Tr. J. of Mathematics
20 (1996) , 61 – 73.

Proceedings of 4^{th} Gökova
Geometry-Topology Conference

Removable singularities and a vanishing theorem for Seiberg-Witten invariants

Dietmar Salamon

1. Introduction

This is an expository paper. The goal is to give a proof of the following vanishing theorem for the Seiberg-Witten invariants of connected sums of smooth 4-manifolds.

Theorem 1.1. *Suppose that X is a compact oriented smooth 4-manifold diffeomorphic to a connected sum $X_1\#X_2$ where*

$$b^+(X_1) \geq 1, \qquad b^+(X_2) \geq 1,$$

and $b^+(X) - b_1(X)$ is odd. Then the Seiberg-Witten invariants of X are all zero.

This result is the Seiberg-Witten analogue of Donaldson's original theorem about the vanishing of the instanton invariants [2] for connected sums. An outline of the proof of Theorem 1.1 was given by Donaldson in [1]. The key ingredient of the proof is a removable singularity theorem for the Seiberg-Witten equations on flat Euclidean 4-space. A proof of Theorem 1.1 was also indicated by Witten in his lecture on 6 December 1994 at the Isaac Newton Institute in Cambridge. The result was used by Kotschick in his proof that (simply connected) symplectic 4-manifolds are irreducible [4].

Seiberg-Witten equations on $\mathbb{R}^4$

Identify $\mathbb{R}^4$ with the quaternions $\mathbb{H}$ via $x = x_0 + ix_1 + jx_2 + kx_3$ and consider the standard spin^c structure $\Gamma : \mathbb{H} = T_x\mathbb{H} \to \mathbb{C}^{4\times 4}$ given by

$$\Gamma(\xi) = \begin{pmatrix} 0 & \gamma(\xi) \\ -\gamma(\xi)^* & 0 \end{pmatrix}, \qquad \gamma(\xi) = \begin{pmatrix} \xi_0 + i\xi_1 & \xi_2 + i\xi_3 \\ -\xi_2 + i\xi_3 & \xi_0 - i\xi_1 \end{pmatrix}.$$

Thus $\gamma(e_0) = 1\!\!1$, $\gamma(e_1) = I$, $\gamma(e_2) = J$, and $\gamma(e_3) = K$ with

$$I = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \qquad J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad K = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}.$$

Given a connection 1-form $A = \sum_j A_j dx_j$ with $A_j : \mathbb{H} \to i\mathbb{R}$ and a spinor $\Phi : \mathbb{H} \to \mathbb{C}^2$ denote

$$\nabla_A \Phi = \sum_{j=0}^{3} \nabla_j \Phi \, dx_j, \qquad \nabla_j \Phi = \frac{\partial \Phi}{\partial x_j} + A_j \Phi$$

The Seiberg-Witten equations have the form

$$D_A \Phi = 0, \quad \rho^+(F_A) = (\Phi\Phi^*)_0 \tag{1}$$

where $D_A = -\nabla_0 + I\nabla_1 + J\nabla_2 + K\nabla_3$ is the Dirac operator associated to the connection A, $F_A = dA = \sum_{i<j} F_{ij} dx_i \wedge dx_j$ is the curvature, and $\rho^+(F_A) \in \mathbb{C}^{2\times 2}$ is given by

$$\rho^+(F_A) = (F_{01} + F_{23})I + (F_{02} + F_{31})J + (F_{03} + F_{12})K.$$

Moreover, $(\Phi\Phi^*)_0$ denotes the traceless part of the matrix $\Phi\Phi^* \in \mathbb{C}^{2\times 2}$ and hence the second equation in (1) is equivalent to $F_{01} + F_{23} = -2^{-1}\Phi^* I \Phi$, $F_{02} + F_{31} = -2^{-1}\Phi^* J\Phi$, and $F_{03} + F_{12} = -2^{-1}\Phi^* K\Phi$. The energy of a pair (A, Φ) on an open set $\Omega \subset \mathbb{R}^4$ is given by

$$E(A, \Phi; \Omega) = \int_\Omega \left(\sum_{i=0}^{3} |\nabla_i \Phi|^2 + \frac{1}{4}|\Phi|^4 + \sum_{i<j} |F_{ij}|^2 \right).$$

It is invariant under the action of the gauge group $\mathrm{Map}(\Omega, S^1)$ by $(A, \Phi) \mapsto (u^*A, u^{-1}\Phi)$ where $u^*A = u^{-1}du + A$. The proof of Theorem 1.1 relies on the following removable singularity theorem for the finite energy solutions of (1). Denote the unit ball in $\mathbb{R}^4$ by $B = B^4 = \{x \in \mathbb{R}^4 \,|\, |x| \le 1\}$. If $\Phi = 0$ then the result reduces to Uhlenbeck's removable singularity theorem for ASD instantons in the case of the gauge group $\mathrm{G} = S^1$ (cf. Uhlenbeck [10] and Donaldson–Kronheimer [2], pp 58–72 and 166-170).

Theorem 1.2 (Removable singularities). *Let $A \in \Omega^1(B - \{0\}, i\mathbb{R})$ and $\Phi \in C^\infty(B - \{0\}, \mathbb{C}^2)$ satisfy (1) with*

$$E(A, \Phi; B) < \infty.$$

Then there exists a gauge transformation $u : B - \{0\} \to S^1$ such that $u(x) = 1$ for $|x| = 1$ and u^*A and $u^{-1}\Phi$ extend to a smooth solution of (1) over B.

