ADELAIDE UNIVERSITY AUSTRALIA

SUB CRUCE LUMEN

Yang–Mills connections on $U(n)$–bundles over compact Riemann surfaces

Peter Lawrence
Pure Mathematics Department
University of Adelaide [1]

January 24, 2003

[1] Masters by research thesis supervised by Dr. Nicholas Buchdahl

Abstract

This thesis was submitted as a part of a Masters by research degree in the School of Mathematical and Computer Sciences at the University of Adelaide during March 2002. Its aim is to provide a detailed dissertation on the solutions to the *Yang–Mills equations* over compact Riemann surfaces analysed in terms of algebro–topological and differential–geometric structures on vector bundles over such manifolds, in the spirit of the paper *The Yang–Mills equations over compact Riemann surfaces* [2] by *Michael Atiyah* and *Raoul Bott.*

The introduction gives a physical motivation for the subject of *Yang–Mills connections* which is aimed at familiarising the reader, at an informal level, with preliminary concepts in differential geometry needed in exploring this topic. Subsequent chapters will make specific the preliminary material and will also serve as the introduction of the main analytical and algebraic methods implemented in the study of Yang–Mills connections on Riemann surfaces.

The project is structured as follows. Following a brief overview of connections, the Yang–Mills functional and the associated equations are given. A subsequent section on *equivariant Morse theory* sets the framework for the thesis, while the following sections on relations to *stable bundles* and certain *moduli spaces* of semi–stable bundles serve as descriptive methods of solutions of the Yang–Mills equations on compact Riemann surfaces.

By restricting our attention to bundles with structure group $U(n)$ we may apply *Morse theory* to the Yang–Mills functional and *stratify* the space of connections. With this we will deduce information about the Yang–Mills minima by computing the equivariant cohomology of the moduli spaces $N(n,k)$ of stable bundles of rank n and degree k in the coprime case $(n,k) = 1$.

The level of complexity of this thesis is that understandable by honours graduate students of differential geometry and algebraic topology who have gone on to specialise in these fields.

1 Statement of submission.

This work contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text.

I consent to this copy of my thesis, when deposited in the University Library, being available for loan and photocopying.

The principal aid given to me in the production of this thesis was by *Dr. Nicholas Buchdahl* acting as primary supervisor, and *Dr. Michael Murray* as outside consultant, both members of the Pure Mathematics department in the University of Adelaide.

Mr. Peter Ernst Lawrence BSc(Ma & Comp Sc)(Hons)

Friday the 24th of January 2003.

Contents

2 Introduction.

All human knowledge thus begins with intuitions, proceeds thence to concepts, and ends with ideas.
Kant, Emmanual (1724 - 1804). Quoted in Hilbert's Foundations of Geometry.

A *connection* is a mathematical object defined on a *vector bundle* over a *manifold* which allows one to manipulate vectors in the different *fibers* in a consistent fashion. These ideas were developed by the efforts of mathematical pioneers such as *Johann Carl Friedrich Gauss*, *Georg Friedrich Bernhard Riemann* and *Elie Joseph Cartan* in the attempt to construct a differential calculus intrinsic to topological structures (manifolds) not necessarily embedded in Euclidean space.

There is often a necessitation of developments in mathematics motivated by physical theories, with this dependence sometimes assuming the reverse role. Such inter-relation between mathematics and physics is none more prevalent than with the theory of fiber bundles and connections; *Albert Einstein* utilised the already existing geometrical theories of Georg Riemann in devising *general relativity* in 1915, while *quantum (gauge) field theory* was later observed to be a theory of connections on the *space-time* manifold. The significance of the development of the theory of connections can be best appreciated in such physical settings. The familiar (intuitive) notion of differential calculus in $\mathbb{R}^n$ is given merit by the power of its implementation in *Newtonian mechanics* which serves as a fairly good approximation to "real-world" phenomena. With the advent of general relativity the luxury of a differential calculus in an ambient Euclidean space was abandoned for (at least, superseded by) a calculus intrinsic to topological structures known as *differentiable manifolds* which were not required to be embeddable in $\mathbb{R}^n$. No longer were physical fields to be considered as globally defined invariant functions but more appropriately as *sections* of vector bundles over manifolds whose representations change according to the choice of *coördinates* one works with. Thus the differential geometry of manifolds and fiber bundles are important disciplines of interest in physical theories which have merited much study independently in Pure Mathematics.

The notion of connections arises naturally when an attempt is made at constructing a differential calculus for the sections $\Gamma(E)$ of a vector bundle E: differentiation of sections takes the form of a differential operator (a connection)

$$\nabla : \Gamma(E) \longrightarrow \mathcal{E}^1(E)$$

whose codomain is the space of differential 1–forms taking values in E. One already has been exposed to connections in undergraduate differential geometry - the familiar *covariant differential* on a regular smooth surface $X : \Omega \subseteq \mathbb{R}^2 \longrightarrow \mathbb{R}^3$

$$\nabla_U V := (U^\alpha V^\beta_{,\alpha} + \Gamma^\gamma_{\alpha\beta} U^\alpha V^\beta) X_{,\gamma},$$

where $U := U^\alpha X_{,\alpha}$, $V := V^\alpha X_{,\alpha}$ are (tangential) vector fields on X and $\Gamma^\gamma_{\alpha\beta} := \frac{1}{2} g^{\gamma\delta}(g_{\beta\delta,\alpha} + g_{\alpha\delta,\beta} - g_{\alpha\beta,\delta})$ are the *Christoffel symbols of* 2^{nd} *kind* associated to the *metric tensor* $g_{\alpha\beta} := X_{,\alpha} \cdot X_{,\beta}$ on X. The differential operators ∇ share many similar traits to the usual derivative in $\mathbb{R}^n$, such as linearity and the Leibnitz rule, but their main difference with the latter is that the order of differentiation of sections is not necessarily interchangeable. This failure of the derivative operators to commute leads to the notion of *curvature*

$$F_\nabla := d_\nabla \circ \nabla : \Gamma(E) \longrightarrow \mathcal{E}^2(E)$$

on vector bundles associated with a connection ∇ and its induced operators $d_\nabla : \mathcal{E}^p(E) \longrightarrow \mathcal{E}^{p+1}(E)$; the richness in the theory of differential geometry on vector bundles is due to this inherent obstruction in differentiation.

Whether borne out of necessitation from physical theories or developed independently in a purely theoretical setting, the theory of connections on fiber bundles has been a very important field of study primarily due to the elegance of the intrinsic nature of these differential operators in the absence of an ambient Banach space structure.

Much in the same way that the choice of metric on a vector bundle is an arbitrary affair, there is no "natural" connection on a vector bundle due to the affine structure of the space of connections. One can refine the choice of connections by requiring compatibility with certain auxiliary structures on the manifold and vector bundle. For instance, there is physical significance in having derivatives preserve mensuration, thus once a vector bundle E is endowed with a metric g we can limit our attention to (*unitary*) connections ∇ preserving g in the sense that

$$dg(\sigma, \tau) = g(\nabla\sigma, \tau) + g(\sigma, \nabla\tau)$$

where d is the usual exterior derivative, σ and τ are sections of our vector bundle E, and the metric appearing on the right–hand side is the natural extension of our given metric to the space of differential 1–forms with values in E. The familiar *Levi–Civita* connection ∇ on the tensor bundle $(\bigotimes^k TX)\bigotimes(\bigotimes^l T^*X)$ over the *space–time manifold* (X^4, g) is an example of such a unitary connection, which appears in the celebrated *Einstein field equations*

$$G_{ab} := R_{ab} - \frac{1}{2}Rg_{ab} = 8\pi T_{ab}$$

relating the curvature $F_\nabla = [R^i_{jab}]dx^a \wedge dx^b$ of ∇ to the matter distribution in (X, g). When a vector bundle E is endowed with a *holomorphic structure* a further "natural" choice of connection is achieved by requiring our connection to have $(0,1)$–component equal to the partial connection $\overline{\partial}_E : \Gamma(E) \longrightarrow \mathcal{E}^{(0,1)}(E)$ associated with the holomorphic structure of E; the unique connection compatible with both the (Hermitian) metric– and holomorphic–structures on E is called the *metric connection* on E. There has been much study of metric connections on holomorphic vector bundles in the 1980's by mathematicians such as *M. Narasimhan*, *C. Seshadri* and *S. Donaldson* who related these connections to the notion of *stability* of vector bundles; a concept that will be of central importance in the differential–geometric analysis of a special class of connections known as *Yang–Mills connections.*

A further constraint can be made on connections, which will be of principle interest in this thesis, by requiring connections to be extrema of certain functionals. One important such setting is the study of the *Yang–Mills* functional on a compact oriented Riemannian manifold (X, g) for connections ∇ on a (smooth) vector bundle E over X with (Lie) structure group G

$$YM(\nabla) := \int_X |F_\nabla|^2 \, vol_X$$

where the norm of the curvature F_∇ of the connection ∇ is associated to the *Cartan–Killing* (bilinear) *form* on the Lie algebra of G and the natural inner–product on differential forms on X arising from the *Hodge–star operator* associated to the metric g. The critical points of YM are called *Yang–Mills connections* which, by the calculus of variations, are solutions to the system of partial differential equations

$$d_\nabla^* F_\nabla = 0$$

called the *Yang–Mills equations* where d_∇^* is the formal adjoint of the operator d_∇ with respect to the natural inner–product on the space of differential forms. In Physics terminology, the Yang–Mills equations are the *Euler–Lagrange equations* associated to the *action* of the *Yang–Mills Lagrangian* $\mathfrak{L}_{YM} := tr(F_\nabla \wedge *F_\nabla)$ where $*$ is the operator induced by the Hodge–star operator on the manifold X.

There is very much interest in the study of Yang–Mills connections due to their relation to *quantum field theory* in which the laws governing matter fields are formulated in a manner (referred to by physicists as "gauge theories") resembling the Yang–Mills equations. Such physical theories include the famous *Maxwell's* (vacuum) electromagnetic field theory represented by the equations

$$dF = 0 \quad \text{and} \quad d^*F = 0$$

for the electromagnetic field–strength tensor F, as well as the *quantum chromodynamical model* (Q.C.D.) with symmetry group $SU(3)$ of 1969 Nobel prize winning physicist *Gell-Mann* which describes the strong interactions, and the *electroweak model* with symmetry group $SU(2) \times U(1)$ of 1979 joint Nobel prize winners *Glashow, Weinberg* and *Salam* combining the electromagnetic and weak–nuclear interactions, paving the way, thanks to 1999 joint Nobel laureates *Hooft* and *Veltman*, for the enormously successful *standard-model* with symmetry group $U(1) \times SU(2) \times SU(3)$ unifying the electromagnetic, weak– and strong–nuclear forces.

In the early 70's it was recognised that the whole setting of gauge field theory was that of connections on vector bundles over the space–time manifold whose Lie structure group is taken to be the symmetry group of the interaction under study. The field strength could then be identified with the curvature of the connection and the action with the L^2–norm of the curvature. For instance, Maxwell's electromagnetic field theory is a Yang–Mills theory on a line bundle over $\mathbb{R}^4$ with structure group $U(1)$ for which the field–strength tensor is interpreted as the curvature $F = F_\nabla$ of a "field–potential" connection ∇ on this bundle with Maxwell's equations being the *Bianchi* and Yang–Mills equations for ∇; the arbitrariness in phase shift of the field–potential arises as the the action of the *gauge group* on the affine space of connections. This further inter–relation between physics and mathematics is a prime reason why the theory of connections, and in particular Yang–Mills connections, are topics of immense interest in mathematics today.

In recent years mathematical gauge theory has been perhaps the most important technique in the study of differentiable structures on four dimensional manifolds. In particular, the study of the *moduli spaces* of (*anti*) *self-dual* connections on on vector bundles over Riemannian four manifolds has yielded the definition, due to *S. K. Donaldson*, of polynomial invariants which have been highly successful at distinguishing smooth structures on homeomorphic manifolds. Thus studying the homotopy type of these moduli spaces continues to be of fundamental importance in algebraic topology.

Although the methods in this thesis are readily extendable to general *Riemannian manifolds*, we shall restrict our attention to the case of two (real) dimensional manifolds. When our base manifold X is of (real) dimension one all connections are *flat*, that is, have zero curvature, so the Yang–Mills theory in that case is not particularly interesting; thus our first "non–trivial" theory arises when our manifold X is of two real dimensions. This prototype theory merits a good deal of study due to the richness of structures naturally occurring on such manifolds, such as a *complex structure* associated (by the *Newlander–Nirenberg theorem* [28]) to the *almost-complex structure* determined by the *Hodge-star operator* $* : \mathcal{E}^p(X) \longrightarrow \mathcal{E}^{2-p}(X)$ on X (since $*^2 = -id$ on such X); with this endowed structure X becomes known as a *Riemann surface*. Moreover, smooth complex Hermitian vector bundles E over Riemann surfaces have inherent holomorphic structures due to the vacuous *integrability condition* on connections on E; this gives a correspondence between unitary connections ∇ and *holomorphic structures* $\overline{\partial}_E$ on E, thus the study of Yang–Mills connections on Riemann surfaces can be put into a complex–analytic framework. In such a setting, a Yang–Mills connection ∇ has *central curvature* $*F_\nabla$ being a *covariantly constant* holomorphic section of the algebra bundle $ad(E)$ with fibers isomorphic to the Lie algebra of the structure group of E. This naturally implies that the eigenvalues of the operator $i\,ad * F_\nabla$ are locally constant, and so if the bundle E is *indecomposable*, that is, has no proper sub–bundles, then $*F_\nabla = -2\pi i\, deg(E) id_E$ where the *degree*, $deg(E) \in \mathbb{Z}$, of E is the evaluation of the first *Chern class* on the fundamental cycle of X. *M. S. Narasimhan* and *C. S. Seshadri*, as did later *Donaldson*, showed that this implied that necessarily and sufficiently E was *stable* in the sense that $deg(F)/rk(F) < deg(E)/rk(E)$ for any non–zero proper sub–bundle $F < E$; (*semi-stability* means the possibility of equality in the preceding inequality).

When our smooth complex vector bundles E have structure group $U(n)$ and our Riemann surface is compact, we may convert the theory of Yang–Mills to the natural setting of *Morse theory* which deals with the analysis of the critical points of functions. The *critical manifolds* of the Yang–Mills functional other than the one corresponding to the minimum for YM can be shown to be expressed in terms of the minima of YM restricted to $U(m)$–sub–bundles for $m < n$.

Calculations in Morse theory are readily simplified if we can decompose a space M into a collection of locally closed submanifolds M_λ known as *strata*; the analysis of singularities of a function on such a space then localise to these strata. Under the aforementioned identification of the space of unitary connections $\mathcal{A}(E)$ with the collection $\mathcal{C}(E)$ of holomorphic structures on a smooth complex vector bundle E over a compact Riemann surface X, a *stratification* of $\mathcal{A}$ is obtained by stratifying $\mathcal{C}$ whose open strata correspond to the semi–stable holomorphic structures on E and the other strata described in terms of *canonical filtrations*: flags associated with holomorphic bundles whose respective quotients are semi–stable. Relative to the group $Aut(E)$ of automorphisms on E this stratification is *perfect* in the sense that

$$P_t(M) = \sum_\lambda t^{k_\lambda} P_t(M_\lambda)$$

where $k_\lambda := codim\, M_\lambda$ and P_t is the *Poincaré series*. Thus we may deduce information about the *equivariant cohomology* of the semi–stable stratum, and hence in the case $(n, k) = 1$, where k is the degree of E, about the cohomology of the *moduli spaces* $N(n, k) := \mathcal{C}_s(E)/Aut(E)$ of stable bundles – we restrict to this class $\mathcal{C}_s$ of bundles in order to avoid non–Hausdorff phenomena becoming prevalent in these moduli spaces. These moduli spaces, in a way, parametrise the solution space of the Yang–Mills equations.

These relations between Morse theory and complex–analytic geometry forged by Yang–Mills theory are testament to the richness of our two–dimensional prototype theory.

The theory of connections is often presented in many (equivalent) manners, usually by appealing to a class of structured fiber bundles known as *principal G–bundles*. For all intents and purposes, manipulations involving connections are made on vector bundles since they are, in a sense, easier objects to deal with, and as there is a direct correspondence between connections defined on principal bundles and vector bundles: explicitly, the matrix representation of a connection (covariant derivative) on a vector bundle E with respect to a trivialisation $\psi : E_{|U} \longrightarrow U \times \mathbb{R}^k$ is the pull–back $\psi^*(A)$ of a Lie algebra valued connection 1–form A on a principal *frame bundle* P associated to E representing the distribution of horizontal subspaces in the fibers of P.

3 The Yang–Mills equations.

In this section we shall introduce the important class of connection on a vector bundle known as *Yang-Mills connections*. These connections are the solutions of the *Yang-Mills equations* which arise when extremising the *Yang-Mills functional* on the space of connections. Although the results of this section could have well been defined on a compact Riemannian manifold, we shall limit our attention to Riemann surfaces for the aim of the thesis.

Let X be a compact Riemann surface, and $E[\mathbb{R}^k]$ a smooth vector bundle (of rank k) over X with compact Lie structure group[1] G whose corresponding Lie algebra is denoted by $\mathfrak{g}$.

For an open set $U \subseteq X$, denote by $\Gamma(U,E)$ the $C^\infty(X)$–module of smooth sections of E over U, and by $\mathcal{E}^p(U)$ the space of $\mathbb{C}$–valued smooth differential p–forms over U. We also write $\mathcal{E}^p(U,E) := \Gamma(U, E \bigotimes \bigwedge^p X) \cong \mathcal{E}^p(U) \bigotimes_{\mathcal{E}^0(U)} \Gamma(U,E)$. When $U = X$ we simply write $\Gamma(E)$ and $\mathcal{E}^p(E)$.

A *connection* (or *covariant derivative*) on E is a $\mathbb{R}$–linear operator

$$\nabla : \Gamma(E) \longrightarrow \mathcal{E}^1(E)$$

satisfying the *Leibnitz rule*

$$\nabla(f\sigma) = df \otimes \sigma + f\nabla\sigma \quad \text{where} \quad f \in \mathcal{E}^0(X) \quad \text{and} \quad \sigma \in \Gamma(E).$$

Given $\{g_{\alpha\beta}\}$ the transition functions representing E with respect to a trivialising cover $\{U_\alpha\}$ over X, the *adjoint map* $Ad : G \longrightarrow Aut(\mathfrak{g}) : g \mapsto \{Ad_g : L \mapsto gLg^{-1}\}$ induces transition functions $Ad(g_{\alpha\beta})$ for the *Lie algebra bundle* $ad(E) \subseteq End(E)$ whose fibers are thus isomorphic to $\mathfrak{g}$. A connection ∇ of E can thus be formally regarded as a differential operator of the form $d + A$ where d is the exterior derivative and $A \in \Gamma(ad(E) \bigotimes T^*X)$. Given a trivialising cover $\{(\psi_\alpha, U_\alpha)\}$ of E with an associated local frame $\{e_i^\alpha\}_{i=1}^k$ of E, for instance $e_i^\alpha = \phi_\alpha^{-1}(\,\cdot\,, \boldsymbol{e}_i)$, this operator acts on local sections $\sigma_i^\alpha e_i^\alpha$ of E by

$$\nabla(\sigma_i^\alpha e_i^\alpha) = d\sigma_i^\alpha \otimes e_i^\alpha + \sigma_i^\alpha A_{ij}^\alpha \otimes e_j^\alpha$$

where $A_{ij}^\alpha \in \mathcal{E}^1(U_\alpha)$ may be thought of as the entries of the matrix $A_{|U_\alpha} = [A_{ij}^\alpha]$ arising via $\nabla e_i^\alpha = A_{ij}^\alpha \otimes e_j^\alpha$.

The collection of all (smooth) connections on E, denoted $\mathcal{A}(E)$, is thus an affine space modelled over $\Gamma(X, ad(E) \bigotimes T^*X)$. One may extend a covariant derivative $\nabla : \Gamma(E) \longrightarrow \mathcal{E}^1(E)$ to a linear map

$$d_\nabla : \mathcal{E}^p(E) \longrightarrow \mathcal{E}^{p+1}(E)$$

by linear application of the formula

$$d_\nabla : \omega \otimes \sigma \mapsto d\omega \otimes \sigma + (-1)^p \omega \otimes \nabla\sigma$$

[1]We shall, in this thesis, assume G to be the *unitary group* $U(k/2)$ when E is a complex vector bundle. When G is a subgroup of the *classical* Lie groups such as $O(k)$, $SO(k)$, or $U(k/2)$, $SU(k/2)$ when E is a complex vector bundle, G is then semi-simple and hence the *Cartan-Killing form* on the Lie algebra $\mathfrak{g}$, used in defining the Yang–Mills functional, is positive–definite.

for $\omega \in \mathcal{E}^p(X)$ and $\sigma \in \Gamma(E)$. The resulting sequence

$$0 \longrightarrow \Gamma(E) \xrightarrow{\nabla} \mathcal{E}^1(E) \xrightarrow{d_\nabla} \mathcal{E}^2(E) \longrightarrow 0$$

is not necessarily a complex since the $\mathcal{E}^0(X)$–linear map

$$F_\nabla := d_\nabla \circ \nabla : \Gamma(E) \longrightarrow \mathcal{E}^2(E)$$

is not zero in general. Under the isomorphisms

$$\begin{aligned} Hom_{\mathcal{E}^0(X)}(\Gamma(E), \mathcal{E}^2(E)) &\cong Hom_{\mathcal{E}^0(X)}(\Gamma(E), \Gamma(E)) \bigotimes_{\mathcal{E}^0(X)} \mathcal{E}^2(E) \\ &\cong \Gamma(Hom(E,E)) \bigotimes_{\mathcal{E}^0(X)} \mathcal{E}^2(E) \\ &\cong \mathcal{E}^2(Hom(E,E)), \end{aligned}$$

we define the *curvature* associated to the connection ∇ as either the map $F_\nabla := d_\nabla \circ \nabla$ or the corresponding 2–form $F_\nabla \in \Gamma(ad(E) \bigotimes T^*X \bigwedge T^*X)$. Observe that if $\nabla = d + A$ then $F_\nabla = dA + A \wedge A$; that is, $F_\nabla(e_i) = dA_{ij} \otimes e_j + A_{ik} \wedge A_{kj} \otimes e_j$ for $\{e_i\}$ a local frame for E.

On the affine space $\mathcal{A}(E)$ of connections on E there is defined an important functional that will be of prime interest in this thesis. Given

$$\begin{aligned} k : \mathfrak{g} \times \mathfrak{g} &\to \mathbb{R} \\ : (L, K) &\mapsto tr(LK^*) \end{aligned}$$

the *Cartan-Killing form*[2] on the Lie algebra $\mathfrak{g}$ we construct an inner–product on the fiber $ad(E)_x \bigotimes \bigwedge_x^p X$ by linearly applying the formula

$$\langle L_1 \otimes \omega_1, L_2 \otimes \omega_2 \rangle := k(L_1, L_2) * (\omega_1 \wedge *\omega_2)$$

where $* : \bigwedge_x^p X \to \bigwedge_x^{2-p} X$ is the *Hodge-Star operator*[3] on X; with this we construct a global inner–product on $\mathcal{E}^p(ad(E))$ by linearity on the formula

$$(L_1 \otimes \omega_1, L_2 \otimes \omega_2) := \int_X \langle L_1(x) \otimes \omega_1(x), L_2(x) \otimes \omega_2(x) \rangle \, vol_X$$

where $vol_X \in \mathcal{E}^2(X)$ is the *volume form* given locally by $\sqrt{det(g)}\, dx_1 \wedge dx_2$ for g a metric on X.

With these preliminary definitions out of the way we can now define the *Yang-Mills functional.*

[2]For a general semi–simple compact Lie structure group G for our vector bundle E, we replace this definition of the Killing form with $k(L, K) := tr(ad_L \circ ad_K^*)$.

[3]This operator exists since X is an oriented manifold.

3.1. Definition.
The *Yang-Mills functional* on E is defined and denoted by

$$YM(\nabla) := \|F_\nabla\|_2^2 := (F_\nabla, F_\nabla)$$

where $\nabla \in \mathcal{A}(E)$ and F_∇ its associated curvature form.

In physics terminology, the Yang-Mills functional is the *action* associated to the *Lagrangian density*[4] $\mathcal{L}_{YM}(\nabla) := tr(F_\nabla \wedge *F_\nabla)$. The *Euler-Lagrange equations* corresponding to this action are the *Yang-Mills equations*; more succinctly:

3.2. Proposition.
A connection $\nabla \in \mathcal{A}(E)$ is extremal for the Yang–Mills functional if and only if

$$d_\nabla * F_\nabla = 0.$$

Proof:
By the methods of variational calculus ∇ will be extremal for YM if and only if the first variation

$$\delta YM(\nabla) := \frac{d}{dt} YM(\nabla_t)_{|t=0}$$

of YM vanishes at ∇, where $\nabla_t := \nabla + t\alpha$, $t \in [0,1]$ and $\alpha \in \mathcal{E}^1(ad(E))$. Given the bracket on $\mathcal{E}^1(ad(E))$ by the formula $[\omega, \gamma](v, w) = [\omega(v), \gamma(w)] - [\omega(w), \gamma(v)]$ for $v, w \in \mathfrak{X}(X)$, the curvature of the variation about ∇ becomes $F_{\nabla_t} = F_\nabla + t d_\nabla \alpha + \frac{1}{2}t^2[\alpha, \alpha]$. Thus $YM(\nabla_t) = YM(\nabla) + 2t(d_\nabla \alpha, F_\nabla) + O\{t^2\}$, and so

$$\delta YM(\nabla) = 2(d_\nabla \alpha, F_\nabla) = 2(\alpha, d_\nabla^* F_\nabla)$$

where $d_\nabla^* := - * \circ d_\nabla \circ *$ is the formal adjoint of d_∇. □

3.3. Definition.
Together with the *Bianchi identity* we have an elliptic system of partial differential equations on E

$$d_\nabla F_\nabla = 0, \qquad d_\nabla * F_\nabla = 0$$

known as the *Yang-Mills equations*. The solutions of the Yang–Mills equations (extrema for YM) are called *Yang-Mills connections*.

---

[4]This is actually a bundle–map between the *jet bundle* associated to $ad(E)$ and the bundle $\bigwedge^2 X$.

4 Equivariant Morse theory.

Yang–Mills theory, in the mathematical point of view, consists of analysing the fixed–points of the Yang–Mills functional YM. Classical *Morse theory* has been used as a means of studying multiple solutions of differential equations which arise in the calculus of variations. The theory is used to estimate the number of solutions by describing the local behaviour of the functional the differential equation came from.

Morse theory relates the analytic information of a smooth *Morse function* on a manifold by means of its critical set to the underlying topology of the manifold. In this way it is a natural framework to analyse the Yang–Mills equations. We shall in particular be interested in the extension of Morse theory to functions with symmetry under a compact Lie group action. In particular, we are interested in the action of the *gauge group* on the affine space of connections since it will be used to parametrise the solution space of YM by means of *moduli spaces.* The proofs of results for this specialised setting requires little modification of those of the classical theory (§7.2 [8]). In this way we give a complete overview of classical Morse theory and assume the same results when we pass over to Lie group *invariant* functions.

We shall in this section assume, unless stated otherwise, (X^n, g) to be a Riemannian manifold.

4.1. Definition.
Given $f \in \mathcal{F}(X)$ a smooth function, $x \in X$ is called a *critical point* for f if $d_x f = 0$ where

$$\begin{aligned} d_x f &: T_x X \to \mathbb{R} \\ &: v_i \left(\frac{\partial}{\partial x_i} \right)_x \mapsto v_i \frac{\partial f}{\partial x_i}(x) \end{aligned}$$

is the derivative of f; this is equivalent to requiring that $\frac{\partial f}{\partial x_i}(x) = 0$ for all $1 \leq i \leq n$ in any coördinate system about x. A real number $c \in \mathbb{R}$ is said to be a *critical value* for f if $f^{-1}(c)$ contains at least one critical point.

Let ∇_g the (unique) *Levi-Civita* connection on TX over the Riemannian manifold (X, g); this connection is *torsion free* in the sense that $(\nabla_g)_v w - (\nabla_g)_w v - [v, w] = 0$ for all $v, w \in \mathfrak{X}(X)$. The symmetric bilinear form

$$H_x f := \nabla_g \, df \, (x)$$

on $T_x X$ is called the *Hessian* of f at x, whose associated local matrix representation is given by

$$\left[\frac{\partial^2 f}{\partial x_i x_j}(x) - \frac{\partial f}{\partial x_k}(x) \Gamma^k_{ij}(x) \right]$$

where $\Gamma^k_{ij} := \frac{1}{2} g^{kl}(g_{il,j} + g_{jl,i} - g_{ij,l})$ are the *Christoffel symbols of* 2^{nd} *kind* associated to ∇_g; at a critical point x of f we clearly have the reduction of this matrix to the familiar form

$$\left[\frac{\partial^2 f}{\partial x_i x_j}(x) \right].$$

4.2. Definition.
A critical point x of f is said to be *non-degenerate* if the (matrix representation of the) Hessian is invertible [5] ; that is, if $detH_xf \neq 0$. One calls f a *Morse function* if all its critical points are non–degenerate. The *index* of a non–degenerate critical point x of f, denoted by λ_xf, is the number of negative eigenvalues of H_xf; this is equivalent to the maximal dimension of a subspace of T_xX upon which H_xf is negative–definite [6].

To a Morse function $f \in \mathcal{F}(X)$, we associate the *Morse counting series*

$$M_t(f) := \sum_{x:d_xf=0} t^{\lambda_xf}.$$

This is only well defined when X is compact, as is apparent from the following results.

4.3. Proposition.
Given $x \in X$ a non–degenerate critical point of a smooth function $f \in \mathcal{F}(X)$, there exists a coördinate chart (U, ϕ) about x with $\phi(x) = 0$ upon which

$$f \circ \phi^{-1}(y) = f(x) + \sum_{i=1}^{n} \delta_i y_i^2$$

where $\delta_i = \pm 1$ for $y \in \phi(U)$.
Proof:
Assume without loss of generality that $f(x) = 0$ (otherwise replace f by $f - f(x)$). For the purpose of local analysis, assume also for the time being that $f \in \mathcal{F}(V)$ where $V \subset \mathbb{R}^n$ is an open convex neighbourhood about 0, and that 0 is the non–degenerate critical point of f in context with $f(0) = 0$.

Given $y \in V$ since $\frac{\partial f}{\partial y_i}(0) = 0$ we have the Taylor series expansion of f about y

$$f(y) = \sum_{i=1}^{n} \sum_{j=1}^{n} y_i y_j f_{ij}(y), \tag{4.3.1}$$

$$\text{where } f_{ij}(y) := \int_0^1 \frac{\partial}{\partial y_j} f_i(ty)\, dt \text{ and } f_i(y) := \int_0^1 \frac{\partial}{\partial y_i} f(ty)\, dt.$$

If necessary, we replace f_{ij} by $\frac{1}{2}(f_{ij} + f_{ji})$ in order to ensure that $[f_{ij}]$ is symmetric. Observe that $[f_{ij}(0)]$ is invertible since $\frac{\partial^2 f}{\partial x_i \partial x_j}(0) = f_{ij}(0)$.

Assume for $1 \leq k \leq n$ our$f \in \mathcal{F}(V)$ has the form

$$f(y) = \sum_{i=1}^{k-1} \delta_i y_i^2 + \sum_{i=k}^{n} \sum_{j=k}^{n} y_i y_j f_{ij}(y) \tag{4.3.2}$$

[5]Non–singularity of the Hessian at a critical point of f is independent of the coördinate system used since $\frac{\partial^2 f}{\partial y_i y_j} = \frac{\partial^2 f}{\partial x_k x_l} \frac{\partial x_k}{\partial y_i} \frac{\partial x_l}{\partial y_j}$.

[6]By *Sylvester's theorem* in linear algebra, the index is independent of the choice of coördinates used in representing H_xf. It may be interpreted as the number of independent directions along which f is decreasing.

where $[f_{ij}] \in M_{(n-k+1)}(\mathcal{F}(V))$ is a symmetric matrix.

We may perform a linear change of the variables $y_k, ..., y_n$ so that equation (4.3.2) holds with $f_{kk}(0) \neq 0$; moreover, by continuity, we may assume $f_{kk}(y)$ is of constant sign $\delta_k = \pm 1$ for all $y \in V$. Setting $p := \sqrt{|f_{kk}|}$ we define

$$z_k := p(y)\left(y_k + \sum_{i=k+1}^{n} y_i \frac{f_{ik}(y)}{f_{kk}(y)}\right)$$
$$z_j := y_j \quad \text{for} \quad j \neq k, \ 1 \leq j \leq n.$$

By the inverse function theorem, $z_1, ..., z_n$ are then local coördinates in a neighbourhood of 0, and the change of variables from z to y defines a diffeomorphism Ψ so that in a neighbourhood of 0 we have $z = \Psi(y)$. Thus

$$\begin{aligned} f \circ \Psi^{-1}(z) &= f(y) \\ &= \sum_{i=1}^{k-1} \delta_i y_i^2 + y_k^2 f_{kk}(y) + 2y_k \sum_{j=k+1}^{n} y_j f_{jk}(y) + \sum_{i=k+1}^{n} \sum_{j=k+1}^{n} y_i y_j f_{ij}(y) \\ &= \sum_{i=1}^{k-1} \delta_i y_i^2 + f_{kk}(y)\left(y_k + \sum_{j=k+1}^{n} y_j \frac{f_{jk}(y)}{f_{kk}(y)}\right)^2 \\ &\quad - f_{kk}(y)\left(\sum_{j=k+1}^{n} y_j \frac{f_{jk}(y)}{f_{kk}(y)}\right)^2 + \sum_{i=k+1}^{n} \sum_{j=k+1}^{n} y_i y_j f_{ij}(y) \\ &=: \sum_{i=1}^{k} \delta_i z_i^2 + \sum_{i=k+1}^{n} \sum_{j=k+1}^{n} y_i y_j H_{ij}(y) \\ &= \sum_{i=1}^{k} \delta_i z_i^2 + \sum_{i=k+1}^{n} \sum_{j=k+1}^{n} z_i z_j H_{ij} \circ \Psi^{-1}(y), \end{aligned}$$

where $[H_{ij}] \in M_{(n-k)}(\mathcal{F}(V))$ is symmetric.

Thus, by induction on $1 \leq k \leq n$ we have that the smooth chart ϕ can be chosen such that $f \circ \phi^{-1}$ is given by (4.3.1) with $[f_{ij}] = diag(\pm 1, ..., \pm 1)_{n \times n}$.

□

By an additional permutation of coördinates in the above proposition we can put $f \in \mathcal{F}(X)$ into the "standard form"

$$f = f(x) - x_1^2 - x_2^2 - \cdots - x_k^2 + x_{k+1}^2 + \cdots + x_n^2$$

with respect to a suitable coördinate system $\phi = (x_1, ..., x_n)$ about x where $k = \lambda_x f$.

4.4. Corollary.
If X is compact, then a Morse function $f \in \mathcal{F}(X)$ has only finitely many critical points.

The relation between analysis and topology in Morse theory is expressed in the so–called *Morse inequalities.* Defining the *Poincaré series* for X relative to a field $\mathbb{F}$

$$P_t(X^n;\mathbb{F}) := \sum_{i=0}^{n} t^i \beta_i$$

where $\beta_i := dim_{\mathbb{F}} H^i(X^n, \mathbb{F})$ are the i^{th}–Betti numbers for a chosen cohomology theory (satisfying the *Eilenberg-Steenrod axioms*) relative to $\mathbb{F}$, the Morse inequalities determine a lower bound for $M_t(f)$ in terms of $P_t(X;\mathbb{F})$.

4.5. Theorem.
Given $f \in \mathcal{F}(X)$ a Morse function on the compact manifold X^n, upon denoting by m_i the number of critical points of f of index i, we then have the following relations

$$\begin{aligned} m_0 &\geq \beta_0, \\ m_0 - m_1 &\leq \beta_0 - \beta_1, \\ m_0 - m_1 + m_2 &\geq \beta_0 - \beta_1 + \beta_2, \\ &\vdots \\ \sum_{j=0}^{q}(-1)^{q-j} m_j &\geq \sum_{j=0}^{q}(-1)^{q-j}\beta_j, \quad 0 \leq q < n, \\ &\vdots \\ \sum_{j=0}^{n}(-1)^{j} m_j &= \sum_{j=0}^{n}(-1)^{j}\beta_j. \end{aligned}$$

These inequalities represent the formal *domination* of $P_t(X;\mathbb{F})$ by $M_t(f)$ in the sense that

$$M_t(f) - P_t(X;\mathbb{F}) = (1+t)Q(t)$$

where $Q(t)$ is a polynomial (formal series) with non–negative coefficients.

The key observation which provides this result is the change of homotopy type of the sublevel sets

$$X_a := \{x \in X \mid f(x) \leq a\}$$

when a crosses a critical value. For each critical point $x \in f^{-1}(a)$ a cell e_λ of dimension $\lambda = \lambda_x f$ is attached to X_a when a crosses the value $f(x)$; that is

$$X_{a+\varepsilon} \sim X_{a-\varepsilon} \cup_{x \in crit(f) \cap f^{-1}(a)} e_{\lambda_x f}$$

where $crit(f)$ is the set of critical points of f.

Starting from the absolute minimum (which exists if X is assumed compact) then one obtains a cell decomposition of X up to homotopy equivalence

$$X \sim \cup_{x \in crit(f)} e_{\lambda_x f}.$$

These observations are made precise as follows.

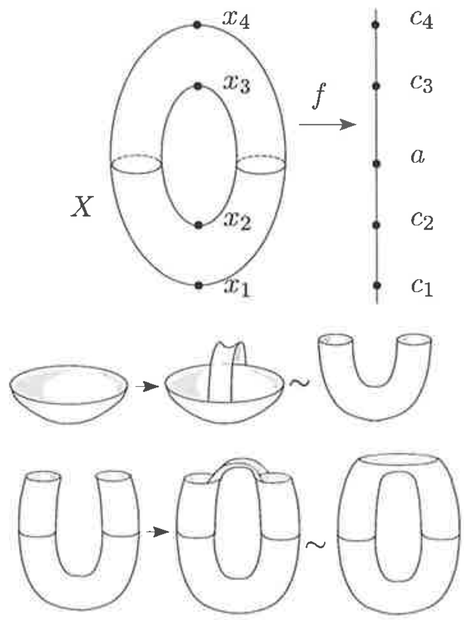


Figure 1: Given X the 2–Torus embedded in $\mathbb{R}^3$ and $f \in \mathcal{F}(X)$ the projection onto the vertical coördinate axis, the following two diagrams depict the crossing over a critical value.

4.6. Theorem.
Given $f \in \mathcal{F}(X)$ and $a < b$, suppose $f^{-1}[a,b]$ is compact and contains no critical points of f. Then X_a is diffeomorphic to X_b. Furthermore, X_a is a deformation retract of X_b, so that the inclusion $X_a \hookrightarrow X_b$ is a homotopy equivalence.

The deformation retracts required to prove this result are constructed by a so-called *1–parameter group of diffeomorphisms* of the manifold X. This is a smooth map

$$\phi : \mathbb{R} \times X \to X$$

such that

(i): for each $t \in \mathbb{R}$, the map $\phi_t : x \in X \mapsto \phi(t,x) \in X$ is a diffeomorphism of X onto itself;

(ii): for all $s, t \in \mathbb{R}$, $\phi_{t+s} = \phi_t \circ \phi_s$.

