ELSEVIER

Journal of Geometry and Physics 27 (1998) 79–98

JOURNAL OF GEOMETRY AND PHYSICS

Degenerating metrics and instantons on the four-sphere

Stuart Jarvis [a,1], Paul Norbury [b,*]

[a] *Merton College, Oxford, UK*

[b] *Mathematics Institute, University of Warwick, Coventry CV4 7AL, UK*

Received 18 February 1997; received in revised form 22 August 1997

Abstract

We give a direct proof of Atiyah's theorem relating instantons over the four-sphere with holomorphic maps from the two-sphere to the loop group. Our approach uses the non-linear heat flow equation for Hermitian metrics as used in the study of Kahler manifolds. The proof generalises immediately to a larger class of four-manifolds. Copyright © 1998 Elsevier Science B.V.

Subj. Class.: Differential geometry

Keywords: Instantons; Four-sphere; Atiyah theorem; Donaldson problem

1. Introduction

It is interesting to both mathematicians and physicists to relate gauge-theoretic constructions over four-manifolds to spaces of holomorphic curves into related manifolds. In physical terms, this amounts to relating the instantons of four-dimensional and two-dimensional theories. One of the earliest results of this type is a theorem of Atiyah that relates Yang–Mills instantons over the four-sphere to holomorphic maps of the two-sphere to the loop group [2].

Theorem 1.1 (Atiyah). *For any classical group G and positive integer k, the following two spaces are diffeomorphic:*

(i) *the parameter space of Yang–Mills k-instantons over S^4 with group G, modulo based gauge transformations,*

(ii) *the parameter space of all based holomorphic maps $S^2 \to \Omega G$ of degree k.*

* Corresponding author. Present address: Department of Mathematics and Statistics, University of Melbourne, Parkville, Vic. 3052, Australia. E-mail: norbs@maths.warwick.ac.uk.

[1] E-mail: jarvis@maths.ox.ac.uk.

PII S0393-0440(97)00067-3

The purpose of this paper is to describe a new isomorphism between the spaces (i) and (ii) of this theorem. Under any such isomorphism, there are interesting relationships between the symmetries of the respective spaces. A description of the particular symmetries that feature in the different isomorphisms would take us too far from the aims of this paper. We will instead settle for a brief comparison confined to this paragraph. In both the isomorphism defined by Atiyah and the one described here, the circular symmetry given by rotating the S^2 of based holomorphic maps $S^2 \to \Omega G$ (infinity is fixed) induces the same symmetry on the space of instantons as the circle of isometries of S^4 given by rotating the first coordinate of $\mathbf{C}^2 \subset S^4$. An extension of the reult described in this paper to include *all* holomorphic maps $S^2 \to \Omega G$ allows the circle symmetry of S^2 to be enlarged to $SO(3)$. (The space of unbased holomorphic maps $S^2 \to \Omega G$ of fixed degree is an infinite dimensionl space that fibres over the loop group with fibres isomorphic to the finite-dimensional instanton spaces.) The space of conformal symmetries of the unit disk $\{|z| \leq 1\}$ that fix $z = 1$ act on the boundary circle and hence on the loop group. This induces an action on the space of holomorphic maps $S^2 \to \Omega G$ which corresponds via the isomorphism of this paper to an action on the space of instantons induced from a family of conformal tranformations of S^4. Using Atiyah's isomorphism the two previous examples of symmetries do not arise from conformal transformations of S^4. Instead, Atiyah's isomorphism gives rise to other symmetry comparisons, including an interesting involution on the space of holomorphic maps $S^2 \to \Omega G$, induced from the involution on S^4 obtained by swapping coordinates in $\mathbf{C}^2 \subset S^4$. For the analogous study of a new isomorphism of the moduli space of monopoles with rational maps and the interesting symmetries that arise see [12,13,20].

Atiyah's proof of his theorem relies on algebraic geometry which uses the special form of the twistor space of the four-sphere. One can view this paper as presenting an alternative proof of Atiyah's theorem more in line with the direct methods used by Dostoglou and Salamon [9,10] in their proof of a relationship between the instantons over a large class of topologically more interesting four-manifolds and pseudo-holomorphic curves inside particular Kahler manifolds. We essentially flow directly from the holomorphic map into the loop group to the instanton over the four-sphere. This method has the advantage that it generalises to a larger class of four-manifolds and loop groups. It also fits in with the homotopy theorists' intuition regarding the respective configuration spaces.

Choose local coordinates $\{(w, z) = (u + iv, x + iy)\}$ for $S^2 \times D$. The map $f : S^2 \to \Omega U(n)$ is holomorphic when $f^{-1}\partial_{\bar{w}} f : \{|z| = 1\} \to \mathbf{gl}(n, \mathbf{C})$ extends to a holomorphic map from the disk $\{|z| \leq 1\}$ to $\mathbf{gl}(n, \mathbf{C})$ for each $w \in S^2$. Put η equal to this extension. Over $S^2 \times D$ define the connection

$$A_f = \eta \, d\bar{w} - \bar{\eta}^{\mathrm{T}} \, dw, \tag{1}$$

so A_f is flat on each $\{w\} \times D$. Furthermore,

$$[\partial_u^A, \partial_x^A] = [\partial_v^A, \partial_y^A], \quad [\partial_u^A, \partial_y^A] = -[\partial_v^A, \partial_x^A], \quad [\partial_x^A, \partial_y^A] = 0, \tag{2}$$

which resembles the anti-self-dual (ASD) equations with respect to the product Kahler metric on $S^2 \times D$:

$$[\partial_u^A, \partial_x^A] = [\partial_v^A, \partial_y^A], \qquad [\partial_u^A, \partial_y^A] = -[\partial_v^A, \partial_x^A],$$
$$[\partial_x^A, \partial_y^A] = \left(\frac{1+u^2+v^2}{1-x^2-y^2}\right)^2 [\partial_u^A, \partial_v^A], \tag{3}$$

where we are using the round metric and the hyperbolic metric on S^2 and D, respectively. It so happens that $S^2 \times D \cong S^4 - S^1$ and the product metric is conformally equivalent to the round metric on S^4. That means that (2) also resembles the ASD equations over S^4. Notice that if we change the product metric non-conformally so that the area of the two-sphere goes to infinity, or equivalently so that the area of the disk goes to zero, then the third of the ASD equations tends to the flat third condition of (2). See Remark 1.3 (iii) below.

Atiyah remarked that his proof, which uses a result of Donaldson [6], merely gives existence without a direct means of associating an instanton to a holomorphic map. In [7], Donaldson suggested that there ought to be some type of adiabatic limit proof that avoids Atiyah's roundabout route. The following theorem addresses these two comments and gives an alternative proof of Atiyah's theorem.

Theorem 1.2. *For each based holomorphic map $f : S^2 \to \Omega U(n)$, there exists a unique gauge equivalence class of ASD connections on a framed $U(n)$-bundle over S^4 and a canonical representative A that is in some sense close to A_f. This correspondence defines a diffeomorphism between the respective moduli spaces.*

Remark 1.3.

(i) The sense in which the connections are close will be made clear later. We will not actually prove that the connections are close, rather it will be sufficient to prove that Hermitian metrics associated to the connections are uniformly close. The precise estimate is given in Lemma 4.7.

(ii) The techniques in this paper generalise to any compact group. For the orthogonal and symplectic groups, we can deduce the corresponding result from the unitary case, rather than using the more general construction. This is because, as subgroups of $U(n)$, the extra structure determined by $O(n)$ and $Sp(n)$ is quite explicit. The objects we use inherit the extra structure by their uniqueness properties.

(iii) We can think of Eqs. (2) as describing the ASD equations with respect to a metric that degenerates in the disk factor. Theorem 1.2 essentially describes the limit of the moduli space of instantons as we stretch the metric on S^4 so that the area of the disk goes to zero.

(iv) The connections invariant under the natural circle action on S^4 can be identified with hyperbolic monopoles. The results in this paper generalise some parts of [13,14].