The following three fundamental identities will play a crucial role in the proof of Theorem 1.2. The first is the Weitzenböck formula

$$D_A{}^* D_A \Phi + \sum_{i=0}^{3} \nabla_i \nabla_i \Phi = \rho^+(F_A)\Phi \tag{2}$$

where $D_A{}^* = \nabla_0 + I\nabla_1 + J\nabla_2 + K\nabla_3$. The second is the energy identity

$$\begin{aligned} E(A, \Phi; \Omega) \;&=\; \int_\Omega \left(|D_A \Phi|^2 + \left|\rho^+(F_A) - (\Phi\Phi^*)_0\right|^2 \right) \\ &\quad + \int_{\partial\Omega} A \wedge dA + \int_{\partial\Omega} \langle \Phi, \nabla_{A,\nu}\Phi + \Gamma(\nu) D_A \Phi \rangle \, \mathrm{dvol}_{\partial\Omega} \end{aligned} \tag{3}$$

for $A \in \Omega^1(\mathbb{R}^4, i\mathbb{R})$ and $\Phi \in C^\infty(\mathbb{R}^4, \mathbb{C}^2)$. Here we use the norm $|T|^2 = \frac{1}{2}\mathrm{trace}(T^*T)$ for complex 2×2-matrices so that $1\!\!1$, I, J, K form an orthonormal basis of $\mathbb{C}^{2\times 2}$. Moreover, $\nu : \partial\Omega \to \mathbb{R}^4$ denotes the outward unit normal vector field, $\nabla_{A,\nu}\Phi = \sum_i \nu_i \nabla_i \Phi$, and $\Gamma(\nu) = -\nu_0 1\!\!1 + \nu_1 I + \nu_2 J + \nu_3 K$. The third equation is

$$\Delta|\Phi|^2 = -2|\nabla_A\Phi|^2 - |\Phi|^4 \tag{4}$$

for solutions of (1) where $\Delta = -\sum_i \partial^2/\partial {x_i}^2$. It is proved by direct computation using (2) and $\rho^+(F_A)\Phi = (\Phi\Phi^*)_0\Phi = |\Phi|^2\Phi/2$. Equation (4) was first noted by Kronheimer and Mrowka in [5] and lies at the heart of their compactness proof for the solutions of (1).

Proof of the energy identity: The proof relies on the familiar equation

$$\int_\Omega \left(|F_A|^2 - 2|F_A^+|^2\right) = \int_\Omega F_A \wedge F_A = \int_{\partial\Omega} A \wedge dA,$$

and on the formula

$$\int_\Omega \left(|\nabla_A\Phi|^2 - |D_A\Phi|^2\right) = \int_{\partial\Omega} \langle \Phi, \nabla_{A,\nu}\Phi + \Gamma(\nu)D_A\Phi\rangle - \int_\Omega \langle \Phi, \rho^+(F_A)\Phi\rangle.$$

This last equation follows from Stokes' theorem and (2). With $|\rho^+(F_A)|^2 = 2|F_A^+|^2$ and $\langle \Phi, \rho^+(F_A)\Phi\rangle = 2\langle \rho^+(F_A), (\Phi\Phi^*)_0\rangle$ the rest of the proof is an easy exercise. □

2. Removable singularities for 1-forms

The first step in the proof of Theorem 1.2 is the following weak removable singularity theorem for 1-forms on $\mathbb{R}^n$. The theorem asserts that if α is a 1-form on the punctured ball $B^n - \{0\}$ such that $d\alpha$ is of class L^2 then there exists a function $\xi : B^n - \{0\} \to \mathbb{R}$ such that $\alpha - d\xi$ is of class $W^{1,2}$ (and $d^*(\alpha - d\xi) = 0$). If $n = 4$ and α is anti-self-dual then it follows easily that $\alpha - d\xi$ extends to a smooth 1-form on B^4. This is Uhlenbeck's removable singularity theorem for ASD instantons in the case $\mathrm{G} = S^1$. Note also that this is the special case $\Phi = 0$ in Theorem 1.2. Even though this result is simply a special case of Uhlenbeck's theorem we give a proof below which is is specific to the abelian case and is considerably simpler than both Ulenbeck's original proof in [10] and the proof given by Donaldson and Kronheimer in [2]. Throughout denote by $B^n(r) = \{x \in \mathbb{R}^n \,|\, |x| \le r\}$ the closed ball in $\mathbb{R}^n$ of radius r and abbreviate $B^n = B^n(1)$ and $A(r_0, r_1) = A^n(r_0, r_1) = \{x \in \mathbb{R}^n \,|\, r_0 \le |x| \le r_1\}$ for $r_0 < r_1$.

Proposition 2.1 (Uhlenbeck). *Assume $n \ge 4$ and let $\alpha \in \Omega^1(B^n - \{0\})$ be a smooth real valued 1-form which satisfies*

$$\int_{B^n} |d\alpha|^2 < \infty.$$

Then there exists a smooth function $\xi : B^n - \{0\} \to \mathbb{R}$ such that $\alpha - d\xi$ is of class $W^{1,2}$ on the (unpunctured) unit ball and satisfies

$$\int_{B^n} \left(|\nabla(\alpha - d\xi)|^2 + \frac{|\alpha - d\xi|^2}{|x|^2}\right) \le 4 \int_{B^n} |d\alpha|^2$$

as well as

$$d^*(\alpha - d\xi) = 0, \qquad \frac{\partial \xi}{\partial \nu} = \alpha(\nu).$$

Here $d\xi/\partial\nu$ denotes the normal derivative on ∂B^n and $\alpha(\nu) = \sum_i \alpha_i(x)x_i$ for $|x| = 1$.

