J. DIFFERENTIAL GEOMETRY
70 (2005) 59-112

FOUR-MANIFOLD INVARIANTS FROM HIGHER-RANK BUNDLES

P.B. Kronheimer

1. Introduction

1.1. Background. Donaldson's polynomial invariants for a smooth, oriented 4-manifold X are defined in [**4**] using the moduli spaces of anti-self-dual connections (or instantons) on a vector bundle $E \to X$. The structure group of the bundle E is generally taken to be either $SU(2)$ or $SO(3)$, both in [**4**] and in the subsequent mathematical literature, even though much of the underlying analysis is applicable to bundles with arbitrary compact structure group. In the physics literature, there is the work of Mariño and Moore, who argue in [**15**] that the $SU(2)$ Donaldson invariants have a well-defined generalization for $SU(N)$ bundles. They also propose a formula for the $SU(N)$ invariants when X has simple type, which suggests that these higher-rank invariants contain no new topological information. Nekrasov [**20**] formulated related conjectures concerning the moduli spaces of $SU(N)$ instantons on $\mathbb{R}^4$. These conjectures have been verified mathematically by Nekrasov and Okounkov [**19**] and by Nakajima and Yoshioka [**18**]. The purpose of this paper is to give a mathematical definition of the $SU(N)$ Donaldson invariants for general 4-manifolds, and to calculate some of them in an interesting family of examples arising from knot complements. The results, we obtain are consistent with the predictions of [**15**].

We recall some background material on the Donaldson invariants. The notion of "simple type" was introduced in [**14**]. For 4-manifolds X with this property, the Donaldson invariants arising from the Lie groups $SU(2)$ and $SO(3)$ were shown to be determined by a rather small amount of data. To state this result, we recall that, in the case that E is an $SU(2)$ bundle with $c_2(E)[X] = k$, the corresponding polynomial invariant q_k is defined whenever $b_2^+(X) - b_1(X)$ is odd and $b_2^+(X)$ is at least 2; the invariant is a homogeneous polynomial function

$$q_k : H_2(X;\mathbb{Z}) \to \mathbb{Z}$$

This work was supported in part by the National Science Foundation, award number DMS 0405271.

Received 02/21/2005.

of degree $d/2$, where $d = 8k - 3(b_2^+ - b_1 + 1)$ is the expected dimension of the instanton moduli space of instanton number k on X. If X has simple type and $b_1(X) = 0$, then the structure theorem from [**14**] establishes that there is a distinguished collection of 2-dimensional cohomology classes $K_i \in H^2(X;\mathbb{Z})$ $(i = 1, \dots, r)$ and rational coefficients a_i (independent of k) such that $q_k/(d/2)!$ is the term of degree $d/2$ in the Taylor expansion of the function

$$\mathcal{D}_X : H_2(X;\mathbb{R}) \to \mathbb{R}$$

given by

$$\mathcal{D}_X(h) = \exp\left(\frac{Q(h)}{2}\right) \sum_{i=1}^{r} a_i \exp K_i(h). \tag{1}$$

Here, Q is the intersection form on $H_2(X;\mathbb{Z})$. There is a similar formula for the $SO(3)$ invariants, involving the same "basic classes" K_i and coefficients a_i, which means that there is no more information in the $SO(3)$ invariants than is already contained in the $SU(2)$ invariants.

Witten's conjecture [**26**] states that for simply-connected 4-manifolds of simple type, the coefficients a_i and basic classes K_i are determined by the Seiberg–Witten invariants of X: we should have

$$a_i = 2^{2+(7\chi(X)+11\sigma(X))/4} SW_X(\mathfrak{s}_i)$$

and

$$K_i = c_1(S^+_{\mathfrak{s}_i}),$$

where the $\mathfrak{s}_i$ are the spin-c structures on X for which the Seiberg–Witten invariant $SW_X(\mathfrak{s}_i)$ is non-zero and $S^+_{\mathfrak{s}_i} \to X$ is the half-spin bundle corresponding to $\mathfrak{s}_i$. This conjecture was extended to the case of $SU(N)$ Donaldson invariants in [**15**]: the conjectured formula for the $SU(N)$ invariants is a homogeneous expression of degree $N-1$ in the quantities $SW_X(\mathfrak{s}_i)$.