The novelty of the decomposition $S^4 = S^1 \times B^3 \cup S^2 \times D^2$ rather than the more usual picture of S^4 as $\mathbf{CP}^2$ with a divisor collapsed allows us to generalise the result. We can replace the loop group and S^4, respectively in Theorem 1.2 by $LGL(n, \mathbf{C})/L_\Sigma^+GL(n, \mathbf{C})$ and $X_\Sigma = S^1 \times B^3 \cup S^2 \times \Sigma$ for a Riemann surface Σ with $\partial\Sigma = S^1$. Precise definitions are given in Section 5.

Theorem 1.4. *The moduli space of instantons on a framed $U(n)$-bundle over X_Σ is diffeomorphic to the space of based holomorphic maps from S^2 to $LGL(n, \mathbf{C})/L^+_\Sigma GL(n, \mathbf{C})$.*

2. Metrics on the four-sphere

In order to define a global metric over S^4 we shall use the identification

$$S^4 \cong \mathbf{HP}^1 = \mathbf{H}^2/\mathbf{H}^*,$$

where the non-zero quaternions $\mathbf{H}^*$ act on the right of $\mathbf{H}^2$. We can cover $\mathbf{HP}^1$ with two affine complex coordinate patches

$$\{(q, 1) | q = a + bj\} \cup \{(1, q^{-1}) | q^{-1} = A + Bj\}.$$

The round metric is then given by

$$\mathrm{d}s^2 = \frac{4(\mathrm{d}\bar{a}\,\mathrm{d}a + \mathrm{d}\bar{b}\,\mathrm{d}b)}{(1 + |a|^2 + |b|^2)^2} = \frac{4(\mathrm{d}\bar{A}\,\mathrm{d}A + \mathrm{d}\bar{B}\,\mathrm{d}B)}{(1 + |A|^2 + |B|^2)^2}.$$

Consider $S^2_\infty \subset \mathbf{HP}^1$ given by $\{b = 0\}$ and $S^1_0 \subset \mathbf{HP}^1$ given by $\{a = 0, |b| = 1\}$. We have notated these two submanifolds with subscripts since we will refer to them again. The open submanifold $S^4 - S^1_0$ can be identified with a trivial disk bundle over S^2_∞. We would prefer to work in the coordinate system that parametrises this disk bundle. Thus, $S^4 - S^1_0 = \{(w, z)\}$ where $w \in \mathbf{C}$ (and $w^{-1} \in \mathbf{C}$) parametrises S^2_∞ and $\{z \in \mathbf{C} \mid |z| < 1\}$ parametrises the disk fibres. We can parametrise all of S^4 by including the over-defined coordinate $\{|z| = 1\}$. With respect to this coordinate system the round metric is given by

$$\mathrm{d}s^2 = \left(\frac{1 - |z|^2}{1 + |z|^2}\right)^2 \frac{4\,\mathrm{d}\bar{w}\,\mathrm{d}w}{(1 + |w|^2)^2} + \frac{4\,\mathrm{d}\bar{z}\,\mathrm{d}z}{(1 + |z|^2)^2}.$$

We will instead work with the conformally equivalent metric

$$\mathrm{d}s^2 = \frac{4\,\mathrm{d}\bar{w}\,\mathrm{d}w}{(1 + |w|^2)^2} + \frac{4\,\mathrm{d}\bar{z}\,\mathrm{d}z}{(1 - |z|^2)^2}, \tag{4}$$

which is the product of the round metric on S^2_∞ with the hyperbolic metric on the disk. In particular, it is a Kahler metric on $S^2 \times D$.

3. Loop group

Let E be a framed $U(n)$-bundle over S^4 with $c_2(E) - \frac{1}{2}c_1(E)^2 = k$. Let A be a smooth unitary connection on E. We will show how to associate to A a smooth map from S^2 to the loop group, $\Omega U(n)$.

Fix $w \in S^2$ and consider the associated fibre, D_w. Over D_w, A defines a holomorphic structure on E. Choose the basepoint over which we frame E to lie on the $S^1_0 \subset S^4$ that gives the common boundary to all the disks.

Proposition 3.1. *There is a unique frame g_w of E over D_w satisfying:*
(i) $\partial_{\bar{z}}^A g_w = 0$;
(ii) g_w *is unitary on* ∂D_w;
(iii) g_w *matches the framing at the basepoint.*
Furthermore, for $U \subset S^2$, g_w is a frame for E over $U \times D$ which is smooth in w.

This is just a restatement of the factorisation theorem for loop groups as observed by Donaldson [7].

Theorem 3.2 [17]. *Any loop $\gamma \in LGL(n, \mathbf{C})$ can be factorised uniquely*

$$\gamma = \gamma_u \cdot \gamma_+,$$

with $\gamma_u \in \Omega U(n)$ and $\gamma_+ \in L^+GL(n, \mathbf{C})$, those loops that are boundary values of holomorphic maps from the disk to $GL(n, \mathbf{C})$. In fact the product map

$$\Omega U(n) \times L^+GL(n, \mathbf{C}) \to LGL(n, \mathbf{C}) \tag{5}$$

is a diffeomorphism.

Proof of Proposition 3.1. Choose a frame $\tilde{g}$ of E over $U \times D$ satisfying $\partial_{\bar{z}}^A \tilde{g} = 0$. That we can do this so that $\tilde{g}$ is smooth in w is proven in [8]. Also, choose a unitary frame of E along $S_0^1 \subset S^4$ that agrees with the framing at the basepoint. Over each disk D_w, Theorem 3.2 enables us to find a unique $\gamma_+(w)$ that maps D_w holomorphically to $GL(n, \mathbf{C})$ so that $g = \tilde{g}\gamma_+$ is unitary on $S_0^1 = \partial D_w$ and agrees with the frame at the basepoint. In fact, since (5) is a *diffeomorphism*, when restricted to $S_0^1 = \partial D_w$, γ_+ is smooth in w. Since $\gamma_+(w)$ is holomorphic in z there is an exact expression for its values on the interior of D_w via a Cauchy integral formula. It follows that γ_+ is smooth in w over all of $U \times D$. Since $\tilde{g}$ was chosen to be smooth in w it follows that $g = \tilde{g}\gamma_+$ is also. □

Equip the space of gauge equivalence classes of connections on a bundle E over S^4, $\mathcal{B}_{S^4}$, with the smooth topology and likewise for the space of smooth maps from the two-sphere to the loop group, $Map^*(S^2, \Omega U(n))$.

Corollary 3.3. *There is a smooth map*

$$\mathcal{F} : \mathcal{B}_{S^4} \to Map^*(S^2, \Omega U(n)).$$

Proof. Given a smooth connection A on E, on each disk in $S^2 \times D \cong S^4 - S_0^1$ restrict the g supplied by Proposition 3.1 to the boundary S_0^1 to get S^2 unitary frames there. Use the frame defined by the disk corresponding to $\infty \in S_\infty^2$ as a background frame. Comparing this to the other frames we get a smooth map

$$\mathcal{F}(A) : S^2 \to \Omega U(n)$$

that sends ∞ to the constant loop I. Furthermore, the factorisation (5) which gives the smoothness of $\mathcal{F}(A)$ also implies that $\mathcal{F}$ is smooth as a map on $\mathcal{B}_{S^4}$. □

Corollary 3.4. *If A satisfies the ASD equations then $\mathcal{F}(A)$ is a holomorphic map.*

Proof. We need only two of the three ASD equations to prove this. In complex coordinates they can be combined to give

$$[\partial^A_{\bar{w}}, \partial^A_{\bar{z}}] = 0. \tag{6}$$

Associate to A the frame g from Proposition 3.1. Since $\partial^A_{\bar{z}} g = 0$ it follows from (6) that $\partial^A_{\bar{z}}(\partial^A_{\bar{w}} g) = 0$ or equivalently that $\partial^A_{\bar{w}} g = g\eta$ for a map $\eta : S^2 \times D \to GL(n, \mathbf{C})$ that is holomorphic in z. Now choose a unitary gauge for E in a neighbourhood of $S^1_0 \subset S^4$ that extends the backgound frame on S^1_0 determined by A over D_∞. The map $u = \mathcal{F}(A)$ is simply the restriction of g to each ∂D_w with respect to the background frame. With respect to this frame $\partial^A_{\bar{w}} = \partial_{\bar{w}}$ simply due to the choice of coordinate system. Thus we have

$$u^{-1}\partial_{\bar{w}} u : S^2 \to L^+\mathbf{gl}(n, \mathbf{C}).$$

But this is exactly the statement that $u = \mathcal{F}(A)$ is a holomorphic map into the loop group. We can see this by looking closely at the complex structure J on $\Omega U(n)$. For $\xi \in \Omega\mathbf{u}(n)$, $J\xi \equiv \mathrm{i}\xi (\mathrm{mod}\ L^+\mathbf{gl}(n, \mathbf{C}))$ and in fact this defines J since each element of $L\mathbf{gl}(n, \mathbf{C})$ has a unique unitary representative in its $L^+\mathbf{gl}(n, \mathbf{C})$ coset. Put $w = x + \mathrm{i}y$, then u is holomorphic when

$$\begin{aligned} 0 &= u^{-1}(\partial u/\partial x + J\partial u/\partial y) \\ &\equiv u^{-1}(\partial u/\partial x + \mathrm{i}\partial u/\partial y)\ (\mathrm{mod}\ L^+\mathbf{gl}(n, \mathbf{C})). \qquad \square \end{aligned}$$

