$ action on X_k.

Lemma 5.2.6. *The Hamiltonian $H : X_k \to \mathbb{R}$ defined by*

$$H(z,w) = -\frac{2}{k} f\left(Im(z)^k |w|^2\right).$$

generates with respect to ω_f the Hamiltonian vector field

$$X_H(z,w) = (0, \boldsymbol{i}w).$$

Its flow generates the S^1 action on X_k which is given by rotation of the fibres.

Proof. By (5.15) it follows

$$\iota(X_H)\omega_f(0,w) = -\frac{2\bar{w}}{k} f'(|w|^2)dw + \mathbf{i}|w|^2 f'(|w|^2)dz - \mathbf{i}|w|^2 f'(|w|^2)d\bar{z}$$

On the other hand holds

$$dH(z,w) = -\frac{2}{k} f'\left(\mathrm{Im}(z)^k|w|^2\right)\left(\mathrm{Im}(z)^k \bar{w}dw + k\mathrm{Im}(z)^{k-1}|w|^2 \frac{dz - d\bar{z}}{2\mathbf{i}}\right).$$

Combining both formulas proves $\iota(X_H)\omega_f = dH$ on the fibre $z = \mathbf{i}$. The general case follows from this, since H, X_H and ω_f are all $\mathrm{SL}(2,\mathbb{R})$-invariant. □

Proposition 5.2.7. *The $SL(2,\mathbb{R})$ action on X_k is Hamiltonian with respect to ω_f. It is generated by the moment map $\mu : X_k \to \mathfrak{sl}^*(2,\mathbb{R})$ defined by*

$$\langle \mu(z,w), \xi\rangle = \left(f(Im(z)^k|w|^2) - 1\right) tr(j(z)\xi), \qquad \text{for } \xi \in \mathfrak{sl}(2,\mathbb{R})$$

where $j : \mathbb{H} \to \mathcal{J}(\mathbb{R}^2)$ is defined by (4.39).

Proof. For $w = 0$, the proposition follows from Lemma 4.3.7. For $0 < r < 1$ consider the circle bundle

$$S_r = \{(z,w) \in X_k \,|\, \mathrm{Im}(z)^k|w|^2 = r^2\}.$$

By Lemma 5.2.6, this is a level set of the Hamiltonian H which generates the S^1 action on X_k, given by rotation of the fibres. Thus S_r/S^1 is a Marsden–Weinstein quotient of X_k and it follows that there exists a function $h(r)$ with

$$(\omega_f)_{(z,w)}((\hat{z}_1,\hat{w}_1),(\hat{z}_1,\hat{w}_1)) = h(r)\omega_z^{\mathbb{H}}(\hat{z}_1,\hat{z}_1) \qquad \text{for all } (\hat{z}_i,\hat{w}_i) \in T_{(z,w)}S_r$$

where $\omega^{\mathbb{H}}$ denotes the hyperbolic area form on $\mathbb{H}$. Evaluating this expression on $(z,w) = (\mathbf{i}, r)$ and the tangent vectors $(1,0)$, $(\mathbf{i}, -\frac{k}{2}r) \in T_{(\mathbf{i},r)}S_r$ yields

$$h(r) = (\omega_f)_{(\mathbf{i},r)}\left((1,0), \left(\mathbf{i}, -\frac{k}{2}r\right)\right) = 1 - f(r^2).$$

It then follows for $(z,w) \in S_r$ and $(\hat{z},\hat{w}) \in T_{(z,w)}S_r$ that

$$\omega_f(L_{(z,w)}\xi, (\hat{z},\hat{w})) = h(r)\omega^{\mathbb{H}}(L_z\xi, \hat{z}) = \mathcal{L}_{(\hat{z},\hat{w})}\left(\left(f(r^2) - 1\right)\mathrm{tr}(j(z)\xi)\right).$$

where $L_{(z,w)} : \mathfrak{sl}(2,\mathbb{R}) \to T_{(z,w)}S_r$ denotes the infinitesimal action and the last step uses Lemma 4.3.7. Hence the moment map equation is satisfied in this case.

It remains to verify the moment map equation along the fibre. For this define $V_r, V_\phi \in \mathrm{Vect}(X_k \backslash (\mathbb{H} \times \{0\}))$ by

$$V_r(z,w) = \left(0, \frac{w}{r}\right), \qquad V_\phi(z,w) = (0, \mathbf{i}w)$$

and fix a point $(z,w) \in X_k \backslash (\mathbb{H} \times \{0\})$. Let $\xi \in \mathfrak{sl}(2,\mathbb{R})$ and assume first that $j(z)\xi = -\xi j(z)$. Then follows $j(z)\xi \in \mathfrak{sl}(2,\mathbb{R})$, $\mathbf{i}L_{(z,w)}\xi = -L_{(z,w)}(j(z))$, and therefore

$$\omega_f\left(L_{(z,w)}\xi, V_r\right) = \omega_f\left(\mathbf{i}L_{(z,w)}\xi, \frac{1}{r}V_\phi\right) = \omega_f\left(-L_{(z,w)}(j(z)\xi), \frac{1}{r}V_\phi\right) = 0$$

Here we used that V_ϕ is tangential to the S^1 action and therefore in the kernel of $(\omega_f)|_{TS_r \times TS_r}$. Since $j(z)\xi$ is complex anti-linear, we also have $\mathrm{tr}(j(z)\xi) = 0$ and the moment map equation is satisfied is this case.

Finally, assume that $\xi \in \mathfrak{sl}(2,\mathbb{R})$ satisfies $j(z)\xi = \xi j(z)$. Then ξ is a multiple of $j(z)$ and we may assume $\xi = j(z)$. We calculate in this case

$$\omega_f\left(L_{(z,w)}j(z), V_r\right) = \omega_f\left(\left(0, \frac{k\mathbf{i}w}{r}\right), (0,w)\right) = -4rf'(r^2)$$

and on the other hand

$$\partial_r \langle \mu(z,w), j(z)\rangle = \partial_r\left(2 - 2f(r^2)\right) = -4rf'(r^2).$$

Hence the moment map equation is again satisfied and this completes the proof of the proposition.

□

5.2.3 Moduli spaces of differentials of order k

Throughout this section, let (Σ, ρ) be a closed 2-dimensional oriented manifold with $\mathrm{genus}(\Sigma) \geq 2$ and fixed area form $\rho \in \Omega^2(\Sigma)$. Denote by $P \to \Sigma$ its $\mathrm{SL}(2,\mathbb{R})$ frame bundle, let $X_k \subset (T^*\mathbb{H})^{k/2}$ be defined by (5.11) and let $\omega_f \in \Omega^2(X_k)$ be a symplectic form as in Lemma 5.2.4.

Geometric description of the sections

Denote by $(j, \gamma_k) : X \to \mathcal{J}(\mathbb{R}^2) \times \mathrm{Hom}((\mathbb{R}^2)^{\otimes k}, \mathbb{C})$ the bijection defined in Lemma 4.3.5 and recall that the fibre maps $\gamma(\zeta, \cdot) : (T^*\mathbb{H})^{k/2}_\zeta \to D_k(j(\zeta))$ are complex anti-linear isometries. This yields an embedding of the associated bundle $P(X_k) := P \times_{\mathrm{SL}(2,\mathbb{R})} X$ into a suitable tensor bundle over Σ which is defined by

$$\begin{aligned} P(X_k) &\hookrightarrow \mathrm{End}(T\Sigma) \times S^k(T^*\Sigma \otimes \mathbb{C}), \\ [(z,\theta),(\zeta,\eta)] &\mapsto \left(\theta j(\zeta)\theta^{-1}, \theta^*\gamma_k(\zeta,\eta)\right) \end{aligned} \tag{5.16}$$

for $z \in \Sigma$, a volume preserving frame $\theta : \mathbb{R}^2 \to T_z\Sigma$ and $(\zeta,\eta) \in X_k$. On the space of section this yields the identification

$$\mathcal{S}(X_k) = \mathcal{D}^1_k(\Sigma) := \{(J,\tau) \mid J \in \mathcal{J}(\Sigma),\, \tau \in D_k(\Sigma, J),\, |\tau|_J < 1\}$$

where $D_k(\Sigma, J)$ denotes the space of complex differentials of order k on (Σ, J).

Lemma 5.2.8.

1. *Any torsion free* $SL(2,\mathbb{R})$ *connection on* $T\Sigma$ *induces connections on* $P(X)$ *and* $End(T\Sigma) \times S^k(T^*\Sigma \otimes \mathbb{C})$ *which are compatible with respect (5.16).*

2. *The inclusion (5.16) is* $Symp(\Sigma,\rho)$*-equivariant*

3. *The symplectic form* ω_f *on* X_k *induces a symplectic form on the bundle* $P(X_k)$ *(again denoted* ω_f*) which satisfies*

$$(\omega_f)_{(J,\tau)}((0,\hat{\tau}_1),(0,\hat{\tau}_2)) = -\frac{4}{k} f'(|\tau|^2)\omega_D(\hat{\tau}_1,\hat{\tau}_2)$$

for $(J,\tau) \in \mathcal{D}^1_k(\Sigma)$ *and* $\hat{\tau}_i \in D_k(J)$*. Here we denote by* ω_D *the pointwise symplectic structure obtained from* $S^k(T^*\Sigma \otimes_J \mathbb{C})$*.*

Proof. The first two claims are a matter of unravelling the definitions and left to the reader. The formula for the symplectic form follow from Lemma 5.2.4. □

Calculation of the moment map

The symplectic form on $P(X_k)$ integrates a symplectic form on $\mathcal{D}^1_k(\Sigma)$ defined by

$$\underline{\omega}_f((\hat{J}_1,\hat{\tau}_1),(\hat{J}_2,\hat{\tau}_2)) := \int_\Sigma \omega_f((\hat{J}_1,\hat{\tau}_1),(\hat{J}_2,\hat{\tau}_2))\rho.$$

The next theorem calculates a moment map for the action of $\mathrm{Ham}(\Sigma,\rho)$ on $\mathcal{D}^1_k(\Sigma)$ with respect to this symplectic form.

Theorem 5.2.9.

1. *Let* $c := 2\pi(2-genus(\Sigma))/vol(\Sigma,\rho)$*. A moment map for the action of* $Ham(\Sigma,\rho)$ *on* $\mathcal{D}^1_k(\Sigma)$ *is given by*

$$\begin{aligned}\underline{\mu}_f(J,\tau) = &\left[\frac{4f'(|\tau|^2)}{k}\left(|\bar{\partial}\tau|^2 - |\partial\tau|^2\right) + 2(1-f(|\tau|^2)K - 2c\right]\rho \\ &- 2\boldsymbol{i}\bar{\partial}\partial f(|\tau|^2).\end{aligned} \tag{5.17}$$

where K_J *is the Gaussian curvature for the metric* $\rho(\cdot,J\cdot)$ *and all covariant derivatives are induced by its Levi-Civita connection.*

2. *Define* $F:[0,1)\to\mathbb{R}$ *by* $F(t) := \int_0^t \frac{f(s)}{ks} - f'(s)\,ds$ *and suppose* $\bar{\partial}\tau = 0$*. Then*

$$\underline{\mu}_f(J,\tau) = \left(2K_J + \Delta F\left(|\tau|^2\right)\right)\rho \tag{5.18}$$

where $\Delta = d^*d$ *is the positive Laplacian of the metric* $\rho(\cdot,J\cdot)$*.*

Proof. Let ∇ be the Levi-Civita connection for the metric $\rho(\cdot,J\cdot)$. The moment map equation (5.17) follows from Theorem 4.2.4 once we have identified the three components of the general moment map in the present context. The constant c follows from the Gauss–Bonnet theorem and guarantees that $\underline{\mu}_f(J,\tau)$ takes values in the space of exact 2-forms.

Step 1.1: $\omega_f(\nabla(J,\tau)\wedge\nabla(J,\tau)) = \frac{4f'(|\tau|^2)}{k}\left(|\bar\partial\tau|^2 - |\partial\tau|^2\right)\rho$.

For the Levi–Civita connection, it holds $\nabla J = 0$ and with Lemma 5.2.8 it follows

$$\omega_f(\nabla_u(J,\tau),\nabla_v(J,\tau)) = -\frac{4f'(|\tau|^2)}{k}\omega_D(\nabla_u\tau,\nabla_v\tau)$$

for all $u,v\in\mathrm{Vect}(\Sigma)$. Using the relation $|\partial_u\tau|^2 - |\bar\partial_u\tau|^2 = \omega_D(\nabla_u\sigma,\nabla_{Ju}\sigma)$ for any $u\in\mathrm{Vect}(\Sigma)$, it follows

$$\left(|\partial\tau|^2 - |\bar\partial\tau|^2\right)\rho = \omega_D(\nabla\tau\wedge\nabla\tau).$$

Combining both expressions proves Step 1.1.

Step 1.2: $\langle\mu_{(J,\tau)},R^\nabla\rangle = 2(f(|\tau|^2)-1)K\rho$.

Let $\mu : X_k \to \mathfrak{sl}^*(2,\mathbb{R})$ be the moment map calculated in Proposition 5.2.7. For $(J,\tau)\in\mathcal{D}^1_k(\Sigma)$, this yields a section of $\mu_{(J,\tau)}\in\mathrm{End}_0(T\Sigma)^*$ defined by

$$\mu_{(J,\tau)}(\Psi) = \left(f\left(|\tau|^2\right)-1\right)\mathrm{tr}\,(J\Psi)$$

for $\Psi\in\Omega^0(\Sigma,\mathrm{End}_0(T\Sigma))$. Step 1.2 follows from this and $R^\nabla = -KJ\otimes\rho$.

Step 1.3: $dc(\nabla\mu_{(J,\tau)}) = 2\mathbf{i}\bar\partial\partial f(|\tau|^2)$.

Since $\nabla J = 0$, we obtain

$$\nabla_u\mu_{(J,\tau)}(\Psi) = \mathcal{L}_u f(|\tau|^2)\mathrm{tr}(J\Psi)$$

for all $\Psi\in\Omega^0(\Sigma,\mathrm{End}_0(T\Sigma))$ and $u\in\mathrm{Vect}(\Sigma)$. Let e_1, $e_2 = Je_1$ be a local orthonormal frame for $T\Sigma$ and write $v\in\mathrm{Vect}(\Sigma)$ as $v = v_1e_1+v_2e_2$. Then

$$\begin{aligned}
c(\nabla\mu_{(J,\tau)})(v) &= -\mathcal{L}_{e_1}f(|\tau|^2)\mathrm{tr}(Je_1^*\otimes v) - \mathcal{L}_{e_2}f(|\tau|^2)\mathrm{tr}(Je_2^*\otimes v)\\
&= \mathcal{L}_{e_1}f(|\tau|^2)v_2 - \mathcal{L}_{e_2}f(|\tau|^2)v_1\\
&= -\mathcal{L}_{Jv}f(|\tau|^2).
\end{aligned}$$

Step 3 follows then from the relation $d(dg\circ J) = 2\mathbf{i}\bar\partial\partial g$ which holds for every smooth function $g:\Sigma\to\mathbb{C}$.

Step 2: *Now assume $\bar\partial\tau = 0$. Then holds $\underline{\mu}_f(J,\tau) = \left(2K+\Delta F(|\tau|^2)\right)\rho$.*

We work around a point where $\tau\neq 0$ and define $h := |\tau|^2$. Since $\bar\partial\tau = 0$, we have $|\partial h| = |\tau||\partial\tau|$ and hence

$$|\partial\tau|^2\rho = h^{-1}|\partial h|^2 = -\frac{\mathbf{i}}{2h}\bar\partial h\wedge\partial h. \tag{5.19}$$

Next choose holomorphic coordinates and write

$$\rho = \lambda dx \wedge dy, \qquad \tau(z) = g(z)dz^k$$

for some positive function λ and a holomorphic function g. The Gaussian curvature K_J can be computed in these coordinates via

$$K_J = -\frac{1}{2}\lambda^{-1}(\partial_x^2 \log(\lambda) + \partial_y^2 \log(\lambda)).$$

Since $f(z)$ is holomorphic, $\log(|f(z)|^2)$ is harmonic and we compute

$$\begin{aligned}
\bar\partial\partial \log(h) &= \frac{1}{4}(\partial_x^2 + \partial_y^2)\log(|f(z)|^2\lambda^{-k})2\mathbf{i}\, dx \wedge dy \\
&= -\frac{k}{2}\mathbf{i}(\partial_x^2 + \partial_y^2)\log(\lambda)\, dx \wedge dy \\
&= \mathbf{i}kK_J\lambda dx \wedge dy.
\end{aligned}$$

This shows

$$K\rho = -\frac{\mathbf{i}}{k}\bar\partial\partial\log(h). \tag{5.20}$$

Now plug (5.19) and (5.20) into (5.17) to obtain

$$\begin{aligned}
\underline{\mu}_f(J,\tau) &= \frac{2\mathbf{i}f'(h)}{kh}\bar\partial h \wedge \partial h + \frac{2\mathbf{i}(f(h)-1)}{k}\bar\partial\partial\log(h) - 2\mathbf{i}\bar\partial\partial f(h) \\
&= 2\mathbf{i}\bar\partial\left[\frac{f(h)-1}{k}\partial\log(h) - \partial f(h)\right] \\
&= 2K\rho + 2\mathbf{i}\bar\partial\left[\left(\frac{f(h)}{kh} - f'(h)\right)\partial h\right] \\
&= 2K\rho + 2\mathbf{i}\bar\partial\partial F(h)
\end{aligned}$$

This proves Step 2 and the theorem, since $2\mathbf{i}\bar\partial\partial F(h) = \Delta F(h)\rho$. □

The moduli space

The space of holomorphic pairs $\{(J,\tau) \in \mathcal{D}_k^1(\Sigma)\,|\,\bar\partial\tau = 0\} \subset \mathcal{D}_k^1(\Sigma)$ is a $\mathrm{Symp}(\Sigma,\rho)$-invariant complex and hence symplectic submanifold. Theorem 5.2.9 shows that its Marsden–Weinstein quotient by $\mathrm{Ham}(\Sigma,\rho)$ is given by

$$\tilde{\mathcal{M}}_f(k) := \left\{(J,\gamma) \in \mathcal{D}_k^1(\Sigma) \,\middle|\, \bar\partial\gamma = 0,\ K_J + \frac{1}{2}\Delta F(|\gamma|^2) = c\right\} \Big/ \mathrm{Ham}(\Sigma,\rho)$$

where $c = 2\pi(2 - 2\mathrm{genus}(\Sigma))/\mathrm{vol}(\Sigma,\rho)$.

Define $H := \mathrm{Symp}_0(\Sigma,\rho)/\mathrm{Ham}(\Sigma,\rho)$. Since $\mathrm{Ham}(\Sigma,\rho) < \mathrm{Symp}_0(\Sigma,\rho)$ is a normal subgroup (see [85], Proposition 10.2) this is indeed a quotient group and the flux homomorphism yields an identification $H \cong H^1(\Sigma,\rho)$, see Remark 4.4.7. It is not hard to see that the H-orbits in $\tilde{\mathcal{M}}_f(k)$ are complex and hence symplectic submanifolds. The argument is essentially the same as in Lemma 4.4.8. It then follows from Lemma 4.4.9 that the quotient

$$\mathcal{M}_f(k) := \left\{ (J,\gamma) \in \mathcal{D}^1_k(\Sigma) \;\middle|\; \bar{\partial}\gamma = 0, \quad K_J + \frac{1}{2}\Delta F(|\gamma|^2) = c \right\} \Big/ \mathrm{Symp}_0(\Sigma,\rho)$$

carries an induced symplectic structure.

For every Riemannian metric $g \in \mathrm{Met}(\Sigma)$ there exists a unique complex structure $J_g \in \mathcal{J}(\Sigma)$ which is compatible with g and we define $\mathcal{D}_k(g) := \mathcal{D}_k(J_g,\Sigma)$. It then follows from standard Moser isotopy arguments that

$$\mathcal{M}_f(k) \cong \left\{ (g,\tau) \;\middle|\; \begin{array}{c} g \in \mathrm{Met}(\Sigma),\ \tau \in D_k(g) \\ \bar{\partial}\tau = 0, \quad |\tau|_g < 1, \quad K_g + \frac{1}{2}\Delta F\left(|\tau|^2\right) = c \end{array} \right\} \Big/ \mathrm{Diff}_0(\Sigma)$$

where the isomorphism is induced by $(J,\tau) \mapsto (\rho(\cdot, J\cdot),\tau)$.

The next proposition provides a simpler model for the moduli space $\mathcal{M}_f(k)$ for a particular choice of f. In the case $k = 2$, this choice corresponds precisely to the hyperkähler metric.

Theorem 5.2.10. *There exists a unique smooth function $F_k : [0,1) \to (0,\infty)$ with*

$$te^{-kF_k(t)} - ke^{-F_k(t)} + (k-1) = 0, \qquad F_k(0) = \log\left(\frac{k-1}{k}\right) \tag{5.21}$$

This is concave, strictly monotone decreasing and satisfies $\lim_{t\to 1} F_k(t) = 0$. Define $f_k : [0,1) \to [0,1)$ by

$$f_k(t) := -\left(\int_0^t F_k'(s)s^{-1/k}\,ds\right) t^{1/k}. \tag{5.22}$$

This is a smooth function with $f_k(0) = 0$ and $f_k' > 0$ and therefore defines a symplectic form $\underline{\omega}_{f_k}$ on $\mathcal{D}^1_k(\Sigma)$. For this symplectic form, there is an isomorphism

$$\mathcal{M}_{f_k}(k) \cong \left\{ (g,\tau) \;\middle|\; \begin{array}{c} g \in Met(\Sigma),\ \tau \in \mathcal{D}_k(g) \\ \bar{\partial}\tau = 0, \quad |\tau|_g < 1, \quad K_g - \frac{c}{k}|\tau|_g^2 = c\frac{k-1}{k} \end{array} \right\} \Big/ Diff_0(\Sigma)$$

induced by the map $(J,\tau) \mapsto (e^{F_k(|\tau|_J^2)}\rho(\cdot,J\cdot),\tau)$.

Proof. One checks first that $F_k(t)$ is well-defined, strictly monotone descreasing and concave. All of this follows by successively applying the implicit function theorem which we leave to the reader. Next we solve the differential equation

$$F_k'(t) = \frac{f_k(t)}{kt} - f_k'(t), \qquad f_k(0) = 0.$$

The homogeneous equation $y'(t) = y(t)/kt$ has the solution $y(t) = ct^{1/k}$. Using variation of constants, we make the ansatz $f_k(t) = c(t)t^{1/k}$ for the inhomogeneous equation. This leads to the formula

$$f_k(t) := -\left(\int_0^t F_k'(s)s^{-1/k}\,ds\right) t^{1/k}.$$

Clearly, $f_k(0) = 0$ and $f_k'(t) > 0$ are satisfied and we need to argue that $f_k(1) \leq 1$. For $k = 2$ this follows from the explicit formulas

$$F_2(t) = \log\left(1 + \sqrt{1-t}\right), \qquad f_2(t) = 1 - \sqrt{1-t}.$$

Now assume $k \geq 3$. By monotonicity and concavity of F_k, it follows

$$-\int_0^{\frac{1}{2}} F_k'(s)s^{-1/s}\,ds \leq -F_k'\left(\frac{1}{2}\right)\frac{k}{k-1} \leq 2\log\left(\frac{k}{k-1}\right)\frac{k}{k-1}.$$

On the other hand, we have

$$-\int_{\frac{1}{2}}^1 F_k'(s)s^{-1/k}\,ds \leq 2^{1/k}\log\left(\frac{k}{k-1}\right)$$

and therefore

$$-\int_0^1 u_k'(s)s^{-1/s}\,ds \leq 2\log\left(\frac{k}{k-1}\right)\frac{k}{k-1} + 2^{1/k}\log\left(\frac{k}{k-1}\right) < 1.$$

The last estimate holds for all $k \geq 3$. We thus have shown that $f_k : [0,1) \to [0,1)$ gives rise to a symplectic form $\underline{\omega}_{f_k}$ on $\mathcal{D}_k^1(\Sigma)$.

The moduli space $\mathcal{M}_{f_k}(k)$ parametrizes pairs $g \in \mathrm{Met}(\Sigma)$ and $\tau \in \mathcal{D}_k(g)$ satisfying

$$\bar{\partial}\tau = 0, \quad |\tau|_g < 1, \quad K_g + \frac{1}{2}\Delta u_k(|\tau|^2) = c.$$

Define $\tilde{g} \in \mathrm{Met}(\Sigma)$ by $\tilde{g} = e^{F_k(|\tau|^2)}g$. Since F_k is monotone decreasing with $\lim_{t\to 1} F_k(t) = 0$, it follows that $|\tau|_{\tilde{g}} < 1$ is equivalent to $|\tau|_g < 1$. The Gaussian curvature transforms under this conformal change as follows

$$K_{\tilde{g}} = e^{-F_k(|\tau|_g^2)}\left(K_g + \frac{1}{2}\Delta_g u_k(|\tau|_g^2)\right) = ce^{-F_k(|\tau|_g^2)}.$$

Using $|\tau|_{\tilde{g}}^2 = |\tau|_g^2 e^{-kF_k(|\tau|^2)}$ and (5.21), this yields

$$K_{\tilde{g}} = ce^{-F_k(|\tau|_g^2)} = c\frac{|\tau|_g^2}{k}e^{-kF_k(|\tau|_g^2)} + c\frac{k-1}{k} = c\frac{|\tau|_{\tilde{g}}^2}{k} + c\frac{k-1}{k}.$$

This completes the proof of the theorem. □

5.3 Moduli spaces of differentials of mixed order

Let $\mathbf{k} = (k_1, k_2, \ldots, k_n) \in \mathbb{Z}_{\geq 2}^n$ be a vector of integers and define

$$X_{\mathbf{k}} := \left\{(z, \mathbf{w}) \in \mathbb{H} \times \mathbb{C}^n \;\middle|\; \mathrm{Im}(z)^{k_i}|w_i|^2 < 1 \text{ for } i = 1, \ldots, n\right\}.$$

We think of $X_{\mathbf{k}}$ as a subset of $(T^*\mathbb{H})^{k_1/2} \oplus \cdots \oplus (T^*\mathbb{H})^{k_n/2}$. This induces a natural action of $\mathrm{SL}(2,\mathbb{R})$ which is given by

$$\mathrm{SL}(2,\mathbb{R}) \times X_{\mathbf{k}} \to X_{\mathbf{k}}, \qquad \begin{pmatrix} a & b \\ c & d \end{pmatrix}(z,w) := \left(\frac{az+b}{cz+d}, (cz+d)^{\mathbf{k}}\mathbf{w}\right)$$

where

$$(cz+d)^{\mathbf{k}}\mathbf{w} := \left((cz+d)^{k_1}w_1, (cz+d)^{k_2}w_2, \dots, (cz+d)^{k_n}w_n\right).$$

The next lemma introduces a particular class of symplectic forms on $X_{\mathbf{k}}$. Recall from Lemma 5.2.4 that a function $f_i : [0,1) \to [0,1)$ with $f_i(0) = 0$ and $f' > 0$ yields a symplectic form $\omega_{f_i} \in \Omega^2(X_{k_i})$ on the space

$$X_{k_i} := \left\{(z,w) \in \mathbb{H} \times \mathbb{C} \,\middle|\, \mathrm{Im}(z)^{k_i}|w|^2 < 1\right\}.$$

Lemma 5.3.1 (Symplectic structures). *Let $\alpha = (\alpha_1, \dots, \alpha_n) \in (0,1)^n$ be a vector of weights with $\sum_{i=1}^n \alpha_i = 1$ and let $\boldsymbol{f} = (f_1, \dots, f_n)$ be a tuple of smooth functions $f_i : [0,1) \to [0,1)$ with $f_i(0) = 0$ and $f_i' > 0$. Denote the canonical projections by*

$$\pi_i : X_{\boldsymbol{k}} \to X_{k_i}, \qquad \pi_i(z, \boldsymbol{w}) := (z, w_i).$$

and define $\omega_{\boldsymbol{f}}^{\alpha} \in \Omega^2(X_{\boldsymbol{k}})$ by

$$\omega_{\boldsymbol{f}}^{\alpha} := \sum_{i=1}^n \alpha_i \pi_i^* \omega_{f_i}.$$

Then $\omega_{\boldsymbol{f}}^{\alpha}$ is a symplectic form and restricts to the hyperbolic area form on $\mathbb{H} \times \{0\}$.

Proof. Since each of the ω_{f_i} is a closed 2-form which restricts to the hyperbolic area form, it follows that $\omega_{\mathbf{f}}^{\alpha}$ is also closed and restricts to the hyperbolic area form on $\mathbb{H} \times \{0\}$.