Note that addition of any exact 1-form on $B^n - \{0\}$ does not alter the L^2-norm of $d\alpha$. Thus the behaviour of α near zero may be extremely singular. The proposition asserts that there exists an exact 1-form $d\xi$ on $B^n - \{0\}$ which *tames* the singularity at 0 in the sense that $\alpha - d\xi$ is of class $W^{1,2}$ on B^n. The function ξ will be constructed as a limit of functions $\xi_\varepsilon : B^n(1) - B^n(\varepsilon) \to \mathbb{R}$ which satisfy $d^*(\alpha - d\xi_\varepsilon) = 0$ with boundary condition $\partial\xi_\varepsilon/\partial\nu = \alpha(\nu)$ on $\partial(B_1 - B_\varepsilon)$. The convergence proof relies on the following four lemmata.

Lemma 2.2. *Assume $n \geq 4$. Then every smooth 1-form $\alpha \in \Omega^1(A^n(\varepsilon,1))$ with $\alpha(\nu) = 0$ on $\partial A^n(\varepsilon,1)$ satisfies the inequality*

$$\int_{A(\varepsilon,1)} \left(|\nabla\alpha|^2 + \frac{|\alpha|^2}{|x|^2}\right) \leq 4\int_{A(\varepsilon,1)} \left(|d\alpha|^2 + |d^*\alpha|^2\right).$$

Proof: Let $\alpha = \sum_i \alpha_i dx_i$ be a smooth 1-form on a domain $\Omega \subset \mathbb{R}^n$ with smooth boundary. Suppose that $\langle \alpha, \nu\rangle = \sum_{i=1}^n \alpha_i\nu_i = 0$ on $\partial\Omega$. This condition is equivalent to $*\alpha|_{\partial\Omega} = 0$. Integration by parts shows that

$$\|\nabla\alpha\|^2 - \|d\alpha\|^2 - \|d^*\alpha\|^2 = \int_{\partial\Omega}\left\langle \alpha, \frac{\partial\alpha}{\partial\nu}\right\rangle \mathrm{dvol}_{\partial\Omega} - \int_{\partial\Omega} \alpha\wedge *d\alpha.$$

Here all norms on the left are L^2-norms on $A(\varepsilon,1)$. Now use the formulae $*dx_i|_{\partial\Omega} = \nu_i\mathrm{dvol}_{\partial\Omega}$ and $dx_i \wedge *(dx_i\wedge dx_j) = -*dx_j$ for $i < j$ to obtain

$$\int_{\partial\Omega} \alpha\wedge *d\alpha - \int_{\partial\Omega}\left\langle \alpha, \frac{\partial\alpha}{\partial\nu}\right\rangle \mathrm{dvol}_{\partial\Omega} = \int_{\partial\Omega}\sum_{i,j}\alpha_i\alpha_j\frac{\partial\nu_j}{\partial x_i}\,\mathrm{dvol}_{\partial\Omega}.$$

This equation uses the fact that $\sum_i \alpha_i\nu_i = 0$ on $\partial\Omega$ and $\alpha = (\alpha_1,\ldots,\alpha_n)$ is tangent to $\partial\Omega$. In the case $\Omega = A(\varepsilon,1)$ the last two identities combine to

$$\|\nabla\alpha\|^2 = \|d\alpha\|^2 + \|d^*\alpha\|^2 + \frac{1}{\varepsilon}\int_{|x|=\varepsilon}|\alpha|^2 - \int_{|x|=1}|\alpha|^2 \tag{5}$$

for 1-forms on $A(\varepsilon,1)$ which satisfy $\langle\alpha,\nu\rangle=0$ on the boundary. Now consider the function $f(x)=x/|x|^2$ with $\mathrm{div}(f)=(n-2)/|x|^2$. Then for every smooth function $u:A(\varepsilon,1)\to\mathbb{R}$

$$\begin{aligned}
\frac{1}{\varepsilon}\int_{|x|=\varepsilon}|u|^2-\int_{|x|=1}|u|^2 &= -\int_{\partial A(\varepsilon,1)}\langle\nu,f\rangle|u|^2\,\mathrm{dvol}\\
&= -\int_{A(\varepsilon,1)}\sum_{i=1}^n\frac{\partial}{\partial x_i}(f_i|u|^2)\\
&= -\int_{A(\varepsilon,1)}\sum_{i=1}^n\left(2f_iu\frac{\partial u}{\partial x_i}+|u|^2\frac{\partial f_i}{\partial x_i}\right)\\
&\le 2\int_{A(\varepsilon,1)}\frac{|u|\,|\nabla u|}{|x|}-\int_{A(\varepsilon,1)}\mathrm{div}(f)|u|^2\\
&= 2\int_{A(\varepsilon,1)}\frac{|u|\,|\nabla u|}{|x|}-(n-2)\int_{A(\varepsilon,1)}\frac{|u|^2}{|x|^2}\\
&\le \delta\int_{A(\varepsilon,1)}|\nabla u|^2-\left(n-2-\frac{1}{\delta}\right)\int_{A(\varepsilon,1)}\frac{|u|^2}{|x|^2}.
\end{aligned}$$