1.2. Summary of results. In this paper, we shall give a quite careful construction of the $SU(N)$ and $PSU(N)$ Donaldson invariants for 4-manifolds X with $b_2^+ \geq 2$. As is the case already with $N = 2$, the $PSU(N)$ invariants are best defined by choosing a bundle $P \to X$ with structure group $U(N)$ (and possibly non-zero first Chern class) and then studying a moduli space of connections A in the associated adjoint bundle. (Note that not every principal $PSU(N)$ bundle can be lifted to a $U(N)$ bundle, so there is a slight loss of generality at this point.)

The new difficulties that arise in the higher-rank case are two-fold. First, we can expect singularities in the moduli space corresponding to *reducible* anti-self-dual connections. We shall bypass this difficulty initially by considering only the case that $c_1(P)$ is not divisible by any prime dividing N in the lattice $H^2(X;\mathbb{Z})/\text{torsion}$. In this case, if $b_2^+(X)$ is non-zero, then for a generic Riemannian metric on X, there will be no reducible solutions to the equations, as we explain in Section 2.3. If $c_1(P)$ is not coprime to N in this way, we can replace X with the blow-up $\tilde{X} = X\#\bar{\mathbb{CP}}^2$, and then define the invariants of X using moduli spaces on $\tilde{X}$. (This is the approach taken in [**14**] for the $SU(2)$ invariants.) The second difficulty is that for $N > 2$, one does not know that the irreducible parts of the $SU(N)$ or $PSU(N)$ moduli spaces will be smooth for generic choice of Riemannian metrics: we do not have the "generic metrics theorem" of Freed and Uhlenbeck [**10**], and the anti-self-duality equations must be artificially perturbed to achieve transversality.

Having defined the higher-rank invariants, we shall also calculate them in the case corresponding to a zero-dimensional moduli space on a particular family of 4-manifolds. Let K be a knot in S^3, let M be the knot complement (a compact 3-manifold with torus boundary), let X be a $K3$ surface with an elliptic fibration, and let $T \subset X$ be a torus arising as one of the fibers. The following construction is considered by Fintushel and Stern in [**8**]. Let X^o denote the 4-manifold with boundary T^3 obtained by removing an open tubular neighborhood of T, and let X_K be the closed 4-manifold formed as

$$X_K = (S^1 \times M) \cup_\phi X^o,$$

where $\phi : S^1 \times \partial M \to \partial X^o$ is an identification of the two 3-torus boundaries, chosen so that ϕ maps (point) $\times$ (longitude) to a curve on $\partial X^o \subset X$ which is the boundary of a 2-disk transverse to T. It is shown in [**8**] that X_K has the same homotopy type as X, but that the Seiberg–Witten invariants of X_K encode the Alexander polynomial of K. If Witten's conjecture holds, then the function $\mathcal{D}$ that determines the $SU(2)$ Donaldson invariants (see (1)) is given by

$$\mathcal{D}_{X_K} = \exp\left(\frac{Q(h)}{2}\right)\Delta_K(\exp(2F(h))),$$

where F is the cohomology class Poincaré dual to the torus fibers and

$$\Delta_K(t) = \sum_{k=-n}^{n} a_k t^k$$

is the symmetrized Alexander polynomial of the knot K.

We shall consider a particular $U(N)$ bundle $P \to X_K$ with $(c_1(P) \cdot F) = 1$ and $c_2(P)$ chosen so that the corresponding moduli space is zero-dimensional. The zero-dimensional moduli space gives us an integer-valued invariant of X_K (the signed count its points), which we are able to calculate when N is odd. The answer we obtain is the integer

$$\prod_{k=1}^{N-1} \Delta_K(e^{2\pi i k/N}),$$

as long as this quantity is non-zero. It seems likely that this answer is correct also when N is even, up to an overall sign (the answer is consistent with Witten's conjecture when $N = 2$ for example); but, we have not overcome an additional difficulty that arises in the even case: see Lemma 6.8 and the remarks at the end of Section 6.3.