It remains to show that $\omega_{\mathbf{f}}^{\alpha}$ is non-degenerate. Along the fibres

$$(X_{\mathbf{k}})_z = \{\mathbf{w} \in \mathbb{C}^n \mid (z, \mathbf{w}) \in X_{\mathbf{k}}\}$$

the 2-form $\omega_{\mathbf{f}}^{\alpha}$ is a product symplectic form and hence non-degenerated. Moreover there exist unique $\mathrm{SL}(2,\mathbb{R})$-equivariant vector fields $V, W \in \mathrm{Vect}(X_{\mathbf{k}})$ with

$$V(\mathbf{i}, \mathbf{w}) := \left(1, \mathbf{i}\frac{k_1}{2}w_1, \mathbf{i}\frac{k_2}{2}w_2, \dots, \mathbf{i}\frac{k_n}{2}w_n\right) \in T_{(\mathbf{i},\mathbf{w})}X_{\mathbf{k}}$$
$$W(\mathbf{i}, \mathbf{w}) := \left(\mathbf{i}, -\frac{k_1}{2}w_1, -\frac{k_2}{2}w_2, \dots, -\frac{k_n}{2}w_n\right) \in T_{(\mathbf{i},\mathbf{w})}X_{\mathbf{k}}$$

along the fibre above $\mathbf{i}$. These span the $\omega_{\mathbf{f}}^{\alpha}$ orthogonal complements of the fibre and yield a horizontal distribution. Finally, one checks

$$\omega_{\mathbf{f}}(V, W) = 1 - \sum_{i=1}^n \alpha_i f_i\left(|w_i|^2\right) > 0$$

and hence $\omega_{\mathbf{f}}$ is non-degenerated. □

Lemma 5.3.2. *Fix the symplectic structure $\omega_{\boldsymbol{f}}^{\alpha} \in \Omega^2(X_{\boldsymbol{k}})$. The $SL(2,\mathbb{R})$ action on $X_{\boldsymbol{k}}$ is then Hamiltonian with moment map*

$$\mu : X_{\boldsymbol{k}} \to \mathfrak{sl}^*(2,\mathbb{R}), \qquad \langle \mu(z, \boldsymbol{w}), \xi \rangle = \left(\sum_{i=1}^n \alpha_i f_i(Im(z)^{k_i}|w_n|^2) - 1\right) tr(j(z)\xi)$$

where $j : \mathbb{H} \to \mathcal{J}(\mathbb{R}^2)$ is defined by (4.39).

Proof. The embedding $X_{\mathbf{k}} \hookrightarrow \prod_{i=1}^n X_{k_i}$ defined by

$$(z, \mathbf{w}) \mapsto ((z, w_1), (z, w_2), \ldots, (z, w_n))$$

is $\mathrm{SL}(2, \mathbb{R})$-equivariant and symplectic with respect to the weighted product symplectic structure $(\alpha_1 \omega_1) \oplus \cdots \oplus (\alpha_n \omega_n)$. The lemma follows thus from Lemma 5.2.7. □

Denote by

$$\mathcal{D}_{\mathbf{k}}(\mathbb{R}^2) := \big\{(J, \gamma_1, \ldots, \gamma_n) \,|\, J \in \mathcal{J}(\mathbb{R}^2),\ \gamma_i \in D_{k_i}(J) \text{ for } i = 1, \ldots, n\big\}$$

Then Lemma 4.3.5 yields a map

$$(j, \gamma_1, \ldots, \gamma_n) : X_{\mathbf{k}} \to \mathcal{D}_{\mathbf{k}}(\mathbb{R}^2)$$

whose image contains all tuples with $|\gamma_i| < 1$ for all i. This yields for the associated bundle the identification

$$\begin{aligned} P(X_{\mathbf{k}}) &\hookrightarrow \mathrm{End}(T\Sigma) \times S^{k_1}(T^*\Sigma \otimes \mathbb{C}) \times \cdots \times S^{k_n}(T^*\Sigma \otimes \mathbb{C}), \\ [(z, \theta), (\zeta, \eta)] &\mapsto \big(\theta j(\zeta)\theta^{-1}, \theta^* \gamma_{k_1}(\zeta, \eta_1), \ldots, \theta^* \gamma_{k_n}(\zeta, \eta_n)\big) \end{aligned} \tag{5.23}$$

for $z \in \Sigma$, a volume preserving frame $\theta : \mathbb{R}^2 \to T_z\Sigma$ and $(\zeta, \eta_1, \ldots, \eta_n) \in X_{\mathbf{k}}$. Its space of sections is then identified with

$$\mathcal{S}(P(X_{\mathbf{k}})) \cong \mathcal{D}^1_{\mathbf{k}}(\Sigma) := \{(J, \tau_1, \ldots, \tau_n) \,|\, J \in \mathcal{J}(\Sigma),\ \tau_i \in D_{k_i}(J, \Sigma)),\ |\tau_i| < 1\}\,.$$

The symplectic form $\omega_{\mathbf{f}} \in \Omega^2(X_{\mathbf{k}})$ integrates to a symplectic form on $\mathcal{D}^1_{\mathbf{k}}(\Sigma)$ defined by

$$\underline{\omega}_{\mathbf{f}}((\hat{J}, \hat{\tau}_1, \ldots, \hat{\tau}_n), (\hat{J}', \hat{\tau}'_1, \ldots, \hat{\tau}'_n)) := \int_\Sigma \omega_{\mathbf{f}}((\hat{J}, \hat{\tau}_1, \ldots, \hat{\tau}_n), (\hat{J}', \hat{\tau}'_1, \ldots, \hat{\tau}'_n))\rho$$

The next theorem shows that the action of $\mathrm{Ham}(\Sigma, \rho)$ on Σ is Hamiltonian and calculates a moment map for it.

Theorem 5.3.3.

1. *The action of $Ham(\Sigma, \rho)$ on $\mathcal{D}^1_{\mathbf{k}}(\Sigma)$ is Hamiltonian with moment map*

$$\underline{\mu}^\alpha_{\boldsymbol{f}}(J, \tau_1, \ldots, \tau_n) = \sum_{i=1}^n \alpha_i \underline{\mu}_i(J, \tau_i) \tag{5.24}$$

where

$$\begin{aligned} \underline{\mu}_i(J, \tau_i) = &\left[\frac{4 f_i'(|\tau_i|^2)}{k}\left(|\bar{\partial}\tau_i|^2 - |\partial\tau_i|^2\right) + 2(1 - f_i(|\tau_i|^2)K_J - 2c\right]\rho \\ &+ 2\boldsymbol{i}\bar{\partial}\partial f_i(|\tau|_i^2). \end{aligned} \tag{5.25}$$

In this formula K_J denotes is the Gaussian curvature for the metric $\rho(\cdot, J\cdot)$ and all covariant derivatives are induced by its Levi-Civita connection.

2. Suppose $\bar\partial\gamma_i = 0$ for all i. Then

$$\underline{\mu}^{\alpha}_{\boldsymbol{f}}(J,\tau) = 2K_J\rho + \sum_{i=1}^{n}\left(\Delta\alpha_i F_i(|\tau|_i^2)\right)\rho \tag{5.26}$$

where $F_i : [0,1) \to \mathbb{R}$ is defined by

$$F_i(t) := \int_0^t \frac{f_i(t)}{kt} - f_i'(t)\,dt$$

and $\Delta = d^*d$ is the positive Laplacian of the metric $\rho(\cdot, J\cdot)$.

Proof. The embedding $\mathcal{D}^1_{\mathbf{k}}(\Sigma) \hookrightarrow \prod_{i=1}^n \mathcal{D}^1_{k_i}(\Sigma)$ defined by

$$(J,\tau_1,\dots,\tau_n) \mapsto ((J,\tau_1),(J,\tau_2),\dots,(J,\tau_n))$$

is $\mathrm{Symp}(\Sigma,\rho)$-equivariant and symplectic with respect to the α-weighted product symplectic structure. The theorem follows thus directly from Theorem 5.2.9. □

After taking the action of $\mathrm{Symp}_0(\Sigma,\rho)/\mathrm{Ham}(\Sigma,\rho)$, this construction gives rise to a symplectic form on the moduli space

$$\mathcal{M}^{\alpha}_{\mathbf{f}}(\mathbf{k}) := \left\{ (J,\tau) \in \mathcal{D}^1_{\mathbf{k}}(\Sigma) \;\middle|\; \begin{array}{c} \bar\partial\gamma = 0, \\ K_J + \frac{1}{2}\Delta\left(\sum_{i=1}^n \alpha_i F_i\left(|\tau_i|^2\right)\right) = c \end{array} \right\} \Big/ \mathrm{Symp}_0(\Sigma,\rho)$$

5.4 Hitchin's higher Teichmüller components

We describe in this section the relation between holomorphic differentials and higher Teichmüller components introduced by Hitchin [59]. For a more complete overview of the subject, we refer to [74, 16].

The Hitchin components

Let G^c be a complex simple Lie group and let G^r be the adjoint group of the split real form of G^c. Denote by $\mathrm{Hom}^*(\pi_1(\Sigma), G^r)$ the space of representations, which act completely reducible on the Lie algebra of G^r. Hitchin [59] showed that the moduli space

$$\mathcal{R}^*_{G^r}(\Sigma) := \mathrm{Hom}^*(\pi_1(\Sigma), G^r)/G^r \tag{5.27}$$

contains a connected component $\mathcal{H}_{G^r}(\Sigma)$ homeomorphic to $\mathbb{R}^{(2\mathrm{genus}(\Sigma)-2)\dim(G^r)}$. He called this the Teichmüller component, which by now is often called the Hitchin component.

There exits a natural inclusion of Teichmüller space $\mathcal{T}(\Sigma)$ into the Hitchin component which relies on the notion of the principle 3-dimensional subgroup introduced by Kostant [72]. This gives rise to a distinguished inclusion $\mathrm{PSL}(2,\mathbb{R}) \hookrightarrow G^r$ and therefore

$$\mathcal{T}(\Sigma) \subset \mathcal{R}^*_{\mathrm{PSL}(2,\mathbb{R})}(\Sigma) \to \mathcal{R}^*_{G^r}(\Sigma) \tag{5.28}$$

where we view Teichmüller space as the space of Fuchsian representations (see Theorem 4.4.13).

Parametrizing the Hitchin component by holomorphic differentials

Fix a complex structure $J \in \mathcal{J}(\Sigma)$ and denote the space of holomorphic differentials of degree k by

$$\mathcal{D}_k(J) := \left\{\tau \in \Omega^0(\Sigma, S^k(T^*\Sigma \otimes_J \mathbb{C})) \,|\, \bar{\partial}_J \tau = 0\right\}$$

Let $p_1, \ldots, p_\ell$ be a basis for the ring of invariant polynomials on the Lie algebra of G^c and denote by $m_i := \deg(p_i)$ their degrees. Hitchin showed that there exists a parametrization of the Hitchin component $\mathcal{H}_{G^r}(\Sigma)$ by the product space of the bundle of holomorphic differentials of degree m_i, i.e.

$$\mathcal{H}_{G^r}(\Sigma) \cong \mathcal{D}_{m_1}(J) \oplus \mathcal{D}_{m_2}(J) \oplus \cdots \oplus \mathcal{D}_{m_\ell}(J).$$

The construction of this map uses the theory of Higgs bundles and we will briefly describe it in the case of $G^r = \mathrm{PSL}(n+1, \mathbb{R})$ below. The general case is somewhat analogue, but requires considerably more Lie theory.

The case $G^r = \mathbf{PSL}(n+1, \mathbb{R})$

The invariant polynomials of $\mathfrak{sl}(n+1, \mathbb{C})$ are generated by the coefficients of the characteristic polynomial

$$\det(t\mathbb{1} - \xi) = t^{n+1} + p_2(\xi)t^{n-2} + \cdots + p_{n+1}(\xi)$$

(since $p_1(\xi) = -\mathrm{tr}(\xi) = 0$ for $\xi \in \mathfrak{sl}(n+1, \mathbb{C})$). Since $\deg(p_i) = i$, we seek a parametrization

$$\mathcal{H}_{\mathrm{PSL}(n+1,\mathbb{R})}(\Sigma) \cong \mathcal{D}_2(J) \oplus \mathcal{D}_3(J) \oplus \cdots \oplus \mathcal{D}_{n+1}(J).$$

One can view the n-fold symmetric power $S^n(\mathbb{C}^2)$ as the space $\mathbb{C}_n[z_1, z_2]$ of homogeneous polynomials of degree n. This has the natural basis $z_1^n, z_1^{n-1}z_2, \ldots, z_2^n$ and is thus isomorphic to $\mathbb{C}^{n+1}$. The action of $\mathrm{SL}(2, \mathbb{C})$ on $\mathbb{C}^2$ induces an action on $S^n(\mathbb{C}^2)$ which gives rise to an irreducible representation $\mathrm{SL}(2, \mathbb{C}) \to \mathrm{SL}(n+1, \mathbb{C})$. The corresponding inclusion $\mathrm{PSL}(2, \mathbb{C}) \hookrightarrow \mathrm{PSL}(n+1, \mathbb{C})$ agrees with the one obtained from Kostant's theory of the principle 3-dimensional subgroup.

Next, choose a holomorphic line bundle $L \to \Sigma$ with $L^2 = T\Sigma$ and define

$$W = S^k(L^{-1} \oplus L) = L^{-n} \oplus \cdots \oplus L^n.$$

We view W as holomorphic vector bundle with structure group $S^1 \subset \mathrm{SL}(n+1, \mathbb{C})$. For $\tau_k \in \mathcal{D}_k(J) \subset \Omega^{0,1}(\Sigma, \mathrm{Hom}(L^n, L^{n-2k})$ define the Higgs field $\Phi \in \Omega^{0,1}(\Sigma, \mathrm{End}_0(W))$ by

$$\Phi := \begin{pmatrix} 0 & \mathbf{1} & 0 & \cdots & 0 \\ 0 & 0 & \mathbf{1} & \cdots & 0 \\ \vdots & \vdots & \ddots & & \vdots \\ 0 & 0 & & 0 & \mathbf{1} \\ \tau_{n+1} & \tau_n & \cdots & \tau_2 & 0 \end{pmatrix}.$$

where $\mathbf{1} \in \Omega^{1,0}(\Sigma, L^2) = \Omega^{1,0}(\Sigma, \mathrm{Hom}(L^i, L^{i+2})$ are canonically defined. Using standard results on stable Higgs bundles, it is not hard to show that the pair (W, Φ) is a stable Higgs bundle. It therefore admits an irreducible unitary connection $A \in \mathcal{A}(W)$ which is compatible with the holomorphic structure of W and satisfies the equation

$$F_A + [\Phi \wedge \Phi^*] = 0.$$

A formal consequence of this equation is that $B := A + \Phi + \Phi^*$ defines an irreducible flat $\mathrm{SL}(n+1, \mathbb{C})$-connections. Hitchin showed that this connection has in fact holonomy in $\mathrm{SL}(n+1, \mathbb{R})$ and hence gives rise to a representation

$$\rho_\Phi : \pi_1(\Sigma) \to \mathrm{SL}(n+1, \mathbb{R})$$

which is uniquely determined by the Higgs bundle (W, Φ) up to conjugation. The induced representations in $\mathcal{R}^*_{\mathrm{PSL}(n+1,\mathbb{R})}(\Sigma)$ then parametrize the Hitchin component $\mathcal{H}_{\mathrm{PSL}(n+1,\mathbb{R})}(\Sigma)$ as $\tau_k \in \mathcal{D}_k(J)$ varies.

Mapping class group invariant structures

In the representational point of view, the mapping class group appears as the group of outer automorphism of the fundamental group

$$\mathrm{MCG}(\Sigma) \cong \mathrm{Out}(\pi_1(\Sigma)) := \mathrm{Aut}(\pi_1(\Sigma)(\mathrm{Inn}(\pi_1(\Sigma))$$

(see Theorem 4.4.12). This group acts naturally on $\mathcal{R}^*_{G^r}(\Sigma)$ and preserves the Hitchin component. This follows from the fact, that the mapping class group preserves the Teichmüller component of Fuchsian representations in $\mathcal{R}^*_{\mathrm{PSL}(2,\mathbb{R})}(\Sigma)$ and the inclusion (5.28) is mapping class group equivariant.

The parametrization of the Hitchin component

$$h_J : \mathcal{D}_{m_1}(J) \oplus \mathcal{D}_{m_2}(J) \oplus \cdots \oplus \mathcal{D}_{m_\ell}(J) \to \mathcal{H}_{G^r}(\Sigma)$$

sketched above depends on the choice of a complex structure $J \in \mathcal{J}(\Sigma)$ and does not yields any mapping class group invariant structures on the Hitchin component.

Now consider the case $G^r = \mathrm{PSL}(n, \mathbb{R})$ and define

$$\mathcal{D} := \{(J, \tau_3, \ldots, \tau_n) \mid J \in \mathcal{J}(\Sigma),\ \tau_k \in \mathcal{D}_k(J)\} / \mathrm{Diff}_0(\Sigma).$$

The parametrizations of Hitchin then combine to a mapping class group equivariant map

$$\mathcal{D} \to \mathcal{H}_{\mathrm{PSL}(n,\mathbb{R})}(\Sigma), \qquad [J, \tau_3, \ldots, \tau_n] \mapsto h_J(0, \tau_3, \ldots, \tau_n). \tag{5.29}$$

Labourie [75] showed that this map is surjective and conjectured that it is in fact a homeomorphism. This conjecture has been verified for $n = 3$ by Labourie [73] and Loftin [79].

Moreover, using (5.29) one obtains mapping class group equivariant maps

$$\mathcal{M}^\alpha_{\mathbf{f}}(3, 4, \ldots, n) \to \mathcal{H}_{\mathrm{PSL}(n,\mathbb{R})}(\Sigma)$$

which are conjecturally embeddings. It would be interesting to understand the push-forward of the Kähler structure on $\mathcal{M}^\alpha_{\mathbf{f}}(3, 4, \ldots, n)$ under this map. This gives rise to mapping class group invariant structures along the image and one might hope that these structures extend to mapping class group invariant structures on the Hitchin component, or more generally, they might suggest how such a structure can be obtained.

Chapter 6

The Ricci form and Calabi–Yau Teichmüller space

This chapter summarizes joint work with Oscar Garcia–Prada and Dietmar A. Salamon [50]. We show that the Ricci form yields a moment map for the action of the group of exact volume preserving diffeomorphims on the space of almost complex structures. This gives rise to an extended Weil–Petersson symplectic form on the Calabi–Yau Teichmüller space of isotopy classes of complex structures with real first Chern class zero and nonempty Kähler cone. We also discuss variants of the theory for Kähler–Einstein pairs which have not been included into our joint paper. The presentation in this chapter is rather brief and we only sketch the arguments for the more technical results. Full details can be found in our joint article [50].

6.1 Introduction

Let (M, ρ) be a closed $2n$-dimensional manifold with fixed volume form ρ. The space $\mathcal{J}(M)$ of almost complex structures on M, compatible with the orientation determined by ρ, carries the natural symplectic form defined by

$$\Omega_{\rho,J}(\hat{J}_1, \hat{J}_2) := \int_M \frac{1}{2}\mathrm{tr}\left(\hat{J}_1 J \hat{J}_2\right)\rho. \tag{6.1}$$

We define the Ricci form $\mathrm{Ric}_{\rho,J} \in \Omega^2(M)$ associated to the volume form ρ and an almost complex structure $J \in \mathcal{J}(M)$ by

$$\mathrm{Ric}_{\rho,J}(u,v) := \frac{1}{4}\mathrm{tr}\left((\nabla_u J)J(\nabla_v J)\right) + \frac{1}{2}\mathrm{tr}\left(JR^\nabla(u,v)\right) + \frac{1}{2}d\lambda_J^\nabla \tag{6.2}$$

for $u, v \in \mathrm{Vect}(M)$, where ∇ is a torsion free ρ-connection on M and the 1-form λ_J^∇ is defined by $\lambda_J^\nabla(u) := \mathrm{tr}\left((\nabla J)u\right)$ for $u \in \mathrm{Vect}(M)$. The next theorem can be derived as a special case of Donaldson's moment map [38] (see Theorem 4.2.4). However, in [50] we give a direct and independent proof of this result.

Theorem A (**Ricci form**). *The Ricci form $Ric_{\rho,J} \in \Omega^2(M)$ does not depend on the choice of the connection ∇ used to define it, represents the cohomology class*

$2\pi c_1(TM, J)$ and agrees with the usual definition of the Ricci form on Kähler manifolds. The map $J \mapsto 2Ric_{\rho,J}$ satisfies the moment map equation for the action of the exact volume preserving diffeomorphism group on the space of almost complex structures.

Proof. See Theorem 6.2.1. □

A useful generalization of the moment map equation involves the 1-form $\Lambda_\rho \in \Omega^1(\mathcal{J}(M), \Omega^1(M))$ defined by

$$\Lambda_\rho(J, \hat{J})(u) := \operatorname{tr}\left((\nabla \hat{J})u + \frac{1}{2}\hat{J}J\nabla_u J\right) \tag{6.3}$$

for $u \in \mathrm{Vect}(M)$, where ∇ is a torsion free ρ-connection on M. The linearisation of $\mathrm{Ric}_{\rho,J}$ when varying J in direction $\hat{J}$ is given by $\frac{1}{2}d\Lambda_\rho(J, \hat{J})$ and we show in Lemma 6.2.5 that

$$\int_M \Lambda_\rho(J, \hat{J}) \wedge \iota(v)\rho = \Omega_{J,\rho}(\hat{J}, \mathcal{L}_v J) \tag{6.4}$$

for all $v \in \mathrm{Vect}(M)$. This setup leads to a new construction of the Weil–Petersson symplectic form on Calabi–Yau Teichmüller space

$$\mathcal{T}_0(M) := \left\{ J \in \mathcal{J}_{\mathrm{int}}(M) \,\middle|\, \begin{array}{c} c_1(TM, J) = 0 \in H^2(M, \mathbb{R}) \\ \text{and } J \text{ admits a Kähler form} \end{array} \right\} \Big/ \mathrm{Diff}_0(M). \tag{6.5}$$

This moduli space has been studied extensively in the polarized cased [64, 90, 98] and for $K3$-surfaces, see [44] Chapter 16. The Bogomolov–Tian–Todorov theorem [11, 111, 113] asserts that $\mathcal{T}_0(M)$ is a smooth manifold. However, it is not Hausdorff in general [54, 120]. The construction of the Weil–Petersson metric involves three main steps:

1. The natural inclusion of

$$\mathcal{T}_0(M, \rho) := \{J \in \mathcal{J}_{\mathrm{int},0}(M) \,|\, \mathrm{Ric}_{\rho,J} = 0\} / \mathrm{Diff}_0(M, \rho) \tag{6.6}$$

   into Teichmüller space $\mathcal{T}_0(M)$ is a bijection.

2. The group $\mathrm{Diff}_0(M, \rho)/\mathrm{Diff}^{\mathrm{ex}}(M, \rho)$ acts trivially on

$$\mathcal{T}_0^{\mathrm{ex}}(M, \rho) := \mathcal{J}_{\mathrm{int},0}(M, \rho)/\mathrm{Diff}^{\mathrm{ex}}(M, \rho). \tag{6.7}$$

   Hence, $\mathcal{T}_0(M, \rho) = \mathcal{T}_0^{\mathrm{ex}}(M, \rho)$ embeds into the Marsden–Weinstein quotient of $\mathcal{J}(M)$ and carries a natural closed 2-form.

3. The space of integrable structures $\mathcal{J}_{\mathrm{int}}(M) \subset \mathcal{J}(M)$ is not a symplectic submanifold and it is not obvious that the closed 2-form on $\mathcal{T}_0(M, \rho)$ is non-degenerated. We give a complete characterization of the kernel of the restriction of the symplectic form which then proves non-degeneracy of the Weil–Petersson symplectic form.

The tangent spaces at the space of integrable complex structures are

$$T_J\mathcal{J}_{\mathrm{int}}(M) = \ker\left(\bar\partial_J : \Omega^{0,1}(M,TM) \to \Omega^{0,2}(M,TM)\right). \tag{6.8}$$

If $\mathrm{Ric}_{\rho,J} = 0$ and $\bar\partial_J \hat J = 0$ then there exist smooth functions $f, g : M \to \mathbb{R}$ such that

$$\Lambda_\rho(J,\hat J) = -df \circ J + dg \tag{6.9}$$

Moreover, for every integrable $J \in \mathcal{J}_{\mathrm{int}}(M)$ with vanishing real first Chern class and non-empty Kähler cone, there exists a unique volume form ρ_J with $\mathrm{Ric}_{\rho_J,J} = 0$ and $\int_M \rho_J = V$.

Theorem B (**Weil–Petersson symplectic form**). *The Weil–Petersson symplectic form on $\mathcal{T}_0(M,\rho)$ is given by*

$$\Omega_J(\hat J_1, \hat J_2) = \int_M \left(\frac{1}{2} tr\left(\hat J_1 J \hat J_2\right) - f_1 g_2 + f_2 g_1\right)\rho_J \tag{6.10}$$

for $J \in \mathcal{J}_{int}(M)$ with vanishing real first Chern class and non-empty Kähler cone, $\hat J_i \in \Omega^{0,1}(M,TM)$ with $\bar\partial_J \hat J_i = 0$ and f_i, g_i defined by (6.9). This symplectic form is $Diff_0(M)$ equivariant and thus the mapping class group acts on $\mathcal{T}_0(M)$ by symplectomorphism.

Proof. See Theorem 6.3.5. □

The Weil–Petersson symplectic form on Teichmüller spaces gives rise to a symplectic connection on the bundle $\mathcal{E}_0(M)$ of isotopy classes of Ricci-flat Kähler structures over the space $\mathcal{B}_0(M)$ of symplectic forms with vanishing first Chern class.

Theorem C (**A symplectic connection**). *The projection $\mathcal{E}_0(M) \to \mathcal{B}_0(M)$ is a submersion and for every Ricci flat Kähler structure (ω, J) on M and for every closed 2-form $\hat\omega$, there exists a unique element $\hat J = \mathcal{A}_{\omega,J}(\hat\omega) \in \Omega^{0,1}_J(M,TM)$ satisfying*

$$\Omega_J(\hat J, \hat J') = 0 \qquad \textit{for all } \hat J' \in \Omega^{0,1}_J(M,TM) \textit{ with } \bar\partial_J \hat J' = 0 \textit{ and } \hat J' = (\hat J')^*$$

and the tangency conditions

$$\bar\partial_J \hat J = 0, \quad \Lambda_\rho(J,\hat J) = -d\langle\hat\omega,\omega\rangle \circ J, \quad \hat\omega(\cdot,\cdot) - \hat\omega(J\cdot, J\cdot) = \langle(\hat J - \hat J^*)\cdot,\cdot\rangle.$$

This connection is $Diff_0(M)$-equivariant and satisfies $\mathcal{A}_{\omega,J}(d\iota(v)\omega) = \mathcal{L}_v J$ for all $v \in Vect(M)$ with $d\iota(Jv)\rho = 0$.

Proof. See Theorem 6.3.6 □

The final section discusses variants of the theory for Kähler–Einstein manifolds which have not been included into our joint paper. Fix a volume form $\rho \in \Omega^{2n}(M)$ and cohomology classes $a, c \in H^2(M)$ such that $2\pi c = \kappa a$ for some $\kappa \in \mathbb{R}$. Denote by $\mathcal{S}_a(M,\rho) \subset \Omega^2(M)$ the space of symplectic forms on M with volume form $\omega^n/n! = \rho$ and denote by $\mathcal{J}_c(M)$ the space of almost complex structures with $c_1(TM,J) = c$. We call $a \in H^2(M,\mathbb{R})$ a Lefschetz class when $\cdot \cup a^{n-1} : H^1(M,\mathbb{R}) \to H^{2n-1}(M,\mathbb{R})$ is an isomorphism. By the hard Lefschetz theorem, every Kähler class is Lefschetz. The converse is not true in general, see [123, 45, 84] and the references therein. We show that under this assumptions $\mathcal{S}_a(M,\rho)$ is a symplectic manifold with the Lefschetz symplectic form

$$\Omega_\omega(\hat\omega_1,\hat\omega_2) := \int_M \lambda_1\wedge\lambda_2\wedge\frac{\omega^{n-1}}{(n-1)!} \tag{6.11}$$

for $\omega\in\mathcal{S}_a(M,\rho)$ and exact 2-forms $\hat\omega\in\Omega^2(M)$ with $\hat\omega\wedge\omega^{n-1}=0$, where $\lambda_i\in\Omega^1(M)$ satisfy $d\lambda_i=\hat\omega$ and $\lambda_i\wedge\omega^{n-1}$ is exact. The motivation for this symplectic form comes from a moment map description of the equation $\omega^n/n!=\rho$.

Theorem D (Kähler–Einstein pairs). *The action of $\mathrm{Diff}^{ex}(M,\rho)$ on the product space $\mathcal{J}_c(M)\times\mathcal{S}_a(M,\rho)$ is Hamiltonian for the product symplectic form*

$$\Omega_{J,\omega}((\hat J_1,\hat\omega_1),(\hat J_2,\hat\omega_2)) := \int_M \frac{1}{2}tr\left(\hat J_1 J\hat J_2\right)\rho - 2\kappa\lambda_1\wedge\lambda_2\wedge\frac{\omega^{n-1}}{(n-1)!} \tag{6.12}$$

where $\lambda_i\in\Omega^1(M)$ satisfy $d\lambda_i=\hat\omega$ and $\lambda_i\wedge\omega^{n-1}$ is exact. A moment map for this action is given by $\mu:\mathcal{J}_c(M)\times\mathcal{S}_a(M,\rho)\to\Omega^2_{ex}(M)$ defined by

$$\mu(J,\omega) = 2(Ric_{\rho,J}-\kappa\omega). \tag{6.13}$$

Proof. See Theorem 6.4.3. □

This leads to a Weil–Petersson metric on the Teichmüller space of Kähler–Einstein manifolds with a fixed symplectic form $\omega\in\mathcal{S}_a(M)$. Although this yields a new perspective on the subject, the symplectic form has been studied extensively, see [71, 98, 105].