Given a smooth vector field $v \in \mathfrak{X}(X)$, by the *Picard–Lindelöf theorem* for ordinary differential equations, we have that for each $x \in X$ there exists an open interval $I_x \subset \mathbb{R}$ with $I_x \ni 0$ and a smooth curve $\gamma_x : I_x \to X$ with

$$\dot{\gamma_x}(t) = v_{\gamma_x(t)}, \qquad \gamma_x(0) = x.$$

Since the solution also depends smoothly on the initial point x, we furthermore have that for each $x \in X$ there exists an open neighbourhood U about x and an open interval $I \ni 0$ with the property that for all $y \in U$, the curve γ_y satisfying $\dot{\gamma_y}(t) = v_{\gamma_y(t)}$, $\gamma_y(0) = y$, called the *integral curve of v* *through* y, is defined on I. Furthermore, the map $(t,y) \in I \times U \mapsto \gamma_y(t) \in X$, called the *local flow of the vector field* v, is smooth.

The family of functions $\{\phi_t\}_{t\in I_y}$ given by $\phi_t(y) := \gamma_y(t)$ is observed to satisfy $\phi_{t+s}(y) = \phi_t \circ \phi_s(y)$ for all $s, t, t+s \in I_y$ since $\gamma_y(t+s) = \gamma_{\gamma_y(s)}(t)$: a "walk" from y along the integral curve for a time $t+s$ gives the same result as walking from $\gamma_y(s)$ for a time t. With this semi-group property, one has that $\phi_{-s} \circ \phi_s(y) = \phi_0(y) = y$ and so ϕ_t, defined on U, maps U diffeomorphically onto its image.

The semi–group $\{\phi_t\}_{t\in I_y}$ is called a *local 1–parameter group of diffeomorphisms of* X *generated by* v. In general, a local 1–parameter group need not be extendable to a group, since the maximal interval of definition I_y of γ_y need not be all of $\mathbb{R}$. The following result gives a condition under which this occurs.

4.7. Lemma.
A smooth vector field on X with compact support has a flow defined for all $t \in \mathbb{R}$ and for all $x \in X$, and the local 1–parameter group of diffeomorphisms becomes a group.
Proof:
By the preceding results, for every $x \in X$ there exists a neighbourhood U and an $\varepsilon > 0$ such that for all $y \in U$, the curve $\gamma_y(t)$ is defined for $|t| < \varepsilon$. Let *supp*$(v) \subset A$, a compact subset of X. A can be covered by finitely many such neighbourhoods, and we choose ε_o to be the smallest such ϵ.

As $v_{|X\setminus A} \equiv 0$, then $\phi_t(y) := \gamma_y(t)$ is defined on $(-\varepsilon, \varepsilon) \times X$, and for $|s|, |t| < \varepsilon_o/2$ we have the semi–group property $\phi_{t+s}(y) = \phi_t \circ \phi_s(y)$.

In order to define ϕ_t for $|t| \geq \varepsilon_o$, we expand t as $t = k(\varepsilon_o/2) + r$, $k \in \mathbb{Z}$ and $|r| < \varepsilon_o/2$. We define

$$\phi_t := \begin{cases} \underbrace{\phi_{\varepsilon_o/2} \circ \phi_{\varepsilon_o/2} \circ \cdots \circ \phi_{\varepsilon_o/2}}_{k-\text{fold}} \circ \phi_r, & k \geq 0 \\ \underbrace{\phi_{-\varepsilon_o/2} \circ \phi_{-\varepsilon_o/2} \circ \cdots \circ \phi_{-\varepsilon_o/2}}_{-k-\text{fold}} \circ \phi_r, & k < 0. \end{cases}$$

In this way ϕ_t is defined for all t, is smooth, and satisfies $\phi_{t+s} = \phi_t \circ \phi_s$ for all $t \in \mathbb{R}$.

□

With this result we can continue with our main task:

Proof of Theorem 4.6:
Given g a Riemannian metric on the Riemannian manifold X, we define the *gradient* vector field *grad* f of f by the equation

$$g(v, grad\, f) = v(f)$$

the directional derivative of f in the direction of $v \in \mathfrak{X}(X)$. One observes that the vector field *grad* f vanishes precisely on the critical set of f.

Let $\rho : X \to \mathbb{R}$ be a smooth function with compact support with

$$\rho = 1/g(grad\, f, grad\, f) \quad \text{on } f^{-1}[a,b].$$

The associated vector field $v \in \mathfrak{X}(X)$ defined by

$$v_x := \rho(x)(grad\, f)_x$$

thus satisfies the requirements of *Lemma 4.7* above, and so generates a 1–parameter group of diffeomorphisms

$$\phi_t : X \to X.$$

If $\phi_t(x) \in f^{-1}[a,b]$ then

$$\frac{d}{dt}f(\phi_t(x)) = g(\frac{d}{dt}\phi_t(x), grad\, f) = g(v, grad\, f) = 1$$

and so the function

$$t \mapsto f(\phi_t(x))$$

is linear with derivative 1 provided $a \leq f(\phi_t(x)) \leq b$. In this way the map ϕ_{b-a} is a diffeomorphism of X_a onto X_b; what we have done here is to "push" X_b "down to" X_a along the orthogonal trajectories of the hypersurfaces $f =$ constant.

Furthermore, the 1–parameter family of maps $r_t : X_b \to X_b$ defined by

$$r_t(x) := \begin{cases} x, & f(x) \leq a \\ \phi_{t(a-f(x))}(x), & a \leq f(x) \leq b. \end{cases}$$

is observed to satisfy $r_0 = 1_{X_b}$ with r_1 a retraction from X_b to X_a; thus X_a is a deformation retract of X_b.

□

4.8. Theorem.
Given $f \in \mathcal{F}(X)$ with non–degenerate critical point $x_o \in X$ of index λ, if $c = f(x_o)$ with $f^{-1}[c-\varepsilon, c+\varepsilon]$ compact and contains no critical point of f other than x for some $\varepsilon > 0$, then for all sufficiently small ε the set $X_{c+\varepsilon}$ is homotopy equivalent to the space obtained by attaching a λ–cell to $X_{c-\varepsilon}$.

Proof:
By *Proposition 4.3* we may choose a coördinate chart (U, ϕ) about x_o upon which f has the form

$$f = c - x_1^2 - x_2^2 - \cdots - x_\lambda^2 + x_{\lambda+1}^2 + \cdots + x_n^2.$$

Choose ε small enough so that $f^{-1}[c-\varepsilon, c+\varepsilon]$ is compact and contains no critical point of f other than x_o, and such that $\phi(U)$ contains the closed ball $\overline{B}_{2\varepsilon}(0) \subset \mathbb{R}^n$.

Define our λ-cell by

$$e_\lambda := \{x \in U \mid x_1^2 + \cdots + x_\lambda^2 \leq \varepsilon \text{ and } x_{\lambda+1} = \cdots = x_n = 0\}.$$

Observe that $e_\lambda \cap X_{c-\varepsilon} = \partial e_\lambda$ so that e_λ is attached to $X_{c-\varepsilon}$ as required. We must prove that $e_\lambda \cup X_{c-\varepsilon}$ is a deformation retract of $X_{c+\varepsilon}$.

Define the smooth function $F : X \to \mathbb{R}$ as f outside of U and by

$$F := f - \mu(x_1^2 + \cdots + x_\lambda^2 + 2x_{\lambda+1}^2 + \cdots + 2x_n^2)$$

in U where $\mu : \mathbb{R} \to \mathbb{R}$ is a smooth bump–function which vanishes outside the interval $(-2\varepsilon, 2\varepsilon)$ and with $\mu(0) > \epsilon$ and $-1 < \frac{d}{dr}\mu(r) \leq 0$ for all r.

For convenience let us write $f = c - \xi + \eta$ and hence

$$F(x) = c - \xi(x) + \eta(x) - \mu(\xi(x) + 2\eta(x))$$

for $\xi, \eta : U \to [0, \infty)$ given by

$$\xi := x_1^2 + \cdots + x_\lambda^2$$
$$\eta := x_{\lambda+1}^2 + \cdots + x_n^2.$$

The region $F^{-1}(-\infty, c+\epsilon]$ is observed to coincide with the region $X_{c+\varepsilon} =: f^{-1}(-\infty, c+\epsilon]$ since outside the region $\xi + 2\eta \leq 2\varepsilon$ the functions f and F coincide and within this region we have

$$F \leq f = c - \xi + \eta \leq c + \frac{1}{2}\xi + \eta \leq c + \varepsilon.$$

Another observation is that the critical points of f and F coincide since, by the above observation, $F(x_o) = c - \mu(0) < c - \varepsilon$ and so $F^{-1}[c-\varepsilon, c+\varepsilon] \subseteq f^{-1}[c-\varepsilon, c+\varepsilon]$ contains no critical points of F. This observation together with *Theorem 4.6* imply that $F^{-1}(-\infty, c-\varepsilon]$ is a deformation retract of $X_{c+\varepsilon}$. For convenience let us denote $F^{-1}(-\infty, c-\varepsilon]$ by $X_{c-\varepsilon} \cup H$ where H is the closure of $F^{-1}(-\infty, c-\varepsilon] \backslash X_{c-\varepsilon}$.

We shall complete the proof by showing that $X_{c-\varepsilon} \cup e_\lambda$ is a deformation retract of $X_{c-\varepsilon} \cup H$, which together with the previous observation proves our theorem. We define the deformation retract $r_t : X_{c-\varepsilon} \cup H \to X_{c-\varepsilon} \cup e_\lambda$ by

$:= id$ outside U;

$: (x_1, \ldots, x_n) \mapsto (x_1, \ldots x_\lambda, tx_{\lambda+1}, \ldots, tx_n)$ in the region $\xi \leq \varepsilon$;

$: (x_1, \ldots, x_n) \mapsto (x_1, \ldots x_\lambda, s_t x_{\lambda+1}, \ldots, s_t x_n)$ in the region $\varepsilon \leq \xi \leq \eta + \varepsilon$ where $s_t := t + (1-t)\sqrt{(\xi - \varepsilon)/\eta}$;

$:= id$ within the region $\eta + \varepsilon \leq \xi$.

□

With the same arguments in the preceding proof, one shows more generally that:

4.9. Theorem.
Given $f \in \mathcal{F}(X)$ with non–degenerate critical points $x_1, ..., x_r$ in $f^{-1}(c)$ of respective indices $\lambda_1, ..., \lambda_r$, suppose that $f^{-1}[c-\varepsilon, c+\varepsilon]$ is compact and does not contain any other critical points besides $x_1, ..., x_r$. Then $X_{c+\varepsilon}$ is homotopy equivalent to $X_{c-\varepsilon} \cup e_{\lambda_1} \cup \cdots \cup e_{\lambda_r}$, a space obtained by attaching cells to $X_{c-\varepsilon}$.

We are now at a position to prove the validity of the Morse inequalities. To this extent we require the following lemma.

4.10. Lemma.
Let $f \in \mathcal{F}(X)$ have non–degenerate critical points $x_1, ..., x_r$ in $f^{-1}(c)$ with respective indices $\lambda_1, ..., \lambda_r$. Suppose $a < c < b$ with $f^{-1}[a, b]$ compact and containing no critical points of f other than those above. Then

$$H^q(X_b, X_a) \cong \mathbb{F}^{m_q}$$

where m_q is the number of critical points of f of index q.
Proof:
We assume $r = 1$ and generalise our results by appealing to *Theorem 4.9*.
For ε as in *Theorem 4.8* and U a coördinate chart as in its proof, set

$$\begin{aligned} H_\delta &:= X_{c-\varepsilon} \cup \{x \in U \mid x_{\lambda+1}^2 + \cdots + x_n^2 \leq \delta^2\} \\ \text{and } H_{\delta,\eta} &:= X_{c-\varepsilon} \cup \{x \in U \cap H_\delta \mid x_1^2 + \cdots + x_\lambda^2 \geq \eta^2\} \end{aligned}$$

where $0 < \delta^2 \leq \varepsilon$ and $0 < \eta^2 < \varepsilon$. $X_{c-\varepsilon}$ is then a deformation retract of $H_{\delta,\eta}$ via the retraction

$$r(x) := \begin{cases} \rho_{\sigma(x)}(x), & x \in H_{\delta,\eta} \backslash X_{c-\varepsilon} \\ x, & \text{otherwise} \end{cases}$$

where $\rho_t(x) := ((1+t)x_1, ..., (1+t)x_\lambda, x_{\lambda+1}, ..., x_n)$ and

$$\sigma(x) := \begin{cases} 0, & x \in X_{c-\varepsilon} \\ \sup_{\rho_t(x) \notin X_{c-\varepsilon}} \{t \geq 0\}, & x \in H_{\delta,\eta} \backslash X_{c-\varepsilon}. \end{cases}$$

By the homotopy invariance of cohomology theories, the fact that $H^q(X, A) \cong 0$ if $A \subseteq X$ is a deformation retract of X, and by the standard relative cohomology sequence, we have that $H^q(X_b, X_a) \cong H^q(H_\delta, H_{\delta,\eta})$ where $0 < \eta^2 \leq \epsilon$ is arbitrary. By the *excision theorem* [7] on the exterior of $H_{\delta,\eta}$ we obtain a pair of spaces homotopy equivalent to $(e_\lambda \times \mathbb{R}^{n-\lambda}, \mathbb{S}^{\lambda-1} \times \mathbb{R} \times \mathbb{R}^{n-\lambda})$ and hence to $(e_\lambda, \mathbb{S}^{\lambda-1})$.
Using the identity

$$H^q(e_n, \mathbb{S}^{n-1}) \cong \begin{cases} \mathbb{F}, & q = n \\ 0, & \text{otherwise} \end{cases}$$

we have our required result.

□

Proof of Theorem 4.5
Let $c_1 < c_2 < \cdots < c_{k-1} < c_k$ be the critical values for f. Choose $b_0 < c_1$, $b_j \in (c_j, c_{j-1})$ for $1 \leq j \leq k-1$ and $b_k > c_k$.
The triple $Z \subseteq Y \subseteq X$ gives rise to the exact relative cohomology sequence

$$\begin{array}{cccccc} 0 & \longrightarrow & H^0(X,Y) & \longrightarrow & H^0(X,Z) & \longrightarrow \\ H^0(Y,Z) & \xrightarrow{\delta} & H^1(X,Y) & \longrightarrow & \cdots & \end{array}$$

[7] The *excision theorem* for singular cohomology states that given $U \subseteq Y \subseteq X$ manifolds with $dim X = dim Y$ and $\overline{U} \subseteq Y^\circ$ then $H^q(X, Y) \cong H^q(X \backslash U, Y \backslash U)$.

Upon defining

$$\beta_i(X,Y) := rk\, H^i(X,Y)$$
$$\beta_i(X,Y,Z) := rk\, im(\delta^i),$$

we have by the exactness of the above sequence that

$$\sum_{i=0}^{q}(-1)^i\Big(\beta_i(X,Y) - \beta_i(X,Z) + \beta_i(Y,Z) - (-1)^q\beta_i(X,Y,Z)\Big) = 0$$

hence

$$\begin{aligned}(-1)^{q-1}\beta_{q-1}(X,Y,Z) =& (-1)^q\beta_q(X,Y,Z) - (-1)^q\beta_q(X,Y) \\ &+ (-1)^q\beta_q(X,Z) - (-1)^q\beta_q(Y,Z).\end{aligned} \tag{4.10.1}$$

We define the following polynomials

$$P(t,X,Y) := \sum_{m\geq 0} t^m\beta_m(X,Y),$$
$$Q(t,X,Y,Z) := \sum_{m\geq 0} t^m\beta_m(X,Y,Z).$$

Multiplying the preceding equation (4.10.1) by $(-1)^m t^m$ and summing over m we obtain

$$Q(t,X,Y,Z) = -tQ(t,X,Y,Z) + P(t,X,Y) - P(t,X,Z) + P(t,Y,Z).$$

Applying these results to the triple $X_{b_0} \subseteq X_{b_{j-1}} \subseteq X_{b_j}$ and observing that $X_{b_0} = \emptyset$ since $c_1 = \min_{x\in X} f(x)$ we have

$$P(t,X_{b_j},X_{b_{j-1}}) = P(t,X_{b_j},\emptyset) - P(t,X_{b_{j-1}},\emptyset) + (1+t)Q(t,X_{b_j},X_{b_{j-1}},\emptyset),$$

and so as $X_{b_k} = X$ since $c_k = \max_{x\in X} f(x)$ we have upon summing the above relations

$$\sum_{l=1}^{k} P(t,X_{b_j},X_{b_{j-1}}) = P(t,X,\emptyset) + (1+t)Q(t)$$

for some $Q(t) \in \mathbb{Z}^+[t]$.

By *Lemma 4.10* we have

$$\sum_{j-1}^{k} P(t,X_{b_j},X_{b_{j-1}}) = \sum_{j-1}^{k} t^j m_j,$$

and as $H^j(X,\emptyset) = H^j(X)$ we also have that

$$P(t,X,\emptyset) = \sum_{j\geq 0} t^j\beta_j(X).$$

□

4.11. Corollary.
Let X^n be a compact smooth manifold and $f \in \mathcal{F}(X)$ a Morse function. Then

(i): $m_j \geq \beta_j$ for all j;

(ii): $\sum_{j=0}^{n}(-1)^j m_j = \chi(X)$ where $\chi(X) := \sum_{j=0}^{n}(-1)^j \beta_j(X)$ is the *Euler characteristic* of X;

(iii): $\sum_{j=0}^{q}(-1)^{q-j} m_j \geq \sum_{j=0}^{q}(-1)^{q-j}\beta_j$, $0 \leq q < n$.

4.12. Definition.
We call a Morse function $f \in \mathcal{F}(X)$ $\mathbb{F}$*-perfect* if

$$M_t(f) = P_t(X; \mathbb{F}),$$

and *perfect* if this relation holds for all fields $\mathbb{F}$.

Clearly a perfect Morse function can only exist on a *torsion free* manifold X; that is, X must have all its finitely generated cohomology modules $H^i(X, \mathbb{F})$ being free abelian groups – in general, if $H^i(X, \mathbb{F})$ is finitely generated, then by the structure theorem for groups we have $H^i(X, \mathbb{F}) \cong F_i \bigoplus T_i$ where F_i (the i^{th}*-Betti group*) is free abelian hence $\cong \bigoplus^k \mathbb{Z}$ and T_i (the i^{th}*-Torsion group*) decomposes uniquely as $\cong \mathbb{Z}_{d_{i1}} \bigoplus \cdots \bigoplus \mathbb{Z}_{d_{im_i}}$ where $d_{ij} | d_{i,j+1}$, $1 \leq j \leq m_i$, are the *torsion coefficients*.

If X were not torsion free with respect to the field $\mathbb{Z}_p$ for p a prime, we would have by the *universal coefficient theorem*[8]

$$\beta_j(X, \mathbb{Z}_p) = \beta_j(X, \mathbb{Z}) + t_j(p) + t_{j-1}(p)$$

where $t_j(p)$ is the number of torsion coefficients of $H_j(X, \mathbb{Z})$ divisible by p. Substituting this relation in the Morse inequalities we have that

$$\sum_{j=0}^{q}(-1)^{q-j} m_j \geq \sum_{j=0}^{q}(-1)^{q-j}\beta_j + t_q(p).$$

Choosing p so that $p | d_{q1}$, d_{ij} the torsion coefficients of $H_i(X, \mathbb{Z})$, then $p | d_{qj}$ for all j and so $t_q(p) = \nu_q$ the number of torsion coefficients in $H_q(X, \mathbb{Z})$. Thus we have

$$\sum_{j=0}^{q}(-1)^{q-j} m_j \geq \sum_{j=0}^{q}(-1)^{q-j}\beta_j + \nu_q$$

and so f cannot be $\mathbb{Z}_p$-perfect.

The following are criteria to ensure a perfect Morse function.

[8]The *universal coefficient theorem* states that for R a principal ideal domain $H_i(X, R) \cong H_i(X) \bigotimes R \bigoplus Tor(H_{i-1}(X), R)$.

4.13. Proposition.
A Morse function $f \in \mathcal{F}(X)$ is perfect if either

(i): (The lacunary principle):
if the set $\{\lambda_x f\}_{x \in crit(f)}$ of indices of f contains no consecutive integers;

(ii): (The completion principle):
if all the critical points x_o of f are *completable* in the sense that, in a suitable coördinate system U about x_o where

$$f = c - x_1^2 - \cdots - x_{\lambda_x f}^2 + x_{\lambda_x f+1}^2 + \cdots + x_n^2,$$

the boundary of the set

$$\nu_{x_o}^- := \{x \in U \mid x_1^2 + \cdots + x_{\lambda_x f}^2 \leq \varepsilon,\ x_{\lambda_x f+1} = \cdots = x_n = 0\}$$

bounds a singular chain in $X_{c-\varepsilon}$ for small enough $\varepsilon > 0$.

Proof:
(i):
If for some r we have $m_{r-1} = m_{r+1} = 0$ then by *Corollary 4.11(i)* we have $\beta_{r-1} = \beta_{r+1} = 0$. As $m_j - \beta_j = q_j + q_{j-1}$ for $r-1 \leq j \leq r+1$, for $Q(t) = \sum_{j \geq 0} t^j q_j$ the polynomial in the Morse inequalities with $q_j \geq 0$, then $q_j = 0$ for $r-2 \leq j \leq r+1$ hence $m_r = \beta_r$.
(ii):
If each $x_o \in crit(f)$ are completable, then in the proof of the Morse inequalities we have $\beta_j(X_{b_j}, X_{b_{j-1}}, X_{b_0}) = 0$ for all j, thus $Q(t) = 0$.

□

Having expounded an overview of classical Morse theory we now extend these principles as follows.

4.14. Definition.
A connected submanifold $Y \subseteq X$ of X is called a *non-degenerate critical manifold* if

$$d_x f_{|x \in Y} = 0 \quad \text{and} \quad H_Y f_{|\nu(Y)} \text{ is non} - \text{degenerate}$$

where $\nu(Y)$ is the *normal bundle* of Y (given by $TX_{|Y} \cong TY \bigoplus \nu(Y)$ with respect to a Riemannian metric on X). A function $f \in \mathcal{F}(X)$ is called *non-degenerate* if its critical set is the union of non-degenerate critical manifolds.

We extend the Morse counting series as follows. If (X, g) is a Riemannian manifold, we have an induced metric h on the normal bundle $\nu(Y)$ of a submanifold $Y \subseteq X$ and thus a self-adjoint endomorphism

$$A_Y : \nu(Y) \to \nu(Y)$$

defined by the formula

$$h(A_Y v, w) = H_Y f(v, w), \quad v, w \in \nu(Y).$$

As $H_Y f$ is non–degenerate A_Y has non–zero eigenvalues and hence decomposes $\nu(Y)$ orthogonally into a direct sum of negative– and positive–eigenspaces

$$\nu(Y) \cong \nu^-(Y) \bigoplus \nu^+(Y).$$

We call the rank of the "negative" bundle $\nu^-(Y)$ the *index* of Y, (as a critical manifold of f), denoted by $\lambda_Y f$.

Given $f \in \mathcal{F}(X)$ a non–degenerate function with $Y \subseteq X$ a non–degenerate critical manifold of f, with respect to a field $\mathbb{F}$ we define the polynomial

$$M_t(f, Y) := \sum_i t^i dim_{\mathbb{F}} H_c^i(\nu^-(Y))$$

where H_c^i denotes the compactly supported cohomology. When $\nu^-(Y)$ is orientable we have, by the *Thom isomorphism* [9], a reduction of this polynomial to

$$t^{\lambda_Y f} P_t(Y).$$

With these results we define

$$M_t(f) := \sum_Y M_t(f, Y)$$

where we sum over the critical manifolds of f.

Given $\varepsilon > 0$, let $\nu_\varepsilon^-(Y)$ denote the set in the exponential image [10] of $\nu^-(Y)$ in X where $f \geq f(Y) - \varepsilon$; if ε is small enough, $\nu_\varepsilon^-(Y)$ is a $\lambda_Y f$–disc bundle over Y.

We have a parallel result to *Theorem 4.9* given as follows.

4.15. Theorem.
Given $f \in \mathcal{F}(X)$ with critical value $c \in \mathbb{R}$ and $f^{-1}(c) = \{Y_1, Y_2, \ldots, Y_r\}$ consisting of non–degenerate critical manifolds, there exists an $\varepsilon > 0$ such that if $f^{-1}[c - \varepsilon, c + \varepsilon]$ is compact then

$$X_{c+\varepsilon} \sim X_{c-\varepsilon} \cup \bigcup_{j=1}^{r} \nu_\varepsilon^-(Y_j).$$

The proof of this result is essentially the same as that for *Theorem 4.9* with the deformation constructed along the fibers. With this result we have that the Morse inequalities also hold with

[9] Given an oriented vector bundle of rank k with total space E over a compact manifold X, the *Thom isomorphism* theorem calculates the compactly supported cohomology $H_c^*(E)$ in terms of $H^*(X)$; namely, $H^p(X) \cong H_c^{p+k}(E)$.

[10] Given X a Riemannian manifold, for each $x \in X$, $v \in T_x X$ there exists a maximal interval $I_v \subset \mathbb{R}$ containing 0 and a geodesic $\gamma_v : I_v \to X$ with $\gamma(0) = x$ and $\dot\gamma(0) = v$. We define the *exponential map* $exp : C \to X : v \mapsto \gamma_v(1)$ on the star–shaped neighbourhood $C := \{ v \in TX \mid 1 \in I_v \}$ of the zero section of TX. Thus exp_x maps a neighbourhood of $0 \in T_x X$ diffeomorphically onto a neighbourhood of $x \in X$.

respect to this extended Morse counting series, and thus we say that a non-degenerate f is an $\mathbb{F}$-*perfect Morse function* if

$$M_t(f) = P_t(X)$$

where the field $\mathbb{F}$ is used in the cohomology modules on both sides of this equation.

Observe that if our critical manifolds are points, that is if $Y = \{x\}$, then $\nu(Y) = T_xX$, and so $\nu^-(Y)$ is the largest subspace of T_xX upon which H_xf is negative-definite, and so $dim\ \nu^-(Y) = \lambda_x f$ thus $M_t(f, Y) = t^{\lambda_x f}$ as in the classical theory.

The completion process is extended via the following commutative diagram where Y is said to be $\mathbb{F}$-*completable* if the map α is zero; here π is the projection of the disc-bundle $\nu_\varepsilon^-(Y)$ and $\widetilde{H}$ the reduced homology modules over $\mathbb{F}$.

$$\begin{array}{ccccc} H_\bullet(\nu_\varepsilon^-(Y)) & \longrightarrow & \widetilde{H}_\bullet(\nu_\varepsilon^-(Y), \partial\nu_\varepsilon^-(Y)) & \xrightarrow{\partial} & \widetilde{H}_{\bullet-1}(\partial\nu_\varepsilon^-(Y)) \\ & & \pi^*\uparrow & & \downarrow \\ & & H_{\bullet-\lambda_Y f}(Y) & \xrightarrow{\alpha} & \widetilde{H}_{\bullet-1}(X_{f(Y)-\varepsilon}) \end{array} \quad (4.15.1)$$

Observe that π^* corresponds to the Thom isomorphism since $H^{q-k}(Y) \cong H_c^q(\nu_\varepsilon^-(Y))$ by the Thom isomorphism (Theorem 2.13 [24]) which in turn is isomorphic to $H^{n-q+k}(\nu_\varepsilon^-(Y))^*$ (§21 [24]), and finally isomorphic to $H_q(\nu_\varepsilon^-(Y), \partial\nu_\varepsilon^-(Y))$ by the *Lefschetz duality* [11].

4.16. Proposition.
If all critical manifolds of f are $\mathbb{F}$-completable then f is an $\mathbb{F}$-perfect Morse function on M.

The extended notion of non-degeneracy to critical manifolds has the advantage of being functorial under pull back. More succinctly,

4.17. Proposition.
Given $\pi : E \to X$ a smooth fibration, $f \in \mathcal{F}(X)$ is non-degenerate if and only is $\pi^* f = f \circ \pi$ is non-degenerate on E. Moreover, if Y is a non-degenerate critical manifold of f then $\lambda_Y f = \lambda_{\pi^{-1}(Y)} \pi^* f$.
Proof:
These results follow from the fact that for $p \in \pi^{-1}(Y)$, $d_p\pi^* = d_{\pi(p)} f \circ d_p\pi$, $\nu(\pi^{-1}(Y)) = (d\pi)^{-1}\nu(Y)$, and $H\pi^* = J \circ Hf \circ d\pi + df \circ H\pi$ where J is the change of coördinate matrix from E to X.

□

We now come to extending the main results of Morse theory to the equivariant setting; namely, for functions possessing a certain symmetry under a compact Lie group action. *Equivariant Morse theory* studies the Morse relations and the Morse handle body theorem (the attaching of cells upon passing critical points) for *G-invariant functions* $f : X \to \mathbb{R}$ for G a compact Lie group acting on a smooth manifold X. These are functions satisfying $f(g \cdot x) = f(x)$ for all $x \in X$ and for all

[11] The *Lefschetz duality* states (Theorem 28.18 [16]) that for M^n a compact manifold with boundary, $H^q(M) \cong H_{n-q}(M, \partial M)$.

$g \in G$. As stated at the beginning of this section, the proofs of results in this setting of Morse theory follow from those of classical Morse theory with the associated deformation retracts being equivariant functions(§7.2 [8]), and so we shall not give any proofs here.

4.18. Definition.
Given G a compact Lie group, a *G–space* X is a topological space (or manifold, depending on context) with continuous G–action. Let X be a G–space. A set $A \subset X$ is called *G–invariant* if $g \cdot x \in A$ for all $x \in A$ and for all $g \in G$. A map $F : X \to Y$ between two G–spaces is called *G–equivariant* if $F(g \cdot x) = g \cdot F(x)$ for all $x \in X,\ g \in G$.

Notice that for a G–invariant function, if x is a critical point then the points on the G orbit containing x are also critical points.

The set of all G–orbits is called the *orbit space*. Endowed with the quotient topology it is denoted by X/G or simply $\overline{X}$. Observe that a G–equivariant function naturally induces a map $\overline{F} : \overline{X} \to \overline{Y}$ between orbit spaces.

A fiber bundle $\pi : E \to X$ with structure group G is called a *G–bundle* if for all $g \in G$ the fiberwise multiplicative map $g : E \to E$ is a differentiable bundle map such that $gE_x = E_{g \cdot x}$ for all $x \in X$. Thus if X is a G–manifold the tangent bundle TX is a G–bundle with $g \cdot v = d_x\psi(g, x)(v)$ for all $v \in T_xX$ for all $x \in X$. A fiber bundle $\pi : E \to X$ is called a *Riemannian G–vector bundle* if it is a G–bundle and possesses a Riemannian metric such that the G–action is an isometry. Assume X is a Riemmanian manifold and $Y \subseteq X$ is a connected compact submanifold. Then $TY \leq TX$ and so the normal bundle $\nu(Y)$ is also a sub–bundle of TX. If in addition X is a G–manifold and Y is G–invariant, then both TY and $\nu(Y)$ are G–bundles.

Let $f \in C^1(X, \mathbb{R})$ be G–invariant. This gives rise to a G–equivariant gradient vector field $grad\, f$ given by

$$\langle grad_{g \cdot x} f, g \cdot v \rangle = \langle d_x f, v \rangle$$

for all $(g, x) \in G \times X$ and for all $v \in T_xX$; that is, $g^* \cdot grad\, f \cdot g = grad\, f$. As the action g on T_xX is unitary, hence $dg^* = g^{-1}$, we obtain $grad\, f \cdot g = g \cdot grad\, f$. Analogously, the Hessian Hf is also G–equivariant if $f \in C^2(X, \mathbb{R})$.

One observes that the sets X_c, $f^{-1}(c)$ and the critical sets $K_c := crit(f) \cap f^{-1}(c)$ are G–invariant. Also, a critical orbit $O := O(x)$ is a G–submanifold of X. It follows that $T_xO \leq ker(H_xf)$ and that the bounded self–adjoint operator $H_xf : \nu_x(O) \to \nu_x(O)$ satisfies $g^* \cdot H_{g \cdot x}f \cdot g = H_xf$.

In applying Morse theory to G–spaces X, there is much advantage in having a free action. In this case the orbit space $\overline{X} := X/G$ is a manifold and $\pi : X \to \overline{X}$ is a smooth fibration [12] with fiber G. Thus we can carry out Morse theory on $\overline{X}$; that is, f is a perfect equivariant Morse function if the induced function $\overline{f}$ on $\overline{X}$ is perfect in the usual sense.

On the other hand, when the action of G on X is *not* free, $\overline{X}$ may possess singularities and one cannot implement Morse theory on such a space as easily. The remedy to this situation is to appeal to homotopy theory and convert to free actions without changing the homotopy of the space on which the group acts; thus we carry out Morse theory of an induced function f_G on a new space X_G called the *homotopy quotient*. These ideas are made precise as follows.

[12] See Appendix B for the definition of a fibration.

Consider a principal G–bundle E over a manifold $B \approx E/G$ [13]; in particular we may choose E to be a *universal G–bundle* which is a bundle unique up to homotopy with contractible total space (a concrete example of such a bundle is given in Appendix B). Now G operates on $E \times X$ *diagonally*, that is $g \cdot (p, x) := (gp, g \cdot x)$, and this action is free since the action of G on E is free. The pull–back of a G–invariant function f on X to a function on $E \times X$ is G–invariant under the diagonal action, and hence induces a smooth function f_G on the *homotopy quotient* $X_G := E \times_G X := (E \times X)/G$, which is a fiber bundle over $BG := B$ with fiber X and structure group G. Homotopy quotients have the properties that $\{x\}_G \sim BG$, $X_G \sim BG$ if X is contractible, and $X_G \sim \overline{X}$ if the action of G on X is free.

Upon extending the Morse relations and the Morse handle body theorem to the equivariant setting we obtain the following main result.

4.19. Proposition.
Given Y a non–degenerate critical manifold of f on X, the corresponding non–degenerate critical manifold of f_G on X_G is Y_G, and $\lambda_Y f = \lambda_{Y_G} f_G$. Moreover, the equivariant Poincaré series is

$$P_t^G(Y) := P_t(Y_G) = \sum_i t^i dim\, H^i(Y_G) =: \sum_i t^i dim\, H^i_G(Y)$$

and Y contributes to $M_t(f_G)$, the *counting series* of f_G on X_G, by $t^{\lambda_Y f} P_t(Y_G)$.

Notice that if the non–degenerate critical manifold Y of f consists of a single G–orbit with stability group (the stabiliser) H, that is, $Y = G/H$, then f_G will have corresponding non–degenerate critical manifold $BH := E/H$, whence Y contributes $t^{\lambda_Y f} P_t(BH)$ to $M_t(f_G)$.

If Y is the orbit of G through p, then H the stability group of p acts on the normal space to Y and, using an H–invariant metric, also on the *negative normal space* $\nu_p^-(Y)$. It follows that $\nu^-(Y)$ is associated with the principle bundle $G/H = Y$, via this representation, and correspondingly that $\nu^-(BH)$ is associated with the universal H–bundle $EH := E$ over $BH := EH/H$, by the same representation.

Given a G–pair (X, Y) and a field $\mathbb{F}$, the cohomology

$$H^\bullet_G(X, Y; \mathbb{F}) := H^\bullet(X_G, Y_G; \mathbb{F})$$

is called the *G–equivariant cohomology. A. Borel* proved that G–equivariant cohomology satisfies most of the properties of general cohomology (§4 [8]), namely homotopy invariance, the excision principle and resultant exact sequences, however the dimension axiom does not hold since $H^\bullet_G(\{x\}) \cong H^\bullet(BG)$.

A method of simplifying calculations in Morse theory, as will be utilised later in the Yang–Mills theory, is to partition a space into a collection of submanifolds on which the singularities of the space are isolated. This methodology is known as *Whitney stratification* of a space into *strata*. This is defined more precisely as follows.

[13]The definition of a principal G–bundle is given in Appendix B.

Given $\mathcal{S}$ a partially ordered set and Y a closed subset of a smooth manifold X, a $\mathcal{S}$*–decomposition* of Y is a locally finite collection of disjoint locally closed subsets $S_i \subset Y$ for each $i \in \mathcal{S}$, called *strata*, such that

$$Y = \bigcup_{i \in \mathcal{S}} S_i \quad \text{and} \quad S_i \cap \overline{S}_j \neq \emptyset \;\Leftrightarrow\; S_i \subset \overline{S}_j \;\Leftrightarrow\; i = j \;\text{ or }\; i < j$$

and we write $S_i < S_j$. This decomposition is called a *Whitney stratification* of Y if:

(i): each stratum S_i is a locally closed smooth submanifold of X;

(ii): whenever $S_\alpha < S_\beta$, the pair satisfies the following *Whitney conditions*: given sequences $\{x_i\} \subseteq S_\beta$ and $\{y_i\} \subseteq S_\alpha$ both converging to some $y \in S_\alpha$, and suppose that (with respect to some coördinate system on X) the secant lines $l_i := \overline{x_i y_i}$ converge to some limiting line l and that $T_{x_i} S_\beta$ converges to some limiting plane τ. Then

(iia): $T_y S_\alpha \subset \tau$;

(iib): $l \subset \tau$.

Now suppose that Y is a compact Whitney stratified subspace of a manifold X and f is the restriction to Y of a smooth function on X. We define a *critical point* of f to be a critical point of the restriction of f to any stratum; in particular, the zero–dimensional strata are critical points.

In stratified Morse theory we consider Whitney stratified spaces X embedded in some smooth manifold M. We say that a function f on X is smooth if it is the restriction to X of a smooth function on M. By definition, Morse functions on a Whitney stratified space X are defined by the following three properties:

(1): the critical values of f must be distinct;

(2): at each critical point x of f, the restriction of f to the stratum S containing x is non–degenerate;

(3): $d_x f$ for $x \in crit(f)$ does not annihilate any limit of tangent spaces to any stratum S' other than the stratum S containing x.

Given a subset X of some smooth manifold M and a function $f : X \to \mathbb{R}$ which is the restriction of a smooth function $\widetilde{f} : M \to \mathbb{R}$ and fixing a Whitney stratification on X, a *critical point* of such a function f is any point $x \in X$ such that $d_x \widetilde{f}_{|T_x S} = 0$ where S is the stratum of X containing x. A *Morse function* $f : X \to \mathbb{R}$ is then the restriction of a smooth function $\widetilde{f} : M \to \mathbb{R}$ such that

(i): $f = \widetilde{f}_{|X}$ is proper and the critical values of f are distinct;

(ii): for each stratum S of X, the critical points of $f_{|S}$ are non–degenerate; that is, if $dim(S) \geq 1$, the Hessian matrix of $f_{|S}$ is non–singular;

(iii): for each such critical point $x \in S$, and for each *generalised tangent space* $Q := \lim_{x_i \to x} T_{x_i} R$ at the point x, for $R > S$ a stratum of X and $\{x_i\} \subset R$ converging to x, $d_x \widetilde{f}(Q) \neq 0$ except for the single case of $Q = T_x S$.

An important example of a Whitney stratification is the *Morse stratification* of a Riemannian manifold X. We associate to a function $f \in \mathcal{F}(X)$ the vector field $grad\, f$ dual to df. The *gradient flow* for f is given by the paths of steepest decent, that is the trajectories (integral curves) of $-grad\, f$. If f is a Morse function, every trajectory converges to some $x \in crit(f)$, and the set of all points on trajectories converging to a given $x \in crit(f)$ form a cell $X^+(x)$, called the *stable manifold* of x. Upon replacing f by $-f$ we similarly get the cell $X^-(x)$ called the *unstable manifold* of x; one observes that $dim\, X^-(x) = codim\, X^+(x) = \lambda_x f$. More generally, if $Y \subseteq X$ is a non-degenerate manifold of f we similarly obtain stable manifolds $X^+(Y)$ which are cell-bundles over Y. We thus have a *stratification*

$$X = \cup_{Y \in crit(f)} X^+(Y)$$

called the *Morse stratification* of X. If f is a G-invariant function for G a compact Lie group, we can always choose a G-invariant metric on TX. The gradient flow if G-invariant so that the above stratification is G-invariant.