2. Gauge theory for $SU(N)$

2.1. The configuration space. Let X be a smooth, compact, connected, oriented 4-manifold. We consider the Lie group $U(N)$ and its adjoint group, $PU(N)$ or $PSU(N)$. Let $P \to X$ be a smooth principal $U(N)$-bundle, and let $\mathrm{ad}(P)$ denote the associated principal bundle with structure group the adjoint group. Fix an integer $l \geq 3$, and let $\mathcal{A}$ denote the space of all connections in $\mathrm{ad}(P)$ of Sobolev class L^2_l. The assumption $l \geq 3$ ensures that $L^2_l \subset C^0$: we will use this assumption in Lemma 3.7, although at most other points $l \geq 2$ would suffice. If we fix a particular connection A_0, we can describe $\mathcal{A}$ as

$$\mathcal{A} = \{\, A_0 + a \mid a \in L^2_l(X; \mathfrak{su}_P \otimes \Lambda^1 X) \,\},$$

where $\mathfrak{su}_P$ is the vector bundle with fiber $\mathfrak{su}(N)$ associated to the adjoint representation. The *gauge group* $\mathcal{G}$ will be the group of all automorphisms of P of Sobolev class L^2_{l+1} which have determinant 1. That is, $\mathcal{G}$ consists of all L^2_{l+1} sections of the fiber bundle $SU(P) \to X$ with fiber $SU(N)$, associated to the adjoint action of $PSU(N)$ on $SU(N)$. The group $\mathcal{G}$ acts on $\mathcal{A}$, but the action is not effective. The subgroup that acts trivially is a cyclic group $Z \subset \mathcal{G}$ that we can identify with the center of $SU(N)$.

Definition 2.1. For a connection $A \in \mathcal{A}$, we write Γ_A for the stabilizer of A in $\mathcal{G}$. We say that A is *irreducible* if $\Gamma_A = Z$.

We write $\mathcal{A}^* \subset \mathcal{A}$ for the irreducible connections. We write $\mathcal{B}$ and $\mathcal{B}^*$ for the quotients of $\mathcal{A}$ and $\mathcal{A}^*$ by the action of $\mathcal{G}$. The space $\mathcal{B}^*$ has the structure of a Banach manifold in the usual way, with coordinate charts being provided by slices to the orbits of the action of $\mathcal{G}$. We write $[A] \in \mathcal{B}$ for the gauge-equivalence class represented by a connection A.

A $U(N)$-bundle P on X is determined up to isomorphism by its first and second Chern classes. We shall write $w \to X$ for the line bundle associated to P by the 1-dimensional representation $g \mapsto \det(g)$ of $U(N)$, so that $c_1(w) = c_1(P)$. We set

$$
\begin{aligned}
\kappa &= \left\langle c_2(P) - ((N-1)/(2N))c_1(P)^2, [X]\right\rangle \\
&= -(1/2N)\left\langle p_1(\mathfrak{su}_P), [X]\right\rangle .
\end{aligned}
\tag{2}
$$

The above expression is constructed so that κ depends only on $\mathrm{ad}(P)$, and so that, in the case that X is S^4, the possible values of κ run through all integers as P runs through all isomorphism classes of $U(N)$-bundles. We may label the configuration spaces $\mathcal{B}$ etc. using w and κ to specify the isomorphism class of the bundle. So, we have $\mathcal{B}^{w,*}_\kappa \subset \mathcal{B}^w_\kappa$ and so on. We refer to κ as the *instanton number*, even though it is not an integer.

There is a slightly different viewpoint that one can take to describe $\mathcal{A}$. Fix a connection θ in the line bundle w. Then, for each A in $\mathcal{A}$, there is a unique $U(N)$-connection $\tilde{A}$ in the $U(N)$-bundle P such that the connections which A induces in $\mathrm{ad}(P)$ and w are A and θ respectively. Thus, we can identify $\mathcal{A}$ with the space $\tilde{\mathcal{A}}$ of $U(N)$-connections inducing the connection θ in w.