Remark 3.5. In Section 4 we will show that $\mathcal{F}$ defines a diffeomorphism from the space of instantons to the space of holomorphic maps. This fact together with the proof of the Atiyah–Jones conjecture and an analogue of the Atiyah–Jones conjecture for maps into the loop group implies that $\mathcal{F}$ defined in Corollary 3.3 is a homotopy equivalence.

4. Existence and uniqueness

In this section we will show that when restricted to $\mathcal{M}_{S^4}$, gauge equivalence classes of ASD connections over S^4, the map $\mathcal{F}$ defines a diffeomorphism

$$\mathcal{F} : \mathcal{M}_{S^4} \to Hol^*(S^2, \Omega U(n)).$$

Associate to any instanton A the pair (H, η) consisting of a metric $H = \bar{g}^{\mathrm{T}} g$ using the frame g supplied by Proposition 3.1 and $\eta : S^2 \times D \to \mathbf{gl}(n, \mathbf{C})$, the holomorphic (in z) extension of the map $\mathcal{F}(A)^{-1}\partial_{\bar{w}}\mathcal{F}(A) : S^2 \to L^+\mathbf{gl}(n, \mathbf{C})$. By construction $H \equiv I$ on $S^1_0 \subset S^4$. We can retrieve A from (H, η) since with respect to the gauge defined by g, we get

$$A_{\bar{w}} = \eta, \quad A_{\bar{z}} = 0, \quad A_w = H^{-1}\partial_w H - H^{-1}\bar{\eta}^{\mathrm{T}} H, \quad A_z = H^{-1}\partial_z H. \tag{7}$$

Notice that gauge-equivalent connections produce the same H. Associate to the pair (H, η) the Hermitian–Yang–Mills tensor

$$B(H,\eta)=(1-|z|^2)^2\partial_{\bar z}(H^{-1}\partial_z H)+(1+|w|^2)^2\{\partial_{\bar w}(H^{-1}\partial_w H)\\ -\partial_{\bar w}(H^{-1}\bar\eta^{\mathrm T}H)-\partial_w\eta+[\eta,H^{-1}\partial_w H-H^{-1}\bar\eta^{\mathrm T}H]\}.$$

This vanishes when (H,η) comes from an instanton. Later we will study more general pairs (H,η) and attempt to solve the equation $B(H,\eta)=0$. This is elliptic in H away from $S^1_0\subset S^4$.

4.1. Uniqueness

Proposition 4.1. *Two instantons A_1 and A_2 are gauge equivalent if and only if $\mathcal{F}(A_1)=\mathcal{F}(A_2)$.*

Proof. Associate to each instanton the pair (H_i,η) (by assumption η is common to both). Set $h=H_1^{-1}H_2$. This is an endomorphism of the bundle over S^4. So far we have been working gauge invariantly. In order to compare A_1 and A_2 we will choose the gauge defined by Proposition 3.1. Thus we identify g_1 and g_2. With respect to this gauge we have

$$\bar\partial^{A_2}=\bar\partial^{A_1},\qquad \partial^{A_2}=\partial^{A_1}+h^{-1}\partial^{A_1}h,$$

where we have separated the connections, respectively, into their (1, 0) and (0, 1) parts. This expression is gauge-invariant and in fact it holds in all gauges. (We have merely used g_1 and g_2 to specify isomorphisms with the bundle E.) Since the connections are ASD we have

$$F_{A_i}=\bar\partial^{A_i}\circ\partial^{A_i}+\partial^{A_i}\circ\bar\partial^{A_i}\quad\Rightarrow\quad F_{A_2}=F_{A_1}+\bar\partial^{A_1}(h^{-1}\partial^{A_1}h).$$

Thus from $B(H_i,\eta)=0$ we get

$$0=(1-|z|^2)^2\partial_{\bar z}^{A_1}(h^{-1}\partial_z^{A_1}h)+(1+|w|^2)^2\{\partial_{\bar w}^{A_1}(h^{-1}\partial_w^{A_1}h)\tag{8}$$

Lemma 4.2. *The function $tr(h)$ is subharmonic.*

Proof. With respect to the metric in (4), the Laplacian is given by

$$\Delta=-(1-|z|^2)^2\partial_{\bar z}\partial_z-(1+|w|^2)^2\partial_{\bar w}\partial_w,$$

so

$$-\Delta tr(h)=(1-|z|^2)^2tr\{(\partial_{\bar z}^{A_1}h)(h^{-1}\partial_z^{A_1}h)+h\partial_{\bar z}^{A_1}(h^{-1}\partial_z^{A_1}h)\}\\ +(1+|w|^2)^2tr\{(\partial_{\bar w}^{A_1}h)(h^{-1}\partial_w^{A_1}h)+h\partial_{\bar w}^{A_1}(h^{-1}\partial_w^{A_1}h)\},$$

which we will show to be non-negative. The two right terms vanish by (8). In order to show that the other two terms are non-negative we will choose a gauge in which ∂^{A_1} and $\bar\partial^{A_1}$ are adjoints. Use the g_1 constructed in Proposition 3.1 to transform from the holomorphic frame to a unitary frame. With respect to this frame $h=(\bar g_1^{\mathrm T})^{-1}H_2g^{-1}$ which is self-adjoint. In fact, since we have the freedom to replace g with ug and thus h with uhu^{-1} where u is a constant unitary transformation, at any point we can arrange that h is diagonal. It has positive eigenvalues since H_2 is a metric. So each term is of the form $tr(\bar{M}^{\mathrm{T}}hM) \geq 0$ and the lemma follows. □

Proof of Proposition 4.1 (continued). By reversing the roles of H_1 and H_2 we see that $tr(h^{-1})$ is also subharmonic. Put $\sigma(h) = tr(h) + tr(h^{-1}) - 2n$. Since the eigenvalues of h are all positive, then $\sigma(h) \geq 0$ everywhere. We also know that $\sigma(h)$ is subharmonic and on the boundary $\sigma(h) = 0$. By the maximum principle $\sigma(h) \leq 0$ so $\sigma(h) \equiv 0$ and $h \equiv I$. Thus $H_1 = H_2$ and A_1 is gauge-equivalent to A_2. □

4.2. The heat flow

We will now prove that every based holomorphic map from the two-sphere to the loop group comes from an instanton over the four-sphere. In order to do this we will prove the existence theorems for instantons in a standard way using a heat flow. We closely follow the approach used in [12] to prove a similar theorem for Euclidean monopoles. Our proof of the long time existence of the flow on a subset of S^4 is equivalent to the proof in [19]. It is necessary that we go through this proof in order to get estimates to extend to S^4 and since our proof will be necessary when we generalise to other Riemann surfaces. All these methods are really variations on Donaldson's proof of the existence of ASD connections on stable holomorphic bundles over a Kahler surface [5].

Away from $|z| = 1$, the Hermitian–Yang–Mills tensor $B(H, \eta)$ is elliptic in H. We wish to find a solution of the equation $B(H, \eta) = 0$ and since η encodes the holomorphic map we will be able to retrieve an ASD connection associated to that map. A solution of the heat flow equation

$$H^{-1}\partial H/\partial t = B(H, \eta), \qquad H(w, z, 0) = I, \tag{9}$$

will converge to the required solution as $t \to \infty$. Later we will explain the significance of the fact that we can choose the constant metric I for an initial condition.