The Morse stratification of a space X by a function f has a natural partial ordering $\prec$ on the critical manifolds of f given by

$$Y_1 \prec Y_2 \iff \partial X^+(Y_1) \cap X^+(Y_2) \neq \emptyset.$$

One observes that $Y_1 \prec Y_2$ implies there exists a trajectory of $grad\, f$ starting on Y_1 and passing within $\varepsilon > 0$ of Y_2. In particular, upon taking $\varepsilon < f(Y_2) - f(Y_1)$ we have that $Y_1 \prec Y_2$ implies that $f(Y_1) < f(Y_2)$. Hence the transitive relation $<$ generated by $\prec$ is a partial ordering and has the property that

$$\overline{X^+(Y)} \subseteq \cup_{Y' \geq Y} X^+(Y').$$

The presence of such a partial ordering of the Morse strata aids us to develop a criterion for which a general stratification is a Morse stratification. This criterion is developed as follows.

Often one is given an explicit stratification of X, say $X = \cup_{\lambda \in \Lambda} X_\lambda$ where each X_λ is a locally closed submanifold of X and the indexing set Λ is strictly[14] partially ordered, for all $\lambda \in \Lambda$ we have

$$\overline{X_\lambda} \subseteq \cup_{\mu \geq \lambda} X_\mu.$$

One can use this stratification to obtain Morse-type information on $H_\bullet(X)$. We begin with open strata, given by minimal λ, and inductively add other strata. At each stage we can write down the exact cohomology sequence for a pair $(U, U \backslash V)$ where V is a closed submanifold of U. We now explain this procedure.

Define a subset I of indices to be

open if $\lambda \in I$ and $\mu \leq \lambda$ imply $\mu \in I$;

closed if $\lambda \in I$ and $\mu \geq \lambda$ imply $\mu \in I$,

[14] A partial ordering $\leq$ is *strict* if $\lambda \leq \mu$ and $\mu \leq \lambda$ imply $\lambda = \mu$.

for which we observe that I is closed if and only if I^c is open.

Moreover the subspace $X_I := \cup_{\lambda \in I} X_\lambda$ of X is open (or closed) if I is open (respectively closed). If I is open and $\lambda \in I^c$ is minimal then $J := I \cup \{\lambda\}$ is open and our inductive step is from X_I to X_J. From the decomposition $\overline{X_\lambda} \subseteq \cup_{\mu \geq \lambda} X_\mu$ we have that $X_\lambda := X_J \backslash X_i$ is a closed submanifold of X_J. Assuming that all the normal bundles to our stratification are orientable, we have by the Thom isomorphism theorem $H^{q-k}(X_\lambda) \cong H^q(X_J, X_I)$, and so we obtain the exact sequence

$$\cdots \longrightarrow H^{q-k}(X_\lambda) \longrightarrow H^q(X_J) \longrightarrow H^q(X_J) \longrightarrow \cdots$$

where $k = k_\lambda = codim\, X_\lambda$. If for a given field $\mathbb{F}$ the exact sequence breaks up into short–exact sequences for all q and λ it follows that

$$P_t(X) = \sum t^{k_\lambda} P_t(X_\lambda);$$

in such a case we say that the stratification if *perfect* over $\mathbb{F}$. If this holds for all prime order fields $\mathbb{F} = \mathbb{Z}_p$ we shall simply refer to the stratification as *perfect*; thus a perfect Morse function defines a perfect stratification.

If the stratification is G–invariant and the corresponding equivariant cohomology sequences break up, we shall call the stratification *G–equivariantly perfect.*

When a manifold is infinite dimensional the strata still have finite codimension. When the following two finiteness properties hold for the stratification we may proceed to compute the cohomology of X as in the finite dimensional case; although the induction never terminates, only finitely many steps will be needed to compute $H^q(X)$.

(F1): For every finite subset I there are a finite number of minimal elements of the complement I^c (so that our inductive procedure still applies);

(F2): for each $q \in \mathbb{Z}$ there are only finitely many indices $\lambda \in I$ with $codim\, X_\lambda < q$.

Given a stratification of X and a function $f \in \mathcal{F}(X)$, the following are axioms for checking whether the stratification is a Morse stratification (by stable manifolds) arising from f (for some metric on X).

4.20. Proposition.
Given $f \in \mathcal{F}(X)$ having only non–degenerate critical manifolds Y_λ, suppose $X = \cup_\lambda X_\lambda$ is a stratification by disjoint locally closed submanifolds X_λ which, subject to a partial ordering $\leq$ on the indexing set $\{\lambda\}$, satisfies

(i): $\lambda \leq \mu$ implies $f(X_\lambda) \leq f(X_\mu)$;

(ii): $\overline{X_\lambda} \subseteq \cup_{\mu \geq \lambda} X_\mu$;

(iii): for any $x \in X$, $grad_x f$ is tangential to the X_λ containing x;

(iv): $Y_\lambda \subseteq X_\lambda$;

(v): $\lambda_{Y_\lambda} f = codim\, X_\lambda$,

then X_λ is the stable manifold $S_\lambda := X^+(Y_\lambda)$ of Y_λ so that we have the Morse stratification.
Proof:
The trajectories $x(t)$ of $-grad\, f$ through any point $x \in X_\lambda$ converges to Y_λ as $t \to \infty$ due to the following reasons.

Condition (iii) implies $x(t)$ remains in X_λ for all finite t, while condition (ii) implies $x(\infty) \in Y_\mu$ for some $\mu \geq \lambda$. If x is sufficiently close to Y_λ, $x(t)$ either converges to or lies "below" Y_λ as $t \to \infty$. As $x(\infty) \in Y_\mu$ for some $\mu \geq \lambda$, $x(\infty)$ cannot lie below Y_λ and so $x(\infty) \in Y_\lambda$; thus locally near Y_λ, $X_\lambda \subset X^+(Y_\lambda)$. By property (v) $dim\, X_\lambda = dim\, X^+(Y_\lambda)$, and so near Y_λ the set X_λ is open in $X^+(Y_\lambda)$. Property (iv) implies X_λ and $X^+(Y_\lambda)$ coincide near Y_λ.

For $x \in X_\lambda$ with $x(\infty) \in Y_\mu$, as t becomes large $x(t)$ gets close to Y_μ in $X^+(Y_\mu)$, and so $x(t)$ lies in X_μ for large t. On the other hand, $x(t) \in X_\lambda$ for all finite t.

As the X_λ are disjoint, $\mu = \lambda$, and this implies what we require to prove.

□

We shall show that a stratification satisfying properties (i)–(v) above is induced by the Yang–Mills functional and hence obtain an explicit means of calculating Morse–type information about YM. If the Morse strata exist, that is, if one can prove good properties about the trajectories $x(t)$ as $t \to \infty$, then Proposition 4.20 will identify the Morse strata with our strata.

The problem in the Yang–Mills setting is that the manifold X in Proposition 4.20 (which is actually the space $\mathcal{A}$ of connections) is infinite–dimensional and the critical sets Y of the Yang–Mills functional YM have singularities. As pointed out in the closing of §1 in [2], due to these problems the connection with Morse theory and our work on Yang–Mills theory is left at a conjectural level, and our stratification will be used directly to compute cohomology.

5 Morse theory for the Yang–Mills functional.

Our primary aim in this thesis is to use Morse theory on the Yang–Mills functional

$$YM : \mathcal{A}(E) \to \mathbb{R},$$

when E is a $U(n)$–bundle over a compact Riemann surface X, in order to obtain information about solutions to the Yang–Mills equations.

There is a natural group acting on the affine space of connections $\mathcal{A}(E)$ called the *gauge group*, denoted by $\mathcal{G} := \Gamma(X, Aut(E))$, whose action is defined by

$$s \cdot \nabla(\sigma) := s^{-1}\nabla(s\sigma), \quad s \in \mathcal{G}, \quad \nabla \in \mathcal{A}(E), \quad \sigma \in \Gamma(E);$$

when $\nabla = d + A$ with respect to a trivialising cover of E over X, the above formula expands as

$$s \cdot \nabla := d + s^{-1}ds + s^{-1}As.$$

The action of the gauge group $\mathcal{G}$ on $\mathcal{A}(E)$ is not necessarily free and so we must carry out ($\mathcal{G}$–equivariant) Morse theory on the induced functional $YM_{\mathcal{G}}$ on the homotopy quotient

$$\mathcal{A}_{\mathcal{G}} \sim B\mathcal{G}$$

since $\mathcal{A}(E)$ is contractible.

To show that YM is an equivariantly perfect Morse function is very difficult and is yet to be established, although some directions have been followed by *K. Uhlenbeck* by analysing the properties of the Yang–Mills flow. Instead we shall concentrate on showing that the Yang–Mills functional induces a Morse stratification on $\mathcal{A}$ and that this stratification is *perfect* in the sense described in §2. For this purpose we now proceed to calculate Poincaré series $P_t(\mathcal{A}_{\mathcal{G}})$ via a well known description of the classifying space [15] $B\mathcal{G}(E)$, namely that[16]

$$B\mathcal{G}(E) \sim Map_E(X, BU(n)), \tag{5.0.1}$$

where the space on the right–hand side is the the space of maps $X \to BU(n)$ pulling back the universal bundle $EU(n) \to BU(n)$ to a bundle over X isomorphic to E; this is explained more precisely as follows. Consider the principal fibration [17] (a fibration whose total space is endowed with a group action)

$$\mathcal{G} \longrightarrow Map_{U(n)}(E, EU(n)) \xrightarrow{\pi} Map_E(X, BU(n))$$

where $Map_{U(n)}(E, EU(n))$ denotes the space of $U(n)$–equivariant maps. This fibration arises from the principal $U(n)$–bundle

$$U(n) \longrightarrow EU(n) \xrightarrow{\Phi} BU(n),$$

[15] The notion of classifying spaces is described in Appendix B.

[16] The formula $B\mathcal{G}(E) \sim Map_E(X, BG)$ holds for vector bundles E with general compact Lie structure group G.

[17] Some notes on fibrations appear in Appendix B for the sake of completeness for readers unfamiliar with this topic; some following locutions originate from this introduction.

itself a principal fibration, since $\mathcal{G}$ acts on the space $Map_{U(n)}(E, EU(n))$ by composition whose space of orbits is naturally homeomorphic to the function space $Map_E(X, BU(n))$ under the assignment $[\alpha] \mapsto \Phi \circ \alpha \circ p^{-1}$ for $p : E \to X$. The classifying space $BU(n)$ may be realised as the topological union $G_n(\mathbb{C}^\infty) := \bigcup_q G_n(\mathbb{C}^q)$ where $G_n(\mathbb{C}^q)$ is the *Grassman variety* whose points are n–dimensional subspaces of $\mathbb{C}^q$; this is a paracompact manifold modelled on an infinite–dimensional complex Hilbert space and is therefore locally contractible, thus so too is the space $Map_E(X, BU(n))$. By *Corollary 7.27, Chapter I* [34] (described in Theorem B.9 in Appendix B), the above fibration is locally fiber homotopically trivial since its base–space is locally contractible. As EG is contractible for any Lie group G (see §5 [25]) then so too is the space $Map_{U(n)}(E, EU(n))$. These facts imply that our fibration is a universal $\mathcal{G}$–bundle, and so by the *classification theorem of universal principle G–bundles* [18] the classifying space of $\mathcal{G}$ is homotopy equivalent to $Map_E(X, BU(n))$.

□

We shall now proceed to analyse the spaces $Map_E(X, BU(n))$ in order to derive the $\mathcal{G}$–equivariant Poincaré series $P_t^{\mathcal{G}}(\mathcal{A}(E)) := P_t(B\mathcal{G})$.

When E is a $U(1)$–bundle the classifying space of the structure group is $BU(1) \cong \mathbb{CP}^\infty$ where

$$\mathbb{CP}^\infty := \coprod_{n\geq 0} \{\mathbb{CP}^n \times [0,1]\}/\{(x,1) \sim (i(x),0)\}$$

where i are the inclusions

$$\{pt\} \hookrightarrow \cdots \overset{i}{\hookrightarrow} \mathbb{CP}^n \overset{i}{\hookrightarrow} \mathbb{CP}^{n+1} \hookrightarrow \cdots$$

Note that $\mathbb{CP}^\infty$ could also be thought of as the space $P(H) := S(H)/\mathbb{S}^1$ where $S(H)$ is the (contractible) unit sphere of an infinite dimensional Hilbert space H over $\mathbb{C}$. The corresponding universal $U(1)$–bundle

$$U(1) \longrightarrow S(H) \overset{\pi}{\longrightarrow} P(H)$$

gives rise to a long–exact homotopy[19] sequence from which results

$$\pi_k(P(H)) \cong \begin{cases} \mathbb{Z}, & k = 2 \\ 0, & k \neq 2. \end{cases}$$

$P(H)$ is thus an *Eilenberg-MacLane space*; for abelian groups τ and integers $n \geq 1$, these are CW–complexes $K(\tau; n)$ with homotopy

$$\pi_k(K(\tau; n)) \cong \begin{cases} \tau, & k = n \\ 0, & k \neq n. \end{cases}$$

Thus $P(H) \sim K(\mathbb{Z}; 2)$.

[18] This classification theorem is given as Theorem B.2 in Appendix B.

[19] The higher homotopy groups of a space X are defined by $\pi_n(X) := [\mathbb{S}^n, X]$ the set of homotopy classes of based maps $\mathbb{S}^n \to X$. These are groups if $n \geq 1$ and abelian if $n \geq 2$.

The following result aids in our calculation of $B\mathcal{G}(E)$ in this case, which is cited directly from [2] (Theorem 2.6, a result of *René Thom*) since it could not be located in available texts on algebraic topology.

5.1. Theorem.
Given X a finite complex, τ an abelian group and $n \geq 1$ an integer we have

$$Map(X, K(\tau; n)) \sim \prod_q K(H^q(X, \tau); n - q).$$

Thus we conclude, for X a Riemann surface of genus g,

$$\begin{aligned} Map(X, BU(1)) &\sim K(\mathbb{Z}; 0) \times K(\underbrace{\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}}_{2g-\text{fold}}; 1) \times K(\mathbb{Z}; 2) \\ &\sim \{pt\} \times \underbrace{\mathbb{S}^1 \times \cdots \times \mathbb{S}^1}_{2g-\text{fold}} \times P(H) \end{aligned}$$

where g is the genus of X; thus $B\mathcal{G}$ has no torsion.

By the *Künneth formula* for cohomology [20] we have that

$$H^i(Map(X, BU(1))) = \bigoplus_{p_0+p_1+\cdots+p_{2g+1}=i} H^{p_0}(\{pt\}) \bigotimes H^{p_1}(\mathbb{S}^1) \bigotimes \cdots \bigotimes H^{p_{2g}}(\mathbb{S}^1) \bigotimes H^{p_{2g+1}}(P(H)).$$

As we know the cohomologies of the spaces $\{pt\}, \mathbb{S}^1$ and $P(H)$, we calculate the following list of coefficients for the $\mathcal{G}$–equivariant Poincaré series $P_t(B\mathcal{G}) := \sum_{i \geq 0} t^i dim\, H^i(B\mathcal{G}, \mathbb{Q})$.

polynomial term order i	coefficient
0	1
1	$2g$
2	$\binom{2g}{2} + 1$
3	$\binom{2g}{3} + \binom{2g}{1}$
⋮	
$2n$	$\sum_{j=0}^{n} \binom{2g}{2j}$
$2n+1$	$\sum_{j=0}^{2n+1} \binom{2g}{j}$
⋮	

Thus[21]

$$P_t(B\mathcal{G}) = \frac{(1+t)^{2g}}{(1-t^2)}. \tag{5.1.1}$$

[20] The Künneth formula states that $H^p(X) \bigotimes H^q(Y) \cong H^{p+q}(X \times Y)$, see §29 [16].

[21] This result can also be understood from the fact that if A and B are graded algebras then $P_t(A \bigotimes B) = P_t(A)P_t(B)$.

Consider now the case $n > 1$ for our structure groups $U(n)$. The classifying space $BU(n)$ is realised as the infinite *Grassmann variety* $G_n(\mathbb{C}^\infty) := \bigcup_{q\geq 0} G_n(\mathbb{C}^q)$ where $G_n(\mathbb{C}^q)$ is the compact manifold of n–planes in $\mathbb{C}^q$ whose points are the n–dimensional subspaces of $\mathbb{C}^q$; the Grassmann variety is given the topology of a union since $G_n(\mathbb{C}^q) \subset G_n(\mathbb{C}^{q+1})$.

Over the rationals $\mathbb{Q}$ the space $BU(n)$ is a product of Eilenberg–MacLane spaces

$$BU(n) \sim_{\mathbb{Q}} K(\mathbb{Z};2) \times K(\mathbb{Z};4) \times \cdots \times K(\mathbb{Z};2n),$$

the reason for this follows from a type of *Postnikov approximation* (Proposition 18.19 [5]) which states that every connected CW–complex can be approximated by a twisted product of Eilenberg–MacLane spaces. The explanation as given in §2 [2] is that each Chern class $c_i \in H^{2i}(BU(n),\mathbb{Z})$ induces a map $BU(n) \to K(\mathbb{Z};2i)$, and as $H^\bullet(BU(n)) \cong \mathbb{Z}[c_1,\ldots,c_n]$ (§7 [25]) then the product of these maps induces a $\mathbb{Q}$-equivalence of these spaces.

As we know that $K(\mathbb{Z};1) = \mathbb{S}^1$ and $K(\mathbb{Z};2) = \mathbb{CP}^\infty$ we find by the *Künneth theorem* [22] that

$$H^i(K(\mathbb{Z};2k)) = \begin{cases} \mathbb{Z}, & i = 2kj \text{ for some } j \in \mathbb{N}\cup\{0\} \\ 0, & \text{otherwise} \end{cases}$$

$$\text{and} \quad H^i(K(\mathbb{Z};2k-1)) = \begin{cases} \mathbb{Z}, & i = 0,\ 2k-1 \\ 0, & \text{otherwise,} \end{cases}$$

and so we have by *Theorem 5.1* for $k \geq 2$ that

$$P_t(Map(X,K(\mathbb{Z};2k))) = \frac{(1+t^{2k-1})^{2g}}{(1-t^{2k-2})(1-t^{2k})}$$

and so as the space $Map(X,BU(n))$ splits up as a product of map spaces $Map(X,K(\mathbb{Z};2k))$ for $k = 1,\ldots,n$ we have with with the result (5.1.1) that

$$P_t(B\mathcal{G}) = \frac{\{(1+t)(1+t^3)\cdots(1+t^{2n-1})\}^{2g}}{\{(1+t^2)(1+t^4)\cdots(1-t^{2n-2})\}^2(1-t^{2n})}.$$

It turns out furthermore that the space $B\mathcal{G}$ is torsion free in this case. This is explained as follows. A compact Riemann surface of genus g is homotopically equivalent to the *bouquet* [23] $\bigvee_{2g}\mathbb{S}^1 \bigvee e^2$, where e^2 is a 2–cell, and so we have a cofibration[24]

$$\bigvee\nolimits_{2g}\mathbb{S}^1 \longrightarrow X \longrightarrow \mathbb{S}^2 \tag{5.1.2}$$

which gives rise to the following fibrations on base–point preserving maps

$$\begin{array}{ccc} Map^*(\mathbb{S}^2,BU(n)) & \longrightarrow & Map^*(X,BU(n)) \\ & & \downarrow \\ & & Map^*(\bigvee_{2g}\mathbb{S}^1,BU(n)) \end{array}$$

[22] The Künneth theorem (chapter V, Theorem 7.8 [34]) states that for abelian groups τ,τ' and n,q positive integers, $H^q(K(\tau\bigoplus\tau';n)) \cong \bigoplus_{r+s=q} H^r(K(\tau;n)) \bigotimes H^s(K(\tau';n)) \bigoplus \bigoplus_{r+s=q-1} Tor\{H^r(K(\tau;n)),H^s(K(\tau';n))\}$.

[23] In the disjoint union $X\coprod Y$ of based manifolds (X,x) and (Y,y) identify the points x and y to obtain the quotient space $X\bigvee Y$, called the it wedge or *bouquet* of X and Y.

[24] See Appendix B for the definition of a cofibration.

and

$$\begin{array}{ccc} Map^*(X, BU(n)) & \longrightarrow & Map(X, BU(n)) \\ & & \downarrow \\ & & BU(n). \end{array} \tag{5.1.3}$$

Now, the classifying space $BU(n)$ is known to be torsion free since (§7 [25]) $H^\bullet(BU(n), \mathbb{Z}) \cong \mathbb{Z}[c_1, \dots, c_n]$. The *loop-space* $\Omega BU(n) \approx U(n)$ and its second loop–space $\Omega U(n) \approx Map^*(\mathbb{S}^1, U(n))$ are also torsion free [25]. This follows from Theorem 4.1 Chapter VII [34] which states that the integral cohomology groups $H^\bullet(U(n), \mathbb{Z})$ are free and of finite rank. Also, as $U(n) \approx SU(n) \times U(1)$, hence $\Omega U(n) \approx \Omega SU(n) \times \Omega U(1)$, given that $SU(n)$ is a semisimple, compact, connected and simply connected Lie group the loop space $\Omega SU(n)$ has no torsion by the result of [4] [26] as does $\Omega U(1)$ since it is topologically equivalent to the free group $\mathbb{Z}$, thus the loop space $\Omega U(n)$ is torsion free.

Given the following identifications

$$\begin{aligned} \Omega U(n) := Map^*(\mathbb{S}^1, U(n)) &= Map^*(\mathbb{S}^1, \Omega BU(n)) \\ &\approx Map^*(\Sigma\mathbb{S}^1, BU(n)) = Map^*(\mathbb{S}^2, BU(n)) \end{aligned}$$

and

$$\begin{aligned} Map^*(\bigvee_{2g} \mathbb{S}^1, BU(n)) &\approx \prod_{2g} Map^*(\mathbb{S}^1, BU(n)) \\ &= \prod_{2g} \Omega BU(n) = \prod_{2g} U(n) \end{aligned}$$

the fibration (5.1.2) becomes

$$\begin{array}{ccc} \Omega U(n) & \longrightarrow & Map^*(X, BU(n)) \\ & & \downarrow \\ & & \prod_{2g} U(n) \end{array}$$

whose fiber and base–space are torsion free. Applying Theorem 5.1 to pointed maps we must have that the Poincaré series of the middle term of the above fibration must be the product of the Poincaré series of the factors. If there were any non–trivial homological twisting, the Poincaré series of the middle term would be smaller than the product of the Poincaré series of the factors; thus $Map^*(X, BU(n))$ must be free of torsion. A similar argument applied to fibration (5.1) then implies that the space $Map(X, BU(n))$ is torsion free.

We shall obtain a Morse stratification for $\mathcal{A}$ by applying *Proposition 4.20* to a particular stratification $\mathcal{A} = \bigcup_\lambda \mathcal{A}_\lambda$ we shall soon meet. In particular, we will require to show that this stratification

[25] The *loop-space* of a space X is the space $\Omega X := Map^*(\mathbb{S}^1, X)$, which satisfies the *adjunction formula* $Map^*(\Sigma X, Y) \cong Map^*(X, \Omega Y)$ for the *suspension* $\Sigma X := X \wedge \mathbb{S}^1$; here the *smash product* of two spaces X and Y is the space $X \bigwedge Y := (X \times Y)/(X \vee Y)$.

[26] The main result (Theorem I) of this paper states that if G is a semisimple compact connected and simply connected Lie group, then the loop space ΩG has no torsion.

satisfies condition (v) of *Proposition 4.20*; namely, that

$$\lambda_{\mathcal{A}'} YM = codim_{\mathbb{C}} \mathcal{A}_\lambda$$

for any non–degenerate critical manifold $\mathcal{A}' \subseteq \mathcal{A}(E)$ of YM contained in a particular stratum $\mathcal{A}_\lambda$. To this extent it is of importance to be able to compute the index $\lambda_{\mathcal{A}'} YM$; in particular, the index $\lambda_\nabla YM$ for a Yang–Mills connection ∇ localised in a certain stratum. We now proceed to calculate this quantity.

Given a connection ∇ that minimizes YM the self–adjoint endomorphism on $\mathcal{E}^0(ad(E))$

$$\Lambda := i\, ad_{*F_\nabla} : \alpha \mapsto i\, [*F_\nabla, \alpha]$$

has locally constant eigenvalues. This follows from the following argument: given v_x an eigenvector of Λ in the fiber $ad(E)_x$ such that $\Lambda(x)v_x = \lambda v_x$, then by the equation for parallel transport for a smooth loop γ on X we can find a local smooth section σ of $ad(E)$ such that $\nabla_{\dot\gamma}\sigma = 0$. So

$$\nabla(\Lambda\sigma - \lambda\sigma) = (\nabla\Lambda)\sigma + \Lambda\nabla\sigma - \lambda\nabla\sigma.$$

The right–hand side of this equation vanishes due to the fact that $d_\nabla * F_\nabla = 0$. Thus we have a first order initial value problem

$$\nabla w = 0, \qquad w(x) = 0$$

where $w := \Lambda\sigma - \lambda\sigma$, which by elementary differential equation theory implies $w \equiv 0$.

Thus Λ decomposes the complexification

$$ad(E)^{\mathbb{C}} \cong ad^-(E) \bigoplus ad^0(E) \bigoplus ad^+(E)$$

corresponding to its negative, zero and positive eigenvalues respectively, with dualities

$$ad^0(E)^* \cong ad^0(E) \quad \text{and} \quad ad^-(E)^* \cong ad^+(E)$$

induced by the Riemannian metric on $ad(E)$.

5.2. Proposition.
Given ∇ a Yang–Mills connection, we have

$$index(\nabla) = 2dim_{\mathbb{C}} H^1(X, \mathcal{O}(ad^-(E))) \tag{5.2.1}$$
$$nullity(\nabla) = 2dim_{\mathbb{C}} H^1(X, \mathcal{O}(ad^0(E))). \tag{5.2.2}$$

These results follow from the formula arising from a spectral estimate for the eigenvalues of the elliptic problem arising in the second–variation of YM. This is described more precisely as follows due to the identification $T_\nabla\mathcal{A}$ with $\mathcal{E}^1(X, ad(E))$.

Although $\mathcal{A}$ is an infinite–dimensional manifold we shall proceed to show that the notion of the index and nullity of a critical point of a map on a Banach manifold is well defined. One potential direction for defining these formulae directly is the suggestion of chapter 11 [21]. Given $f : X \to \mathbb{R}$ a smooth function on a manifold X modelled over a Banach space, for $x \in X$ a critical point of f let β_x denote the directed set of finite–dimensional subspaces of T_xX, and let $\lambda_x^E f$ denote the maximal dimension of a subspace of $E \in \beta_x$ on which the Hessian H_xf is negative–definite. If ν_ε denotes the exponential pre–image in T_xX of the ball $B_\varepsilon(x)$, then we define the index of x to be

$$\lambda_x f := \lim_{\varepsilon \to 0} \inf_{E \in \beta_x} \lambda_x^{E \cap \nu_\varepsilon} f$$

with a similar definition for the nullity of x.

5.3. Proposition.
The quadratic form $Q : \mathcal{E}^1(X, ad(E)) \times \mathcal{E}^1(X, ad(E)) \to \mathbb{R}$ induced by the Hessian of the Yang–Mills functional YM at a Yang–Mills connection ∇ is given by

$$Q(\eta, \eta) = (d_\nabla^* d_\nabla \eta + *[*F_\nabla, \eta] \ , \ \eta).$$

Proof:
Given the curve $\nabla_t := \nabla + t\eta$ for $\eta \in \mathcal{E}^1(X, ad(E))$, we have that the norm of its curvature expands to second order as

$$\|F_{\nabla_t}\|_2^2 = \|F_\nabla\|_2^2 + 2t(d_\nabla \eta, F_\nabla) + t^2\{\|d_\nabla \eta\|_2^2 + (F_\nabla, [\eta, \eta])\} + O\{t^3\}$$

where $(\cdot\,,\,\cdot)$ is the global inner–product on $\mathcal{E}^2(ad(E))$ constructed, as in §2, by the Hodge-Star operator on X.

The Hessian of YM at the extremum ∇ is obtained from this expansion to yield

$$\begin{aligned} Q(\eta, \eta) &= \frac{1}{2}\frac{d^2}{dt^2}\|F_{\nabla_t}\|_{2|t=0}^2 \\ &= \|d_\nabla \eta\|_2^2 + (F_\nabla, [\eta, \eta]). \end{aligned}$$

Observe that $\|d_\nabla \eta\|_2^2 = (d_\nabla^* d_\nabla \eta, \eta)$ since d_∇^* is the formal adjoint of d_∇, and that

$$\begin{aligned} (F_\nabla, [\eta, \eta]) &:= \int_X [\eta, \eta] \wedge *F_\nabla \\ &= \int_X \eta \wedge [\eta, *F_\nabla] \\ &= (-1)^{dim(X)+1} \int_X \eta \wedge * *^{-1} [*F_\nabla, \eta]. \end{aligned}$$

Using the formula for $*^{-1}$ the last equation reduces to $(\eta, *[*F_\nabla, \eta])$.

□

The endomorphism

$$\widehat{F}_\nabla : \eta \mapsto *[*F_\nabla, \eta]$$

is a degree zero operator on $\mathcal{E}^1(X, ad(E))$, for which we have observed

$$(\widehat{F}_\nabla \eta, \xi) = (F_\nabla, [\eta, \xi]).$$

5.4. Proposition.
The index and nullity of a Yang–Mills connection ∇ are finite and equal to, respectively, the index and nullity of the quadratic form

$$\widehat{Q}(\eta) := (\triangle_\nabla \eta + \widehat{F}_\nabla \eta \, , \, \eta)$$

on the kernel of d^*_∇ in $\mathcal{E}^1(X, ad(E))$ where $\triangle_\nabla := d^*_\nabla d_\nabla + d_\nabla d^*_\nabla$ is the Laplacian associated to ∇.
Proof:
The finiteness of the nullity follows from the following argument. Given the *Jacobi operator*

$$L_\nabla := d^*_\nabla d_\nabla + \widehat{F}_\nabla$$

of a Yang–Mills connection ∇, and $J_\nabla \subset \mathcal{E}^1(X, ad(E))$ the collection of *Jacobi fields* of L_∇, that is, elements $\eta \in \mathcal{E}^1(X, ad(E))$ such that $L_\nabla \eta = 0$, we call the quotient $N_\nabla := J_\nabla / im(d_\nabla)$ the *null space* of Q and its dimension the *nullity* of ∇. In the usual norm on $\mathcal{E}^1(X, ad(E))$ the ortho–complement of $im(d_\nabla)$ in the exact sequence

$$\mathcal{E}^0(X, ad(E)) \xrightarrow{d_\nabla} J_\nabla \longrightarrow N_\nabla \longrightarrow 0$$

is $ker(d^*_\nabla)$. Thus we may identify N_∇ with the space

$$\{ \ \eta \in \mathcal{E}^1(X, ad(E)) \mid L_\nabla \eta = 0, \quad d^*_\nabla \eta = 0 \ \}$$

whose conditions are also equivalent to

$$(\triangle_\nabla + \widehat{F}_\nabla)\eta = 0, \quad d^*_\nabla \eta = 0. \tag{5.4.1}$$

As $\widehat{F}_\nabla$ is a degree zero operator and the Laplacian $\triangle_\nabla$ is elliptic, the solution space to (5.4.1) is finite dimensional, and thus so too is the nullity.

As the Morse index of a Yang–Mills ∇ connection is defined as the dimension of a maximal negative subspace of Q, or equivalently, as the dimension of a maximal subspace in $ker(d^*_\nabla)$ on which the form

$$\widehat{Q}(\eta) := (\, \triangle_\nabla \eta + \widehat{F}_\nabla \eta \, , \, \eta \,)$$

is negative definite, we may extend the preceding argument to yield the finiteness of the Morse index.

□

We finally proceed to derive quantities (5.2.1) and (5.2.2). Recall that the operator $\Lambda := i \, ad * F_\nabla$ has locally constant eigenvalues λ, and so forms a decomposition

$$ad(E) \bigotimes \mathbb{C} = \bigoplus_\lambda ad_\lambda(E)$$

into orthogonal sub-bundles on which this operator reduces to the constant matrix $\lambda \, id$.

As $\widehat{F}_\nabla = -i * \Lambda$ we can reduce the analysis of $\widehat{Q}$ to the cases where Λ is zero or a positive or negative scalar multiple of the identity matrix. In the case $\Lambda = 0$, which corresponds to the bundle

$ad^0(E)$, we have $\widehat{Q}(\eta) := (\triangle_\nabla \eta, \eta)$ which is semi–definite and so its index is zero. The nullity of $\widehat{Q}$ is equal to the dimension of the subspace of ∇–harmonic forms in $ker(d_\nabla^*)$. As $\triangle_\nabla \eta = 0$ for $\eta \in ker(d_\nabla^*)$ implies $d_\nabla \eta = 0$ and $d_\nabla^* \eta = 0$, then the nullity of $\widehat{Q}$ equals dimension of the subspace of harmonic forms in $\mathcal{E}^1(X, ad^0(E))$; the space $\mathcal{H}^1(ad^0(E))$ of harmonic forms is isomorphic to the DeRham space $H^1_{d_\nabla}(ad^0(E))$ (p.152 [17]), which in turn is isomorphic to $H^1(X, \mathcal{O}(ad^0(E)))$ by the *Dolbeault theorem* [27].

For the case when $\Lambda = \lambda\, id$ for $\lambda > 0$ we require the following results.

5.5. Lemma.
Given an eigenvalue $\lambda > 0$ of the operator Λ, the Laplacian $\triangle_\nabla$ preserves the spaces $\mathcal{E}^{(1,0)}(ad_\lambda(E))$ and $\mathcal{E}^{(0,1)}(ad_\lambda(E))$. Moreover, the first positive eigenvalue of $\triangle_{\nabla|\mathcal{E}^{(1,0)}(ad_\lambda(E))}$ is greater than or equal to 2λ.

Proof:
The Laplacian splits up on the natural decomposition of the complexification by the Hodge–star operator $*$

$$\mathcal{E}^1_{\mathbb{C}}(ad_\lambda(E)) = \mathcal{E}^{(1,0)}(ad_\lambda(E)) \bigoplus \mathcal{E}^{(0,1)}(ad_\lambda(E))$$

into operators on the partial connections

$$\begin{aligned} \triangle_{\partial_\nabla} &:= \partial_\nabla \partial_\nabla^* + \partial_\nabla^* \partial_\nabla \\ \text{and } \triangle_{\overline{\partial}_\nabla} &:= \overline{\partial}_\nabla \overline{\partial}_\nabla^* + \overline{\partial}_\nabla^* \overline{\partial}_\nabla. \end{aligned}$$

where $\overline{\partial}_\nabla$ and ∂_∇ are the $(0,1)$ and $(1,0)$ components respectively of the covariant derivative d_∇ given locally and respectively by $\overline{\partial} + A_1$ and $\partial + A_2$ for $\overline{\partial}$ and ∂ the standard Cauchy–Riemann operators and A_1 and A_2 the respective $(0,1)$ and $(1,0)$ components of the local connection matrix for ∇.

By the formula $d_\nabla^* = - * d_\nabla *$ it follows that $\triangle_{\partial_\nabla}$ and $\triangle_{\overline{\partial}_\nabla}$ induce the same operator $\frac{1}{2}\triangle_\nabla$ on $\mathcal{E}^{(1,0)}$ and $\mathcal{E}^{(0,1)}$ which preserves these spaces, and $\triangle_\nabla = \triangle_{\partial_\nabla} + \triangle_{\overline{\partial}_\nabla}$ on the spaces $\mathcal{E}^{(0,0)}$ and $\mathcal{E}^{(1,1)}$. Also, on $\mathcal{E}^{(0,0)}$ observe that $i\, d_\nabla^2 \alpha = (\partial_\nabla \overline{\partial}_\nabla + \overline{\partial}_\nabla \partial_\nabla)\alpha = *[\Lambda, \alpha]$, and so $\triangle_{\partial_\nabla} - \triangle_{\overline{\partial}_\nabla} = \lambda$ on this space.

As ∂_∇ and $\overline{\partial}_\nabla$ are elliptic operators their associated Laplacians are compact self–adjoint operators. Thus we have a Hilbert space decomposition (p.95 [17]) of $\mathcal{E}^1(ad_\lambda(E))$ into a direct sum of eigenspaces of these operators. As we have shown that the partial Laplacians are linearly related on the spaces $\mathcal{E}^{(0,0)}, \mathcal{E}^{(1,1)}, \mathcal{E}^{(1,0)}$ and $\mathcal{E}^{(0,1)}$, the eigenspaces of these operators in $\mathcal{E}^1(ad_\lambda(E))$ are the same. Thus the positive eigenvalues (spectra) of the two partial Laplacians are in one–to–one correspondence, and so the positive spectra of $\triangle_{\partial_\nabla}$ on the spaces $\mathcal{E}^{(0,0)}$ and $\mathcal{E}^{(1,1)}$ coincide. By the preceding remark, as the two partial Laplacians differ by λ on $\mathcal{E}^{(0,0)}$, the spectrum is bounded below by λ since $\triangle_{\overline{\partial}_\nabla}$ is semi–definite. As $\triangle_\nabla = 2\triangle_{\partial_\nabla}$ on $\mathcal{E}^{(1,0)}$ we have the required result.

□

[27] The *Dolbeault theorem* states that $H^q(X, \Omega^p(E)) \cong H^{(p,q)}_{\overline{\partial}_E}(E)$ where $\Omega^p(E)$ is the sheaf of germs of holomorphic p–forms on E (p.151 [17]).

5.6. Corollary.
The quadratic form $\widehat{Q}$ has nullity zero and $index(\widehat{Q}_{|ker(d^*_\nabla)})$ equals the dimensions of the subspace of harmonic forms in $\mathcal{E}^{(1,0)}(ad_\lambda(E))$ for $\lambda > 0$.
Proof:
As $\widehat{F}_\nabla = i * \Lambda$ and as $* = -i$ on $\mathcal{E}^{(1,0)}$ and $* = i$ on $\mathcal{E}^{(0,1)}$ we have that

$$\widehat{F}_\nabla|\mathcal{E}^{(1,0)} = -\lambda \quad \text{and} \quad \widehat{F}_\nabla|\mathcal{E}^{(0,1)} = \lambda$$

and so the operator $\triangle_\nabla + \widehat{F}_\nabla$ is positive on $\mathcal{E}^{(0,1)}$, and has single negative eigenvalue $-\lambda$ on the space $\mathcal{E}^{(1,0)}$ with multiplicity the dimension of the harmonic forms in $\mathcal{E}^{(1,0)}$.

□

Proof of Proposition 5.2:
With the above derived expressions for the nullity and index of the quadratic form $\widehat{Q}$ on $ker(d^*_\nabla)$, we proceed to put the theory of harmonic forms into a sheaf–theoretic framework.

By *Hodge theory* [28] on bundles we have

$$\begin{aligned} H^i(X, \mathcal{O}(ad_\lambda(E))) &\cong ker(\triangle_{\overline{\partial}_\nabla}|\mathcal{E}^{(0,i)}(ad_\lambda(E))) \\ \text{and} \quad H^i(X, \mathcal{O}(ad_\lambda(E) \bigotimes T'^*X)) &\cong ker(\triangle_{\overline{\partial}_\nabla}|\mathcal{E}^{(1,i)}(ad_\lambda(E))) \end{aligned}$$

where T'^*X is the holomorphic cotangent space of X. Also, the *Kodaira-Serre duality* [29] gives

$$H^i(X, \mathcal{O}(ad_\lambda(E) \bigotimes T'^*X)) \cong H^{1-i}(X, \mathcal{O}(ad_\lambda(E)^*)).$$

Applying these results and our earlier mentioned duality $ad^-(E)^* \cong ad^+(E)$ with *Corollary 5.6* for $\lambda > 0$ gives

$$index(Q) = dim\, H^1(X, \mathcal{O}(ad_{-\lambda}(E))).$$

By applying the above arguments to the case $\lambda < 0$ gives a completely analogous result

$$index(Q) = dim\, H^1(X, \mathcal{O}(ad_\lambda(E))),$$

and we obtain the general formula for the index once we sum over the eigenvalues λ.

