Journal of Geometry and Physics 8 (1992) 89–122
North-Holland

JOURNAL OF
GEOMETRY AND
PHYSICS

Boundary value problems for Yang–Mills fields

S.K. Donaldson

The Mathematical Institute, 24–29 St. Giles, Oxford, UK

Received 23 September 1991

This paper investigates boundary value problems for Hermitian Yang–Mills equations over complex manifolds. The main result is the unique solubility of the Dirichlet problem for the Hermitian Yang–Mills equation. Connections with a number of topics are found, including the link with loop groups.

Keywords: *Yang–Mills fields, boundary value problems*
1991 MSC: 81T13

Dedicated to Roger Penrose

1. Introduction

It is natural to hope that differential-geometric results dealing with closed manifolds will extend to yield interesting information for manifolds with boundary. A good instance of this occurs in two-dimensional Riemannian geometry. Each conformal class of metrics on a closed surface contains a metric of constant curvature: On a surface with boundary we might seek conformal constant curvature metrics with an additional boundary condition, for example, that the boundary have constant geodesic curvature. Then the positive solution of this boundary value problem is equivalent to a version of the Riemann mapping theorem. Another instance occurs in Hamilton's results [14] for harmonic maps.

In this paper our main concern is not with Riemannian metrics or harmonic maps (although we shall use Hamilton's result in section 3, and we return to Riemannian metrics briefly in section 4) but with Yang–Mills fields. More specifically, we will investigate boundary value problems for the Hermitian Yang–Mills equations over complex manifolds. There is a sizeable literature on these equations over closed manifolds, and in particular on the link between the equations and the holomorphic geometry of stable vector bundles [28,7,8], and our aim here is to see what kind of counterparts we can obtain in the boundary case. In section 2 we begin by recalling some of the salient theory and prove the main result of this paper: the unique solubility of the Dirichlet problem for the Hermitian Yang–Mills equation. The proof of this is quite straightforward and the main thrust of the paper is thus in the interpretation of the result, rather than the proof. We find connections with a number of topics, which are discussed in section 3. Perhaps the most notable of these is the link with the literature on the geometry of loop groups. It is well known that the space ΩG of based maps from the circle to a compact group G has a natural complex Kähler metric, related to the Kähler metrics on coadjoint orbits of G. The complex structure can be obtained from a basic factorisation theorem which gives an alternative description

$$\Omega G \cong LG^{c}/\mathrm{Hol}(D, G^{c}),$$

of (free) loops in the complexified group G^c, modulo the boundary values of holomorphic maps from the disc to G^c. We shall see that this factorisation is equivalent to a special case of our main theorem, and from this perspective we shall obtain generalisations of the above results in the form of *hyper-Kähler* metrics on spaces

$$\mathrm{Maps}(\partial Z, G^{c})/\mathrm{Hol}(Z, G^{c}),$$

where Z is a convex domain in $\mathbb{C}^2$. We shall also study a variant of the equations due to Hitchin, and we shall see that these lead to hyper-Kähler metrics on the "complex loop groups" ΩG^c, where G^c is the complexification of G. We explain how these are related to the hyper-Kähler metrics on complex co-adjoint orbits which have been found by Kronheimer. Other topics we discuss in section 3 are variational problems, constant mean curvature surfaces in hyperbolic space, and a "Neumann problem" for the Hermitian Yang–Mills equations. We finish the paper, in section 4, by mentioning a number of further directions which might be explored.

Section 2. The main result

2.1. THE HERMITIAN YANG–MILLS EQUATIONS

We consider a holomorphic vector bundle E over a complex manifold Z, and a Hermitian metric H on the fibres of E. It is a simple fact that there is then a preferrred unitary connection induced on E, with curvature F_H say. In a local holomorphic trivialisation of E, by sections s_i, we can represent the metric by a Hermitian matrix (which we also denote by H) $H_{ij} = (s_i, s_j)_H$. Then the connection matrix in this trivialisation is $H^{-1}\partial H$ and the curvature is

$$F_H = \overline{\partial}(H^{-1}\partial H) = H^{-1}(\overline{\partial}\partial H - \overline{\partial} H H^{-1}\partial H). \qquad (1)$$

We now suppose that the base manifold Z has a fixed Kähler metric, with Kähler form ω, and we let $\Lambda : \Omega_Z^{1,1} \to \Omega_Z^0$ be the contraction $\Lambda(\theta) = (\omega, \theta)$. The Hermitian Yang–Mills tensor of a metric H on E is defined to be $\mathrm{i}\Lambda F_H$. The *Hermitian Yang–Mills (HYM) equation* for the metric H is the condition:

$$\mathrm{i}\Lambda F_H = 0. \tag{2}$$

We will call the solutions Hermitian Yang–Mills metrics. We may equally refer to a unitary connection whose curvature has type $(1,1)$ and satisfies the condition (2) as an HYM connection; this causes no confusion, since the $\bar{\partial}$-operator of such a connection defines a holomorphic structure. We could develop the whole theory for G-connections, where G is a general compact Lie group. This would involve little more than a change in notation. The proofs are slightly easier to write down in the basic case of a unitary group, but it is useful to have the general case in mind, so we will adopt the policy of sometimes stating results for a general group G, but assuming that $G = \mathrm{U}(n)$ in the proofs. We should recall that these equations have special significance in the cases when the complex dimension N of the base manifold is 1 or 2. When $N = 1$ the map Λ is an isomorphism and the HYM connections are just the flat connections. When $N = 2$ the kernel of Λ consists of the anti-self-dual forms, which depend only on the Riemannian structure of the base four-manifold, and the solutions are the anti-self-dual connections, or instantons.

In a local holomorphic trivialisation, as above, the HYM equation takes the form:

$$H\mathrm{i}\Lambda\bar{\partial}(H^{-1}\partial H) = \tfrac{1}{2}\Delta H - \mathrm{i}\Lambda\bar{\partial}HH^{-1}\partial H = 0. \tag{3}$$

Here we have used the fact that the Laplacian Δ (on functions) on a Kähler manifold can be written $\Delta = 2\partial^*\partial = 2\mathrm{i}\Lambda\bar{\partial}\partial$. We see then that our equation is a non-linear, second order PDE whose leading term is the Laplace equation. As we have mentioned in the introduction, it is known that when Z is a closed Kähler manifold the necessary and sufficient condition for the existence of the solution is that E be a "stable" holomorphic bundle. In the present paper we will consider the case when Z is the interior of a compact manifold $\overline{Z}$ with non-empty boundary ∂Z, and the Kähler metric ω is smooth and non-degenerate on the boundary. We consider a holomorphic bundle E which extends to the boundary. (That is, E can be given by a system of transition functions on overlaps $U_\alpha \cap U_\beta \subset \overline{Z}$ which are smooth up to the boundary and holomorphic on the interior $U_\alpha \cap U_\beta \cap Z$.) In this situation we will say that E is a holomorphic bundle over $\overline{Z}$. We consider the Dirichlet problem for the Hermitian Yang–Mills equations, in which the metric H is specified on the boundary. Our main result is very easy to state.

Theorem 1. *Let E be a holomorphic vector bundle over the compact Kähler manifold* $\overline{Z}$, *with non-empty boundary* ∂Z. *For any Hermitian metric f on the restriction of E to* ∂Z *there is a unique metric H on E such that*
(i) $H = f$ *over* ∂Z,
(ii) $\mathrm{i}\Lambda F_H = 0$ *in Z.*

This result is simpler to state than that for closed manifolds, since the delicate notion of stability does not enter. The reason for this will emerge clearly in the course of the proof. The result for manifolds with boundary is indeed not at all difficult; it is more or less an exercise in adapting the techniques which already appear in the literature on the Hermitian Yang–Mills equation. In particular, the result lies very close to some of those obtained by Simpson in ref. [26], who considers boundary value problems in an auxiliary way, although Simpson's direction is rather different and he does not state precisely the result we need.

While it is, perhaps, disappointing that the boundary value problem requires little new input on the analytical side, it seems worth giving the proof, since it is so simple. For the sake of exposition we will present two methods of proof. The first applies in the restricted case (which is where our main interest lies in this paper), when the bundle E is holomorphically trivial and we can use the continuity method. The second proof treats the general case using the "heat equation" method. Before beginning these proofs we will briefly recall some of the important properties of the Hermitian Yang-Mills equation.

2.2. PROPERTIES OF THE HERMITIAN YANG–MILLS EQUATION

We will use two key properties of the Hermitian Yang–Mills equations. Both depend on a global analogue of the formula (1). If H and K are two metrics on a holomorphic bundle E then $\eta = K^{-1}H$ is a section of the bundle of endomorphisms $\operatorname{End} E$, self-adjoint with respect to either metric. We have then

$$\partial_H = \partial_K + \eta^{-1}\partial_K\eta,$$

$$F_H = F_K + \overline{\partial}(\eta^{-1}\partial_K\eta). \qquad (4)$$

Here ∂_H is the $(1,0)$ part of the covariant derivative ∇_H of the connection compatible with H, and similarly for ∂_K. [The $(0,1)$ part of each connection is the $\overline{\partial}$ operator associated to the holomorphic bundle E.] We see from (4) that the linearisation of the HYM equations about a solution K is represented by the covariant Laplace equation:

$$\Delta_K\rho \equiv 2\mathrm{i}\Lambda\overline{\partial}\partial_K\rho = 0. \qquad (5)$$

That is, if η_t is a one-parameter family of K-self-adjoint endomorphisms with $\eta_0 = 1$ and if $\eta_t K$ is a solution of the Hermitian Yang–Mills equation for each t, then $\Delta_K \rho = 0$ where ρ is the t-derivative of η_t at $t = 0$. This follows immediately by applying $\mathrm{i}\Lambda$ to (4) and using the fact that the covariant Laplacian $\Delta_K = \nabla_K^* \nabla_K$ can be written

$$\Delta_K = 2\partial_K^* \partial_K = 2\mathrm{i}\Lambda \overline{\partial} \partial_K,$$

when K is an HYM solution. (In general, $\Delta_K \rho = 2\partial_K^* \partial_K \rho + 2\mathrm{i}[\Lambda F_K, \rho]$.)

The second property we need brings in the non-linear term in the equation. Let $\mathcal{H}$ be the space of $n \times n$ positive Hermitian matrices, which we may think of as the quotient $\mathrm{GL}(n, \mathbb{C})/\mathrm{U}(n)$. (In the case of a general group G, we would work with the quotient $\mathcal{H} = G^c/G$.) Hermitian metrics on a trivial bundle are represented by maps into $\mathcal{H}$, and for a general bundle E by cross-sections of a bundle $\mathcal{H}_E$ with fibre $\mathcal{H}$. There is a natural complete, $\mathrm{GL}(n, \mathbb{C})$-invariant Riemannian metric on $\mathcal{H}$ given by $\mathrm{Tr}(h^{-1}\delta h)^2$. For any two points h, k in $\mathcal{H}$ we set

$$\sigma(h,k) = \mathrm{Tr}(k^{-1}h) + \mathrm{Tr}(h^{-1}k) - 2n.$$

If we diagonalise h with respect to k, with relative eigenvalues λ_a, then

$$\sigma(h,k) = \sum_a (\lambda_a^{-1} + \lambda_a - 2),$$

from which we see that $\sigma \geq 0$, with equality only if $h = k$, and that

$$\sigma(h,k) \sim d(h,k)^2 \tag{6}$$

as h tends to k, where $d(\ ,\ )$ is the Riemannian distance function on $\mathcal{H}$. We can express d as a monotone function of σ, and for each fixed k the balls in $\mathcal{H}$, $\{h | \sigma(h,k) \leq R\}$ are compact.