It is disappointing that we have not been able to solve (9) on the compact manifold S^4 without cutting it open and solving a sequence of boundary-value problems. It seems that the metrics we use are not $C^{2,\alpha}$ or $W^{3,2}$ as the existing methods require. Probably the metrics are $W^{2,p}$ which make it seem likely that there is a way around the boundary-value problem.

A word on existent methods. The round metric on $S^4 - S^1_0$ is conformally equivalent to an infinite volume Kahler metric on $S^2 \times D$. The H we use differs from the Hermitian–Yang–Mills $\tilde{H}$ by a complex gauge transformation, $\tilde{H} = \bar{p}^{\mathrm{T}}Hp$ where $p : U \times D \to GL(n, \mathbf{C})$, $U \subset S^2$ satisfies

$$\partial_{\bar{z}}p = 0, \qquad -\partial_{\bar{w}}p \cdot p = \eta.$$

The existence of p follows from the existence of a universal holomorphic bundle over $\Omega U(n) \times S^2$ which requires explicit knowledge of the cell decomposition of the loop group [17]. By restricting to a compact subset of $S^2 \times D$ we can use Simpson's results [19] to get long-time existence of the heat flow for $\tilde{H}$. We would still need to go through the proof to get precise estimates of how far the metric flows from the initial choice in order to extend to S^4 as well as interpret the result as in Section 6. For a more general Riemann surface, we do not have the existence of the complex gauge transformation p that relates the metrics. For this reason we do not use Simpson's results. Still, once we have the ASD connection then we can produce the required complex gauge transformation, so the two methods are related. Essentially a corollary of our result is a theorem about holomorphic disks in loop groups related to general Riemann surfaces. In particular we get an alternative proof of the existence of the universal holomorphic bundle over the loop group.

Put

$$X_\epsilon = \{(w, z) \in S^2 \times D \mid |z| \le \epsilon\}$$

so the X_ϵ exhaust $S^2 \times D$ as $\epsilon \to 1$.

Proposition 4.3. *Over each X_ϵ there is a unique solution, H_t^ϵ, of the boundary-value problem*

$$H^{-1}\partial H/\partial t = B(H, \eta), \quad H(w, z, 0) = I, \quad H_{|\partial X_\epsilon} = I \tag{10}$$

defined for all t and converging to a smooth metric H_∞^ϵ that satisfies $B(H_\infty^\epsilon, \eta) = 0$.

Proof. Since we have fixed X_ϵ for the moment we will omit the superscript in H_t^ϵ during this proof. Short-time existence of a solution of (10) is automatic since $B(H, \eta)$ is elliptic in H and we have Dirichlet boundary conditions. In order to extend this to long-time existence we will take the approach given by Donaldson [5] and extended by Simpson [19] and show that a solution on $[0, T)$ gives a limit at T which is a good initial condition to start the flow again. The lemmas we need to prove on the way use the details of our particular case and allow us to proceed with Donaldson's proof.

Lemma 4.4. *If H_1 and H_2 are two solutions of the heat equation then*

$$\partial_t\sigma + \Delta\sigma \le 0 \tag{11}$$

for $\sigma = tr(H_1^{-1}H_2) + tr(H_1H_2^{-1}) - 2n$.

Proof. We can generalise the proof of Proposition 4.1 as follows. So far we have shown that

$$tr\{hB(H_2, \eta) - hB(H_1, \eta)\} \le -\Delta tr(h).$$

Now,

$$tr(\partial_t h) = tr(hH_2^{-1}\partial_t H_2) - tr(H_1^{-1}\partial_t H_1 h),$$

so from the flow equation we get $\partial_t tr(h) + \Delta tr(h) \le 0$ and by reversing the roles of H_1 and H_2, (11) follows. □

Apply (11) to H_t and $H_{t+\tau}$, the flow at two times. Since they obey the same boundary conditions on X_ϵ, σ vanishes on the boundary. By the maximum principle $\sup_{X_\epsilon} \sigma$ is a non-increasing function of t. By continuity, for any $\rho > 0$ there exists a δ small enough so that

$$\sup_{X_\epsilon} \sigma(H_t, H_{t'}) < \rho$$

for $0 < t, t' < \delta$. It follows from the non-increasing property of σ that

$$\sup_{X_\epsilon} \sigma(H_t, H_{t'}) < \rho$$

for $T - \delta < t, t' < T$. Since ρ can be made arbitrarily small, H_t is a Cauchy sequence in the C^0 norm as $t \to T$. The metrics take their values in a complete metric space (described below) and the function σ acts like the metric so there is a *continuous* limit H_T of the sequence. Notice also that (11) and the maximum principle show that this short-time solution to the heat flow equation is unique.

A metric H takes its values in the space $GL(n, \mathbf{C})/U(n)$ which comes equipped with the complete metric d which is given locally by $tr(H^{-1}\delta H)^2$. Thus

$$\mathrm{d}(H(w, z, t), H(w, z, 0)) = \int_0^t |B(H_s, \eta)|\,\mathrm{d}s,$$

where $|B(H_s, \eta)|^2 = tr(B^* B)$ and the adjoint is taken with respect to the metric H_s. Notice that $B^* = B$ so $|B(H_s, \eta)|^2 = tr(B^2)$.

Lemma 4.5. *If H_t is a solution of the heat equation then*

$$(\mathrm{d}/\mathrm{d}t + \Delta)|B(H_t, \eta)| \leq 0 \quad \text{whenever } |B| > 0. \tag{12}$$

Proof. First notice that it is only ∂^A, the holomorphic part of connection (7), that depends on t, so

$$\partial_t B(H, \eta) = (1 - |z|^2)^2 \partial_{\bar{z}}^A(\partial_t(\partial_z^A)) + (1 + |w|^2)^2 \partial_{\bar{w}}^A(\partial_t(\partial_w^A))$$

and since $\partial_t(\partial^A) = \partial^A(H^{-1}\partial_t H)$, we have

$$\begin{aligned}
\partial_t B(H, \eta) &= \{(1 - |z|^2)^2 \partial_{\bar{z}}^A \partial_z^A + (1 + |w|^2)^2 \partial_{\bar{w}}^A \partial_w^A\}(H^{-1}\partial_t H) \\
&= \{(1 - |z|^2)^2 \partial_{\bar{z}}^A \partial_z^A + (1 + |w|^2)^2 \partial_{\bar{w}}^A \partial_w^A\} B(H, \eta).
\end{aligned}$$

This last expression looks quite like the Laplacian and in fact

$$\begin{aligned}
\partial_t |B|^2 &= \partial_t tr(B^2) = 2\,tr(B \partial_t B) \\
&= 2\,tr\{(1 - |z|^2)^2 B \partial_{\bar{z}}^A \partial_z^A + (1 + |w|^2)^2 B \partial_{\bar{w}}^A \partial_w^A\} B \\
&= -\Delta tr(B^2) - 2(1 - |z|^2)^2 tr(\partial_{\bar{z}}^A B \partial_z^A B) \\
&\quad - 2(1 + |w|^2)^2\, tr(\partial_{\bar{w}}^A B \partial_w^A B)
\end{aligned}$$

so

$$2|B|(\partial_t + \Delta)|B| = (\partial_t + \Delta)|B|^2 + 2(1 - |z|^2)^2(|\partial_z|B||^2 + |\partial_{\bar z}|B||^2) + 2(1 + |w|^2)^2(|\partial_w|B||^2 + |\partial_{\bar w}|B||^2)$$

is non-positive by Kato's inequality $|\partial_x|f|| \le |\partial_x^A f|$. □

It follows from (12) and the maximum principle that if there is a function $f(w, z, t)$ defined on $X_\epsilon \times \mathbf{R}$ that satisfies $(\partial_t + \Delta)f = 0$ and $|B_0| = |B(I, \eta)| \le f(w, z, 0)$ then $B(H_t, \eta)| \le f(w, z, t)$ for all t.