□

The index is *stable* in the sense that the above formula can be put in purely topological terms. Recall that the *Riemann-Roch (Hirzebruch) theorem* states that for a compact complex manifold X^n a holomorphic bundle E over X we have $\chi(E) = \{ch(E) \cdot td(TX)\}[X]$ where ch is the *Chern*

[28] Let X be a complex manifold and E a holomorphic bundle over X. Given the Laplacian $\triangle := \overline{\partial}_E\overline{\partial}^*_E + \overline{\partial}^*_E\overline{\partial}_E$: $\mathcal{E}^{(p,q)}(E) \to \mathcal{E}^{(p,q)}(E)$, the *harmonic space* $\mathcal{H}^{(p,q)}(E) := ker(\triangle)$ is finite dimensional and isomorphic to $H^{(p,q)}_{\overline{\partial}_E}(E)$. See p.152 [17].

[29] Given X^n a complex manifold and E a holomorphic bundle over X, the $*$–operator gives isomorphisms $H^q(X, \Omega^p(E)) \cong H^{n-q}(X, \Omega^{n-p}(E^*))^*$ where $\Omega^p(E)$ is the sheaf of germs of holomorphic p–forms on E. For $p = 0$ this gives $H^q(X, \mathcal{O}(E)) \cong H^{n-q}(X, \mathcal{O}(E^* \otimes K_X))^*$ where $K_X := \bigwedge^n T'^*X$ is the *canonical line bundle* and T'^*X the holomorphic cotangent space. See p.153 [17].

character, *td* is the *Todd class*, $\chi(E) := \sum_{i=0}^{n}(-1)^i dim\, H^i(X, \mathcal{O}(E))$, and $\{\omega\}[X]$ denotes the evaluation of the degree n component of the DeRham class $\omega \in H^{\bullet}_{dR}(X;\mathbb{C})$ on the fundamental cycle of X. When X is a Riemann surface of genus g, this formula reduces to $dim_{\mathbb{C}} H^0(X, \mathcal{O}(E)) - dim_{\mathbb{C}} H^1(X, \mathcal{O}(E)) = deg(E) + (g-1)rk(E)$, and as $c_1(E^*) = -c_1(E)$ we have

$$index(\nabla) = 2\big\{deg(ad^+(E)) + rk(ad^+(E))(g-1)\big\}.$$

The fact that $H^0(X, \mathcal{O}(ad_\lambda(E))) = 0$ follows since $\triangle_{\partial_\nabla} - \triangle_{\overline{\partial}_\nabla} = -\lambda$ on $\mathcal{E}^{(0,0)}(ad_\lambda(E))$ for $\lambda > 0$. In this way for $s \in H^0(X, \mathcal{O}(ad_\lambda(E)))$ we have $\overline{\partial}_\nabla s = 0$ and thus $\triangle_{\overline{\partial}_\nabla} s = \overline{\partial}^*_\nabla \overline{\partial}_\nabla s = 0$. Using the natural inner product $(\,\cdot\,,\,\cdot\,)$ on $\mathcal{E}^*(ad(E))$ induced by the Catan–Killing form on $\mathfrak{u}(k)$ and the Hodge–star operator on X, we have that as $\triangle_{\partial_\nabla} s - \triangle_{\overline{\partial}_\nabla} s = -\lambda s$ then $(\partial^*_\nabla \partial_\nabla s\,,\, s) - (\overline{\partial}^*_\nabla \overline{\partial}_\nabla\,,\, s) = -\lambda(s\,,\, s)$. As ∂^*_∇ is the formal adjoint of ∂_∇ this results in $\|\partial_\nabla s\|^2 - 0 = -\lambda \|s\|^2$ which implies that $s = 0$ since $\lambda > 0$.

On the other hand, the nullity is not stable in this sense.

As a point of interest, the induced Yang–Mills functional $YM_{\mathcal{G}}$ on the homotopy quotient $\mathcal{A}_{\mathcal{G}} \sim B\mathcal{G}$ is defined as *perfect* if

$$P_t^{\mathcal{G}}(\mathcal{A}) = M_t(YM_{\mathcal{G}})$$

where the Morse counting M_t series is given by

$$M_t(YM_{\mathcal{G}}) = \sum_{\mathcal{A}' \in crit(YM)} t^{\lambda_{\mathcal{A}'} YM} P_t^{\mathcal{G}}(\mathcal{A}')$$

due to our results in *Proposition 4.19*, and the $\mathcal{G}$–equivariant Poincaré series $P_t^{\mathcal{G}}(\mathcal{A}')$ is computable owing to results in a subsequent chapter when we construct a Morse stratification for $\mathcal{A}$. As we have been able to explicitly compute the index $\lambda_{\mathcal{A}'} YM$ of a non–degenerate critical manifold $\mathcal{A}' \subseteq \mathcal{A}(E)$ of YM we can compute the Morse counting series $M_t(YM_{\mathcal{G}})$. In the case of rank $n = 2$ $U(2)$–bundles E over X with degree $k = 1$, *Bott* [3] shows that

$$M_t(YM_{\mathcal{G}}) = \frac{P_t(N(2,1))}{1-t^2} + \frac{t^{2g+4}(1+t^2)^{4g}}{(1-t^2)^2(1-t^4)}$$

where $N(n,k)$ denotes the isomorphism classes of rank n holomorphic bundles of degree k over X.

If we hazard to assume that YM is perfect in the equivariant sense, we have upon equating $M_t(YM_{\mathcal{G}})$ and $P_t(B\mathcal{G})$ that

$$P_t(N(2,1)) = \frac{\{(1+t)(1+t^3)\}^{2g}}{(1-t^2)(1-t^4)} - \frac{t^{2g}(1+t)^{4g}}{(1-t^2)(1-t^4)}.$$

G. Harder and *M. Narasimhan* [19] derived this result purely in the context of algebraic geometry by defining these varieties over a finite field whose number of rational points were found with the aid of number theory, and then by applying the *Weil* conjectures (see §11 [2]).

6 Stable bundles.

In this section we shall introduce the notion of *stability* of a holomorphic vector bundle over a Riemann surface X which will aid in our construction of a Morse stratification of the affine space $\mathcal{A}(E)$ of connections on a $U(n)$–bundle E over X. To this extent we shall expound results on the stability of holomorphic vector bundles over our Riemann surface X relating to the form of the curvature of a metric connection on such bundles.

Let X be a compact Riemann surface (without boundary) with Hermitian metric g chosen so that $\int_X vol_X = 1$ where $vol = vol_X$ is locally given by $\sqrt{det[g_{ij}]}\, dx_1 \wedge dx_2$ for the local representation $g_{ij}dx_i \otimes dx_j$ of g, and let E be a smooth complex vector bundle over X of rank $rk(E) = k$ with compact Lie structure–group $G \subseteq U(k)$.

Define the *slope* of E to be the rational number

$$\mu(E) := deg(E)/rk(E)$$

where the *degree* of E is the integer $deg(E) := c_1(E)[X] = c_1(\bigwedge^k E)[X]$ where $\bigwedge^k E$ is the *determinant* (line) *bundle*. We shall take $c_1(E) = [-\frac{1}{2\pi i}\, tr(F_\nabla)] \in H^2_{dR}(X)$ for any smooth connection ∇ on E where the *curvature* of ∇, $F_\nabla := d_\nabla \circ \nabla$, is a smooth section of $ad(E) \bigotimes_{\mathbb{C}} T^*X \bigwedge T^*X$; this DeRham cohomology class being independent of the choice of ∇ since two connections on E differ by an element of $\Gamma(X, ad(E) \bigotimes_{\mathbb{C}} T^*X)$ where $ad(E) \subseteq End(E)$ is the *Lie algebra bundle* on X with fibers isomorphic to the Lie algebra $\mathfrak{g} \subseteq \mathfrak{u}(k)$ of G.

6.1. Definition.
A holomorphic vector bundle $\mathcal{E}$ over X is said to be *indecomposable* if it cannot be decomposed into a proper direct sum of holomorphic sub-bundles. $\mathcal{E}$ is said to be *stable* (respectively, *semi-stable*) if for all non–zero proper holomorphic sub-bundles $\mathcal{F} < \mathcal{E}$ we have

$$\mu(\mathcal{F}) < \mu(\mathcal{E}) \quad \text{(respectively, } \mu(\mathcal{F}) \leq \mu(\mathcal{E})\text{)}.$$

Due to the fact that $c_1(\mathcal{E}/\mathcal{F} \bigoplus \mathcal{F}) = c_1(\mathcal{E}/\mathcal{F}) + c_1(\mathcal{F})$ we have that these conditions are equivalent to

$$\mu(\mathcal{E}/\mathcal{F}) > \mu(\mathcal{E}) \quad \text{(respectively, } \mu(\mathcal{E}/\mathcal{F}) \geq \mu(\mathcal{E})\text{)}.$$

In 1965 *M. S. Narasimhan & C. S. Seshadri* proved (*Theorem 2*(A) §12[27]) that a holomorphic vector bundle $\mathcal{E}$ on a compact Riemann surface X of genus $g \geq 2$ is *stable* if and only if $\mathcal{E}$ arises from an irreducible projective unitary representation of the fundamental group $\pi_1(X)$. An equivalent result was proved by *S. K. Donaldson* [9] in 1983 by a method expounded in this section which is "self-contained" as opposed to the proof given in [27]. We shall later demonstrate the equivalence between *Theorem 2*(A) §12 [27] and Donaldson's version appearing as *Theorem 6.2* below.

6.2. Theorem.
An indecomposable holomorphic hermitian vector bundle $(\mathcal{E}, h)$ over (X, g) is stable if and only if there is a unitary connection ∇ on $(\mathcal{E}, h)$ having curvature satisfying $*F_\nabla = -2\pi i \mu(\mathcal{E})$. Such a connection is furthermore unique up to isomorphism.

Here a connection ∇ on $(\mathcal{E}, h)$ is called *unitary (with respect to* $h := (\cdot\,,\,\cdot)$*)* if $d(\sigma,\tau) = (\nabla\sigma,\tau)+(\sigma,\nabla\tau)$ for $\sigma,\tau \in \Gamma(X,\mathcal{E})$. This means that with respect to a unitary frame of $(\mathcal{E},h)$ the connection matrix of ∇ is skew–adjoint.

6.3. Definition.
A connection ∇ on (E,h) over (X,g) whose *central curvature* $*F_\nabla$ equals $-2\pi i\mu(E)$ is referred to as an *Hermitain–Einstein (H–E) connection* with *factor* $-2\pi i\mu(E)$.

As we are predominately working with the differential–geometric structures of connections on holomorphic bundles, we will utilise an equivalence between holomorphic structures $\mathcal{E}$ and unitary connections on smooth complex vector bundles E over X in order to simplify the proof of *Theorem 6.2.* This correspondence is made precise as follows.

When analysing connections one naturally discusses their coördinate representations; this being the motivation for the introduction of the *gauge group.* The *complex gauge group* $\mathcal{G}^{\mathbb{C}}$ of general linear automorphisms on E acts on $\mathcal{A}(E)$ by

$$g\cdot\nabla := \nabla - (\overline{\partial}_\nabla g)g^{-1} + ((\overline{\partial}_\nabla g)g^{-1})^*, \quad g\in\mathcal{G}^{\mathbb{C}}, \quad \nabla\in\mathcal{A}(E),$$

thus extending the action $u\cdot\nabla := \nabla-(d_\nabla u)u^{-1}$of the *unitary gauge group* $\mathcal{G} := \{u\in\mathcal{G}^{\mathbb{C}} \mid u^*u = 1\}$.

The existence of holomorphic structures $\mathcal{E}$ on a smooth complex vector bundle E over X is related to a class of *partial connections* on E called $\overline{\partial}$*–operators.* These are $\mathbb{C}$–linear operators

$$\overline{\partial}_{\mathcal{E}} : \Gamma(E) \longrightarrow \mathcal{E}^{(0,1)}(E)$$

satisfying the $\overline{\partial}$*–Leibnitz rule* $\overline{\partial}_{\mathcal{E}}(f\sigma) = \overline{\partial}f\otimes\sigma + f\overline{\partial}_{\mathcal{E}}(\sigma)$ for $f\in\mathcal{E}^0(X)$ and $\sigma\in\Gamma(E)$ and satisfying the *integrability condition* $\overline{\partial}_{\mathcal{E}}^{\,2} = 0$ (where we have used the same symbol for the extension of $\overline{\partial}_{\mathcal{E}}$ to the operator $\mathcal{E}^{(p,q)}(E)\to\mathcal{E}^{(p,q+1)}(E)$); this latter condition is vacuous on a Riemann surface X since $\bigwedge^{(2,0)}X = \bigwedge^{(0,2)}X = 0$.

Given a holomorphic structure $\mathcal{E}$ and a local holomorphic frame $\{e_i\}$ on E, we define a natural $\overline{\partial}$*–operator* by the local formula

$$\overline{\partial}_{\mathcal{E}}(\sum_i \omega_i\otimes e_i) := \sum_i \overline{\partial}\omega_i\otimes e_i.$$

This construction is independent of the choice of holomorphic frame since the transition functions relating these frames are holomorphic maps.

Conversely, given a partial connection $\overline{\partial}_\alpha$ satisfying the $\overline{\partial}$–Leibnitz rule and the (vacuous) integrability condition, we have that its local components $\frac{\partial}{\partial\bar{z}}+\alpha$ (with respect to an open set $U\subseteq X$) commute. Thus by the *Newlander–Nirenberg theorem* for almost complex manifolds [28] we have that there is a complex gauge transformation $g : U\to GL(k,\mathbb{C})$ such that $g(\frac{\partial}{\partial\bar{z}}+\alpha)g^{-1} = \frac{\partial}{\partial\bar{z}}$; that is, there is local trivialisation of E upon which $\alpha = 0$. In such a local trivialisation the solutions to $\overline{\partial}_\alpha s = 0$ are just the holomorphic vector functions, so the sheaf of germs of local solutions to this equation is a locally free sheaf of $\mathcal{O}_X$–modules which thus corresponds to a holomorphic vector bundle $\mathcal{E}_\alpha$ over X. This latter result can be alternatively understood to mean the holomorphic vector bundle $\mathcal{E}_\alpha$ constructed from the holomorphic transition functions which connect any two local trivialisations of E on which $\alpha = 0$.

With this understanding, a unitary connection ∇ on E induces a holomorphic structure $\mathcal{E}_\nabla$ on E via its $(0,1)$–component $\overline{\partial}_\nabla : \Gamma(E) \to \mathcal{E}^{(0,1)}(E)$. Conversely, given a holomorphic structure $\mathcal{E}$ on (E,h) and an associated partial–connection $\overline{\partial}_\mathcal{E}$, there exists a unique unitary connection ∇ (with respect to h) which is *compatible* with the holomorphic structure $\mathcal{E}$ on E; that is, with $\overline{\partial}_\nabla = \overline{\partial}_\mathcal{E}$. For, within a local unitary trivialisation of E in which the partial connection $\overline{\partial}_\mathcal{E}$ is represented by a matrix α of $(0,1)$–forms, the connection matrix A of 1–forms of such a ∇ must satisfy $A = -A^*$ and must have $(0,1)$–component α. These requirements uniquely determine the connection matrix as $A = \alpha - \alpha^*$, which gives rise to a well defined global object due to the compatibility conditions satisfied by α. Alternatively, given a local holomorphic frame $\{e_1, \cdots, e_k\}$ for $\mathcal{E}$ upon which our Hermitian metric h is locally represented by the matrix $h_{ij} := h(e_i, e_j)$, in this trivialisation for $\mathcal{E}$ our compatible connection is given by the matrix of $(1,0)$–forms $h^{-1}\partial h$, where we have abused notation here and written $h = (h_{ij})$ for the local representation of h. As before, ∇ is determined as a well defined global object by the conditions of compatibility of h.

6.4. Definition.
The unique connection on a holomorphic Hermitian bundle $(\mathcal{E}, h)$ over (X, g) which is compatible with both the metric– and holomorphic–structures on the bundle is called the *metric connection* on $(\mathcal{E}, h)$.

Clearly, two $\overline{\partial}$–operators give isomorphic holomorphic structures if and only if they are conjugate by an automorphism of the underlying smooth bundle; that is, $\mathcal{E} \cong \mathcal{F}$ as holomorphic structures on E if and only if there exists a $g \in \Gamma(X, Aut(E))$ such that $\overline{\partial}_\mathcal{E}\, g = g\, \overline{\partial}_\mathcal{F}$; or equivalently, if $\overline{\partial}_\mathcal{F} = \overline{\partial}_\mathcal{E} - (\overline{\partial}_\mathcal{E}\, g)g^{-1}$.

Connections thus define isomorphic holomorphic structures precisely when they lie in the same $\mathcal{G}^\mathbb{C}$–orbit. Given $\mathcal{E}$ a holomorphic vector bundle denote by $O(\mathcal{E})$ the orbit of connections $\{g \cdot \nabla \mid g \in \mathcal{G}^\mathbb{C}\}$ such that $\mathcal{E}_\nabla \cong \mathcal{E}$.

In the proof of our main theorem we shall also require to generalise the class of connections to incorporate the L_1^2 connections. Connections of this class are described as follows.

Let $\mathcal{A}(E)$ denote the affine space, modelled over the space $\Gamma(X, ad(E) \bigotimes_\mathbb{C} T^*X)$, of connections on E. Choosing a "base" connection $\nabla_o \in \mathcal{A}(E)$, we obtain a *Sobolev–norm* on sections $\sigma \in \Gamma(U, E)$ for $U \subseteq X$ open

$$\|\sigma\|_{p,k} := \sum_{\alpha=0}^{k} \|{\nabla_o}^\alpha \sigma\|_p$$

where ${\nabla_o}^\alpha$ means α–fold composition of the operator ∇_o with the convention that ${\nabla_o}^0 := 1_{\Gamma(E)}$. The completion of the space $\Gamma(U, E)$ in the $\|\cdot\|_{p,k}$–norm is defined to be the *Sobolev space* $L_k^p(U, E)$. Different choices of ∇_o yield equivalent norms so the definition of this space is independent of the choice of "base" connection. We define a connection $\nabla \in \mathcal{A}(E)$ to be of class L_k^p if, given $\nabla = d + A$ locally over U, $A \in L_k^p(U, ad(E) \bigotimes T^*X)$; the affine space of such connections is denoted $\mathcal{A}_k^p(E)$.

For necessity in later calculations the gauge group acting on $\mathcal{A}_k^p(E)$ is taken to be $\mathcal{G}_{k+1}^p := L_{k+1}^p(X, Aut(E))$.

The proof of *Theorem 6.2* incorporates the minimisation of a certain functional J on the space of L_1^2 connections satisfying $J(\nabla) = 0$ if and only if the unitary connection ∇ on the holomorphic Hermitian bundle $(\mathcal{E}, h)$ is of type required by the theorem. This functional is constructed so as to have the *lower semi-continuity property* which allows one to utilise the weak–compactness result of *K. Uhlenbeck* (*Theorem 1.5* [32]) for connections on bundles (which shall be expounded in a later appendix) in obtaining a limiting connection $\widetilde{\nabla}$ for a given minimising sequence for $J_{|O(\mathcal{E})}$. This connection will either lie in $O(\mathcal{E})$ or define a different holomorphic structure $\mathcal{E}_{\widetilde{\nabla}} \ncong \mathcal{E}$ of same rank and degree as $\mathcal{E}$. In either case we have a non–zero sheaf homomorphism $\alpha : \mathcal{E}_{\widetilde{\nabla}} \longrightarrow \mathcal{E}$. If α is not an isomorphism we will show that $J(\widetilde{\nabla}) \geq J_1 := \nu_{\mathcal{E}}(ker(\alpha)) - \nu_{\mathcal{E}}(im(\alpha))$ where $\nu_{\mathcal{E}}(\mathcal{F}) := rk(\mathcal{F})(\mu(\mathcal{E}) - \mu(\mathcal{F}))$ for $\mathcal{F} < \mathcal{E}$. By a result of *N. Buchdahl* (*Lemma 2* [7]) we will also show that there exists a connection ∇ on $\mathcal{E}$ compatible with $\overline{\partial}_{\mathcal{E}}$ such that $J(\nabla) < J_1$. By the lower semi–continuity property of J we have that $J(\widetilde{\nabla}) = inf\ J_{|O(\mathcal{E}_{\widetilde{\nabla}})} \leq inf\ J_{|O(\mathcal{E})} \leq J(\nabla)$ thus yielding a contradiction; thus concluding that $\mathcal{E}_{\widetilde{\nabla}} \cong \mathcal{E}$ and $\widetilde{\nabla}$ minimises $J_{|O(\mathcal{E})}$. Finally, by considering small variations within $O(\mathcal{E})$ we deduce that $\widetilde{\nabla}$ minimises $J_{|O(\mathcal{E})}$ precisely when $J(\widetilde{\nabla}) = 0$ which is the condition that $\widetilde{\nabla}$ is of type required by our main *Theorem 6.2.*

The functional J is constructed as follows.

On the space of Hermitian matrices define the functional

$$\nu(M) := tr(\sqrt{M^*M}).$$

Any Hermitian matrix can be diagonalised by a unitary matrix, for instance $U^{-1}MU = \Lambda$ where U is a unitary matrix whose columns are eigenvectors of M constituting an orthonormal basis of $\mathbb{C}^n$ and $\Lambda := diag(\lambda_i)$ the diagonal matrix of eigenvalues of M. This leads one to find that $\nu(M) = \sum_{i=1}^n |\lambda_i|$. Moreover, $\nu(M) = max_{\{e_i\}} \sum_i |\langle Me_i, e_i\rangle|$ for $\{e_i\}$ orthonormal bases of $\mathbb{C}^n$. These formulae for $\nu(M)$ show that ν is a norm on the space of Hermitian matrices. One observes by the latter formula for $\nu(M)$ that block matrices in $\mathbb{C}^{2n}$ of the form

$$M := \begin{bmatrix} A & B \\ B^* & D \end{bmatrix}$$

satisfy $\nu(M) \geq |tr(A)| + |tr(D)|$.

Given $(\mathcal{E}, h)$ a holomorphic Hermitian vector bundle over X, by applying ν fiber–wise on $\mathcal{E}$ we define on the space of self–adjoint sections $\sigma \in \Gamma(End(\mathcal{E}))$ the functional

$$N(\sigma) := \left[\int_X \nu(\sigma)^2\, vol\right]^{1/2}.$$

As ν is a norm, so too is N on the space of such sections. Moreover, N is norm-equivalent to the usual L^2 norm $\|\sigma\|_2^2 := \int_X |\sigma|^2\, vol := \int_X tr(\sigma^*\sigma)\, vol$ on the same space; for if $\sigma = U\Lambda U^{-1}$, for U unitary and Λ the diagonal matrix of eigenvalues of σ, then $tr(\sigma^*\sigma) = tr(\sigma^2) = \sum_i \lambda_i^2$, and as $\nu(\sigma)^2 = \{\sum_i |\lambda_i|\}^2 = \sum_i \lambda_i^2 + \sum_{i\neq j} |\lambda_i\lambda_j|$, then one observes $\|\sigma\|_2 \leq N(\sigma)\ \ \forall\sigma$. Conversely, $N(\sigma) \leq k\|\sigma\|_2$ where $k = rk(\mathcal{E})^2$. This means we can extend the norm N to L^2–sections.

With this norm we define the functional

$$J(\nabla) := N\left(\frac{*F_\nabla}{2\pi i} + \mu(\mathcal{E}) \cdot 1\right)$$

for L^2_1 unitary connections ∇ on $\mathcal{E}$ where $F_\nabla \in \mathcal{E}^2(End(\mathcal{E}))$ is the associated curvature and 1 in the expression denotes the section of $End(\mathcal{E})$ given by $x \mapsto I_k$ the $k \times k$ identity matrix.

This functional is in fact lower semi-continuous; that is, for any ∇ we have $J(\nabla) \leq \liminf_j J(\nabla_j)$ for $\nabla_j \xrightarrow{L^2_1} \nabla$. To show this we must prove the equivalent condition that given $\epsilon > 0$ there exists an $n \in \mathbb{N}$ such that $J(\nabla) \leq J(\nabla_j) + \epsilon$ for all $j \geq n$, which we shall now establish by considering the reverse of this inequality. Given $\epsilon > 0$ consider the set C_ϵ of all $\alpha \in L^2(X, End(\mathcal{E}))$ satisfying $N(\alpha + \mu \cdot 1) \leq J(\nabla) - \epsilon$. The closed set C_ϵ is convex since ν is a norm. As $\frac{*F_\nabla}{2\pi i} \notin C_\epsilon$ we can *separate* C_ϵ and $\frac{*F_\nabla}{2\pi i}$; that is, there is a *hyperplane* H (a linear functional $H \in L^2(X, End(\mathcal{E}))^*$) and a constant $c \in \mathbb{R}$ such that $H(\alpha) > c$ for all $\alpha \in C_\epsilon$ and $H(\alpha) < c$ when α lies in a neighbourhood of $\frac{*F_\nabla}{2\pi i}$. As $\nabla_j \xrightarrow{L^2_1} \nabla$ then $F_{\nabla_j} \xrightarrow{L^2} F_\nabla$ so for sufficiently large j we have that $H(\frac{*F_{\nabla_j}}{2\pi i}) < c$ hence there exists an $n \in \mathbb{N}$ such that $\frac{*F_{\nabla_j}}{2\pi i} \notin C_\epsilon$. That is, for all $j \geq n$ we have $J(\nabla_j) \geq J(\nabla) - \epsilon$.

We shall also require explicit reference to the structure of connections and their curvature on sub–bundles and quotient bundles. For this we briefly consider *extension classes* of holomorphic vector bundles.

Given holomorphic bundles $\mathcal{E}$, $\mathcal{E}'$, $\mathcal{E}''$ We say that $\mathcal{E}$ is given by an *extension* of $\mathcal{E}''$ by $\mathcal{E}'$ if there is an exact sequence

$$0 \longrightarrow \mathcal{E}' \xrightarrow{\;i\;} \mathcal{E} \xrightarrow{\;p\;} \mathcal{E}'' \longrightarrow 0$$

We furthermore say that two extensions of $\mathcal{E}''$ by $\mathcal{E}'$ are *equivalent* if there exists a commutative diagram of the form

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \mathcal{E}' & \xrightarrow{i_1} & \mathcal{E}_1 & \xrightarrow{p_1} & \mathcal{E}'' & \longrightarrow & 0 \\ & & \| & & \downarrow & & \| & & \\ 0 & \longrightarrow & \mathcal{E}' & \xrightarrow{i_2} & \mathcal{E}_2 & \xrightarrow{p_2} & \mathcal{E}'' & \longrightarrow & 0 \end{array}$$

Upon applying the functor $Hom(\mathcal{E}'', \cdot\,)$ to an extension of $\mathcal{E}''$ by $\mathcal{E}'$ we obtain the exact sequence

$$0 \longrightarrow Hom(\mathcal{E}'', \mathcal{E}') \longrightarrow Hom(\mathcal{E}'', \mathcal{E}) \longrightarrow Hom(\mathcal{E}'', \mathcal{E}'') \longrightarrow 0.$$

Forming the associated long exact Čech cohomology sequence we have the connecting homomorphism

$$\delta^* : \check{H}^1(X, \mathcal{O}(Hom(\mathcal{E}'', \mathcal{E}''))) \longrightarrow \check{H}^1(X, \mathcal{O}(Hom(\mathcal{E}'', \mathcal{E}'))).$$

For an extension to split holomorphically we require the presence of a holomorphic map $\alpha : \mathcal{E} \longrightarrow \mathcal{E}'$ such that $p \circ \alpha = 1_\mathcal{E}$, which translates to $\delta^*(1_{\mathcal{E}''}) = 0$. Thus the Čech cohomology class $\delta^*(1_{\mathcal{E}''})$ is the "obstruction" to the holomorphic splitting of an extension of $\mathcal{E}''$ by $\mathcal{E}'$, which we call the *extension class*. In fact we have [30]

[30] In the general setting, the splitting of an extension of coherent sheaves of $\mathcal{O}_X$–modules $\mathcal{E}''$ by $\mathcal{E}'$ over a compact complex manifold X is characterised by the *hypercohomology* group $Ext^1(X; \mathcal{E}'', \mathcal{E}') := \mathbb{H}^1(X, Hom_\mathcal{O}(E.(\mathcal{E}''), \mathcal{E}'))$ where $E.(\mathcal{E}'')$ is a *global syzygy* of the sheaf $\mathcal{E}''$. When $\mathcal{E}''$ is locally free then $Ext^1(X; \mathcal{E}'', \mathcal{E}') \cong \check{H}^1(X, \mathcal{O}(Hom(\mathcal{E}'', \mathcal{E}')))$. See Chapter 5§4 [17] for details; the proof of *Lemma 6.5* appears on p.725.

6.5. Lemma.
There is a natural bijection $(\mathcal{E}, i, p) \mapsto \delta^*(1_{\mathcal{E}''})$ between the equivalence classes of extensions of $\mathcal{E}''$ by $\mathcal{E}'$ and the cohomology group $\check{H}^1(X, \mathcal{O}(Hom(\mathcal{E}'', \mathcal{E}')))$.

Using the operator $\overline{\partial} = \overline{\partial}_{\mathcal{E}}$ defining the holomorphic structure $\mathcal{E}$, we have by the *Dolbeault isomorphism theorem* [31] that $\beta \in \mathcal{E}^{(0,1)}(Hom(\mathcal{E}'', \mathcal{E}'))$, satisfying $\overline{\partial}(\alpha(s)) - \alpha(\overline{\partial}(s)) = i(\beta(s))$ for all $s \in \Gamma(\mathcal{E}'')$, corresponds to a representative of the extension class $\delta^*(1_{\mathcal{E}''})$. Upon fixing a Hermitian metric h on $\mathcal{E}$ this Dolbeault representative can be expressed in terms of connections by choosing unitary connections ∇', ∇'' on $\mathcal{E}', \mathcal{E}''$ respectively and comparing a connection ∇ on the smooth splitting $\mathcal{E}' \bigoplus \mathcal{E}''$ with $\nabla' \oplus \nabla''$. The representative β then corresponds to the tensor in the following expression

$$A = \begin{bmatrix} A' & \beta \\ -\beta^* & A'' \end{bmatrix}$$

where A' and A'' are the connection matrices of the respective unitary connections ∇' and ∇'' with respect to local unitary frames $\gamma_{\mathcal{E}'}$ and $\gamma_{\mathcal{E}''}$ for $\mathcal{E}'$ and $\mathcal{E}''$ respectively with respect to the induced metrics, and A the connection matrix of ∇ with respect to the local unitary frame (with respect to h) $\{\gamma_{\mathcal{E}'}, \gamma_{\mathcal{E}''}\}$ of $\mathcal{E}$. Moreover,

$$\begin{aligned} F_\nabla &= dA + A \wedge A \\ &= \begin{bmatrix} dA' & d\beta \\ -d\beta^* & dA'' \end{bmatrix} + \begin{bmatrix} A' \wedge A' - \beta \wedge \beta^* & A' \wedge \beta + \beta \wedge A'' \\ -\beta^* \wedge A' - A'' \wedge \beta^* & -\beta^* \wedge \beta + A'' \wedge A'' \end{bmatrix} \\ &= \begin{bmatrix} F_{\nabla'} - \beta \wedge \beta^* & d_{\nabla_{\mathcal{E}''^* \otimes \mathcal{E}'}} \beta \\ -d_{\nabla_{\mathcal{E}''^* \otimes \mathcal{E}'}} \beta^* & F_{\nabla''} - \beta^* \wedge \beta \end{bmatrix} \end{aligned}$$

where $d_{\nabla_{\mathcal{E}''^* \otimes \mathcal{E}'}}$ is the linear operator $\mathcal{E}^1(\mathcal{E}''^* \bigotimes \mathcal{E}') \longrightarrow \mathcal{E}^2(\mathcal{E}''^* \bigotimes \mathcal{E}')$ induced by the connection $\nabla_{\mathcal{E}''^* \otimes \mathcal{E}'}$ which is itself constructed from the connections ∇'' and ∇', which acts on elements $\alpha := \omega \otimes (\sigma \otimes \tau)$ for $\omega \in \mathcal{E}^1(X)$, $\sigma \in \Gamma(\mathcal{E}''^*)$ and $\tau \in \Gamma(\mathcal{E}')$ by

$$d_{\nabla_{\mathcal{E}''^* \otimes S}} \alpha := d\omega \wedge (\sigma \otimes \tau) - \omega \wedge \nabla_{\mathcal{E}''^* \otimes \mathcal{E}'}(\sigma \otimes \tau)$$

where

$$\nabla_{\mathcal{E}''^* \otimes \mathcal{E}'}(\sigma \otimes \tau) := \nabla_{\mathcal{E}''^*} \sigma \wedge \tau + \sigma \wedge \nabla' \tau.$$

As the bulk of the work done in this section resides in proving the necessary condition of *Theorem 6.2*, we shall firstly prove its sufficient condition.

6.6. Proposition.
An indecomposable holomorphic Hermitian vector bundle $(\mathcal{E}, h)$ over a compact Riemann surface (X, g) endowed with an H–E unitary connection ∇ with factor $-2\pi i \mu(\mathcal{E})$ is stable.
Proof:

[31] The Dolbeault isomorphism theorem states that for a holomorphic vector bundle E over a complex manifold X we have $\check{H}^1(X, \mathcal{O}(E)) \cong \mathcal{E}^{(0,1)}(E)/\overline{\partial}\Gamma(E)$. See p.45[17]

Let $\mathcal{E}'$ be a holomorphic sub–bundle of $\mathcal{E}$ and $\mathcal{E}''$ its orthogonal complement in $(\mathcal{E}, h)$. If $\nabla_{\mathcal{E}'}$ and $\nabla_{\mathcal{E}''}$ are the respective metric connections on $(\mathcal{E}', h_{|\mathcal{E}'})$ and $(\mathcal{E}'', h_{|\mathcal{E}''})$, our H–E connection ∇ on $(\mathcal{E}, h)$ will have curvature (as calculated earlier) of the form

$$F_\nabla = \begin{bmatrix} F_{\nabla_{\mathcal{E}'}} - \beta \wedge \beta^* & d_{\nabla_{\mathcal{E}''^* \otimes \mathcal{E}'}} \beta \\ -d_{\nabla_{\mathcal{E}''^* \otimes \mathcal{E}'}} \beta^* & F_{\nabla_{\mathcal{E}''}} - \beta^* \wedge \beta \end{bmatrix}$$

for some $\beta \in \mathcal{E}^{(0,1)}(\mathcal{E}''^* \bigotimes \mathcal{E}')$.

If $\{e_i\}$ is a unitary frame for $(\mathcal{E}, h)$ such that $\{e_1, \cdots, e_p\}$ is a unitary frame for $(\mathcal{E}', h_{|\mathcal{E}'})$ where $p := rk(\mathcal{E}')$, then we write $\beta e_a = \sum_\lambda \omega_a^\lambda \otimes e_\lambda$ where $\omega_a^\lambda \in \mathcal{E}^{(0,1)}(X)$

From the above decomposition of F_∇ we obtain vector bundle analogues of the *Gauss–Codazzi equations*

$$\Omega'^a_b = \Omega^a_b - \sum_{b \leq p < \lambda \leq r} \omega_b^\lambda \wedge \overline{\omega}_a^\lambda$$

where $a \geq 1$, $r := rk(\mathcal{E})$, and $F_\nabla = (\Omega^a_b)$, $F_{\nabla_{\mathcal{E}'}} = (\Omega'^a_b)$.

Representatives of the first Chern classes of $\mathcal{E}$ and $\mathcal{E}'$ are given respectively by

$$c_1(\mathcal{E}, h) := -\frac{1}{2\pi i} \sum_{j=1}^{r} \Omega^j_j,$$

$$c_1(\mathcal{E}', h) := -\frac{1}{2\pi i} \sum_{\alpha=1}^{p} \Omega'^\alpha_\alpha.$$

That our connection ∇ is H–E is equivalent to $\Omega^i_j = \gamma \delta^i_j$ for the $(1,1)$–form $\gamma := -2\pi i \mu(\mathcal{E})\, vol.$ Therefore

$$deg(\mathcal{E}) = \int_X c_1(\mathcal{E}, h) = -\frac{1}{2\pi i} \int_X r\gamma$$

$$\text{and} \quad deg(\mathcal{E}') = \int_X c_1(\mathcal{E}', h) = -\frac{1}{2\pi i} \int_X (p\gamma - \sum_{1 \leq a \leq p < \lambda \leq r} \omega_a^\lambda \wedge \overline{\omega}_a^\lambda),$$

$$\text{hence} \quad \mu(\mathcal{E}) = -\frac{1}{2\pi i} \int_X \gamma$$

$$\text{and} \quad \mu(\mathcal{E}') = -\frac{1}{2\pi i} \int_X \gamma + \frac{1}{2\pi i p} \int_X \sum_{1 \leq a \leq p < \lambda \leq r} \omega_a^\lambda \wedge \overline{\omega}_a^\lambda.$$

Thus $\mu(\mathcal{E}') \leq \mu(\mathcal{E})$ with equality holding if and only if all the $\omega_a^\lambda = 0$. The vanishing of β implies $\mathcal{E} = \mathcal{E}' \bigoplus \mathcal{E}''$ both holomorphically and orthogonally (since an extension splits holomorphically if and only if the Dolbeault representative of the extension class is $\overline{\partial}$–exact) ; however we have assumed that $\mathcal{E}$ is indecomposable, thus $\mu(\mathcal{E}') < \mu(\mathcal{E})$.

□

The following lemma provides the setting for our earlier discussed methodology of proving *Theorem 6.2.*

6.7. Lemma.
Let $\mathcal{E}$ be a holomorphic vector bundle over X. Then either $\inf J_{|O(\mathcal{E})}$ is attained in $O(\mathcal{E})$ or there is a holomorphic bundle $\mathcal{F} \not\cong \mathcal{E}$ of same degree and rank as $\mathcal{E}$ with $\inf J_{|O(\mathcal{F})} \leq \inf J_{|O(\mathcal{E})}$ and $Hom(\mathcal{E}, \mathcal{F}) \neq 0$.
Proof:
Let $\{\nabla_j\}$ be a sequence of smooth connections minimising $J_{|O(\mathcal{E})}$. As N is equivalent to the L^2 norm we have $\|F_{\nabla_j}\|_2$ bounded for all j. By *Theorem 1.5* [32] there is a subsequence $\{\nabla_{j'}\}$ of our sequence and L^2_2 gauge transformations $g_{j'}$ such that $g_{j'} \cdot \nabla_{j'} \stackrel{L^2_1}{\rightharpoonup} \widetilde{\nabla}$, say. As $\{g_{j'} \cdot \nabla_{j'}\}$ is also a minimising sequence for $J_{|O(\mathcal{E})}$ we may assume without loss of generality that $\nabla_{j'} \stackrel{L^2_1}{\rightharpoonup} \widetilde{\nabla}$.

As J has the lower semi-continuity property, we have

$$J(\widetilde{\nabla}) \leq \liminf_{j'} J(\nabla_{j'}) = \inf J_{|O(\mathcal{E})}.$$

We proceed to show that the alternative bundle $\mathcal{F}$ in the statement of the lemma is actually $\mathcal{E}_{\widetilde{\nabla}}$.