Now if H, K are two metrics on a bundle E over Z we can define a positive function $\sigma(H,K)$ on Z fibrewise. Then if H and K are solutions of the Hermitian Yang–Mills equation we claim that σ is sub-harmonic: $\Delta(\sigma) \leq 0$. To see this we apply $\mathrm{i}\Lambda$ to (4), and also take the trace in the bundle E to get:

$$\Delta\ \mathrm{Tr}(K^{-1}H) = \Delta\ \mathrm{Tr}\ \eta = 2\mathrm{i}\Lambda\ \mathrm{Tr}(\overline{\partial}\eta\ \eta^{-1}\ \partial_H \eta).$$

Choose a bundle trivialisation which, at a given point $z \in Z$, is K orthonormal and in which η is diagonal with eigenvalues λ_a, and let π_{ab} be the matrix entries of $\overline{\partial}\eta$, at z. Then $\partial_H \eta$ has entries $\overline{\pi}_{ba}$ and

$$\mathrm{i}\Lambda\ \mathrm{Tr}(\overline{\partial}\eta\ \eta^{-1}\ \partial_H\eta) = i\Lambda \sum_{a,b} \lambda_a^{-1} \pi_{ab} \overline{\pi}_{ab} = -\sum_{a,b} \lambda_a^{-1} |\pi_{ab}|^2 \leq 0.$$

Thus Δ Tr $(K^{-1}H) \leq 0$. Interchanging H and K we see that $\mathrm{Tr}(H^{-1}K)$ is likewise sub-harmonic, and hence also $\sigma(h,k)$.

This calculation leads immediately to the *uniqueness* of the solution to the Dirichlet problem. For if H, K are two Hermitian Yang–Mills solutions on E with the same boundary value f, then $\sigma = \sigma(H,K)$ is a smooth function on Z with

$$\sigma \geq 0, \quad \Delta\sigma \leq 0, \quad \sigma = 0 \text{ on } \partial Z,$$

and it follows immediately from the maximum principle that σ is everywhere 0, and so $H = K$. [Alternatively, we can see this by writing

$$\int_Z |\nabla\sigma|^2 = \int \sigma\Delta\sigma \leq 0,$$

which shows that $\nabla\sigma = 0$.]

2.3. THE PROOF OF THEOREM 1 FOR TRIVIAL BUNDLES

If the holomorphic bundle E in theorem 1 is trivial over Z we can prove existence by the continuity method. We represent a metric in a fixed holomorphic trivialisation by a map $H : Z \to \mathcal{H}$. Any constant map gives a flat solution to the Hermitian Yang–Mills equations, so the solution to the boundary value problem exists if f is a constant. Since the space $\mathcal{H}$ is contractible, it suffices to show that, if f_s is a continuous family of smooth maps from ∂Z to $\mathcal{H}$, for $s \in [0,1]$, and a solution H_0 to the HYM equations with boundary value f_0 exists, then it extends to a family of solutions H_s with boundary values f_s for all s. Following the standard pattern, we show that the set $S \subset [0,1]$ for which a solution exists is both open and closed.

One sees that S is closed by obtaining bounds on the solutions and their derivatives. Let $H : Z \to \mathcal{H}$ be a solution with boundary value f, and K be any constant map. Then we have, by the maximum principle,

$$\max_Z \sigma(H,K) = \max_{\partial Z} \sigma(f,K).$$

This is an *a priori* bound which confines a solution to a compact sub-set of $\mathcal{H}$, determined by the boundary values. This elementary step, giving a C^0 bound on the solution, is the crucial one in which we see the difference between our present boundary value problem and the more difficult problem over closed manifolds. It is a fairly routine matter to go on to obtain bounds on higher derivatives, and we describe two approaches to this.

For the first approach we follow ref. [8] and suppose s_∞ is the limit of a sequence s_i in S, so there are solutions $H_i = H_{s_i}$. By the maximum principle again, we see that the maximum value of $\sigma(H_i, H_j)$ is attained on ∂Z, and it follows that H_i is a Cauchy sequence in the C^0 metric on $\mathrm{Maps}(Z, \mathcal{H})$, and so converges uniformly to a limit H_∞—a continuous metric on E. The argument of ref. [8] gives a C^1 bound on H_∞ (and shows that the H_i converge in C^1) through a *reductio ad absurdum*. One considers points $z_i \in Z$ where the supremum $D_i = \sup_Z |\nabla H_i|$, is attained, and rescales a ball of radius D_i^{-1} about z_i to a fixed unit ball, in which one applies elliptic estimates for the Laplacian. This argument extends without difficulty to the case of the boundary value problem, see ref. [26]. Once one has the C^1 bound it is very easy to obtain bounds on all higher derivatives: "bootstrapping" by substituting into the equation

$$\Delta H = 2\mathrm{i}\Lambda\bar{\partial}HH^{-1}\partial H.$$

This approach has the disadvantage of being nonconstructive—it does not yield any explicit bound on the derivatives of solutions. To do this we now consider a second approach, in which we assume for simplicity that Z is a domain in $\mathbb{C}^N$, with the Euclidean metric. We will explain a scheme which will give, in principle, explicit estimates on all derivatives of a solution H to the HYM equations in terms of its boundary values f.

Fix a point w in ∂Z and let K_w be the constant metric with constant value $f(w)$. We know that $\sigma = \sigma(H, K_w)$ is sub-harmonic, so if ϕ_w is the harmonic function in Z with boundary value:

$$\phi_w|_{\partial Z} = \sigma|_{\partial Z} = \sigma(f, K_w);$$

the maximum principle gives

$$0 \leq \sigma(z) \leq \phi_w(z) \quad \text{for all } z \in Z.$$

Now ϕ_w is smooth up to the boundary, so for z near w,

$$\phi_w(z) \leq C|z - w|,$$

so by (6) we get a Holder bound

$$d(H_z, H_w) \leq \text{const.}|z - w|^{1/2},$$

for all w on the boundary.

We will now go from this Holder condition on the boundary, to a first derivative estimate over all of Z. For any vector v in $\mathbb{C}^n$ the endomorphism $\rho_v = H^{-1}\nabla_v H$ satisfies $\Delta_H \rho_v = 0$. This follows from (5) by considering the one-parameter family of solutions obtained by translating in the direction of v, $H_t(z) = H(z + tv)$. It follows that the square norm $|\rho_v|_H^2 = \operatorname{Tr} \rho_v^2$ is sub-harmonic,

$$\Delta|\rho_v|^2 = 2(\Delta_H \rho_v, \rho_v) - 2|\nabla_H \rho_v|^2 = -2|\nabla_H \rho_v|^2 \leq 0.$$

Summing over a real orthonormal basis v_i for $\mathbb{C}^n$ we find that

$$|H^{-1}\nabla H|_H^2 = \sum_i |\eta_{v_i}|_H^2$$

is sub-harmonic, and so attains its maximum value on the boundary. Thus it suffices to estimate the derivative of the solution on the boundary.

We now appeal to a scaling argument, much as in the first approach. For $r \geq 0$ let $\Omega(r)$ be the half-ball

$$\Omega(r) = B^{2N}(r) \cap \{z_i \in \mathbb{C}^N | \operatorname{Re}(z_1) \geq 0\}.$$

We consider a solution $\tilde{H}$ of the HYM equations over $\Omega(1)$ with boundary value $\tilde{f}$ on $\{\operatorname{Re}(z_1) = 0\}$. We suppose that the modulus of the derivative $|\tilde{H}^{-1}\nabla\tilde{H}|_H$ is bounded by 1 say over all of $\Omega(1)$, and that $\tilde{H}$ takes values in a fixed compact set in $\mathcal{H}$. Then we can substitute into the HYM equation, written in the form

$$\Delta\tilde{H} = 2\mathrm{i}\Lambda(\bar{\partial}\tilde{H}\tilde{H}^{-1}\partial\tilde{H}),$$

and use standard elliptic estimates for the Laplace operator to obtain bounds on all the derivatives of $\tilde{H}$ in an interior region $\Omega(r)$, for $r < 1$, in terms of the boundary data $\tilde{f}$ and its derivatives. In particular, by bounding the second derivative and using the intermediate value theorem, we get an estimate for the derivative $[\tilde{H}^{-1}\nabla\tilde{H}]_0$ at the point $z = 0$:

$$|[\tilde{H}^{-1}\nabla\tilde{H}]_0| \leq CV(\epsilon), \tag{7}$$

where $V(\epsilon, 0)$ is the supremum of $d(\tilde{H}(\zeta), \tilde{H}(0))$ as ζ runs over the hemisphere $\{|\zeta| = \epsilon,\ \operatorname{Re}(\zeta_1) \geq 0\}$, and ϵ and $C > 0$ are determined by $\tilde{f}$. We may suppose the same estimate holds for solutions over a slightly deformed copy of $\Omega(1)$.

We can now scale this inequality down, and apply it to small half-balls on the boundary of Z. Let D be the supremum of $H^{-1}|(\nabla H)|$ over Z, and w be a point in ∂Z where this supremum is attained. We rescale a ball of radius D^{-1} about w by a factor D, and translate the centre to the origin. Then we deduce from (7) that for some suitable constants C, ϵ, there is a point z in Z with $d(z, w) = \epsilon D^{-1}$ such that

$$\begin{aligned} D = |[H^{-1}\nabla H]_w| &\leq C\ d(H(z), H(w))(|z-w|)^{-1} \\ &\leq C\epsilon D d(H(z), H(w)), \end{aligned}$$

i.e., $d(H(z), H(w)) \geq (C\epsilon)^{-1}$. But the Holder bound above gives, since $|z-w| = \epsilon D^{-1}$, the bound $d(H(z), H(w)) \leq$ const. $(\epsilon D^{-1})^{1/2}$, so combining the two inequalities we obtain an upper bound on D. Thus we have a first derivative bound on solutions H, which could be written out explicitly in terms of the boundary data f. From this point it is straightforward to bootstrap, using the HYM equation to obtain bounds on all higher derivatives. This completes our discussion of the "closedness" step in the continuity method.

To see that S is open one applies the implicit function theorem. Let K be a solution of the Hermitian Yang–Mills equations in Z (smooth up to the boundary). Recall that if we parametrise the metrics $H = \eta K$ by endomorphisms η of E, self-adjoint relative to H, then

$$\mathrm{i}\Lambda F_H = \mathrm{i}\Lambda\bar{\partial}(\eta^{-1}\partial_K\eta).$$

This is self-adjoint with respect to H, so the conjugate $\Phi(\eta) \equiv \eta^{1/2}\mathrm{i}\Lambda F_H\eta^{-1/2}$ is self-adjoint with respect to K. We fix a large index l (greater than the complex dimension of Z) and consider the following spaces:

(i) the space U_l of L^2_l sections of $\operatorname{End} E$ over Z, self-adjoint with respect to K;

(ii) the space $U^{\partial Z}_{l-1/2}$ of $L^2_{l-1/2}$, self-adjoint sections of $\operatorname{End} E$ over ∂Z.

We define a map

$$N : U_l \to U_{l-2} \times U^{\partial Z}_{l-1/2},$$

by $N(\rho) = (\Phi(\exp(\rho)), \rho|_{\partial Z})$. We claim that N gives an isomorphism between neighbourhoods of 0 in the two spaces. This follows from the inverse function theorem. The derivative of N at 0 is the map

$$DN : L^2_l(Z) \to L^2_{l-2}(Z) \times L^2_{l-1/2}(\partial Z)\ ,\ DN(\rho) = (\tfrac{1}{2}\Delta_K\rho, \rho|_{\partial Z}).$$

The usual Fredholm alternative for boundary value problems tells us that DN is an isomorphism so long as there is no non-zero τ with $\Delta_H\tau = 0, \tau|_{\partial Z} = 0$, and this is immediate since for any such τ,

$$\int_Z |\nabla_H\tau|^2 = \int_Z (\Delta_H\tau, \tau) = 0,$$

so $\nabla_H\tau = 0$ and the boundary condition forces τ to be zero over Z.