6.2 The Ricci form

6.2.1 Linear complex structures

Denote the space of linear complex structures on $\mathbb{R}^{2n}$ which are compatible with the standard orientation by

$$\mathcal{J}(\mathbb{R}^{2n}) := \{J\in\mathrm{SL}(2n,\mathbb{R})\,|\,J^2=-\mathbb{1}\}. \tag{6.14}$$

The group $\mathrm{SL}(2n,\mathbb{R})$ acts transitively on $\mathcal{J}(\mathbb{R}^{2n})$ by conjugation. Since every $J\in\mathcal{J}(\mathbb{R}^{2n})$ has trace zero, one can view $\mathcal{J}(\mathbb{R}^{2n})\subset\mathfrak{sl}(2n,\mathbb{R})$ as an adjoint orbit. It follows from this setup that $\mathcal{J}(\mathbb{R}^{2n})$ carries a canonical symplectic form for which the $\mathrm{SL}(2n,\mathbb{R})$ action is Hamiltonian. More explicitly, the tangent space at $J\in\mathcal{J}(\mathbb{R}^{2n})$ is given by

$$T_J(\mathbb{R}^{2n}) := \left\{\hat J\in\mathbb{R}^{2n\times 2n}\ \mid J\hat J+\hat JJ=0\} = \{[\xi,J]\,|\,\xi\in\mathfrak{sl}(2n,\mathbb{R})\right\}. \tag{6.15}$$

and the symplectic form $\tau\in\Omega^2(\mathcal{J}(\mathbb{R}^{2n}))$ is defined by

$$\tau_J(\hat J_1,\hat J_2) := \frac{1}{2}\mathrm{tr}\left(\hat J_1 J\hat J_2\right) = -\mathrm{tr}\left([\xi_1,\xi_2]J\right) \tag{6.16}$$

for $\xi_i \in \mathfrak{sl}(2n,\mathbb{R})$ and $\hat{J}_i = [\xi_i, J]$. The symplectic from τ is of type $(1,1)$ with respect to the complex structure $\hat{J} \mapsto -J\hat{J}$. The corresponding symmetric form is indefinite and thus turns $\mathcal{J}(\mathbb{R}^{2n})$ into a pseudo-Kähler manifold. It follows from the general setup that the $\mathrm{SL}(2n,\mathbb{R})$ action on $\mathcal{J}(\mathbb{R}^{2n})$ is Hamiltonian with moment map $\mu : \mathcal{J}(\mathbb{R}^{2n}) \to \mathfrak{sl}(2n,\mathbb{R})^*$ defined by

$$\langle \mu(J), \xi \rangle := -\mathrm{tr}\,(J\xi) \qquad \text{for } \xi \in \mathfrak{sl}(2n,\mathbb{R}). \tag{6.17}$$

6.2.2 The Ricci form as a moment map

Let (M,ρ) be a closed oriented $2n$-dimensional manifold and let $P \to M$ denote its $\mathrm{SL}(2n,\mathbb{R})$ frame bundle. The space of sections of the associated bundle $P \times_{\mathrm{SL}(2n,\mathbb{R})} \mathcal{J}(\mathbb{R}^{2n})$ admits a canonical identification with the space of almost complex structures

$$\mathcal{J}(M) := \left\{ J \in \Omega^0(M, \mathrm{End}(TM)) \;\middle|\; \begin{array}{c} J^2 = -\mathbb{1} \text{ and } J \text{ is compatible} \\ \text{with the orientation of } M \end{array} \right\}. \tag{6.18}$$

This is a symplectic submanifold with the induced symplectic form

$$\Omega_{J,\rho}(\hat{J}_1, \hat{J}_2) := \frac{1}{2}\int_M \mathrm{tr}\left(\hat{J}_1 J \hat{J}_2\right)\rho \tag{6.19}$$

for $\hat{J}_1, \hat{J}_2 \in \Omega^{0,1}_J(M,TM)$. We are now in the general framework considered by Donaldson [38] and the next theorem can be obtained as a special case of Theorem 4.2.4. In [50] we give a direct and independent proof of this result.

Theorem 6.2.1 (The Ricci form). *Define $Ric_{\rho,J} \in \Omega^2(M)$ by (6.2).*

1. *$Ric_{\rho,J} \in \Omega^2(M)$ is closed and independent of the torsion free ρ-connection ∇ used to defined. It satisfies the naturality condition*

$$Ric_{\phi^*\rho,\phi^*J} = \phi^* Ric_{\rho,J}, \qquad \text{for all } \phi \in Diff(M) \tag{6.20}$$

   and the scaling property

$$Ric_{e^f\rho,J} = Ric_{\rho,J} + \frac{1}{2}d(df \circ J), \qquad \text{for all } f \in \Omega^0(M). \tag{6.21}$$

2. *The map $J \mapsto 2Ric_{\rho,J}$ satisfies the moment map equation for the action of the group $Diff^{ex}(M,\rho)$ of exact volume preserving diffeomorphism on the space $\mathcal{J}(M)$.*

Remark 6.2.2. The Ricci form is only in the Calabi–Yau case a honest moment map. Otherwise it takes values in the space of closed 2-forms, which is not quite the dual space of the space of exact divergence free vector fields. Nevertheless, the moment map equation for the Ricci form is well-posed and satisfied in every case, see Remark 4.2.5.

Proof. It can be deduced from Donaldson's moment map in Theorem 4.2.4 that $\mathrm{Ric}_{\rho,J}$ is closed, independent of the connection ∇ used to define it and satisfies the moment map equation, see Theorem 4.2.4. In [50] we give direct arguments for all these properties without invoking Theorem 4.2.4.

(6.20) follows easily from (6.2), once we know that it is independent of the connection ∇. Equation (6.21) is equivalent to

$$\left.\frac{d}{dt}\right|_{t=0} \mathrm{Ric}_{\rho_t,J} = d\left(d(\hat{\rho}/\rho)\circ J\right)$$

for any smooth path $t \mapsto \rho_t$ of positive volume forms with $\rho_0 = \rho$ and $\partial_t|_{t=0}\rho_t = \hat{\rho}$. This equation can be shown by a direct computation which we leave to the reader. □

Now suppose $\omega \in \Omega^2(M)$ is a closed 2-form compatible with J such that $\omega^n/n! = \rho$. Let ∇ be the Levi-Civita connection of the Riemannian metric $\omega(\cdot, J\cdot)$ and define $\tilde{\nabla} := \nabla - \frac{1}{2}J\nabla J$. One can check that $\tilde{\nabla}$ preserves ρ, J and the metric. Moreover, the Ricci form of (ρ, J) is given by $\mathrm{Ric}_{\rho,J} = \frac{1}{2}\mathrm{tr}(JR^{\tilde{\nabla}})$. It follows from this that (6.2) coincides with the usual definition of the Ricci form in the Kähler setting and $\mathrm{Ric}_{\rho,J}$ represents the cohomology class $2\pi c_1(TM, J)$.

Theorem 6.2.3. *Let (M, J, ω) be a $2n$-dimensional Kähler manifold and define $\rho := \frac{\omega^n}{n!} \in \Omega^{2n}(M)$.*

1. *There exists a diffeomorphism $\phi \in \mathrm{Diff}_0(M)$ such that $\mathrm{Ric}_{\rho,\phi^*J} = 0$ if and only if $c_1(TM, J) = 0$.*
2. *Let $\phi : M \to M$ be an orientation preserving diffeomorphism. If $\mathrm{Ric}_{\rho,\phi^*J} = \mathrm{Ric}_{\rho,J}$, then $\phi^*\rho = \rho$.*

Proof. Suppose first $c_1(TM, J) = 0$. Then $\mathrm{Ric}_{\rho,J}$ is exact and there exists a function f with

$$d(df \circ J) = \frac{1}{2}\mathrm{Ric}_{\rho,J}, \qquad \int_M e^{-f}\rho = \int_M \rho.$$

By Moser isotopy, there exists $\phi \in \mathrm{Diff}_0(M)$ with $\phi^*(e^{-f}\rho) = \rho$ and it follows with (6.20) and (6.21) that

$$\mathrm{Ric}_{\rho,\phi^*J} = \mathrm{Ric}_{\phi^*(e^{-f}\rho),\phi^*J} = \phi^*\mathrm{Ric}_{e^{-f}\rho,J} = \phi^*\left(\mathrm{Ric}_{\rho,J} - \frac{1}{2}d(df \circ J)\right) = 0.$$

This proves the first part. For the second part assume $\mathrm{Ric}_{\rho,\phi^*J} = \mathrm{Ric}_{\rho,J} = 0$ and define f by $\phi_*\rho = e^{-f}\rho$. Then follows with (6.20) and (6.21)

$$\frac{1}{2}d(df \circ J) = \mathrm{Ric}_{\rho,J} - \mathrm{Ric}_{e^{-f}\rho,J} = -\mathrm{Ric}_{\phi_*\rho,J} = -\phi_*\mathrm{Ric}_{\rho,\phi^*J} = 0$$

Thus f is constant and $\int_M e^{-f}\rho = \int_M \rho$ then yields $f \equiv 0$. □

Remark 6.2.4 (Bott–Chern cohomology). The Kähler assumption in the previous theorem was made for convenience. One can show for any integrable complex structure J that $\mathrm{Ric}_{\rho,J}$ represents the first Bott–Chern class of the holomorphic tangent bundle in $H^{1,1}_{BC}(M, J) := \left(\ker d \cap \Omega^{1,1}_J(M)\right) / \left\{d(df \circ J) \mid f \in \Omega^0(M)\right\}$. Then there exists a diffeomorphism ϕ with $\mathrm{Ric}_{\rho,\phi^*J} = 0$ if and only if the first Bott–Chern class of (TM, J) vanishes.

Theorem 6.2.3 implies that Teichmüller space admits the description

$$\mathcal{T}_0(M) \cong \mathcal{T}_0(M,\rho) := \{J \in \mathcal{J}_{\mathrm{int},0}(M) \,|\, \mathrm{Ric}_{\rho,J} = 0\} \,/ \mathrm{Diff}_0(M,\rho). \tag{6.22}$$

For this we used $\mathrm{Diff}_0(M,\rho) = \mathrm{Diff}_0(M) \cap \mathrm{Diff}(M,\rho)$ which follows from a parametrized version of Moser isotopy.

6.2.3 The 1-form $\Lambda_\rho(J,\hat{J})$

Define the 1-form $\Lambda_\rho \in \Omega^1(\mathcal{J}(M), \Omega^1(M))$ by (6.3). One can check that the linearisation of $\mathrm{Ric}_{\rho,J}$, when varying J in direction $\hat{J}$, is given

$$\widehat{\mathrm{Ric}}_\rho(J,\hat{J}) = \frac{1}{2} d\Lambda_\rho(J,\hat{J}). \tag{6.23}$$

The next lemma can thus be view as generalization of the moment map equation.

Lemma 6.2.5. *The* 1*-form* $\Lambda_\rho \in \Omega^1(\mathcal{J}(M), \Omega^1(M))$ *does not depend on the torsion free* ρ*-connection* ∇ *used to define it, satisfies the naturality condition*

$$\Lambda_{\phi^*\rho}(\phi^* J, \phi^* \hat{J}) = \phi^* \Lambda_\rho(J,\hat{J}), \qquad \textit{for all } \phi \in \mathit{Diff}(M) \tag{6.24}$$

and the generalized moment map equation

$$\int_M \Lambda_\rho(J,\hat{J}) \wedge \iota(v)\rho = \Omega_{J,\rho}(\hat{J}, \mathcal{L}_v J) \tag{6.25}$$

for every $v \in \mathit{Vect}(M)$ *and* $\hat{J} \in \Omega^{0,1}_J(M,TM)$.

Proof. We prove (6.25). This equation implies in turn that Λ_ρ is indeed independent of the connection ∇ used to defined. Using

$$(\mathcal{L}_v J)u = J\nabla_u v - \nabla_{Ju} v + (\nabla_v J)u$$

for $u, v \in \mathrm{Vect}(M)$, we obtain

$$\mathrm{tr}\left(\hat{J}J\mathcal{L}_v J\right) = \mathrm{tr}\left(-\hat{J}\nabla v - \hat{J}J\nabla_{J\cdot} v + \hat{J}J\nabla_v J\right) = \mathrm{tr}\left(-2\hat{J}\nabla v + \hat{J}J\nabla_v J\right)$$

and thus

$$\begin{aligned}\Lambda_\rho(J,\hat{J})(v) &= \mathrm{tr}\left((\nabla\hat{J})v + \frac{1}{2}\hat{J}J\nabla_v J\right) = \mathrm{tr}\left(\nabla(\hat{J}v) - \hat{J}\nabla v + \frac{1}{2}\hat{J}J\nabla_v J\right) \\ &= \mathrm{tr}\left(\nabla(\hat{J}v)\right) + \frac{1}{2}\mathrm{tr}\left(\hat{J}J\mathcal{L}_v J\right)\end{aligned}$$

for every $v \in \mathrm{Vect}(M)$. Since $\mathrm{tr}\left(\nabla(\hat{J}v)\right)\rho = d\iota(\hat{J}v)\rho$ is exact, it follows

$$\int_M \Lambda_\rho(J,\hat{J}) \wedge \iota(v)\rho = \int_M \Lambda_\rho(J,\hat{J})(v)\rho = \frac{1}{2}\int_M \mathrm{tr}\left(\hat{J}J\mathcal{L}_v J\right)\rho = \Omega_{J,\rho}(\hat{J}, \mathcal{L}_v J)$$

and this completes the proof of the lemma. $\square$

6.2.4 Scalar curvature

Let $\omega \in \Omega^2(M)$ be a symplectic form with $\frac{\omega^n}{n!} = \rho$ and denote by

$$\mathcal{J}(M,\omega) := \{J \in \mathcal{J}(M) \mid J \text{ is compatible with } \omega\} \tag{6.26}$$

the space of compatible complex structures. The scalar curvature of the pair (J,ω) is defined by

$$S_{\omega,J} := 2\langle \mathrm{Ric}_{\omega,J}, \omega\rangle = \frac{2\mathrm{Ric}_{\omega,J} \wedge \omega^{n-1}/(n-1)!}{\omega^n/n!} \in \Omega^0(M). \tag{6.27}$$

As a Corollary of Theorem 6.2.1, we then obtain the following well-known theorem of Quillen, Fujiki [48] and Donaldson [34].

Corollary 6.2.6 (Donaldson–Fujiki–Quillen). *The map $J \mapsto S_{\omega,J}$ is a moment map for the action of Ham(M,ω) on $\mathcal{J}(M,\omega)$.*

Proof. Define $v_H \in \mathrm{Vect}(M)$ by $\iota(v_H)\omega = dH$. Then $\iota(v_H)\rho := d\left(H\frac{\omega^{n-1}}{(n-1)!}\right)$ and it follows from Theorem 6.2.1 that the differential of the map

$$J \mapsto \int_M S_{\omega,J} H \frac{\omega^n}{n!} = \int_M 2\mathrm{Ric}_{\omega,J} \wedge H \frac{\omega^{n-1}}{(n-1)!}$$

is given by $\hat{J} \mapsto \Omega_{J,\rho}(\hat{J}, \mathcal{L}_v J)$. This proves the Corollary. □

6.3 The Weil–Petersson symplectic form

We establish first some important properties for Ricci-flat Kähler manifolds and then uses these to investigate the subsubspace $\mathcal{J}_{int,0}(M) \subset \mathcal{J}(M)$ of integrable complex structures with non-empty Kähler cone and vanishing real first Chern class. The restriction of the symplectic form (6.19) to $\mathcal{J}_{int,0}(M)$ is not symplectic and we completely characterize its kernel in Proposition 6.3.3. After this preparatory work, we derive an explicit formula for the Weil–Petersson symplectic form on Teichmüller space in Theorem 6.3.5. We briefly indicate how this gives rise to a symplectic connection on the space of isotopy classes of Ricci flat Kähler structures in Theorem 6.3.6.

6.3.1 Ricci-flat Kähler manifolds

For $v \in \mathrm{Vect}(M)$ define $f_v \in \Omega^0(M)$ by $f_v\rho = d\iota(v)\rho$. The starting point for our discussion is the identity

$$\Lambda_\rho(J, \mathcal{L}_v J) = 2\iota(v)\mathrm{Ric}_{\rho,J} - df_v \circ J + df_{Jv} \tag{6.28}$$

which holds for every integrable complex structure $J \in \mathcal{J}_{\mathrm{int}}(M)$ and $v \in \mathrm{Vect}(M)$. Our derivation of this identity in [50] goes by a lengthy computation which we omit here. We also need the following technical lemma.

Lemma 6.3.1. *Let (M, J, ω) be a $2n$-dimensional closed Kähler manifold with $\rho := \frac{\omega^n}{n!}$ and let $v \in \mathrm{Vect}(M)$. Then the following holds*

1. *$\mathcal{L}_v J \in \Omega^0(M, \mathrm{End}(TM))$ is self-adjoint if and only if $d\iota(v)\omega \in \Omega^{1,1}_J(M)$.*

2. *$\iota(J\bar{\partial}_J^* \hat{J}^*)\omega = \Lambda_\rho(J, \hat{J})$.*

3. *$\iota(v)\omega$ is harmonic if and only if $d\iota(v)\omega = d\iota(Jv)\omega = 0$.*

Proof. Define $\hat{\omega} = d\iota(v)\omega$. From the identity

$$\hat{\omega}(u, u') - \hat{\omega}(Ju, Ju') = \langle (\mathcal{L}_v J)u - (\mathcal{L}_v J)^* u, u' \rangle$$

we see that $\mathcal{L}_v J \in \Omega^0(M, \mathrm{End}(TM))$ is self-adjoint if and only if $d\iota(v)\omega$ is of type $(1, 1)$ and this proves the first part.

For the second part, note first that $\mathcal{L}_v J = 2J(\bar{\partial}_J v)$. By (6.25) we then have

$$\begin{aligned}\int_M \Lambda_\rho(J, \hat{J}) \wedge \iota(v)\rho &= \frac{1}{2} \int_M \mathrm{tr}\left(\hat{J} J \mathcal{L}_v J\right) \rho = -\langle \hat{J}^*, \bar{\partial}_J v \rangle \\ &= \int_M \omega(J\bar{\partial}_J^* \hat{J}^*, v)\rho = \int_M \iota(J\bar{\partial}_J^* \hat{J}^*)\omega \wedge \iota(v)\rho\end{aligned}$$

for all $v \in \mathrm{Vect}(M)$ and this proves the second part.

We prove the third part. For every 1-form $\lambda \in \Omega^1(M)$ it holds

$$*\lambda = -\lambda \circ J \wedge \frac{\omega^{n-1}}{(n-1)!}. \tag{6.29}$$

and hence $*\iota(v)\omega = \iota(Jv)\rho$. Therefore, $d\iota(v)\omega = d\iota(Jv)\omega = 0$ implies that $\iota(v)\omega$ is harmonic. Assume conversely that $\iota(v)\omega$ is harmonic. Then $d\iota(v)\omega = 0$ and $d\iota(Jv)\rho = 0$ by (6.29). Hence $\mathcal{L}_v J$ is self-adjoint, by the first part of the lemma. Since $\mathcal{L}_{Jv} J = J\mathcal{L}_v J$ is also self-adjoint, $d\iota(Jv)\omega$ is an exact $(1, 1)$ form and there exists $f \in \Omega^0(M)$ with $d\iota(Jv)\omega = d(df \circ J)$ and

$$d(df \circ J) \wedge \frac{\omega^{n-1}}{(n-1)!} = d\iota(Jv)\rho = 0.$$

By (6.29) it follows that $d^*df = 0$. Hence f is constant and $d\iota(Jv)\omega = 0$. This complete the proof of the lemma. □

Proposition 6.3.2. *Let (M, J, ω) be a $2n$-dimensional closed Kähler manifold with $\rho := \frac{\omega^n}{n!}$ and $Ric_{\rho,J} = 0$.*

1. *Let $v \in \mathrm{Vect}(M)$. Then $\mathcal{L}_v J = 0$ if and only if $\iota(v)\omega$ is harmonic.*

2. *Let $\hat{J} \in \Omega^{0,1}_J(M, TM)$ with $\bar{\partial}_J \hat{J} = 0$. Then there exist unique smooth functions $f, g : M \to \mathbb{R}$ such that*

$$\Lambda_\rho(J, \hat{J}) = -df \circ J + dg, \qquad \int_M f\rho = \int_M g\rho = 0. \tag{6.30}$$

Proof. Assume first that $\iota(v)\omega$ is harmonic. Then $\mathcal{L}_v J$ is self-adjoint by Lemma 6.3.1 and (6.28) yields

$$\iota(\bar\partial_J^*\bar\partial_J v)\omega = -\frac{1}{2}\iota(J\bar\partial_J^*\mathcal{L}_v J)\omega = -\frac{1}{2}\Lambda_\rho(J,\mathcal{L}_v J) = 0.$$

Hence $\bar\partial_J v = 0$ and so $\mathcal{L}_v J = 0$. Assume conversely $\mathcal{L}_v J = 0$. Then follows from (6.28) that $f_v = f_{Jv} = 0$ and by (6.29) $d * \iota(v)\omega = d\iota(Jv)\rho = 0$. Moreover, it follows Lemma 6.3.1, that $d\iota(v)\omega$ is an exact $(1,1)$-form and hence there exists $f \in \Omega^0(M)$ with $d(df \circ J) = d\iota(v)\omega$. Since $d(df \circ J) \wedge \omega^{n-1} = 0$, it follows from (6.29) that $d^*df = 0$. Hence f is constant and $d\iota(v)\omega = 0$. This completes the proof of the first part.

For the second part, we show first that

$$\bar\partial_J \hat J = 0, \quad \bar\partial_J^* \hat J = 0 \qquad \Longrightarrow \qquad \Lambda_\rho(J,\hat J) = 0. \tag{6.31}$$

For this define $v := \bar\partial_J^* \hat J^* \in \mathrm{Vect}(M)$. By Lemma 6.3.1, it follows $\iota(v)\omega = -\Lambda_\rho(J, J\hat J)$. The holomorphic Poincaré lemma and (6.28) show then that $d\iota(v)\omega$ is an exact $(1,1)$-form. Using Lemma 6.3.1 again, it follows that $\mathcal{L}_v J$ is self-adjoint. Hence $\bar\partial_J\bar\partial_J^*(\hat J^* - \hat J) = \bar\partial_J v$ is also self-adjoint and therefore

$$0 = \langle \bar\partial_J\bar\partial_J^*(\hat J^* - \hat J), (\hat J^* - \hat J)\rangle_{L^2} = \langle \bar\partial_J^*(\hat J^* - \hat J), \bar\partial_J^*(\hat J^* - \hat J)\rangle_{L^2} = ||\bar\partial_J^*\hat J^*||_{L^2}^2.$$

This shows $\bar\partial_J^*\hat J^* = 0$ and hence $\Lambda_\rho(J,\hat J) = 0$ by Lemma 6.3.1. This completes the proof of (6.31).

Now let $\hat J \in \Omega_J^{0,1}(M,TM)$ with $\bar\partial_J \hat J = 0$ be given and choose $v \in \mathrm{Vect}(M)$ such that $\bar\partial_J^*(\hat J - \mathcal{L}_v J) = 0$. Then follows from (6.31) that $\Lambda_\rho(J,\hat J) = \Lambda_\rho(J,\mathcal{L}_v J)$ and (6.30) follows from (6.28) with $f = f_v$ and $g = f_{Jv}$. □

6.3.2 The space of integrable complex structures

Denote by $\mathcal{J}_{\mathrm{int}}(M)$ the space of integrable complex structures which are compatible with the orientation of M. The Newlander–Nierenberg theorem shows that an almost complex structure is integrable if and only if the Nijenhuis tensor $N_J(u,v) := [u,v] + J[Ju,v] + J[u,Jv] - [Ju,Jv]$ vanishes. Differentiating this conditions yields

$$T_J\mathcal{J}_{\mathrm{int}}(M) = \left\{ \hat J \in \Omega^{0,1}(M,TM) \,|\, \bar\partial_J \hat J = 0 \right\}. \tag{6.32}$$

In the next proposition we assume that $J \in \mathcal{J}_{\mathrm{int}}(M)$ admits a Ricci-flat Kähler form. We show that $T_J\mathcal{J}_{\mathrm{int}}(M) \subset T_J\mathcal{J}(M)$ is not a symplectic subspace and more precisely the kernel of the restriction of the symplectic form $\Omega_{\rho,J}$ defined by (6.19) to $T_J\mathcal{J}_{\mathrm{int}}(M)$ is the space $\{\mathcal{L}_v J \,|\, d\iota(v)\rho = d\iota(Jv)\rho = 0\}$.

Proposition 6.3.3. *Let (M,J,ω) be a $2n$-dimensional closed Kähler manifold. Define $\rho := \frac{\omega^n}{n!}$ and assume that $Ric_{\rho,J} = 0$. Let $\hat J \in \Omega^{0,1}(M,TM)$ with $\bar\partial_J \hat J = 0$ be given. Then*

$$\Omega_{J,\rho}(\hat J, \hat J') = 0 \qquad \text{for all } \hat J' \in \Omega^{0,1}(M,TM) \text{ with } \bar\partial_J \hat J' = 0$$

if and only if $\hat J = \mathcal{L}_v J$ for some $v \in Vect(M)$ with $d\iota(v)\rho = d\iota(Jv)\rho = 0$.