The last inequality holds for any constant $\delta>0$. If $n\ge4$ we can choose $1/(n-2)<\delta<1$. For example, with $\delta=3/4$ we obtain from (5)

$$\|\nabla\alpha\|^2\le\|d\alpha\|^2+\|d^*\alpha\|^2+\frac{3}{4}\|\nabla\alpha\|^2-\left(n-2-\frac{4}{3}\right)\int_{A(\varepsilon,1)}\frac{|\alpha|^2}{|x|^2}.$$

This holds for all n. But for $n\ge4$ the last term on the right is negative and the desired inequality follows. □

Lemma 2.3 (Poincaré's inequality). *There is a constant $c=c(n)>0$ such that every smooth function $\xi:A^n(1/2,1)\to\mathbb{R}$ with mean value zero satisfies the inequality*

$$\int_{A(1/2,1)}|\xi|^2\le c\int_{A(1/2,1)}|d\xi|^2.$$

Lemma 2.4. *Every smooth function $\xi:A^n(r_0,r_1+t)\to\mathbb{R}$ satisfies*

$$\int_{A(r_0,r_1)}|\xi|^2\le2\int_{A(r_0+t,r_1+t)}|\xi|^2+\int_{A(r_0,r_1+t)}|d\xi|^2$$

for $0<r_0<r_1\le1$ and $0\le t\le1$.

Proof: Consider the identity

$$\xi(rx)=\xi((t+r)x)-\int_0^t\langle\nabla\xi((r+s)x),x\rangle\,ds$$

and use the Cauchy-Schwartz inequality to obtain

$$|\xi(rx)|^2\le2\,|\xi((t+r)x)|^2+\frac{2}{(n-2)r^{n-2}}\int_r^{r+t}s^{n-1}|d\xi(sx)|^2\,ds$$

for $|x| = 1$ and $n \geq 3$. In the case $n = 2$ there is a similar inequality with $1/(n-2)r^{n-2}$ replaced by $\log(r+t) - \log r \leq r - \log r$. Now multiply by r^{n-1} and integrate over S^{n-1} and over $r_0 \leq r \leq r_1$. □

Lemma 2.5. *Let $u : B^n - \{0\} \to \mathbb{R}$ be a smooth function such that*

$$\int_{B^n} |\nabla u(x)|^2 < \infty.$$

Then u is of class $W^{1,2}$ on B^n, i.e. its distributional derivatives exist and agree with the ordinary derivatives.

Proof: For any compactly supported test function $\varphi : B^n \to \mathbb{R}$ integrate the function $u\partial_i\varphi + \varphi\partial_i u$ over the annulus $\varepsilon \leq |x| \leq 1$ and show that the boundary integral over $|x| = \varepsilon$ converges to zero as $\varepsilon \to 0$. □

Proof of Proposition 2.1: For every $\varepsilon > 0$ there exists a smooth function $\xi_\varepsilon : A^n(\varepsilon, 1) \to \mathbb{R}$ which satisfies

$$d^*(\alpha - d\xi_\varepsilon) = 0, \qquad \frac{\partial \xi_\varepsilon}{\partial \nu} = \langle \alpha, \nu \rangle$$

where the last equation holds on the boundary. The function ξ_ε is only determined up to a constant which can be fixed by the normalization condition

$$\int_{1/2 \leq |x| \leq 1} \xi_\varepsilon(x)\, dx = 0.$$

It follows from Lemma 2.2 that

$$\|\nabla(\alpha - d\xi_\varepsilon)\|^2_{L^2(A(\varepsilon,1))} + \int_{\varepsilon \leq |x| \leq 1} \frac{|\alpha - d\xi_\varepsilon|^2}{|x|^2} \leq 4\, \|d\alpha\|^2_{L^2(A(\varepsilon,1))}.$$

Fix some number $\delta > 0$. Then for $\varepsilon < \delta$

$$\|\nabla d\xi_\varepsilon\|_{L^2(A(\delta,1))} \leq 2\, \|d\alpha\|_{L^2} + \|\nabla\alpha\|_{L^2(A(\delta,1))},$$

$$\|d\xi_\varepsilon\|_{L^2(A(\delta,1))} \leq 2\, \|d\alpha\|_{L^2} + \|\alpha\|_{L^2(A(\delta,1))}.$$

Now use Lemma 2.3 and the mean value condition to control the L^2-norm of ξ_ε on $A(1/2, 1)$ and Lemma 2.4 to control this norm on $A(\delta, 1/2)$. This shows that for every $\delta > 0$ there exists a constant $c_\delta > 0$ such that

$$\|\xi_\varepsilon\|_{W^{2,2}(A(\delta,1))} \leq c_\delta$$

for every $\varepsilon \in (0, \delta)$. Now the usual diagonal sequence argument shows that there exists a sequence $\varepsilon_i \to 0$ such that ξ_{ε_i} converges strongly in $W^{1,2}(K)$ and weakly in $W^{2,2}(K)$ for every compact subset $K \subset B^n - \{0\}$. The limit function $\xi : B^n - \{0\} \to \mathbb{R}$ is of class $W^{2,2}$ on every compact subset away from 0 and satisfies $d^*(\alpha - d\xi) = 0$ and $\langle \alpha - d\xi, \nu \rangle = 0$. Hence Lemma 2.2 shows that

$$\int_K \left(|\nabla(\alpha - d\xi)|^2 + \frac{|\alpha - d\xi|^2}{|x|^2} \right) \leq 4 \int_{B^n} |d\alpha|^2$$

for every compact subset $K \subset B^n - \{0\}$. By Lemma 2.5, $\alpha - d\xi$ is of class $W^{1,2}$ on B^n. This proves the proposition. □