By the inverse function theorem, then, N is a local isomorphism and in particular for any small ψ there is a ρ with $N(\rho) = (0, \psi)$. But this is just the assertion that the set of boundary values f for which an L^2_l solution H exists is open. If the boundary data are smooth then so is the solution by elliptic regularity. Thus it follows that in our path f_s the set S for which there exists a solution is open. This completes the first proof, for the case when the bundle E is trivial.

2.4. THE PROOF OF THEOREM 1 FOR THE GENERAL CASE

In the general case when E need not be trivial we can prove theorem 1 by using the "heat equation" method to deform an arbitrary initial metric to the desired solution. While this is slightly more complicated technically than the continuity method considered above, the main points in the discussion are very similar. For example, we make heavy use of the maximum principle for the linear heat equation rather than for the Laplace equation.

For our given data f on ∂Z we consider the evolution equation

$$H^{-1}\frac{\partial H}{\partial t} = -2\mathrm{i}\Lambda F_H, \quad H|_{\partial Z} = f, \tag{8}$$

as in refs. [7,8,26], with some arbitrary smooth initial condition H_0 extending f. This is a parabolic equation, so standard theory gives short-time existence. The nub of the proof is to show that the solution persists for all time and converges to a limit. The proof of long-time existence is given in ref. [26], and is not significantly different from the corresponding proof in the case of a closed base manifold. It depends on uniform estimates for the solutions and their derivatives, and one could make these estimates quite explicit in the same manner as in the second discussion of section 2.3 above. It remains then to show that the solution to the evolution equation converges as $t \to \infty$, and it is here that we see clearly how the Dirichlet problem on a manifold with non-empty boundary differs from the problem on a closed manifold.

The essential observation is the following:

Lemma. *Suppose $\theta \geq 0$ is a sub-solution of the heat equation on $Z \times [0, \infty)$, i.e., $\partial\theta/\partial t + \Delta\theta \leq 0$. If $\theta = 0$ on ∂Z for all time then θ decays exponentially*

$$\sup_Z \theta(z,t) \leq C\mathrm{e}^{-\mu t}$$

(where μ depends only on Z and C on the initial value of θ).

This standard fact can be seen from the spectral representation of the solution to the heat equation. By the maximum principle for the heat operator it suffices to establish the bound when $\partial\theta/\partial t + \Delta\theta = 0$. This solution can be written

$$\theta = \sum_\lambda a_\lambda \mathrm{e}^{-\lambda t} g_\lambda,$$

where λ, g_λ are eigenvalues and normalised eigenfunctions of the Laplacian with Dirichlet boundary conditions on Z. On the other hand, for $k \geq 1/2\ \dim_{\mathbf{C}} Z$

the Sobolev embedding theorem gives an inequality, for any function $\phi = \sum b_\lambda g_\lambda$:

$$\|\phi\|^2_{C^0} \leq \text{Const.}\|\Delta^k \phi\|^2_{L^2} = \text{Const.} \sum \lambda^k |b_\lambda|^2.$$

So

$$|\theta|^2 \leq \text{Const.} \sum a_\lambda^2 \lambda^k \mathrm{e}^{-2\lambda t},$$

which gives the desired estimate, with any μ bigger than the first eigenvalue of the Laplacian.

To apply this lemma to the solution H of the non-linear equation (8) we consider the function $\mathcal{E} = \|\mathrm{i}\Lambda F_H\|^2_H$ on $Z \times [0, \infty)$. An easy calculation [7,10, ch. 6] shows that $\partial\mathcal{E}/\partial t + \Delta\mathcal{E} \leq 0$, and the boundary condition satisfied by H implies that, for $t > 0$, $\mathcal{E}$ vanishes on the boundary of Z. Thus the lemma tells us that $\mathcal{E}$ decays exponentially, and in particular that

$$\int_0^\infty \sqrt{\mathcal{E}}\, \mathrm{d}t < \infty \tag{9}$$

at each point of Z. But, pointwise on Z, the time derivative of the family of metrics $H_t(z)$, considered in the fibre $\mathcal{H}_z$ of the bundle of metrics $\mathcal{H}_E$, is the endomorphism $\mathrm{i}\Lambda F_H$, so $\sqrt{\mathcal{E}}$ is the velocity of $H_t(z)$ in $\mathcal{H}_z$, measured in the invariant metric on $\mathcal{H}_z \cong \mathcal{H}$. The bound (9) tells us that the path $H_t(z)$ in $\mathcal{H}_z$ has finite length. Since $\mathcal{H}_z$ is complete there is a limiting metric $H_\infty(z)$. It is easy then to show, as in ref. [7], that a sub-sequence of the H_t converge in C^∞ to H_∞, and since $\mathrm{i}\Lambda F_H$ tends to zero with t, this limit is the desired solution to the HYM equations.

Note that we do not have this exponential decay for the heat equation on a closed manifold, since the Laplacian has a zero eigenvalue due to the constants. For the HYM heat equation over a closed manifold we can only say that the function $\mathcal{E}$ is bounded, which is useful in establishing the long-time existence of the solution H but allows the possibility that as $t \to \infty$ the metrics H_t "escape"to infinity. This is precisely what happens when the bundle E is "unstable", and it is the analysis of this phenomenon which makes for the complication in the proofs.

Section 3. Discussion of the theorem

3.1. FACTORISATION IN LOOP SPACES

Suppose that the bundle E in theorem 1 is topologically trivial over ∂Z. For any metric on this bundle over the boundary we can choose a orthonormal framing, and our result can be expressed in a slightly different way.

Theorem 1′. *The natural forgetful map sets up a 1–1 correspondence between the equivalence classes of:*

(i) holomorphic bundles over Z with a framing over ∂Z,
(ii) unitary Hermitian Yang–Mills connections over Z on a bundle with a unitary framing over ∂Z.

For a general group G we get a correspondence between holomorphic G^c bundles over Z, trivialised over the boundary, and HYM G-connections, with a G-trivialisation over the boundary. For simplicity we will use the expression "framed bundle" to mean a bundle with a trivialisation over the boundary.

We shall now examine the content of this in the simplest possible case, when Z is the unit disc in $\mathbb{C}$. In complex dimension one the Hermitian Yang–Mills equations just say that the connection should be flat, so we have a correspondence between framed holomorphic G^c bundles and framed flat G-connections. Since the disc is simply connected any flat connection is trivial, and the trivialisation is unique up to the action of G. We see then that the equivalence classes of flat connections on a framed G-bundle can be identified with the quotient LG/G of the free loop space

$$LG = \mathrm{Maps}(S^1, G),$$

under the translation action by G. Of course this is the same as the based loop space ΩG. Turning to the holomorphic side: any bundle over D is holomorphically trivial (as we can see, if we like, from the existence of a flat connection). The framings on the trivial bundle are just elements of LG^c and plainly the equivalence classes of framed holomorphic bundles can be identified with the quotient LG^c/L^+G^c where $L^+G^c \subset LG^c$ is the image under restriction of the group $\mathrm{Hol}(D, G^c)$ of holomorphic maps from the disc to G^c which are smooth up to the boundary. So the content of our result is the natural identification $LG/G \equiv LG^c/L^+G^c$, which is one of the known "factorisation theorems" in loop group theory [22, ch. 8]. The assertion is that any loop in LG^c can be written as the product of a unitary loop in LG and a loop in L^+G^c, and the factorisation is unique up to a constant element of G. This is the analogue for loop groups of the isomorphism, for any compact group G, $G/T = G^c/B$, where T is a maximal torus and B is a Borel sub-group (a maximal parabolic sub-group of G^c).

It is instructive to recall the proof of the factoristaion theorem [in the case $G = \mathrm{U}(n)$] given in ref. [22], which is quite simple and direct. We recast the problem in another form: suppose E is a holomorphic vector bundle over the disc D with a metric on the restriction $E|\partial D$; we want to find a basis of holomorphic sections $s_1, \ldots, s_n$ of E over D which give an orthonormal frame over ∂D. Let H be the Hilbert space of L^2 sections of E over ∂D, and $H^+ \subset H$ the closed sub-space made up of the boundary values of holomorphic sections. We consider the action of multiplication by the holomorphic function

z (i.e. the identity function), which plainly takes H^+ to itself. One shows by index and continuity considerations that $zH^+ \subset H^+$ has codimension n. Let V be the orthogonal complement of zH^+ in H^+. We claim that any orthonormal basis $s_1, \ldots, s_n$ for V has the desired properties. First, the s_a extend holomorphically to the disc since they are in H^+. Second, it suffices to show that the pointwise inner products (s_a, s_b), for each pair a, b, are constant as functions over the boundary circle, for then the fact that the s_a are orthonormal in L^2 implies that they are pointwise orthonormal. To do this we consider the Fourier coefficients of the function (s_a, s_b):

$$p_k = \int_{S^1} (s_a, s_b) e^{ik\theta} \, d\theta,$$

where of course $z = e^{i\theta}$ on the circle. If $k > 0$ we can write this as the L^2 inner product

$$p_k = \langle z^k s_a, s_b \rangle = \int_{S^1} (e^{ik\theta} s_a, s_b) \, d\theta.$$

This vanishes since $z^k s_a$ is contained in zH^+ and s_b is in the orthogonal complement V. Similarly, for negative values $k = -l < 0$ we write $p_{-l} = \langle s_a, z^l s_b \rangle$, and this vanishes for the same reason. So the only non-vanishing Fourier coefficient is p_0 and thus (s_a, s_b) is constant. Finally we have to show that the extensions s_a are linearly independent over the disc. By a change of basis it suffices to show that s_1, say, does not vanish anywhere in D. Suppose on the contrary that s_1 vanishes at a point ζ, with $|\zeta| < 1$. Then we could write $s_1 = (z - \zeta)\tau$ where τ is in H. Then the inner product

$$\langle s_1, z\tau \rangle = \int_{S^1} ((z - \zeta)\tau, z\tau) \, d\theta = \int_{S^1} (1 - \zeta e^{-i\theta}) |\tau|^2 d\theta$$

is non-zero, since $|\zeta| < 1$, and this contradicts the assumption that s_1 is orthogonal to zH^+.

Given this simple alternative argument, it might seem perverse to reach the same conclusion in the model case $Z = D$ through our approach of section 2, using non-linear analysis. The virtue of this latter approach is of course that it gives natural generalisations of the factorisation theorem, which may not be accessible otherwise. In one direction, we can consider a general Riemann surface Σ with smooth boundary $\partial \Sigma = S^1$ say. We let $\mathcal{M}_\Sigma$ be the moduli space of equivalence classes of framed flat G connections over Σ, so we have a map from $\mathcal{M}_\Sigma$ to the representation variety R_Σ of conjugacy classes of representations of $\pi_1(\Sigma)$ in G. The fibre of this map (over the trivial representation) is the loop space LG/G. On the other hand M_Σ can be identified with the framed holomorphic G^c bundles and, since all these bundles are holomorphically trivial, we get a description $M_\Sigma = LG^c / L^{+,\Sigma} G^c$, say, where $L^{+,\Sigma}$ denotes loops which extend holomorphically over Σ. (It would be interesting to see if one can obtain this picture by extending the Fourier theory arguments of the previous paragraph.)

If the boundary of Σ is empty the same results hold, with the technical modification that we must restrict to stable bundles. This is the content of the theorem of Narasimhan and Seshadri [21,3]. We get a finite-dimensional moduli space M_Σ which can be described either as the representation variety (of flat connections), or as the moduli space of stable holomorphic bundles. It is natural to regard the theorem of Narasimhan and Seshadri and the factorisation theorem for loops as companion results: they deal with the two extremes in the general class of framed moduli spaces of flat connections—in the one case the framing data are vacuous and in the other the flat connections are trivial.