Proof. Let $v \in \mathrm{Vect}(M)$ with $d\iota(v)\rho = d\iota(Jv)\rho = 0$ and let $\hat{J} \in \Omega^{0,1}(M, TM)$ with $\bar{\partial}_J \hat{J} = 0$. With $f, g : M \to \mathbb{R}$ defined by (6.30), it follows from (6.25) that

$$\begin{aligned}\Omega_{\rho,J}(\hat{J}, \mathcal{L}_v J) &= \int_M \Lambda_\rho(J, \hat{J}) \wedge \iota(v)\rho = \int_M (-df \circ J + dg) \wedge \iota(v)\rho \\ &= \int_M f d\iota(Jv)\rho - g d\iota(v)\rho = 0.\end{aligned}$$

This shows that $\mathcal{L}_v J$ is in the kernel of $\Omega_{\rho,J}$ restricted to $T_J \mathcal{J}_{\mathrm{int}}(M)$. Conversely, let $\hat{J} \in \Omega^{0,1}(M, TM)$ be given with $\bar{\partial}_J \hat{J} = 0$ and $\Omega_J(\hat{J}, \hat{J}') = 0$ for all $\hat{J}' \in \Omega^{0,1}_J(M, TM)$ with $\bar{\partial}_J \hat{J}' = 0$. Choose $v \in \mathrm{Vect}(M)$ with $\bar{\partial}_J^*(\hat{J} - \mathcal{L}_v J) = 0$. Then $\bar{\partial}_J(\hat{J} - \mathcal{L}_v J)^* = 0$ by Lemma 6.3.4 below. Hence there exists a 2-form $\sigma \in \Omega^{0,2}_J(M, TM)$ with $\bar{\partial}_J^* \sigma = (\hat{J} - \mathcal{L}_v J)^*$. It follows $\bar{\partial}_J \bar{\partial}_J^* \sigma = 0$ and hence $\bar{\partial}_J^* \sigma = 0$. This shows $\hat{J} = \mathcal{L}_v J$. Moreover, (6.28) yields

$$\Omega_{\rho,J}(\mathcal{L}_u J, \mathcal{L}_v J) = \int_M (2\mathrm{Ric}_{\rho,J}(u, v) + f_u f_{Jv} - f_{Ju} f_v)\, \rho \tag{6.33}$$

for all $u \in \mathrm{Vect}(M)$. Since $\mathrm{Ric}_{\rho,J} = 0$, (6.33) vanishes for all $u \in \mathrm{Vect}(M)$ if and only if $f_v = f_{Jv} = 0$. This completes the proof of the proposition. □

Lemma 6.3.4. *Let (M, J, ω) be a closed $2n$-dimensional Kähler manifold. Define $\rho := \frac{\omega^n}{n!}$ and assume that $\mathrm{Ric}_{\rho,J} = 0$. Let $\hat{J} \in \Omega^{0,1}_J(M, TM)$ such that $\bar{\partial}_J \hat{J} = 0$ and $\bar{\partial}_J^* \hat{J} = 0$. Then $\bar{\partial}_J \hat{J}^* = 0$ and $\bar{\partial}_J^* \hat{J}^* = 0$.*

Sketch of proof: The proof is inspired by [40]. Choose a hermitian line bundle $L \to M$ with $c_1(L) = c_1(TM, J)$, a Hermitian connection ∇_L, and an n-form $\theta \in \Omega^{0,n}_J(M, L)$ which satisfies $d^{\nabla_L}\theta = 0$ and $c_n \langle \theta \wedge \theta \rangle = \rho$ where $c_n = 1$ if n is even and $c_n = -\mathbf{i}$ when n is odd. Now there exists for every $\hat{J} \in \Omega^{0,1}_J(M, TM)$ a unique $\beta = \beta_{\hat{J}} \in \Omega^{n-1,1}_J(M, L)$ satisfying

$$\mathbf{i}\iota(u)\beta - \iota(Ju)\beta = \iota(\hat{J}u)\theta \qquad \text{for all } u \in \mathrm{Vect}(\Sigma). \tag{6.34}$$

Then one can show that

$$\begin{aligned}\bar{\partial}_J^* \hat{J} = 0 \qquad &\Longleftrightarrow \qquad (\bar{\partial}_J^{\nabla_L})^* \beta = 0 \\ \bar{\partial}_J \hat{J} = 0 \qquad &\Longleftrightarrow \qquad \bar{\partial}_J^{\nabla_L} \beta = 0\end{aligned}$$

and $\mathrm{Ric}_{\rho,J} = \mathbf{i}F^{\nabla_L}$. Thus $\mathrm{Ric}_{\rho,J} = 0$, $\bar{\partial}_J \hat{J} = 0$ and $\bar{\partial}_J^* \hat{J} = 0$ imply $\mathbf{i}F^{\nabla_L} = 0$, $\bar{\partial}_J^{\nabla_L} \beta = 0$ and $(\bar{\partial}_J^{\nabla_L})^* \beta = 0$. It now follows from the Akizuki–Nakano theorem (see [29], Chapter VII) that

$$(\partial_J^{\nabla_L})^* \partial_J^{\nabla_L} \beta + \partial_J^{\nabla_L} (\partial_J^{\nabla_L})^* \beta = (\bar{\partial}_J^{\nabla_L})^* \bar{\partial}_J^{\nabla_L} \beta + \bar{\partial}_J^{\nabla_L} (\bar{\partial}_J^{\nabla_L})^* \beta = 0.$$

This shows $(\partial_J^{\nabla_L})^* \beta = 0$ and $\partial_J^{\nabla_L} \beta = 0$ and therefore $(\bar{\partial}_J^{\nabla_L})^*(*\beta) = 0$ and $\bar{\partial}_J^{\nabla_L}(*\beta) = 0$. Since $\mathbf{i}\iota(u) * \beta - \iota(Ju) * \beta = -c_n \iota(\hat{J}^* u)\theta$ it follows that $*\beta$ corresponds to $\hat{J}^*$ up to the factor $-c_n$. Hence $\bar{\partial}_J \hat{J}^* = 0$ and $\bar{\partial}_J^* \hat{J}^* = 0$ and this completes the proof. □

6.3.3 The Weil–Petersson symplectic form

Denote the space of integrable complex structures on M with vanishing real first Chern class and nonempty Kähler cone by

$$\mathcal{J}_{\mathrm{int},0}(M) := \left\{ J \in \mathcal{J}_{\mathrm{int}}(M) \;\middle|\; \begin{array}{c} c_1(TM,J) = 0 \in H^2(M,\mathbb{R}) \\ \text{and } J \text{ admits a Kähler form} \end{array} \right\} \tag{6.35}$$

Fix a volume form $\rho \in \Omega^{2n}(M)$ and denote the space of Ricci-flat complex structures by

$$\mathcal{J}_{\mathrm{int},0}(M,\rho) := \{ J \in \mathcal{J}_{\mathrm{int},0}(M) \,|\, \mathrm{Ric}_{\rho,J} = 0 \}. \tag{6.36}$$

The inclusion $\mathcal{J}_{\mathrm{int},0}(M,\rho) \subset \mathcal{J}_{\mathrm{int},0}(M)$ yields a bijection between the spaces

$$\mathcal{T}_0(M) := \mathcal{J}_{\mathrm{int},0}(M)/\mathrm{Diff}_0(M) \tag{6.37}$$

$$\mathcal{T}_0(M,\rho) := \mathcal{J}_{\mathrm{int},0}(M,\rho)/\mathrm{Diff}_0(M,\rho) \tag{6.38}$$

by Theorem 6.2.3. It follows from Theorem 6.2.3 and (6.20) that for every $J \in \mathcal{J}_{\mathrm{int},0}(M)$ there exists a unique volume form $\rho_J \in \Omega^{2n}(M)$ such that

$$\mathrm{Ric}_{\rho_J,J} = 0 \quad \text{and} \quad \int_M \rho_J = V := \int_M \rho. \tag{6.39}$$

By Yau's theorem [124, 125], there exists a Kähler form $\omega \in \Omega^2(M)$ with $\frac{\omega^n}{n!} = \rho_J$. Hence we are in position to apply the various results on Ricci-flat Kähler manifolds. By Propositon 6.3.2, $\mathrm{Diff}_0(M,\rho)/\mathrm{Diff}^{\mathrm{ex}}(M,\rho)$ acts trivially on

$$\mathcal{T}_0^{\mathrm{ex}}(M,\rho) := \mathcal{J}_{\mathrm{int},0}(M,\rho)/\mathrm{Diff}^{\mathrm{ex}}(M,\rho). \tag{6.40}$$

Therefore, by Theorem 6.2.1, $\mathcal{T}_0(M) \cong \mathcal{T}_0^{\mathrm{ex}}(M,\rho)$ embeds into the Marsden–Weinstein quotient of $\mathcal{J}(M)$ by $\mathrm{Diff}^{\mathrm{ex}}(M,\rho)$. The symplectic form on this quotient is obtained by restriction of the symplectic form $\Omega_{J,\rho}$ defined by (6.19). It follows from Proposition 6.3.3 that this yields a non-degenerated symplectic form on Teichmüller space.

Theorem 6.3.5 (**The Weil–Petersson symplectic form**). *Let M be a closed connected orientend $2n$-dimensional manifold and fix $V > 0$. For a complex structure $J \in \mathcal{J}_{int,0}(M)$ define $\rho_J \in \Omega^{2n}(M)$ by (6.39). For $\hat{J}_1, \hat{J}_2 \in \Omega^{0,1}_J(M,TM)$ with $\bar{\partial}_J \hat{J}_i = 0$ define $f_i, g_i : M \to \mathbb{R}$ as in Proposition 6.3.2. The Weil–Petersson symplectic form on $\mathcal{T}_0(M)$ is then given by*

$$\Omega_J(\hat{J}_1,\hat{J}_2) = \int_M \left(\frac{1}{2} tr\left(\hat{J}_1 J \hat{J}_2 \right) - f_1 g_2 + f_2 g_1 \right) \rho_J. \tag{6.41}$$

This symplectic form is $\mathit{Diff}_0(M)$ equivariant and thus the mapping class group acts on $\mathcal{T}_0(M)$ by symplectomorphism.

Proof. It follows from (6.20) and (6.24) that the Weil–Petersson form (6.41) is $\mathrm{Diff}_0(M)$-equivariant. We show next that it is descends to the quotient

$$T_{[J]}\mathcal{T}_0(M) = \frac{\{ \hat{J} \in \Omega^{0,1}_J \,|\, \bar{\partial}_J \hat{J} = 0 \}}{\{ \mathcal{L}_v J \,|\, v \in \mathrm{Vect}(M) \}}. \tag{6.42}$$

By equivariance, we may assume in the following that $\mathrm{Ric}_{\rho,J} = 0$ and $\rho = \rho_J$. Then follows from (6.28)

$$\Lambda_\rho(J, \mathcal{L}_v J) = -df_v \circ J + df_{Jv} \tag{6.43}$$

for every $v \in \mathrm{Vect}(M)$, where $f_v\rho = d\iota(v)\rho$ and $f_{Jv}\rho = d\iota(Jv)\rho$. Now let $\hat{J} \in \Omega^{0,1}_J(M, TM)$ be given and define $f, g : M \to \mathbb{R}$ as in Proposition 6.3.2. Then follows from (6.25)

$$\begin{aligned}\Omega_J(\hat{J}, \mathcal{L}_v J) &= \int_M \Lambda_\rho(J, \hat{J}) \wedge \iota(v)\rho - gd\iota(v)\rho + fd\iota(Jv)\rho \\ &= \int_M -df(Jv) + dg(v)\rho + dg \wedge \iota(v)\rho - df \wedge \iota(Jv)\rho = 0\end{aligned}$$

Hence (6.41) descends to a well-defined 2-form on the quotient space $T_{[J]}\mathcal{T}_0(M)$. It follows from Proposition 6.3.3 that it is non-degenerated. Finally, consider the quotient

$$T_{[J]}\mathcal{T}_0(M, \rho) = \frac{\{\hat{J} \in \Omega^{0,1}_J \,|\, \bar{\partial}_J \hat{J} = 0, \widehat{Ric}_\rho(J, \hat{J}) = 0\}}{\{\mathcal{L}_v J \,|\, v \in \mathrm{Vect}(M) \text{ with } d\iota(v)\rho = 0\}} \tag{6.44}$$

where $\widehat{Ric}_\rho(J, \hat{J})$ is the linearization of $\mathrm{Ric}_{\rho,J}$, when varying J in direction $\hat{J}$. It follows from (6.23) and (6.43) that

$$\Omega_J(\hat{J}_1, \hat{J}_2) = \int_M \frac{1}{2}\mathrm{tr}\left(\hat{J}_1 J \hat{J}_2\right)\rho$$

for $\hat{J}_i \in \Omega^{0,1}_J(M, TM)$ with $\bar{\partial}_J \hat{J} = 0$ and $\widehat{Ric}_\rho(J, \hat{J}) = 0$. Hence the Weil–Petersson symplectic form (6.41) corresponds on $\mathcal{T}_0(M, \rho)$ to the canonical symplectic form on the Marsden–Weinstein quotient and it is therefore closed. □

Let $\omega \in \Omega^2(M)$ be a symplectic form with vanishing real first Chern. Define

$$\begin{aligned}\mathcal{J}_{\mathrm{int},0}(M, \omega) &:= \{J \in \mathcal{J}_{\mathrm{int},0}(M) \,|\, J \text{ is compatible with } \omega\} \\ \mathcal{T}(M, \omega) &:= \mathcal{J}_{\mathrm{int},0}(M, \omega)/\left(\mathrm{Diff}_0(M) \cap \mathrm{Symp}(M, \omega)\right)\end{aligned} \tag{6.45}$$

Then $\mathcal{T}(M, \omega) \subset \mathcal{T}_0(M)$ embeds as a complex submanifold with respect to the complex structure $\hat{J} \mapsto -J\hat{J}$. The Weil–Petersson symplectic form restrics to a Kähler form along $\mathcal{T}(M, \omega)$ and the symmetric form

$$\left\langle \hat{J}_1, \hat{J}_2 \right\rangle := \int_M \left(\frac{1}{2}\mathrm{tr}\left(\hat{J}_1 \hat{J}_2\right) - f_1 f_2 - g_1 g_2\right)\rho_J \tag{6.46}$$

is positive definite on $T_{[J]}\mathcal{T}(M, \omega)$ and negative definite on its symplectic complement in $T_{[J]}\mathcal{T}(M)$.

6.3.4 A symplectic connection

Denote by $\mathcal{S}_0(M) \subset \Omega^2(M)$ the space of symplectic forms on M with vanishing first Chern class which admit an integrable compatible complex structure. Define

$$\mathcal{E}_0(M) := \left\{ (\omega, J) \;\middle|\; \begin{array}{c} \omega \in \mathcal{S}_0(M),\ J \in \mathcal{J}_{\mathrm{int}}(M), \\ J \text{ compatible with } \omega \text{ and } \mathrm{Ric}_{\omega,J} = 0 \end{array} \right\} \Big/ \mathrm{Diff}_0(M) \tag{6.47}$$

and

$$\mathcal{B}_0(M) := \mathcal{S}_0(M)/\mathrm{Diff}_0(M). \tag{6.48}$$

Then $\mathcal{E}_0(M)$ and $\mathcal{B}_0(M)$ are finite dimensional manifolds and the canonical projection $\mathcal{E}_0(M) \to \mathcal{B}_0(M)$ is a surjective submersion. The fibre over $[\omega] \in \mathcal{B}_0(M)$ is the Teichmüller space $\mathcal{T}_0(M, \omega)$ defined by (6.45). Hence the pullback of the Weil–Petersson symplectic form defines a closed 2-form on $\mathcal{E}_0(M)$ which is non-degenerated along the fibres. Such a 2-form defines a symplectic connection on the fibration. The horizontal subspaces of this connection are defined as the symplectic complements of the vertical subspaces along the fibres. This can be described by a $\mathrm{Diff}_0(M)$-equivariant family of 1-forms

$$\mathcal{A}_{\omega,J} : \Omega^2_{\mathrm{cl}}(M) \to \Omega^{0,1}_J(M, TM) \tag{6.49}$$

where $\Omega^2_{\mathrm{cl}}(M)$ denotes the space of closed 2-forms on M. The horizontal lift of $[\hat\omega] \in T_{[\omega]}\mathcal{B}_0(M)$ is then given by $[(\hat\omega, \mathcal{A}_{\omega,J}(\hat\omega))] \in T_{[\omega,J]}\mathcal{E}_0(M)$.

Theorem 6.3.6 (A symplectic connection). *Let (ω, J) be a Ricci-flat Kähler structure on M. For every closed 2-form $\hat\omega$, there exists a unique element $\hat J = \mathcal{A}_{\omega,J}(\hat\omega) \in \Omega^{0,1}_J(M, TM)$ satisfying*

$$\Omega_J(\hat J, \hat J') = 0 \qquad \text{for all } \hat J' \in \Omega^{0,1}_J(M, TM) \text{ with } \bar\partial_J \hat J' = 0 \text{ and } \hat J' = (\hat J')^*$$

and the tangency conditions

$$\bar\partial_J \hat J = 0, \quad \Lambda_\rho(J, \hat J) = -d\langle \hat\omega, \omega\rangle \circ J, \quad \hat\omega(\cdot,\cdot) - \hat\omega(J\cdot, J\cdot) = \langle (\hat J - \hat J^*)\cdot, \cdot\rangle.$$

This connection is $\mathrm{Diff}_0(M)$-equivariant and satisfies $\mathcal{A}_{\omega,J}(d\iota(v)\omega) = \mathcal{L}_v J$ for all $v \in \mathrm{Vect}(M)$ with $d\iota(Jv)\rho = 0$.

Proof. We omit the proof. We give a complete proof of this theorem, including a calculation of the curvature of this connection, in [50] Theorem 4.3. □

6.4 Kähler–Einstein manifolds

Let (M, ρ) be a $2n$-dimensional closed oriented manifold with volume form ρ. For $a \in H^2(M, \mathbb{R})$ denote by

$$\mathcal{S}_a(M, \rho) := \left\{ \omega \in \Omega^2(M) \;\middle|\; d\omega = 0,\ [\omega] = a,\ \frac{\omega^n}{n!} = \rho \right\} \tag{6.50}$$

the space of symplectic forms in the cohomology class of a with volume form ρ. Let $c \in H^2(M, \mathbb{R})$ with $2\pi c = \kappa a$ for some $\kappa \in \mathbb{R}$ and denote by

$$\mathcal{J}_c(M) := \{J \in \mathcal{J}(M) \mid c_1(TM, J) = c\} \tag{6.51}$$

the space of almost complex structures with first Chern class c. Ultimately one is interested in the subspace of compatible pairs (J, ω) which implies the additional constraint $c_1(\omega) = c$. It is convenient for our discussion and for the derivation of the moment map equation to omit this assumption in the general setup. The Kähler–Einstein equation for $(J, \omega) \in \mathcal{J}_c(M) \times \mathcal{S}_a(M, \rho)$ is given by

$$\mathrm{Ric}_{\omega,J} - \kappa\omega = 0. \tag{6.52}$$

The goal of this section is to provide a moment map description for the Kähler–Einstein equation. We call $a \in H^2(M, \mathbb{R})$ a Lefschetz class when $\cdot\cup a^{n-1} : H^1(M, \mathbb{R}) \to H^{2n-1}(M, \mathbb{R})$ is an isomorphism. The starting point is a moment map description for the volume constraint $\frac{\omega^n}{n!} = \rho$ in Proposition 6.4.1. This gives rise to a natural symplectic structure on $\mathcal{S}_a(M, \rho)$ when a is Lefschetz class. The left-hand side of (6.52) yields then a moment map for the action of $\mathrm{Diff}^{\mathrm{ex}}(M, \rho)$ on $\mathcal{J}_c(M) \times \mathcal{S}_a(M, \rho)$ for a suitably weighted product symplectic form. This leads to a Weil–Petersson metric on the Teichmüller space of Kähler–Einstein manifolds with a fixed symplectic form $\omega \in \mathcal{S}_a(M)$.

6.4.1 The volume form as moment map

Assume first that $a \in H^2(M, 2\pi\mathbb{Z})$ and let $P \to M$ be a principal S^1 bundle with $2\pi c_1(P) = a$. Denote by

$$\mathcal{A}_{\mathrm{Symp}}(P) := \{A \in \mathcal{A}(P) \mid (\mathbf{i}F_A)^n \neq 0\} \tag{6.53}$$

the space of S^1-connections on P with symplectic curvature. The next proposition shows that the volume form $\frac{(\mathbf{i}F_A)^n}{n!}$ provides a moment map for the action of the gauge group.

Proposition 6.4.1 (The volume form as moment map). *The 2-form*

$$\Omega_A(\hat{A}_1, \hat{A}_2) := \int_M \hat{A}_1 \wedge \hat{A}_2 \wedge \frac{(\boldsymbol{i}F_A)^{n-1}}{(n-1)!} \tag{6.54}$$

is a symplectic form $\mathcal{A}_{Symp}(P)$. The action of the gauge group $\mathcal{G}(P)$ on $\mathcal{A}_{Symp}(P)$ is Hamiltonian with respect to this symplectic structure with moment map

$$\langle \mu(A), \xi \rangle = \int_M Im(\xi) \frac{(\boldsymbol{i}F_A)^n}{n!} \tag{6.55}$$

for all $A \in \mathcal{A}_{Symp}(P)$ and $\xi \in \Omega^0(M, \boldsymbol{i}\mathbb{R})$.

Proof. The exterior differential of (6.54) is given by Cartan's formula as

$$\begin{aligned}(d\Omega)_A(\hat{A}_1,\hat{A}_2,\hat{A}_3) &= \int_M \hat{A}_2\wedge\hat{A}_3\wedge d\mathbf{i}\hat{A}_1\wedge\frac{(\mathbf{i}F_A)^{n-2}}{(n-2)!}\\ &\quad+\int_M \hat{A}_3\wedge\hat{A}_1\wedge d\mathbf{i}\hat{A}_2\wedge\frac{(\mathbf{i}F_A)^{n-2}}{(n-2)!}\\ &\quad+\int_M \hat{A}_1\wedge\hat{A}_2\wedge d\mathbf{i}\hat{A}_3\wedge\frac{(\mathbf{i}F_A)^{n-2}}{(n-2)!}\\ &= \mathbf{i}\int_M d\left(\hat{A}_1\wedge\hat{A}_2\wedge\hat{A}_3\right)\wedge\frac{(\mathbf{i}F_A)^{n-2}}{(n-2)!}=0\end{aligned}$$

Hence (6.54) defines a closed 2-form and it is clearly non-degenerated, since $(\mathbf{i}F_A)^n \neq 0$. This proves that (6.54) defines indeed a symplectic form on $\mathcal{A}_{\mathrm{Symp}}(P)$. Next, differentiating the moment map equation yields

$$\langle d\mu(A)\hat{A},\xi\rangle = \int_M \xi d\hat{A}\wedge\frac{(\mathbf{i}F_A)^{n-1}}{(n-1)!} = \int_M -d\xi\wedge\hat{A}\wedge\frac{(\mathbf{i}F_A)^{n-1}}{(n-1)!} = \Omega_A(-d\xi,\hat{A}).$$

This completes the proof of the proposition, since $-d\xi$ is the infinitesimal action of A for the left action of the gauge group. □

We call $a=[\omega]\in H^2(M,\mathbb{R})$ a **Lefschetz class** when $a^n\neq 0$ and

$$L_a^{n-1}: H^1(M)\to H^{2n-1}(M),\quad [\lambda]\mapsto[\lambda]\cup a^{n-1}=[\lambda\wedge\omega^{n-1}] \tag{6.56}$$

is an isomorphism. The hard Lefschetz theorem asserts that every Kähler class is Lefschetz, while the opposite is generally not true (see [123, 45, 84] and the references therein). The Lefschetz condition is needed in order to obtain a symplectic form of the space $\mathcal{S}_a(M,\rho)$: By Proposition 6.4.1, it follows that

$$\mathcal{M} := \left\{A\in\mathcal{A}_{Symp}(M)\;\middle|\;\frac{(\mathbf{i}F_A)^n}{n!}=\rho\right\}\Big/\mathcal{G}(P) \tag{6.57}$$

is a Marsden–Weinstein quotient for the action of the gauge group and thus carries a natural symplectic structure induced by (6.54). This fibres over the space $\mathcal{S}_a(M,\rho)$ of symplectic forms where the projection map is defined by

$$\pi:\mathcal{M}\to\mathcal{S}_a(M,\rho),\qquad \pi([A]):=\mathbf{i}F_A. \tag{6.58}$$

The tangent spaces of $\mathcal{M}$ and $\mathcal{S}_a(M,\rho)$ are given by

$$T_{[A]}\mathcal{M} = \frac{\{\hat{A}\in\Omega^1(M,\mathbf{i}\mathbb{R})\,|\,d\hat{A}\wedge F_A^{n-1}=0\}}{\{d\xi\,|\,\xi\in\Omega^0(M,\mathbf{i}\mathbb{R})\}} \tag{6.59}$$

$$T_\omega\mathcal{S}_a(M,\rho) := \{\hat{\omega}\in\Omega^2(M)\,|\,\hat{\omega}\text{ is exact, }\hat{\omega}\wedge\omega^{n-1}=0\}. \tag{6.60}$$

and the derivative of (6.58) is

$$d\pi([A]): T_{[A]}\mathcal{M}\to T_{\mathbf{i}F_A}\mathcal{S}_a(M,\rho),\qquad d\pi([A])[\hat{A}]=\mathbf{i}d\hat{A}$$

Hence, the fibres $p^{-1}(\omega) \cong H^1(M, \mathbf{i}\mathbb{R})$ are symplectic submanifolds of $\mathcal{M}$ if and only if a is a Lefschetz class. Under this assumption, the symplectic form on $\mathcal{M}$ induces a natural symplectic form on $\mathcal{S}_a(M, \rho)$ which we describe in the next lemma. It turns out that the formula for the Lefschetz symplectic form remains well-defined for any Lefschetz class a and does not require rationality.

Lemma 6.4.2 (The Lefschetz symplectic form). *Suppose $a \in H^2(M, \mathbb{R})$ is a Lefschetz class.*

1. *Let $\omega \in \mathcal{S}_a(M, \rho)$ and $\hat{\omega} \in \Omega^2(M)$ be exact with $\hat{\omega} \wedge \omega^{n-1} = 0$. Then there exists $\lambda \in \Omega^1(M)$ such that*

$$d\lambda = \hat{\omega} \qquad \text{and} \qquad \lambda \wedge \omega^{n-1} \text{ is exact.} \tag{6.61}$$

   If $\lambda_1, \lambda_2 \in \Omega^1(M)$ are two solutions of (6.61), then $\lambda_1 - \lambda_2$ is exact.

2. *The Lefschetz symplectic form on $\mathcal{S}_a(M, \rho)$ is defined by*

$$\Omega_\omega(\hat{\omega}_1, \hat{\omega}_2) = \int_M \lambda_1 \wedge \lambda_2 \wedge \frac{\omega^{n-1}}{(n-1)!} \tag{6.62}$$

   for $\omega \in \mathcal{S}_a(M, \rho)$ and $\hat{\omega}_i \in \Omega^2(M)$ exact with $\hat{\omega}_i \wedge \omega^{n-1} = 0$, where $\lambda_i \in \Omega^1(M)$ satisfy (6.61).

Proof. Let $\hat{\omega} = d\lambda_0 \in \Omega^2(M)$ be an exact 2-form with $\hat{\omega} \wedge \omega^{n-1} = 0$. Then $\tau := \lambda_0 \wedge \omega^{n-1}$ is closed and surjectivity of (6.56) implies that there exists a closed $\lambda_1 \in \Omega^1(M)$ such that $\tau - \lambda_1 \wedge \omega^{n-1}$ is exact. Hence $\lambda := \lambda_0 - \lambda_1$ satisfies (6.61). Injectivity of (6.56) shows that any two solutions of (6.61) differ by an exact 1-form. This proves the first part.

Suppose $\lambda \in \Omega^1(M)$ satisfies (6.61) and let $f \in \Omega^0(M)$. Then

$$\int_M df \wedge \lambda \wedge \frac{\omega^{n-1}}{(n-1)!} = -\int_M f \wedge d\lambda \wedge \frac{\omega^{n-1}}{(n-1)!} = 0$$

shows that (6.62) is well-defined. Its exterior differential is given by

$$\begin{aligned}
(d\Omega)_\omega(\hat{\omega}_1, \hat{\omega}_2, \hat{\omega}_3) &= \int_M \lambda_2 \wedge \lambda_3 \wedge \hat{\omega}_1 \wedge \frac{\omega^{n-2}}{(n-2)!} \\
&\quad + \int_M \lambda_3 \wedge \lambda_1 \wedge \hat{\omega}_2 \wedge \frac{\omega^{n-2}}{(n-2)!} \\
&\quad + \int_M \lambda_1 \wedge \lambda_2 \wedge \hat{\omega}_3 \wedge \frac{\omega^{n-2}}{(n-2)!} \\
&= \int_M d\left(\lambda_1 \wedge \lambda_2 \wedge \lambda_3\right) \wedge \frac{\omega^{n-2}}{(n-2)!} = 0
\end{aligned}$$

and hence (6.62) is closed. Finally, let $\lambda \in \Omega^1(M)$ satisfy (6.61) and suppose

$$\int_M \lambda \wedge \lambda' \wedge \frac{\omega^{n-1}}{(n-1)!} = 0 \qquad \text{for all } \lambda' \in \Omega^1(M) \text{ satisfying (6.61).}$$

It follows from linear algebra that every $(2n-1)$-form on M can be written as $\lambda'' \wedge \omega^{n-1}$ and therefore

$$\int_M \lambda \wedge \eta = 0 \qquad \text{for all exact } \eta \in \Omega^{2n-1}(M).$$

This implies that λ is closed and thus (6.62) is non-degenerated. □

We thus have shown that the space of symplectic forms $\mathcal{S}_a(M,\rho)$ is a symplectic manifold. This is the main ingredient for the moment map interpretation of the Kähler–Einstein equations in the next subsection.

6.4.2 The Kähler–Einstein condition as moment map

Let $c \in H^2(M,\mathbb{R})$ with $2\pi c = \kappa a$ for some $\kappa \in \mathbb{R}$ and define $\mathcal{J}_c(M)$ by (6.51). Equip the product space $\mathcal{J}_c(M) \times \mathcal{S}_a(M,\rho)$ with the symplectic form

$$\Omega_{J,\omega}\Big((\hat{J}_1,\hat{\omega}_1),(\hat{J}_1,\hat{\omega}_1)\Big) = \int_M \frac{1}{2}\mathrm{tr}\left(\hat{J}_1 J \hat{J}_2\right)\rho - 2\kappa\lambda_1 \wedge \lambda_2 \wedge \frac{\omega^{n-1}}{(n-1)!} \tag{6.63}$$

where $\lambda_i \in \Omega^1(M)$ satisfy (6.61).

Theorem 6.4.3 (**Kähler–Einstein pairs**). *Let $a \in H^2(M,\mathbb{R})$ be a Lefschetz class, such that (6.56) is an isomorphism. Let $\kappa \in \mathbb{R}$ and $c \in H^2(M,\mathbb{R})$ with $2\pi c = \kappa a$. Then the action of $Diff^{ex}(M,\rho)$ on $\mathcal{J}_c(M) \times \mathcal{S}_a(M,\rho)$ is Hamiltonian with respect to (6.63) and with moment map*

$$\underline{\mu} : \mathcal{J}_c(M) \times \mathcal{S}_a(M,\rho) \to \Omega^2_{ex}(M), \qquad \underline{\mu}(J,\omega) = 2\left(Ric_{\rho,J} - \kappa\omega\right). \tag{6.64}$$

This map is equivariant and takes values in the space of exact 2-forms. For every $v \in Vect(M)$ and $\alpha_v \in \Omega^{2n-2}(M)$ with $d\alpha_v = \iota(v)\rho$ is holds

$$\partial_t \int_M 2\left(Ric_{\rho,J_t} - \kappa\omega_t\right) \wedge \alpha_v = \Omega_{J_t,\omega_t}\left((\partial_t J_t, \partial_t\omega_t),(\mathcal{L}_v J_t, d\iota(v)\omega_t)\right) \tag{6.65}$$

for every smooth path $\mathbb{R} \to \mathcal{S}_a(M,\rho) \times \mathcal{J}_c(M)$, $t \mapsto (J_t,\omega_t)$.