Lemma 4.6. *Since η is the holomorphic extension of $u^{-1}\partial_{\bar w}u$, for a given holomorphic map $u : S^2 \to \Omega U(n)$, there exists a constant M such that $|B(I, \eta)| \le M(1 - |z|)$ on $S^2 \times D$.*

Proof. The map

$$B(I, \eta) = -(1 + |w|^2)^2(\partial_w \eta + \partial_{\bar w}\bar\eta^{\mathrm{T}} + [\eta, \bar\eta^{\mathrm{T}}])$$

is continuous on $S^2 \times D$. (Notice that it is invariant under the change $w \mapsto w^{-1}$.) Thus, if we can show that $B(I, \eta)/(1 - |z|)$ is continuous on $S^2 \times D$ then it must be bounded since its domain is a compact set. Away from $|z| = 1$ this is clear. At $|z| = 1$, $B(I, \eta) = 0$ since $\eta = u^{-1}\partial_{\bar w}u$ there so $\partial_w \eta + \partial_{\bar w}\bar\eta^{\mathrm{T}} + [\eta, \bar\eta^{\mathrm{T}}]$ is the curvature of a flat connection and hence 0. Thus, a continuous limit of $B(I, \eta)/(1 - |z|)$ as $|z| \to 1$ is the same as a continuous derivative $\partial_{|z|}B(I, \eta)$ at $|z| = 1$. Away from $|z| = 1$, η satisfies $\partial_{\bar z}\eta = 0$. Thus

$$|z|\partial_{|z|}\eta = -\mathrm{i}\partial_\theta \eta, \qquad |z|\partial_{|z|}\bar\eta^{\mathrm{T}} = \mathrm{i}\partial_\theta \bar\eta^{\mathrm{T}}$$

and these derivatives extend continuously to $|z| = 1$. Therefore

$$\begin{aligned} |z|\partial_{|z|}B(I, \eta) &= -(1 + |w|)^2\{\partial_w |z|\partial_{|z|}\eta + \partial_{\bar w}|z|\partial_{|z|}\bar\eta^{\mathrm{T}} \\ &\quad + [|z|\partial_{|z|}\eta, \bar\eta^{\mathrm{T}}] + [\eta, |z|\partial_{|z|}\bar\eta^{\mathrm{T}}]\} \\ &= -(1 + |w|)^2\{-\mathrm{i}\partial_w\partial_\theta\eta + \mathrm{i}\partial_{\bar w}\partial_\theta\bar\eta^{\mathrm{T}} \\ &\quad - \mathrm{i}[\partial_\theta\eta, \bar\eta^{\mathrm{T}}] + \mathrm{i}[\eta, \partial_\theta\bar\eta^{\mathrm{T}}]\} \end{aligned}$$

and this last expression extends continuously to the boundary since the derivatives with respect to θ exist. □

Lemma 4.7. *There exists a function, $C(|z|)$, depending on η but independent of ϵ, continuous on $[0, 1]$ with $C(1) = 0$, and such that*

$$d(H(w, z, t), I) \le C(|z|).$$

Proof. Use the maximum principle with $f(w, z, 0) = M(1 - |z|)$. Notice that $f(w, z, 0) = f(|z|)$, so $f(w, z, t) = f(|z|, t)$ since the Laplacian reduces to the one-dimensional Laplacian. From the flow equation (10) we have

$$d(H_t, H_0) = \int_0^t B(H_\tau)\,d\tau \le \int_0^t f(w,z,\tau)\,d\tau \le \int_0^\infty f(w,z,\tau)\,d\tau. \tag{13}$$

Now, $f(|z|, t) = \int f(s, 0)k(|z|, s, t)\,ds$ where k is the one-dimensional heat kernel operator. Since $\int_0^\infty k(|z|, s, t)\,dt = G(|z|, s)$, Green's operator, is finite, Fubini's theorem allows us to interchange the order of integration in (13). So

$$d(H_t(w,z), H_0(w,z)) \le M \int_0^\epsilon (1-s)G(|z|, s)\,ds \le M \int_0^1 (1-s)G(|z|, s)\,ds.$$

With respect to the Laplacian $\Delta = -(1-|z|^2)^2\partial^2_{|z|}$,

$$G(|z|, s) = -\max\{\ln(|z|), \ln(s)\}/(1-s^2)^2.$$

Actually, this Green's operator is only valid for the entire interval ($\epsilon = 1$) and Fubini's theorem does not apply there. There is a monotone property of heat kernels which means that our choice of G is simply an overestimate when $\epsilon < 1$, so the calculation is valid.

Thus it remains to estimate the quantity

$$-\int_0^{|z|} (1-s)\ln(|z|)/(1-s^2)^2\,ds - \int_{|z|}^1 (1-s)\ln(s)/(1-s^2)^2\,ds. \tag{14}$$

The finite integral $-\int_{|z|}^1 \ln(1-s)/s\,ds$ dominates (14) so the lemma follows. □

Proof of Proposition 4.3 (conclusion). The preceding lemmas have shown that there is a solution to the heat equation that satisfies $H_t \to H_T$ in C^0 and $B(H_t, \eta)$ is uniformly bounded. These are the conditions required to use Simpson's extension of Donaldson's result to show that H_t are bounded in $W^{2,p}$ uniformly in t. Hamilton's methods [11] then give control of all higher Sobolev norms. Thus we get a solution, H_t, of (10) for all t that converges to a smooth limit H_∞ defined on X_ϵ and satisfying $B(H_\infty, \eta) = 0$ and $H_\infty = I$ on ∂X_ϵ.

Proposition 4.8. *For each holomorphic map $u : S^2 \to \Omega SU(n)$ there is an ASD connection A on S^4 such that $\mathcal{F}(A) = u$.*

Proof. We have proven the existence of a family of metrics H^ϵ, respectively, defined over X_ϵ and satisfying $B(H^\epsilon, \eta) = 0$. Since $\sigma(H^\epsilon, H^{\epsilon'})$ is subharmonic its maximum occurs at the boundary of the set on which it is defined. For $\epsilon < \epsilon'$, the common set is X_ϵ. From Lemma 4.7,

$$d(H^{\epsilon'}(w,z), H^\epsilon(w,z)) \le C(\epsilon)$$

since the initial value of the flow for $H^{\epsilon'}$ is given by $H^\epsilon = I$ on ∂X_ϵ. Since $C(\epsilon) \to 0$ as $\epsilon \to 1$, the sequence $\{H^\epsilon\}$ is Cauchy as $\epsilon \to 1$. Thus it converges uniformly to a limit

H. Moreover, on each X_ϵ the limit satisfies $B(H,\eta)=0$ so by regularity is smooth. This comes from a remark of Simpson [19]. The difference between this situation and that in the proof of Proposition 4.3 is that the metrics no longer satisfy Dirichlet boundary conditions so we need to work with $W^{2,p}_{\mathrm{loc}}$. This argument applies to all X_ϵ so the limit H is smooth on $S^4-S^1_0$ and continuous on all of S^4, converging to I on S^1_0. It remains to show that this metric H produces an ASD connection using (7). The connection A is defined and ASD on $S^4-S^1_0$. By the following lemma, A has finite charge. Since codimension three singularities of finite charge ASD connections can be removed [18], A is smooth on all of S^4.

Lemma 4.9. *The curvature of the limiting connection A has finite L^2 norm.*

Proof. We will show that on $X=\lim_{\epsilon\to 0}X_\epsilon$ the heat flow decreases the total charge (which is just an explicit version of the fact that the heat flow is the same as the Yang–Mills flow), and that the charge of the initial connection is bounded. Lemmas 4.5 and 4.6 show that the self-dual (SD) part of the curvature decreases. In order to show that the integral of the ASD part of the curvature and hence the charge decreases, it is sufficient to show that $k(E)=c_2(E)-\frac{1}{2}c_1(E)^2$, the Chern number of the bundle restricted to X, is constant throughout the flow. Then any decrease in the integral of the SD part of the curvature will be matched by the same decrease in the integral of the ASD part of the curvature. The fractional part of $k(E)$ is given by the Chern–Simons invariant of the connection restricted to ∂X. Since $k(E)$ varies continuously with t it is sufficient to show that its fractional part remains constant in order to deduce that it remains constant. The derivative of the Chern–Simons invariant has quite a simple form:

$$\partial_t k(E)=\int\limits_{\partial X}F_A\wedge\partial_t A=\int\limits_{\partial X}F_A\wedge\partial^A B(H,\eta).$$

Here we have used the fact that $\partial_t A=\partial^A B(H,\eta)$, where ∂^A is the holomorphic part of d^A. Since $B(H,\eta)$ vanishes on ∂X then $\partial^A B(H,\eta)=0$ also vanishes there since it is constant in t and in the limit $B(H,\eta)\equiv 0$. Thus the Chern number is constant.