In another direction, we can move to higher dimensions, for example to a strictly pseudoconvex domain with smooth boundary $Z \subset \mathbb{C}^N$. Here we encounter a small technical difficulty. The space Z is a Stein manifold so, according to a theorem of Grauert [12], all topologically trivial holomorphic vector bundles over Z are also holomorphically trivial. We would like to apppeal to a variant of this theorem, dealing with bundles over $\overline{Z}$: we would like to say that any holomorphic bundle over $\overline{Z}$, in the sense of section 2.1, has a holomorphic trivialisation which is smooth and non-degenerate up to the boundary. The author has unfortunately not been able to find such a result in the literature. In the appendix we will give an *ad hoc* argument covering the case we are most interested in, when $N = 2$. But the result is almost certainly true in general so for the rest of the discussion in this paragraph we will not explicitly restrict to two variables. Proceeding then by analogy with the model case, we consider the space of framed (topologically trivial) holomorphic bundles over $\overline{Z}$. On the one hand we can write this as a quotient: $\mathrm{Maps}(\partial Z, G^c)/\mathrm{Hol}(Z, G^c)$, where Hol( ) denotes holomorphic maps smooth up to the boundary, and we have supressed the obvious map induced by restriction to the boundary. On the other hand we can apply our theorem, which now takes the form:

$$\mathrm{Maps}(\partial Z, G^c)/\ \mathrm{Hol}(Z, G^c) \cong \text{Isomorphism classes of framed Hermitian Yang–Mills connections over } Z. \tag{10}$$

3.2. KÄHLER AND HYPER-KÄHLER METRICS

Another aspect of loop group theory which fits in naturally with our discussion involves the existence of a Kähler metric on LG/G (see ref. [22, section 8.9]). We will first digress to recall basic facts about moduli spaces of connections, and the metrics on them (compare ref. [10, ch. 4]). Let us consider Z

just as a Riemannian manifold with boundary for the moment, and let $\mathcal{B}_Z$ be the space of framed G-connections over Z. This can be made into an infinite dimensional manifold, and the tangent space at a framed connection A can be identified with the kernel of the operator $d_A^* : \Omega_Z^1 \to \Omega_Z^0$. To see this, we note that $\mathcal{B}_{\overline{Z}}$ may be regarded as the quotient of the space $\mathcal{A}_{\overline{Z}}$ of connections over Z (smooth up to the boundary) by the group $\mathcal{G}_{\overline{Z}}$ of gauge transformations equal to the identity over the boundary. We wish to identify a neighbourhood of an equivalence class $[A] \in \mathcal{B}_{\overline{Z}}$ with a neighbourhood of 0 in $\ker d_A^*$. Just as in the usual set up for closed manifolds one needs to see that for small $a \in \Omega_Z^1(\operatorname{End} E)$ there is a unique gauge transformation g, close to the identity, with

$$d_A^*(g^{-1}d_Ag + gag^{-1}) = 0\,, \quad g|_{\partial Z} = 1. \tag{11}$$

This follows from the implicit function theorem, starting with the unique solubility of the Dirichlet problem

$$d_A^* d_A \phi = \rho\,, \ \phi|_{\partial Z} = 0,$$

which is the linearisation of (11). (Here we can introduce appropriate Sobolev spaces and construct a quotient $\mathcal{B}_{\overline{Z}}$ as a Banach manifold, or stay with C^∞ connections and obtain a Frechet manifold.)

We will now discuss the structure of the moduli space of framed Hermitian Yang-Mills solutions $\mathcal{M}_{\overline{Z}} \subset \mathcal{B}_{\overline{Z}}$. We begin with a formal treatment. Let $[A] \in \mathcal{M}_{\overline{Z}}$ be a solution and consider a nearby connection $A + a$. We write $a = \alpha - \alpha^*$, where $\alpha \in \Omega^{0,1}(\operatorname{End} E)$, $\alpha^* \in \Omega^{1,0}(\operatorname{End} E)$. The Hermitian Yang–Mills equations comprise the pair

$$F^{0,2}(A + a) = \overline{\partial}_A\alpha + \alpha \wedge \alpha = 0,$$

$$\mathrm{i}\Lambda F(A + a) = \mathrm{i}\Lambda(\partial_A\alpha - \overline{\partial}\alpha^* - \alpha^* \wedge \alpha - \alpha \wedge \alpha^*) = 0,$$

whose linearisations are

$$\overline{\partial}_A\alpha = 0\,, \quad \mathrm{i}\Lambda\partial\alpha - \mathrm{i}\Lambda\overline{\partial}\alpha^* = 0.$$

However, by the Kähler identities, the gauge fixing condition $d_A^* a = 0$ can be written as

$$d_A^*(\alpha - \alpha^*) = \mathrm{i}\Lambda\partial_A\alpha + \mathrm{i}\Lambda\overline{\partial}_A\alpha^* = 0,$$

so the linearised equations and the gauge fixing condition combine into the complex-linear equations $\overline{\partial}_A\alpha = \overline{\partial}_A^*\alpha = 0$.

We might expect then that $\mathcal{M}_{\overline{Z}}$ is an infinite dimensional manifold whose tangent space at $[A]$ is identified with the kernel of $\overline{\partial}_A \oplus \overline{\partial}_A^*$, a vector space with an obvious complex structure. This is certainly correct if the complex dimension $N = 1$ or 2, which are our primary concerns; in the higher dimensional case, when one encounters over-determined equations, it is possible that $\mathcal{M}_{\overline{Z}}$ might have singularities. In these two cases we have alternative expressions for the operator $\overline{\partial}_A \oplus \overline{\partial}_A^*$. When $N = 1$ it can be identified with the operator $d_A \oplus d_A^*$ and when $N = 2$ with the operator $d_A^+ \oplus d_A^*$, both acting on one-forms with values in the bundle of skew adjoint endomorphisms of E. For definiteness, let us fix attention on the case $N = 2$, i.e the case of real dimension 4. (The ensuing discussion applies equally well to any Riemannian four-manifold with boundary.) The symbol of the operator $d_A^+ : \Omega^1 \to \Omega^2_+$ is surjective and this means that we can find a right inverse $P : \Omega^2_+ \to \Omega^1$ which is bounded on Sobolev spaces:

$$d^+ P\sigma = \sigma\ , \quad \|P\sigma\|_{L^2_k} \le \text{Const.}\|\sigma\|_{L^2_{k-1}},$$

and with $d^* P\sigma = 0$. One way to construct P is to embed Z isometrically in a closed Riemannian manifold X and choose an extension operator E, from forms on Z to forms on X, which is bounded on Sobolev spaces (see ref. [18, p. 23]). Over X there is a bounded operator P_X which is an inverse to d_A^+ modulo a finite dimensional space, i.e., we have $d^+ P_X \tau = \tau + H(\tau)$, where H is a finite-rank map. Then it is easy to see that we can choose P_X so that $H(\tau)$ is supported in the complement of Z, so we can just take $P = P_X \circ E$ and the estimates for P follow immediately from those for P_X and E.

A local model for $\mathcal{M}_Z$ can now be constructed in just the same way as in the case of a closed base manifold. For simplicity we work in fixed Sobolev spaces: we seek solutions in the form $A + a + P\sigma$, where $d_A^* a = d_A^+ a = 0$. The implicit fuunction theorem in Banach spaces shows that for any small a there is unique small solution σ_a to the equation

$$F^+(A + a + P\sigma) \equiv \sigma + ((a + P\sigma)(a + P\sigma))^+ = 0,$$

and that all nearby solutions are obtained in this way. So we conclude that, when $N = 2$, $\mathcal{M}_{\overline{Z}}$ is a manifold, with tangent space

$$T\mathcal{M}_{\overline{Z}} \cong \ker(\overline{\partial}_A \oplus \overline{\partial}_A^*) = \ker(d_A^* \oplus d_A^+).$$

A similar, but easier, discussion applies when $N = 1$.

We now move on to Riemannian metrics. The L^2 norm on one-forms induces a metric on the tangent space of $\mathcal{M}_Z$:

$$\|a\|^2 = \int_Z |a|^2\, \mathrm{d}\mu, \quad a \in \ker \overline{\partial}_A \oplus \overline{\partial}_A^*.$$

This makes $\mathcal{M}_Z$ into an infinite dimensional Riemannian manifold, and in fact a Kähler manifold. The Kähler property follows from the general symplectic reduction principle, which has been extensively discussed in similar Yang–Mills problems. First, the tangent space $\ker \bar{\partial} \oplus \bar{\partial}^*$ has an obvious complex structure, since $\bar{\partial} \oplus \bar{\partial}^*$ is a complex linear operator. The skew form corresponding to the L^2 metric under this complex structure is

$$\Omega(a,b) = \int_Z \mathrm{Tr}(a \wedge b) \wedge \omega^{N-1}.$$

The map $\mathcal{A} \to \Omega_Z^0(\mathrm{End}\, E)$ defined by $A \mapsto \Lambda F(A)$ is a moment map for the action of the gauge group $\mathcal{G}_{\bar{Z}}$. The calculation is just the same as in the case of a closed manifold [3,17,10]—the fact that the gauge transformations are the identity on the boundary permits the integration by parts used in the calculation. Then the form Ω appears as the usual induced form on the "Marsden–Weinstein quotient" (the quotient of the zeros of the moment map by the action of the symmetry group). In this way one sees that Ω is a closed two-form on $\mathcal{M}_Z$. Similarly, one sees that the almost-complex structure on $\mathcal{M}_Z$, visible from the description of the tangent space, is integrable by appealing to the alternative description of $\mathcal{M}_{\bar{Z}}$ as a moduli space of framed holomorphic bundles.

In the model case, when Z is the disc and we have seen that the moduli space can be identified with the loop space ΩG, it is easy to verify that the metric we have defined agrees with the Kähler metric considered in ref. [22]. So we can again regard our discussion as providing a generalisation of this piece of loop group theory. Using our identification (10) we obtain, for example, Kähler metrics on spaces $\mathrm{Maps}(\partial Z, G^c)/\mathrm{Hol}(Z, G^c)$ when Z is a pseudoconvex domain. On the other hand we can say more in the four-dimensional case. Suppose that Z is a domain in Euclidean space $\mathbb{R}^4$, with smooth boundary. The moduli space $\mathcal{M}_Z$ together with its Riemannian metric depends only on the metric on $\mathbb{R}^4$. For any choice of complex structure, $\mathbb{R}^4 \cong \mathbb{C}^2$, Z becomes a complex manifold and $\mathcal{M}_Z$ acquires a Kähler structure through the discussion above. Putting this together we see that $\mathcal{M}_Z$ has a *hyper-Kähler* structure: an action of the quaternions I, J, K on $T\mathcal{M}_Z$, each one of which defines a complex Kähler structure. On the level of tangent spaces, we see the action of the quaternions on $\ker d_A^* \oplus d_A^+$ from the fact that this operator is quaternion-linear. (This is, again, a familiar principle in four-dimensional Yang–Mills theory, see, for example, ref. [19].) In particular, we see that if $Z \subset \mathbb{C}^2$ is convex (so pseudoconvex for any choice of complex structure) then there is a hyper-Kähler metric on $\mathrm{Maps}(\partial Z, G^c)/\mathrm{Hol}(Z, G^c)$, which we can regard as a four-dimensional analogue of the Kähler metric on LG^c/L^+G^c.