Proof. For $(J,\omega) \in \mathcal{J}_c(M) \times \mathcal{S}_a(M,\rho)$, it holds

$$[\mathrm{Ric}_{\rho,J} - \kappa\omega] = 2\pi c - \kappa a = 0 \in H^2(M,\mathbb{R})$$

and hence $\underline{\mu}(J,\omega)$ is exact. Equivariance of $\underline{\mu}(J,\omega)$ follows from (6.20). Next, let $v \in \mathrm{Vect}(M)$ and $\alpha_v \in \Omega^{2n-2}(M)$ with $d\alpha_v = \iota(v)\rho$ be given. For $\omega \in \mathcal{S}_a(M,\rho)$ and $\hat{\omega} \in \Omega^2(M,\mathbb{R})$ exact with $\hat{\omega} \wedge \omega^{n-1} = 0$ choose $\lambda \in \Omega^1(M)$ satisfying (6.61). Then follows

$$\int_M \hat{\omega} \wedge \alpha_v = \int_M d\lambda \wedge \alpha_v = \int_M \lambda \wedge \iota(v)\omega \wedge \frac{\omega^{n-1}}{(n-1)!} = \Omega_\omega(\hat{\omega}, d\iota(v)\omega).$$

The moment map equation (6.65) follows from this and Theorem 6.2.3. □

Remark 6.4.4 (A variant using the extended gauge group.). There is a variant of this story which yields the volume constraint $\omega^n/n! = \rho$ and the Kähler–Einstein equation $\mathrm{Ric}_{\rho,J} - \kappa\omega = 0$ simultaneously as moment map equations. For this assume that $2\pi a \in H^2(M,\mathbb{Z})$ and let $P \to M$ be the principal S^1 bundle with Chern class $2\pi c_1(P) = a$. Denote by $\tilde{\mathcal{G}}(P)$ the group of all bundle automorphism of P which cover exact volume preserving diffeomorphism on (M,ρ). One can then verify that the action of $\tilde{\mathcal{G}}(P)$ on the product space $\mathcal{J}_c(M) \times \mathcal{A}_{\mathrm{Symp}}(P)$ is Hamiltonian for a suitably weighted product symplectic from and it is generated by the moment map

$$\langle \underline{\mu}(J,A), V\rangle = \int_M 2\left(\mathrm{Ric}_{\rho,J} - \kappa(\mathbf{i}F_A)\right) \wedge \alpha_v - 2\mathbf{i}\kappa \int_M A(V)\left(\frac{(\mathbf{i}F_A)^n}{n!} - \rho\right)$$

for $(J,A) \in \mathcal{J}_c(M) \times \mathcal{A}_{\mathrm{Symp}}(P)$ and $V \in \mathrm{Lie}(\tilde{\mathcal{G}}(P)) \subset \mathrm{Vect}(P)$, where $A(V) \in \Omega^0(M,\mathbf{i}\mathbb{R})$ denotes its A-vertical part, $v \in \mathrm{Vect}(M)$ is the exact divergence free vector field covered by V, and $\alpha_v \in \Omega^{2n-2}(M)$ satisfies $d\alpha_v = \iota(v)\rho$.

Denote the subspace of compatible pairs in $\mathcal{J}_c(M) \times \mathcal{S}_a(M,\rho)$ by

$$\mathcal{C}_{c,a}(M,\rho) := \{(J,\omega) \in \mathcal{J}_c(M) \times \mathcal{S}_a(M,\rho) \,|\, J \text{ compatible with } \omega\}. \tag{6.66}$$

Differentiating the compatibility condition shows that

$$T_{(J,\omega)}\mathcal{C}_{c,a}(M,\rho) := \left\{(\hat{J},\hat{\omega}) \,|\, \hat{\omega}(\cdot,\cdot) - \hat{\omega}(J\cdot,J\cdot) = \langle(\hat{J} - \hat{J}^*)\cdot,\cdot\rangle\right\}.$$

Lemma 6.4.5 (Compatible pairs). *Equip $\mathcal{J}_c(M) \times \mathcal{S}_a(M,\rho)$ with the symplectic form (6.63) and let $(J,\omega) \in \mathcal{C}_{c,a}(M,\rho)$. The kernel of the restriction of the symplectic form to $T_{(J,\omega)}\mathcal{C}_{c,a}(M,\rho)$ is the subspace*

$$\left\{(\hat{J},\hat{\omega}) \in T_J\mathcal{J}_c(M) \times \mathcal{S}_a(M,\rho) \;\middle|\; \begin{array}{c} \hat{J}^* = -\hat{J},\ \widehat{Ric}_\rho(J,\hat{J}) - \kappa\hat{\omega} = 0 \\ \hat{\omega}(\cdot,\cdot) - \hat{\omega}(J\cdot,J\cdot) = \langle(\hat{J} - \hat{J}^*)\cdot,\cdot\rangle \end{array}\right\}$$

where $\hat{J}^*$ is the adjoint with respect to the metric $\langle\cdot,\cdot\rangle = \omega(\cdot,J\cdot)$.

Proof. First, suppose that $(\hat{J},\hat{\omega}) \in T_{(J,\omega)}\mathcal{C}_{c,a}(M,\rho)$ satisfies

$$\Omega_{J,\omega}\Big((\hat{J},\hat{\omega}),\, (\hat{J}',\hat{\omega}')\Big) = 0 \qquad \text{for all } (\hat{J}',\hat{\omega}') \in T_{(J,\omega)}\mathcal{C}_{c,a}(M,\rho)$$

For every $v \in \mathrm{Vect}(M)$ with $\iota(v)\rho = d\alpha_v$, Theorem 6.4.3 shows

$$\Omega_{J,\omega} = \Big((\hat{J},\hat{\omega}),\, \mathcal{L}_v(J,\omega)\Big) = 2\int_M \Big(\widehat{Ric}_\rho(J,\hat{J}) - \kappa\hat{\omega}\Big) \wedge \alpha_v.$$

This implies $\widehat{Ric}_\rho(J,\hat{J}) - \kappa\hat{\omega} = 0$. For every self-adjoint $(\hat{J}')^* = \hat{J}'$, we have

$$\Omega_{(J,\omega)}\Big((\hat{J},\hat{\omega}),\, (\hat{J}',0)\Big) = \frac{1}{2}\int_M \mathrm{tr}\left(\hat{J}J\hat{J}'\right)\rho.$$

This implies $\hat{J} = -\hat{J}^*$ and proves one inclusion.

For the other inclusion, we show that every $(\hat{J}', \hat{\omega}') \in T_{(J,\omega)}\mathcal{C}_{c,a}(M,\rho)$ can be written as

$$(\hat{J}', \hat{\omega}') = \mathcal{L}_v(J,\omega) + (\hat{J}'', 0)$$

where $v \in \mathrm{Vect}(M,\rho)$ is an exact divergence free vector field and $(\hat{J}'')^* = \hat{J}''$. Indeed, by Lemma 6.4.2 there exists $\lambda \in \Omega^1(M)$ such that $\hat{\omega}' = d\lambda$ and $\lambda \wedge \omega^{n-1}$ is exact. Define $v \in \mathrm{Vect}(M)$ by $\iota(v)\omega = \lambda$. Then $\hat{\omega}' = \mathcal{L}_v\omega$ and $\iota(v)\rho$ is exact. The claim follows now from the fact, that $(\hat{J}'', 0) \in T_{(J,\omega)}\mathcal{C}_{c,a}(M,\rho)$ is equivalent to $\hat{J}'' = (\hat{J}'')^*$. This completes the proof of the lemma. □

In order to obtain a finite dimensional moduli spaces, consider the space of integrable pairs

$$\mathcal{E}_{c,a}(M,\rho) := \{(J,\omega) \in \mathcal{J}_{\mathrm{int},c}(M) \times \mathcal{S}_a(M,\rho) \,|\, J \text{ compatible with } \omega\}. \tag{6.67}$$

Then a natural set of questions are the following.

Question 6.4.6.

1. Is $\mathcal{C}_{c,a}(M,\rho)$ a symplectic submanifold? By Lemma 6.4.5 this is equivalent to the following: Let $(J,\omega) \in \mathcal{C}_{c,a}(M,\rho)$, let $\hat{J} \in \Omega^{0,1}_J(M,TM)$ and let $\hat{\omega} \in \Omega^2_{ex}(M)$ be exact. Suppose

$$\hat{\omega} \wedge \omega^{n-1} = 0, \qquad \hat{\omega}(\cdot,\cdot) - \hat{\omega}(J\cdot, J\cdot) = \langle(\hat{J} - \hat{J}^*)\cdot, \cdot\rangle,$$

$$\hat{J}^* = -\hat{J}, \qquad \widehat{Ric}_\rho(J,\hat{J}) - \kappa\hat{\omega} = 0.$$

   Does this imply $\hat{J} = 0$ and $\hat{\omega} = 0$?

2. Is $\mathcal{E}_{c,a}(M,\rho)$ a symplectic submanifold?

We did not succeed in ansering these questions at the time of writing and hope to come back to this in the future. In any case, both spaces $\mathcal{C}_{c,a}(M,\rho)$ and $\mathcal{E}_{c,a}(M,\rho)$ can be viewed as symplectic fibrations over $\mathcal{S}_a(M,\rho)$.

It follows from similar arguments as in Theorem 6.2.3 that every $\mathrm{Diff}_0(M)$-orbit in $\mathcal{E}_{c,a}(M,\rho)$ contains a solution of the Kähler–Einstein equations which is unique up to the action of $\mathrm{Diff}_0(M,\rho)$. Define the space of Kähler–Einstein pairs by

$$\mathcal{K}_{c,a}(M,\rho) := \{(J,\omega) \in \mathcal{E}_{c,a}(M,\rho) \,|\, \mathrm{Ric}_{\rho,J} - \kappa\omega = 0\}. \tag{6.68}$$

The inclusion of $\mathcal{K}_{c,a}(M,\rho) \subset \mathcal{E}_{c,a}(M,\rho)$ yields then a bijective correspondence

$$\mathcal{E}_{c,a}(M,\rho)/\mathrm{Diff}_0(M) \cong \mathcal{K}_{c,a}(M,\rho)/\mathrm{Diff}_0(M,\rho) \tag{6.69}$$

A much more difficult question in this context is the following: Let (M,J,ω) be a Kähler manifold. Does there exists a Kähler potential $h : M \to \mathbb{R}$ such that the corresponding Kähler form $\omega_h := \omega + \mathbf{i}\partial\bar{\partial}h$ satisfies the Kähler–Einstein equations? In the cases where the canonical bundle of M is trivial or ample, it was proven by Yau [124, 125] and Aubin [5] that there always exist solutions for the Kähler–Einstein equations with $\kappa \leq 0$. On the contrary, when the anti-canonical bundle is ample, then there are known obstructions to the existence of Kähler–Einstein metrics.

Donaldson–Chen–Sun [20, 21, 22] proved in this case that the existence of Kähler–Einstein metrics is equivalent to a particular algebraic geometric notion of K-stability.

The Marsden–Weinstein quotient of $\mathcal{J}_c(M) \times \mathcal{S}_a(M, \rho)$ by $\mathrm{Diff}^{\mathrm{ex}}(M, \rho)$ is

$$\mathcal{M}^{\mathrm{ex}} := \{(J, \omega) \in \mathcal{J}_c(M) \times \mathcal{S}_a(M, \rho) \,|\, \mathrm{Ric}_{\rho,J} - \kappa\omega = 0\} \,/\mathrm{Diff}^{\mathrm{ex}}(M, \rho) \tag{6.70}$$

and it follows from general principles that the symplectic form (6.63) descends to a symplectic form on $\mathcal{M}^{\mathrm{ex}}$. In order to relate this to the moduli space (6.68), we need to understand the further quotient

$$\mathcal{M} := \{(J, \omega) \in \mathcal{J}_c(M) \times \mathcal{S}_a(M, \rho) \,|\, \mathrm{Ric}_{\rho,J} - \kappa\omega = 0\} \,/\mathrm{Diff}_0(M, \rho) \tag{6.71}$$

Suppose $\mathrm{Ric}_{\rho,J} - \kappa\omega = 0$ and let $u, v \in \mathrm{Vect}(M)$ be given such that $\iota(u)\omega$ and $\iota(v)\omega$ are harmonic 1-forms. It follows from (6.25), (6.28) and Lemma 6.3.1 that

$$\Omega_{(J,\omega)}(\mathcal{L}_u(J, \omega), \mathcal{L}_v(J, \omega)) = 2\kappa \int_M \iota(u)\omega \wedge \iota(v)\omega \wedge \frac{\omega^{n-2}}{(n-2)!} \tag{6.72}$$

This shows that the $\mathrm{Diff}_0(M, \rho)/\mathrm{Diff}^{\mathrm{ex}}(M, \rho)$ orbits in $\mathcal{M}^{\mathrm{ex}}$ are symplectic submanifolds when $\kappa \neq 0$. In the case $\kappa = 0$, we saw in Proposition 6.3.2 that this action is trivial. It hence follows that $\mathcal{M}$ always carries a natural symplectic structure.

The next lemma shows that $\mathrm{Diff}_0(M, \rho)/\mathrm{Diff}^{\mathrm{ex}}(M, \rho)$ acts freely when $\kappa < 0$ and $\mathrm{Diff}_0(M, \rho) = \mathrm{Diff}^{\mathrm{ex}}(M, \rho)$ when $\kappa > 0$.

Lemma 6.4.7. *Let (M, ω, J) be a closed connected $2n$-dimensional Kähler manifold and define $\rho := \frac{\omega^n}{n!}$. Assume there exists $\kappa \in \mathbb{R}$ such that $\mathrm{Ric}_{\rho,J} - \kappa\omega = 0$.*

1. *Assume $\kappa > 0$. Then $H^1(M, \mathbb{R}) = 0$ and $\mathrm{Diff}_0(M, \rho) = \mathrm{Diff}^{ex}(M, \rho)$.*
2. *Assume $\kappa < 0$. Then for every $0 \neq v \in \mathrm{Vect}(M)$ it holds $\mathcal{L}_v J \neq 0$.*

Proof. Assume first $\kappa > 0$ and choose $v \in \mathrm{Vect}(M)$ such that $\iota(v)\omega$ is harmonic. It follows from Lemma 6.3.1, that $(\mathcal{L}_v J)^* = \mathcal{L}_v J$, $d\iota(v)\rho = d\iota(Jv)\rho = 0$ and

$$\begin{aligned} \iota(\kappa v + \bar{\partial}_J^* \bar{\partial}_J v)\omega &= \iota(v)\mathrm{Ric}_{\rho,J} + \frac{1}{2}\iota(\bar{\partial}_J^*((\mathcal{L}_v J)J)^*)\omega \\ &= \iota(v)\mathrm{Ric}_{\rho,J} - \frac{1}{2}\Lambda(J, \mathcal{L}_v J) = 0. \end{aligned}$$

where the last equation follows from (6.28). Hence $\kappa v + \bar{\partial}_J^* \bar{\partial}_J v = 0$ and therefore $v = 0$. This completes the proof of the first claim.

Next, assume $\kappa < 0$ and let $v \in \mathrm{Vect}(M)$ with $\mathcal{L}_v J = 0$. It follows from Lemma 6.3.1 that $d\iota(v)\omega$ is an exact $(1,1)$-form and hence there exists $F : M \to \mathbb{R}$ with $d(dF \circ J) = d\iota(v)\omega$. Then (6.28) implies

$$0 = \Lambda_\rho(J, \mathcal{L}_v J) = 2\iota(v)\mathrm{Ric}_{\rho,J} - df_v \circ J + df_{Jv} = 2\kappa\iota(v)\omega - df_v \circ J + df_{Jv}$$

and therefore $d(d(2\kappa F - f_v) \circ J) = 0$. Hence $2\kappa F - df_v$ is constant. Using (6.29), it follows that $f_v = -d^*dF$ and thus $2\kappa F - d^*dF = 0$. The maximum principle then implies that F is constant and therefore $f_v = 0$. By the same argument, since $\mathcal{L}_{Jv} J = 0$, we obtain $f_{Jv} = 0$. Thus $0 = \Lambda_\rho(J, \mathcal{L}_v J) = 2\kappa\iota(v)\omega$ by (6.28) and this yields $v = 0$. This proves the second part of the lemma. □

For a fixed symplectic form $\omega \in \mathcal{S}_a(M,\rho)$ consider the Teichmüller space

$$\begin{aligned}\mathcal{J}_{\mathrm{int},c}(M,\omega) &:= \{J \in \mathcal{J}_{\mathrm{int},c}(M) \,|\, J \text{ compatible with } \omega \text{ and } \mathrm{Ric}_{\rho,J} = \kappa\omega\} \\ \mathcal{T}_c(M,\omega) &:= \mathcal{J}_{\mathrm{int},c}(M,\omega)/\mathrm{Symp}(M,\rho) \cap \mathrm{Diff}_0(M,\rho).\end{aligned} \tag{6.73}$$

This embeds into the moduli space $\mathcal{M}$ defined by (6.71) and the symplectic form on $\mathcal{M}$ restricts to a Kähler form on $\mathcal{T}_c(M,\omega)$ for the complex structure $\hat{J} \mapsto -J\hat{J}$. More generally, the moduli spaces (6.69) yield a fibration

$$\mathcal{K}_{a,c}(M)/\mathrm{Diff}_0(M,\rho) \to \mathcal{S}_a(M,\rho)/\mathrm{Diff}_0(M,\rho) \tag{6.74}$$

with fibres $\mathcal{T}_c(M,\omega)$. The embedding of $\mathcal{K}_{a,c}(M)/\mathrm{Diff}_0(M,\rho)$ into $\mathcal{M}$ gives then rise to a closed 2-form on the total space of this fibration, which restricts to the Weil–Petersson symplectic form along the fibres. It gives therefore rise to a symplectic fibration on (6.74). Although this yields a new perspective on the subject, the Weil–Petersson metric on Teichmüller space (6.73) and its curvature properties have been studied extensively, see [71, 98, 105] and the references therein.

Chapter 7

Universal Hitchin moduli spaces

This chapter contains joint work with Oscar Garcia–Prada, Luis Álvarez-Consul and Mario Garcia-Fernandez. We investigate variants of Hitchin's equations [58] on a Riemann surface Σ. In contrast to the classical theory, we do not fix the complex structure on the surface and calculate moment maps for the action of the extended gauge group. This yields various universal Hitchin moduli spaces which fibre naturally over Teichmüller space with fibre being the corresponding Hitchin moduli space. Most of the material is still work in progress and has not yet been explored in full detail.

7.1 Introduction

The Hitchin's self-duality equations [58] on a Riemann surface Σ can be viewed as hyperkähler extension of the Atiyah–Bott picture for the Yang–Mills equations [4]. Let G be a compact Lie group, let (Σ, J) be a closed Riemann surface and let $P \to \Sigma$ a principal G bundle. The space

$$\mathcal{A}(P) \times \Omega^{1,0}_J(\Sigma, \mathrm{ad}(P) \otimes \mathbb{C}) \tag{7.1}$$

is isomorphic to the cotangent bundle $T^*\mathcal{A}(P)$ and hence carries a natural hyperkähler structure. The Hitchin equations for a pair $(A, \phi) \in \mathcal{A}(P) \times \Omega^{1,0}_J(\Sigma, \mathrm{ad}(P) \otimes \mathbb{C})$ are given by

$$F_A + [\phi \wedge \phi^*] = 0, \qquad \bar{\partial}_A \phi = 0. \tag{7.2}$$

These occur as hyperkähler moment map for the action of the gauge group $\mathcal{G}(P)$ and the moduli space of solutions to Hitchin's equation carries therefore a natural hyperkähler structure. All this requires a fixed complex structure on Σ as background data. The Hitchin moduli spaces for different complex structures are all diffeomorphic, but carry different hyperkähler metrics. We investigate in the following variants of this setup where the complex structure on Σ is not fixed and the gauge group is extended by a subgroup of the diffeomorphism group. This leads to moduli spaces which naturally fibre over Teichmüller space with fibre being the corresponding Hitchin moduli space.

We consider three variants which we describe in the following. In all these cases let (Σ, ρ) be a closed 2-dimensional surfaces equipped with a fixed area form ρ. For a principal bundle $P \to \Sigma$ we denote by $\tilde{\mathcal{G}}(P)$ its extended gauge group . This fits into the exact sequence

$$1 \to \mathcal{G}(P) \to \tilde{\mathcal{G}}(P) \to \mathrm{Ham}(\Sigma, \rho) \to 1. \tag{7.3}$$

and consists of bundle isomorphisms covering Hamiltonian diffeomorphisms on Σ. Every connection $A \in \mathcal{A}(P)$ defines a splitting of the Lie algebras

$$\begin{gathered}\mathrm{Lie}(\tilde{\mathcal{G}}(P)) \xrightarrow{\cong} \Omega^0(\Sigma, \mathrm{ad}(P)) \oplus \{v \in \mathrm{Vect}(\Sigma) \,|\, d\iota(v)\rho \text{ is exact}\}, \\ V \mapsto (A(V),\, v := \pi_* V)\end{gathered} \tag{7.4}$$

which corresponds to the splitting of $\mathrm{Lie}(\tilde{\mathcal{G}}(P)) \subset \mathrm{Vect}(P)$ into its A-horizonal and A-vertical component.

Real reductive groups. Suppose $G = (G, H, \theta, B)$ is a real reductive group, that is a quadruple consisting of a real Lie group G with reductive Lie algebra $\mathfrak{g}$, a maximal compact subgroup $H \subset G$, a Cartan involution $\theta : \mathfrak{g} \to \mathfrak{g}$ which defines a splitting $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{m}$ and a θ- and G-invariant bilinear form $B : \mathfrak{g} \times \mathfrak{g} \to \mathbb{R}$ (see Section 7.3 for more details). The adjoint representation of G restricts to the so-called isotropy representation $\iota : H \to \mathrm{Aut}(\mathfrak{m})$. Now let $P \to \Sigma$ be a principal H bundle and denote by $P(\mathfrak{m}) := P \times_\iota \mathfrak{m}$ the associated $\mathfrak{m}$-bundle. In this setting, the holomorphicity condition of the Higgs field has no interpretation in terms of moment maps and we consider the configuration space

$$\mathcal{X}_1 := \left\{ (J, A, \phi) \;\middle|\; \begin{array}{c} J \in \mathcal{J}(\Sigma),\, A \in \mathcal{A}(P) \\ \phi \in \Omega^{1,0}_J(\Sigma, P(\mathfrak{m}) \otimes \mathbb{C}),\, \bar{\partial}_{A,J}\phi = 0 \end{array} \right\}. \tag{7.5}$$

This carries the following symplectic structure induced by the bilinear form B

$$\Omega_{(A,\phi)}\left((\hat{A}_1, \hat{\phi}_1), (\hat{A}_2, \hat{\phi}_2)\right) := -\int_\Sigma B(\hat{A}_1 \wedge \hat{A}_2) + 2\mathrm{Re}\left(B(\hat{\phi}_1^* \wedge \hat{\phi}_2)\right) \tag{7.6}$$

Theorem A. *The natural action of $\tilde{\mathcal{G}}(P)$ on $\mathcal{X}_1$ is Hamiltonian with moment map*

$$\langle \mu(J, A, \phi), V \rangle = -\int_\Sigma B\left(A(V),\, F_A - [\phi^* \wedge \phi]\right) + \int_\Sigma H\left(2K_J - c\right)\rho$$

where $V \in Lie(\tilde{\mathcal{G}}(P))$, with $v_H = \pi_* V \in Vect(\Sigma)$ and $A(V) \in \Omega^0(\Sigma, ad(P))$ defined by (7.4) and v_H is the Hamiltonian vector field for $H : \Sigma \to \mathbb{R}$, and $c := \frac{4\pi(2g-2)}{vol(\Sigma,\rho)}$.

Proof. This is established in Proposition 7.3.2, which is obtained from a more general moment map calculation in Theorem 7.3.1. □

Complex reductive groups. Assume that G is a compact group and $P \to \Sigma$ a principal G bundle. We consider the space

$$\mathcal{X}_2 := \mathcal{J}(\Sigma) \times \mathcal{A}(P) \times \Omega^1(\Sigma, \mathbf{i}\mathrm{ad}(P)). \tag{7.7}$$

For every fixed J, there is a natural identification of $\mathcal{A}(P) \times \Omega^1(\Sigma, \mathbf{i}\mathrm{ad}(P))$ with $\mathcal{A}(P) \times \Omega^{0,1}_J(\Sigma, \mathrm{ad}(P) \otimes \mathbb{C})$. The hyperkähler structure on (7.1) then gives rise the following three symplectic structures on $\mathcal{X}_2$:

$$\begin{aligned}
&(\Omega_1)_{(J,A,\psi)}\Big((\hat{J}_1, \hat{A}_1, \hat{\psi}_1), (\hat{J}_2, \hat{A}_2, \hat{\psi}_2)\Big)\\
&\qquad := \frac{1}{2}\int_\Sigma \mathrm{tr}\left(\hat{J}_1 J \hat{J}_2\right)\rho - \int_\Sigma \mathrm{tr}\left(\hat{A}_1 \wedge \hat{A}_2\right) - \int \mathrm{tr}\left(\hat{\psi}_1 \wedge \hat{\psi}_2\right)\\
&(\hat{\Omega}_2)_{(J,A,\psi)}\Big((\hat{J}_1, \hat{A}_1, \hat{\psi}_1), (\hat{J}_2, \hat{A}_2, \hat{\psi}_2)\Big)\\
&\qquad := \frac{1}{2}\int_\Sigma \mathrm{tr}\left(\hat{J}_1 J \hat{J}_2\right)\rho + \mathbf{i}\int_\Sigma \mathrm{tr}\left(\hat{A}_1 \wedge *^J \hat{\psi}_2 - \hat{A}_2 \wedge *^J \hat{\psi}_1\right)\\
&\qquad\quad + \mathbf{i}\int_\Sigma \mathrm{tr}\left(\hat{A}_1 \wedge \psi \circ (-\hat{J}_2) - \hat{A}_2 \wedge \psi \circ (-\hat{J}_1)\right).\\
&(\Omega_3)_{(J,A,\psi)}\Big((\hat{J}_1, \hat{A}_1, \hat{\psi}_1), (\hat{J}_2, \hat{A}_2, \hat{\psi}_2)\Big)\\
&\qquad := \frac{1}{2}\int_\Sigma \mathrm{tr}\left(\hat{J}_1 J \hat{J}_2\right)\rho - \mathbf{i}\int_\Sigma \mathrm{tr}\left(\hat{A}_1 \wedge \hat{\psi}_2 - \hat{A}_2 \wedge \hat{\psi}_1\right)
\end{aligned} \tag{7.8}$$

The last term in the definition of $\hat{\Omega}_2$ is needed for obtaining a closed form.

Theorem B. *The action of $\tilde{\mathcal{G}}(P)$ on $\mathcal{X}_2$ is the Hamiltonian for all three symplectic forms with moment maps*

$$\mu_1(J, A, \psi) = \left(\left(F_A + \frac{1}{2}[\psi \wedge \psi]\right),\ (2K_J - 2c)\rho + d\,tr\,(\psi\Lambda_\rho(d_A\psi))\right) \tag{7.9}$$

$$\hat{\mu}_2(J, A, \psi) = (\boldsymbol{i}d_A * \psi,\ (2K_J - 2c)\rho + \boldsymbol{i}d * tr\,(\psi\Lambda_\rho(F_A))) \tag{7.10}$$

$$\mu_3(J, A, \psi) = \left(\boldsymbol{i}d_A\psi,\ (2K_J - 2c)\rho + \boldsymbol{i}d\,tr\left(\psi\Lambda_\rho\left(F_A + \frac{1}{2}[\psi \wedge \psi]\right)\right)\right) \tag{7.11}$$

where $c := 2\pi(2genus(\Sigma) - 2)/vol(\Sigma, \rho)$ and

$$\Lambda_\rho : \Omega^2(\Sigma, ad(P) \otimes \mathbb{C}) \to \Omega^0(\Sigma, ad(P) \otimes \mathbb{C})$$

is the natural map induced by ρ. All three moment maps take values in the space $\Omega^2(\Sigma, ad(P)) \oplus \Omega^2_{ex}(\Sigma)$ which we identify with the dual space of $Lie(\tilde{\mathcal{G}}(P))$ using (7.4).

Proof. See Theorem 7.4.5. □

Note that this theorem does not quite yields a hyperkähler moment map on $\mathcal{X}_2$ Nevertheless, we have

$$(J, A, \psi) \in \mu_1^{-1}(0) \cap \hat{\mu}_2^{-1}(0) \cap \mu_3^{-1}(0) \qquad \Longleftrightarrow \qquad \left\{\begin{array}{c} d_A\psi = 0,\ d_A^*\psi = 0 \\ F_A + \frac{1}{2}[\psi \wedge \psi] = 0 \\ 2K_J = c \end{array}\right. .$$

The equations $d_A\psi = 0$, $d_A^*\psi = 0$, and $F_A + \frac{1}{2}[\psi \wedge \psi] = 0$ correspond to the Hitchin equations (7.2) under the identification of $\mathcal{A}(P) \times \Omega^1(\Sigma, \mathbf{i}\mathrm{ad}(P))$ with $\mathcal{A}(P) \times \Omega^{0,1}_J(\Sigma, \mathrm{ad}(P) \otimes \mathbb{C})$.