Let ∇_o be a L^2_1 unitary connection on the underlying smooth bundle E that induces a holomorphic structure isomorphic $\mathcal{E}$; here we have implicitly used the result of *Lemma 8* [7] which states that L^2_1 unitary connections on E induce holomorphic structures. Construct a connection on $Hom(E, E) \cong E^* \bigotimes E$ as earlier composed of the (unitary) connection ∇'' on E^*, induced by ∇_o, and $\widetilde{\nabla}$. Denote the $(0,1)$-component of this connection by $\overline{\partial}_{\nabla_o \widetilde{\nabla}}$. This differential operator is 1^{st}–order elliptic since $\overline{\partial}$ is. As $\overline{\partial}_{\nabla_o \widetilde{\nabla}} s = 0$ means that s is a holomorphic section of $Hom(E, E)$, solutions of $\overline{\partial}_{\nabla_o \widetilde{\nabla}} s = 0$ correspond exactly to elements of $Hom(\mathcal{E}, \mathcal{E}_{\widetilde{\nabla}})$. So if we assume $Hom(\mathcal{E}, \mathcal{E}_{\widetilde{\nabla}}) = 0$ then $\overline{\partial}_{\nabla_o \widetilde{\nabla}}$ has no kernel, so we obtain the *elliptic estimate* [32]

$$\|\overline{\partial}_{\nabla_o \widetilde{\nabla}} s\|_2 \geq c\|s\|_{2,1} \quad c > 0, \text{ for all } s.$$

The Sobolev imbedding $L^2_1 \subset L^4$ gives rise to the inequality $\|s\|_{2,1} \geq c'\|s\|_4$ for some $c' > 0$. Moreover, the Sobolev imbedding $L^2_1 \hookrightarrow L^4$ is compact [33] thus $\{\nabla_{j'}\}$ has a subsequence $\{\nabla_{j''}\}$ converging (strongly) to the L^4 connection $\widetilde{\nabla}$.

As

$$(\overline{\partial}_{\nabla_o \widetilde{\nabla}} - \overline{\partial}_{\nabla_o \nabla_{j''}})(\sigma \otimes \tau) = (B - A_{j''})_{(0,1)} \otimes (\sigma \otimes \tau)$$
$$\text{where } \nabla_{j''} = d + A_{j''} \text{ and } \widetilde{\nabla} = d + B,$$

the Hölder inequality gives

$$\|(\overline{\partial}_{\nabla_o \widetilde{\nabla}} - \overline{\partial}_{\nabla_o \nabla_{j''}}) s\|_2 \leq c''\|A_{j''} - B\|_4 \|s\|_4 \quad \text{for some } c'' > 0.$$

Using this inequality and the previous elliptic estimate and Sobolev inequality, we have

$$\|\overline{\partial}_{\nabla_o \nabla_{j''}} s\|_2 \geq (c' - c''\|A_{j''} - B\|_4)\|s\|_4 \quad \text{for each } j'' \text{ and for all } s.$$

[32] If L is an elliptic operator of order l on $\Gamma(\xi)$ for some vector bundle ξ then for each $k \geq 0$ there exists a $c = c(k) > 0$ such that for all $s \in \Gamma(\xi)$ we have $\|s\|_{2,k+l} \leq c(\|Ls\|_{2,k} + \|s\|_{2,k})$. Furthermore, if s is L^2_k–orthogonal to $ker(L)$ we can omit the $\|s\|_{2,k}$ term. See *Theorem 4.1 Appendix*, §4 [23].

[33] The *Rellich-Kondrakov theorem* states that given a compact Riemannian manifold X^n and integers integers $j \geq 0$, $m \geq 1$ and real numbers p and $q \geq 1$ such that $1 \leq p < nq/(n - mq)$ the imbedding of L^q_{j+m} in L^p_j is compact.

As $\nabla_{j''} \stackrel{L^4}{\to} \widetilde{\nabla}$, then given $\epsilon > 0$ there exists an $n \in \mathbb{N}$ such that $\|A_{j''} - B\|_4 < \epsilon \ \forall j'' \geq n$. Upon choosing $\epsilon = c'/c''$ then for large enough j'' we have $\|\overline{\partial}_{\nabla_o \nabla_{j''}} s\|_2 > 0$ for all s. As elements of $Hom(\mathcal{E}, \mathcal{E}_{\nabla_{j''}})$ correspond to solutions of $\overline{\partial}_{\nabla_o \nabla_{j''}} s = 0$, our result implies $Hom(\mathcal{E}, \mathcal{E}_{\nabla_{j''}}) = 0$, which is a contradiction since $\mathcal{E} \cong \mathcal{E}_{\nabla_{j''}}$ as the ∇_j were chosen to lie in the orbit $O(\mathcal{E})$.

Thus we conclude $Hom(\mathcal{E}, \mathcal{E}_{\widetilde{\nabla}}) \neq 0$.

We therefore conclude:

(i):

if $\widetilde{\nabla} \in O(\mathcal{E})$, that is $\mathcal{E}_{\widetilde{\nabla}} \cong \mathcal{E}$, then by the lower semi-continuity of J we have $J(\widetilde{\nabla}) \leq \liminf_{j'} J(\nabla_{j'}) = \inf J_{|O(\mathcal{E})} \leq J(\widetilde{\nabla})$, therefore $J(\widetilde{\nabla}) = \inf J_{|O(\mathcal{E})}$ which means that $\inf J_{|O(\mathcal{E})}$ is attained in $O(\mathcal{E})$.

otherwise

(ii):

if $\mathcal{E}_{\widetilde{\nabla}} \not\cong \mathcal{E}$, as $\nabla_{j'} \stackrel{L^2_1}{\rightharpoonup} \widetilde{\nabla}$, whence $F_{\nabla_{j'}} \stackrel{L^2}{\rightharpoonup} F_{\widetilde{\nabla}}$, therefore $\lim_{j' \to \infty} deg(\mathcal{E}_{\nabla_{j'}}) = deg(\mathcal{E}_{\widetilde{\nabla}})$ since $c_1(\xi) := [-\frac{1}{2\pi i} tr(F_\nabla)]$ for any connection ∇ on E. As $deg(\mathcal{E}_{\nabla_{j'}}) = deg(\mathcal{E})$ since the $\nabla_{j'}$ were chosen in $O(\mathcal{E})$, we have that $\mathcal{E}_{\widetilde{\nabla}}$ is of the same degree (and rank) as $\mathcal{E}$.

For the limiting connection $\widetilde{\nabla}$ we have by the semi-continuity property of J that $\inf J_{|\mathcal{O}(\mathcal{E}_{\widetilde{\nabla}})} = J(\widetilde{\nabla}) \leq \inf J_{|O(\mathcal{E})}$.

□

We now proceed to show that if $\mathcal{E}$ is a stable holomorphic bundle then the second alternative of the above lemma does not occur.

6.8. Lemma.
If $\mathcal{F}$ is a holomorphic vector bundle over X expressible as an extension
$0 \to \mathcal{M} \to \mathcal{F} \to \mathcal{N} \to 0$, and if $\mu(\mathcal{M}) \geq \mu(\mathcal{F})$ (whence $\mu(\mathcal{F}) \geq \mu(\mathcal{N})$), then for any unitary connection ∇ on $\mathcal{F}$

$$\begin{aligned} J(\nabla) &\geq rk(\mathcal{M})\{\mu(\mathcal{M}) - \mu(\mathcal{F})\} + rk(\mathcal{N})\{\mu(\mathcal{F}) - \mu(\mathcal{N})\} \\ &=: J_o, \end{aligned}$$

with equality holding only if the extension splits (that is, only if $\mathcal{F} \cong \mathcal{M} \bigoplus \mathcal{N}$ holomorphically).

Proof:
Fixing an hermitian metric on $\mathcal{F}$ and letting $\nabla_\mathcal{M}$ and $\nabla_\mathcal{N}$ be the metric connections on $\mathcal{M}$ and $\mathcal{N}$ respectively with respect to the naturally induced metrics on these bundles, a unitary connection ∇ on $\mathcal{F}$ with respect to the chosen metric will have curvature of the form

$$F_\nabla = \begin{bmatrix} F_{\nabla_\mathcal{M}} - \beta \wedge \beta^* & d_{\nabla_{\mathcal{N}^* \otimes \mathcal{M}}} \beta \\ -d_{\nabla_{\mathcal{N}^* \otimes \mathcal{M}}} \beta^* & F_{\nabla_\mathcal{N}} - \beta^* \wedge \beta \end{bmatrix}$$

where $\beta \in \mathcal{E}^{(0,1)}(\mathcal{N}^* \bigotimes \mathcal{M})$ is a representative of the extension class. Using the properties of the trace-norm ν on block matrices we have that

$$\nu\left(\frac{*F_\nabla}{2\pi i} + \mu(\mathcal{F})1\right) \geq \left| tr\left\{ \frac{*F_{\nabla_\mathcal{M}} - *(\beta \wedge \beta^*)}{2\pi i} + \mu(\mathcal{F})1 \right\} \right| + \left| tr\left\{ \frac{*F_{\nabla_\mathcal{N}} - *(\beta^* \wedge \beta)}{2\pi i} + \mu(\mathcal{F})1 \right\} \right|.$$

Therefore

$$J(\nabla) \geq \int_X \nu\left(\frac{*F_\nabla}{2\pi i} + \mu(\mathcal{F})1\right) vol \tag{6.8.1}$$

$$\geq \left|\int_X \left\{tr\left(\frac{*F_{\nabla_\mathcal{M}}}{2\pi i} + \mu(\mathcal{F})1\right) - |\beta|^2\right\} vol\right| + \left|\int_X \left\{tr\left(\frac{*F_{\nabla_\mathcal{N}}}{2\pi i} + \mu(\mathcal{F})1\right) + |\beta|^2\right\} vol\right| \tag{6.8.2}$$

the first inequality following from the *Jensen inequality* [34] since the operation of squaring is convex; the second inequality follows from the fact that $*tr(\beta^* \wedge \beta) = -2\pi i|\beta|^2$.

As

$$\int_X \left\{tr\left(\frac{*F_{\nabla_\mathcal{M}}}{2\pi i} + \mu(\mathcal{F})1\right) - |\beta|^2\right\} vol = \int_X tr\left(\frac{*F_{\nabla_\mathcal{M}}}{2\pi i}\right) vol + \mu(\mathcal{F}) \int_X tr(1_\mathcal{M})\, vol - \|\beta\|_2^2$$
$$= -deg(\mathcal{M}) + \mu(\mathcal{F})rk(\mathcal{M}) - \|\beta\|_2^2$$

As $\mu(\mathcal{F}) \leq \mu(\mathcal{M})$ by assumption the first term on the right hand side of inequality (6.8.2) is thus

$$deg(\mathcal{M}) - \mu(\mathcal{F})rk(\mathcal{M}) + \|\beta\|_2^2$$
$$= rk(\mathcal{M})\big(\mu(\mathcal{M}) - \mu(\mathcal{F})\big) + \|\beta\|_2^2.$$

Similarly, the second term in inequality (6.8.2) is

$$\mu(\mathcal{F})rk(\mathcal{N}) - deg(\mathcal{N}) + \|\beta\|_2^2$$
$$= rk(\mathcal{N})\big(\mu(\mathcal{F}) - \mu(\mathcal{N})\big) + \|\beta\|_2^2,$$

thus

$$J(\nabla) \geq rk(\mathcal{M})\big(\mu(\mathcal{M}) - \mu(\mathcal{F})\big) + rk(\mathcal{N})\big(\mu(\mathcal{F}) - \mu(\mathcal{N})\big) + 2\|\beta\|_2^2$$
$$=: J_o + 2\|\beta\|_2^2$$
$$\geq J_o$$

One observes that equality holds in this expression only if $\|\beta\|_2 = 0$, that is $\beta = 0$ almost everywhere which implies that the extension splits.

□

We wish to use the above lemmae to prove our main *Theorem 6.2* by induction. In order to proceed inductively, we must show that *Theorem 6.2* holds for lines bundles.

Clearly, every line bundle is stable and indecomposable. Let $\mathcal{H}^1(X)$ be the subspace of $\mathcal{E}^1(X)$ consisting of *harmonic forms*; that is, 1–forms ω satisfying $d^*d\omega = 0$. The *Hodge decomposition theorem* states that on a compact Riemann surface X

$$\mathcal{E}^1(X) \cong \mathcal{H}^1(X) \bigoplus \partial\mathcal{E}^1(X) \bigoplus \overline{\partial}\mathcal{E}^1(X)$$

[34] The *Jensen inequality* states that $\int_X (\phi \circ f)\, d\mu \geq \phi(\int_X f\, d\mu)$ for $\phi : \mathbb{R} \longrightarrow \mathbb{R}$ a convex function and $f : X \longrightarrow \mathbb{R}$ integrable.

From this one concludes that a 2–form $\omega \in \mathcal{E}^2(X)$ satisfying $\int_X \omega = 0$ is expressible as $\omega = \partial\overline{\partial} f$ for some $f \in \mathcal{E}^0(X)$. Upon taking $\omega := F_\nabla + 2\pi i\mu(\mathcal{E})\, vol$ for ∇ any unitary connection on our holomorphic line bundle $\mathcal{E}$ we observe that $\int_X \omega = 0$ (as we have assumed $\int_X vol = 1$), hence $F_\nabla + 2\pi i\mu(\mathcal{E})\, vol = \partial\overline{\partial} f$ for some $f \in \mathcal{E}^0(X)$. By the affine structure of $\mathcal{A}(\mathcal{E})$ we may write $\nabla = \nabla' + a$ for some other unitary connection ∇' on $\mathcal{E}$ where $a \in \mathcal{E}^1(X)$. If ∇' is chosen so that $a = \overline{\partial} f$ then we have that $F_\nabla = F_{\nabla'} + \partial\overline{\partial} f$. Thus we have shown that there exists a unitary connection ∇' on $\mathcal{E}$ such that $*F_{\nabla'} = -2\pi i\mu(\mathcal{E})$.

Thus *Theorem 6.2* holds for holomorphic line bundles, and we use this start in induction to prove the following Lemma. We utilise *Lemma 2* [7] in order to avoid calculations involving *semi-stable filtrations* [2]; that is, a nested sequence of proper sub–bundles $0 =: \mathcal{P}_o < \mathcal{P}_1 < \cdots < \mathcal{P}_k := \mathcal{P}$ with each $\mathcal{P}_i/\mathcal{P}_{i-1}$ semi-stable and associated slopes $\mu(\mathcal{P}_i/\mathcal{P}_{i-1})$ decreasing with i. These filtrations are used in [9] to sequentially decompose an arbitrary extension of a stable bundle into a sequence of stable bundles. *Lemma 2* [7] avoids this by showing that there exists an extension of a stable bundle by stable bundles. With this adoption, Donaldson's[9] *Lemma 3* is modified to the following:

6.9. Lemma.
Given $\mathcal{E}$ a stable holomorphic bundle, and assuming *Theorem 6.2* holds true for bundles of rank less than $rk(\mathcal{E})$, if $\mathcal{E}$ can be expressed as an extension $0 \to \mathcal{P} \to \mathcal{E} \to \mathcal{L} \to 0$ with $\mathcal{P}$ and $\mathcal{L}$ stable, then there is a smooth unitary connection $\nabla \in O(\mathcal{E})$ with

$$\begin{aligned} J(\nabla) &< rk(\mathcal{P})\{\mu(\mathcal{E}) - \mu(\mathcal{P})\} + rk(\mathcal{L})\{\mu(\mathcal{L}) - \mu(\mathcal{E})\} \\ &=: J_1. \end{aligned}$$

Proof:
On $\mathcal{P}$, $\mathcal{L}$ fix the *H–E* connections which exist by the inductive hypothesis, and set $\beta \in \mathcal{E}^{(0,1)}(\mathcal{L}^* \bigotimes \mathcal{P})$ to be a representative of the extension class.

The operator $Q := - * \partial_{Hom(\mathcal{L},\mathcal{P})}\overline{\partial}_{Hom(\mathcal{L},\mathcal{P})}$ acting on smooth sections of $Hom(\mathcal{L},\mathcal{P})$ satisfies $Q^* + Q = \triangle - *F_\nabla$ where ∇ is the metric connection on $Hom(\mathcal{L},\mathcal{P})$ and $\triangle := d^*_\nabla d_\nabla$ is the Laplacian. With the induced H–E connections on $\mathcal{P}$ and $\mathcal{L}$ the metric connection on $Hom(\mathcal{L},\mathcal{P}) \cong \mathcal{L}^* \bigotimes \mathcal{P}$ has factor $-2\pi i(\mu(\mathcal{P}) - \mu(\mathcal{L}))$. As $\mathcal{E}$ is stable, then $\mu(\mathcal{P}) < \mu(\mathcal{L})$, thus Q^* has no kernel whence Q is surjective. Thus there exists a $\gamma \in Hom(\mathcal{L},\mathcal{P})$ such that $\partial(\beta + \overline{\partial}\gamma) = 0$. Modify β in this way so that $\partial\beta = 0$ and rescale so that $\|\beta\|_2 = 1$ (possible since $\beta \neq 0$ since $\mathcal{E}$ is stable).

Triples $(\nabla_\mathcal{P}, \nabla_\mathcal{L}, t\beta)$ for $t \neq 0$ gives us a connection $\nabla^t \in O(\mathcal{E})$ with curvature

$$F_{\nabla^t} = \begin{pmatrix} F_{\nabla_\mathcal{P}} - t^2\beta \wedge \beta^* & 0 \\ 0 & F_{\nabla_\mathcal{L}} - t^2\beta^* \wedge \beta \end{pmatrix}$$

By the inductive hypothesis $F_{\nabla_\mathcal{P}} = -2\pi i\mu(\mathcal{P})\, vol$ and $F_{\nabla_\mathcal{L}} = -2\pi i\mu(\mathcal{L})\, vol$. As $\mu(\mathcal{P}) < \mu(\mathcal{E}) < \mu(\mathcal{L})$, then for t small enough we have that the eigenvalues of $(\mu(\mathcal{E}) - \mu(\mathcal{P}))1 - \frac{t^2}{2\pi i} * (\beta \wedge \beta^*)$ are positive, and those of $(\mu(\mathcal{E}) - \mu(\mathcal{L}))1 - \frac{t^2}{2\pi i} * (\beta^* \wedge \beta)$ are negative. Therefore

$$\begin{aligned} \nu\left(\frac{*F_{\nabla^t}}{2\pi i} + \mu(\mathcal{E})1\right) &= rk(\mathcal{P})\{\mu(\mathcal{E}) - \mu(\mathcal{P})\} + rk(\mathcal{L})\{\mu(\mathcal{L}) - \mu(\mathcal{E})\} - 2t^2|\beta|^2 \\ &=: J_1 - 2t^2|\beta|^2, \end{aligned}$$

and so

$$J(\nabla^t)^2 = \int_X (J_1 - 2t^2\|\beta\|_2)^2$$
$$= {J_1}^2 - 4J_1 t^2\|\beta\|_2^2 + 4t^4 \int_X |\beta|^4.$$

Choosing t small enough so that $t^4 \int_X |\beta|^4 \ll t^2\|\beta\|_2^2 = t^2$ we then have the required result $J(\nabla^t) < J_1$. □

Observe that the last clause of *Lemma 6.8* says that if a holomorphic bundle $\mathcal{F}$ is indecomposable, that is, every extension $0 \to \mathcal{M} \to \mathcal{F} \to \mathcal{N} \to 0$ does not split, then $J(\nabla) > J_o$ for any unitary connection ∇ on $\mathcal{F}$.

Assuming ∇ satisfies *Theorem 6.2*, that is, $J(\nabla) = 0$, then $0 > J_o := rk(\mathcal{M})\{\mu(\mathcal{M}) - \mu(\mathcal{F})\} + rk(\mathcal{N})\{\mu(\mathcal{F}) - \mu(\mathcal{N})\}$. By assumption of *Lemma 6.8*, $\mu(\mathcal{M}) \geq \mu(\mathcal{F})$, thus $\mu(\mathcal{F}) \geq \mu(\mathcal{N})$, so if $\mathcal{N} \subset \mathcal{F}$ for all such extensions then $\mathcal{F}$ is stable.

Conversely, if $\mathcal{E}$ is a stable holomorphic bundle and *Theorem 6.2* holds for all bundles of rank less than that of $\mathcal{E}$ then the second alternative of *Lemma 6.7* does not occur; thus $\inf J_{|O(\mathcal{E})}$ is attained in $O(\mathcal{E})$. For, if the second alternative of *Lemma 6.7* did hold true, namely, if $\exists \mathcal{F} \not\cong \mathcal{E}$ with $rk(\mathcal{F}) = rk(\mathcal{E})$, $deg(\mathcal{E}) = deg(\mathcal{E})$ and $Hom(\mathcal{E}, \mathcal{F}) \neq 0$ such that $\inf J_{|O(\mathcal{F})} \leq \inf J_{|O(\mathcal{E})}$ then the following argument shows that we attain a contradiction.

As $Hom(\mathcal{E}, \mathcal{F}) \neq 0$, choose a non-zero sheaf map $\alpha : \mathcal{E} \longrightarrow \mathcal{F}$. We have extensions of these bundles and consequently a diagram

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \mathcal{P} & \longrightarrow & \mathcal{E} & \longrightarrow & \mathcal{L} & \longrightarrow & 0 \\ & & & & {\scriptstyle\alpha}\big\downarrow & & {\scriptstyle\beta}\big\downarrow & & \\ 0 & \longleftarrow & \mathcal{N} & \longleftarrow & \mathcal{F} & \longleftarrow & \mathcal{M} & \longleftarrow & 0 \end{array}$$

with exact rows, $rk(\mathcal{L}) = rk(\mathcal{M})$, $det(\beta) \not\equiv 0$, and $deg(\mathcal{L}) \leq deg(\mathcal{M})$. This diagram is constructed as follows. Given the exact sequences of sheaves $0 \to ker(\alpha) \to \mathcal{E} \to im(\alpha) \to 0$ and $0 \to im(\alpha) \to \mathcal{F} \to coker(\alpha) \to 0$. All the constituent sheaves, save $coker(\alpha)$, are locally free thus correspond to holomorphic bundles. By factoring the *analytic* sheaf $\mathcal{N}' := coker(\alpha)$ by its *torsion subsheaf*

$$\tau(\mathcal{N}') := \amalg_{x \in X}\{s \in \mathcal{N}'_x \mid fs = 0 \text{ for some } f \in \mathcal{O}_x\}$$

we obtain the sheaf $\mathcal{N} := \mathcal{N}'/\tau(\mathcal{N}')$ which is torsion free everywhere hence a locally free sheaf since X is a Riemann surface [35]. We thus obtain an exact sequence of locally free sheaves $0 \to \mathcal{M} \to \mathcal{F} \to \mathcal{N} \to 0$ for some locally free sheaf $\mathcal{M}$ and a map $\beta : im(\alpha) \longrightarrow \mathcal{M}$ with $deg(\beta) \not\equiv 0$ by "chasing" the preceding commutative diagram.

As $\mathcal{E} \cong \mathcal{P} \bigoplus \mathcal{L}$ and $\mathcal{F} \cong \mathcal{N} \bigoplus \mathcal{M}$ smoothly, then $rk(\mathcal{E}) = rk(\mathcal{P}) + rk(\mathcal{L})$ and $rk(\mathcal{F}) = rk(\mathcal{N}) + rk(\mathcal{M})$, whence as $rk(\mathcal{L}) = rk(\mathcal{M})$ from the diagram we have $rk(\mathcal{P}) = rk(\mathcal{N})$. We similarly obtain the inequality $deg(\mathcal{P}) \geq deg(\mathcal{N})$ from the properties of the first Chern class and the information from the diagram. Thus we conclude $J_1 \leq J_o$. Applying *Lemma 6.8* to the bottom

[35] In fact if X were a general complex manifold, the set of all points $x \in X$ on which a *coherent analytic sheaf* is not locally free is *analytic* and nowhere dense in X. See Chapter 4§4 [15] for details.

row of the preceding commutative diagram we have $J(\nabla) \geq J_o$ for any unitary connection ∇ on $\mathcal{F}$; in particular, $\inf J_{|O(\mathcal{F})} \geq J_o$. Also applying *Lemma 6.9* to the top row of our diagram, there exists a unitary connection ∇ on $\mathcal{E}$ with $J(\nabla) < J_1$; whence $\inf J_{|O(\mathcal{E})} < J_1$. However, we have assumed the second alternative of *Lemma 6.7*, namely that $\inf J_{|O(\mathcal{F})} \leq \inf J_{|O(\mathcal{E})}$, so $J_o < J_1$ hence a contradiction. Thus, $\inf J_{|O(\mathcal{E})}$ is attained in $O(\mathcal{E})$.

Consider the operator $d^*_\nabla d_\nabla$ acting on L^2_2 sections of $End(\mathcal{E})$. Recall that a *normal* linear operator on a complex inner-product space V with distinct eigenvalues $\lambda_1, ..., \lambda_k$ and corresponding eigenspaces W_j will decompose V into $W_1 \bigoplus \cdots \bigoplus W_k$. If $\sigma \in \Gamma(End(\mathcal{E}))$ is a non-constant element (that is, one not of the form $x \mapsto cI_{rk(\mathcal{E})}$) of $ker(d^*_\nabla d_\nabla)$ it would have eigenspaces that decompose $\mathcal{E}$ holomorphically contrary to the assumption that $\mathcal{E}$ is indecomposable. Thus the kernel of $d^*_\nabla d_\nabla$ consists only of the constant scalars. By the *Hodge decomposition theorem* for bundles, given $\mathcal{H} : \Gamma(End(\mathcal{E})) \longrightarrow \mathcal{H}^0(End(\mathcal{E}))$ the orthogonal projection and $G : \Gamma(End(\mathcal{E})) \longrightarrow \Gamma(End(\mathcal{E}))$ the *Green's* operator, we have $id = \mathcal{H} + \triangle G$ on $\Gamma(End(\mathcal{E}))$ where $\triangle$ is the *Laplacian* $d^*_\nabla d_\nabla + d_\nabla d^*_\nabla$. As $d_\nabla d^*_\nabla = 0$ on $\Gamma(End(\mathcal{E}))$ then $\triangle = d^*_\nabla d_\nabla$, hence $\mathcal{H}^0(End(\mathcal{E})) := ker(\triangle) = \{cI \mid c \in \mathbb{C}\}$, whence $-i * F_\nabla = -2\pi\mu(\mathcal{E})1 + d^*_\nabla d_\nabla G(-i * F_\nabla)$; that is, there exists $h \in \mathrm{L}^2_2(End(\mathcal{E}))$ such that $d^*_\nabla d_\nabla h = 2\pi\mu(\mathcal{E})1 - i * F_\nabla$.

For this element h and small $t \in \mathbb{R}$ define the gauge transformation $g_t := 1+th$, and subsequently define the connection

$$\begin{aligned}\nabla_t &:= g_t \cdot \nabla \in O(\mathcal{E}) \\ &:= \nabla - (\overline{\partial}_\nabla g_t)g_t^{-1} + ((\overline{\partial}_\nabla g_t)g_t^{-1})^* \\ &=: d + A_t\end{aligned}$$

whence

$$\begin{aligned}F_{\nabla_t} &:= dA_t + A_t \wedge A_t \\ &= F_\nabla - \partial_\nabla(\overline{\partial}_\nabla(g_t)g_t^{-1}) + \overline{\partial}_\nabla(g_t^{-1}\partial_\nabla(g_t)) - \overline{\partial}_\nabla(g_t)g_t^{-2}\partial_\nabla(g_t) + g_t^{-1}\partial_\nabla(g_t)\overline{\partial}_\nabla(g_t)g_t^{-1} \\ &= F_\nabla - t(\partial_\nabla\overline{\partial}_\nabla - \overline{\partial}_\nabla\partial_\nabla)h + q(t,h)\end{aligned}$$

where for small t $\|q(t,h)\|_2 \leq c\|h\|_{2,2}t^2$.

As $d^*_\nabla d_\nabla = i * (\overline{\partial}_\nabla\partial_\nabla - \partial_\nabla\overline{\partial}_\nabla)$ and as $d^*_\nabla d_\nabla h = 2\pi\mu(\mathcal{E})1 - i * F_\nabla$ then $-t(\partial_\nabla\overline{\partial}_\nabla - \overline{\partial}_\nabla\partial_\nabla)h = -t2\pi i\mu(\mathcal{E})vol - tF_\nabla$ whence

$$\frac{*F_{\nabla_t}}{2\pi i} + \mu(\mathcal{E})1 = \left\{\frac{*F_\nabla}{2\pi i} + \mu(\mathcal{E})1\right\}(1-t) + *q(t,h).$$

Therefore $J(\nabla_t) = J(\nabla)(1-t) + O\{t^2\}$.

As $J(\nabla) = \inf J_{|O(\mathcal{E})}$, then by the *calculus of variations* the *first variation* $\frac{d}{dt}J(\nabla_t)_{|t=0}$ must vanish, hence concluding $J(\nabla) = 0$; that is, $*F_\nabla = -2\pi i\mu(\mathcal{E})1$.

We must finally prove that our connection $\nabla \in O(\mathcal{E})$ is unique up to isomorphism. For this we firstly need the following lemma (see *Proposition 3(ii)* [10] for proof).

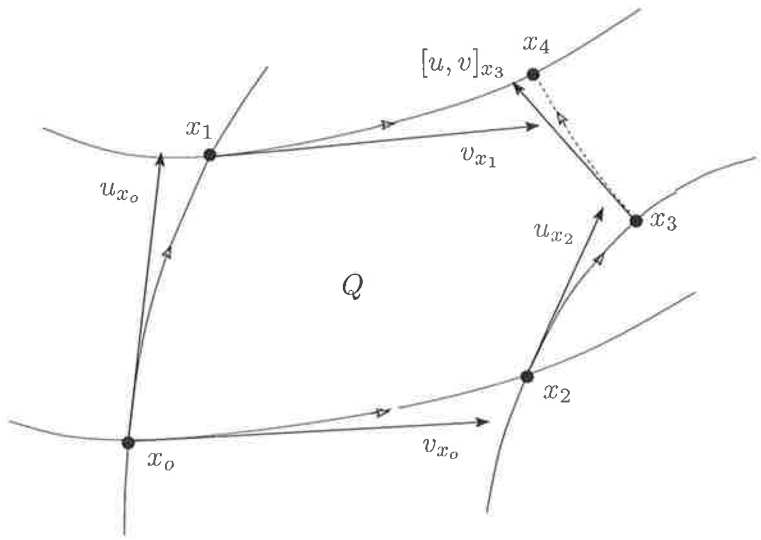
6.10. Lemma.
If $\mathcal{E}$ is a holomorphic bundle over X endowed with a H–E connection and $\mu(\mathcal{E}) = 0$ then any holomorphic section is covariantly constant.

If we have two unitary connections ∇_0, ∇_1 on E inducing isomorphic holomorphic structures $\mathcal{E}_{\nabla_0}$ and $\mathcal{E}_{\nabla_1}$ respectively, there must exist a complex automorphism g of E such that $\overline{\partial}_{\nabla_1} = g \circ \overline{\partial}_{\nabla_0} \circ g$ and $\partial_{\nabla_1} = g \circ \partial_{\nabla_0} \circ g$. Applying a unitary change of gauge to one of our connections, say by $g(g^*g)^{-1/2}$, g may be assumed to be positive self–adjoint. If ∇_0 and ∇_1 are our H–E connections on E, then the holomorphic isomorphism $g : \mathcal{E}_{\nabla_0} \longrightarrow \mathcal{E}_{\nabla_1}$ is covariantly constant by the above proposition. This means $0 = \partial_{\nabla_0}(g^*g) = \partial_{\nabla_0}(g^2)$ and $\overline{\partial}_{\nabla_0}(g^2) = 0$. As $\mathcal{E}_{\nabla_0}$ is indecomposable, $g^2 = c1_E$ where c is a constant. Furthermore, as g is positive self–adjoint then $g = c'1_E$ for c' a constant.

The equivalence between *Theorem 2*(A) §12 [27] of Narasimhan & Seshadri and Donaldson's *Theorem 6.2* is observed as follows.

A unitary connection ∇ on E with curvature of the form $F_\nabla = -2\pi i\mu(E)\, vol_X$ corresponds to a projective unitary representation $\rho : \pi_1(X) \to \mathbb{P}U(k)$ of the fundamental group where $\mathbb{P}U(k) := U(k)/U(1)I_k$ is the *projective unitary group.* This follows from the following observations.

Choosing vector fields $u, v \in \mathfrak{X}(X)$ their *integral curves* (paths γ being solutions of the differential equations $\dot{\gamma} = u$ or v) can be intersected to make a quadrilateral Q on X, being closed by a streamline to the commutator field $[v, u]$.



Observe here that the gap "$x_4 - x_3$" in the four legged curve is characterised by the difference $f(x_4) - f(x_3)$ for any $f \in \mathcal{F}(X)$, and in a given coördinate basis we have the Taylor's expansion $f(x_4) - f(x_3) = ((u_\beta v_{\alpha,\beta} - v_\beta u_{\alpha,\beta})\frac{\partial f}{\partial x_\alpha})_{x_3} +$ "cubic errors"$=: [u, v](f)_{x_3} +$ "cubic errors". Thus the vector $[u, v]$ describes the separation between points x_3 and x_4; its description gets arbitrarily accurate when u and v get arbitrarily short. Thus we keep the lengths of legs of the quadrilateral comparable/equal to the lengths of the tangential vectors, and these latter lengths are taken "small enough" so that the second–order terms in the expansion of the parallel transport equation $P(Q, \nabla)$ about Q are negligible.

As

$$\nabla_u \sigma(y) = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon}\{P(-\gamma, \nabla)\sigma(\gamma(\varepsilon)) - \sigma(y)\}$$

for $\sigma \in \Gamma(E)$ and γ a curve with $\gamma(0) = y$ and $\dot{\gamma} = u$, then to second order we have, for $p \in E_{|Q}$,

$$\begin{aligned} P(Q, \nabla)p &= p + \{\nabla_u \nabla_v - \nabla_v \nabla_u - \nabla_{[u,v]}\}p\,|u||v| \\ &=: p + F_\nabla(u, v)p\,|u||v|. \end{aligned}$$

For an arbitrary loop γ on X, break the region bounded by γ into a number of contingent quadrilaterals Q as defined above, whence for $\sigma \in \Gamma(E)$

$$\begin{aligned} P(\gamma, \nabla)\sigma &= \sigma + \sum_Q F_\nabla(u_Q, v_Q)\sigma\,|u_Q||v_Q| \\ &= \sigma + F_\nabla(u_\gamma, v_\gamma)\sigma \sum_Q |u_Q||v_Q| \end{aligned}$$

which is only valid for curves γ of small compass; the difference $P(\gamma, \nabla)\sigma - \sigma$ doubles when the area bounded by γ doubles, but the error increases by a factor of $\sim 2^{3/2}$.

Parallel transport about two different small curves γ_1 and γ_2 (based at the same point $x_o \in X$) for our "central" connection ∇ differs by a constant in $U(1)$ and so $[P(\gamma_1, \nabla)] = [P(\gamma_2, \nabla)]$ as equivalence classes in $\mathbb{P}U(k)$, and so we have a representation $\rho : \pi_1(X) \to \mathbb{P}U(k)$. Given a universal covering of $\widetilde{X} \to X$, the associated bundle $\widetilde{E} := \widetilde{X} \times_\rho \mathbb{CP}^{k-1}$ is isomorphic to the projectivisation $P(E)$.

Conversely, given a projective unitary representation $\rho : \pi_1(X) \to \mathbb{P}U(k)$ of the fundamental group we have a projectively flat connection ∇ on E. Given the natural homomorphism $\mu : U(k) \to \mathbb{P}U(k)$ with associated Lie algebra homomorphism $\mu' : \mathfrak{u}(k) \to \mathfrak{pu}(k)$ the curvature of the induced connection on $P(E)$ is given by $\mu'(F_\nabla)$, and so F_∇ takes on scalar multiples of id_E.

7 Stratifying the space $\mathcal{A}(E)$ of connections.

Assume X is a compact Riemann surface with volume normalised to unity and E is a smooth complex vector bundle over X of degree k with structure group $G = U(n)$.

We shall induce a Morse stratification on the space of unitary connections on E, which by abuse of notation is also denoted by $\mathcal{A}(E)$, by constructing one for the space $\mathcal{C}(E)$ of holomorphic structures on E via the identification of $\mathcal{A}$ with $\mathcal{C}$ as expounded in the previous chapter owing to the isomorphism $\mathcal{E}^1(\mathfrak{u}(n)) \cong \mathcal{E}^{(0,1)}(\mathfrak{gl}(n,\mathbb{C}))$. The stratification of $\mathcal{C}$ involves working with *canonical filtrations* of non–semi–stable holomorphic vector bundles over X. To this end we need some preparatory material on the existence and uniqueness of such flags.

7.1. Definition.
Given E a holomorphic vector bundle over X which is not semi–stable, a non–zero proper holomorphic sub–bundle F of E is said to be *strongly contradicting semi-stability* (SCSS) if it satisfies the following conditions

(C1): F is semi–stable;

(C2): for every sub–bundle F' of E containing F as a proper sub–bundle, we have $\mu(F) > \mu(F')$.

Note that condition (C2) and the following conditions are equivalent:

(C2$'$): for any non–zero sub–bundle Q of E/F we have $\mu(Q) < \mu(F)$;

(C2$''$): for any stable non–zero sub–bundle Q of E/F we have $\mu(Q) < \mu(F)$.

7.2. Lemma.
Given F_1 and F_2 sub–bundles of E such that F_1 is semi–stable and F_2 satisfies condition (C2), if F_1 is not contained in F_2 then $\mu(F_2) > \mu(F_1)$.
Proof:
Given that the canonical sheaf map map $f : F_1 \to E/F_2$ is non–zero by assumption, as expounded in the previous chapter we have a factorisation

$$\begin{array}{ccccc} F_1 & \longrightarrow & F_1' & \longrightarrow & 0 \\ f\big\downarrow & & g\big\downarrow & & \\ E/F_2 & \longleftarrow & F"_1 & \longleftarrow & 0 \end{array}$$

with $rk(F_1') = rk(F_1'')$, $det(g) \not\equiv 0$ and $deg(F_1') \leq deg(F_1'')$. As F_1 is semi–stable, $\mu(F_1) \leq \mu(F_1')$ and as F_2 satisfies condition (C2), $\mu(F_1'') < \mu(F_2)$. As $deg(F_1') \leq deg(F_1'')$ and $rk(F_1') = rk(F_1'')$ then $\mu(F_1') \leq \mu(F_1'')$ and so $\mu(F_1) < \mu(F_2)$.

□

7.3. Lemma.
Given sub–bundles F_1 and F_2 of E which are SCSS, then $F_1 = F_2$.
Proof:

If $F_1 \not\subset F_2$, *Lemma 7.2* implies $\mu(F_2) > \mu(F_1)$ and that we must have $F_2 \subsetneq F_1$. We immediately obtain a contradiction since $\mu(F_2) \leq \mu(F_1)$ owing to the semi–stability of F_1, and so $F_1 \subseteq F_2$. A similar argument shows $F_2 \subseteq F_1$.

□

7.4. Proposition.
If E is not semi–stable then it contains a unique SCSS sub–bundle.
Proof:
The Uniqueness of a SCSS sub–bundle follows from *Lemma 7.3.*

Let $m := \sup_{0 \neq F \subset E} \mu(F)$ whish exists since the values of the degrees are discrete and are bounded from above. As E is not semi–stable we have $m \geq \mu(E)$.

Choose a sub–bundle F_o of E of maximal rank such that $\mu(F_o) = m$. If F' is a non–zero sub–bundle of F_o we have $\mu(F') \leq m = \mu(F_o)$ so that F_o is semi–stable. On the other hand, a sub–bundle F' of E containing F_o as a proper sub–bundle satisfies $\mu(F') < \mu(F_o)$ and so F_o also satisfies condition (C2); that is, F_o is SCSS.

□

7.5. Lemma.
If a vector bundle bundle E is not semi–stable we have a flag

$$0 = F_0 \subsetneq F_1 \subsetneq \cdots \subsetneq F_k = E$$

satisfying conditions

(F1): F_i/F_{i-1} is semi–stable for $i = 1, \ldots, k$,

(F2): F_i/F_{i-1} is SCSS in E/F_{i-1} for $i = 1, \ldots, k-1$.