This viewpoint is helpful in understanding the stabilizer Γ_A. Pick a basepoint $x_0 \in X$, and let A be a connection in $\mathrm{ad}(P)$. An element g in Γ_A is completely determined by its restriction to the fiber over x_0, so we can regard Γ_A as a subgroup of the copy of a copy of $SU(N)$, namely $SU(P)_{x_0}$. Let $\tilde{A}$ be the $U(N)$-connection in $P \to X$, corresponding to A and θ as above, and let $\tilde{H}$ be the group of automorphisms of P_{x_0} generated by the holonomy of $\tilde{A}$ around all loops based at x_0. Then, Γ_A is the centralizer of $\tilde{H}$ in $SU(P)_{x_0}$. The structure of Γ_A can then be described as follows. For any isomorphism of inner product spaces

$$\psi : \bigoplus_{i=1}^{r} (\mathbb{C}^{n_i} \otimes \mathbb{C}^{m_i}) \to \mathbb{C}^N,$$

we get a representation

$$\rho : \prod_{i=1}^{r} U(n_i) \to U(N).$$

The group $\Gamma_A \subset SU(P)_{x_0} \cong SU(N)$ is isomorphic to

$$\rho\left(\prod_{i=1}^{r} U(n_i)\right) \cap SU(N)$$

for some such ψ. In particular, we make the important observation that Γ_A has positive dimension unless $r = 1$ and $n_1 = 1$: that is to say, Γ_A has positive dimension unless A is irreducible.

2.2. The moduli space. Let $M^w_\kappa \subset \mathcal{B}^w_\kappa$ be the space of gauge-equivalence classes of anti-self-dual connections:

$$M^w_\kappa = \{\, [A] \in B^w_\kappa \mid F^+_A = 0 \,\}.$$

If A is anti-self-dual, then it is gauge-equivalent to a smooth connection. Supposing that A itself is smooth, we have a complex

$$\Omega^0(X; \mathfrak{su}_P) \xrightarrow{d_A} \Omega^1(X; \mathfrak{su}_P) \xrightarrow{d^+_A} \Omega^+(X; \mathfrak{su}_P).$$

We denote by H^0_A, H^1_A and H^2_A the cohomology groups of this elliptic complex. If A is not smooth, but only of class L^2_l, we can define the groups H^i_A in the same way after replacing the spaces of smooth forms above by Sobolev completions.

Lemma 2.2. *The vector space H^0_A is zero if A and only if A is irreducible.*

Proof. In general, H^0_A is the Lie algebra of Γ_A. We have already observed that Γ_A has positive dimension if A is reducible. q.e.d.

We say that A is *regular* if H^2_A vanishes. If A is regular and irreducible, then M^w_κ is a smooth manifold in a neighborhood of $[A]$, with tangent space isomorphic to H^1_A. The dimension of the moduli space near this point is then given by minus the index of the complex above, which is

$$d = 4N\kappa - (N^2 - 1)(b^+_2(X) - b_1(X) + 1). \tag{3}$$

We say that the moduli space is regular if A is regular for all $[A]$ in M^w_κ.

Uhlenbeck's theorem allows us to compactify the moduli space M^w_κ in the usual way. We define an ideal anti-self-dual connection of instanton number κ to be a pair $([A], \mathbf{x})$, where A is an anti-self-dual connection belonging to $M^2_{\kappa-m}$, and $\mathbf{x}$ is a point in the symmetric product X^m/S_m. The space of ideal anti-self-dual connections is compactified in the usual way, and we define the Uhlenbeck compactification of M^w_κ to be the closure of the M^w_κ in the space of ideal connections:

$$\bar{M}^w_\kappa \subset \bigcup_m M^w_{\kappa-m} \times (X^m/S_m).$$

2.3. Avoiding reducible solutions. We now turn to finding conditions to ensure that the moduli space contains no reducible solutions. We shall say that a class $c \in H^2(X;\mathbb{R})$ is *integral* if it is in the image of $H^2(X;\mathbb{Z})$ (or equivalently if it has integer pairing with every class in $H_2(X;\mathbb{Z})$). We shall use $\mathcal{H}^-$ to denote the space of real anti-self-dual harmonic 2-forms, which we may regard as a linear subspace of $H^2(X;\mathbb{R})$.