Fibrations over Donaldson's moduli space. Consider the configuration space

$$\mathcal{X}_3 := \mathcal{Q}_1(\Sigma) \times \mathcal{A}(P) \times \Omega^1(\Sigma, \mathbf{i}\mathrm{ad}(P)). \tag{7.12}$$

where $\mathcal{Q}_1(\Sigma)$ denotes the space of pairs (J, σ) consisting of a complex structure $J \in \mathcal{J}(\Sigma)$ and a quadratic differential $\sigma \in \Omega^0(\Sigma, S^2(T^*\Sigma \otimes_J \mathbb{C}))$ with pointwise norm $|\sigma|_J < 1$. Donaldson [38] observed that the space $\mathcal{Q}_1(\Sigma)$ carries a hyperkähler structure whose hyperkähler quotient (after taking the action of the flux group into account) yields the Feix–Kaledin hyperkähler extension $\mathcal{M}$ of Teichmüller space. See Chapter 4.5 in this thesis for a detailed exposition of this hyperkähler structure and the moment map calculation. We expect that the space $\mathcal{X}_3$ carries three symplectic forms for which the action of the extended gauge group is Hamiltonian. This should give rise to a moduli space which fibres over $\mathcal{M}$ with fibres being the corresponding Hitchin moduli spaces.

This is still work in progress and has not yet been written up. There are two intriguing aspects which we would like to mention: First, there is a construction of Donaldson [38] which associates to every element in $\mathcal{M}$ a solution of the SU(2) Hitchin equations over Σ and thus $\mathcal{M}$ really parametrizes pairs of solutions to Hitchin's equation (see Lemma 4.6.11). Second, the resulting moduli space is naturally a hyperkähler fibration over the hyperkähler space $\mathcal{M}$. It is probably too optimistic to expect that they combine to a hyperkähler structure on the whole moduli space, but this is certainly something to be investigated more closely.

7.2 The extended gauge group

Let (Σ, ρ) be a closed oriented surface with fixed area form ρ, let G be a real Lie group, and let $P \to \Sigma$ be a principal G bundle. The purpose of this section is to define various extensions of the gauge group of P by subgroups of the diffeomorphism group of Σ and to introduce our notation for it. We also establish Cartan's fromula in this context.

7.2.1 Hamiltonian extension

The extended gauge group $\tilde{\mathcal{G}}(P)$ is defined as the group of bundle automorphisms $\psi : P \to P$ which cover Hamiltonian diffeomorphisms $\varphi : \Sigma \to \Sigma$. It fits into an exact sequence with the gauge group $\mathcal{G}(P)$ of the P

$$1 \to \mathcal{G}(P) \to \tilde{\mathcal{G}}(P) \to \mathrm{Ham}(\Sigma, \rho) \to 1. \tag{7.13}$$

The Lie algebra of the extended gauge group $\tilde{\mathcal{G}}(P) \subset \mathrm{Vect}(P)$ fits into an exact sequence with the Lie algebras of the gauge group and the Hamiltonian group:

$$1 \to \mathrm{Lie}(\mathcal{G}(P)) \to \mathrm{Lie}(\tilde{\mathcal{G}}(P)) \to \mathrm{Lie}(\mathrm{Ham}(\Sigma, \rho)) \to 1. \tag{7.14}$$

and every connection $A \in \mathcal{A}(P)$ defines a splitting

$$\begin{gathered} \mathrm{Lie}(\tilde{\mathcal{G}}(P)) \xrightarrow{\cong} \Omega^0(\Sigma, \mathrm{ad}(P)) \oplus \{v \in \mathrm{Vect}(\Sigma) \,|\, d\iota(v)\rho \text{ is exact}\}, \\ V \mapsto (A(V),\, v = \pi_* V). \end{gathered} \tag{7.15}$$

which is defined by decomposing $V \in \mathrm{Lie}(\tilde{\mathcal{G}}(P)) \subset \mathrm{Vect}(P)$ into its A-horizontal and A-vertical part. Here we identify in the usual way the Lie algebra of the gauge group $\mathrm{Lie}(\mathcal{G}(P)) = \Omega^0(\Sigma, \mathrm{ad}(P))$ with equivariant vertical vector fields on P. We identify $\mathrm{Lie}(\mathrm{Ham}(\Sigma, \rho)) = C^\infty(\Sigma)/\mathbb{R}$ with Hamiltonian vector fields following the sign conventions $dH = \rho(v_H, \cdot)$.

We always use the splitting (7.15) defined by some connection $A \in \mathcal{A}(P)$ to describe the dual Lie algebra of the extended gauge group. The dual space of $\mathrm{Lie}(\mathrm{Ham}(\Sigma, \rho)) = C^\infty(\Sigma)/\mathbb{R}$ can be identified with the space $\Omega^2_{\mathrm{ex}}(\Sigma)$ of exact 2-forms on Σ, where the dual pairing is defined

$$\Omega^2_{\mathrm{ex}}(\Sigma) \times C^\infty(\Sigma)/\mathbb{R} \to \mathbb{R}, \qquad \langle \tau, [H] \rangle := \int_\Sigma H\tau \tag{7.16}$$

Suppose $\mathfrak{g}$ carries an invariant inner product $\langle \cdot, \cdot \rangle$. The dual space of $\mathrm{Lie}(\mathcal{G}(P)) = \Omega^0(\Sigma, \mathrm{ad}(P))$ can then be identified with $\Omega^2(\Sigma, \mathrm{ad}(P))$ via the pairing

$$\Omega^2(\Sigma, \mathrm{ad}(P)) \times \Omega^0(\Sigma, \mathrm{ad}(P)) \to \mathbb{R}, \qquad \langle \eta, \xi \rangle := \int_\Sigma \langle \eta, \xi \rangle. \tag{7.17}$$

7.2.2 Extension by diffeomorphism groups

The extension of the gauge group by $\mathrm{Symp}_0(\Sigma, \rho)$ and $\mathrm{Diff}_0(\Sigma)$ are defined completely analogue to the Hamiltonian case. They fit into the exact sequences

$$1 \to \mathcal{G}(P) \to \tilde{\mathcal{G}}_{\mathrm{Symp}_0}(P) \to \mathrm{Symp}(\Sigma, \rho) \to 1 \tag{7.18}$$

$$1 \to \mathcal{G}(P) \to \tilde{\mathcal{G}}_{\mathrm{Diff}_0}(P) \to \mathrm{Diff}_0(\Sigma) \to 1. \tag{7.19}$$

The various extensions of the gauge are naturally embedded into each other by $\tilde{\mathcal{G}}(P) \subset \tilde{\mathcal{G}}_{\mathrm{Symp}_0}(P) \subset \tilde{\mathcal{G}}_{\mathrm{Diff}_0}(P)$. As before, every connection $A \in \mathcal{A}(P)$ provides a splitting of the corresponding Lie algebra sequences

$$\mathrm{Lie}(\tilde{\mathcal{G}}_{\mathrm{Symp}_0}(P)) \xrightarrow{\cong} \Omega^0(\Sigma, \mathrm{ad}(P)) \oplus \{v \in \mathrm{Vect}(\Sigma) \,|\, d\iota(v)\rho = 0\} \tag{7.20}$$

$$\mathrm{Lie}(\tilde{\mathcal{G}}_{\mathrm{Diff}_0}(P)) \xrightarrow{\cong} \Omega^0(\Sigma, \mathrm{ad}(P)) \oplus \mathrm{Vect}(\Sigma) \tag{7.21}$$

which we both denote by $V \mapsto (A(V),\, v = \pi_* V)$.

7.2.3 A Cartan formula

Let W be a real vector space, let $\gamma : G \to \mathrm{GL}(W)$ a representation and consider the associated vector bundle

$$\gamma(P) := P \times_\gamma W := (P \times W)/G \tag{7.22}$$

where G acts diagonally on $P \times W$ by $g_*(p, w) := (pg^{-1}, \gamma(g)w)$. Any $\psi \in \tilde{\mathcal{G}}_{\mathrm{Diff}_0}(P)$ defines a bundle map

$$\gamma(\psi) : \gamma(P) \to \gamma(P), \qquad \gamma(\psi)[p, w] = [\psi(p), w] \tag{7.23}$$

covering the same diffeomorphism as ψ. This induces a natural action of $\tilde{\mathcal{G}}_{\mathrm{Diff}_0}(P)$ on the space of k-forms with values in $\gamma(P)$ given by

$$\begin{gathered}\tilde{\mathcal{G}}_{\mathrm{Diff}_0}(P) \times \Omega^k(\Sigma, \gamma(P)) \to \Omega^k(\Sigma, \gamma(P)) \\ (\psi^*\alpha)(z, \hat{z}_1, \dots, \hat{z}_k) := \gamma(\psi)^{-1}\alpha(\varphi(z), d\varphi(z)\hat{z}_1, \dots, d\varphi(z)\hat{z}_k)\end{gathered} \tag{7.24}$$

where $\varphi \in \mathrm{Diff}_0(\Sigma)$ denotes the diffeomorphism covered by ψ. The following lemma calculates the infinitesimal action of (7.24).

Lemma 7.2.1. *Fix a connection $A \in \mathcal{A}(P)$ and denote by*

$$\dot{\gamma} := d\gamma(\mathbb{1}) : \mathfrak{g} \to End(W) \tag{7.25}$$

the infinitesimal representation. Let $V \in Lie(\tilde{\mathcal{G}}_{Diff_0}(P))$ and denote its flow by $\psi_t \in \tilde{\mathcal{G}}_{Diff_0}(P)$. For $\alpha \in \Omega^k(\Sigma, \gamma(P))$ it holds

$$\mathcal{L}_V\alpha := \left.\frac{d}{dt}\right|_{t=0} (\psi_t)^*\alpha = -\dot{\gamma}(A(V))\alpha + d_A\iota(v)\alpha + \iota(v)d_A\alpha \tag{7.26}$$

where $A(V) \in \Omega^0(\Sigma, ad(P))$ and $v = \pi_*V \in Vect(\Sigma)$ are defined by (7.19).

Proof. By linearity, it suffices to establish the formula for vertical and horizontal vector field separately. Suppose first V is vertical and $\pi_*V = 0$. Then

$$\mathcal{L}_V\alpha := \left.\frac{d}{dt}\right|_{t=0} \gamma(e^{-tA(V)})\alpha = -\dot{\gamma}(A(V))(\alpha)$$

and (7.26) is satisfied.

Next let $s \in \Omega^0(\Sigma, \gamma(P))$ and assume that V is horizontal. Denote by $\gamma_*V \in \mathrm{Vect}(\gamma(P))$ the composition of V with $TP \to TP \times W \to T\gamma(P)$. Then γ_*V is horizontal for the induced connection on $\gamma(P)$ and its flow is $\gamma(\psi_t)$. Hence

$$\mathcal{L}_V s := \left.\frac{d}{dt}\right|_{t=0} \gamma(\psi_t)^{-1} \circ s \circ \varphi_t = -(\gamma_*V)(s) + ds(v) = \iota(v)d_A s \tag{7.27}$$

where $\varphi_t \in \mathrm{Diff}_0(\Sigma, \rho)$ is the diffeomorphism covered by ψ_t and $v := \pi_*V \in \mathrm{Vect}(\Sigma)$. This proves (7.26) for $k = 0$.

For the general case let V be horizontal and define as before $v := \pi_*V \in \mathrm{Vect}(\Sigma)$. By linearity we may assume that $\alpha = s \otimes \tau$ with $s \in \Omega^0(\Sigma, \gamma(P))$ and $\tau \in \Omega^k(\Sigma)$. From (7.27) and Cartan's formula on differential forms it follows

$$\begin{aligned}\mathcal{L}_V\,(s \otimes \tau) &= (\mathcal{L}_V s) \otimes \tau + s \otimes (\mathcal{L}_v \tau) \\ &= (\iota(v)d_A s) \otimes \tau + s \otimes (\iota(v)d\tau + d\iota(v)\tau) \\ &= \iota(v)\,(d_A s \wedge \tau + s \otimes d\tau) + (d_A s) \wedge (\iota(v)\tau) + s \otimes (d\iota(v)\tau) \\ &= \iota(v)d_A(s \otimes \tau) + d_A(s \otimes \iota(v)\tau).\end{aligned}$$

This establishes (7.26) for $k > 0$ and completes the proof. $\square$

7.3 Higgs bundles for real reductive groups

Following Knapp [70], a real reductive Lie group is a quadruple (G, H, θ, B) consisting of the following data:

- G is a real Lie group with reductive Lie algebra $\mathfrak{g}$
- $H \subset G$ is a maximal compact subgroup with Lie algebra $\mathfrak{h} \subset \mathfrak{g}$
- $\theta : \mathfrak{g} \to \mathfrak{g}$ is a Lie algebra involution which defines a decomposition

$$\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{m}$$

  where $\mathfrak{h} = \mathrm{Lie}(H)$ is the $+1$ eigenspace of θ and $\mathfrak{m}$ is the -1 eigenspace. Moreover, the multiplication map from $H \times \exp(\mathfrak{m}) \to G$ is a diffeomorphism
- $B : \mathfrak{g} \times \mathfrak{g} \to \mathbb{R}$ is a G-invariant and θ-invariant symmetric nondegenerate bilinear form such that the associated symmetric form

$$B_\theta : \mathfrak{g} \times \mathfrak{g} \to \mathbb{R}, \qquad B_\theta(\xi, \eta) := -B(\xi, \theta(\eta))$$

  is positive definite.

Let $P \to \Sigma$ be a principal H bundle. The adjoint representation $\mathrm{Ad} : G \to \mathfrak{g}$ restricts to the so-called isotropy representation

$$\iota : H \to \mathrm{GL}(\mathfrak{m}). \tag{7.28}$$

Its infinitesimal representation $d\iota(\mathbb{1}) : \mathfrak{h} \to \mathrm{GL}(\mathfrak{m})$ is given by the Lie bracket on $\mathfrak{g}$, i.e. $(d\iota(\mathbb{1})\eta)\xi = [\eta, \xi]$. Denote the associated bundle with fibre $\mathfrak{m}$ by

$$P(\mathfrak{m}) := P \times_\iota \mathfrak{m}. \tag{7.29}$$

The bilinear form B extends uniquely to a complex bilinear from on $\mathfrak{m} \otimes \mathbb{C}$, which we still denote B. The space $\mathcal{A}(P) \times \Omega^1(\Sigma, P(\mathfrak{m}) \otimes \mathbb{C})$ carries then the natural symplectic form

$$\Omega_{(A,\phi)}\Big((\hat{A}_1, \hat{\phi}_1), (\hat{A}_2, \hat{\phi}_2)\Big) := -\int_\Sigma B(\hat{A}_1 \wedge \hat{A}_2) + 2\mathrm{Re}\Big(B(\hat{\phi}_1^* \wedge \hat{\phi}_2)\Big) \tag{7.30}$$

The factor 2 in introduced for consistency with the original setting considered by Hitchin for complex reductive groups. Morover, the adjoint ϕ^* is defined by the same formula as in the unitary case. To be precise, in local coordinates (x, y) of Σ we can write

$$\hat{\phi} = (\xi_1 + \mathbf{i}\eta_1)dx + (\xi_2 + \mathbf{i}\eta_2)dy.$$

with ξ_j, η_j taking values in $\mathfrak{m}$.The adjoint is then given by

$$\hat{\phi}^* = (-\xi_1 + \mathbf{i}\eta_1)dx + (-\xi_2 + \mathbf{i}\eta_2)dy.$$

This does not depend on the conformal structure of Σ.

7.3.1 Moment map calculation for a fixed conformal structure

We show that the action of $\tilde{\mathcal{G}}_{\mathrm{Diff}_0}(P)$ on the space $\mathcal{A}(P) \times \Omega^1(\Sigma, P(\mathfrak{m}) \otimes \mathbb{C})$ is Hamiltonian for the symplectic form (7.30).

Theorem 7.3.1. *Let $V \in Lie(\tilde{\mathcal{G}}_{Diff_0}(P)) \subset Vect(P)$ and denote its flow by $\psi_t \in \tilde{\mathcal{G}}_{Diff_0}(P)$. Then holds for all $(A, \phi) \in \mathcal{A}(P) \times \Omega^1(\Sigma, P(\mathfrak{m}) \otimes \mathbb{C})$*

$$\left.\frac{d}{dt}\right|_{t=0} \psi_t^*(A, \phi) = \begin{pmatrix} d_A A(V) + \iota(v) F_A \\ -[A(V), \phi] + d_A \iota(v)\phi + \iota(v) d_A \phi \end{pmatrix} \tag{7.31}$$

where $A(V) \in \Omega^0(\Sigma, ad(P))$ and $v := \pi_* V \in Vect(\Sigma)$ are defined by (7.21). The $\tilde{\mathcal{G}}_{Diff_0}(P)$ action on $\mathcal{A}(P) \times \Omega^1(\Sigma, P(\mathfrak{m}) \otimes \mathbb{C})$ is Hamiltonian with moment map

$$\begin{aligned} \left\langle \mu_{Diff_0}(A, \phi), V \right\rangle = & \int_\Sigma B\left(A(V),\, -F_A + [\phi^* \wedge \phi]\right) \\ & + 2 \int_\Sigma B\left(\phi^* \wedge \iota(v) d_A \phi\right) \end{aligned} \tag{7.32}$$

where $V \in Lie(\tilde{\mathcal{G}}_{Diff_0}(P))$, and $A(V) \in \Omega^0(\Sigma, ad(P))$ and $v := \pi_* V \in Vect(\Sigma)$ are defined as above by (7.21).

Proof. The proof consists of three steps: We calculate the moment map for the action $\tilde{\mathcal{G}}(P)$ on $\mathcal{A}(P)$ and $\Omega^1(\Sigma, P(\mathfrak{m}) \times \mathbb{C})$ separately. The formula for the second moment map is expressed by choosing a connection $A \in \mathcal{A}(P)$ and we verify in the last step that it is independent of this choice.

Step 1: *The infinitesimal action on the space of connections is given by*

$$\mathcal{L}_V A := \left.\frac{d}{dt}\right|_{t=0} \psi_t^* A = d_A A(V) + \iota(v) F_A$$

and the equation

$$\int_\Sigma B\left((d_A A(V) + \iota(v) F_A) \wedge \hat{A}\right) = -\left.\frac{d}{dt}\right|_{t=0} \int_\Sigma B\left(A_t(V),\, F_{A_t}\right)$$

holds for any smooth family $\{A_t\}_{t \in \mathbb{R}}$ with $A_0 = A$ and $\partial_t|_{t=0} A_t = \hat{A}$.

The infinitesimal action of $V \in \mathrm{Lie}(\tilde{\mathcal{G}}_{\mathrm{Diff}_0}(P)) \subset \mathrm{Vect}(P)$ on the connection 1-forms $A \in \mathcal{A}(P) \subset \Omega^1(P, \mathfrak{h})$ is given by Cartan's formula as:

$$\begin{aligned} \mathcal{L}_V A &= d\iota(V) A + \iota(V) dA = dA(V) + [A, A(V)] + \iota(V)\left(dA + \frac{1}{2}[A \wedge A]\right) \\ &= d_A A(V) - \iota(v) F_A \end{aligned}$$

For $\hat{A} \in \Omega^1(\Sigma, \mathrm{ad}(P))$ it then follows

$$\begin{aligned}\int_\Sigma B\left((d_A A(V) + \iota(v) F_A) \wedge \hat{A}\right) &= \int_\Sigma B\left(d_A A(V) \wedge \hat{A}\right) + \int_\Sigma B\left(\iota(v) F_A \wedge \hat{A}\right) \\ &= -\int_\Sigma B\left(A(V) \wedge d_A \hat{A}\right) - \int_\Sigma B\left(\iota(v)\hat{A},\, F_A\right) \\ &= -\partial_{\hat{A}} \int_\Sigma B\left(A(V),\, F_A\right)\end{aligned}$$

Step 2: *Fix a connection $A \in \mathcal{A}(P)$. The infinitesimal action of $V \in Lie\left(\tilde{\mathcal{G}}_{Diff_0}(P)\right)$ on the space $\Omega^1(\Sigma, P(\mathfrak{m}) \otimes \mathbb{C})$ is given by*

$$\mathcal{L}_V \phi := \left.\frac{d}{dt}\right|_{t=0} \psi_t^* \phi = -[A(V), \phi] + d_A \iota(v)\phi + \iota(v) d_A \phi.$$

Moreover,

$$\begin{aligned}&Re \int_\Sigma 2B\left((-\mathcal{L}_V \phi)^* \wedge \hat{\phi}\right) \\ &\quad = \left.\frac{d}{dt}\right|_{t=0} \left(\int_\Sigma B\left(A(V),\, [\phi_t^* \wedge \phi_t]\right) + 2\int_\Sigma B\left(\phi_t^*,\, \iota(v) d_A \phi_t\right)\right)\end{aligned}$$

for any smooth family $\{\phi_t\}_{t\in\mathbb{R}}$ with $\phi_0 = \phi$ and $\partial_t|_{t=0}\, \phi_t = \hat{\phi}$.

The formula for the infinitesimal action follows from Lemma 7.2.1. We then calculate

$$\begin{aligned}&\mathrm{Re} \int_\Sigma 2B\left((-\mathcal{L}_V \phi)^* \wedge \hat{\phi}\right) \\ &\quad = 2\mathrm{Re} \int_\Sigma B\left([A(V), \phi]^* \wedge \hat{\phi}\right) - 2\mathrm{Re} \int_\Sigma B\left((d_A \iota(v)\phi + \iota(v) d_A \phi)^* \wedge \hat{\phi}\right) \\ &\quad = 2\int_\Sigma B\left(A(V),\, [\phi^* \wedge \phi]\right) + 2\mathrm{Re} \int_\Sigma B\left(\phi^* \wedge (\iota(v) d_A \hat{\phi}) + \hat{\phi}^* \wedge (\iota(v) d_A \phi)\right) \\ &\quad = \partial_{\hat{\phi}} \int_\Sigma B\left(A(V),\, [\phi^* \wedge \phi]\right) + 2\partial_{\hat{\phi}} \int_\Sigma B\left(\phi^* \wedge \iota(v) d_A \phi\right)\end{aligned}$$

This proves the moment map equation and Step 2.

Step 3: *The moment map in Step 2 does not depend on A, i.e.*

$$\left.\frac{d}{dt}\right|_{t=0} \left(\int_\Sigma B\left(A_t(V),\, [\phi^* \wedge \phi]\right) + 2\int_\Sigma B\left(\phi^* \wedge \iota(v) d_{A_t} \phi\right)\right) = 0$$

for any smooth family $\{A_t\}_{t\in\mathbb{R}}$.

A direct calculation yields

$$\begin{aligned}
\left.\frac{d}{dt}\right|_{t=0} & \left(\int_\Sigma B\left(A_t(V),\, [\phi^* \wedge \phi]\right) + 2\int_\Sigma B\left(\phi^* \wedge \iota(v) d_{A_t}\phi\right)\right) \\
&= \int_\Sigma B\left(\hat{A}(v),\, [\phi^* \wedge \phi]\right) + 2\mathrm{Re}\int_\Sigma B\left(\phi^* \wedge \iota(v)[\hat{A} \wedge \phi]\right) \\
&= \int_\Sigma B\left(\hat{A} \wedge \iota(v)[\phi^* \wedge \phi]\right) + 2\mathrm{Re}\int_\Sigma B\left([\iota(v)\phi^*, \phi] \wedge \hat{A}\right) \\
&= \int_\Sigma B\left(\hat{A} \wedge \iota(v)[\phi^* \wedge \phi]\right) + \int_\Sigma B\left(\iota(v)[\phi^*, \phi] \wedge \hat{A}\right) \\
&= 0
\end{aligned}$$

where the penultimate equation follows from the identity

$$2\mathrm{Re}\left([\iota(v)\phi^*, \phi]\right) = \iota(v)[\phi^* \wedge \phi].$$

This completes the proof of Step 3 and the theorem. □

7.3.2 Integrability conditions

Denote by $\mathcal{J}(\Sigma)$ the space of complex structures on (Σ, ρ), which are compatible with ρ. This can be viewed as an infinite dimensional symplectic manifold, where the symplectic structure is given by

$$\Omega_J(\hat{J}_1, \hat{J}_2) := \frac{1}{2}\int_\Sigma \mathrm{tr}\left(\hat{J}_1 J \hat{J}_2\right) \tag{7.33}$$

(see Section 4.4 in this thesis for more details). Consider the space of triple (J, A, ϕ), consisting of a complex structure J, a connection A and a holomorphic Higgs field ϕ:

$$\mathcal{X} := \left\{ (J, A, \phi) \;\middle|\; \begin{array}{c} J \in \mathcal{J}(\Sigma),\ A \in \mathcal{A}(P) \\ \phi \in \Omega^{1,0}_J(\Sigma, P(\mathfrak{m}) \otimes \mathbb{C}),\ \bar{\partial}_{A,J}\phi = 0 \end{array} \right\}. \tag{7.34}$$

This is a symplectic submanifold of $\mathcal{J}(\Sigma) \times \left(\mathcal{A}(P) \times \Omega^1(\Sigma, P(\mathfrak{m}) \otimes \mathbb{C})\right)$ for the product symplectic form obtained from (7.33) and (7.30). The extended gauge group $\tilde{\mathcal{G}}(P)$ acts naturally on this space in a Hamiltonian fashion and we have the following.

Proposition 7.3.2. *The extended gauge group $\tilde{\mathcal{G}}(P)$ acts in a Hamiltonian way on the space $\mathcal{X}$ defined by (7.34) with moment map*

$$\langle \mu(J, A, \phi), V \rangle = -\int_\Sigma B\left(A(V),\, F_A - [\phi^* \wedge \phi]\right) + \int_\Sigma H\left(2K_J - 2c\right)\rho$$

where $V \in Lie(\tilde{\mathcal{G}}(P))$, with $v_H = \pi_* V \in Vect(\Sigma)$ and $A(v) \in \Omega^0(\Sigma, ad(P))$ defined by (7.15) and v_H is the Hamiltonian vector field for some function $H : \Sigma \to \mathbb{R}$, and $c := \frac{2\pi(2g-2)}{vol(\Sigma,\rho)}$.

Proof. It follows from Theorem 7.3.1 and integration by parts, that the action of $\tilde{\mathcal{G}}(P)$ on $\mathcal{A}(P) \times \Omega^1(\Sigma, P(\mathfrak{m}) \otimes \mathbb{C})$ is Hamiltonian with moment map

$$\langle \tilde{\mu}(A, \phi), V \rangle = \int_\Sigma B\left(A(V),\, -F_A + [\phi^* \wedge \phi]\right) + \int_\Sigma 2H dB\left(\phi^*,\, \Lambda(\iota(v) d_A \phi)\right)$$

where $V \in \mathrm{Lie}(\tilde{\mathcal{G}}(P))$ with $v_H = \pi_* V \in \mathrm{Vect}(\Sigma)$ and $\Lambda_\rho : \Omega^2(\Sigma, P(\mathfrak{m}) \otimes \mathbb{C}) \to \Omega^0(\Sigma, P(\mathfrak{m}) \otimes \mathbb{C})$ is the natural map defined by ρ. The holomorphicity condition $\bar{\partial}_{A,J}\phi = 0$ is equivalent to $d_A\phi = 0$ and thus the second integral vanishes.

Moreover, it follows from Theorem 4.4.2 that $J \mapsto (2K_J - 2c)\rho$ is a moment map for the action of $\mathrm{Ham}(\Sigma, \rho)$ on $\mathcal{J}(\Sigma)$. The proposition follows by combining these two moment maps. □

The zero set of the moment map equation consists of triples (J, A, ϕ) where $\rho(\cdot, J\cdot)$ defines a constant curvature metric on Σ and (A, ϕ) satisfies the Hitchin equations $\bar{\partial}_A\phi = 0$, $F_A = [\phi^* \wedge \phi]$ for the complex structure J.

7.4 Higgs bundles for complex reductive groups

First, we recall for fixed $J \in \mathcal{J}(\Sigma)$ the hyperkähler structure introduced by Hitchin on the space

$$\mathcal{A}(P) \times \Omega^{1,0}_J(\Sigma, \mathrm{ad}(P) \otimes \mathbb{C}) \cong \mathcal{A}(P) \times \Omega^1(\Sigma, \mathbf{i}\mathrm{ad}(P)) \cong \mathcal{A}(P^c).$$

On $\mathcal{A}(P) \times \Omega^1(\Sigma, \mathbf{i}\mathrm{ad}(P))$, it turns out that two of the hyperkähler symplectic forms are independent of J. The action of the extended gauge group $\tilde{\mathcal{G}}_{\mathrm{Diff}_0}(P)$ is Hamiltonian for these symplectic structures and we calculate moment maps for this action in Theorem 7.4.2.