3. Proof of the removable singularity theorem

By Proposition 2.1 there exists a smooth function $\xi : B^4 - \{0\} \to i\mathbb{R}$ such that $A - d\xi$ is of class $W^{1,2}$ on the closed ball B^4 and $d^*(A - d\xi) = 0$. Hence we may assume from now on that $A \in W^{1,2}$ and $d^*A = 0$. Moreover, by the finite energy condition, we have $\Phi \in L^4$ and $\nabla_i\Phi \in L^2$. The Sobolev embedding theorem shows that $A \in L^4$ and hence

$$\partial_i\Phi = \nabla_i\Phi - A_i\Phi \in L^2$$

for $i = 0, 1, 2, 3$. By Lemma 2.5, this shows that $\Phi \in W^{1,2}$. Thus we have a solution (A, Φ) of (1) which is smooth on the punctured ball $B^4 - \{0\}$ and on the closed ball satisfies

$$A \in W^{1,2}, \qquad \Phi \in W^{1,2}, \qquad d^*A = 0.$$

We shall prove in three steps that there exists a constant $c > 0$ such that

$$E_0(A, \Phi; B_r) = \int_{|x|\le r} \left(|\nabla_A\Phi|^2 + \frac{1}{2}|\Phi|^4\right) \le cr^2. \tag{6}$$

Step 1: *For every* $r \in (0, 1]$

$$E_0(A, \Phi; B_r) = \int_{|x|=r} \sum_i \langle \Phi, \nabla_i\Phi\rangle \frac{x_i}{r}.$$

Let $\Omega \subset \mathbb{R}^4$ be any open domain with smooth boundary such that A and Φ are defined on its closure. (Thus $0 \notin \bar{\Omega}$.) Consider the energy

$$E_0(A, \Phi; \Omega) = \int_\Omega \left(|\nabla_A\Phi|^2 + \frac{1}{4}|\Phi|^4 + 2|F_A^+|^2\right) = \int_{\partial\Omega} \langle \Phi, \nabla_{A,\nu}\Phi\rangle.$$

The first equality follows from the fact that $|\Phi|^4 = 8|F_A^+|^2$ for solutions of (1) and the second equality follows from the energy identity (3). Abbreviate

$$f(r) = \int_{|x|=r} \sum_i \langle \Phi, \nabla_i\Phi\rangle \frac{x_i}{r}.$$

Then $f : (0, 1] \to \mathbb{R}$ is a smooth function and the previous identity shows that

$$E_0(A, \Phi; B_r - B_\varepsilon) = f(r) - f(\varepsilon).$$

Hence f is monotonically increasing and bounded below. This shows that the limit $f(0) := \lim_{\varepsilon\to 0} f(\varepsilon)$ exists. Now it follows from the finiteness of the energy that $\Phi \in L^4$ and $\nabla_i\Phi \in L^2$ and hence $\langle \Phi, \nabla_i\Phi\rangle \in L^{4/3}$ for all i. Moreover, by Hölder's inequality,

$$|f(r)|^{4/3} \le (2\pi^2)^{1/3} r \int_{|x|=r} (|\Phi|\,|\nabla_A\Phi|)^{4/3}$$

and hence

$$\int_0^1 \frac{|f(r)|^{4/3}}{r}\, dr < \infty.$$

This shows that there must be a sequence $\varepsilon_i \to 0$ with $f(\varepsilon_i) \to 0$ and it follows that $f(0) = 0$. This implies $f(r) = E_0(A, \Phi; B_r)$ as claimed.

Step 2: *Every smooth function* $u : \mathbb{R}^4 - \{0\} \to \mathbb{R}$ *satisfies the identity*

$$-\int_{\rho\le|x|\le r} \frac{\Delta u}{|x|^2} = \int_{|x|=r} \frac{2u + \langle \nabla u, x\rangle}{r^3} - \int_{|x|=\rho} \frac{2u + \langle \nabla u, x\rangle}{\rho^3}.$$

This is Stokes' theorem on the annulus $\rho \le |x| \le r$ with $\Delta v = -\sum_i \partial^2 v/\partial {x_i}^2 = 0$ for $v(x) = 1/|x|^2$.

Step 3: *Proof of (6).*

Recall from (4) that $\Delta|\Phi|^2 = -2|\nabla_A\Phi|^2 - |\Phi|^4$. Moreover, note that

$$\int_{|x|=r} \langle \nabla|\Phi|^2, x\rangle = 2\int_{|x|=r} \sum_i \langle \Phi, \nabla_i\Phi\rangle x_i = 2rf(r).$$

Hence it follows from Step 2 with $u = |\Phi|^2$ that

$$\int_{\rho\le|x|\le r} \frac{2|\nabla_A\Phi|^2 + |\Phi|^4}{|x|^2}\, dx = \int_{|x|=r} \frac{2|\Phi|^2}{r^3} + \frac{2f(r)}{r^2} - \int_{|x|=\rho} \frac{2|\Phi|^2}{\rho^3} - \frac{2f(\rho)}{\rho^2}.$$

This implies

$$\frac{f(\rho)}{\rho^2} \le \frac{f(r)}{r^2} + \frac{1}{r^3}\int_{|x|=r} |\Phi|^2$$

for $0 < \rho \le r$ and (6) follows.