Proposition 2.3. *If M^w_κ contains a reducible solution, then there is an integer n with $0 < n < N$ and an integral class $c \in H^2(X;\mathbb{R})$ such that*

$$c - \frac{n}{N} c_1(w) \in \mathcal{H}^-.$$

Proof. Suppose that $[A]$ is a reducible solution in M^w_κ, and let $\tilde{A}$ be the corresponding $U(N)$-connection in the bundle $P \to X$, inducing the connection θ in w. The connection $\tilde{A}$ respects a reduction of P to $U(n_1) \times U(n_2)$, with $n_1 + n_2 = N$: we have

$$P \supset P_1 \times_X P_2,$$

where P_i is a bundle with structure group $U(n_i)$, and $n_i > 0$. Let $\tilde{A}_1$ and $\tilde{A}_2$ be the connections obtained from $\tilde{A}$ in the bundles P_1 and P_2. Let $\iota : \mathfrak{u}(1) \to \mathfrak{u}(N)$ be the inclusion that is the derivative of the inclusion of $U(1)$ as the center of $U(N)$. Because $\tilde{A}$ induces the connection θ on w and A is anti-self-dual, we have

$$F^+_{\tilde{A}} = (1/N)\iota(F^+_\theta).$$

Because $\tilde{A} = \tilde{A}_1 \oplus \tilde{A}_2$, we have

$$F^+_{\tilde{A}_1} = (1/N)\iota_1(F^+_\theta),$$

where $\iota_1 : \mathfrak{u}(1) \to \mathfrak{u}(n_1)$ is the inclusion again. Let w_1 be the determinant line bundle of P_1, and θ_1 the connection that it obtains from $\tilde{A}_1$. From the equality above, we obtain

$$F^+_{\theta_1} = (n_1/N)F^+_\theta.$$

Since F_θ and F_{θ_1} represent $-2\pi i c_1(w)$ and $-2\pi i c_1(w_1)$ respectively, we have proved the proposition: for the class c in the statement, we take $c_1(w_1)$, and for n, we take n_1. q.e.d.

Definition 2.4. We will say that $c_1(w)$ is *coprime to* N if there is no n with $0 < n < N$ such that $(n/N)c_1(w)$ is an integral class. This is equivalent to saying that there is a class in $H_2(X;\mathbb{Z})$ whose pairing with $c_1(w)$ is coprime to N.

As corollaries to the proposition, we have:

Corollary 2.5. *Let w be a fixed line bundle on X, and suppose $c_1(w)$ is coprime to N. Let Met denote the set of all C^r metrics on X, and let $\Xi \subset$ Met denote the set of metrics for which there exists a κ such that M^w_κ contains reducible solutions. Then, Ξ is contained in a countable union of smooth submanifolds of Met, each of which has codimension b^+_2 in Met.*

Proof. If the coprime condition holds, then for each integral class c and each n less than N, the class $c - (n/N)c_1(w)$ is non-zero. The set of metrics g such that

$$c - (n/N)c_1(w) \in \mathcal{H}^-(g)$$

is therefore a smooth submanifold $Met_{c,n} \subset Met$, of codimension b_2^+. (See [6], section 4.3, for example.) The set Ξ is contained in the union of these submanifolds $Met_{c,n}$ as c runs through the integral classes and n runs through the integers from 1 to $N-1$. q.e.d.

Corollary 2.6. *If $c_1(w)$ is coprime to N and b_2^+ is non-zero, then for a residual set of Riemannian metrics g on X, the corresponding moduli spaces M^w_κ contain no reducible solutions.*

3. Perturbing the equations

When $N = 2$, it is known that for a residual set of Riemannian metrics, the irreducible solutions in the corresponding moduli spaces M^w_κ are all regular. This is the generic metrics theorem of Freed and Uhlenbeck [10]. Unfortunately, no such result is known for larger N, and to achieve regularity, we must perturb the anti-self-duality equations. The holonomy perturbations which we use follow a plan which appears also in [3], [9] and [25] among other places. A discussion of bubbles in the presence of such perturbations occurs in [5].

3.1. A Banach space of perturbations. Let X be a Riemannian 4-manifold as before, let B be an embedded ball in X, and let $q : S^1 \times B \to X$ be a smooth map with the following two properties:

1) the map q is a submersion;
2) the map $q(1, -)$ is the identity: $q(1, x) = x$ for all $b \in B$.