We then incorporate the conformal structure on Σ into our data and consider triples (J, A, ψ) consisting of a complex structure J and a Higgs pair (A, ψ). The hyperkähler structure on the space of Higgs pairs and the symplectic form on $\mathcal{J}(\Sigma)$ give rise to three symplectic structures. The action of the extended gauge group $\tilde{\mathcal{G}}_{\mathrm{Diff}_0}(P)$ is Hamiltonian for all three symplectic forms and we calculate the corresponding moment in Theorem 7.4.5. This gives rise to a moduli space which naturally fibres over Teichmüller space with the corresponding Hitchin moduli space as fibre.

Throughout this section, we assume that $G \subset U(n)$ is a compact real Lie group with Lie algebra $\mathfrak{g} \subset \mathfrak{u}(n)$ and invariant inner product

$$\langle \xi, \eta \rangle_{\mathfrak{g}} = \mathrm{tr}(\xi^* \eta) = -\mathrm{tr}(\xi\eta) \tag{7.35}$$

induced from the unitary group.

7.4.1 The space of Higgs pairs – the unitary point of view

Fix a complex structure $J \in \mathcal{J}(\Sigma)$. The space of Higgs pairs

$$\mathcal{A}(P) \times \Omega^{1,0}_J(\Sigma, \mathrm{ad}(P) \otimes \mathbb{C}). \tag{7.36}$$

is an affine space over the linear space $\Omega^1(\Sigma, \mathrm{ad}(P)) \times \Omega^{1,0}(\Sigma, \mathrm{ad}(P) \otimes \mathbb{C})$ and Hitchin [58] introduced the following hyperkähler structure on the space.

The Riemannian structure is given by

$$g_{(A,\phi)}((\hat{A}_1, \hat{\phi}_1), (\hat{A}_2, \hat{\phi}_2)) := -\int_\Sigma \operatorname{tr}\left(\hat{A}_1 \wedge *\hat{A}_2\right) + \operatorname{Im}\int 2\operatorname{tr}\left(\hat{\phi}_1^* \wedge \hat{\phi}_2\right) \tag{7.37}$$

To be more explicit, choose a local holomorphic coordinate $z = x + \mathbf{i}y$ (with respect to J) and write $\hat{A}$ and $\hat{\phi}$ as

$$\hat{A} = adx + bdy, \qquad \hat{\phi} = (\xi + \mathbf{i}\eta)dz = (\xi + \mathbf{i}\eta)dx + (-\eta + \mathbf{i}\xi)dy.$$

with a, b, ξ, η taking values in the real Lie algebra $\mathfrak{g} \subset \mathfrak{u}(n)$. Then

$$\hat{\phi}^* = (-\xi + \mathbf{i}\eta)d\bar{z} = (-\xi + \mathbf{i}\eta)dx + (\eta + \mathbf{i}\xi)dy$$

and the integrand appearing in the formula for the metric is given by

$$\begin{aligned} &-\operatorname{tr}\left(\hat{A}_1 \wedge *\hat{A}_2\right) + \operatorname{Im}\left(2\operatorname{tr}\left(\phi_1^* \wedge \phi_2\right)\right) \\ &\qquad = -\operatorname{tr}\left(a_1a_2 + b_1b_2\right) dx \wedge dy - 4\operatorname{tr}\left(\xi_1\xi_2 + \eta_1\eta_2\right) dx \wedge dy \end{aligned}$$

The inner product on $\Omega^{1,0}(\Sigma, \operatorname{ad}(P) \otimes \mathbb{C})$ agrees up to a factor of 2 with the standard Riemannian structure on $\Omega^1(\Sigma, \operatorname{ad}(P) \otimes \mathbb{C})$ induced by J and the invariant inner product of $\mathfrak{g}$.

The hyperkähler structure consists of the three complex structures

$$\mathcal{I}_1(\hat{A}, \hat{\phi}) := (*\hat{A}, \mathbf{i}\hat{\phi}) \tag{7.38}$$

$$\mathcal{I}_2(\hat{A}, \hat{\phi}) := (\mathbf{i}(\hat{\phi} + \hat{\phi}^*), \mathbf{i}\hat{A}^{1,0}) \tag{7.39}$$

$$\mathcal{I}_3(\hat{A}, \hat{\phi}) := (\hat{\phi} - \hat{\phi}^*, -\hat{A}^{1,0}) \tag{7.40}$$

where $\hat{A}^{1,0}, \hat{A}^{0,1} \in \Omega^1(\Sigma, \operatorname{ad}(P) \otimes \mathbb{C})$ are defined by

$$\hat{A}^{0,1} = \frac{\hat{A} - \mathbf{i}(*\hat{A})}{2}, \qquad \hat{A}^{1,0} = \frac{\hat{A} + \mathbf{i}(*\hat{A})}{2}.$$

Both of them determine $\hat{A}$ uniquely through the relations $\hat{A} = \hat{A}^{1,0} + \hat{A}^{0,1}$ and $\hat{A}^{1,0} = -(\hat{A}^{0,1})^*$. The corresponding Kähler forms are

$$(\Omega_1)_{(A,\phi)}((\hat{A}_1, \hat{\phi}_1), (\hat{A}_2, \hat{\phi}_2)) := -\int_\Sigma \operatorname{tr}\left(\hat{A}_1 \wedge \hat{A}_2\right) - \operatorname{Re}\int 2\operatorname{tr}\left(\hat{\phi}_1^* \wedge \hat{\phi}_2\right) \tag{7.41}$$

$$(\Omega_2)_{(A,\phi)}((\hat{A}_1, \hat{\phi}_1), (\hat{A}_2, \hat{\phi}_2)) := \operatorname{Re}\int_\Sigma 2\operatorname{tr}\left(\hat{A}_1 \wedge \hat{\phi}_2 - \hat{A}_2 \wedge \hat{\phi}_1\right) \tag{7.42}$$

$$(\Omega_3)_{(A,\phi)}((\hat{A}_1, \hat{\phi}_1), (\hat{A}_2, \hat{\phi}_2)) := \operatorname{Im}\int_\Sigma 2\operatorname{tr}\left(\hat{A}_1 \wedge \hat{\phi}_2 - \hat{A}_2 \wedge \hat{\phi}_1\right) \tag{7.43}$$

The first symplectic form naturally extends to a symplectic form defined on the larger space $\mathcal{A}(P) \times \Omega^1(\Sigma, \mathrm{ad}(P) \otimes \mathbb{C})$ via the same formula:

$$(\tilde{\Omega}_1)_{(A,\phi)}((\hat{A}_1, \hat{\phi}_1), (\hat{A}_2, \hat{\phi}_2)) := -\int_\Sigma \mathrm{tr}\left(\hat{A}_1 \wedge \hat{A}_2\right) - \mathrm{Re} \int 2\mathrm{tr}\left(\hat{\phi}_1^* \wedge \hat{\phi}_2\right) \tag{7.44}$$

for $A \in \mathcal{A}(P)$, $\hat{A}_i \in \Omega^1(\Sigma, \mathrm{ad}(P))$ and $\phi, \hat{\phi}_i \in \Omega^1(\Sigma, \mathrm{ad}(P) \otimes \mathbb{C})$.

Proposition 7.4.1. *The natural $\tilde{\mathcal{G}}_{Diff_0}(P)$ action on $\mathcal{A}(P) \times \Omega^1(\Sigma, ad(P) \otimes \mathbb{C})$ is Hamiltonian with respect to the $\tilde{\Omega}_1$ and with moment map*

$$\langle \mu(A, \phi), V \rangle = -\int_\Sigma tr(A(V)(F_A + [\phi^* \wedge \phi])) - 2\int_\Sigma tr(\phi^* \wedge \iota(v) d_A \phi) \tag{7.45}$$

where $V \in Lie(\tilde{\mathcal{G}}_{Diff_0}(P))$, $A \in \mathcal{A}(P)$, $\phi \in \Omega^1(\Sigma, ad(P) \otimes \mathbb{C})$, and $A(V)$ and $v := \pi_* V$ are defined by (7.21).

Proof. This is a special case of Theorem 7.3.1. For this note that G^c is real reductive group with maximal compact subgroup G, bilinear form $B(\xi, \eta) = \mathrm{tr}(\xi\eta)$ and Cartan involution $\theta(\zeta) = -\zeta^*$. We obtain different signs in front of the Higgs fields in the moment map equation, which are due to the identification $\mathfrak{m} = \mathbf{i}\mathfrak{g}$. □

7.4.2 The space of Higgs pairs – the self-adjoint point of view

Fix $J \in \mathcal{J}(\Sigma)$ and identify the space Higgs pairs (7.36) with

$$\mathcal{A}(P) \times \Omega^1(\Sigma, \mathbf{i}\mathrm{ad}(P)). \tag{7.46}$$

The identification is obtained by the map $(A, \phi) \mapsto (A, \psi := \phi^* + \phi)$ with inverse $(A, \psi) \mapsto (A, \phi = \psi^{1,0})$. We summarize in the following the hyperkähler structure on (7.46) which corresponds to the hyperkähler structure on (7.36).

The Riemannian structure is given by

$$g_{(A,\psi)}\left((\hat{A}_1, \hat{\psi}_1), (\hat{A}_2, \hat{\psi}_2)\right) := -\int_\Sigma \mathrm{tr}(\hat{A}_1 \wedge *\hat{A}_2) + \int_\Sigma \mathrm{tr}(\hat{\psi}_1 \wedge *\hat{\psi}_2) \tag{7.47}$$

Let $P^c := P \times_G G^c$ denote the complexified bundle. Then there is a natural isomorphism

$$\mathcal{A}(P) \times \Omega^1(\Sigma, \mathbf{i}\mathrm{ad}(P)) \to \mathcal{A}(P^c), \qquad (A, \psi) \mapsto A + \psi \tag{7.48}$$

and this is an isometry with respect to the standard Riemannian structure on $\mathcal{A}(P^c)$. This amounts to the formula

$$g_{(A,\psi)}((\hat{A}_1, \hat{\psi}_1), (\hat{A}_2, \hat{\psi}_2)) = \mathrm{Re} \int_\Sigma \mathrm{tr}\left((\hat{A}_1 + \hat{\psi}_1)^* \wedge *(\hat{A}_2 + \hat{\psi}_2)\right).$$

The hyperkähler structure consists of the three complex structures

$$\mathcal{I}_1(\hat{A},\hat{\psi}) := (*\hat{A}, -*\hat{\psi}) \tag{7.49}$$
$$\mathcal{I}_2(\hat{A},\hat{\psi}) := (\mathbf{i}\hat{\psi}, \mathbf{i}\hat{A}) \tag{7.50}$$
$$\mathcal{I}_3(\hat{A},\hat{\psi}) := (*\mathbf{i}\hat{\psi}, -*\mathbf{i}\hat{A}) \tag{7.51}$$

which satisfy the quaternionic relations. Their corresponding Kähler forms are

$$(\Omega_1)_{(A,\psi)}((\hat{A}_1,\hat{\psi}_1),(\hat{A}_2,\hat{\psi}_2)) := -\int_\Sigma \mathrm{tr}\left(\hat{A}_1\wedge\hat{A}_2+\hat{\psi}_1\wedge\hat{\psi}_2\right) \tag{7.52}$$
$$(\Omega_2)_{(A,\psi)}((\hat{A}_1,\hat{\psi}_1),(\hat{A}_2,\hat{\psi}_2)) := \mathbf{i}\int_\Sigma \mathrm{tr}\left(\hat{A}_1\wedge *\hat{\psi}_2-\hat{A}_2\wedge *\hat{\psi}_1\right) \tag{7.53}$$
$$(\Omega_3)_{(A,\psi)}((\hat{A}_1,\hat{\psi}_1),(\hat{A}_2,\hat{\psi}_2)) := (-\mathbf{i})\int_\Sigma \mathrm{tr}\left(\hat{A}_1\wedge\hat{\psi}_2-\hat{A}_2\wedge\hat{\psi}_1\right) \tag{7.54}$$

Let $P^c := P\times_G G^c$ denote the complexified bundle. Under the isomorphism (7.48), the complex structure $\mathcal{I}_2$ and Ω_2 corresponds to the standard hermitian structure on $\mathcal{A}(P^c)$, since $\mathcal{I}_2(A+\psi)=\mathbf{i}(A+\psi)$ and

$$(\Omega_2)_{(A,\psi)}((\hat{A}_1,\hat{\psi}_1),(\hat{A}_2,\hat{\psi}_2)) = \mathrm{Im}\int_\Sigma \mathrm{tr}\left((\hat{A}_1+\hat{\psi}_1)^*\wedge *(\hat{A}_2+\hat{\psi}_2)\right).$$

The two remaining symplectic forms Ω_1 and Ω_3 correspond to the standard holomorphic symplectic form on $\mathcal{A}(P^c)$ by the following relation

$$(\Omega_1-\mathbf{i}\Omega_3)_{(A,\psi)}\left((\hat{A}_1,\hat{\psi}_1),(\hat{A}_2,\hat{\psi}_2)\right) = -\int_\Sigma \mathrm{tr}\left((\hat{A}_1+\hat{\psi}_1)\wedge(\hat{A}_2+\hat{\psi}_2)\right).$$

Moreover, the symplectic forms Ω_1 and Ω_3 are independent of the conformal structure on Σ and they are preserved by the natural action of the extended Gauge group $\tilde{\mathcal{G}}_{\mathrm{Diff}_0}(P)$. The next proposition shows that this action is in fact Hamiltonian and calculates the corresponding moment maps.

Proposition 7.4.2. *Let $V\in Lie(\tilde{\mathcal{G}}_{Diff_0}(P))\subset Vect(P)$ and denote its flow by $g_t\in\tilde{\mathcal{G}}_{Diff_0}(P)$. Then holds*

$$\mathcal{L}_V(A,\psi) := \left.\frac{d}{dt}\right|_{t=0} g_t^*(A,\psi) = \begin{pmatrix} d_A A(V)+\iota(v)F_A \\ -[A(V),\psi]+\iota(v)d_A\psi+d_A\iota(v)\psi \end{pmatrix}$$

for all $(A,\psi)\in\mathcal{A}(P)\times\Omega^1(\Sigma,\boldsymbol{i}ad(P)$ where $A(V)$ and $v:=\pi_*V$ are defined by (7.21). The $\tilde{\mathcal{G}}_{Diff_0}(P)$ action on $\mathcal{A}(P)\times\Omega^1(\Sigma,\boldsymbol{i}ad(P)$ is Hamiltonian with respect to Ω_1 and Ω_3 and with moment maps

$$\langle(\mu_1-\boldsymbol{i}\mu_3)(A,\psi),\,V\rangle = -\int_\Sigma tr((A(V)+\iota(v)\psi)F_{A+\psi}) \tag{7.55}$$

where $V\in Lie(\tilde{\mathcal{G}}_{Diff_0}(P))$, $A\in\mathcal{A}(P)$, and $\psi\in\Omega^1(\Sigma,\boldsymbol{i}ad(P))$.

Proof. The formula for the infinitesimal action follows from Step 1 in the proof of Theorem 7.3.1 and Lemma 7.2.1. Using the isomorphism (7.48), the infinitesimal action can be expressed as

$$\begin{aligned}\mathcal{L}_V(A,\psi) &= d_A A(V) + \iota(v)F_A - [A(V),\psi] + \iota(v)d_A\psi + d_A\iota(v)\psi \\ &= \iota(v)F_{A+\psi} + d_{A+\psi}(A(V)+\iota(v)\psi)\end{aligned}$$

It then follows

$$\begin{aligned}&(\Omega_1 - \mathbf{i}\Omega_3)_{(A,\psi)}((\hat{A},\hat{\psi}),\mathcal{L}_V(A,\psi)) \\ &\quad= \int_\Sigma \operatorname{tr}\Big((\iota(v)F_{A+\psi} + d_{A+\psi}(A(V)+\iota(v)\psi))\wedge(\hat{A}+\hat{\psi})\Big) \\ &\quad= -\int_\Sigma \operatorname{tr}\Big((\iota(v)\hat{A}+\iota(v)\hat{\psi})F_{A+\psi}\Big) \\ &\qquad - \int_\Sigma \operatorname{tr}\Big((A(V)+\iota(v)\psi)d_{A+\psi}(\hat{A}+\hat{\psi})\Big) \\ &\quad= \partial_{(\hat{A},\hat{\psi})}\int_\Sigma -\operatorname{tr}\left((A(V)+\iota(v)\psi)F_{A+\psi}\right).\end{aligned}$$

and this proves the moment map equation. □

The action of $\tilde{\mathcal{G}}(P)$ does not preserve the second symplectic structure Ω_2 which depends on the conformal structure of Σ. For the action of the (not extended) gauge group $\mathcal{G}(P)$ it was observed by Hitchin [58] that this action is Hamiltonian with moment map $(A,\psi)\mapsto -\mathbf{i}d_A * \psi$.

7.4.3 The full configuration space

Three symplectic forms

We consider the configuration space

$$\mathcal{J}(\Sigma)\times\mathcal{A}(P)\times\Omega^1(\Sigma,\mathbf{i}\mathrm{ad}(P)). \tag{7.56}$$

This carries three non-degenerate 2-forms which arise as combination of the symplectic structure on $\mathcal{J}(\Sigma)$ and the three symplectic structures on $\mathcal{A}(P)\times\Omega^1(\Sigma,\mathbf{i}\mathrm{ad}(P))$.

$$\begin{aligned}&(\Omega_1)_{(J,A,\psi)}\Big((\hat{J}_1,\hat{A}_1,\hat{\psi}_1),(\hat{J}_2,\hat{A}_2,\hat{\psi}_2)\Big) \\ &\quad:= \frac{1}{2}\int_\Sigma \operatorname{tr}\Big(\hat{J}_1 J\hat{J}_2\Big)\rho - \int_\Sigma \operatorname{tr}\Big(\hat{A}_1\wedge\hat{A}_2\Big) - \int \operatorname{tr}\Big(\hat{\psi}_1\wedge\hat{\psi}_2\Big) \\ &(\Omega_2)_{(J,A,\psi)}\Big((\hat{J}_1,\hat{A}_1,\hat{\psi}_1),(\hat{J}_2,\hat{A}_2,\hat{\psi}_2)\Big) \\ &\quad:= \frac{1}{2}\int_\Sigma \operatorname{tr}\Big(\hat{J}_1 J\hat{J}_2\Big)\rho + \mathbf{i}\int_\Sigma \operatorname{tr}\Big(\hat{A}_1\wedge *^J\hat{\psi}_2 - \hat{A}_2\wedge *^J\hat{\psi}_1\Big) \\ &(\Omega_3)_{(J,A,\psi)}\Big((\hat{J}_1,\hat{A}_1,\hat{\psi}_1),(\hat{J}_2,\hat{A}_2,\hat{\psi}_2)\Big) \\ &\quad:= \frac{1}{2}\int_\Sigma \operatorname{tr}\Big(\hat{J}_1 J\hat{J}_2\Big)\rho - \mathbf{i}\int_\Sigma \operatorname{tr}\Big(\hat{A}_1\wedge\hat{\psi}_2 - \hat{A}_2\wedge\hat{\psi}_1\Big)\end{aligned}$$

These forms are clearly non-degenerated and Ω_1 and Ω_3 are closed as product forms. Since the Hodge $*$-operator in the definition of Ω_2 depends on J, this is not a product symplectic form and Lemma 7.4.3 shows that it is not closed! However, we can slightly modify Ω_2 and define

$$\begin{aligned}(\hat{\Omega}_2)_{(J,A,\psi)}&\left((\hat{J}_1,\hat{A}_1,\hat{\psi}_1),(\hat{J}_2,\hat{A}_2,\hat{\psi}_2)\right)\\ &:=\frac{1}{2}\int_\Sigma \operatorname{tr}\left(\hat{J}_1 J\hat{J}_2\right)\rho+\mathbf{i}\int_\Sigma \operatorname{tr}\left(\hat{A}_1\wedge *^J\hat{\psi}_2-\hat{A}_2\wedge *^J\hat{\psi}_1\right)\\ &\quad+\mathbf{i}\int_\Sigma \operatorname{tr}\left(\hat{A}_1\wedge\psi\circ(-\hat{J}_2)-\hat{A}_2\wedge\psi\circ(-\hat{J}_1)\right).\end{aligned}$$

Lemma 7.4.4 below shows that $\hat{\Omega}_2$ is indeed a symplectic form which agrees with Ω_2 along the slices $\{J\}\times\mathcal{A}(P)\times\Omega^1(\Sigma,\mathbf{i}\mathrm{ad}(P))$ and $\mathcal{J}(\Sigma)\times\{(A,\psi)\}$.

Lemma 7.4.3. *The exterior derivative of Ω_2 is the three-form:*

$$\begin{aligned}(d\Omega_2)_{(J,A,\psi)}&\left((\hat{J}_1,\hat{A}_1,\psi_1),(\hat{J}_2,\hat{A}_2,\psi_2),(\hat{J}_3,\hat{A}_3,\psi_3)\right)\\ &=(-\boldsymbol{i})\int_\Sigma tr\left(\sum_{j=1}^{3}\hat{A}_{j+1}\wedge(\hat{\psi}_{j+2}\circ\hat{J}_j)-\hat{A}_{j+2}\wedge(\hat{\psi}_{j+1}\circ\hat{J}_j)\right)\end{aligned}$$

where the indices are understood cyclic modulo 3.

Proof. Fix a point $(J,A,\psi)\in\mathcal{J}(\Sigma)\times\mathcal{A}(P)\times\Omega^1(\Sigma,\mathbf{i}\mathrm{ad}(P))$ and choose

$$(\hat{J}_j,\hat{A}_j,\psi_j)\in T_J\mathcal{J}(\Sigma)\oplus\Omega^1(\Sigma,\mathrm{ad}(P))\oplus\Omega^1(\Sigma,\mathbf{i}\mathrm{ad}(P))$$

for $j=1,2,3$. Moreover, choose a 3-parameter familiy $(J_{r,s,t},A_{r,s,t},\psi_{r,s,t})$ with

$$(J_{0,0,0},A_{0,0,0},\psi_{0,0,0})=(J,A,\psi)$$

$$\partial_r(J_{r,s,t},A_{r,s,t},\psi_{r,s,t})|_{(r,s,t)=(0,0,0)}=(\hat{J}_1,\hat{A}_1,\psi_1)$$

$$\partial_s(J_{r,s,t},A_{r,s,t},\psi_{r,s,t})|_{(r,s,t)=(0,0,0)}=(\hat{J}_2,\hat{A}_2,\psi_2)$$

$$\partial_t(J_{r,s,t},A_{r,s,t},\psi_{r,s,t})|_{(r,s,t)=(0,0,0)}=(\hat{J}_3,\hat{A}_3,\psi_3).$$

Then the exterior derivative is given by

$$\begin{aligned}(d\Omega_2)_{(J,A,\psi)}&\left((\hat{J}_1,\hat{A}_1,\psi_1),(\hat{J}_2,\hat{A}_2,\psi_2),(\hat{J}_3,\hat{A}_3,\psi_3)\right)\\ &=\partial_r\left[(\Omega_2)_{(J_{r,s,t},A_{r,s,t},\psi_{r,s,t})}\left(\partial_s(J_{r,s,t},A_{r,s,t},\psi_{r,s,t}),\partial_t(J_{r,s,t},A_{r,s,t},\psi_{r,s,t})\right)\right]\\ &+\partial_s\left[(\Omega_2)_{(J_{r,s,t},A_{r,s,t},\psi_{r,s,t})}\left(\partial_t(J_{r,s,t},A_{r,s,t},\psi_{r,s,t}),\partial_r(J_{r,s,t},A_{r,s,t},\psi_{r,s,t})\right)\right]\\ &+\partial_t\left[(\Omega_2)_{(J_{r,s,t},A_{r,s,t},\psi_{r,s,t})}\left(\partial_r(J_{r,s,t},A_{r,s,t},\psi_{r,s,t}),\partial_s(J_{r,s,t},A_{r,s,t},\psi_{r,s,t})\right)\right]\end{aligned}$$

where the right hand side is evaluated at $(r,s,t)=(0,0,0)$. Since $*^J\psi=\psi\circ(-J)$, its differential in direction $\hat{J}$ is given by $\partial_{\hat{J}}(*^J\psi)=\psi\circ(-\hat{J})$. It thus follows from the chain rule that the first term is given by

$$\begin{aligned}\frac{1}{2}\int_\Sigma \mathrm{tr}\left((\partial_r\partial_s J)J\hat{J}_3+\hat{J}_2\hat{J}_1\hat{J}_3+\hat{J}_1 J(\partial_r\partial_s J)\right)\\ +\mathbf{i}\int_\Sigma \mathrm{tr}\left(\partial_r\partial_s A\wedge *^J\psi_3+\hat{A}_2\wedge *^J\partial_r\partial_t\psi+\hat{A}_2\wedge\psi_3\circ(-\hat{J}_1)\right)\\ -\mathbf{i}\int_\Sigma \mathrm{tr}\left(\partial_r\partial_t A\wedge *^J\psi_2+\hat{A}_3\wedge *^J\partial_r\partial_s\psi-\hat{A}_3\wedge\psi_2\circ(-\hat{J}_1)\right)\end{aligned}$$

The other two terms are given by cyclic permutations. All terms involving second order partial derivatives cancel out when summing all three equations up. Moreover, $\mathrm{tr}(\hat{J}_1\hat{J}_2\hat{J}_3)=0$, since $\hat{J}_1\hat{J}_2\hat{J}_3$ anti-commutes with J. The remaining terms yield the claimed formula for the exterior derivative. □

Lemma 7.4.4. *$\hat{\Omega}_2$ defines a symplectic form on $\mathcal{J}(\Sigma)\times\mathcal{A}(P)\times\Omega^1(\Sigma,\boldsymbol{i}ad(P))$ and it agrees with Ω_2 along the slices $\mathcal{J}(\Sigma)\times\{(A,\psi)\}$ and $\{J\}\times\mathcal{A}(P)\times\Omega^1(\Sigma,\boldsymbol{i}ad(P))$.*

Proof. It follows along the lines of the proof of Lemma 7.4.3 that $\hat{\Omega}_2$ is closed. It clearly restricts to Ω_2 along the slices $\mathcal{J}(\Sigma)\times\{(A,\psi)\}$ and $\{J\}\times\mathcal{A}(P)\times\Omega^1(\Sigma,\mathbf{i}\mathrm{ad}(P))$. For non-degeneracy let $(J,A,\psi)\in\mathcal{J}(\Sigma)\times\mathcal{A}(P)\times\Omega^1(\Sigma,\mathbf{i}\mathrm{ad}(P))$ and $(\hat{J},\hat{A},\hat{\psi})$ in the tangent space of (J,A,ψ) be given. Define

$$(\hat{J}',\hat{A}',\hat{\psi}'):=(-J\hat{J},\lambda\mathbf{i}\hat{\psi},\lambda\mathbf{i}\hat{A})$$

where $\lambda\in\mathbb{R}$ is a constant to be determined. Then

$$\begin{aligned}(\hat{\Omega}_2)_{(J,A,\psi)}\left((\hat{J},\hat{A},\hat{\psi}),(\hat{J}',\hat{A}',\hat{\psi}')\right)=\frac{1}{2}\int_\Sigma\mathrm{tr}(\hat{J}^2)\rho+\lambda\int_\Sigma\mathrm{tr}\left(-\hat{A}\wedge *\hat{A}+\hat{\psi}\wedge *\hat{\psi}\right)\\ -\mathbf{i}\lambda\int_\Sigma\mathrm{tr}\left(\hat{A}\wedge(\psi\circ J\hat{J})-\mathbf{i}\hat{\psi}\wedge(\psi\circ\hat{J})\right)\end{aligned}$$

This is strictly positive as $\lambda\to 0$ and thus $\hat{\Omega}_2$ is non-degenerated. □

Three moment maps

The following theorem calculates moment maps for the action of $\tilde{\mathcal{G}}(P)$ with respect to all three symplectic forms.

Theorem 7.4.5. *Let $V\in Lie(\tilde{\mathcal{G}}_{Diff_0}(P))\subset Vect(P)$ and denote its flow by $g_t\in\tilde{\mathcal{G}}_{Diff_0}(P)$. Then holds*

$$\mathcal{L}_V(J,A,\psi)=\left.\frac{d}{dt}\right|_{t=0}g_t^*(J,A,\psi)=\begin{pmatrix}2\bar{\partial}_J(v)\\ d_A A(V)\iota(v)F_A\\ -[A(V),\psi]+\iota(v)d_A\psi+d_A\iota(v)\psi\end{pmatrix}$$

for all $(J,A,\psi)\in\mathcal{J}(\Sigma)\times\mathcal{A}(P)\times\Omega^1(\Sigma,\boldsymbol{i}ad(P)$, where $v:=\pi_*V\in Vect(\Sigma)$ and $A(V)\in\Omega^0(\Sigma,ad(P))$ are defined by (7.21). The $\tilde{\mathcal{G}}(P)$ action on $\mathcal{J}(\Sigma)\times\mathcal{A}(P)\times$

$\Omega^1(\Sigma, \boldsymbol{i}ad(P)$ is Hamiltonian with respect to Ω_1, $\hat{\Omega}_2$ and Ω_3 and with moment maps

$$\mu_1(J, A, \psi) = \left(\left(F_A + \frac{1}{2}[\psi \wedge \psi]\right), (2K_J - 2c)\rho + dtr(\psi \Lambda_\rho(d_A\psi))\right) \tag{7.57}$$

$$\hat{\mu}_2(J, A, \psi) = (\boldsymbol{i}d_A * \psi, (2K_J - 2c)\rho + \boldsymbol{i}d * tr(\psi \Lambda_\rho(F_A))) \tag{7.58}$$

$$\mu_3(J, A, \psi) = \left(\boldsymbol{i}d_A\psi, (2K_J - 2c)\rho + \boldsymbol{i}dtr\left(\psi \Lambda_\rho\left(F_A + \frac{1}{2}[\psi \wedge \psi]\right)\right)\right) \tag{7.59}$$

where $c := 2\pi(2genus(\Sigma) - 2)/vol(\Sigma, \rho)$ and

$$\Lambda_\rho : \Omega^2(\Sigma, ad(P) \otimes \mathbb{C}) \to \Omega^0(\Sigma, \boldsymbol{i}ad(P) \otimes \mathbb{C})$$

is the natural map induced by ρ. All three moment map take values in the space $\Omega^2(\Sigma, ad(P)) \oplus \Omega^2_{ex}(\Sigma)$ which we identify with the dual space of $Lie(\tilde{\mathcal{G}}(P))$ using (7.16) and (7.17).