By (4), the function $x \mapsto |\Phi(x)|^4$ is subharmonic and hence

$$|\Phi(x)|^4 \le \frac{2}{\pi^2 r^4}\int_{B_r(x)} |\Phi|^4 \le \frac{2}{\pi^2 r^4} E_0(A, \Phi; B_{2r}) \le \frac{8c}{\pi^2 r^2}$$

for $r = |x|$. The first inequality is the mean value inequality for subharmonic functions, the second follows from the definition of E_0, and the last follows from (6). Thus

$$|\Phi(x)|^4 \le \frac{8c}{\pi^2 |x|^2}$$

and, since the function $x \mapsto 1/|x|^\alpha$ is integrable in a neighbourhood of zero whenever $\alpha < 4$, it follows that $|\Phi|^p$ is integrable for every $p < 8$. Thus we have proved that $|\Phi|^2 \in L^p$ for any $p < 4$. Since $d^+A = \sigma^+((\Phi\Phi^*)_0)$ this shows that $d^+A \in L^p$ for any $p < 4$. Now recall that $d^*A = 0$ and hence

$$\Delta A = d^*dA = 2d^*d^+A = 2d^*\sigma^+((\Phi\Phi^*)_0).$$

Note that A is a weak solution of this equation on the closed (unpunctured) ball and hence it follows that $A \in W^{1,p}$ for any $p < 4$. Thus $A \in L^q$ for any $q < \infty$. The formula

$$0 = D_A\Phi = D\Phi - \Gamma(A)\Phi$$

with $\Gamma(A)\Phi \in L^p$ now shows that $\Phi \in W^{1,p}$ for any $p < 4$. Thus $\Phi \in L^q$ for some $q > 4$ and using the last equation again with $\Gamma(A)\Phi \in L^q$ we find that $\Phi \in W^{1,q}$ for some $q > 4$. This implies $d^*\sigma^+((\Phi\Phi^*)_0) \in L^q$ and, by the previous equation $A \in W^{2,q}$. Using the two equations alternatingly we conclude that A and Φ are smooth on B_1. This is a standard elliptic bootstrapping argument and completes the proof of Theorem 1.2.

4. Proof of the vanishing theorem

The goal of this section is to prove Theorem 1.1. The proof given here was outlined by Donaldson in [1]. It is based on choosing a sequence of metrics g_ν on the connected sum $X_1 \# X_2$ which *pinches* the neck to a point and has the property that the scalar curvature s_ν is bounded below by a constant independent of ν. Note, however, that the scalar curvature will diverge to $+\infty$ near the *pinched neck*. More precisely, the following remark shows how to construct a metric on the unit disc in $\mathbb{R}^4$ which agrees with the standard metric outside a ball of radius δ and with the pullback metric from $\mathbb{R} \times \varepsilon S^3$ under the diffeomorphism $x \mapsto (\varepsilon \log |x|, \varepsilon x/|x|)$ inside a punctured ball of radius δ^{m+1} for some integer m.

Remark 4.1. Consider the diffeomorphism

$$f : \mathbb{R}^4 - \{0\} \to \mathbb{R} \times \varepsilon S^3, \qquad f(x) = \left(\varepsilon \log |x|, \varepsilon \frac{x}{|x|}\right).$$

It is easy to see that the pullback of the standard product metric g_ε on $\mathbb{R} \times \varepsilon S^3$ under this diffeomorphism is given by

$$f^* g_\varepsilon(\xi, \eta) = \frac{\varepsilon^2}{|x|^2}\langle \xi, \eta\rangle$$

for $|x| \leq \varepsilon^2$. Now choose a function $\lambda : (0,1] \to [1, \infty)$ which satisfies

$$\lambda(r) = \begin{cases} \varepsilon/r & \text{if } r \leq \delta^{m+1}, \\ 1 & \text{if } r \geq \delta. \end{cases} \tag{7}$$

Consider the metric

$$g_\lambda(\xi, \eta) = \lambda(|x|)^2 \langle \xi, \eta \rangle.$$

Note that for $|x| \leq \delta^{m+1}$ this metric agrees with the above pullback metric $f^* g_\varepsilon$. The scalar curvature of g_λ is given by

$$s_\lambda = 6\frac{\Delta\lambda}{\lambda^3} = -6\frac{\lambda'' + 3\lambda'/r}{\lambda^3}.$$

One can choose λ decreasing and thus $\lambda'(r) \leq 0$ for all r. It remains to prove that λ can be chosen such that (7) is satisfied and, say,

$$\frac{\lambda''(r)}{\lambda(r)} + 3\frac{\lambda'(r)}{r\lambda(r)} \leq 1. \tag{8}$$

Here the constant 1 is an arbitrary choice and can be replaced by any positive number. We must prove that for every $\delta > 0$ there exists a function $\lambda : [0,1] \to [0,\infty)$ which satisfies (7) and (8) for some constant $\varepsilon > 0$. Following Micallef and Wang [7] we introduce a function $\alpha = \alpha(r)$ by

$$\frac{\lambda'}{\lambda} = -\frac{\alpha}{r}, \qquad \frac{\lambda''}{\lambda} = -\frac{\alpha'}{r} + \frac{\alpha + \alpha^2}{r^2}.$$