Let

$$\omega \in \Omega^+(X; \mathbb{C})$$

be a smooth complex-valued self-dual 2-form whose support is contained in B. For each x in B, let

$$q_x : S^1 \to X$$

be the map given by $q_x(z) = q(z, x)$. Given a smooth connection A in $\mathcal{A}$, let $\tilde{A}$ be the corresponding connection in $P \to X$ with determinant θ, and let $\mathrm{Hol}_{q_x}(\tilde{A})$ denote the holonomy of $\tilde{A}$ around q_x. As x varies over B, the holonomy defines a section of the bundle of unitary groups $U(P)$ over B. We regard $U(P)$ as a subset of the vector bundle $\mathfrak{gl}_P$, with fiber $\mathfrak{gl}(N)$. We can then multiply the holonomy by the 2-form ω, and extend the resulting section by zero to define a section on all of X:

$$\omega \otimes \mathrm{Hol}_q(\tilde{A}) \in \Omega^+(X; \mathfrak{gl}_P).$$

We apply the orthogonal bundle projection $\pi : \mathfrak{gl}_P \to \mathfrak{su}_P$ to this section, and finally obtain a smooth section

$$V_{q,\omega}(A) \in \Omega^+(X; \mathfrak{su}_P).$$

Recall that $\mathcal{A}$ is defined as the L^2_l completion of the space of smooth connections.

Proposition 3.1. *For fixed q and ω, the map $V_{q,\omega}$ extends to a smooth map of Banach manifolds,*

$$V_{q,\omega} : \mathcal{A} \to L^2_l(X; \Lambda^+ \otimes \mathfrak{su}_P).$$

Furthermore, if we fix a reference connection A_0, then there are constants K_n, depending only on q and A_0, such that the n-th derivative

$$D^n V_{q,\omega}|_A : L^2_l(X; \Lambda^1 \otimes \mathfrak{su}_P)^n \to L^2_l(X; \Lambda^+ \otimes \mathfrak{su}_P)$$

satisfies

$$\big\| D^n V_{q,\omega}|_A(a_1, \ldots, a_n) \big\|_{L^2_{l,A_0}} \leq K_n \big\|\omega\big\|_{C^l} \prod_{i=1}^{n} \|a_i\|_{L^2_{l,A_0}}.$$

Proof. For smooth connections A and A_0, consider first the pull-backs $q^*(A)$ and $q^*(A_0)$ on $S^1 \times B$. Because q is a submersion, the pull-back map is continuous in the L^2_l topology: there is a constant c depending on q and A_0, such that

$$\|q^*(A) - q^*(A_0)\|_{L^2_{l,q^*(A_0)}} \leq c\|A - A_0\|_{L^2_{l,A_0}}.$$

In other words, the pull-back map q^* extends continuously to $\mathcal{A}$.

We wish to regard $q^*(A)$ as providing a family of connections on S^1, parametrized by B. To this end, for each $x \in B$, let $\mathcal{H}_x$ denote the Hilbert space of square-integrable $\mathfrak{su}_P$-valued 1-forms on $S^1 \times \{x\}$:

$$\mathcal{H}_x = L^2(S^1 \times \{x\}, \Lambda^1(S^1) \otimes q^*(\mathfrak{su}_P)).$$

These form a Hilbert bundle $\mathcal{H} \to B$, and $q^*(A_0)$ supplies $\mathcal{H}$ with a connection, $A_0^{\mathcal{H}}$. For A in $\mathcal{A}$, we have just observed that the 1-form $q^*(A) - q^*(A_0)$ lies in $L^2_l(S^1 \times B)$; and we therefore obtain from it, by restriction, a section of the bundle $\mathcal{H}$ of class L^2_l:

$$\begin{aligned} &a : x \mapsto a_x \\ &a \in L^2_l(B; \mathcal{H}). \end{aligned} \tag{4}$$

(This is just the statement that an element of $L^2_l(X_1 \times X_2)$ defines an L^2_l map from X_1 to $L^2(X_2)$.) Furthermore, we have an inequality

$$\|a\|_{L^2_{l,A_0^{\mathcal{H}}}} \le \|q^*(A) - q^*(A_0)\|_{L^2_{l,q^*(A_0)}}.$$

We can apply these constructions also to a unitary connection $\tilde{A}$ with determinant θ, to obtain $\tilde{a}$ in a similar way.