Proof. The formula for the infinitesimal action and the moment map equations for μ_1 and μ_3 follow from Theorem 4.4.2, Proposition 7.4.2 and integration by parts.

We verify the moment map equation for $\hat{\mu}_2$. Let $V \in \mathrm{Lie}(\tilde{\mathcal{G}}(P))$ with $\pi_* V = v_H$ for some Hamiltonian $H : \Sigma \to \mathbb{R}$. Using integration by parts, we have

$$\begin{aligned}
&\langle \hat{\mu}_2(J, A, \psi), V \rangle \\
&\quad = \mathbf{i} \int_\Sigma \mathrm{tr}\left(A(V) d_A *^J \psi\right) + \int_\Sigma H\left(2K_J - 2c\right)\rho - \mathbf{i} \int_\Sigma \mathrm{tr}\left(\iota(v_H) F_A \wedge *^J \psi\right)
\end{aligned}$$

The proof consists of differentiation this expression in all three arguments.

We first differentiate the moment map into direction $(\hat{J}, 0, 0)$. The first order change of $*^J\psi = \psi \circ (-J)$, when varying J in direction $\hat{J}$, is $\psi \circ (-\hat{J})$ and thus

$$\begin{aligned}
&\partial_{(\hat{J},0,0)} \langle \hat{\mu}_2(J, A, \psi), V \rangle \\
&\quad = \partial_{\hat{J}} \int_\Sigma H(2K_J - c)\rho + \mathbf{i} \int_\Sigma \mathrm{tr}\left(A(V)(d_A\psi \circ (-\hat{J})) - \iota(v_H) F_A \wedge \psi \circ (-\hat{J})\right) \\
&\quad = \frac{1}{2} \int_\Sigma (-2\bar{\partial}_J v_H) J \hat{J} \rho + \mathbf{i} \int_\Sigma \mathrm{tr}\left((-d_A A(V) - \iota(v_H) F_A) \wedge \psi \circ (-\hat{J})\right) \\
&\quad = (\hat{\Omega}_2)_{J,A,\psi}((-\mathcal{L}_V(J, A, \psi)), (\hat{J}, 0, 0))
\end{aligned}$$

where the penultimate equation follows from Theorem 4.4.2.

Next, we differentiate the moment map into direction $(0, \hat{A}, 0)$.

$$\begin{aligned}
&\partial_{(0,\hat{A},0)} \langle \hat{\mu}_2(J, A, \psi), V \rangle \\
&\quad = \mathbf{i} \int_\Sigma \mathrm{tr}\left(\iota(v_H) \hat{A} d_A * \psi + A(V)[\hat{A} \wedge *\psi] - \iota(v_H) d_A \hat{A} \wedge *\psi\right) \\
&\quad = \mathbf{i} \int_\Sigma \mathrm{tr}\left(-\hat{A} \wedge *[A(V), \psi]\right) + \mathbf{i} \int_\Sigma \mathrm{tr}\left((-d_A \iota(v_H)\hat{A} - \iota(v_H) d_A \hat{A}) \wedge *\psi\right)
\end{aligned}$$

Using Lemma 7.2.1 we can express the last term by Lie derivatives. We use the notation $\tilde{v}_H \in \mathrm{Vect}(\mathrm{ad}(P))$ for the A-horizontal lift of v_H and identify horizontal equivariant differential forms of the total space of $\mathrm{ad}(P)$ with differential forms on Σ taking values in $\mathrm{ad}(P)$. Then follows

$$\begin{aligned}
&\mathbf{i}\int_\Sigma \mathrm{tr}\left(-d_A\iota(v_H)\hat{A} - \iota(v_H)d_A\psi) \wedge *\psi\right)\\
&\quad= \mathbf{i}\int_\Sigma \mathrm{tr}\left(\mathcal{L}_{\tilde{v}_H}(\hat{A}) \wedge \psi \circ J\right)\\
&\quad= \mathbf{i}\int_\Sigma -\mathrm{tr}\left(\hat{A} \wedge \mathcal{L}_{\tilde{v}_H}(\psi \circ J)\right)\\
&\quad= \mathbf{i}\int_\Sigma -\mathrm{tr}\left(\hat{A} \wedge (((\mathcal{L}_{\tilde{v}_H}\psi) \circ J) + \psi \circ \mathcal{L}_{v_H} J)\right)\\
&\quad= \mathbf{i}\int_\Sigma \mathrm{tr}\left(\hat{A} \wedge \left(*d_A\iota(v_H)\psi * \iota(v_H)d_A\psi + \psi \circ (2\bar{\partial}_J v_H)\right)\right)\\
&\quad= \mathbf{i}\int_\Sigma \mathrm{tr}\left((d_A\iota(v_H)\psi + \iota(v_H)d_A\psi) \wedge *\hat{A}\right) - \mathbf{i}\int_\Sigma \mathrm{tr}\left(\hat{A} \wedge \psi \circ (-2\bar{\partial}_J v_H)\right)
\end{aligned}$$

Combining this with the computation above shows

$$\partial_{(0,\hat{A},0)}\langle \hat{\mu}_2(J,A,\psi), v\rangle = (\hat{\Omega}_2)_{J,A,\psi}((-\mathcal{L}_v(J,A,\psi)),\, (0,\hat{A},0)).$$

Finally, we differentiate the moment map into direction $(0,0,\hat{\psi})$.

$$\begin{aligned}
\partial_{(0,0,\hat{\psi})}\langle \hat{\mu}_2(J,A,\psi), V\rangle &= \mathbf{i}\int_\Sigma \mathrm{tr}\left(A(V)d_A * \hat{\psi} - \iota(v_H)F_A \wedge *\hat{\psi}\right)\\
&= \mathbf{i}\int_\Sigma \mathrm{tr}\left((-d_A A(V) - \iota(v_H)F_A) \wedge *\hat{\psi}\right)\\
&= (\hat{\Omega}_2)_{J,A,\psi}((-\mathcal{L}_V(J,A,\psi)),\, (0,0,\hat{\psi}))
\end{aligned}$$

This completes the proof of the moment map equation. □

The moduli space

The three moment maps μ_1, $\hat{\mu}_2$ and μ_3 calculated in Theorem 7.4.5 do not combine to a hyperkähler moment map. Nevertheless, when one considers their joint vanishing locus, the moment map equations greatly simplify and uncouple:

$$(J,A,\psi) \in \mu_1^{-1}(0) \cap \hat{\mu}_2^{-1}(0) \cap \mu_3^{-1}(0) \qquad \Longleftrightarrow \qquad \begin{cases} d_A\psi = 0,\ d_A^*\psi = 0\\ F_A + \frac{1}{2}[\psi \wedge \psi] = 0\\ 2K_J = c \end{cases}$$

By uniformization and Moser isotopy, Teichmüller space has the symplectic description

$$\mathcal{T}(\Sigma) := \mathcal{J}(\Sigma)/\mathrm{Diff}_0(\Sigma) \cong \{J \in \mathcal{J}(\Sigma) \,|\, K_J = c\} /\mathrm{Symp}_0(\Sigma,\rho)$$

(see Chapter 4.4 for more details). Consequently,

$$\left(\mu_1^{-1}(0) \cap \hat{\mu}_2^{-1}(0) \cap \mu_3^{-1}(0)\right) /\tilde{\mathcal{G}}_{\mathrm{Symp}_0}(P)$$

fibres over Teichmüller space, with the fibre being the corresponding Hitchin moduli space.

Bibliography

[1] Adams, J., and Fournier, J. *Sobolev Spaces*, 2nd edition ed., vol. 140 of *Pure and Applied Mathematics.* Academic Press, 2003.

[2] Anchouche, B., and Biswan, I. Einstein–Hermitian connections on polystable principal bundles over a compact Kähler manifold. *American Journal of Mathematics 123* (2001), 207–228.

[3] Asok, A., Doran, B., and Kirwan, F. Yang–Mills theory and Tamagawa numbers: the fascination of unexpected links in mathematics. *Bulletin London Mathematical Socienty 40* (2008), 533–567.

[4] Atiyah, M., and Bott, R. The Yang-Mills equations over Riemann surfaces. *Philosophical Transactions of the Royal Society of London A 308* (1983), 523–615.

[5] Aubin, T. *Nonlinear Analysis on manifolds. Monge–Ampere equations.* Springer, 1982.

[6] Bando, S., and Siu, Y. *Stable sheaves and Einstein–Hermitian metrics.* Geometry and Analysis on Complex Manifolds, World Scientific Publishing, River Edge, NJ, 1994.

[7] Banfield, D. Stable pairs and pricipal bundles. *The Quarterly Journal of Mathematics 51*, 4 (1999), 417–436.

[8] Bers, L. Simultaneous uniformization. *Bulletin of the American Mathematical Society 66* (1960), 94–97.

[9] Bers, L. *Spaces of Kleinian Groups*, vol. 155 of *Lecture Notes on Mathematics.* Springer, 1970.

[10] Biswan, I., and Wilkin, G. Morse theory for the space of Higgs G-bundles. *Geometriae Dedicata 149*, 1 (2010), 189–203.

[11] Bogomolov, F. Kähler manifolds with trivial canonical class. *Institute des Hautes Études Scientifiques Preprint* (1981), 1–32.

[12] Bradlow, S. Special metrics and stability for holomorphic bundles with global sections. *Journal of Differential Geometry 33* (1991), 169–213.

[13] Bradlow, S., Daskalopoulos, G., García-Prada, O., and Wentworth, R. Stable augmented bundles over Riemann surfaces. *Vector bundles in Algebraic Geometry, London Matematical Society Lecture Notes Series 208* (1995), 15–68.

[14] Bruasse, L. Optimal destabilizing vectors in some gauge theoretical moduli problems. *Annales de l'Institute Fourier 56* (2006), 1805–1826.

[15] Bruasse, L., and Teleman, A. Harder–Narasimhan filtrations and optimal destablilizing vectors in complex geometry. *Annales de l'Institute Fourier 55* (2005), 1017–1053.

[16] Burger, M., Iozzi, A., and Wienhard, A. Heigher Teichmüller spaces: form $SL(2, R)$ to other liegroups. handbook on Teichmüller theory. *IRMA Lectures in Mathematics and Theoretical Physics IV*, 19 (2014), 539–618.

[17] Calabi, E., and Chen, X. The space of Kähler metrics II. *Journal of Differential Geometry 61* (2002), 173–193.

[18] Chen, X. Space of Kähler metrics IV - on the lower bound of the K-energy. *http://arxiv.org/abs/0809.4081v2* (2008).

[19] Chen, X. Space of Kähler metrics III - one the lower bound of the Calabi energy and geodesic distance. *Inventiones Mathematicae 175*, 3 (2009), 453–503.

[20] Chen, X., Donaldson, S., and Sun, S. Kähler–Einstein metrics on Fano manifolds, I: Approximation of metrics with cone singularities. *Journal of the American Mathematical Society 28*, 1 (2015), 183–197.

[21] Chen, X., Donaldson, S., and Sun, S. Kähler–Einstein metrics on Fano manifolds, II: Limits with cone angle les then 2π. *Journal of the American Mathematical Society 28*, 1 (2015), 199–234.

[22] Chen, X., Donaldson, S., and Sun, S. Kähler–Einstein metrics on Fano manifolds, III: Limits as cone angle approaches 2π and completion of the main proof. *Journal of the American Mathematical Society 28*, 1 (2015), 235–278.

[23] Chen, X., and Sun, S. Calabi flow, geodesic rays, and uniqueness of constant scalar curvature Kähler metrics. *Annals of Mathematics 180* (2014), 407–454.

[24] Cieliebak, K., Gaio, A., i Riera, I. M., and Salamon, D. The symplectic vortex equations and invariants of Hamiltonian group actions. *The Journal of Symlectic Geometry 1* (2002), 543–645.

[25] Cieliebak, K., Gaio, A., and Salamon, D. J-holomorphic curves, moment maps, and invariants of Hamiltonian group actions. *International Mathematical Research Notes 10* (2000), 831–882.

[26] Cieliebak, K., i Riera, I. M., and Salamon, D. Equivariant moduli problems, branched manifolds, and the Euler class. *Topology 32* (2003), 641–700.

[27] Daskalopoulos, G. The topology on the space of stable bundles on a compact Riemann surface. *Journal of Differential Geomerty 36*, 3 (1992), 699–746.

[28] Daskalopoulos, G., and Wentworth, R. Convergence properties of the Yang–Mills flow on Kähler surfaces. *Journal für reine und angewandte Mathematik 575* (2004), 69–99.

[29] Demailly, J.-P. Complex analytic and differential geometry. *Monograph (https://www-fourier.ujf-grenoble.fr/~demailly/manuscripts/agbook.pdf)* (2012).

[30] Donaldson, S. A new proof of a theorem of Narasimhan and Seshadri. *Journal of Differential Geometry 18* (1983), 269–277.

[31] Donaldson, S. Anti self-dual Yang–Mills connections over complex algebraic surfaces and stable vector bundles. *Proceedings of the London Mathematical Society 50* (1985), 1–26.

[32] Donaldson, S. Infinite determinants, stable bundles and curvature. *Duke Mathematical Journal 54*, 1 (1987), 231–247.

[33] Donaldson, S. Twisted harmonic maps and the self-duality equation. *Proceedings of the London Mathematical Society 55*, 3 (1987), 127–131.

[34] Donaldson, S. Remarks on gauge theory, complex geometry, and 4-manifolds topology. *Fields medalist lecture, World Scientific* (1997), 384–403.

[35] Donaldson, S. Moment maps and diffeomorphism. *Surveys in Differential geometry - International Press 7* (2000), 107–127.

[36] Donaldson, S. Conjectures in Kähler geometry. *Clay Mathematical Proceedings (AMS) 3* (2002), 71–78.

[37] Donaldson, S. Scalar curvature and stability of toric varieties. *Journal of Differential Geomerty 62* (2002), 289–349.

[38] Donaldson, S. Moment maps in differential geometry. *Surveys in Differential Geometry - International Press 8* (2003), 171–190.

[39] Donaldson, S. Lower bounds on the Calabi functional. *Journal of Differential Geometry 70*, 3 (2005), 453–472.

[40] Donaldson, S. The Ding functional, Berndtsson convexity and moment maps. *Geometry, Analysis and Probability, Progress in Mathematics, Springer 310* (2017), 57–67.

[41] Early, C., and Eells, J. A fibre bundle description of Teichmüller theory. *Journal of Differential Geometry 3* (1969), 19–43.

[42] Farb, B., and Margalit, D. *A Primer on Mapping Class groups.* Princeton Mathematical, 2011.

[43] Feix, B. Hyperkähler metrics on cotangent bundles. *Journal für die reine und angewandte Mathematik 532* (2001), 33–46.

[44] Fels, G., Huckleberry, A., and Wolf, J. *Cycle Space of Flag Domains: A Complex Geometric Viewpoint.* Springer, 2005.

[45] Fernández, M., de León, M., and M.Saralegui. A six-dimension compact symplectic solvmanifold without Kähler structures. *Osaka Journal of Mathematics 33* (1996), 19–35.

[46] Fischer, A., and Tromba, A. On a purely Riemannian proof of the structure and dimension of the unramified moduli space of a compact Riemann surface. *Mathematische Annalen 267* (1984), 311–345.

[47] Friedman, A. *Partial differential equations of parabolic type.* Prentice-Hall, 1964.

[48] Fujiki, A. The moduli space of Kähler metrics on polarized manifolds. *Sugaku 42*, 1990 (1990), 231–243. (English trans. Sugaku Expositions 5 (1992) 173-191).

[49] García-Prada, O., Gothen, P., and i Riera, I. M. The Hitchin-Kobayashi correspondence, Higgs pairs and surface group representations. *http://arxiv.org/abs/0909.4487* (2009).

[50] García-Prada, O., Salamon, D., and Trautwein, S. Complex structures, moment maps, and the Ricci form. *https://arxiv.org/pdf/1805.00536.pdf* (2018).

[51] Georgoulas, V., Robbin, J., and Salamon, D. The moment-weight inequality and the Hilbert-Mumford criterion. *http://arxiv.org/abs/1311.0410* (2013).

[52] Goldman, W. The complex-symplectic geometry of $\mathrm{SL}(2,\mathbb{C})$-characters over surfaces. *Algebraic groups and arithmetic, Tata Institute of Fundamental Research Studies in Mathematics* (2004), 375–407.

[53] Griffiths, P., and Harris, J. *Principles of Algebraic Geometry.* Wiley Classics Library Edition, 1994.

[54] Gritsenko, V., Hulek, K., and Sankaran, G. Moduli of K3 surfaces and irreducible symplectic manifolds. *Handbook of Moduli, Vol. I, Advanced Lectures in Mathematics, International Press 24* (2013), 459–526.

[55] Guo, R., Huang, Z., and Wang, B. Quasi-Fuchsian 3-manifolds and metrics on Teichmüller space. *Asian Journal of Mathematics 14*, 2 (2010), 243–256.

[56] Hamilton, R. *Harmonic Maps of Manifolds with Boundary*, vol. 471 of *Lecture notes in Mathematics.* Springer, 1975.

[57] Hass, J. Minimal fibtrations of hyperbolic 3-manifolds. *https://arxiv.org/abs/1512.04145* (2015).

[58] HITCHIN, N. The self duality equations on a Riemann surface. *Proceedings of the London Mathematical Society 55* (1987), 59–126.

[59] HITCHIN, N. Lie groups and Teichmüller space. *Topology 31*, 3 (1992), 449–473.

[60] HITCHIN, N., KARLHEDE, A., LINDSTRÖM, U., AND ROCEK, M. Hyper-Kähler metrics and supersymmetry. *Communications in Mathematical Physics 108*, 4 (1987), 535–589.

[61] HODGE, T. *Hyperkähler geometry and Teichmüller space.* PhD thesis, Imperial College London, 2005.

[62] HUANG, Z., AND LUCIA, M. Minimal immersions of closed surfaces in hyperbolic three-manifolds. *Geometriae Dedicata 158*, 1 (2012), 397–411.

[63] HUANG, Z., AND WANG, B. On almost-Fuchsian manifolds. *Transactions of the American Mathematical Socienty 365* (2013), 4679–4698.

[64] HUYBRECHTS, D. *Lectures on K3 surfaces*, vol. 158 of *Cambridge Studies in Advanced Mathematics.* Cambridge University Press, 2016.

[65] JACOB, A. The limit of the Yang-Mills flow on semi-stable bundles. *Journal für reine und angewandte Mathematik 2015*, 709 (2013), 1–13.

[66] JACOB, A. The Yang-Mills flow and the Atiyah-Bott formula on compact Kähler manifolds. *American Journal of Mathematics 138*, 2 (2016), 329–365.

[67] KALEDIN, D. Hyperkähler metrics on total spaces of cotangent bundles. *Mathematical Physics Series, International Press 12* (1999).

[68] KEMPF, G. Instability in invariant theory. *Annals of Mathematics 108* (1978), 299–317.

[69] KIRWAN, F. *Cohomology of Quotients in Symplectic and Algebraic Geometry.* Princeton University Press, Mathematical Notes 31, 1985.

[70] KNAPP, A. *Lie Groups Beyond an Introduction.* Birkhäuser, 1996.

[71] KOISO, M. Einstein metrics and complex structures. *Inventiones Mathematicae 73* (1983), 71–106.

[72] KOSTANT, B. The principal three-dimensional subgroup and the betti numbers of a complex simple lie group. *American Journal of Mathematics 81* (1959). 973–1032.

[73] LABOURIE, F. $\mathbb{R}P^2$-structures et differentielles cubiques holomorphes. *Proceedings of the GARC Conference in Differential Geometry, Seoul Nationlal University* (1997).

[74] LABOURIE, F. Flat projective structures on surfaces and cubic holomorphic differentials. *Pure and applied mathematics quaterly 3*, 4 (2007), 1057–1099.

[75] LABOURIE, F. Cross ratios, anosov representations and the energy functional on Teichmüller space. *Annales Scientifiques de l'ENS 41*, 3 (2008).

[76] Lerman, E. Gradient flow of the norm squared of the moment map. *L'Enseignement Mathématique 51* (2005), 117–127.

[77] Lin, L. The heat flow for Kähler fibrations. *http://arxiv.org/abs/1211.5696* (2012).

[78] Lions, J., and Magenes, E. *Non-Homogeneous Boundary Value Problems and Application I.* Die Grundlehren der mathematischen Wissenschaften, 1972.

[79] Loftin, J. Affine spheres and convex $\mathbb{R}P^n$-manifolds. *American Journal of Mathematics 123*, 2 (2001), 255–274.

[80] Łojasiewicz, S. Une propriété topologique des sous-ensembles analytiques réels. *Colloques du CNRS, Les Équations du Centre National de la Recherche Scientifique, Paris 117* (1963), 87–89.

[81] Marden, A. The geometry of finitely generated Kleinian groups. *Annals of Mathematics 99*, 2 (1974), 383–462.

[82] Marsden, J., and Weinstein, A. Reduction of symplectic manifolds with symmetry. *Reports on Mathematical Physics 5* (1974), 121–130.

[83] Marsden, J., and Weinstein, A. *Quantization of Singular Symplectic Quotients*, vol. 198 of *Progress in Mathematics.* Birkhäuser, Basel, 2001, ch. Comments on the history, theory, and applications of symplectic reduction, pp. 1–19.

[84] McDuff, D. The moment map for circle actions on symplectic manifolds. *Journal of Geometry and Physics 5*, 2 (1988), 149–160.

[85] McDuff, D., and Salamon, D. *Introduction to Symplectic Topology.* Oxford Science Publications, 1998.

[86] McDuff, D., and Salamon, D. *J-holomorphic Curves and Symplectic Topology*, second ed., vol. 52. American Mathematical Society - Colloquium Publications, 2012.

[87] Mumford, D. Projective invariants of projective structures and applications. *Proceedings of the Internatational Congress of Mathematicians (Stockholm, 1962)* (1963).

[88] Mumfrod, D., Fogarty, J., and Kirwan, F. *Geometric Invariant Theory.* Springer, 1994.

[89] Mundet i Riera, I. A Hitchin–Kobayashi correspondence for Kähler fibrations. *Journal für reine und angewandte Mathematik 528* (2000), 41–80.

[90] Nanicinni, A. Weil–Petersson metric in the space of compact polarized Kähler–Einstein manifolds of zero first chern class. *Manuscripta Mathematica 54* (1986), 405–438.

[91] Narasimhan, M., and Seshadri, C. Stable and unitary vector bundles on a compact Riemann surface. *Annals of Mathematics Second Series 82* (1965), 540–567.

[92] Ness, L., and Mumford, D. A stratification of the null cone via the moment map. *American Journal of Mathematics 106*, 6 (1984), 1281–1329.

[93] Palais, R. *Foundations of global non-linear analysis.* W.A.Benjamin, Inc., 1968.

[94] Platis, I. Complex symplectic geometry of quasi-Fuchsian space. *Geometriae Dedicata 87* (2001), 17–34.

[95] Ramanathan, A. Stable principal bundles on a compact Riemann surface. *Mathematische Annalen 213* (1975), 129–152.

[96] Robbin, J., and Salamon, D. Introduction to differential topology. `https://people.math.ethz.ch/~salamon/PREPRINTS/difftop.pdf`.

[97] Råde, J. On the Yang-Mills heat equation in two and three dimension. *Journal für reine und angewandte Mathematik 431* (1992), 123–164.

[98] Schumacher, G. On the geometry of moduli spaces. *Manuscripta Mathematica 50* (1985), 229–267.

[99] Serre, J.-P. Représentations linéaires et espaces homogènes Kählériens des groupes de Lie compacts. *Séminaire Bourbaki 2* (1954), 447–454.

[100] Sibley, B. Asymptotics of the Yang-Mills flow for holomorphic vector bundles of Kähler manifolds: The canonical structure of the limit. *Journal für die reine und angewandte Mathematik 706* (2015), 123–191.

[101] Simon, L. Asymptotics for a class of non-linear evolution equations, with application to geometric problems. *Annals of Mathematics, Second Series 118* (1983), 525–571.

[102] Simpson, C. Yang–Mills theory and uniformisation. *Letters in Mathematical Physics* (1987).

[103] Simpson, C. Constructing variations of Hodge structure using Yang-Mills theory and applications to uniformization. *Journal of the American Mathematical Society 1*, 4 (1988), 867–918.

[104] Singer, I. The geometric interpretation of a special connection. *Pacific Journal of Mathematics 9* (1959), 585–590.

[105] Siu, Y. Curvature of the Weil–Petersson metric in the moduli space of Kähler–Einstein manifolds of negative first Chern class. *Contributions to Several Complex Variables (In Honour of Wilhelm Stoll), Aspekte der Mathematik, Springer* (1986), 261–298.

[106] Siu, Y. *Lectures on Hermitian-Einstein metrics for stable bundles and Kähler-Einstein metrics.* Birkhäuser, DMV Seminar, Band 8, 1987.

[107] Stein, E. *Singular Integrals and Differentiability Properties of Functions.* Princeton University Press, 1970.

[108] TAUBES, C. Minimal surfaces in germs of hyperbolic 3-manifolds. *Geometry & Topology Monographs Volume 7: Proceedings of the Casson Fest* (2004), 69–100.

[109] THOMAS, R. Notes on GIT and symplectic reductions for bundles and varieties. *Surveys in Differential Geometry 10* (2006), A Tribute to Profesor S. S. Chern.

[110] THURSTON, W. *The geometry and topology of three-manifolds.* Princeton University Press, 1980.

[111] TIAN, G. Smoothness of the universal deformation space of compact Calabi–Yau manifolds and its Petersson–Weil metric. *Mathematical aspects of String Theory, World Scientific* (1988).

[112] TIAN, G. Kähler–Einstein metrics with positive scalar curvature. *Inventiones Mathematicae 130* (1997), 1–37.

[113] TODOROV, A. The Weil–Petersson geometry of the moduli space of SU($n \geq 3$) (Calabi–Yau) manifolds I. *Communications in Mathematical Physics 126* (1989), 325–346.

[114] TRAUTWEIN, S. A survey of the GIT picture for the Yang-Mills equation over Riemann surfaces. *L'Enseignement Mathématique 63*, 2 (2017), 61–151.

[115] TRAUTWEIN, S. Convergence of the Yang-Mills-Higgs flow on gauged holomorphic maps and application. *International Journal of Mathematics 29*, 4 (2018), 1850024.

[116] TRIEBEL, H. *Interpolation Theory, Function Spaces, Differential Operators*, 2nd ed. Johann Ambrosius Barth Verlag, 1994.

[117] UHLENBECK, K. Closed minimal surfaces in hyperbolic 3-manifolds. *Seminar on Minimal Submanfiolds, Princeton University Press* (1983).

[118] UHLENBECK, K., AND YAU, S. On the existence of Hermitian Yang-Mills connections in stable vector bundles. *Communications on Pure and Applied Mathematics 36* (1986), 257–293.

[119] VENUGOPALAN, S. Yang–Mills heat flow on gauged holomorphic maps. *Journal of Symplectic Geometry 14*, 3 (2016), 903–981.

[120] VERBITSKI, M. Mapping class group and a global Torelli theorem for hperkähler manifolds. *Duke Mathematical Journal 162* (2013), 2929–2986.

[121] WANG, B. Minimal surfaces in quasi-fuchsian 3-manifolds. *Mathematische Annalen 354*, 3 (2012), 955–966.

[122] WEHRHEIM, K. *Uhlenbeck compactness.* EMS Series of Lectures in Mathematics, 2004.

[123] YAMADA, T. Examples of compact Lefschetz solvmanifolds. *Tokyo Journal of Mathematics 25*, 2 (2002), 261–283.

[124] Yau, S.-T. Calabi's conjecture and some new results in algebraic geometry. *Proceedings of the National Academy of Sciences USA 74*, 5 (1977), 1789–1799.

[125] Yau, S.-T. On the Ricci curvature of a compact Kähler manifold and the complex Monge–Ampere equation, I. *Communciations on Pure and Applied Mathematics 16* (1978), 339–411.