Moreover, such a flag is uniquely determined.
Proof:
Existence follows from *Proposition 7.4*. More explicitly, given F_1 a SCSS sub–bundle of E, if E/F_1 is not semi–stable we find another SCSS sub–bundle $F_2' \subset E/F_1$ and define F_2 to be the inverse of F_2' by the map $E \to E/F_1$. Repeating this procedure if necessary yields a flag satisfying conditions (F1) and (F2).

The uniqueness follows by induction on the rank of E by applying *Proposition 7.4* and noting that the sub–bundles F_i/F_{i-1} for $i \geq 2$ from a flag of E/F_1 satisfying conditions (F1) and (F2).

□

7.6. Lemma.
Given a flag

$$0 = F_0 \subsetneq F_1 \subsetneq \cdots \subsetneq F_k = E,$$

conditions (F1) and (F2) are equivalent to the conditions

(F1′): F_i/F_{i-1} is semi–stable for $i = 1, \ldots, k$,

(F2′): $\mu(F_i/F_{i-1}) > \mu(F_{i+1}/F_i)$ for $i = 1, \ldots, k-1$.

Proof:
If conditions (F1) and (F2) hold, we have an exact sequence

$$0 \longrightarrow F_i/F_{i-1} \longrightarrow E/F_{i-1} \longrightarrow E/F_i \longrightarrow 0.$$

As $F_{i+1}/F_i \subset E/F_i$ and F_i/F_{i-1} are SCSS in E/F_{i-1}, we have $\mu(F_i/F_{i-1}) > \mu(F_{i+1}/F_i)$.

Now, if conditions (F1$'$) and (F2$'$) are satisfied, we have the exact sequence

$$0 \longrightarrow F_{k-1}/F_{k-2} \longrightarrow E/F_{k-2} \longrightarrow E/F_{k-1} \longrightarrow 0,$$

whence $\mu(F_{k-1}/F_{k-2}) > \mu(E/F_{k-1})$ by condition (F2'). As E/F_{k-1} and F_{k-1}/F_{k-2} are semi–stable, we have that F_{k-1}/F_{k-2} is SCSS in E/F_{k-2}; this follows because if A is a bundle that is not semi–stable and B a non–zero sub–bundle of A satisfying the following conditions

(i): B and A/B are not semi–stable;

(ii): $\mu(B) > \mu(A/B)$,

then B is SCSS in A. In fact for any non–zero sub–bundle Q of A/B we have $\mu(Q) \leq \mu(A/B) < \mu(B)$, so that condition (C2$'$) is satisfied.

We prove that condition (F1) is satisfied by downward induction on i. Consider the exact sequences

$$0 \longrightarrow F_i/F_{i-1} \longrightarrow E/F_{i-1} \longrightarrow E/F_i \longrightarrow 0,$$

and

$$0 \longrightarrow F_{i+1}/F_i \longrightarrow E/F_i \longrightarrow E/F_{i+1} \longrightarrow 0.$$

To prove that F_i/F_{i-1} is SCSS in E/F_{i-1} we must show that for any stable non–zero sub–bundle Q of E/F_i we have $\mu(F_i/F_{i-1}) > \mu(Q)$; the required result will then follow from condition (C2$''$).

Given $Q \subset F_{i+1}/F_i$ we have $\mu(Q) \leq \mu(F_{i+1}/F_i)$ since F_{i+1}/F_i is semi–stable, and by hypothesis $\mu(F_{i+1}/F_i) < \mu(F_i/F_{i-1})$ so that $\mu(Q) < \mu(F_i/F_{i-1})$. If Q is not contained in F_{i+1}/F_i, then by induction we may assume F_{i+1}/F_i is SCSS in E/F_i. As Q is semi–stable and $Q \subset F_{i+1}/F_i$, we have by *Lemma 7.2* that $\mu(Q) < \mu(F_{i+1}/F_i)$. As $\mu(F_{i+1}/F_i) < \mu(F_i/F_{i-1})$ by hypothesis, it follows that $\mu(Q) < \mu(F_i/F_{i-1})$.

□

Upon combining *Lemma 7.5* and *Lemma 7.6* we obtained our desired result.

7.7. Proposition.
If E is a holomorphic bundle which is not semi–stable, then E contains a uniquely determined flag

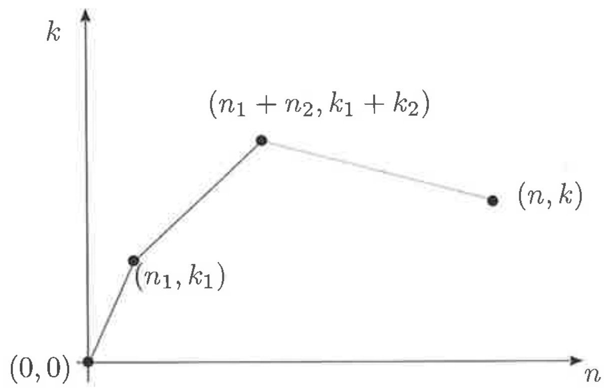
$$0 = E_0 < E_1 < \cdots < E_r = E \tag{7.7.1}$$

with $D_i := E_i/E_{i-1}$ semi–stable and $\mu(D_1) > \mu(D_2) > \cdots > \mu(D_r)$. Such a flag is called the *canonical filtration* associated to E.

Note that if E is semi–stable then $r = 1$. Given $rk(D_i) =: n_i$ and $deg(D_i) =: k_i$, let $(\mu_1, \dots, \mu_n) \in \mathbb{Q}^n$ be a vector such that $\mu_1 \geq \mu_2 \geq \dots \geq \mu_n$ with its first n_1 components equal to k_1/n_1, the next n_2 equal to k_2/n_2, etc. We call μ the *type* of E.

Let $\mathcal{C}_\mu$ denote the subspace of $\mathcal{C}$ of holomorphic bundles of type μ. Note that if all the components μ_i of the vector μ equal k/n then clearly $\mathcal{C}_\mu = \mathcal{C}_{ss}$, the subspace of $\mathcal{C}$ of semi-stable holomorphic bundles. Furthermore, as $\mathcal{C}_\mu$ is preserved by the action of $Aut(E)$ it is then a union of orbits.

We partially order the types $\mu \in \mathbb{Q}^n$ in the following standard manner (following §7 [2]). Associate to μ the convex polygon P_μ in $\mathbb{R}^2$ whose vertices are given by $(0,0)$, (n_1, k_1), $(n_1 + n_2, k_1 + k_2)$, ... The convexity of P_μ reflects the monotonicity of the k_i/n_i.



We define a partial ordering $\leq$ on the set of types $\mu \in \mathbb{Q}^n$ by

$$\lambda \geq \mu \quad \text{if and only if} \quad P_\lambda \text{ lies above } P_\mu. \tag{7.7.2}$$

Upon considering P_μ as the graph of a concave function $p_\mu : i \in \mathbb{Z} \mapsto \sum_{j \leq i} \mu_j$ which interpolates linearly between integers, then

$$\lambda \geq \mu \quad \text{if and only if} \quad \sum_{j \leq i} \lambda_j \geq \sum_{j \leq i} \mu_j, \quad j - 1, \dots, n-1$$
$$\text{and} \quad \sum_{i=1}^{n} \lambda_i = \sum_{i=1}^{n} \mu_i = k.$$

Let $\mathcal{A}_\mu$ denote the stratification of $\mathcal{A}$ induced by the stratification $\mathcal{C}_\mu$ of $\mathcal{C}$ under the earlier mentioned identification $\mathcal{A} \to \mathcal{C}$. Denote by $\mathcal{N}_\mu$ the space of Yang–Mills connections ∇ on E whose curvature if of *type* μ, that is $*F_\nabla = -2\pi i\, diag(\mu_1, \dots, \mu_n)$. We shall proceed to show that the stratification $\mathcal{A} = \bigcup_\mu \mathcal{A}_\mu$ is a Morse stratification with respect to the critical manifolds $\mathcal{N}_\mu$ by satisfying the conditions of *Proposition 4.20.*

Let $\mathcal{N}_s$ [36] denote the *irreducible* [37] Yang–Mills connections on E with $*F_\nabla$ having entries $-2\pi i k/n$. $\nabla \in \mathcal{N}_s$ induces the (absolute) minimum $4\pi^2 k^2/n$ for YM. This is because if ∇'

[36] In the notation of [2], $\mathcal{N}_s$ denotes the set of connections giving the minimum $4\pi^2 k^2/n$ for YM arising from irreducible projective unitary representations of $\pi_1(X)$. By the last paragraph of the preceding chapter we have that such representations induce connections ∇ with $*F_\nabla$ a diagonal matrix with entries $-2\pi i k/n$.

[37] A connection ∇ is said to be *irreducible* if its (central) curvature $*F_\nabla$ a diagonal block matrix.

was another Yang–Mills connection on E, we can decompose E as $E_1 \bigoplus \cdots \bigoplus E_r$ upon which $*F_{\nabla'}$ is a diagonal matrix with entries $-2\pi i k_j/n_j$ where $n_j := rk(E_j)$ and $k_j := deg(E_j)$, thus $YM(\nabla') = 4\pi^2 \sum_{j=1}^{r} k_j^2/n_j$. Now $4\pi^2 k^2/n \leq YM(\nabla')$ by applying the *Cauchy-Schwarz* inequality $(u \cdot v)^2 \leq \|u\|^2 \|v\|^2$ to the vectors $u := (k_1/\sqrt{n_1}, \ldots, k_r/\sqrt{n_r})$ and $v := (\sqrt{n_1}, \ldots, \sqrt{n_r})$; moreover, the Cauchy–Schwarz inequality is an equality if and only if u and v are proportional, hence $k_j/n_j = k/n$ in this case. By Donaldson's *Theorem 6.2*, the existence of such $\nabla \in \mathcal{N}_s$ on a holomorphic bundle E implies E is a direct sum of stable irreducible sub–bundles. This implies connections $\nabla \in \mathcal{N}_\mu$ are direct sums of of connections in $\mathcal{N}_s$ on sub–bundles, thus $\mathcal{N}_\mu \subset \mathcal{A}_\mu$. This establishes condition (iv) of *Proposition 4.20.*

As we have assumed the normalisation $\int_X vol_X = 1$ we have for $\nabla \in \mathcal{N}_\mu$

$$YM(\nabla) = 4\pi^2 \phi(diag(\mu_1, \ldots, \mu_n)) = 4\pi^2 \sum_{i=1}^{n} \mu_i^2$$

where $\phi(x) := tr(x^* x)$ is the convex invariant integrand of the Yang–Mills functional on $\mathfrak{u}(n)$; we denote this evaluation by $\phi(\mu)$. By defining

$$YM(E) := \inf_{\nabla} YM(\nabla)$$

where ∇ runs over the metric connections in E, then by induction on Donaldson's result (*Theorem 6.2*) we have that $YM(E) = \phi(\mu)$ for stable bundles E. If E is an arbitrary holomorphic bundle with filtration (7.7.1) then

$$YM(E) \leq YM(\bigoplus_j D_j).$$

As any semi–stable bundle has a filtration with stable quotients then

$$YM(E) = \phi(\mu) \quad \text{for} \quad E \in \mathcal{C}_\mu.$$

From this result it follows that $\mathcal{C}_\lambda \subset \overline{\mathcal{C}_\mu}$ implies $\phi(\lambda) \leq \phi(\mu)$ since $YM(\mathcal{C}_\lambda) \leq YM(\mathcal{C}_\mu)$ (for instance, $YM(\mathcal{C}_\lambda) = 4\pi^2 \sum_{j=1}^{n} \lambda_j^2$). As the components of the types are increasing then $\phi(\lambda) \leq \phi(\mu)$ implies $\lambda \leq \mu$. It follows therefore that

$$\overline{\mathcal{C}_\mu} \subset \bigcup_{\lambda \geq \mu} \mathcal{C}_\lambda.$$

The associated strata $\{\mathcal{A}_\lambda\}$ for $\mathcal{A}$ and the non–degenerate critical manifolds $\mathcal{N}_\lambda$ of YM thus satisfy the Morse stratification conditions (i), (ii) and (iv) of *Proposition 4.20.* To prove condition (v) for Morse stratification we require the following results.

If E is a smooth complex vector bundle of rank n over X, the elements $\sigma \in Aut(E)$ are given locally (with respect to a trivialising cover $\{U_\alpha\}$ for E) by smooth maps $\lambda_\alpha : U_\alpha \to GL(n, \mathbb{C})$. In this way $Aut(E)$ acts on $\mathcal{C}$ as follows: if $\mathcal{E} \in \mathcal{C}$ has transition functions $g_{\alpha\beta}$ with respect to the trivialising cover $\{U_\alpha\}$ then $\sigma \cdot \mathcal{E}$ is defined to be the bundle with transition functions $\lambda_\alpha^{-1} g_{\alpha\beta} \lambda_\beta$. So the orbits of this action are the isomorphism classes of holomorphic structures on E. The orbit in $\mathcal{C}$

corresponding to a holomorphic structure $\mathcal{E}$ on E can be identified with the $\mathcal{G}^{\mathbb{C}}$–orbit $O(\mathcal{E})$ of unitary connections on E. As $TO(\mathcal{E}) \cong \{ \overline{\partial}_{End(\mathcal{E})}\beta \mid \beta \in \mathcal{E}^0(End(\mathcal{E})) \}$ and $T\mathcal{C}_{|O(\mathcal{E})} \cong \mathcal{E}^{(0,1)}(End(\mathcal{E}))$ (under the identification $\mathcal{A} \to \mathcal{C}$), then the normal bundle to the orbit $O(\mathcal{E})$ is given by

$$\begin{aligned} \nu O(\mathcal{E}) &\cong \frac{T\mathcal{C}_{|O(\mathcal{E})}}{TO(\mathcal{E})} \\ &\cong \frac{\mathcal{E}^{(0,1)}(End(\mathcal{E}))}{im(\overline{\partial}_{End(\mathcal{E})})} \\ &\cong H^1(X, \mathcal{O}(End(\mathcal{E}))) \end{aligned}$$

where the last isomorphism follows from the *Dolbeault isomorphism theorem.*[38] This Čech cohomology group is finite dimensional (Appendix [13]), so the orbit in $\mathcal{C}$ corresponding to a holomorphic structure $\mathcal{E}$ on E is, locally, a manifold of finite codimension in $\mathcal{C}$.

As the filtration (7.7.1) is canonical, the subspaces $\mathcal{C}(E)_\mu$ are preserved by the action of $Aut(E)$, thus $\mathcal{C}_\mu$ is a union of orbits and so its conormal (the dual of its normal bundle) should be a quotient of
$H^1(X, \mathcal{O}(End(E)))$. Letting $End'(E)$ be the bundle of endomorphisms of E which preserve its filtration, and $End''(E)$ the bundle arising from the exact sequence

$$0 \longrightarrow End'(E) \longrightarrow End(E) \longrightarrow End''(E) \longrightarrow 0$$

the associated long–exact cohomology sequence implies $H^1(X, \mathcal{O}(End''(E)))$ is a quotient of $H^1(X, \mathcal{O}(End(E)))$ and so is considered the conormal of $\mathcal{C}_\mu$. The fact that $H^0(X, \mathcal{O}(End''(E))) = \mathcal{O}(End''(E))(X) = 0$ follows by applying *Corollary 7.9* below to the filtration (7.7.1).

7.8. Lemma.
Let E is a semi–stable bundle over X and D a holomorphic bundle over X with $\mu(E) > \mu(D)$ and $f : E \to D$ a non–zero homomorphism. If D_1 is the sub–bundle of D *generated by the image of* f then $\mu(D_1) > \mu(D)$.
Proof:
As explained in the previous section f has the following canonical factorization

$$\begin{array}{ccccccccc} 0 & \longrightarrow & E_1 & \longrightarrow & E & \longrightarrow & E_2 & \longrightarrow & 0 \\ & & & & \downarrow f & & \downarrow g & & \\ 0 & \longleftarrow & D_2 & \longleftarrow & D & \longleftarrow & D_1 & \longleftarrow & 0 \end{array}$$

with exact rows, $rk(E_2) = rk(D_1)$, $det(g) \not\equiv 0$, and $deg(E_2) \leq deg(D_1)$, where D_1 is referred to as the sub–bundle of D *generated by the image of* f.

As $\mu(E_2) \geq \mu(E)$ by the semi–stability of E, if $\mu(E) > \mu(D)$ then $\mu(D_1) \geq \mu(E_2) \geq \mu(E) > \mu(D)$.

□

[38] Given T a complex space, a *holomorphic family* of vector bundles on X parametrised by T is a holomorphic vector bundle $\mathcal{F}$ on $T \times X$, or the collection $\{F_t\}_{t \in T}$ where F_t is the pullback of $\mathcal{F}$ under the map $x \mapsto (t, x)$. By *Lemma 2.1 (ii)* [27], given $\mathcal{E}$ a holomorphic bundle over a compact Riemann surface, there exists a family of holomorphic vector bundles $\{\mathcal{E}_m\}_{m \in M}$ parametrised by a complex manifold M such that $\mathcal{E}_{m_o} = \mathcal{E}$ for some $m_o \in M$ and such that the *infinitesimal deformation map* $T_m M \to H^1(X, \mathcal{O}(End(\mathcal{E}_m)))$ is an isomorphism at m_o. Thus infinitesimal deformations of a holomorphic bundle $\mathcal{E}$ are classified by the elements of $H^1(X, \mathcal{O}(End(\mathcal{E})))$.

7.9. Corollary.
If E and D are semi–stable bundles over X with $\mu(E) > \mu(D)$ then every homomorphism $E \to D$ is zero.

Now $End'(E)$ is a direct sum of the $Hom(D_i, D_j)$ for $i \geq j$ arising from the filtration (7.7.1) (cited from (7.13) [2]), and so

$$End''(E) \cong \bigoplus_{i<j} Hom(D_i, D_j).$$

By Donaldson's result *Theorem 6.2* we have that for a Yang–Mills connection $\nabla \in \mathcal{N}_\mu$ on E of type μ we have $E \cong \bigoplus_i D_i$ where the D_i are semi–stable with $\mu(D_i)$ the components of the vector μ. By *Corollary 7.9* above we have that

$$ad^-(E) \cong \bigoplus_{i<j} Hom(D_i, D_j)$$

and so $ad^-(E) \cong End''(E)$. Thus for our Yang–Mills connection $\nabla \in \mathcal{N}_\mu$ we have

$$\begin{aligned} index(\nabla) &= 2dim_{\mathbb{C}} H^1(X, \mathcal{O}(End''(E))) \\ &= codim_{\mathbb{C}} \mathcal{C}_\mu \\ &= codim_{\mathbb{C}} \mathcal{A}_\mu \end{aligned}$$

where $\mathcal{A}_\mu$ is the stratum containing ∇. The codimension $d_\mu := codim_{\mathbb{C}} \mathcal{C}_\mu$ can be calculated explicitly by applying the Riemann–Roch theorem to the bundle $\bigoplus_{i<j} Hom(D_i, D_j)$: as $deg(D_i^* \bigotimes D_j) = -k_i n_j + k_j n_i$ and $H^0(X, \mathcal{O}(End''(E))) = 0$ then

$$\begin{aligned} d_\mu &= \sum_{i>j} \{(n_i k_j - n_j k_i) + n_i n_j (g-1)\} \\ &= \sum_{\mu_i > \mu_j} \{\mu_i - \mu_j + (g-1)\} \end{aligned} \tag{7.9.1}$$

where g is the genus of X.

Thus we have shown that condition (v) of *Proposition 4.20* holds for $Y_\lambda = \mathcal{N}_\lambda$ and $X_\lambda = \mathcal{A}_\lambda$. One also notes that the finiteness property (F2) in §3 holds for the indexing set of types $\{\lambda\}$ by the above explicit formula for d_λ.

Finally, the gradient field of YM is given by

$$grad_\nabla YM = - * d_\nabla * F_\nabla.$$

As the tangent space to the $\mathcal{G}$–orbit at ∇ consists of vectors $d_\nabla \alpha$ for $\alpha \in \mathcal{E}^0(X, ad(E))$ and the tangent space to the $\mathcal{G}^{\mathbb{C}}$–orbit at ∇ consists of $\overline{\partial}_\nabla \beta$ for $\beta \in \mathcal{E}^0(X, ad(E^{\mathbb{C}}))$ where $E^{\mathbb{C}}$ denotes the complexification of E, upon identifying $\mathcal{E}^1(X, ad(E))$ with $\mathcal{E}^{(0,1)}(X, ad(E^{\mathbb{C}}))$ on which $* = i$ we have that the tangent space to the $\mathcal{G}^{\mathbb{C}}$–orbit at ∇ consists of vectors $d_\nabla \alpha_1 + *d_\nabla \alpha_2$ for $\alpha_1, \alpha_2 \in$

$\mathcal{E}^0(X, ad(E))$. In particular, from the above formula we see that $grad_\nabla YM$ is tangential to the $\mathcal{G}^{\mathbb{C}}$–orbits through ∇ and thus also tangential to the stratum $\mathcal{A}_\lambda$ containing ∇; this corresponds to property (iii) of *Proposition 4.20.*

We have thus shown that the stratification $\mathcal{A}_\lambda$ of $\mathcal{A}$ is a Morse stratification relative to the Yang–Mills functional YM according to *Proposition 4.20.* Moreover, we have

7.10. Theorem.
The stratification $\mathcal{C} = \bigcup_\lambda \mathcal{C}_\lambda$ is $\mathcal{G}$–equivariantly perfect in the sense that

$$P_t^{\mathcal{G}}(\mathcal{C}) = \sum_\lambda t^{2d_\lambda} P_t^{\mathcal{G}}(\mathcal{C}_\lambda), \tag{7.10.1}$$

where $d_\lambda := codim_{\mathbb{C}} \mathcal{C}_\lambda$.

In order to prove this result that our stratification is perfect we require the following results. For the sake of completeness we shall repeat some definitions from the section on Morse theory.

Recall that a *non-degenerate critical manifold* Y of a smooth function f on a compact smooth manifold X is a connected submanifold $Y \subset X$ such that both $d_y f_{|y \in Y} = 0$ and the Hessian $H_Y f_{|\nu(Y)}$ is non–degenerate, where $\nu(Y)$ is the normal bundle of Y. The function f is called *non-degenerate* if its critical set is the union of non–degenerate critical manifolds. In such a case, given $\nu^-(Y)$ the *negative bundle* associated to Y, the *index* of Y as a critical manifold of f is defined and denoted by$\lambda_Y f := rk(\nu^-(Y))$. Given the $\lambda_Y f$–disc bundle $\nu_\varepsilon^-(Y)$ over a critical manifold Y of f, which is the exponential image of $\nu^-(Y)$ in X with $f \geq f(Y) - \varepsilon$, we have the following commutative diagram.

$$\begin{array}{ccccc} H_\bullet(\nu_\varepsilon^-(Y)) & \longrightarrow & \widetilde{H}_\bullet(\nu_\varepsilon^-(Y), \partial\nu_\varepsilon^-(Y)) & \xrightarrow{\partial} & \widetilde{H}_{\bullet-1}(\partial\nu_\varepsilon^-(Y)) \\ & & \pi^* \uparrow & & \downarrow \\ & & H_{\bullet-\lambda_Y f}(Y) & \xrightarrow{\alpha} & \widetilde{H}_{\bullet-1}(X_{f(Y)-\varepsilon}) \end{array} \tag{7.10.2}$$

For a given field $\mathbb{F}$ we say that Y is $\mathbb{F}$-*completable* if the map α in the preceding figure is zero. Here π is the projection of the disc–bundle $\nu_\varepsilon^-(Y)$ and $\widetilde{H}$ the reduced homology modules over $\mathbb{F}$. The map π^* corresponds to the Thom isomorphism.

7.11. Definition.
Given f a smooth non–degenerate function on a compact smooth manifold X, a critical manifold Y of f is called *self–completing* if given a class $s \in H_{\bullet-\lambda_Y f}(Y)$ goes to zero under the dashed arrow in the commutative diagram (7.10.2) provided $\pi^* s$ is in the image of $H_\bullet(\nu_\varepsilon^-(Y))$.

The following result is cited from Proposition 1.9 §1 [2].

7.12. Proposition.
Given G a Lie group, if the equivariant Euler class of the normal bundle to Y is not a zero–divisor in $H_G^\bullet(Y, \mathbb{F})$ for $\mathbb{F}$ any field, then Y is equivariantly self–completing for $\mathbb{F}$.

If all critical manifolds satisfy the hypothesis of Proposition 7.12 then f will be equivariantly perfect over $\mathbb{F}$, so that the equivariant Morse and Poincaré series coincide.

Now given G a compact connected Lie group without torsion in its cohomology, and given T a maximal torus of G, the fibration

$$G/T \longrightarrow BT \longrightarrow BG$$

behaves like a product for integral cohomology and all the spaces involved have no torsion. For any G–space X the induced fibration

$$G/T \longrightarrow X_T \longrightarrow X_G$$

is multiplicative for integral cohomology

$$H^\bullet(X_T) \cong H^\bullet(X_G) \bigotimes H^\bullet(G/T),$$

so that $H^\bullet_G(X)$ is a direct summand of $H^\bullet_T(X)$; or equivalently, that for all primes p, the map

$$H^\bullet_G(X, \mathbb{Z}_p) \longrightarrow H^\bullet_T(X, \mathbb{Z}_p)$$

is injective.

If $T = T_o \times T_1$ a product of two subtori with T_o acting trivially on the connected T–space X, then $X_T = BT_o \times X_{T_1}$ and so

$$H^\bullet_T(X) \cong H^\bullet(BT_o) \bigotimes H^\bullet_{T_1}(X).$$

From these results, by restricting from G to a maximal torus $T \supset T_o$, we have

7.13. Proposition.
Let X be a connected G–space on which some subtorus T_o acts trivially and let Y be a G–vector bundle on X. Assume that the representation of T_o on the fibre of Y is primitive and let $H^\bullet(G)$ have no torsion. Then the multiplication by the top Chern class $c_n(Y_G)$ on $H^\bullet_G(X, \mathbb{Z}_p)$ is injective for all primes p.

We shall use both Propositions 7.12 and 7.13 to prove Theorem 7.10. To this end we require the following result and concepts.

7.14. Proposition.
Given a smooth complex vector bundle E of rank n and degree k over a compact Riemann surface X and given D_i the quotients associated to E via its canonical filtration (7.7.1) whose degrees are the elements of a type vector λ, we have the following identity

$$H^\bullet_{\mathcal{G}}(\mathcal{C}(E)_\lambda) = \bigotimes_i H^\bullet_{\mathcal{G}}(\mathcal{C}(D_i)_{ss}).$$

Proof:
Let $\mathcal{F}_\mu$ denote the space of all smooth filtrations of E ot type μ; that is, the collection of all sections of the fibre bundle over X with fibre the manifold $GL(n, \mathbb{C})/B_\mu$ where B_μ is the parabolic subgroup

preserving a fixed flag of subspaces of $\mathbb{C}^n$ of dimensions $n_1, n_1+n_2, \ldots$. Since the filtration (7.7.1) is canonical we have a map

$$\mathcal{C}_\mu \longrightarrow \mathcal{F}_\mu.$$

Using a fixed base–point of $\mathcal{F}_\mu$ corresponding to a certain smooth filtration E_μ of E, the fibre of this map over this point is the subspace $\mathcal{B}_\mu \subset \mathcal{C}_\mu$ of complex structures compatible with the given filtration. If $Aut(E_\mu)$ is the group of smooth automorphisms of E preserving the filtration then $Aut(E_\mu)$ acts on $\mathcal{B}_\mu$, and $\mathcal{F}_\mu$ is the homogeneous space $Aut(E)/Aut(E_\mu)$ and $\mathcal{C}_\mu$ can be identified with the associated bundle. Thus the equivariant cohomology of the pairs

$$(Aut(E), \mathcal{C}_\mu) \quad \text{and} \quad (Aut(E_\mu), \mathcal{B}_\mu)$$

are equivalent [39]

Now, upon choosing splittings of the filtration E_μ so that we get a direct sum decomposition E^o_μ of E

$$\begin{aligned} E &= D_1 \bigoplus \cdots \bigoplus D_r \\ \text{with} \quad E_i &= D_1 \bigoplus \cdots \bigoplus D_i, \end{aligned}$$

and let $Aut(E^o_\mu)$ and $\mathcal{B}^o_\mu$ be the automorphisms and complex structures (in $\mathcal{B}_\mu$), respectively, compatible with this decomposition. We then have

$$\begin{aligned} Aut(E^o_\mu) &\cong \prod_{i=1}^{r} Aut(D_i), \\ \mathcal{B}^o_\mu &\cong \prod_{i=1}^{r} \mathcal{C}_{ss}(D_i). \end{aligned} \tag{7.14.1}$$

Also, as the homomorphism

$$Aut(E_\mu) \longrightarrow Aut(E^o_\mu)$$

is a homotopy equivalence, and the fibration

$$\mathcal{B}_\mu \longrightarrow \mathcal{B}^o_\mu$$

has a vector space as fibre and is compatible with the group actions, then it follows that the pairs

$$(Aut(E_\mu), \mathcal{B}_\mu) \quad \text{and} \quad (Aut(E^o_\mu), \mathcal{B}^o_\mu)$$

have equivalent equivariant cohomology. Using this and (7.14.1) we have the required result. □

[39] This follows from §13 [2] which states that for G a topological group, $K \triangleleft G$, and X a G–space upon which K acts freely, then $H^\bullet_G(X) \cong H^\bullet_{G/K}(X/K)$.

Proof of Theorem 7.10:
Given N_μ the conormal bundle to $\mathcal{C}_\mu$ in $\mathcal{C}$ and N_μ^o the restriction of N_μ to $\mathcal{B}_\mu^o$, then by a similar argument to the previous result we can replace the triple $(Aut(E), \mathcal{C}_\mu, N_\mu)$, in the equivariant cohomology sense, by the triple $(Aut(E_\mu^o), \mathcal{B}_\mu^o, N_\mu^o)$. From (7.14.1) $Aut(E_\mu^o)$ is observed to contain the r–dimensional torus T^r which acts trivially on $\mathcal{B}_\mu^o$.

We now use Propositions 7.12 and 7.14 to show that the representation of T^r on the fibre of N_μ is primitive.

At a point of $\mathcal{B}_\mu^o$ the bundle E is a holomorphic direct sum of the D_i and so the bundle of endomorphisms preserving the filtration

$$End'(E) \cong \bigoplus_{i \geq j} Hom(D_i, D_j),$$

and hence

$$End''(E) \cong \bigoplus_{i<j} Hom(D_i, D_j). \tag{7.14.2}$$

On $Hom(D_i, D_j)$ the element $(t_1, \ldots, t_r) \in T^r$ acts by $t_i^{-1} t_j$ and so it acts by the same character on $H^1(X, \mathcal{O}(Hom(D_i, D_j)))$. As the fibre of N_μ is $H^1(X, \mathcal{O}(End''(E)))$ it follows from (7.14.2) that the representation of T^r on N_μ is primitive.

□

In conclusion, the $\mathcal{A}_\mu$ play the role of the Morse strata for the Yang–Mills functional YM [40]. This is in the sense that our strata $\mathcal{A}_\mu$ satisfy all the properties of Proposition 4.20 relative to YM, which in "good cases" (that is, for good properties of the trajectories $x(t)$ of $-gradYM$ as $t \to \infty$) characterize the Morse strata. Although we have shown that the stratification of $\mathcal{A}$ by the $\mathcal{A}_\mu$ is equivariantly perfect, we have not proved that the Yang–Mills functional is an equivariantly perfect Morse function. [2] suggests that this could follow from sufficiently good properties about the Yang–Mills flow $gradYM$.

Notice that if all the components of a type vector λ are all equal (to k/n), then $\mathcal{C}_\lambda = \mathcal{C}_{ss}$. Thus knowledge of $P_t^{\mathcal{G}}(\mathcal{C})$ from our results in §4 leads to formulae for $P_t^{\mathcal{G}}(\mathcal{C}_{ss})$ since the equivariant cohomology of all the "un–stable" strata $\mathcal{C}_\lambda$ can be calculated via *Proposition 7.14*. We have shown in §4 that $B\mathcal{G}$ is torsion free in its equivariant cohomology, and so it follows that the equivariant cohomology of the strata $\mathcal{C}_\lambda$ also have no torsion. When $(n, k) = 1$ we have $\mathcal{C}_{ss} = \mathcal{C}_s$, the subspace of $\mathcal{C}$ of stable bundles, and so we can deduce results about the torsion of the moduli space $N(n, k) := \mathcal{C}/Aut(E)$. This is the content of the next section.

[40] This is also the case if YM is replaced with any other functional defined by a strongly convex invariant function on the Lie algebra $\mathfrak{u}(n)$ of $U(n)$

8 The moduli space of semi–stable bundles.

In the language of *varieties* and *schemes* given X a non–singular projective algebraic curve over $\mathbb{C}$ (in our case a Riemann surface) and E a holomorphic bundle over X of rank n and degree k we shall consider the *moduli scheme*

$$N(n,k) := \mathcal{C}_{ss}(E)/Aut(E)$$

the quotient of the space of semi–stable bundles over X of rank n and degree k by the action of the automorphism group $Aut(E)$.

We shall further restrict to the coprime case $(n,k) = 1$ so that these moduli spaces do not exhibit non–Hausdorff phenomena. In this case $\mathcal{C}_{ss} = \mathcal{C}_s$, the subspace of $\mathcal{C}$ of stable bundles, and $Aut(E)$ acts on $\mathcal{C}_s(E)$ with only the constant central scalars as its isotropy group; moreover, in this case the moduli spaces $N(n,k)$ are (compact) *Kähler manifolds* [41].

These moduli spaces, as in the study of varieties and schemes, parametrise the solution space of the Yang–Mills equations and so merit analysis. Upon calculating the ordinary cohomology of the moduli space $N(n,k)$ and thus obtaining a formula for its Poincaré series we shall show that this space has no torsion.

Let $\overline{\mathcal{G}} := \mathcal{G}/U(1)id$ where $U(1)id$ is the constant central $U(1)$–subgroup of $\mathcal{G}$, and let $\overline{\mathcal{G}}^{\mathbb{C}} := \mathcal{G}^{\mathbb{C}}/\mathbb{C}^{\times} id$ which acts freely on $\mathcal{C}_s$ with quotient $N(n,k)$, hence resulting in the formula

$$H^{\bullet}(N(n,k)) \cong H^{\bullet}_{\overline{\mathcal{G}}}(\mathcal{C}_s). \tag{8.0.3}$$

Here we have used $\overline{\mathcal{G}}$–cohomology since it gives the same result as $\overline{\mathcal{G}}^{\mathbb{C}}$–cohomology (Proposition 2.16 [2]).

The fibration

$$BU(1) \longrightarrow B\mathcal{G} \longrightarrow B\overline{\mathcal{G}} \tag{8.0.4}$$

is trivial in rational cohomology for the following reasons. As the composition $U(1) \longrightarrow \mathcal{G} \longrightarrow U(1)$ arising from taking determinants has degree n, the map

$$H^{\bullet}(B\mathcal{G},\mathbb{Q}) \longrightarrow H^{\bullet}(BU(1),\mathbb{Q})$$

is surjective thus $B\mathcal{G} \sim_{\mathbb{Q}} B\overline{\mathcal{G}} \times BU(1)$ and so

$$P_t(N(n,k)) = (1-t^2)P_t^{\mathcal{G}}(\mathcal{C}_s) \tag{8.0.5}$$

where $P_t^{\mathcal{G}}(\mathcal{C}_s)$ is calculated inductively by the formula

$$P_t(\mathcal{C}) = \sum_{\mu} t^{2d_\mu} P_t(\mathcal{C}_\mu) \tag{8.0.6}$$

[41] A compact complex manifold is called *Kähler* if it admits a Hermitian metric g, given locally by $\sum g_{ij} dz_i \otimes d\overline{z}_j = \sum \phi_i \otimes \overline{\phi}_i$, whose associated $(1,1)$–form ω, patched from the local representatives $i/2 \sum \phi_i \wedge \overline{\phi}_i$, is d–closed.

where $d_\mu := codim_{\mathbb{C}}\mathcal{C}_\mu$ [42] whose first term arises from the semi–stable bundles and all the remaining terms can be calculated inductively from the fact that the equivariant cohomology of the stratum $\mathcal{C}_\mu$ is isomorphic to the tensor product of the equivariant cohomology of the semi–stable strata for the quotients of the canonical filtration for the bundle E.

Consider the map arising from taking determinants of bundles

$$det : N(n,k) \to J_k(X)$$

where J_k is the *Jacobian* of X [43] which parametrizes line–bundles of degree k over X . Let $N_o(n,k)$ denote the fibre of this map, which is the moduli space of stable bundles with fixed determinant. If Γ denotes the group of components of $\mathcal{G}$ (which is $H^1(X,\mathbb{Z}) \cong \mathbb{Z}^{2g}$) and define $\Gamma_n := \Gamma/n\Gamma \cong H^1(X,\mathbb{Z}_n)$ we have that (equation (9.5) [2])

$$N(n,k) = (N_o(n,k) \times J_k)/\Gamma_n.$$

Letting $\mathcal{G}' \subset \mathcal{G}$ correspond to the lattice $n\Gamma \subset \Gamma$ so that $\mathcal{G}/\mathcal{G}' = \Gamma/n\Gamma = \Gamma_n$, then

$$H^\bullet(N_o(n,k) \times J_k) \cong H^\bullet_{\overline{\mathcal{G}}'}(\mathcal{C}_s).$$

Since $\overline{\mathcal{G}}$ and $\overline{\mathcal{G}}'$ give the same equivariant cohomology of $\mathcal{C}_s$ over $\mathbb{Q}$ due to the fact that the same is true for $\mathcal{G}$ and $\mathcal{G}'$, we have by the Künneth formula that

$$H^\bullet(\mathcal{C}_s) \cong H^\bullet(N_o(n,k)) \bigotimes H^\bullet(J_k),$$

or equivalently that

$$P_t(N(n,k)) = P_t(N_o(n,k))(1+t)^{2g} \tag{8.0.7}$$

since $J_k \cong \mathbb{C}^g/\mathbb{Z}^{2g} \cong \prod_{2g} \mathbb{S}^1$.

We now proceed to prove that the integral cohomology of the moduli space $N(n,k)$ has no torsion.

As $\mathcal{C}$ is contractible then the homotopy quotient $\mathcal{C}_\mathcal{G}$ is of the same homotopy type as the classifying space $B\mathcal{G}$. We had found earlier that $B\mathcal{G} \sim Map_E(X, BU(n))$ and that this map space had no torsion, thus the space $\mathcal{C}$ and the strata $\mathcal{C}_\mu$ have no torsion in their $\mathcal{G}$–equivariant cohomology. In particular, the same is true for the semi–stable stratum $\mathcal{C}_{ss}$. We wish to deduce a similar result for the cohomology module $H^\bullet_{\overline{\mathcal{G}}}(\mathcal{C}_s)$. Thus it is sufficient to show that the fibration (8.0.4) is a product and so

$$H^\bullet_{\mathcal{G}}(\mathcal{C}_s) \cong H^\bullet_{\overline{\mathcal{G}}}(\mathcal{C}_s) \bigotimes H^\bullet(BU(1)).$$

Now $BU(1) = K(\mathbb{Z};2)$ and so (8.0.4) has characteristic class in $H^2(B\overline{\mathcal{G}},\mathbb{Z})$ whose vanishing will imply the triviality of the fibration. Therefore we equivalently show that the map

$$H^2(B\mathcal{G},\mathbb{Z}) \longrightarrow H^2(BU(1),\mathbb{Z})$$

[42] The dimension d_μ can be calculated by means of formulae (7.9.1).

[43] See Appendix B for the definition of the Jacobian of a Riemann surface.

associated to the cohomology sequence of the fibration (8.0.4) is surjective. Due to time constraints I have assumed the result of Proposition 2.21 [2] which states exactly this.