We will calculate the initial Chern number of the connection and then together with the estimate of Lemma 4.6 we have a bound on the initial charge.

$$k(E)=\frac{1}{8\pi^2}\int\limits_{S^2\times D}tr(F_A^2)=-\frac{1}{8\pi^2}\int\limits_{S^2\times D}tr(\partial_{\bar z}\bar\eta^{\mathrm{T}}\partial_z\eta)\,\mathrm{d}\bar z\,\mathrm{d}z\,\mathrm{d}\bar w\,\mathrm{d}w,$$

since only the $F_{z\bar w}$ and $F_{\bar z w}$ terms contribute. Since η is holomorphic in z, then on the disk $\mathrm{d}\{tr(\bar\eta^{\mathrm{T}}\partial_z\eta)\,\mathrm{d}z\}=tr(\partial_{\bar z}\bar\eta^{\mathrm{T}}\partial_z\eta)\,\mathrm{d}\bar z\,\mathrm{d}z$. So

$$k(E)=-\frac{1}{8\pi^2}\int\limits_{S^2}\int\limits_{|z|=1}tr(\bar\eta^{\mathrm{T}}\partial_z\eta)\,\mathrm{d}z\,\mathrm{d}\bar w\,\mathrm{d}w.$$

On $|z| = 1$, $\eta = u^{-1}\partial_{\bar{w}}u$ so $\bar{\eta}^{\mathrm{T}} = -u^{-1}\partial_{w}u$. Since u is holomorphic, we can put $u^{-1}\partial_{\bar{w}}u = \xi + \mathrm{i}J\xi$, where $\xi = \frac{1}{2}u^{-1}\partial_{\mathrm{Re}(w)}u$ and J is the complex structure on $\Omega SU(n)$. Therefore

$$\begin{aligned}\frac{1}{2\pi}\int\limits_{|z|=1} tr(\bar{\eta}^{\mathrm{T}}\partial_z\eta)\,\mathrm{d}z &= \frac{-1}{2\pi}\int\limits_{S^1} tr\{(\xi - \mathrm{i}J\xi)\partial_\theta(\xi + \mathrm{i}J\xi)\}\,\mathrm{d}\theta \\ &= \frac{-1}{2\pi}\int\limits_{S^1} \mathrm{i}\,tr(\xi\partial_\theta J\xi - J\xi\partial_\theta\xi)\,\mathrm{d}\theta \\ &= \frac{\mathrm{i}}{2\pi}\int\limits_{S^1} -tr(\xi\partial_\theta J\xi + (J\xi)\partial_\theta J(J\xi))\,\mathrm{d}\theta \\ &= \mathrm{i}(g(\xi,\xi) + g(J\xi, J\xi)),\end{aligned}$$

where g is the Kahler metric on $\Omega U(n)$. Thus

$$k(E) = \frac{1}{4\pi}\int\limits_{S^2}(g(\xi,\xi) + g(J\xi, J\xi))\frac{\mathrm{d}\bar{w}\,\mathrm{d}w}{\mathrm{i}},$$

which is the charge of u. □

Proof of Proposition 4.8 (*conclusion*). The fact that $\mathcal{F}(A) = u$ is immediate and the proof of the proposition is complete. □

Corollary 4.10.

$$\mathcal{F} : \mathcal{M}_{S^4} \to Hol^*(S^2, \Omega U(n))$$

is a diffeomorphism.

Proof. We have shown that $\mathcal{F}$ is smooth, one-to-one and onto. A linearisation of the uniqueness argument shows that $D\mathcal{F}$ is an isomorphism at each point. With respect to any topology that makes the two spaces into Banach manifolds (say, the C^k topology), we can invoke the inverse function theorem to get the required result. □

5. General Σ

The results over S^4 generalise immediately to a family of four-manifolds obtained from general Riemann surfaces. Let $\tilde{\Sigma}$ be a compact Riemann surface. Construct the four-manifold X by performing surgery on $S^2 \times \{\infty\} \subset S^2 \times \tilde{\Sigma}$ – replace a neighbourhood of $S^2 \times \{\infty\}$ with $B^3 \times S^1$. When $\tilde{\Sigma} = S^2$, $X = S^4$. Label the core of $B^3 \times S^1$ by $S^1_0 \subset X$. The open manifold $X - S^1_0$ is foliated by a family of Riemann surfaces $\Sigma = \tilde{\Sigma} - D$ parametrised by S^2 with common boundary $\partial\Sigma = S^1_0$. Put a metric on X that is conformally equivalent to the product Kahler metric on $S^2 \times \Sigma$ given by the round metric on S^2 and a hyperbolic metric on Σ. Such a conformal compactification of the product metric exists in general. In fact, there is a metric with constant scalar curvature in the conformal class. (Since the ends of a complete hyperbolic two-manifold have been classified this is a local problem over $S^1 \times B^3$. For our purposes, it is only necessary that a metric exists locally since then we can use elliptic regularity to remove singularities. The metric in a neighbourhood of a point on $\partial\Sigma$ is isometric to a neighbourhood of a point on the boundary of the hyperbolic disk so the S^4 case gives the required local metric.) Theorem 1.4 follows from the following proposition combined with the implicit function theorem.

Proposition 5.1. *There is a smooth map from $\mathcal{M}_X$, the space of instantons on a framed $U(n)$-bundle over X, to the space of based holomorphic maps from S^2 to $LGL(n,\mathbf{C})/L^+_\Sigma GL(n,\mathbf{C})$ that is one-to-one and onto.*

Proof. We have set up the argument for $X = S^4$ so that it adapts easily to this more general situation. In order to define the map

$$\mathcal{F} : \mathcal{M}_X \to Hol^*(S^2, LGL(n,\mathbf{C})/L^+_\Sigma GL(n,\mathbf{C})),$$

we appeal to a generalisation of the factorisation of Theorem 3.2 due to Donaldson [7]. When we restrict a connection on E over X to $\{x\} \times \Sigma$, it defines a holomorphic structure on E there. The restriction is holomorphically trivial and Donaldson proves that amongst the holomorphic trivialisations there is a trivialisation that is unitary when restricted to $\partial\Sigma$. Unlike when $\Sigma = D$, such a trivialisation is not unique. The frame, u, it defines on the boundary is well-defined only as a section of a flat $U(n)$-bundle over S^1. Thus u takes its values inside the twisted loop group. The frame u is smooth in w and

$$u^{-1}\partial_{\bar{w}} u : S^2 \to L^+_\Sigma GL(n,\mathbf{C}). \tag{15}$$

Notice that (15) is a true map without any of the twisting of a section because the flat structure on $E_{|S^1_0}$ is independent of w. Let $\eta : S^2 \times \Sigma \to GL(n,\mathbf{C})$ be the holomorphic (in z) extension of (15). As before we wish to solve the equation $B(H,\eta) = 0$ where B is the Hermitian–Yang–Mills tensor over $S^2 \times \Sigma$. The Kahler metric and the Laplacian over $S^2 \times \Sigma$ are the same as those over $S^2 \times D$ since the hyperbolic metric and Laplacian over D are invariant under $SU(1,1)$. Thus the argument for uniqueness of instantons with the same holomorphic map goes through as before. We use $H \equiv I$ for the initial metric in the heat flow equation over $X_\epsilon \subset X$. The sets X_ϵ are obtained by removing neighbourhoods of S^1_0. Short-time existence of the flow comes from ellipticity again. Except for the use of Green's function, the long-time existence argument goes through as before. We still get a bound, f, on $|B(I,\eta)|$ that vanishes like O(1) near $\partial\Sigma$ so

$$d(H_t(w,z), I) \le \int_\Sigma f(s)G(z,s)\,\mathrm{d}s$$

for Green's function $G(z,s)$ over Σ. Away from $\partial\Sigma$ this is finite as required. We need to know that it vanishes as z approaches the boundary so that the Cauchy sequence argument of Proposition 4.8 goes through. This follows from the fact that in a neighbourhood of a point on $S^1_0 \subset X$, the situation is isometric to that for S^4 where we have already proven the required vanishing as z approaches the boundary. Finally, the limiting connection on $X - S^1_0$ has finite charge because the Chern number on each X_ϵ is constant throughout the flow as in Lemma 4.9 and the initial charge is finite because again the interest only lies near S^1_0 where the situation mimics that of S^4. Since there is a conformally equivalent metric that extends over X, regularity gives smoothness of the connection over all of X. □

Remark 5.2. Consider the complete hyperbolic surface Σ that looks like a punctured unit disk with metric

$$\mathrm{d}s^2 = \mathrm{d}\bar{z}\,\mathrm{d}z/(|z| \ln |z|)^2.$$

In terms of the upper-half-space model of the hyperbolic plane, it is obtained by quotienting out by the action $\zeta \mapsto \zeta + 2\pi$. The $U(1)$-invariant instantons on $S^2 \times \Sigma$ correspond to Euclidean monopoles. More generally, we get periodic instantons or calorons. The proof in this section does not apply to punctured Riemann surfaces. It is necessary to generalise the results here in order to use these methods in the study of calorons [16].