3.3. VARIATIONAL ASPECTS, THE WESS-ZUMINO-WITTEN ACTION

The Hermitian Yang-Mills equations over a closed base manifold can be viewed as Euler-Lagrange equations, but for a non-local Lagrangian functional. Let $\mathrm{Met}(E)$ be the space of Hermitian metrics on a holomorphic bundle E over a Kähler manifold Z, and suppose initially that Z is closed. We define a one-form θ on $\mathrm{Met}(E)$ by assigning to a tangent vector η at a point $H \in \mathcal{H}_E$ the number

$$\theta(\eta) = \int_Z \mathrm{Tr}(\eta\ \mathrm{i}\Lambda F_H).$$

The formula (4) for the variation F shows that the exterior derivative of θ, at the point H, is

$$d\theta(\eta_1, \eta_2) = \int_Z \mathrm{Tr}\,(\eta_1 \Delta_H' \eta_2 - \eta_2 \Delta_H' \eta_1), \tag{12}$$

where $\Delta_H' = \partial_H^* \partial_H$. This expression vanishes, since Δ_H' is self-adjoint. This means that θ is the derivative of a function R on $\mathrm{Met}(E)$, unique up to a constant, and the Euler-Lagrange equation $\delta R = 0$ is, by construction, the Hermitian Yang-Mills condition $\Lambda F_H = 0$. Now if Z has a boundary the same discussion goes through when we restrict attention to metrics with a fixed boundary value. For the expression (12) can then be written as a boundary integral

$$\int_{\partial Z} \mathrm{Tr}(\eta_1 \nabla_\nu \eta_2 - \eta_2 \nabla_\nu \eta_1),$$

where ∇_ν denotes the normal component of ∇_H on the boundary, and this vanishes for variations η_i which are zero on ∂Z.

The indirect definition of these functionals (which are related to determinants of differential operators) makes them rather mysterious; we refer to refs. [7,8,26] for further discussion. We wish here to examine the case when E is the trivial bundle, and we regard our metrics as maps from Z to the homogeneous space $\mathcal{H} = G^c/G$. In this section we consider only the case of one complex variable and suppose that Z is a domain in $\mathbb{C}$. We compare the Hermitian Yang-Mills condition $\bar{\partial}(H^{-1}\partial H) = 0$, with the harmonic map equation, the Euler-Lagrange equations for the energy functional

$$\mathcal{E}(H) = \int |H^{-1}\nabla H|_H^2 = \int \mathrm{Tr}(H^{-1}\nabla H)^2.$$

A short calculation shows that the harmonic map equation takes the form $\bar{\partial}(H^{-1}\partial H) - \partial(H^{-1}\bar{\partial} H) = 0$. The two equations differ by the term

$$H^{-1}\partial H H^{-1}\bar{\partial} H + H^{-1}\bar{\partial} H H^{-1}\partial H.$$

This is the commutator $[\eta_x, \eta_y]$ of the endomorphisms $\eta_1 = H^{-1}\nabla_x H$, $\eta_2 = H^{-1}\nabla_y H$ representing the derivative of H in the x, y directions in $Z \subset \mathbb{C}$. Now let ϕ be the canonical G^c-invariant three-form on the homogeneous space $\mathcal{H}$. At the identity this is represented by

$$\phi(\xi_1, \xi_2, \xi_3) = (\mathrm{Tr})(\xi_1\ [\xi_2, \xi_3]).$$

We consider a reference map H_0, with boundary value f. For any other map H with this boundary value we choose an arbitrary extension $\overline{H} : Z \times [0,1] \to \mathcal{H}$ with $\overline{H}(z,0) = H_0, \overline{H}(z,1) = H$ and $\overline{H}(z,t) = f(z)$ for $z \in \partial Z$ and we set

$$I(H) = \int_{Z\times[0,1]} \overline{H}^* \phi.$$

This is independent of the extension $\overline{H}$, since ϕ is closed and $\mathcal{H}$ is contractible, so we can regard I as a functional on the space of maps from Z to $\mathcal{H}$ with the given boundary value. The variation of I can be written

$$\delta I = \int_Z \mathrm{Tr}[H^{-1}\delta H\ (H^{-1}\partial H H^{-1}\overline{\partial} H + H^{-1}\overline{\partial} H H^{-1}\partial H)],$$

so we see that the Hermitian Yang–Mills equation, in this set up, is the Euler–Lagrange equation for the functional $\mathcal{E} + I$. Indeed $\mathcal{E} + I$ is just a local expression for the functional R. The merit of this expression is that it displays a link with the Wess–Zumino–Witten (WZW) action, which has been studied in conformal field theory [29,30]. The WZW action is a perturbation of the harmonic map energy, for maps from a two-dimensional domain Z into a Lie group G, obtained by the procedure above using the canonical three-form ϕ' on G. We fix a reference map $g_0 : Z \to G$ and for any other map g with the same boundary value we choose an extension $\overline{g} : Z \times [0,1] \to G$. Then define

$$\mathcal{E}_{\mathrm{WZW}}(g) = \int_Z |g^{-1}\nabla g|^2 + \int_{Z\times[0,1]} \overline{g}^*(\phi').$$

The Euler–Lagrange equations for this WZW action are $\overline{\partial}(g^{-1}\partial g)$, which have just the same form as the Hermitian Yang–Mills equation except now the independent variable takes values in the group G rather than the homogeneous space $\mathcal{H} = G^c/G$. (The space G^c/G is the non-compact dual of G, regarded as a symmetric space.) More precisely, since ϕ' is not exact in G, the WZW action is only defined up to a multiples of a period of ϕ', but this does not affect the local discussion.

To extend this analogy we consider the solutions of the two Euler–Lagrange equations. First we consider the Hermitian Yang–Mills equation, $\overline{\partial}(\partial H H^{-1})$

$= 0$, for a metric on a flat bundle. This is, in a sense, a trivial equation since we know that the general local solution has the form $H = u^* u$, where u is a *holomorphic* map into G^c. More generally, if we consider the complexified version of this: the equation $\overline{\partial}(g^{-1}\partial g) = 0$ where g takes values in the group G^c, then the general local solution has the form $g = vu$, where u is holomorphic and v is anti-holomorphic. The WZW equation for maps into the original group G then reduces to the algebraic condition $u^* u = (v^* v)^{-1}$ for holomorphic u and anti-holomorphic v. The relation with the Hermitian Yang–Mills case appears more vividly if we consider the WZW equation for a two-dimensional Lorentzian space. This is obtained from the same Lagrangian but using the Lorentzian inner product to define the harmonic maps energy $|dg\, g^{-1}|^2$. The Euler–Lagrange equations are $\partial_+ (g^{-1}\partial_- g) = 0$, where $\partial_+ = \partial/\partial x + \partial/\partial t$, $\partial_- = \partial/\partial x - \partial/\partial t$. This is a trivial equation, in the same sense as the Hermitian Yang–Mills equation above, in that the general solution can be immediately written down: $g(x,t) = u(x+t)v(x-t)$, where u and v map into G.

We will not push this parallel between the Hermitian Yang–Mills theory and the WZW action (which may well exist, in some form, already in the physics literature) any further. From the differential-geometric point of view it provides a useful new slant on the HYM equations. To illustrate this we will describe a link with constant mean curvature surfaces in hyperbolic space. In this case we take our group G to be $SU(2)$, so $G^c = SL(2,\mathbb{C})$ and the quotient space $\mathcal{H}$ (of metrics with determinant 1) is a model for hyperbolic space of three dimensions. The three-form ϕ on $\mathcal{H}$ is just the volume form, and so the integral $I(H)$ is the (signed) volume enclosed between the two surfaces $H(Z)$, $H_0(Z)$ in the three-space. In particular, this term is independent of the parametrisation of the surface $H(Z)$—it only depends on the map H through its image. Let us shift then to a Lagrangian $\mathcal{L}$ on unparametrised surfaces $\Sigma \subset \mathcal{H}$, with fixed boundary values and using a reference surface Σ_0. We put

$$\mathcal{L}(\Sigma) = \text{Area}(\Sigma) + I(\Sigma),$$

where $I(\Sigma)$ is defined by the volume as above. It is obvious geometrically that the critical points of $\mathcal{L}$ are just the surfaces of constant mean curvature 1. On the other hand, the same relationship holds between this functional and the Hermitian Yang–Mills function $\mathcal{E} + I$ as between the usual area and the harmonic maps energy: $\mathcal{L}(\text{Image}(H))$ is a lower bound for $\mathcal{E} + I$, which is attained precisely when the map H is conformal. It follows then that the constant mean curvature surfaces in $\mathcal{H}$ are just the images of maps $H : Z \to \mathcal{H}$ which are both conformal and satisfy the Hermitian Yang–Mills equations $\overline{\partial}(H^{-1}\partial H) = 0$. We know that the general solution of the latter have the form $H = u^* u$, where $u : Z \to SL(2,\mathbb{C})$ is holomorphic. We now observe that if V is a real Euclidean space and $D : \mathbb{C} \to V \otimes \mathbb{C}$ is a complex linear map then the composite of D with the linear projection from $V \otimes \mathbb{C}$ to V is conformal if and only if the image of D is a null vector for the complexified quadratic form on $V \otimes \mathbb{C}$. Applying this to the derivative of u we see finally that constant mean curvature surfaces in hyperbolic three-space $SL(2,\mathbb{C})/SU(2)$ are obtained as the projections of null holomorphic curves in $SL(2,\mathbb{C})$, i.e., curves whose tangent vectors are null for the invariant quadratic form on $SL(2,\mathbb{C})$. This is precisely the description due to Bryant [5], extending the Weierstrass construction for minimal surfaces in $\mathbb{R}^3$.

3.4. HITCHIN'S EQUATIONS

A interesting extension of our main existence theorem is obtained by adding an auxiliary "Higgs field", as done by Hitchin in ref. [15]. We have seen that our main theorem is an analogue, for manifolds with boundary, of the Narasimhan and Seshadri theorem, which identifies stable bundles over a closed surface with flat bundles. Hitchin extends the Narasimhan and Seshadri theory by considering the equations, for a unitary connection A on a bundle E over a closed Riemann surface, and a Higgs field $\phi \in \Omega^{1,0}(\operatorname{End} E)$:

$$\overline{\partial}_A \phi = 0, \qquad F_A + [\phi, \phi^*] = 0. \tag{13}$$

Notice that these equations are conformally invariant; they do not depend on a metric on Σ. chin (and, in more generality, Corlette [6] and Simpson [26]) showed that, on a closed surface, the moduli space M^c of solutions of this equation had two different descriptions:

(i) as a moduli space of "stable pairs" (E, ϕ) where E is a holomorphic bundle and ϕ is a holomorphic section of $\operatorname{End} E \otimes K_\Sigma$;

(ii) as a space of conjugacy classes of irreducible representations of $\pi_1(\Sigma)$ in G^c.

In the first description M^c appears naturally as a completion of the cotangent bundle of the moduli space $M = M_\Sigma$, and in the second description as a complexification of M. These two descriptions depend on two existence theorems, and we will now explain how to obtain analogues of these results in which, following the scheme of section 3.1, the space M is replaced by the loop group ΩG. The proofs will involve, of course, the solution of appropriate boundary value problems.

We consider first Hitchin's equations on a fixed holomorphic bundle E with a fixed holomorphic section ϕ of $\operatorname{End} E \otimes K$. For simplicity we will just work here with the case when the base manifold is the disc D, and fix a holomorphic trivialisation, so a metric on E is a map from D to $\mathcal{H}$ as before and $\phi = \psi\, dz$ where ψ is a holomorphic matrix-valued function, or more abstractly a holomorphic map from D to the Lie algebra of G^c. We consider the Hitchin equation

$$F_H + [\phi, \phi^{*H}] = \overline{\partial}(H^{-1}\partial H) + [\phi, H^{-1}\phi^* H] = 0 \tag{14}$$

as an equation for the map $H : D \to \mathcal{H}$. We suppose that ϕ is smooth up to the boundary. The first result we need is then

Theorem 2. *For any metric f on ∂D, and any ϕ, there is a unique solution H to the Hitchin equation (14), with $H = f$ on ∂D.*

Of course, if $\phi = 0$ this just reduces to our main theorem, in the case of the disc. This extension can be proved by a straightforward modification of either of the methods used in section 2. Once again, the result is almost contained in the work of Simpson [26]. We will be content here to explain how the two key properties of section 2.1 extend to the Hitchin equations.