Given the commutative diagram of fibrations

$$\begin{array}{ccccc} BU(1) & \longrightarrow & (\mathcal{C}_s)_{\mathcal{G}'} & \longrightarrow & N_o(n,k) \times J_k \\ \| & & \downarrow & & \downarrow \\ BU(1) & \longrightarrow & (\mathcal{C}_s)_{\mathcal{G}} & \longrightarrow & N(n,k) \end{array}$$

where the map $N_o(n,k) \times J_k \to N(n,k)$ is the finite Γ_n–covering. As the bottom row has been shown to be a product, the same must be true for the top row. As we know that the $\mathcal{G}'$–equivariant cohomology of $\mathcal{C}_s$ is torsion free it follows that $N_o(n,k) \times J_k$ has no torsion, hence the same being true for $N_o(n,k)$.To this extent we have proved the following results.

8.1. Theorem.
If $(n,k) = 1$, the moduli spaces $N(n,k)$ and $N_o(n,k)$ of stable bundles have no torsion.

We end with the calculation of the Poincaré series of the moduli space $N(n,k)$ for $(n,k) = (2,1)$ so that we can get a result more concrete than that posed at the end of §4 on a conjectural level. Given equations (7.10.1) and (7.9.1) we have in this case

$$P_t^{\mathcal{G}}(\mathcal{C}) = \sum_{m=0}^{\infty} t^{2(2m+g)} P_t^{\mathcal{G}}(\mathcal{C}_{(m+1,-m)}) + P_t^{\mathcal{G}}(\mathcal{C}_s).$$

The equivariant Poincaré series of the "un–stable" strata $\mathcal{C}_{(m+1,-m)}$ are calculated with *Proposition 7.14* to obtain

$$P_t^{\mathcal{G}}(\mathcal{C}_{(m+1,-m)}) = \left\{ \frac{(1+t)^{2g}}{1-t^2} \right\}^2.$$

Formula (8.0.5) derived above gives

$$P_t^{\mathcal{G}}(\mathcal{C}_s) = \frac{P_t(N(2,1))}{1-t^2}$$

and given the following result derived in §4

$$P_t^{\mathcal{G}}(\mathcal{C}) = P_t(B\mathcal{G}) = \frac{\{(1+t)(1+t^3)\}^{2g}}{(1+t^2)^2(1-t^4)}$$

we obtain

$$\begin{aligned} P_t(N(2,1)) &= \frac{\{(1+t)(1+t^3)\}^{2g}}{(1-t^2)(1-t^4)} - \frac{(1+t)^{4g}}{1-t^2} \sum_{r=0}^{\infty} t^{2(2m+g)} \\ &= \frac{\{(1+t)(1+t^3)\}^{2g}}{(1-t^2)(1-t^4)} - \frac{t^{2g}(1+t)^{4g}}{(1-t^2)(1-t^4)}, \end{aligned}$$

where the infinite summation is reduced by geometric series.

A L^2–connections over Riemann surfaces.

Here we shall give a complete description of the result of *K. Uhlenbeck* [32] used in Donaldson's theorem on stable bundles expounded in §6.

In 1982 Uhlenbeck derived (*Theorem A.8* [32]) a weak–compactness result for connections on bundles. This global theorem rests on the existence of a gauge on the trivial bundle $E[\mathbb{R}^l]$ over the ball $B^2 := \overline{B_1(0)} \subset \mathbb{R}^2$ upon which the *Lorentz-gauge condition* $d^*A = 0$ and the *axial-gauge condition* $(x \cdot A)_{|\mathbb{S}^1} = 0$ hold for a given connection $d + A \in \mathcal{A}_1^2(E)$ with a bound on $\int_{B^2} |F_\nabla| dx$.

Such gauges are often assumed constructable in terms of so called *path–ordered integrals*; however, the convergence of these integrals is too suspect to hold any merit, so we prove the above result analytically by appealing to the *implicit function theorem.*

Let X^n be a compact Riemannian manifold, and $E[\mathbb{R}^l]$ a vector bundle with compact structure group $G \subseteq SO(l)$ whose corresponding Lie algebra is denoted by $\mathfrak{g}$.

As before, denote by $L_k^p(U, E)$ the completion of the space $\Gamma(U, E)$ with respect to the $\|\cdot\|_{p,k}$–norm, and define by $\mathcal{A}_k^p(E)$ the affine space of connections $\nabla \in \mathcal{A}(E)$ of class L_k^p; that is, given $\nabla = d + A$ locally over U, $A \in L_k^p(U, ad(E) \bigotimes T^*X)$. The curvature F_∇ of a connection ∇ extends to a L_k^p section when ∇ is of class L_{k+1}^p.

We define the *smooth gauge group* on E by $\mathcal{G} := \Gamma(X, Aut(E))$. This group acts on $\mathcal{A}(E)$ by

$$s \cdot \nabla(\sigma) := s^{-1}\nabla(s\sigma), \quad s \in \mathcal{G}, \quad \nabla \in \mathcal{A}(E), \quad \sigma \in \Gamma(E).$$

When $\nabla = d + A$ with respect to a trivialising cover of E over X, the above formula expands as

$$s \cdot \nabla := d + s^{-1}ds + s^{-1}As.$$

Also, define as earlier the gauge group acting on $\mathcal{A}_k^p(E)$ by $\mathcal{G}_{k+1}^p := L_{k+1}^p(Aut(E))$. By multiplication theorems on Sobolev spaces [44] this group is a smooth manifold and Lie group for $p(k+1) > 2$.

We shall assume at this point that $X = B^2 := \overline{B_1(0)} \subset \mathbb{R}^2$ endowed with the flat (standard) metric $g_{ij} = \delta_{ij}$, and $E = X \times \mathbb{R}^l$. In this instance $ad(E) = B^2 \times \mathfrak{g}$, so $ad(E) \bigotimes T^*X = B^2 \times (\mathbb{R}^2 \bigotimes \mathfrak{g})$, hence $\mathcal{A}_k^2 = \{\nabla = d + A \mid A \in L_k^2(B^2, \mathbb{R}^2 \bigotimes \mathfrak{g})\}$ and $\mathcal{G}_{k+1}^2 = L_{k+1}^2(B^2, G)$.

For $\Omega \subseteq \mathbb{R}^2$ given the space $C_o^\infty(\Omega)$ of C^∞ functions on Ω with compact support, let $L_{k,o}^2(\Omega)$ denote the closure of $C_o^\infty(\Omega)$ in $L_k^2(\Omega)$. We say that $u \in L_k^2(\Omega)$ *is* 0 *on* $\partial\Omega$ if $u \in L_{k,o}^2(\Omega)$. This space is used to overcome ambiguities associated with making statements such as $u_{|\partial\Omega} = 0$ since u is in fact an equivalence class of measure-almost-everywhere equal functions and such a statement is vacuous since $\partial\Omega$ is a set of measure zero.

Upon defining $\mathcal{A}_{1,\kappa}^2 := \{\nabla \in \mathcal{A}_1^2(E) \mid \int_{B^2} |F_\nabla| dx \le \kappa\}$ the local theorem we aim to prove is the following:

[44] For Sobolev spaces defined on vector bundles over compact n–Riemannian manifolds the multiplication operator $L_{k_1}^{p_1} \bigotimes L_{k_2}^{p_2} \longrightarrow L_k^p$ is defined and continuous if $(k_1 - n/p_1) + ((k_2 - n/p_2) \ge k - n/p$ and $k_1, k_2 \ge k$, $p_1k_1, p_2k_2 < n$, $p_1, p_2 > 1$, $p \ne \infty$. See *Theorem 9.5* [29].

A.1. Theorem.
There exist $\kappa, c > 0$ such that every $\nabla \in \mathcal{A}^2_{1,\kappa}$ is gauge-equivalent to a connection $d + A \in \mathcal{A}^2_1$ where A satisfies:

$$\begin{aligned}
&(a): \quad d^*A = 0 \quad \text{ the } \textit{Lorentz} - \textit{gauge condition};\\
&(b): \quad x \cdot A \in L^2_{1,o}(B^2, \mathfrak{g}) \quad \text{ the } \textit{axial} - \textit{gauge condition};\\
&(c): \quad \|A\|_{1,1} \leq c \int_{B^2} |F_\nabla| dx;\\
&(d): \quad \|A\|_{2,1} \leq c \left(\int_{B^2} |F_\nabla|^2 dx \right)^{1/2}.
\end{aligned}$$

Here the L^p norms are defined with respect to the natural inner–product on the spaces $\bigwedge^q_x X$ and the *Cartan-Killing form* $k(A, B) := tr(AB^*)$ on $\mathfrak{g}$.

Our method of proof of this theorem is as follows. Let $\mathcal{S}$ denote the set of all connections in $\mathcal{A}^2_{1,\kappa}$ which are gauge–equivalent to connections in $\mathcal{A}^2_1$ satisfying conditions $(a) - (d)$ of the above theorem. Upon showing that $\mathcal{A}^2_{1,\kappa}$ is connected and that the set $\mathcal{S}$ is both open and closed in $\mathcal{A}^2_{1,\kappa}$, $\mathcal{S}$ is then equal to the only topological component of $\mathcal{A}^2_{1,\kappa}$, namely $\mathcal{A}^2_{1,\kappa}$ itself.

We begin by establishing the connectivity of $\mathcal{A}^2_{1,\kappa}$ by showing that this space is path–connected.

A.2. Lemma. The space $\mathcal{A}^2_{1,\kappa} \subset \mathcal{A}^2_1$ is connected.
Proof:
Given $\nabla = d + A \in \mathcal{A}^2_{1,\kappa}$, for $t \in [0, 1]$ consider the *dilation at* $0 \in B^2$ *with ratio* t

$$\begin{aligned}
\delta_t : B^2 &\longrightarrow B^2\\
: x &\mapsto tx,
\end{aligned}$$

and define the pull–back of ∇, $\nabla_t := \delta_t^*(\nabla) = d + tA(tx)$.

The L^1–curvature of the path ∇_t is controlled by that of ∇, for

$$\begin{aligned}
F_{\nabla_t} &= t^2 dA(tx) + t^2 A(tx) \wedge A(tx)\\
&= t^2 F_\nabla(tx)\\
\text{hence} \quad \int_X |F_{\nabla_t}| dx &= t \int_{\overline{B_t(0)}} |F_\nabla| dy\\
&\leq t\kappa \leq \kappa \quad \forall t
\end{aligned}$$

where $B_t(0) := \{\ x \in \mathbb{R}^2 \mid |x| < t\ \}$. Therefore the curve ∇_t lies in $\mathcal{A}^2_{1,\kappa}$.

□

A.3. Lemma.
The set $\mathcal{S}$ is closed in $\mathcal{A}^2_{1,\kappa}$.
Proof:
The methodology adopted here is to show that for any sequence $\{\nabla_i\}$ in $\mathcal{S}$, the limiting connection $\nabla := d + \widetilde{A} \in \mathcal{A}^2_{1,\kappa}$ must lie in $\mathcal{S}$.

Each $\nabla_i := d + \widetilde{A}_i \in \mathcal{S}$ is gauge–equivalent to a connection $d + A_i \in \mathcal{A}^2_1$ where the A_i satisfy conditions $(a) - (d)$ of *Theorem A.1*; in particular, $\|A_i\|_{2,1} \leq c\|F_{\nabla_i}\|_2$. As $\nabla_i \xrightarrow{L^2_1} \nabla$, thus $F_{\nabla_i} \xrightarrow{L^2} F_{\nabla}$, the A_i are uniformly bounded in L^2_1. As L^2_1 is a reflexive Banach space there exists a subsequence of $\{A_i\}$ which converges weakly in L^2_1, say to A.

Conditions $(a) - (d)$ of *Theorem A.1* are preserved under weak limits since
(a) $\Lambda \circ d^* \in L^2_1(B^2, \mathbb{R}^2 \bigotimes \mathfrak{g})^* \quad \forall \Lambda \in L^2_1(B^2, \mathbb{R}^2 \bigotimes \mathfrak{g})^*$;
(b) $(x\cdot)$ is a linear functional on $L^2_1(B^2, \mathbb{R}^2 \bigotimes \mathfrak{g})$;
(c) & (d) follow since $\|A_i\|_2 \to \|A\|_2$ and $\|F_{\nabla_i}\|_2 \to \|F_\nabla\|_2$.

To complete the proof we require to show that there exists a gauge transformation between $\nabla = d + \widetilde{A}$ and $d + A$. As $s_i \cdot \nabla_i = d + A_i$ for some $s_i \in \mathcal{G}^2_2$, or equivalently that

$$ds_i = s_i A_i - \widetilde{A}_i s_i \tag{A.3.1}$$

we have that $\|ds_i\|_2 \leq \|A_i\|_2 + \|\widetilde{A}_i\|_2$ since $G \subseteq SO(l)$, hence the ds_i are bounded in L^2 thus the s_i are bounded in L^2_1. As G is compact the s_i are uniformly bounded in L^2_1 hence there is a subsequence $\{s_{i'}\}$ such that $s_{i'} \xrightarrow{L^2_1} s$. The equation (A.3.1) is preserved under weak limits to become $ds = sA - \widetilde{A}s$ since $\Lambda \circ d \in L^2_1(B^2, G)^* \quad \forall \Lambda \in L^2_1(B^2, \mathbb{R}^2 \bigotimes \mathfrak{g})^*$ and upon setting $\Lambda(\widetilde{A}_i s_i) = \Lambda_1(\widetilde{A}_i)\Lambda_2(s_i)$ for $\Lambda_1 \in L^2_1(B^2, \mathbb{R}^2 \bigotimes \mathfrak{g})^*$ and $\Lambda_2 \in L^2_1(B^2, \mathfrak{g})^*$.

As s takes its values in G its L^∞–norm is well defined and bounded, thus

$$\|sA\|_2 \leq \|s\|_\infty \|A\|_2 \quad \text{and} \quad \|\widetilde{A}s\|_2 \leq \|\widetilde{A}\|_2 \|s\|_\infty$$

thus $ds \in L^2$ since

$$\|ds\|_2 \leq \|s\|_\infty (\|A\|_2 + \|\widetilde{A}\|_2).$$

A similar calculations are adopted in showing $\nabla_o ds \in L^2$ where ∇_o is a connection on the bundle $Aut(E) \bigotimes T^*X$, thus concluding that $ds \in L^2_1$. By the elliptic estimate [45]

$$\|s\|_{2,2} \leq c\|ds\|_{2,1}$$

we have that $s \in L^2_2$.

□

We now show that $\mathcal{S}$ is open in $\mathcal{A}^2_{1,\kappa}$, namely to establish the following:

A.4. Lemma.
If $\nabla \in \mathcal{A}^2_{1,\kappa}$ is gauge-equivalent to $d + A$ where A satisfies conditions $(a) - (d)$ in *Theorem A.1*, there exists an open neighbourhood of ∇ in $\mathcal{A}^2_{1,\kappa}$ provided κ is sufficiently small.

We save ourselves some difficulty in the proof of *Lemma A.4* by observing that the "closed" conditions (c) and (d) of *Theorem A.1* are apriori valid estimates on solutions to equations (a) and (b). This is described in the following result.

[45] If L is an elliptic operator of order l on $\Gamma(\xi)$ for some vector bundle ξ over a compact Riemannian manifold, then for each $k \geq 0$ there exists a $c = c(k) > 0$ such that for all $s \in \Gamma(\xi)$ we have $\|s\|_{2,k+l} \leq c(\|Ls\|_{2,k} + \|s\|_{2,k})$. Furthermore, if s is L^2_k–orthogonal to $ker(L)$ we can omit the $\|s\|_{2,k}$ term. See *Theorem 4.1* Appendix, §4 [23].

A.5. Lemma.
Given $A \in L_1^2(B^2, \mathbb{R}^2 \bigotimes \mathfrak{g})$ satisfying conditions (a) and (b) of *Theorem A.1*, there exists a $k > 0$ such that if $\|A\|_2 \leq k$ then $\|A\|_{2,1} \leq c\|F_{d+A}\|_2$ and $\|A\|_{1,1} \leq c\|F_{d+A}\|_1$.
Proof:
As $H^1(B^2) = 0$ the basic elliptic estimate for the 1^{st}–order elliptic operator $d + d^*$ on 1–forms gives

$$\|A\|_{2,1} \leq k'\|dA\|_2 \quad \text{since} \quad d^*A = 0. \tag{A.5.1}$$

From the equation for curvature $F_{d+A} = dA + A \wedge A$ we have

$$\|dA\|_2 \leq \|F_{d+A}\|_2 + \|A \wedge A\|_2. \tag{A.5.2}$$

Moreover,

$$\|A \wedge A\|_2 \leq k''\|A\|_2\|A\|_{2,1}; \tag{A.5.3}$$

this following from the fact that $\|A \wedge A\|_2 \leq c'\|A \wedge A\|_{1,1}$ since $L_1^1 \subseteq L^2$ on 2–dimensional manifolds, and as

$$\begin{aligned}\|A \wedge A\|_{1,1} &\leq c''(\|A \wedge \nabla_o A\|_1 + \|A\|_2^2) \\ &\leq k''(\|A\|_2\|\nabla_o A\|_2 + \|A\|_2^2)\end{aligned}$$

by the Hölder inequality where ∇_o is a connection on $ad(E) \bigotimes T^*X$.

Combining equations (A.5.1), (A.5.2) and (A.5.3) we attain

$$\|A\|_{2,1}(1 - k'k''\|A\|_2) \leq k'\|F_{d+A}\|_2.$$

Upon taking $k = \frac{1}{2k'k''}$ we obtain the required result.

Upon replacing the index "1" for the index "2" in equation (A.5.1) and replacing equation (A.5.3) with the basic Hölder estimate $\|A \wedge A\|_1 \leq \|A\|_2\|A\|_2$, these equations combine with the Sobolev inequality $\|A\|_2 \leq k''\|A\|_{1,1}$ to give $\|A\|_{1,1} \leq c\|F_{d+A}\|_1$.

□

We will also require the following two lemmae for the proof of *Lemma A.4*.

A.6. Lemma.
There exists a linear operator $P : L_1^2(B^2) \longrightarrow L_2^2(B^2)$ such that for $f \in L_1^2(B^2)$ we have $P(f) \in L_{2,o}^2(B^2)$ and $(x \cdot dP(f) - f) \in L_{1,o}^2(B^2)$.
Proof:
Given $f \in L_1^2(B^2)$ let $\widetilde{P}(f)$ be the solution of the inhomogeneous heat equation boundary–value problem on $\mathbb{S}^1$ with the radial coördinate r in B^2 is taken as the "time" variable:

$$\triangle_{\mathbb{S}^1} u - \frac{\partial u}{\partial r} = f, \quad 0 < r \leq 1, \quad u(1) = 0$$

where $\triangle_{\mathbb{S}^1} := (d^*d + dd^*)_{|C^\infty(\mathbb{S}^1)}$ be the usual *Laplacian* on $\mathbb{S}^1$.

The *regularity theorem* [46] for this partial differential equation gives $\widetilde{P}(f) \in L^2_2(B^2\backslash\{0\})$ for $f \in L^2_1(B^2)$. By the nature of the problem $(x \cdot d\widetilde{P}(f) - f) \in L^2_{1,o}(B^2\backslash\{0\})$ and $\widetilde{P}(f) \in L^2_{2,o}(B^2\backslash\{0\})$. Let $P(f) := \widetilde{P}(f)\cdot\phi$ where ϕ is a smooth cut-off function with $\phi(x) = 0$ near $x = 0$ and $\phi(x) = 1$ near $|x| = 1$. Thus $P(f) \in L^2_2(B^2)$ for $f \in L^2_1(B^2)$. Furthermore $P(f) \in L^2_{2,o}(B^2)$ and $(x \cdot dP(f) - f) \in L^2_{1,o}(B^2)$.

□

Assume we have a connection $\nabla \in \mathcal{A}^2_{1,\kappa}$ gauge–equivalent to a connection $d + A \in \mathcal{A}^2_1$ satisfying conditions (a)–(d) of *Theorem A.1*; the *product-connection* d clearly establishes the existence of such a connection. The following lemma will show that in a neighbourhood of $d + A$ there are connections also satisfying the *Lorentz–gauge* condition (a); this will be the foundation of the proof for the "openess" of $\mathcal{S}$. If we start with our intuitive notion of "neighbourhood" connections, namely those of the form $d + (A + \lambda)$ where $\lambda \in L^2_1(B^2, \mathbb{R}^2 \bigotimes \mathfrak{g})$ is small (in the L^2_1 sense), the use of the *implicit function theorem* in the following lemma will show that there exist connections of the form $d + (s^{-1}ds + s^{-1}(A + \lambda)s)$ which satisfy the *Lorentz–gauge* condition (a) for a suitable gauge transformation $s \in \mathcal{G}^2_2$. In order that such a neighbourhood connection also satisfies the *axial–gauge* condition (b), we make the assumption that $x \cdot \lambda$, $x \cdot ds \in L^2_{1,o}(B^2, \mathfrak{g})$. The validity of this assumption rests in the use of the preceding technical lemma which modifies neighbourhood connections $d + (A + \lambda)$ to a form where $x \cdot \lambda \in L^2_{1,o}(B^2, \mathfrak{g})$. In showing that conditions (a) and (b) are satisfied in a neighbourhood of $d + A$, we use the apriori results of *Lemma A.5* to imply satisfaction of the remaining conditions (c) and (d) by this neighbourhood.

To facilitate this methodology we introduce the following spaces

$$\begin{aligned} L^2_{1,\nu} &:= \{\lambda \in L^2_1(B^2, \mathbb{R}^2 \bigotimes \mathfrak{g}) \mid x \cdot \lambda \in L^2_{1,o}(B^2, \mathfrak{g})\}, \\ \mathcal{G}^2_{2,\nu} &:= \{s \in \mathcal{G}^2_2 \mid x \cdot ds \in L^2_{1,o}(B^2, \mathfrak{g})\}, \\ L^{2\perp}_{2,\nu} &:= \{U \in L^2_2(B^2, \mathfrak{g}) \mid \int_{B^2} U\,dx = 0,\ x \cdot dU \in L^2_{1,o}(B^2, \mathfrak{g})\} \\ \text{and}\quad L^{2\perp}_0 &:= \{V \in L^2(B^2, \mathfrak{g}) \mid \int_{B^2} V\,dx = 0\}. \end{aligned}$$

A.7. Lemma.
Given $d + A \in \mathcal{A}^2_{1,\kappa}$ satisfying conditions (a) and (b) of *Theorem A.1* and such that $\|A\|_2 \leq k$ for some $k > 0$, there exists an $\epsilon > 0$ such that for $\lambda \in L^2_{1,\nu}$ with $\|\lambda\|_{2,1} \leq \epsilon$, the non–linear equation

$$d^*(s^{-1}ds + s^{-1}(A + \lambda)s) = 0$$

has a solution $s(\lambda) \in \mathcal{G}^2_{2,\nu}$ depending smoothly on λ.

Proof:
As gauge transformations on E can be expressed as $s = e^U$ for $U \in L^2_2(B^2, \mathfrak{g})$ we consider the smooth map

$$\begin{aligned} \Phi\ &: L^{2\perp}_{2,\nu} \bigotimes L^2_{1,\nu} \longrightarrow L^{2\perp}_0 \\ &: (U, \lambda) \mapsto d^*(e^{-U}de^U + e^{-U}(A + \lambda)e^U). \end{aligned}$$

[46] See *Proposition 2.4* chapter 15 §2 [31].

Suppose $\Phi(U,\lambda) = 0$. Then by the *implicit function theorem* if the linearization on the first variable about (U,λ) is an isomorphism from $L^{2\perp}_{2,\nu}$ to $L^{2\perp}_0$ there is a neighbourhood $\mathcal{N}(\lambda)$ of λ, a neighbourhood $\mathcal{N}(U)$ of U and a smooth function $\Psi : L^2_{1,\nu} \longrightarrow L^{2\perp}_{2,\nu}$ such that $\Psi(\lambda) = U$, $\Psi(\mathcal{N}(\lambda)) \subseteq \mathcal{N}(U)$ and for each $\lambda' \in \mathcal{N}(\lambda)$ the equation $\Phi(U',\lambda') = 0$ is uniquely solved by $U' = \Psi(\lambda') \in \mathcal{N}(U)$.

The linearization of Φ on the first variable about $(U,\lambda) = (0,0)$ is given as $L\Phi(h) := d/dt\,\Phi(th,0)|_{t=0}$ for $h \in L^2_2(B^2,\mathfrak{g})$. As $e^{th} = \sum_{n=0}^{\infty} t^n h^n/n!$ then $e^{-th}de^{th} + e^{-th}Ae^{th} = tdh + A + tAh - thA + O\{t^2\}$, whence $L\Phi(h) = d^*(dh + [A,h])$.

This map is self-adjoint, for $G \subseteq SO(l)$ choose the trace inner–product on $\mathfrak{g}$, thus an inner–product on $L^2(B^2,\mathfrak{g})$ is $(h_1,h_2) := \int_{B^2} tr(h_1h_2^*)\,dx$; the result $(L\Phi(h_1),h_2) = (h_1, L\Phi(h_2))$ follows by application of *Stokes' theorem*.

To show this linearization is invertible, we firstly require to prove that $\|L\Phi(h)\|_2 \geq c\|h\|_{2,2}$ for some $c > 0$ in order to establish that $L\Phi$ is $1-1$.
Observe that $d^*[A,h] = -\partial(A_ih)/\partial x_i + \partial(hA_i)/\partial x_i = -A_i\partial h/\partial x_i + \partial h/\partial x_i A_i$ (since $d^*A = 0$), thus $d^*[A,h] = - * [A,dh]$. So by the Hölder inequality,

$$\begin{aligned}\|L\Phi(h)\|_2 &\geq \|d^*dh\|_2 - \|[A,dh]\|_2\\ &\geq k''\|dh\|_{2,1} - \|dh\|_{2,1}(c_1\|A\|_2 + c_2\|A\|_{2,1})\\ &\geq \|dh\|_{2,1}(k'' - c_1k - c_2c\kappa)\end{aligned}$$

since $\|A\|_2 \leq k$ and hence $\|A\|_{2,1} \leq c\|F_{d+A}\|_2$ by *Lemma A.5.* Taking $k = (c_2c\kappa - k'')/2c_1$ gives the required result.

To show that $L\Phi$ is onto we must deduce that $im(L\Phi) = L^{2\perp}_0$. As $L\Phi$ is linear, $ker(L\Phi) = \{0\}$ whence $ker(L\Phi)^\perp = L^{p\perp}_0$. Also, $ker(L\Phi)^\perp = im(L\Phi^*)$, but $im(L\Phi^*) = im(L\Phi)$ since $L\Phi$ is self-adjoint, thus attaining the result.
Thus the *implicit function theorem* implies a gauge transformation $s(\lambda) := e^{U(\lambda)} \in \mathcal{G}^2_{2,\nu}$ solving our non-linear equation locally (that is, within a neighbourhood of λ, namely $\|\lambda\|_{2,1} \leq \epsilon$), smoothly dependent on λ.

□

Proof of Lemma A.4:
Suppose $\widehat{s} \in \mathcal{G}^2_2$ is the gauge transformation taking $\nabla \in \mathcal{A}^2_{1,\kappa}$ to $d + A$. Our procedure is as follows: we work within a neighbourhood of $d + A$ of elements satisfying conditions $(a) - (d)$ of *Theorem A.1* and then map this neighbourhood back to one about ∇ by $\widehat{s}^{-1}$. Intuitively, neighbourhood elements about $d+A$ will have the form $d+(A+\lambda)$. In order to satisfy $(a) - (d)$ we use *Lemma A.7* to map these elements to those of the form $d + (s^{-1}ds + s^{-1}(A+\lambda)s)$ by a gauge transformation s which satisfies $d^*(s^{-1}ds + s^{-1}(A+\lambda)s) = 0$. In order for *Lemma A.7* to be validly used to imply the existence of such a s, we must require $x\cdot\lambda \in L^2_{1,o}(B^2,\mathfrak{g})$; we thus use *Lemma A.6* to construct such an element. Let $U = U(\lambda) := P(x\cdot\lambda)$ for P as in *Lemma* 2.4. Making the gauge transformation $e^{-U}(d+(A+\lambda))e^U = d + e^{-U}de^U + e^{-U}Ae^U + e^{-U}\lambda e^U =: d + (A+\widetilde{\lambda})$, as $U = P(x\cdot\lambda) \in L^2_{2,o}(B^2,G)$ then $de^U - dU \in L^2_{1,o}(B^2,\mathfrak{g})$ (by use of the identity $e^U = \sum_{i=0}^{\infty} U^i/i!$). As $(x\cdot dU - x\cdot\lambda) \in L^2_{1,o}(B^2,\mathfrak{g})$ (by the results of *Lemma A.6*) then $x\cdot\widetilde{\lambda} \in L^2_{1,o}(B^2,\mathfrak{g})$. Moreover, as $\|U\|_{2,2} \leq \widetilde{c}\|x\cdot\lambda\|_{2,1}$ we can make $\|\widetilde{\lambda}\|_{2,1}$ sufficiently small by doing so with $\|\lambda\|_{2,1}$. *Lemma A.7* can now be used on the neighbourhood

elements of the form $d+(A+\widetilde{\lambda})$. Observe that $\|s^{-1}ds+s^{-1}(A+\widetilde{\lambda})s\|_2 \leq \|ds\|_2+\|A\|_2+\|\widetilde{\lambda}\|_2 \leq k'$ since $\|\widetilde{\lambda}\|_{2,1} \leq \epsilon'$, s takes its values in $G \subseteq SO(l)$, and as A satisfies condition (c) with $\|F_\nabla\|_1 \leq \kappa$. We can thus use the fact that $L^1_1 \subseteq L^2$ and *Lemma A.5* to imply the conditions (c) and (d) for these neighbourhood elements.

□

We have thus shown that the space $\mathcal{S}$ is identically $\mathcal{A}^2_{1,\kappa}$.

We now piece together these results to prove our main theorem.

We now assume X is a compact Riemann surface and and $E[\mathbb{R}^l]$ a vector bundle with compact structure group $G \subseteq SO(l)$. We will use the result *Theorem A.1* in establishing the following result:

A.8. Theorem.
Given $\{\nabla_i\}$ a sequence of connections whose corresponding curvatures F_{∇_i} satisfy $\int_X |F_{\nabla_i}|^2 vol \leq B$ for some $B > 0$, there exists a subsequence $\{\nabla_{i'}\}$ and gauge transformations $s_{i'} \in \mathcal{G}^2_2$ such that $\{s_{i'} \cdot \nabla_{i'}\}$ is weakly convergent in $\mathcal{A}^2_1$. The weak limit ∇ satisfies $\int_X |F_\nabla|^2 vol \leq B$.

We now establish a technical result A.11 involving trivialising covers for the bundle E over X which will aid in the proof of the above main theorem.

Let $\widetilde{G}$ be a fixed neighbourhood of $e \in G$ in the domain of the map exp^{-1} where $exp : \mathfrak{g} \longrightarrow G$. To this extent we require the following 2 results.

A.9. Lemma.
Given G a compact group with equivariant metric, $\exists f_o > 0$ such that if $h, g, \rho \in G$, $|exp^{-1}hg| \leq f_o$ and $|exp^{-1}\rho| < f_o$, then $h\rho g \in \widetilde{G}$ and $|exp^{-1}h\rho g| \leq 2(|exp^{-1}hg| + |exp^{-1}\rho|)$.
Proof:
Define the map $Q : \mathfrak{g} \times \mathfrak{g} \longrightarrow \mathfrak{g}$ by the formula $exp(Q(k,u) = exp(k)exp(u)$.
Q is smooth in both variables in a neighbourhood of 0 in $\mathfrak{g}$. Observe that $Q(0,0) = 0$ and as $exp(Q(k,u))dQ(k,u) = [exp(k)exp(u), exp(k)exp(u)]$ then $|dQ(0,0)| = 1$.
Let $\mathcal{O} := \{k, u \in \overline{B}_{f_o}(0) \subset \mathfrak{g} \mid |dQ(k,u)| \leq 2\}$. As $\mathcal{O}$ is convex, then by the *mean value theorem* we have $|Q(k,u)| \leq 2(|k| + |u|)$ for $|k|, |u| \leq f_o$.
Let $k := exp^{-1}(hg)$ and $u := Ad_g exp^{-1}(\rho)$ then $Q(k,u) = exp^{-1}(gh\, exp^{-1}(Ad_g exp^{-1}(\rho))) = exp^{-1}(hg\rho)$. Thus $|Q(k,u)| \leq 2(|exp^{-1}(hg)| + |Ad_g exp^{-1}(\rho)|) = 2(|exp^{-1}(hg)| + |exp^{-1}(\rho)|)$.

□

A.10. Proposition.
Let $\{U_\alpha\}^l_{\alpha=1}$ be a finite cover of X.
Let $h_{\alpha\beta} : U_\alpha \cap U_\beta \longrightarrow G$ and $g_{\alpha\beta} : U_\alpha \cap U_\beta \longrightarrow G$ be sets of continuous transition functions describing two vector bundles over X.
$\exists f_l$ such that if

$$m := max_{\alpha,\beta \,|\, x \in U_\alpha \cap U_\beta} |exp^{-1}(h_{\alpha\beta}(x) g_{\alpha\beta}(x))| \leq f_l$$

then there exists a refinement $\{V_\alpha\}$ of the cover $\{U_\alpha\}$ and continuous maps $\rho_\alpha : V_\alpha \longrightarrow G$ such that $h_{\alpha\beta} = \rho_\alpha g_{\alpha\beta} \rho_\beta^{-1}$ on $V_\alpha \cap V_\beta$. Moreover, $max_{x \in V_\alpha} |exp^{-1}(\rho_\alpha)| \leq c_l m$.
Proof:

We proceed by induction on l.
For $l = 1$ we have the required result by setting $\rho_1 = e \in G$.
Suppose we have constructed $U_{\alpha,k} \subset U_\alpha$ and $\rho_\alpha : U_{\alpha,k} \longrightarrow G$ satisfying $h_{\alpha\beta} = \rho_\alpha g_{\alpha\beta}\rho_\beta^{-1}$ on $U_{\alpha,k} \cap U_{\beta,k}$ for $1 \leq \alpha, \beta \leq k$; and assume $X \subset \left(\cup_{\alpha\leq k} U_{\alpha,k}\right) \cup \left(\cup_{\alpha>k} U_\alpha\right)$ and $|exp^{-1}\rho_\alpha| \leq c_k m$.
For m sufficiently small we can proceed in induction from $j = k$ to $j = k+1$. We do this as follows.

Define the maps $u_j := exp^{-1}(h_{j\alpha}\rho_\alpha g_{\alpha j}) : U_{\alpha,k} \cap U_j \longrightarrow \mathfrak{g}$ for $\alpha \leq k = j-1$.
If $m \leq f_o/c_k$ then $|exp^{-1}\rho_\alpha(x)| \leq c_k m \leq f_o$ and $|exp^{-1}h_{j\alpha}(x)g_{\alpha j}(x)| \leq m \leq f_o$.
Lemma A.9 implies the existence of the u_j and establishes that $2(1+c_k)m = c_j m$. From the overlap condition on the transition functions we have that the u_j are defined on $U_j \cap \left(\cup_{\alpha\leq k} U_{\alpha,k}\right)$.

Choose a smooth partition of unity $\{\phi_j\}$ subördinate to $\{U_j\}$ which is 0 on $U_j\backslash\left(\cup_{\alpha\leq k} U_{\alpha,k}\right)$.
Make the choice of ϕ_j so that the sets $U_{\alpha,j} = U_{\alpha,k} \cap \{x \mid \phi_j(x) = 1\}^\circ$ cover $X\backslash\overline{\cup_{\alpha>k} U_\alpha}$.
Define $\rho_j = exp(\phi_j u_j)$ on $U_j \cap \left(\cup_{\alpha\leq k} U_{\alpha,k}\right)$ and $\rho_j = 1$ on $U_j\backslash\left(\cup_{\alpha\leq k} U_{\alpha,k}\right)$.
Then $|exp^{-1}\rho_j(x)| \leq |\phi_j(x)u_j(x)| \leq 2(1+c_k)m = c_j m$. This and the iterative equation $c_{k+1} = 2(1+c_k)$ complete the proof.

□

A.11. Corollary.
Let $h_{\alpha\beta},\ g_{\alpha\beta} \in L^2_2(U_\alpha \cap U_\beta, G)$ be transition functions for two vector bundles over X.
Suppose

$$m := max_{\alpha,\beta\,|\,x\in U_\alpha\cap U_\beta}\,|exp^{-1}(h_{\alpha\beta}(x)g_{\alpha\beta}(x))| \leq f_l.$$

Then the ρ_α constructed in *Proposition A.10* are elements of $L^2_2(V_\alpha, G)$.
Furthermore, if $\|h_{\alpha\beta}\|_{2,2|U_\alpha\cap U_\beta}, \|g_{\alpha\beta}\|_{2,2|U_\alpha\cap U_\beta} \leq m'\ \forall \alpha, \beta$, there exists a $k(m')$ such that $\|exp^{-1}\rho_\alpha\|_{2,2|V_\alpha} \leq k(m')$.
Proof:
The maps $\rho_j := exp(\phi_j\, exp^{-1}h_{j\alpha}\rho_\alpha g_{\alpha,j})$ are bounded in $L^2_2(U_j, G)$ by the rules of multiplication and composition of functions in the Sobolev spaces L^2_2.

□

If ∇ is a connection on the bundle E such that $\int_X |F_\nabla|^2 vol \leq B$ for some $B > 0$, in order to apply our local results to the proof of *Theorem A.8* we require to show that this implies $\int_{U_\alpha} |F_\nabla|\, vol \leq \kappa$ over a particular trivialising cover $\{U_\alpha\}$ of E over X. This is done as follows.

A.12. Lemma.
Let ∇ be a connection on the bundle $ad(E)$ such that $\int_X |F_\nabla|^2 vol \leq B$ for some $B > 0$. There is a finite atlas $\{(U_\alpha, \phi_\alpha)\}$ of X with $\phi_\alpha : U_\alpha \longrightarrow (B^2)^\circ$ such that $\int_{B^2} |F_{\widetilde{\nabla}}|^2\, dx \leq \kappa'$ where $\widetilde{\nabla}$ is the pull-back of ∇ restricted to these covers.
Proof:
Choose an atlas $\{(U_\alpha, \widetilde{\phi}_\alpha)\}$ of X with $\widetilde{\phi}_\alpha : U_\alpha \longrightarrow (B^2)^\circ$ which trivialises E, and U_α chosen such that the Riemannian metrics on X compare uniformly to the Euclidean metrics. As $\nabla|_{U_\alpha} = d + A_\alpha$ for A_α a $\mathfrak{g}$-valued 1-form on X, let $\widetilde{A}_\alpha := (\widetilde{\phi}_\alpha^{-1})^* A_\alpha$ and $\widetilde{\nabla} := d + \widetilde{A}_\alpha$. Observe that $(\widetilde{\phi}_\alpha^{-1})^* F_\nabla = (\widetilde{\phi}_\alpha^{-1})^*(dA_\alpha) + (\widetilde{\phi}_\alpha^{-1})^*[A_\alpha, A_\alpha]$ and $F_{\widetilde{\nabla}} = d((\widetilde{\phi}_\alpha^{-1})^* A_\alpha) + [(\widetilde{\phi}_\alpha^{-1})^* A_\alpha, (\widetilde{\phi}_\alpha^{-1})^* A_\alpha] = (\widetilde{\phi}_\alpha^{-1})^*(dA_\alpha)+$

$(\widetilde{\phi}_\alpha^{-1})^*[A_\alpha, A_\alpha] = (\widetilde{\phi}_\alpha^{-1})^* F_\nabla$ (by the properties of pull-backs; namely, that $df^*(\omega) = f^*(d\omega)$ and $f^*(\omega \wedge \gamma) = f^*(\omega) \wedge f^*(\gamma)$).
Let $\chi : x \mapsto \kappa'/B\, tx$ for $t \in [0,1]$, $x \in (B^2)^\circ$ and define $\phi_\alpha := \chi^{-1} \circ \widetilde{\phi}_\alpha$. Then $\phi_\alpha^* := \chi^* \circ (\widetilde{\phi}_\alpha^{-1})^* : \mathcal{E}^p(X) \longrightarrow \mathcal{E}^p((B^2)^\circ)$.
$(\phi_\alpha^{-1})^* F_\nabla = (\kappa'/B)^2 t^2 F_\nabla \circ \widetilde{\phi}_\alpha^{-1}(\kappa'/Btx)$.
$\int_{B^2} |F_{\widetilde{\nabla}}|^2\, dx \leq \int_{B^2} |\phi_\alpha^* F_\nabla|^2\, dx \leq (\kappa'/B)^2 \int_{|x|\leq t} |F_{\widetilde{\nabla}}|^2\, dx \leq \kappa'$.
As X is compact, there is a finite subcover of $\{U_\alpha\}$, and our required coördinate maps are the ϕ_α.