6. Stretching the metric

In this section we will explain the significance of the fact that we can choose $H \equiv I$ as an initial condition in the flow equation. It shows that the connection defined by (I, η) is approximately an instanton and can be interpreted as an instanton with respect to a very singular metric.

The round metric on S^4 is conformally equivalent to the metric

$$\mathrm{d}s^2 = \frac{\mathrm{d}\bar{a}\,\mathrm{d}a + (\mathrm{d}|b|)^2}{|b|^2} + \mathrm{d}\theta^2 = \frac{\mathrm{d}\bar{A}\,\mathrm{d}A + (\mathrm{d}|B|)^2}{|B|^2} + \mathrm{d}\Theta^2,$$

where $b = |b|\mathrm{e}^{\mathrm{i}\theta}$, $B = |B|\mathrm{e}^{\mathrm{i}\Theta}$. Consider, instead, the metric

$$\mathrm{d}s^2 = \frac{\mathrm{d}\bar{a}\,\mathrm{d}a + (\mathrm{d}|b|)^2}{|b|^2} + \kappa^2\,\mathrm{d}\theta^2 = \frac{\mathrm{d}\bar{A}\,\mathrm{d}A + (\mathrm{d}|B|)^2}{|B|^2} + \kappa^2\,\mathrm{d}\Theta^2 \tag{16}$$

for $\kappa > 0$. This metric is not defined over $S^2_\infty \subset S^4$. Still, we will study $W^{1,2}$ instantons with respect to this metric over $S^4 - S^2_\infty$. Really we are working over $H^3 \times S^1 \cong S^4 - S^2_\infty$. We can interpret κ as the length of the circle or the curvature of hyperbolic space.

In a sense, as we let $\kappa \to \infty$, the instantons with respect to the metric (16) converge to connections of the form

$$A = \eta\,\mathrm{d}\bar{w} - \bar{\eta}^{\mathrm{T}}\,\mathrm{d}w,$$

which are the initial values for the heat flow for the round metric over S^4. Since this idea serves only to illuminate the proof of Theorem 1.1, we are being rather loose with this notion of convergence. There are four points on this issue we should note.

(i) It will be clear that the class of instantons over $S^4 - S^2_\infty$ that we consider here lies inside the space of finite energy connections. For the converse – that we get all finite energy instantons – we rely on a recent result of Mazzeo and Rade [15].
(ii) In the limit, the connections actually concentrate at S^2_∞ so we have to reparametrise the normal bundle of S^2_∞ to allow for this.
(iii) We will prove something weaker than convergence of the connections. We will prove that the associated metrics H converge in C^0 rather than in C^1.
(iv) Atiyah and Murray [1] studied the non-renormalised zero mass limit of hyperbolic monopoles and conjectured a relationship with the Yang–Baxter equations. The results here suggest that one can similarly pursue a connection between instantons and solutions of the Yang–Baxter equation.

With respect to the parametrisation of $S^2 \times D$, (16) is

$$\mathrm{d}s^2 = \sinh^2(\rho)\frac{4\,\mathrm{d}\bar{w}\,\mathrm{d}w}{(1+|w|^2)^2} + \mathrm{d}\rho^2 + \kappa^2\,\mathrm{d}\theta^2,$$

where ρ gives the hyperbolic distance from the centre of H^3. Previously, we put $z = \mathrm{e}^{-\rho+\mathrm{i}\theta}$, so $4\,\mathrm{d}\bar{z}\,\mathrm{d}z/(1-|z|^2)^2 = (\mathrm{d}\rho^2 + \mathrm{d}\theta^2)/\sinh^2(\rho)$. To rescale, put $\rho = \kappa\tau$ and $z = \mathrm{e}^{-\tau+\mathrm{i}\theta}$. Now, (16) is conformally equivalent to the Kahler metric

$$\mathrm{d}s^2 = \frac{4\,\mathrm{d}\bar{w}\,\mathrm{d}w}{(1+|w|^2)^2} + \frac{\kappa^2\sinh^2(\tau)}{\sinh^2(\kappa\tau)}\frac{4\,\mathrm{d}\bar{z}\,\mathrm{d}z}{(1-|z|^2)^2}, \tag{17}$$

which makes it clear that the metric degenerates on S^2_∞ (for $\kappa > 1$). The rescaling $\rho = \kappa\tau$, besides exactly compensating for the move of the charge towards S^2_∞, is quite natural when we intepret κ as the curvature of H^3 – the new coordinate τ gives the distance with respect to the new hyperbolic metric.

The Hermitian–Yang–Mills tensor with respect to the new metric is

$$\begin{aligned} B_\kappa(H,\eta) = {} & \frac{\sinh^2(\kappa\tau)}{\kappa^2\sinh^2(\tau)}(1-|z|^2)^2\partial_{\bar{z}}(H^{-1}\partial_z H) \\ & + (1+|w|^2)^2\{\partial_{\bar{w}}(H^{-1}\partial_w H) - \partial_{\bar{w}}(H^{-1}\bar{\eta}^{\mathrm{T}}H) - \partial_w\eta \\ & + [\eta, H^{-1}\partial_w H - H^{-1}\bar{\eta}^{\mathrm{T}}H]\}. \end{aligned}$$

The equation $B_\kappa(H,\eta) = 0$ is elliptic away from S^1_0 and S^2_∞. For $\epsilon_1 > \epsilon_2$, define

$$X_{\epsilon_1,\epsilon_2} = \{(w,z) \in S^2 \times D \mid \epsilon_1 \geq |z| \geq \epsilon_2\}$$

so the $X_{\epsilon_1,\epsilon_2}$ exhaust $S^4 - S^2_\infty \cup S^1_0$.

Proposition 6.1. *Over each $X_{\epsilon_1,\epsilon_2}$ there is a unique solution, $H_t^{\epsilon_1,\epsilon_2}$, of the boundary-value problem*

$$H^{-1}\partial H/\partial t = B_\kappa(H,\eta), \quad H(w,z,0) = I, \quad H_{|\partial X_{\epsilon_1,\epsilon_2}} = I \tag{18}$$

defined for all t and converging to a smooth metric $H_\infty^{\epsilon_1,\epsilon_2}$ that satisfies $B_\kappa(H_\infty^{\epsilon_1,\epsilon_2},\eta) = 0$.

Proof. Since the change in the Laplacian for the new metric mimics the change in the Hermitian–Yang–Mills tensor, the proof of Proposition 4.3 works for this boundary-value problem except that we have to modify Green's function in Lemma 4.7. It is now given by

$$G(|z|, s) = -\frac{\max\{\ln(|z|), \ln(s)\}}{(1-s^2)^2}\,\frac{\kappa^2 \sinh^2(\tau)}{\sinh^2(\kappa\tau)}.$$

For $\kappa > 1$ the integral (14) with the new Green's function is dominated by the finite integral

$$-\int_{|z|}^{1} \frac{\ln(1-s^\kappa)}{s}\, \mathrm{d}s = -\frac{1}{\kappa}\int_{|z|^\kappa}^{1} \frac{\ln(1-s^\kappa)}{s}\, \mathrm{d}s$$

so we can replace $C(|z|)$ in Lemma 4.7 by $C(|z|^\kappa)/\kappa$. □

We can now state the analogue of Proposition 4.8.