Then the conditions (7) and (8) take the form

$$\alpha(r) = \begin{cases} 1, & \text{for } r \leq \delta^{m+1}, \\ 0, & \text{for } r \geq \delta, \end{cases} \tag{9}$$

$$\frac{\alpha'}{r} + \frac{\alpha(2-\alpha)}{r^2} \geq -1. \tag{10}$$

Consider the curve $\gamma(t) = \alpha(\delta e^{-t})$. Then (10) translates into

$$\dot\gamma \leq (2-\gamma)\gamma + \delta^2 e^{-2t}$$

and (9) reads $\gamma(t) = 1$ for $t \geq T = \log(\delta^{-m})$ and $\gamma(t) = 0$ for $t \leq 0$. A solution of the differential equation $\dot\gamma = (2-\gamma)\gamma$ is given by the explicit formula

$$\gamma(t) = \frac{2\delta^{2m}e^{2t}}{1+\delta^{2m}e^{2t}}.$$

This solution satisfies $\gamma(0) = 2\delta^{2m}/(1+2\delta^{2m}) \ll 1$ and $\gamma(T) = \gamma(\log(\delta^{-m})) = 1$. Perturbing this function slightly near $t = 0$ and $t = T$ gives a smooth solution of the required differential inequality provided that m is sufficiently large. Note that essentially the same argument can be used to prove the theorem of Gromov and Lawson about positive scalar curvature for connected sums [3]. □

Recall that the solutions of the Seiberg-Witten equations for a spin^c structure $\Gamma : TX \to \mathrm{End}(W)$ form a moduli space space $\mathcal{M}(X,\Gamma,g,\eta)$ which, for a generic perturbation η, is a finite dimensional compact manifold of dimension

$$\dim \mathcal{M}(X,\Gamma,g,\eta) = \frac{c\cdot c}{4} - \frac{2\chi + 3\sigma}{4}$$

where $\chi = \chi(X)$ and $\sigma = \sigma(X)$ denote the Euler characteristic and signature of X and $c = c_1(L_\Gamma) \in H^2(X,\mathbb{Z})$ is the characteristic class of the spin^c structure. It is convenient to think of the connected sum as follows. Fix two points $x_1 \in X_1$ and $x_2 \in X_2$ and choose a metric g_i on X_i which is flat in a neighbourhood of x_i. Now construct a sequence of manifolds $X_\nu = X_1 \#_\nu X_2$ by removing arbitrarily small discs from X_1 and X_2, centered at x_1 and x_2 respectively, modifying the metrics g_i as in Remark 4.1 above, and then identifying two annuli which are isometric to $[0,1] \times \varepsilon_\nu S^3$. Given two spin^c structures Γ_1

over X_1 and Γ_2 over X_2 one obtains a corresponding sequence of spin^c structures Γ_ν over X_ν by identifying Γ_1 and Γ_2 in suitable trivializations over the two annuli. Let us choose a sequence of perturbations η_ν on X_ν which vanish near the *neck* and are independent of ν on the complement of the neck. Any such sequence determines two fixed perturbations η_1 and η_2 on X_1 and X_2, respectively, which vanish in the given neighbourhoods of x_1 and x_2. In [8], Chapter 9, it is proved that the perturbation can be chosen such that the moduli spaces $\mathcal{M}(X_1, \Gamma_1, g_1, \eta_1)$ and $\mathcal{M}(X_2, \Gamma_2, g_2, \eta_2)$ are regular.

Assume first that the moduli space $\mathcal{M}(X_\nu, \Gamma_\nu, g_\nu, \eta_\nu)$ is zero dimensional. We prove that this space must be empty for ν sufficiently large. Suppose otherwise that for every ν there exists a solution (A_ν, Φ_ν) of the Seiberg-Witten equations for the metric g_ν and the perturbation η_ν. In [5] Kronheimer and Mrowka proved that the spinors Φ_ν satisfy the inequality

$$\sup_X |\Phi_\nu| \leq -\frac{1}{2} \inf_X s_\nu.$$

where s_ν denotes the scalar curvature of g_ν (see also [8]). The previous exercise shows that there exists a constant $c > 0$ such that $s_\nu(x) \geq -c$ for all $x \in X$ and all ν. Hence the Φ_ν are uniformly bounded. Now A_ν and Φ_ν restrict to solutions of the Seiberg-Witten equations on X_1 (for the metric g_1 and the perturbation η_1) outside any neighbourhood of x_1. Hence it follows from the compactness theorem in [5] (see also [8], Chapter 9) that there exists a subsequence which converges in the C^∞-topology on every compact subset of $X_1 - \{x_1\}$ to a solution (A_1, Φ_1) of the Seiberg-Witten equations which is defined on $X_1 - \{x_1\}$ and has finite energy. Since g_1 is flat and η_1 vanishes near x_1 the removable singularity theorem 1.2 asserts that A_1 and Φ_1 extend to a smooth solution over all of X_1. This shows that the moduli space $\mathcal{M}_1 = \mathcal{M}(X_1, \Gamma_1, g_1, \eta_1)$ is nonempty. Obviously, the same argument applies to X_2. Now the perturbation η was chosen such that η_1 and η_2 are regular for g_1 and g_2. But the dimension formula shows that

$$0 = \dim \mathcal{M} = \dim \mathcal{M}_1 + \dim \mathcal{M}_2 + 1.$$

Hence one of the moduli spaces must have negative dimension. Since both moduli spaces are regular it follows that one of them must be empty, a contradiction. This shows that the assumption that $\mathcal{M}(X_\nu, \Gamma_\nu, g_\nu, \eta_\nu)$ was nonempty for all ν must have been false. But if there is a metric for which the moduli space is empty then the Seiberg-Witten inveriant is zero. Thus we have proved that the Seiberg-Witten invariant must vanish whenever the moduli space is zero dimensional.