The calculations here can most easily be understood by going up to four dimensions. Hitchin's equations can be obtained as the dimension reduction of the instanton equations, for translation-invariant solutions. We write our Higgs field as $\phi = \psi\, dz = \frac{1}{2}(\psi_1 + \psi_2)\, dz$, where ψ_1, ψ_2 are skew-adjoint sections of $\operatorname{End} E$. Then, taking standard coordinates p_1, p_2 on $\mathbb{R}^2$, we consider the connection

$$\tilde{A} = A_x\, dx + A_y\, dy + \psi_1\, dp_1 + \psi_2\, dp_2$$

over $Z \times \mathbb{R}^2 \subset \mathbb{C} \times \mathbb{R}^2 \cong \mathbb{R}^4$. It is easy to check that (A, ϕ) satisfy Hitchins equations if and only if $\tilde{A}$ is an instanton. Similarly, eq. (14) for a metric H over Z goes over to the HYM equation for a translation-invariant metric $\tilde{H}$. Thus we obtain immediately from our calculations that the linearisation of eq. (14) about a solution H is given by $\Delta_{\tilde{H}}\rho = 0$. But this can be written back on Z as $(\Delta_H + \phi^{*H}\phi)\rho = 0$. The operator $\Delta_H + \phi^{*H}\phi$ is self-adjoint and positive on sections ρ which vanish on the boundary, so we can use the implicit function theorem just as before to deform solutions. In the same way, if H and K are two solutions to (14) we extend them to HYM solutions $\tilde{H}, \tilde{K}$, so $\sigma(\tilde{H}, \tilde{K}) = \sigma(H, K)$, and the inequality $\Delta\sigma(H, K) \leq 0$ follows from that for HYM solutions. Using these remarks we can obtain theorem 2 by copying the proofs of section 2.

We now move on to the analogue of Hitchin's second main result. Let (A, ϕ) be a solution of Hitchin's equations and consider the G^c connection $A + \phi + \phi^*$ on E. The curvature of this connection is

$$F(A + \phi + \phi^*) = F(A) + \overline{\partial}_A\phi + \partial_A(\phi^*) + [\phi, \phi^*],$$

which vanishes. So E has a flat G^c structure, which may be trivialised over the disc. The metric on E appears as a map H from D to $\mathcal{H}$, and in this trivialisation Hitchin's equations are the condition that H be a *harmonic* map, with respect to the invariant metric on $\mathcal{H}$ (see refs. [6,9]). The other existence theorem we need involves the Dirichlet problem for harmonic maps.

Proposition 3. *For any map $f : \partial D \to \mathcal{H}$ there is a unique harmonic map $h : D \to \mathcal{H}$ with $h = f$ on ∂D.*

This is a special case of the general results of Hamilton [14], for maps into spaces of negative curvature. (The space $\mathcal{H}$ is the Riemannian product of a Euclidean space and a space of strictly negative curvature.) Note that Hamilton's proof, using the heat equation method and following that of Eells and Sampson for closed manifolds, is a model for the heat equation technique we have applied in section 2 to the Hermitian Yang–Mills problem.

In sum these solutions of these two boundary value problems give us two ways of describing the solutions of Hitchin's equations over the disc, and we will now unravel the threads between them. Let X be the space of equivalence classes of solutions of Hitchin's equations on a G bundle E over the disc, with a G trivialisation over the boundary. Any solution endows E with a holomorphic structure, which we may trivialise. Then, as in section 3.1, the boundary data yield a map λ from ∂D to G^c. The Higgs field is represented by a holomorphic map $\psi : D \to \mathrm{Lie}(G^c)$. The choice of holomorphic trivialisation goes over to the action of the group $\mathrm{Hol}(D, G^c)$. The group action is: $\lambda \mapsto \gamma\lambda$, $\psi \mapsto \gamma\psi\gamma^{-1}$ for $\gamma \in \mathrm{Hol}(D, G^c)$. Our first existence theorem is the assertion

$$X \cong \frac{LG^c \times \mathrm{Hol}(D,\ \mathrm{Lie}(G^c))}{L^+G^c}.$$

The right hand side of this equivalence is a complex vector bundle $\mathcal{V}$ over LG^c/L^+G^c , which we have identified in section 3.1 with the loop space ΩG. The fibres are modelled on the complex vector space $V = \mathrm{Hol}(D, \mathrm{Lie}(G^c))$. We would like to identify $\mathcal{V}$, and hence X with the cotangent bundle of ΩG, or more precisely with a dense sub-bundle of the cotangent bundle. (This complication arises from the infinite dimensionality of the space, and we do not want to get involved in technicalities of topological vector spaces.) That is to say, we want to define a complex bilinear pairing of the fibres $\mathcal{V}_\gamma \times (T\Omega G)_\gamma \to \mathbb{C}$, which is non-degenerate, in that no non-zero vector in $\mathcal{V}_\gamma$ annihilates all of $T\Omega_\gamma$. We consider first the fibres over the identity. We have a bilinear form on the space of maps from S^1 in $\mathrm{Lie}(G^c)$:

$$\langle f, g \rangle = \int_{\mathrm{S}^1} \mathrm{Tr}\, f g \,\mathrm{d}z,$$

and the sub-space of boundary values of holomorphic maps is isotropic, by Cauchy's theorem. So we get an induced pairing on the quotient:

$$V \times \frac{\mathrm{Maps}(S^1, \mathrm{Lie}(G^c))}{V} \to \mathbb{C}.$$

This is non-degenerate and is the required pairing, since V is the fibre of $\mathcal{V}$ over 1 and the second term is the tangent space of the quotient $T\Omega G = T(LG^c/L^+G^c)$.

Then it is easy to check that this pairing transforms properly under the action of LG^c, so that it can be extended to all of the fibres of $\mathcal{V}$ by translation. This description of $X \subset T^*\Omega G$ is of course the analogue of Hitchin's first description of M^c.

Going on to the other description, any solution (A, ϕ) endows E with a flat G^c structure, and the solution can then be recovered, using our second existence theorem, from a map $f : \partial D \to G^c$. A G-trivialisation over the boundary corresponds to a lift of this map to G^c. The change of flat trivilialisation goes over to the natural action of G^c. We see then that our second existence theorem is the assertion that $X \cong LG^c/G^c = \Omega G^c$, the space of based loops in the complexified group G^c. So we have two descriptions of the same space X, which are precise analogues of Hitchin's two descriptions of the space M^c in the case of a closed surface.

We may read off some interesting facts from these alternative descriptions, just as in Hitchin's work. For example, using the first description, we see a natural $\mathbb{C}^*$ action on X, in which we multiply the Higgs field ϕ by a complex scalar, i.e., the action of the scalars on the fibres of the vector bundle $\mathcal{V}$. The fixed point set of this action is the zero section, the copy of ΩG embedded inside $X = \mathcal{V} = \Omega G^c$.

We also see that X has a hyper-Kähler structure. This follows just the same pattern as the four-dimensional case discussed above: we can regard it as a variant of the four-dimensional case if we interpret the solutions of Hitchin's equations over D as invariant instantons on $D \times \mathbb{R}^2$, in the manner above. The different complex structures are visible in the two descriptions of X: one, I say, is the natural complex structure on the vector bundle $\mathcal{V}$, and this is preserved by the $\mathbb{C}^*$ action. Another, J say, is the natural complex structure on the complexification ΩG^c. This is not preserved by the circle action, and conjugation by this action generates a family of complex structures, one of which is the third element K of the hyper-Kähler triple. All of this is exactly parallel to Hitchin's discussion of the space M^c.

We shall now find some finite dimensional manifolds, with induced hyper-Kähler metrics, inside the infinite dimensional space X. Recall that if H is any Lie group there is a circle action on the based loop space ΩH with the action of $\lambda \in S^1$ given by

$$[\lambda(\gamma)](z) = \gamma(\lambda z)\gamma(\lambda)^{-1}, \quad \gamma \in \Omega H, z \in S^1.$$

A loop γ is fixed by this action if $\gamma(\lambda z) = \gamma(\lambda)\gamma(z)$, i.e., if it is a closed one-parameter sub-group in H. (More geometrically, if we think of ΩH as the framed flat H connections on the disc then the action is just the obvious one induced by rotations of the disc. The one-parameter sub-group appears in this picture as the action of S^1 on the central fibre of a bundle over the disc.) Thus the fixed point set falls into different components, labelled by the conjugacy class of the representation $S^1 \to G$, and each component is a copy of an adjoint orbit of G in its Lie algebra—the orbit of the tangent vector to the one-parameter sub-group.

We wish to consider this picture in two cases. The well-known case is when $H = G$ is a compact group. The adjoint orbits then have homogeneous Kähler structures, with Kähler form the Kostant–Kirilov form (we use the Killing form to identify adjoint and co-adjoint orbits), and these are isometrically embedded in ΩG as the fixed points of the circle action. More precisely, the "algebraic " orbits, for which the Kähler form is integral, embed in the loop space.

The more novel case is when $H = G^c$ is the complexified group and the fixed point set consists of certain complex adjoint orbits. (We should be careful not to confuse this circle action with the restriction of the $\mathbb{C}^*$ action above.) We identify ΩG^c with the space X of framed solutions to Hitchin's equations. Then the circle action is again induced by rotations of the disc, and it is easy to see that these preserve the metric and all three complex structures on X. Thus the fixed point set inherits a hyper-Kähler structure. Hyper-Kähler structures on complex co-adjoint orbits have been found previously by Kronheimer [19] (who considers a larger family of orbits) and we shall now see that our construction recovers his metrics.

Kronheimer obtains hyper-Kähler metrics in ref. [19] from the solutions of Nahm's equations

$$\mathrm{d}T_i/\mathrm{d}t = [T_j, T_k], \quad (i,j,k) \text{ cyclic},$$

for Lie(G)-valued functions T_1, T_2, T_3 of a real variable t. He considers the set N of solutions over the half-line $(-\infty, 0]$ which converge as $t \to -\infty$ to a fixed commuting triple τ_1, τ_2, τ_3. He shows that, under appropriate non-degeneracy conditions, N has a hyper-Kähler metric and that the map which assigns $T_1(0) + \mathrm{i}T_2(0)$ to a solution identifies N with the adjoint orbit, under G^c, of $\tau_1 + \mathrm{i}\tau_2$. Now solutions of Nahm's eqations can be viewed as instantons which are invariant under translation in three variables, so Kronheimer's solutions can be identified with certain invariant instantons on $(-\infty, 0] \times \mathbb{R}^3$, or, dividing by a discrete set of translations in one direction, instantons on $(-\infty, 0] \times S^1 \times \mathbb{R}^2$ which are invariant under rotation in the S^1 factor and translation in the $\mathbb{R}^2$ factor. But we know that instantons invariant under two translations go over to solutions of Hitchin's equations, so Kronheimer's solutions yield S^1-invariant solutions of Hitchin's equation over the cylinder $(-\infty, 0] \times S^1$. Finally, using the conformal invariance of Hitchin's equations we interpret these as rotation-invariant solutions over the punctured disc $D \setminus \{0\}$. Explicitly, from a solution T_i of Nahm's equations we can write down a solution

$$\begin{aligned} A_r &= 0 \\ A_\theta &= T_1(\log r) \\ \phi &= \big((T_2 + \mathrm{i}T_3)(\log r)\big) z^{-1}\,\mathrm{d}z, \end{aligned}$$

to Hitchin's equations. In general these solutions will be singular at 0. The singularity is removable if $\tau_2 = \tau_3 = 0$ and τ_1 satisfies the integrality condition $\exp(\tau_1) = 1$. We see then that in this case the solutions considered by Kronheimer match up precisely with the rotation-invariant solutions over the disc, and it is straightforward to check that the hyper-Kähler metrics are the same.