□

A.13. Lemma.
Let $\{\nabla_i\}$ be a sequence of connections in $\mathcal{A}_1^2$ such that $\int_X |F_{\nabla_i}|^2 vol \leq B$.
There exists a trivialising cover $\{(U_\alpha, \psi_\alpha^i)\}$ of E such that
(i): given $\nabla_{i|U_\alpha} = d + A_\alpha^i$ with respect to this trivialisation of E, the A_α^i satisfy conditions $(a) - (d)$ of *Theorem A.1*;
(ii): the transition functions $g_{\alpha\beta}^i$ associated with this trivialisation are uniformly bounded in $L_2^2(U_\alpha \cap U_\beta, G)$;
(iii): for an appropriate subsequence, we have weak convergence $A_\alpha^{i'} \overset{L_1^2(U_\alpha, \mathbb{R}^2 \bigotimes \mathfrak{g})}{\rightharpoonup} A_\alpha$ and $g_{\alpha\beta}^{i'} \overset{L_2^2(U_\alpha \cap U_\beta, G)}{\rightharpoonup} g_{\alpha\beta}$;
(iv): the A_α represent a connection ∇ on E with respect to a trivialisation whose transition functions are $g_{\alpha\beta}$.
Proof:
As $L^2 \subseteq L^1$ then by *Lemma A.12* we have that $\|F_{\nabla_i}\|_{1|U_\alpha} \leq \kappa$ thus by *Theorem A.1* we have that (i) follows. (ii) follows from the computations similar to those in *Lemma A.2*. As $\|A_\alpha^i\|_{2,1|U_\alpha} \leq c(\alpha)|F_{\nabla_i}\|_{2|U_\alpha}$ by *Theorem A.1* we have that $\{A_\alpha^i\}$ are uniformly bounded in L_1^2; with this fact and the result of part (ii) we have weak convergence in (iii) of appropriate subsequences. (iv) follows from the fact that the overlap conditions are preserved under weak limits.

□

We can finally prove our main *Theorem A.8*.

Proof of Theorem A.8:
Let us assume the results of *Lemma A.13*, and renumerate so that $i' = i$.
As $L_2^2(X) \subset C^0(X)$ is a compact embedding, we have $g_{\alpha\beta}^i \longrightarrow g_{\alpha\beta}$ (strongly) in $C^0(U_\alpha \cap U_\beta, G)$. So upon applying *Proposition A.10* and *Corollary A.11* to $g_{\alpha\beta}^i, g_{\alpha\beta}^j = g_{\alpha\beta}$ $\exists j$ such that for $j < i \leq \infty$ we have a refinement $\{V_\alpha\}$ of the cover $\{U_\alpha\}$ with $\rho_\alpha^i \in L_2^2(V_\alpha, G)$ and $g_{\alpha\beta}^i = \rho_\alpha^i g_{\alpha\beta}^j {\rho_\alpha^i}^{-1}$.
Moreover, $\rho_\alpha^i \in L_2^2(V_\alpha, G)$ is bounded and so converges to ρ_α in $C^0(V_\alpha, G)$, which is equivalent to $\rho_\alpha^i \overset{L_2^2(V_\alpha, G)}{\rightharpoonup} \rho_\alpha$.
Define the global gauge transformation $s_i \in \mathcal{G}_2^2$ on U_α by $s_i := \rho_\alpha^i$. Observe that $(s_i^{-1} \circ \nabla_i \circ s_i)_{|V_\alpha} = s_i^{-1} \circ (\nabla_i)_{|V_\alpha} \circ s_i = {\rho_\alpha^i}^{-1} \circ (d + A_\alpha^i) \circ \rho_\alpha^i = d + \{{\rho_\alpha^i}^{-1} d\rho_\alpha^i + {\rho_\alpha^i}^{-1} A_\alpha^i \rho_\alpha^i\}$.
As the A_α^i converge weakly in $L_1^2(V_\alpha, \mathbb{R}^2 \bigotimes \mathfrak{g})$ and the ρ_α^i converge weakly in $L_2^2(V_\alpha, G)$, then by

the rules of multiplication in Sobolev spaces we have that the above connection converges weakly in $\mathcal{U}_1^2$ on V_α.

□

B Principal bundles, fibrations and the Jacobian variety.

Let G be a group and X a set. Recall that a *(left) group action* of G on X is a map

$$T : G \times X \longrightarrow X$$

such that $T(e,x) = x$ and $T(g, T(h,x)) = T(gh, x)$ for all $g, h \in G$ and $x \in X$ where $e \in G$ is the identity element of the group; we usually write $g \cdot x$ for $T(g,x)$. A space X together with a group action is called a *G–space*. As there is a bijective correspondence between left and right G–space structures we need only concern ourselves with one type of group action. Given

$$G_o := \{g \in G \mid g \cdot x = x \ \forall x \in X\} \lhd G,$$

a group action on a space is called *effective* if $G_o = \langle e \rangle$.

Given an action of G on X, the *orbits*

$$O_x := \{g \cdot x \mid g \in G\}$$

corresponding to $x \in X$ partition X. We say an action is *transitive* if given any $x_1, x_2 \in X$ there exists a $g \in G$ such that $x_2 = g \cdot x_1$, which is equivalent to requiring that X consists of a single orbit.

We also define the *stability* (or *isotropy*) *group* of a $x \in X$ to be the subgroup

$$G_x := \{g \in G \mid g \cdot x = x\} \leq G.$$

We call the action on X *free* (respectively, *almost free*) if $G_x = \langle e \rangle$ for all $x \in X$ (respectively, if G_x is discrete for each $x \in X$).

Consider a fiber bundle $\xi[F]$ over a (topological) space with typical fiber F having $\{g_{ij}\}$ as its transition functions relative to a trivialising cover $\{U_i\}$ of X. Suppose we have a(n effective) group action of G on F with corresponding homomorphism $\rho : G \longrightarrow Aut(F)$. Suppose $g_{ij}(x) \in \rho(G)$ for all $x \in U_i \cap U_j \neq \emptyset$. In this case G is the structure group of the bundle ξ which is referred to as a G–bundle in this case; the maps g_{ij} are called *G–transition maps*. Furthermore we call a bundle ξ a *G–bundle* if it is isomorphic to a bundle $(E, p, E/G)$ where E/G is the space of orbits corresponding to the action of G on a space E and $p : E \to E/G$ is the natural quotient map.

A G–space X is called *principal* if the action of G on X is effective with a continuous *translation map*; namely, the map $\tau : \{(x, g \cdot x) \in X \times X \mid \forall g \in G\} \to G$ such that $\tau(x, x') \cdot x = x'$. Let G be a topological group. A *principal G–bundle*, denoted PG, is a G–bundle (E, π, X) with principal G–space E. More formally, it is a fiber bundle whose fibers are affine spaces on which the action of G is free and transitive. One observes that a principal G–bundle is a fiber bundle with fiber G.

For brevity we give the following definition of a principal G–bundle; for the purposes of this thesis we shall restrict our attention to the case where the base space is a manifold and the group a Lie group.

B.1. Definition.
Let X be a smooth manifold and G a Lie group. A (smooth) *principal G-bundle* over X is a fiber bundle $PG := (E, \pi, X)$ whose total space E is a G–space such that

(i): the total space E is a smooth manifold;

(ii): the projection $\pi : E \to X$ is surjective satisfying $\pi(g \cdot p) = \pi(p)$ for all $p \in P$ and $g \in G$;

(iii): the action $(g,p) \mapsto g \cdot p$ is a smooth map $G \times P \to G$ with $(hg) \cdot p = h \cdot (g \cdot p)$ for all $p \in P$ and $g, h \in G$ such that the following "local triviality" condition is satisfied: there exists a trivialising cover $\{(\psi_i, U_i)\}$ of PG over X such that $\pi^{-1}(U_i)$ is diffeomorphic to $U_i \times G$ via ψ_i where $\psi_i(p) = (\pi(p), \phi_i(p))$ such that $\phi_i(g \cdot p) = g\phi_i(p)$.

An important G–bundle is a *universal G–bundle* $p : E \to B$ which is characterised by a contractible total space E. The main *classification theorem* (§8 [25]) for universal principal G–bundles for a topological group G, which is given below for completeness, states that the base spaces of any two such bundles are homotopy equivalent, and we usually write BG for any space in this homotopy class, known as the *classifying space* of the group G.

For the statement of the following theorem, let $\mathcal{P}G(B)$ denote the set of equivalence classes of principal G–bundles over the space B; via pullbacks of bundles this is a contravariant set–valued functor on the homotopy category of topological spaces. We also denote by $[X, Y]$ the collection of homotopy equivalence classes of maps $X \to Y$ between topological spaces X and Y.

B.2. Theorem. (Classification theorem for universal bundles)
Given $p : E \to E/G$ a universal principal G–bundle, the natural transformation

$$\Phi : [\,\cdot\,,\, E/G] \to \mathcal{P}G(\,\cdot\,),$$

obtained by sending the homotopy class of a map $f : B \to E/G$ to the equivalence class of the principal G–bundle f^*E, is a natural isomorphism of functors.

What will follow is an explicit construction of a classifying space for a topological group G which will give a concrete example of a universal principal G–bundle.

Let G be a topological group and set

$$E_n(G) := G^{n+1} \quad \text{and} \quad B_n(G) := G^n$$

with $p_n : G^{n+1} \to G^n$ the projection on the first n coördinates.

Define the the *face-* and *degeneracy-operators* on $E_n(G)$ $0 \le i \le n$ to be the respective maps

$$\begin{aligned} d_i &: E_n(G) \to E_{n-1}(G) \\ &: (g_1, \dots, g_{n+1}) \mapsto \begin{cases} (g_2, \dots, g_{n+1}), & i = 0 \\ (g_1, \dots, g_{i-1}, g_i g_{i+1}, g_{i+2}, \dots, g_{n+1}), & 1 \le i \le n, \end{cases} \end{aligned}$$

and

$$\begin{aligned} s_i &: E_n(G) \to E_{n+1}(G) \\ &: (g_1, \dots, g_{n+1}) \mapsto (g_1, \dots, g_{i-1}, e, g_i, \dots, g_{n+1}) \end{aligned}$$

which satisfy the relations $d_i \circ d_j = d_{j-1} \circ d_i$ if $i < j$, $s_i \circ s_j = s_{j+1} \circ s_i$ if $i \leq j$ and

$$d_i \circ s_j = \begin{cases} s_{j-1} \circ d_i & \text{if } i < j, \\ id & \text{if } i = j \text{ or } i = j+1, \\ s_j \circ d_{i-1} & \text{if } i > j+1. \end{cases}$$

The face– and degeneracy–operators are defined on $B_n(G)$ in a similar way except that $d_n(g_1, \ldots, g_n) = (g_1, \ldots, g_{n-1})$.

Given the convex hull $\triangle_n := \{(t_0, \ldots, t_n) \in \mathbb{R}^{n+1} \mid 0 \leq t_i \leq 1, \sum t_i = 1\}$ and operators

$$\begin{aligned} \delta_i &: \triangle_{n-1} \to \triangle_n \\ &: (t_0, \ldots, t_{n+1}) \mapsto (t_0, \ldots, t_{i-1}, 0, t_i, \ldots, t_{n-1}) \end{aligned}$$

and

$$\begin{aligned} \sigma_i &: \triangle_{n+1} \to \triangle_n \\ &: (t_0, \ldots, t_{n+1}) \mapsto (t_0, \ldots, t_{i-1}, t_i + t_{i+1}, \ldots, t_{n+1}) \end{aligned}$$

we define the equivalence relations

$$(g, \delta_i(u)) \sim (d_i(g), u) \quad \text{and} \quad (g, \sigma_i(v)) \sim (s_i(g), v)$$

for $u \in \triangle_{n-1}$ and $v \in \triangle_{n+1}$. With this we construct the quotient spaces

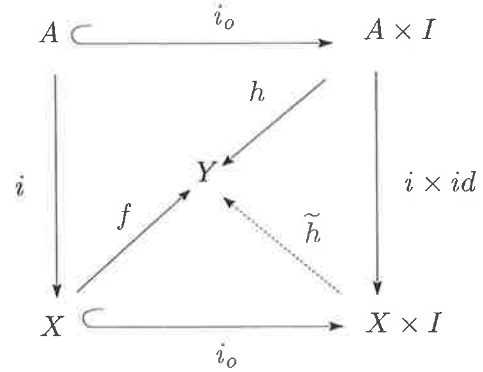
$$E(G) := \coprod_{n \geq 0} (E_n(G) \times \triangle_n) \Big/ \sim \quad \text{and} \quad B(G) := \coprod_{n \geq 0} (B_n(G) \times \triangle_n) \Big/ \sim$$

with topologies given by the union of the constituent quotient topologies. Associated with these spaces is the natural map $p : E(G) \to B(G)$ induced by the maps p_n.

Note that $E(G)$ inherits a free right action by G and that $B(G)$ is the orbit space $E(G)/G$ called the classifying space of G. It can be shown (§5 [25]) that $E(G)$ is a contractible space, and thus for any Lie group G the projection $p : E(G) \to B(G)$ constructed above is a universal principal G–bundle.

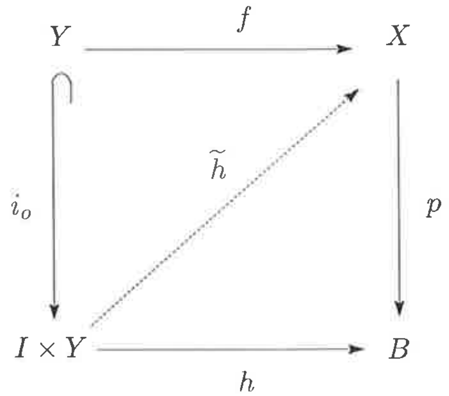
Two important notions in homotopy theory are those of *cofibrations* and *fibrations*; their significance being that all exact sequences that feature in the study of homotopy, homology and cohomology groups can be derived homotopically from the theory of cofiber and fiber sequences. We shall provide the definition of cofibrations and fibrations and follow with some theorems involving fibrations as appear in chapter IV [34]; these results were assumed in the proof of (5.0.1).

B.3. Definition.
A map $i : A \to X$ is known as a *cofibration* if it satisfies the *homotopy extension property* (HEP) which entails the existence of a map $\widetilde{h}$ which makes the following diagram commute if $h \circ i_o = f \circ i$ where $I := [0,1]$ and $i_o(x) = (x, 0)$.



One calls a topological space X *compactly generated* if it is a Hausdorff space and if each $A \subseteq X$ with the property that $A \cap C$ is closed for every compact subset $C \subseteq X$ is itself closed.

B.4. Definition.
Let X and B be topological spaces, Y a compactly generated space and $p : X \to B$ a surjective map. A *homotopy lifting problem* for (p, Y) is symbolised by a commutative diagram



where $I := [0,1]$, $i_o(y) = (0, y)$, and the maps f and h are said to constitute the *data* for the problem in question. The map h is a homotopy of $p \circ f$ and a *solution* to the problem is a map $\widetilde{h} : I \times Y \to X$ such that the the above diagram commutes; thus $\widetilde{h}$ *lifts* the homotopy h of $p \circ f$ to a homotopy of f. The map p is said the have the *homotopy lifting property* (HLP) with respect to Y if every homotopy lifting problem has a solution; a map $p : X \to B$ is said to be a *fibration* if it has the HLP with respect to every space Y.

Given a fibration $p : X \to B$, one calls the space $F_b := p^{-1}(b)$ the *fibre* over the point $b \in B$. It can be shown (chapter IV [34]) that if B is pathwise connected then all the fibres have the same homotopy type; a candidate for this homotopy class is called the *typical* fiber. One often uses the following notation to indicate that the map $p : X \to B$ is a fibration with typical fiber F

$$F \xrightarrow{i} X \xrightarrow{p} B$$

where i is an inclusion.

B.5. Theorem.
If $p : X \to Y$ and $q : Y \to B$ are fibrations, then so too is the composition $q \circ p : X \to B$.

An open cover $\mathcal{O}$ of a space B is called *locally finite* if each $b \in B$ is a point of only finitely many $U \in \mathcal{O}$. The cover $\mathcal{O}$ is furthermore called *numerable* if in addition to it being locally finite there exist continuous maps $\lambda_U : B \to I$ for each $U \in \mathcal{O}$ such that $\lambda_U^{-1}(0,1] = U$; any open cover of a paracompact space is numerable.

B.6. Theorem. (Hurewicz)
Given $p : X \to B$ a continuous map with B a paracompact space. Suppose $\mathcal{O}$ is an open cover of B such that for each $U \in \mathcal{O}$ the maps $p_{|p^{-1}(U)} : p^{-1}(U) \to U$ are fibrations. Then p is a fibration.

B.7. Corollary.
Let $p : X \to B$ be the projection map associated to a fibre bundle whose base–space B is paracompact. Then p is a fibration.

B.8. Theorem.
If $p : X \to B$ is a fibration and $B' \subset B$ with $X' := p^{-1}(B')$, then $p_{|X'} : X' \to B'$ is a fibration.

Two fibrations $p : X \to B$ and $p' : X' \to B$ over the same base space B are said to have the same *fibre homotopy type* if there are maps $\lambda : X \to X', \mu : X' \to X$ such that $p' \circ \lambda = p, p \circ \mu = p'$, and if there are homotopies $\Lambda : I \times X \to X, \Lambda' : I \times X' \to X'$ between pairs $(\mu \circ \lambda, 1_X)$ and $(\lambda \circ \mu, 1_{X'})$ respectively such that $p \circ \Lambda = p \circ \Lambda(0, \cdot)$ and $p' \circ \Lambda' = p' \circ \Lambda'(0, \cdot)$. This means the maps λ and μ are homotopy inverses thus the total spaces of the fibrations have the same homotopy type.

Given spaces F and B, the projection $F \times B \to B$ is designated as the *trivial fibration* over B with typical fiber F. A fibration $p : X \to B$ is said to be *fibre-homotopically trivial* if it has the same fibre homotopy type as such a trivial fibration.

B.9. Theorem.
Let $p : X \to B$ be a fibration with B contractible. Then p is fibre–homotopically trivial.

We end this appendix with the definition of the *Jacobian* of a Riemann surface. We begin with the notion of *lattices*.

B.10. Definition.
Given V a n–dimensional vector space over $\mathbb{R}$, an additive subgroup $\Gamma \subset V$ is called a *lattice* if there exists n $\mathbb{R}$–linearly independent vectors $v_1, \ldots, v_n \in V$ such that

$$\Gamma = \mathbb{Z}v_1 + \cdots + \mathbb{Z}v_n.$$

The following theorem gives conditions for a subgroup to be a lattice.

B.11. Theorem.
A subgroup $\Gamma \subset V$ is a lattice if

(i): Γ is discrete; that is, there exists a neighbourhood U of 0 such that $\Gamma \cap U = \{0\}$;

(ii): Γ is contained in no proper vector subspace of V.

Given X a compact Riemann surface of genus $g \geq 1$ and $\omega_1, \dots, \omega_g$ a basis of the vector space of holomorphic 1–forms $\Omega^1(X)$ on X, define a subgroup

$$Per(\omega_1, \dots, \omega_g) \subset \mathbb{C}^g$$

consisting of all vectors

$$\{ (\int_\alpha \omega_1, \dots, \int_\alpha \omega_g) \in \mathbb{C}^g \mid \alpha \in \pi_1(X) \},$$

or equivalently,

$$\{ (\int_\alpha \omega_1, \dots, \int_\alpha \omega_g) \in \mathbb{C}^g \mid \alpha \in H_1(X) \}.$$

$Per(\omega_1, \dots, \omega_g)$ is a lattice in $\mathbb{C}^g$ (Chapter 2, §21 [13]) considered as a $2g$–dimensional $\mathbb{R}$–vector space, called the *period lattice of X relative to the basis* $\{\omega_1, \dots, \omega_g\}$. Thus, there are $2g$ closed curves $\alpha_1, \dots, \alpha_{2g}$ on X such that the vectors

$$\gamma_\nu := (\int_{\alpha_\nu} \omega_1, \dots, \int_{\alpha_\nu} \omega_g) \in \mathbb{C}^g \quad \nu = 1, \dots, 2g$$

are linearly independent over $\mathbb{R}$, and

$$Per(\omega_1, \dots, \omega_g) = \mathbb{Z}\gamma_1 + \dots + \mathbb{Z}\gamma_{2g}.$$

B.12. Definition.
The *Jacobian* (or *Jacobi variety*) of X is the space

$$J(X) := \mathbb{C}^g / Per(\omega_1, \dots, \omega_g).$$

This is an abelian group which also has the structure of a compact complex manifold (a complex g–dimensional torus). Different choices of bases of $\Omega^1(X)$ lead to isomorphic spaces $J(X)$.

Given the group $Div_k(X)$ of *divisors* on X of degree k, define the group $Pic_k(X) := Div_k(X)/(\text{linear equivalence})$ where two divisors ϑ, ϑ' are *linearly equivalent* if $\vartheta = \vartheta' + (f)$ for some non–vanishing meromorphic function f on X where (f) denotes the divisor associated to f; this is equivalent to requiring that the associated line bundles $[\vartheta], [\vartheta']$ are equal, and so $Pic_k(X)$ may be considered the group of holomorphic line bundles on X of degree k. The map sending a divisor $\vartheta \in Div_k(X)$ to the equivalence class of $J(X)$ with representative $(\int_\alpha \omega_1, \dots, \int_\alpha \omega_g)$ for some $\alpha \in H_1(X)$ with $\partial\alpha = \vartheta$ induces a map $Pic_k(X) \to J(X)$, whose image is defined to be $J_k(X)$.

C Symbol index.

Given here is a list of mathematical symbols used in the thesis with their description and page number of their first mention. Any symbols of single appearance in the thesis are not mentioned if the are in the immediate vicinity of their definition.

Symbol	**Description**	**Page**
$:=$	reads "the expression to the left is defined to be equal the expression on the right"	
$\|\cdot\|$	a vector space norm	
$(\cdot,\cdot)$, $\langle\cdot,\cdot\rangle$	an inner product	
$\mathbb{R}$, $\mathbb{C}$	the field of real– and complex–numbers respectively	
$\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{N}$	the ring of integers,	
the field of rationals	and the set of natural numbers respectively	
$\mathbb{Z}_p$	the field of integers modulo p	
R^+	the subclass of non–negative numbers when $R = \mathbb{R}, \mathbb{Z}, \mathbb{Q}$	
$\mathbb{F}$	an arbitrary field	
$R[x_1, \ldots, x_n]$	the ring of polynomials over the ring R in variables $x_1, \ldots, x_n$	
$\mathbb{S}^n$	the n–sphere $\{x \in \mathbb{R}^{n+1} \,\|\, \|x\| = 1\}$	
$B^n_r(x), B_r(x)$	the ball $\{y \in \mathbb{R}^n \mid \|y - x\| < r\}$, the second case where n is implicit	
$\{x\}$	the set consisting of the single point x	
$\emptyset$	the empty set	
A^c	set complement	
$\overline{A}$	set closure or orbit space if A a G–space	
A^o	set interior	
∂A	set boundary	
$A \backslash B$	the set difference of A with $B \subseteq A$	
$\subsetneq$	proper set inclusion	
$\times$, $\prod$	the direct product	
$\coprod$	disjoint union	
$\approx$	either a homeomorphism or diffeomorphism	
$\cong$	the isomorphism relation	
$\sim$	homotopy equivalence	
$\cup$, $\bigcup$	set union	
$\cap$, $\bigcap$	set intersection	
$\hookrightarrow$	set inclusion	
$dim_{\mathbb{F}} V$	the dimension of the vector space V or manifold over $\mathbb{F}$	
$codim_{\mathbb{F}} V$	the codimension of the vector subspace or submanifold V over $\mathbb{F}$	
$im(f)$	the image space of a linear function f	
$ker(f)$	the kernel of a linear function f	
$rk(f)$, $rk(E)$	the rank of the linear function f, respectively of the bundle E	
V/W	the quotient group, space or bundle; the orbit space if W is a group and V a W–space	

Symbol	Description	Page
$\bigoplus$, $\oplus$	direct sum	
$\bigotimes$, $\otimes$	tensor (or exterior) product	
V^*, E^*	the dual vector space, respectively bundle	
X^n	either a smooth–real or complex manifold of dimension n	
$\partial/\partial x_i$, dx_i	basis vectors, respectively covectors, in a local coördinate chart	
$1, 1_X, id, id_X$	the identity map on a space X	
$c\|d$	for positive integers c and d reads "c divides d", meaning $d(mod\, c) \equiv 0$	
$\chi(X)$	the Euler characteristic of the space X	
$f_{\|U}$	the restriction of a map f to a subset U of its domain	
$E_{\|U}$, E_x	the pre–image of a set U in the base–space of a bundle E; when $U = \{x\}$ this is the fiber of E over x	
$g \cdot x$	the (left) action of a group element g with a point x	
$x_i \rightharpoonup y$	weak limit of a sequence $\{x_i\}$ in a normed space V, meaning $lim_{i\to\infty}\Lambda(x_i) = \Lambda(y)$ for all $\Lambda \in V^*$	
$\bigwedge$, $\wedge$	wedge (or alternating) product or smash product	35
$\bigvee$, $\vee$	the wedge or bouquet product	34
∇, ∇_E	a connection on a vector bundle E	8
∇_v	the evaluation of the connection ∇ on the vector field v	
d_∇	the covariant derivative associated with the connection ∇	8
F_∇	the curvature associated to the connection ∇	9
$\mathcal{E}^p(U)$	the space of smooth p–forms with complex coefficients on an open subset U of a manifold	8
$\mathcal{E}^{(p,q)}(U)$	the space of (p,q)–forms on the space U	
$\mathcal{E}^p(U,E)$, $\mathcal{E}^p(E)$	the space of smooth p–forms with coefficients in the bundle E over a manifold X on an open subset U of X; the latter case when $U = X$	8
$E[\mathbb{F}^k]$	a vector bundle E with fibers isomorphic to the vector space $\mathbb{F}^k$	8
tr	the trace functional	
M^*	the conjugate transpose of a square matrix M	
$\mathfrak{g}$	the Lie algebra corresponding to an arbitrary Lie group	8
$GL(n,\mathbb{F}), SO(n,\mathbb{F})$	the general linear and special orthogonal and unitary groups of dimension n over the field $\mathbb{F}$	
$U(n)$	the unitary group of dimension n	
$\mathfrak{gl}(n,\mathbb{F}), \mathfrak{u}(n)$	the respective Lie algebras of the general linear and unitary groups	
$\Gamma(U,E)$, $\Gamma(E)$	the module of smooth sections of the bundle E over a manifold X on an open subset U of X; the latter case when $U = X$	8
$C^k(U)$	the space of C^k functions on an open subset U of a manifold	8
TX	the (real) tangent bundle to the manifold X whose fiber at $x \in X$ is the (real) tangent space T_xX	
T^*X	the (real) cotangent bundle to the manifold X whose fiber T_x^*X at $x \in X$ is the dual of the tangent space	
$T'X$, T'^*X	the holomorphic tangent, respectively cotangent, bundle	
$T''X$, T''^*X	the antiholomorphic tangent, respectively cotangent, bundle	

Symbol	Description	Page
$\bigwedge^p X$	the bundle of p–forms on a manifold X	
$ad(E)$	the adjoint bundle whose fibers are isomorphic to the Lie algebra of the Lie structure group of the bundle E	8
$V_{\mathbb{C}}$, $V^{\mathbb{C}}$	the complexification $V \bigotimes_{\mathbb{R}} \mathbb{C}$ of a real vector space or bundle V	
$Hom_{C^k(X)}(E, F)$	the bundle of homomorphisms from the bundle E over X into the bundle F over X; its fiber at $x \in X$ is the $C^k(X)$–module $Hom(E_x, F_x)$	8
$End(E)$, $Aut(E)$	the bundle of endomorphisms, respectively automorphisms, on the bundle E	
$*$	the Hodge–star operator	9
YM	the Yang–Mills functional	9
Ad_g, ad_g	the adjoint action on a Lie group, respectively Lie algebra	
$\mathcal{A}(E)$, $\mathcal{A}$	the affine space of connections on the bundle E, the latter case when E is implicit	8
$\mathcal{G}(E)$, $\mathcal{G}$	the gauge group on the bundle E, the latter case when E is implicit	31
$[X, Y]$	the space of homotopy classes of based maps $X \to Y$	
$\pi_k(X)$	the k^{th}–homotopy group $[\mathbb{S}^k, X]$ of X	32
$K(\tau; n)$	the Eilenberg–MacLane space for an abelian group τ and an integer $n \geq 1$	32
$\mathbb{RP}^n$, $\mathbb{CP}^n$	the real–, respectively complex–projective n–space	
$Map(A, B)$	the space of maps $A \to B$	
$Map^*(A, B)$	the space of base–point preserving maps $A \to B$	34
vol, vol_X	the volume form on the manifold X	9
$\mathcal{F}(X)$	the space $C^\infty(X)$ of smooth functions on a manifold X	11
$\mathfrak{X}(X)$	the space $\Gamma(TX)$ of smooth vector fields on a manifold X	
$d_x f$	the (usual) derivative of a function $f \in \mathcal{F}(X)$ at a point $x \in X$	11
$H_x f$	the Hessian of a function $f \in \mathcal{F}(X)$ at a point $x \in X$	11
(X, g)	a Riemannian manifold with metric tensor g	11
g_{ij}, g^{ij}	a metric tensor, respectively its inverse, on a manifold	11
Γ^k_{ij}	the Christoffel symbols of second kind	11
$\lambda_x f$, $\lambda_Y f$	the index of a non–degenerate critical point $x \in X$, respectively manifold $Y \subset X$, of $f \in \mathcal{F}(X)$	12,23
$M_t(f)$	the Morse counting series of a Morse function	
$f \in \mathcal{F}(X)$	12,23	
$P_t(X)$	the Poincaré series for a manifold X	14
β_i	the i^{th}–Betti number	12
m_i	the the number of critical points of a function $f \in \mathcal{F}(X)$ of index i	12
$crit(f)$	the set of critical points for the function $f \in \mathcal{F}(X)$	14
X_a	the sublevel set $\{x \in X \mid f(x) \leq a\}$ corresponding to a function $f \in \mathcal{F}(X)$	14
$grad\, f$	the gradient vector field associated to the function $f \in \mathcal{F}(X)$	16

Symbol	Description	Page
e_n	an n–dimensional cell ($\approx \mathbb{S}^n$)	14,17
$H_n(X;R),\ H_n(X)$	n^{th}–homology module over the ring R, the latter instance when R is implicit	
$H_n(X,Y;R),\ H_n(X,Y)$	n^{th}–relative homology module over the ring R where $Y \subseteq X$, the latter instance when R is implicit	
$H^n(X;R),\ H^n(X)$	n^{th}–cohomology module over the ring R, the latter instance when R is implicit	
$H^n(X,Y;R),\ H^n(X,Y)$	n^{th}–relative cohomology module over the ring R where $Y \subseteq X$, the latter instance when R is implicit	
$\widetilde{H}$	reduced (co)homology	
$\nu(Y)$	the normal bundle of the submanifold Y of a Riemannian manifold X	22
$\nu^-(Y),\ \nu^+(Y)$	the positive, respectively negative, bundles of the submanifold Y	23
$\nu_\varepsilon^-(Y)$	the exponential image of $\nu^-(Y)$ in X	23
X_G	the homotopy quotient of the G–space X	25
f_G	the induced function on the homotopy quotient from a function $f \in \mathcal{F}(X)$	25
BG	the classifying space of the group G - also the base space of a universal G–bundle	25,83
EG	the total space of a universal G–bundle	83
P_t^G	the equivariant Poincaré series	26
$H_G^\bullet$	equivariant cohomology	26
ΩX	the loop–space of the space X	35
$\mathcal{O}(E)$	the sheaf of germs of holomorphic sections of the bundle E	
$H^p(X,\mathcal{S}), \check{H}^p(X,\mathcal{S})$	the p^{th} Čech cohomology module with respect to the sheaf $\mathcal{S}$	
f^*	the pull–back of a smooth function f or the formal adjoint of a linear map f on an inner–product space	
$[\cdot,\cdot]$	the Lie bracket	
∂_∇	the $(1,0)$–component of the covariant derivative d_∇	39
$\overline{\partial}_\nabla$	the $(0,1)$–component of d_∇	39,44
(n,k)	the greatest common divisor of the natural numbers n and k	
$N(n,k)$	the moduli space of stable bundles of rank n and degree k; only defined when $(n,k)=1$	
$c_k(E)$	the k^{th}–Chern class of the complex vector bundle E	
$c_1(E)[X]$	the integration of the first Chern class on the *fundamental cycle* of the Riemann surface base–space X of the bundle E: this is the generator of the homology group $H_1(X)$	42
$deg(E)$	the degree of the complex vector bundle E	42
$\mu(E),\ \mu$	the slope, respectively type of the complex vector bundle E	42,63
$\overline{\partial},\ \overline{\partial}_\mathcal{E}$	a partial connection corresponding to a holomorphic structure $\mathcal{E}$ on a smooth complex vector bundle	43
$\mathcal{E}_\nabla$	the holomorphic vector bundle induced by the locally free sheaf of germs of local solutions to to the equation $\overline{\partial}_\nabla s = 0$	44

Symbol	Description	Page
$O(E)$	the orbit of connections $\{g \cdot \nabla \mid g \in \mathcal{G}(E)^{\mathbb{C}}\}$ such that $E_\nabla \cong E$	44
$\mathcal{C}$, $\mathcal{C}(E)$	the space of holomorphic structures on the smooth bundle E	60
$\mathcal{C}_s$, $\mathcal{C}_{ss}$	the subspace of $\mathcal{C}$ of stable, respectively semi–stable bundles	68,60
$\mathcal{C}_\mu$, $\mathcal{A}_\mu$	the strata of $\mathcal{C}$, respectively, the induced strata of $\mathcal{A}$	60
$\triangle$	the Laplacian of a complex	
$\| \cdot \|_{p,k}$	the Sobolev (p,k)–norm	44
L^p_k	the Sobolev space	44,71
$\mathcal{A}^p_k$, $\mathcal{G}^p_k$	the Sobolev space of connections, respectively gauge transformations	44,71
d_μ	the codimension $codim_{\mathbb{C}}\mathcal{C}_\mu$	63
$J(X)$, $J_k(X)$	the Jacobian variety of a Riemann surface X, the latter case parametrising the line bundles over X of degree k	69,86
C^∞_o	the space of smooth functions with compact support	71
$L^p_{k,o}$	the closure of C^∞_o in L^p_k	71

References

[1] Atiyah, M. *Collected works, volume 5: Gauge theories.* New-York: Oxford Uni. press (1987).

[2] Atiyah, M. F. & Bott, R. *The Yang-Mills equations over Riemann surfaces.* Philos. Trans. Roy. Soc. London **A 308**, 523-615 (1982).

[3] Bott, R. *Differential geometric methods in mathematical physics.* Lecture notes in mathematics, **836**, New-York: Springer-Verlag (1980).

[4] Bott, R. *On torsion of Lie groups.* Proc. Nat. Acad. Sci. USA **40**, 586-588 (1954).

[5] Bott, R. & Tu, R. *Differential forms in algebraic topology.* New-York: Springer–Verlag (1982).

[6] Bourguignon, J. & Lawson, H. B. *Stability and isolation phenomena for Yang–Mills fields.* Commun. Math. Phys. **79**, 189-230 (1981).

[7] Buchdahl, N. P. *Hermitian–Einstein connections and stable vector bundles over compact complex surfaces.* Math. Ann. **280**, 625-628 (1988).

[8] Chang, K. *Infinite-dimensional Morse theory and multiple solution problems.* Boston: Birkhäuser (1993).

[9] Donaldson, S. K. *A new proof of a theorem of Narasimhan and Seshadri.* J. Diff. Geom. **18**, 269-277 (1983).

[10] Donalsdon, S. K. *Anti self-dual Yang–Mills connections over complex algebraic surfaces and stable vector bundles.* Proc. London Math. Soc. **3 50**, A1-A29 (1985).

[11] Donalsdon, S. K. & Kronheimer, P. B. *The geometry of four–manifolds.* New-York: Oxford press (1990).

[12] Eells, J. *Singularities of smooth maps.* New-York: Gordon & Breach (1967).

[13] Forster, O. *Lectures on Riemann surfaces.* New-York: Springer–Verlag (1981).

[14] Goresky, M. & MacPherson, R. *Stratified Morse theory.* Heidelberg: Springer–Verlag (1988).

[15] Grauert, H. & Remmert, R. *Coherent analytic sheaves.* Berlin: Springer–Verlag (1984).

[16] Greenberg, M. J. & Harper, J. R. *Algebraic topology, a first course (revised).* New-York: Addison–Wesley (1981).

[17] Griffiths, P. & Harris, J. *Principles of algebraic geometry.* New-York: John Wiley (1978).

[18] Gunning, R. C. *Lectures on Riemann surfaces.* New-Jersey: Princeton Uni. press (1966).

[19] Harder, G. & Narasimhan, M. S. *On the cohomology groups of moduli spaces of vector bundles over curves.* Math. Annln. **212**, 215-248 (1975).

[20] Hebey, E. *Sobolev spaces on Riemannian manifolds.* Lecture notes in Mathematics **1635**, (1996).

[21] Istrăţescu, V. I. *Fixed point theory: an introduction.* Dordrecht: D. Reidel publishing (1981).

[22] Kobayashi, S. *Differential geometry of complex vector bundles.* Kanô Memorial Lectures **5**, (1987).

[23] Kodaira, K. *Complex manifolds and deformation of complex structures.* New-York: Springer-Verlag (1986).

[24] Madsen, I. & Tornehave, J. *From calculus to cohomology.* Cambridge: Cambridge press (1997).

[25] May, J. P. *A concise course in algebraic topology.* Chicago: Chicago Uni. press (1999).

[26] Milnor, J. *Morse theory.* Ann. of Math. **51**, 1-153 (1963).

[27] Narasimhan, M. S. & Seshadri, C. S. *Stable and unitary vector bundles on a compact Riemann surface.* Ann. of Math. **82**, 540-567 (1965).

[28] Newlander, A. & Nirenberg, L. *Complex analytic coördinates in almost complex manifolds.* Ann. of Math. **65**, 391-404 (1957).

[29] Palais, R. S. *Foundations of global non-linear analysis.* New-York: Benjamin (1968).

[30] Schwarz, M. *Morse homology.* Boston: Birkhäuser (1993).

[31] Taylor, M. E. *Partial differential equations III: nonlinear equations.* New-York: Springer–Verlag (1996).

[32] Uhlenbeck, K. K. *Connections with L^p bounds on curvature.* Comm. Math. Phys. **83**, 31-42 (1982).

[33] Uhlenbeck, K. K. *Removable singularities in Yang–Mills fields.* Comm. Math. Phys. **83**, 11-29 (1982).

[34] Whitehead, G. W. *Elements of homotopy theory.* New-York: Springer–Verlag (1978).