Proposition 6.2. *For each holomorphic map $u : S^2 \to \Omega SU(n)$ there is a unique finite charge connection A on $S^4 - S^2_\infty$ that is ASD with respect to the metric (17), with*

$$\partial_{\bar z}^A = \partial_{\bar z}, \qquad \partial_{\bar w}^A = \partial_{\bar w} + \eta$$

and such that its associated metric H_κ is bounded.

Proof. We will simply modify the proofs of Propositions 4.1 and 4.8. For uniqueness, we again use subharmonicity. For any two metrics H_1, H_2 that come from two connections satisfying the properties above, put $h = tr(H_1^{-1}H_2)$ and $\sigma = tr(h) + tr(h^{-1}) - 2n$. Then σ is subharmonic on $S^4 - S^2_\infty$, equal to 0 on S^1_0 and *bounded*. For any $\lambda > 0$, $\sigma + \lambda \ln(|z|)$ is also subharmonic and 0 on S^1_0, but now negative near S^2_∞. By the maximum principle $\sigma + \lambda \ln(|z|) \le 0$ on all of $S^4 - S^2_\infty$. Since this is true for all λ, $\sigma \le 0$. By construction $\sigma \ge 0$ so $\sigma \equiv 0$ and $H_1 = H_2$.

For existence we need to show that the sequence of metrics $H_\infty^{\epsilon_1,\epsilon_2}$ is Cauchy as $\epsilon_1 \to 1$ and $\epsilon_2 \to 0$. For $\epsilon_1 > \epsilon_2 > \epsilon_3 > \epsilon_4$, associate σ to the two metrics $H_\infty^{\epsilon_1,\epsilon_4}$ and $H_\infty^{\epsilon_2,\epsilon_3}$. Again, σ is subharmonic so it takes its maximum on the boundary of the set over which both are defined, $\partial X_{\epsilon_2,\epsilon_3}$. Since the boundary values of $H_\infty^{\epsilon_2,\epsilon_3}$ give the initial values of $H_\infty^{\epsilon_1,\epsilon_4}$ on $\partial X_{\epsilon_2,\epsilon_3}$, Proposition 6.1 shows that σ is less than a constant times $C(\epsilon_2^\kappa)/\kappa$ and $C(\epsilon_3^\kappa)/\kappa$ on the respective boundary components. (The constant enters since the distance function on $GL(n, \mathbf{C})/U(n)$ dominates σ times a constant.) If we label these two maximum values by M_2 and M_3 then the function

$$M_2 + M_3 \ln(|z|/\epsilon_2)/\ln(\epsilon_3/\epsilon_2) \tag{19}$$

is harmonic and takes on the values M_2 and M_3 on the respective boundary components. Thus σ is less than (19) on all of $X_{\epsilon_2,\epsilon_3}$. As $\epsilon_2 \to 1$ and $\epsilon_3 \to 0$, (19) goes to zero. Hence the sequence is Cauchy and converges in C^0 to a limit H_κ. As before, the convergence is smooth on the $S^4 - S^2_\infty \cup S^1_0$ so H_κ is smooth there also and satisfies $B_\kappa(H, \eta) = 0$. The ASD connection produced by (H_κ, η) extends across S^1_0 by regularity since the charge is finite as before. The same is not true near S^2_∞ since the metric degenerates. □

Corollary 6.3. *As $\kappa \to \infty$, $H_\kappa \to I$ uniformly on compact subsets of $S^4 - S^2_\infty$.*

Proof. This simply follows from the fact that $C/\kappa \to 0$ uniformly since C is bounded. □

We might try to analyse the convergence more closely to see if it remains true on the level of the connections. Alternatively, we might learn from this that by using the heat flow without stretching the metric we can avoid some of the difficult analytic issues involved in similar problems. Dostoglou and Salamon [9,10] solved the Atiyah–Floer conjecture for a mapping cylinder $Y = S^1 \times_h \Sigma$ by stretching the metric in the S^1 direction. The techniques in this paper suggest an alternative approach. In one direction, associate to an instanton over $\mathbf{R} \times Y$ a holomorphic curve into $\mathcal{M}$, the space of flat connections over Σ, by taking the unique (up to conjugation) flat connection over each $\{(t, \theta)\} \times \Sigma$ that defines the same holomorphic structure as the restriction of the instanton. The Hermitian–Yang–Mills tensor is very natural in this problem and will give uniqueness. It will also enable us to use the heat flow to go in the other direction and obtain an instanton from a holomorphic map.

In another direction, we might hope to use the techniques here to study hyperbolic monopoles over a general hyperbolic manifold Y [4]. We have observed here that as the curvature of hyperbolic space tends to $-\infty$, the instantons concentrate at the boundary. It seems reasonable to guess that this would occur for general Y, particularly in light of the conjecture of Austin and Braam [3] that a hyperbolic monopole on Y is determined by its boundary values. Rather than actually take the limit, we can use this as intuition for a good initial guess for the heat flow. Since we reparametrise the normal bundle in the limit, it would mean that the initial metric need only be defined on infinite tubes at the boundary and set to be trivial on the interior of Y.

Acknowledgements

The second author would like to thank David Epstein and Mario Micallef for useful conversations.

References

[1] M.F. Atiyah, Magnetic monopoles and the Yang–Baxter equations, Internat. J. Modern Phys. 6 (1991) 2761–2774.
[2] M.F. Atiyah, Instantons in two and four dimensions, Comm. Math. Phys. 93 (1984) 437–451.
[3] D.M. Austin, P.J. Braam, Boundary values of hyperbolic monopoles, Nonlinearity 3 (1990) 809–823.
[4] P.J. Braam, A Kaluza–Klein approach to hyperbolic three-manifolds, Ens. Math. 34 (1988) 275–311.
[5] S.K. Donaldson, Anti-self-dual Yang–Mills connections over complex algebraic surfaces and stable vector bundles, Proc. Lond. Math. Soc. 30 (1985) 1–26.

[6] S.K. Donaldson, Instantons and geometric invariant theory, Comm. Math. Phys. 93 (1984) 453–460.
[7] S.K. Donaldson, Boundary-value problems for Yang–Mills fields, J. Geom. Phys. 8 (1992) 89–122.
[8] S.K. Donaldson, P.B. Kronheimer, The Geometry of Four-Manifolds, Oxford University Press, Oxford, 1990.
[9] S. Dostoglou, D. Salamon, Self-dual instantons and holomorphic curves, Ann. of Math. 139 (1994) 581–640.
[10] S. Dostoglou, D. Salamon, Instanton homology and symplectic fixed points, in: D. Salamon (Ed.), Symplectic Geometry, LMS Lecture Note Series, Cambridge Univ. Press, Cambridge, 1993, pp. 57–94.
[11] R.S. Hamilton, Harmonic Maps of Manifolds with Boundary, Lecture Notes in Mathematics, vol. 471, Springer, New York, 1975.
[12] S. Jarvis, Monopoles to rational maps via radial scattering, preprint, 1996.
[13] S. Jarvis, P. Norbury, Compactification of hyperbolic monopoles, Nonlinearity 10 (1997) 1073–1092.
[14] S. Jarvis, P. Norbury, Zero and infinite curvature limits of hyperbolic monopoles, Bull. LMS 29 (1997) 737–744.
[15] R. Mazzeo, J. Rade, Private Communication, 1996.
[16] P. Norbury, Periodic instantons, preprint, 1997.
[17] A.N. Pressley, G.B. Segal, Loop Groups, Oxford University Press, Oxford, 1984.
[18] L.M. Sibner, R.J. Sibner, Classification of singular Sobolev connections by their holonomy, Comm. Math. Phys. 144 (1992) 337–350.
[19] C.T. Simpson, Constructing variations of Hodge structure using Yang–Mills theory and applications to uniformization, J. Amer. Math. Soc. 1 (1988) 867–918.
[20] P.M. Sutcliffe, C.J. Houghton, N.S. Manton, Rational maps, monopoles and skyrmions, hep-th/9705151.