It would be interesting to extend this discussion to include the other solutions found by Kronheimer. For this one would have to consider singular solutions of Hitchin's equation, in which ϕ has a pole and A has a non-trivial limiting holonomy around the singularity. These kinds of singularities arise naturally in various contexts and have been considered by a number of authors recently [4,20].

3.5. CONNECTIONS OVER THE BOUNDARY

We will now discuss our main result, in the case of two complex variables, from a different perspective. In general, if W is an oriented Riemannian four-manifold, with boundary Y, it is reasonable to hope that any connection over Y extends to a solution of the full (second order) Yang–Mills equations over W. For connections close to the trivial connection one can see this using the implicit fuction theorem to pass to the linearised problem. For the first order instanton equations the picture is quite different. The sub-set L_W of boundary values of instantons has infinite codimension in the space $\mathcal{B}_Y$ of gauge equivalence classes of connections over Y. In fact it is natural to think of L_W as having roughly "half the dimension" of $\mathcal{B}_Y$, a point of view which underlies the theory of Floer homology [2,11]. We will now see what information our results give about these sub-manifolds in the case when $W = Z$ is a complex Kähler surface.

To interpret our results in this way we introduce two new notions, adapted to the induced CR structure on the boundary $Y = \partial Z$. This is a complex line bundle $V = TY \cap ITY \subset TY$. We define a *partial connection* on a vector bundle E over Y to be a linear map

$$\nabla^{\mathrm{b}} : \Gamma(E) \to \Gamma(E \otimes_{\mathbb{R}} V^*)$$

satisfying a Leibnitz rule

$$\nabla^{\mathrm{b}}(fs) = d^{\mathrm{b}} f s + f \nabla^{\mathrm{b}} s,$$

where d^{b} is the projection of the exterior derivative to V. If E has a fibre metric we can consider partial connections compatible with this, in the obvious way (and more generally we could define partial connections on principal bundles with any structure group). Note that the definition of a partial connection does not use the complex structure on V. We now use this complex structure to decompose $V^* \otimes_{\mathbb{R}} \mathbb{C}$ into linear and anti-linear pieces $V^{1,0} \oplus V^{0,1}$, say. The projection of d^{b} to $V^{0,1}$ yields the $\overline{\partial}_{\mathrm{b}}$ operator of the complex structure—the operator which annihilates the restriction of any holomorphic function on Z. We define a CR structure on a bundle E over Y to be an operator

$$\overline{\partial}_{\mathrm{b}}^{E} : \Gamma(E) \to \Gamma(E \otimes_{\mathbb{C}} V^{0,1}),$$

such that $\overline{\partial}_{\mathrm{b}}^{E}(fs) = \overline{\partial}_{\mathrm{b}}(f)s + f\overline{\partial}_{\mathrm{b}}^{E} s$. A holomorphic bundle over Z restricts to a bundle with a CR structure on ∂Z, and any partial connection over Y yields a CR structure as its $(0,1)$ component. Just as for holomorphic bundles, on a bundle with a metric and CR structure there is a unique compatible partial connection. Obviously any genuine connection over Y, and in particular any connection over Z, defines a partial connection by restriction.

We will consider the following boundary value problem for the first order instanton equations. When can a given partial connection over Y be extended to an instanton over Z ? A naive dimension count indicates that this is a sensible problem. For a G-connection over Y is specified locally by three dim G functions, but the gauge group accounts for dim G of these, so we think of a connection modulo gauge as being specified by $2 \times \dim G$ functions. Similarly, a partial connection is specified by $2 \times \dim G$ functions, and we again subtract dim G for the gauge group, so we expect the partial connections modulo equivalence to be specified by dim G functions. Thus, when account is taken of the gauge symmetry, the partial connection fixes roughly half the data on the boundary, which is appropriate to a boundary value problem for a first order equation.

Let us say that a partial connection ∇^{b} on a bundle E is "holomorphically trivial" if the corresponding CR structure $\overline{\partial}_{\mathrm{b}}^{E}$ is equivalent to the trivial structure, i.e., if there is a trivialisation of E which takes $\overline{\partial}_{\mathrm{b}}^{E}$ to a sum of copies of $\overline{\partial}_{\mathrm{b}}$. Then we have

Theorem. *Let Z be a pseudoconvex domain in $\mathbb{C}^2$ with smooth boundary Y and a Kähler metric which is smooth up to the boundary. A unitary partial connection over Y extends to an instanton (on a topologically trivial bundle) over Z if and only if it is holomorphically trivial, and the extension is then unique up to gauge equivalence.*

This is essentially a reformulation of our main result. First, if a partial connection ∇^{b} extends to a (topologically trivial) instanton, the corresponding CR structure $\overline{\partial}_{\mathrm{b}}^{E}$ extends to a holomorphic bundle over Z, also topologically trivial. But we have explained in section 3.1 that any such holomorphic bundle is trivial so *a fortiori* $\overline{\partial}_{\mathrm{b}}^{E}$ is holomorphically trivial. In the other direction, suppose a partial connection ∇^{b} on a Hermitian bundle E over Y is holomorphically trivial and choose a trivialisation of E in which $\overline{\partial}_{\mathrm{b}}^{E} = \overline{\partial}_{\mathrm{b}}$. The Hermitian metric on E is represented in this trivialisation by a map h from Y to $\mathcal{H}$. The partial connection is determined by its $(0,1)$ part and the metric, so ∇^{b} has "connection matrix" $h^{-1}\partial_{\mathrm{b}}h$. We now apply our main theorem to the trivial holomorphic bundle over Z, to find a metric $H : Z \to \mathcal{H}$ which satisfies the Hermitian Yang–Mills equation and with $H = h$ on the boundary. Then $H^{-1}\partial H$ restricts to $h^{-1}\partial_{\mathrm{b}}h$, and this connection gives the desired extension of the original partial connection.

Finally, suppose we have two instantons A_1, A_2 on bundles E_1, E_2 extending the same partial connection ∇^{b}. The instantons define trivial holomorphic structures on E_1, E_2. In holomorphic trivialisations the isomorphism intertwining the partial connections over the boundary becomes a matrix-valued function χ with $\overline{\partial}_{\mathrm{b}}\chi = 0$. By the extension theorem for pseudoconvex domains [18] this extends to a holomorphic function over all of Z. The determinant cannot vanish in Z, since the zero set would have complex codimension 1, and would be forced to meet the boundary. So we conclude that the intertwining map extends to a holomorphic isomorphism of E_1 and E_2 over the boundary. Comparing the metrics by this isomorphism then, we may suppose that $E_2 = E_1$ is the trivial holomorphic bundle, and we have two Hermitian metrics H and K satisfying the HYM equation and with $H^{-1}\partial_{\mathrm{b}}H = K^{-1}\partial_{\mathrm{b}}K$ on the boundary. But this means that the endomorphism $\eta = H^{-1}K$ is "covariant constant" with respect to the partial connection on the boundary, and its eigenspaces decompose the trivial bundle over the boundary into a direct sum of orthogonal factors $E^{(i)}$, with $H = \lambda_i K$, say, on $E^{(i)}$ for constant λ_i. (This follows from the fact that a function f on Y with $d^b f = 0$ is constant.) We apply the extension theorem again to extend η as a holomorphic function over Z: The eigenvalues of η are constant on the boundary so also in Z, and hence the eigenspace decomposition $\mathcal{O}^N = \bigoplus E^{(i)}$ also extends holomorphically over Z. Then the uniqueness of the solution to the boundary value problem implies that these factors remain orthogonal and $H = \lambda_i K$ on $E^{(i)}$ over Z. Thus the two connections $H^{-1}\partial H$, $K^{-1}\partial K$ are equal over Z.

To close this section we wish to draw attention to another boundary value problem for instantons, which is in a sense a "Neumann problem" to complement the Dirichlet problem that we have studied in this paper. We seek instantons over Z whose curvature vanishes in the two-plane V on the boundary. Thus locally we are solving the equation $\Lambda(\overline{\partial}(H^{-1}\partial H)) = 0$ with boundary condition $\overline{\partial}(H^{-1}\partial H)|_V = 0$. This boundary condition mixes normal and tangential derivatives, since, if the Levi form is non-zero, the restriction of $\overline{\partial}\partial H$ to V involves the first order normal derivative of H. While we are not able to say much about the solutions to this problem, it can be motivated in a number of ways. First, in terms of the geometry of the space $\mathcal{B}_Y$ of equivalence classes of connections over Y. We have a restiction map $\pi : \mathcal{B}_Y \to \mathcal{B}_Y^V$ from connections to partial connections, and a sub-set $T \subset \mathcal{B}_Y^V$ of partial connections which are holomorphically trivial. Our result above says that the boundary values L_Z, in the case when Z is pseudoconvex, form a section of π over T. Since the Levi form of Y is non-degenerate there is another cross-section Σ of π (over all of $\mathcal{B}_Y^V$) consisting of the connections whose curvature in the V plane is zero; that is, any partial connection has a unique extension to a full connection with V curvature zero. To see this we can work locally and choose a pair of vector fields v_1, v_2 spanning V. The curvature of a connection ∇ in the V plane is given by

$$F(v_1, v_2)s = \nabla_{v_1}\nabla_{v_2}s - \nabla_{v_2}\nabla_{v_1}s - \nabla_{[v_1,v_2]}s.$$

For a connection ∇ extending a given partial connection ∇^b the first two terms on the right hand side are determined by ∇^b, so the condition that $F(v_1, v_2) = 0$ determines $\nabla_{[v_1,v_2]}$ in terms of ∇^b. But if the Levi form is non-degenerate $[v_1, v_2]$ generates TY/V, so the connection ∇ is completely determined. In this picture, our "Neumann problem" asks for the intersection of the two sub-manifolds Σ, L_Z.

On the other hand, we can motivate this boundary condition by a variational problem. We suppose that the Kähler metric ω can be written as $i\overline{\partial}\partial\phi$, where ϕ is a function which is positive on Z and vanishes on the boundary. For example, this is the case if Z is the unit ball in $\mathbb{C}^2$, when $\phi(z) = 1 - |z|^2$. We consider the functional

$$S(H) = \int_Z \mathrm{Tr}(F_H \wedge F_H)\phi.$$

For a variation ηH in the metric the variation of S is

$$\delta S = \int_Z \mathrm{Tr}(\overline{\partial}\partial_H \eta \wedge F_H)\phi.$$

One integration by parts expresses this as

$$\int_Z \mathrm{Tr}(\partial_H \eta \wedge F_H) \wedge \overline{\partial}\phi + \int_{\partial Z} \mathrm{Tr}(\partial_H \eta \wedge F)\phi,$$

and the boundary term vanishes since ϕ is zero there. We integrate by parts again to write

$$\delta S = \int_Z \mathrm{Tr}(\eta \wedge F_H \wedge \overline{\partial}\partial\phi) + \int_{\partial Z} \mathrm{Tr}(\eta \wedge F_H \wedge \partial\phi).$$

Note that $\overline{\partial}\partial\phi = -\mathrm{i}\omega$ and $F_H \wedge \omega = \Lambda F_H$ vol., so S is stationary with respect to compactly supported variations if and only if the Hermitian Yang–Mills condition $\mathrm{i}\Lambda F_H = 0$ holds. We consider now two cases: In the first we fix the metric on the boundary, so η vanishes there and the boundary integral is zero. Then the critical points are just the solution of the Dirichlet problem we have considered before, and the function S can be identified with the functional R we discussed in section 3.3. In the other case we allow unrestricted variations η. Then S is stationary if the boundary integral also vanishes, which requires that $F_H \wedge \partial\phi = 0$, but this is precisely the boundary condition $\langle F_H, V\rangle = 0$ that we want.