A similar argument applies to the cut-down moduli spaces when dim $\mathcal{M} > 0$. For this case it is useful to intersect the moduli space $\mathcal{M}_1$, say, with suitable submanifolds of the form

$$\mathcal{N}_h = \left\{ [A, \Phi] \,\Big|\, \int_{X_1} \langle h(A), \Phi \rangle \mathrm{dvol} = 0 \right\} \subset \mathcal{C}(\Gamma_1) = \frac{\mathcal{A}(\Gamma_1) \times C^\infty(X, W_1^+)^*}{\mathrm{Map}(X, S^1)}$$

where $h : \mathcal{A}(\Gamma_1) \to C^\infty(X, W_1^+)^*$ satisfies

$$h(u^* A) = u(y) u^{-1} h(A)$$

for every gauge transformation $u : X_1 \to S^1$ and some $y \in X_1$. The map h can be localized near y as follows. For every 1-form $\alpha \in \Omega^1(X, i\mathbb{R})$ and every smooth path $\gamma : [0,1] \to X$ consider the *holonomy* $\rho_\alpha(\gamma) \in S^1$ defined by

$$\rho_\alpha(\gamma) = \exp\left(\int_\gamma \alpha\right).$$

For each point $x \in X_1$ near y let $\gamma_x : [0,1] \to X_1$ denote the path running from x to y in a straight line in a local chart. Fix a reference connection A_0 and a nonzero section $\Psi \in C^\infty(X_1, W_1^+)$ with support in the given neighbourhood of y. Then the map

$$h(A)(x) = \rho_{A-A_0}(\gamma_x)\Psi(x)$$

has the required properties. Now, as before, dim $\mathcal{M}$ = dim $\mathcal{M}_1$ + dim $\mathcal{M}_2 + 1$ and hence one of the moduli spaces must have dimension strictly smaller than $\mathcal{M}$. Suppose without loss of generality that

$$\dim \mathcal{M}_1 < \dim \mathcal{M} = 2d$$

and choose d functions $h_1, \ldots, h_d : \mathcal{A}(\Gamma_1) \to C^\infty(X, W_1^+)^*$ as above which are localized somewhere on X_1 away from x_1. Then, for a generic perturbation η_1,

$$\mathcal{M}(X_1, \Gamma_1, g_1, \eta_1) \cap \mathcal{N}_{h_1} \cap \cdots \cap \mathcal{N}_{h_d} = \emptyset.$$

On the other hand the h_i determine functions

$$h_{i,\nu} : \mathcal{A}(\Gamma_\nu) \to C^\infty(X, W_\nu^+)^*$$

(defined by the same formula) and one can examine the moduli spaces

$$\mathcal{M}(X_\nu, \Gamma_\nu, g_\nu, \eta_\nu) \cap \mathcal{N}_{h_{1,\nu}} \cap \cdots \cap \mathcal{N}_{h_{d,\nu}}.$$

If these are nonempty for all ν then it follows as above that the space $\mathcal{M}_1 \cap \mathcal{N}_{h_1} \cap \cdots \cap \mathcal{N}_{h_d}$ is nonempty contradicting the choice of the perturbation η_1. Hence these moduli spaces are empty for large ν and thus the Seiberg-Witten invariants are zero.

References

[1] S.K. Donaldson, The Seiberg-Witten equations and 4-manifold topology, Preprint, June 1995.

[2] S.K. Donaldson and P.B. Kronheimer, *The Geometry of Four-Manifolds*, Oxford University Press, 1990.

[3] M. Gromov and H.B. Lawson, The classification of simply connected manifolds of positive scalar curvature, *Ann. of Math.* **111** (1980), 423–434.

[4] D. Kotschick, On irreducible 4-manifolds, Preprint, 1995.

[5] P. Kronheimer and T.S. Mrowka, The genus of embedded surfaces in the projective plane, *Math. Res. Letters* **1** (1994), 797–808.

[6] D. McDuff and D. Salamon, *Introduction To Symplectic Topology*, to appear in Oxford University Press, 1995.

[7] M. Micallef and M.Y. Wang, Metrics with nonnegative isotropic curvature, *Duke Mathematics Journal* **72** (1993), 649–672.

[8] D. Salamon, *Spin geometry and Seiberg-Witten invariants*, in preparation.

[9] C.H. Taubes, The Seiberg-Witten invariants and symplectic forms, *Math. Res. Letters* **1** (1994), 809–822.

[10] K. Uhlenbeck, Removable singularities in Yang-Mills fields, *Commun. Math. Phys.* **83** (1982), 11–29.
[11] K. Uhlenbeck, Connections with L^p bounds on curvature, *Commun. Math. Phys.* **83** (1982), 31–42.
[12] E. Witten, Monopoles and 4-manifolds, Preprint, hep-th/9411102, November, 1994.

MATH. INST., UNIV. OF WARWICK, COVENTRY CV4 7AL, GREAT BRITAIN
E-mail address: das@maths.warwick.ac.uk