4. Conclusion

In this paper we have seen how the Dirichlet problem for Hermitian Yang–Mills connections yields pleasant extensions of a number of different results. We have seen that the loop groups ΩG appear as natural companions of the moduli spaces of flat G-connections over Riemann surfaces, with generalisations on the one hand via Hitchin's equations to hyper-Kähler structures on ΩG^{c} which are counterparts of the flat G^{c} connections on a closed surface, and on the other hand to higher dimensions. There are a number of other questions which these ideas suggest, but which we have not pursued further.

Harmonic maps into G^c/G. We have see that the hyper-Kähler structure on the complexified loop space ΩG^c relies on the solution of the Dirichlet problem for harmonic maps from the disc into the symmetric space G^c/G. There is an extensive literature on the construction of harmonic maps into a compact group G, the dual symmetric space of G^c/G, and their relation with Loop groups [24,27]. It would be interesting to know if there were similar explicit constructions for harmonic maps into G^c/G; these might yield explicit formulae for the metric on ΩG^c.

Instantons and holomorphic maps. Atiyah has shown how, following a rather roundabout route, instantons on $\mathbb{R}^4$ may be identified with holomorphic curves in loop groups ΩG [1]. One of the key steps in this identification theorem is the factorisation theorem which we discussed in section 3.1. On the other hand there is a much more transparent relationship between holomorphic maps from one surface S to the moduli space $M(\Sigma)$ of flat connections on another (closed) surface Σ, and the instantons on $S \times \Sigma$. We can see this either by going through holomorphic bundles or by taking an "adiabatic limit" of the instanton equations, when the metric on the product is shrunk in the Σ factor [23]. The affinity between the loop groups and moduli spaces of flat connections suggests that it may be possible to obtain Atiyah's correspondence by some sort of adiabatic limit.

The Riemann mapping theorem. We have seen that the prototype of our Dirichlet problem is equivalent to the factorisation theorem in loop groups, which may be written in the form

$$LG/G = LG^c/L^+G^c,$$

closely analogous to the two descriptions of "flag" manifolds $G/T = G^c/B$. Segal has explained [25] how to extend this analogy, taking in place of G the group of diffeomorphisms of the circle $\mathrm{Diff}(S^1)$. As a substitute for the complexification Segal takes the semigroup $\mathrm{Diff}^c(S^1)$ of maps from S^1 the disc, with a composition law defined by gluing the resulting complex annuli. The analogue of a maximal parabolic sub-group is the set $\mathrm{Hol} \subset \mathrm{Diff}^c(S^1)$ of maps which are the boundary values of holomorphic embeddings of the disc in itself. Then the factorisation formula

$$\mathrm{Diff}(S^1)/\mathrm{PSL}(2,\mathbb{R}) = \mathrm{Diff}^c(S^1)/\mathrm{Hol}$$

is correct, and a moment's thought shows that this is a restatement of the Riemann mapping theorem (for compact domains in $\mathbb{C}$ with smooth boundary). On the other hand we can formulate this statement in another way as the solution of a boundary value problem for a metric of constant curvature, with a boundary of constant geodesic curvature. This is, of course, completely in line with our formulation of the factorisation theorem for loop groups. It is natural to hope that, extending this pattern, there may be some generalisation of the circle of ideas we have discussed in this paper to Riemannian metrics in higher dimensions.

Appendix A. Vector bundles over pseudoconvex domains

In this appendix we will prove that, if Z is a bounded strictly pseudoconvex domain in $\mathbb{C}^2$ with smooth boundary, then any (topologically trivial) holomorphic bundle over $\overline{Z}$, as defined in section 2.1, is trivial. As we mentioned in section 3.1, Grauert proved the corresponding result for bundles over the open manifold: the difficulty is that *a priori* the trivialisations furnished by Grauert's result may not extend to the boundary. The key step in our proof will be to show that a rank-2 vector bundle over $\overline{Z}$ is the restriction of a bundle over a larger domain (and in fact over all of $\mathbb{C}^2$); then Grauert's result, applied to this larger domain, immediately gives the desired conclusion. For simplicity we will assume that all bundles over Z are topologically trivial, as would occur, for example, if Z is homeomorphic to a ball.

Our strategy is to reduce the proof to the linear theory. We have the following general result:

Proposition 4. *If E is a holomorphic bundle over $\overline{Z}$ the Dolbeault cohomology group $H^1(\overline{Z};E)$ is trivial and the holomorphic sections of E over $\overline{Z}$ generate E at each point of $\overline{Z}$.*

We emphasise that in this proposition we are considering data —holomorphic sections and the forms defining the Dolbeault cohomology— which are smooth up to the boundary. The result is essentially a standard application of the solution of the $\overline{\partial}$ Neumann problem [18] and can be proved by combining the methods of ref. [18] (for the case when E is trivial) with the use of a weight function as in ref. [15, section 5.6], to absorb the extra terms coming from the auxiliary bundle.

We will now proceed by induction on the rank of the bundle E. The case of a line bundle reduces immediately to cohomology, since the isomorphism classes of line bundles over $\overline{Z}$ correspond to $H^1(\overline{Z},\mathcal{O}^*)$. We have an exact sequence

$$H^1(\overline{Z};\mathcal{O}) \to H^1(\overline{Z};\mathcal{O}^*) \to H^2(\overline{Z};\mathbb{Z}),$$

in which the first term vanishes by proposition 4, applied to the trivial bundle. So the holomorphic classification of line bundles coincides with the topological classification by the first Chern class.

Now suppose the result is known for rank $n-1 \geq 2$ and let E be a bundle of rank n. Since E is generated by its sections a generic section s of E will not vanish anywhere in $\overline{Z}$, so we can fit E into an exact sequence:

$$0 \to \mathcal{O} \xrightarrow{s} E \to E' \to 0,$$

where E' has rank $n-1$. These extensions are parametrised by the cohomology group $H^1(\overline{Z}; (E')^*)$ which vanishes by our proposition. So the exact sequence can be split, and $E = \mathcal{O} \oplus E'$. Now, by our inductive hypothesis E' and hence also E is trivial.

It remains then to close the gap in the induction with the case when E has rank 2. In this case a generic section s of E has no zeros on the boundary ∂Z, and a finite set of transverse zeros $z_1, \ldots, z_d$ in the interior. We obtain an exact sequence

$$0 \to \mathcal{O} \xrightarrow{s} E \to \mathcal{I} \to 0,$$

where I is the ideal sheaf of the set of points z_i. It is straightforward to adapt the theory descibed in ref. [13, ch. 5] to show that such extensions over $\overline{Z}$ are parametrised by a group A which fits into an exact sequence

$$H^1(\overline{Z}; \mathcal{O}) \to A \to \oplus \lambda_i \to H^2(\overline{Z}; \mathcal{O}),$$

where λ_i is a copy of $\mathbb{C}$ associated to the point z_i. Since $H^1(\overline{Z}; \mathcal{O}) = 0$ the extension class is determined entirely by its image (l_i) in the local part $\oplus \lambda_i$. Now, going backwards, since $H^2(\mathbb{C}^2; \mathcal{O}) = 0$ we can construct an extension

$$0 \to \mathcal{O} \to V \to \mathcal{I} \to 0$$

over the whole of $\mathbb{C}^2$ which maps to the same local extension data (l_i). Hence E is the restriction to $\overline{Z}$ of the bundle V over $\mathbb{C}^2$ and we can apply Grauert's result.

The author is grateful to Peter Kronheimer and Grahame Small for a number of useful suggestions and discussions.

References

[1] M. F. Atiyah, Instantons in two and four dimensions, Commun. Math. Phys. 93 (1984) 437–451.

[2] M. F. Atiyah, New invariants in three and four dimensions, Proc. Symp. Pure Math. (The Mathematical Heritage of Hermann Weyl) 48 (1985) 285–299.

[3] M. F. Atiyah and R. Bott, The Yang–Mills equations over Riemann surfaces, Philos. Trans. R. Soc. London A 308 (1982) 523–615.

[4] O. Biguard, Fibreés paraboliques stables et connections singulières plates, Ecole Polytechnique preprint.

[5] R. L. Bryant, Surfaces of mean curvature 1 in hyperbolic space, Asterisque 154 (1987) 321–347.
[6] K. Corlette, Flat G-bundles with canonical metrics, J. Diff. Geom. 28 (1988) 361–382.
[7] S. K. Donaldson, Anti-self-dual Yang–Mills connections over complex algebraic surfaces and stable vector bundles, Proc. London Math. Soc. 50 (1985) 1–26.
[8] S. K. Donaldson, Infinite determinants, stable bundles and curvature, Duke Math. J. 54 (1987) 231–247.
[9] S. K. Donaldson, Twisted harmonic maps and the self-duality equations, Proc. London Math. Soc. 55 (1987).
[10] S. K. Donaldson and P. B. Kronheimer, *The Geometry of Four-Manifolds* (Oxford U.P., (1990).
[11] A. Floer, An instanton invariant for three-manifolds, Commun. Math. Phys. 118 (1987) 215–240.
[12] H. Grauert, Analytische Faserungen und holomorph-vollständigen Räumen, Math. Annalen 135 (1958) 267–273.
[13] P. Griffiths and J. Harris, *Principles of Algebraic Geometry* (Wiley, New York, 1978).
[14] R. S. Hamilton, *Harmonic Maps of Manifolds with Boundary* (LNM 471) (Springer, Berlin, 1975).
[15] N.J. Hitchin, The self-duality equations over Riemann surfaces, Proc. London Math. Soc. 55 (1987) 59–126.
[16] L. Hormander, *An Introduction to Complex Analysis in Several Variables* (North-Holland, Amsterdam, 1979).
[17] S. Kobayashi, *Differential Geometry of Complex Vector Bundles* (Princeton U.P., Princeton, 1987).
[18] J.J. Kohn, *The Neumann Problem for the Cauchy–Riemann Complex* (Princeton U.P., Princeton, 1972).
[19] P. B. Kronheimer, A hyper-Kähler structure on co-adjoint orbits of a semi-simple complex group, J. London Math. Soc. 42 (1990) 193–208.
[20] P. B. Kronheimer, Embedded surfaces in 4-manifolds in: *Proc. Intern. Congr. of Mathematicians (Kyoto, 1990)*, to appear.
[21] M. S. Narasimhan and C.S. Seshadri, Stable and unitary vector bundles over a compact Riemann surface, Ann. Math. 82 (1965) 540–564.
[22] A. Pressley and G. Segal, *Loop Groups* (Oxford U.P., Oxford, 1986).
[23] D. Salamon and S. Dostoglou, to appear.
[24] G.B. Segal, Loop groups and harmonic maps, in: *Advances in Homotopy Theory*, Eds. Salamon, Steer and Sutherland, Lecture Notes in Mathematics 139 (1988) 153–165.
[25] G. B. Segal, unpublished manuscript.
[26] C. T. Simpson, Constructing variations of Hodge structure using Yang–Mills theory, and applications to uniformisation, J. Am. Math. Soc. 1 (1989) 867–918.
[27] K. K. Uhlenbeck, Harmonic maps into Lie groups (classical solutions of the chiral model), J. Diff. Geom. 19 (1984) 431–452.
[28] K.K. Uhlenbeck and S-T. Yau, The existence of Hermitian Yang–Mills connections on stable bundles over Kähler manifolds, Commun. Pure Appl. Math. 39 (1986) 257–293.
[29] E. Witten, Non-abelian bosonisation in two dimensions, Commun. Math. Phys. 92 (1984) 455.
[30] E. Witten, On holomorphic factorisation of WZW and coset models, IAS Princeton preprint (1991).