Lemma 3.2. *Let A_0 be a smooth connection in a bundle $P \to S^1$, and write a general connection as $A = A_0 + a$. Identify the structure group $U(N)$ with the group of automorphisms of the fiber of P over $1 \in S^1$. For each a, let $\mathrm{Hol}(a) \in U(N)$ denote the holonomy of $A_0 + a$ around the loop S^1 based at 1. Then,* Hol *extends to a smooth map*

$$\mathrm{Hol} : L^2(S^1; \Lambda^1 \otimes \mathfrak{su}_P) \to U(N).$$

Furthermore, there are constants c_n, independent of A_0 and A, such that the n-th derivative $D^n\mathrm{Hol}|_A$ satisfies

$$\big\|D^n\mathrm{Hol}|_A(a_1, \dots, a_n)\big\| \le c_n \prod_{i=1}^{n} \|a_i\|_{L^2}.$$

Proof. The connection A_0 does not appear in the inequality and serves only to define the domain of Hol. As a statement about A and the a_i, the inequality to be proved is gauge-invariant. It will suffice to show that, for each A, there is a connection $c_n(A)$ such that the inequality holds for all a_i: the fact that c_n can eventually be taken to be independent of A will follow automatically, because the space of connections modulo gauge is compact in the case of S^1. So, we are only left with the first assertion: that Hol extends to a smooth map on L^2. This is proved in [**25**, **7**]. (In fact, L^1 suffices.) q.e.d.

To return to the proof of Proposition 3.1, we apply the lemma to the family of circles parametrized by B. For each x in B, we regard the automorphism group of P_x as subset of $(\mathfrak{gl}_P)_x$, and have the holonomy map

$$\mathrm{Hol} : \mathcal{H}_x \to (\mathfrak{gl}_P)_x.$$

We regard this as giving us a smooth bundle map

$$\mathrm{Hol} : \mathcal{H} \to (\mathfrak{gl}_P)|_B.$$

This map is non-linear on the fibers; but the uniform bounds on derivatives of all orders mean that it still gives a map

$$\mathrm{Hol} : L^2_{l,A_0^{\mathcal{H}}}(B; \mathcal{H}) \to L^2_{l,A_0}(B; \mathfrak{gl}_P). \tag{5}$$

This Hol satisfies uniform bounds and its derivatives also. (For example, if E and F are Hilbert spaces and $h : E \to F$ is a uniformly Lipschitz map, then h defines a uniformly Lipschitz map $L^2(B; E) \to L^2(B; F)$.)

The section $V_{q,\omega}(A)$ is obtained by applying the map (5) to the section a from (4), then multiplying by the C^l 2-form ω and applying a linear projection in the bundle. q.e.d.

Now, fix once and for all a countable collection of balls B_α in X and maps $q_\alpha : S^1 \times B_\alpha \to X$, $\alpha \in \mathbb{N}$, each satisfying the two conditions laid out above. Let $K_{n,\alpha}$ be constants corresponding to q_α, as in the proposition above. Let C_α be a sequence of positive real numbers, defined by a diagonalization, so that

$$C_\alpha \geq \sup\{\, K_{n,\alpha} \mid 0 \leq n \leq \alpha \,\}.$$

For each α, let ω_α be a self-dual complex-valued form with support in B_α, and suppose that the sum

$$\sum_\alpha C_\alpha \|\omega_\alpha\|_{C^l}$$

is convergent. Then, for each n, the series

$$\sum_\alpha V_{q_\alpha,\omega_\alpha} \tag{6}$$

converges uniformly in the C^n topology on bounded subsets of $\mathcal{A}$. That is, for each R, if $B_R(\mathcal{A})$ denotes the Sobolev ball of radius R in $\mathcal{A}$ centered at A_0,

$$B_R(\mathcal{A}) = \{\, A_0 + a \mid \|a\|_{L^2_{l,A_0}} \leq R \,\},$$

then the series (6) converges in the uniform C^n topology of maps $B_R(\mathcal{A}) \to L^2_{l,A_0}(X : \Lambda^+ \otimes \mathfrak{su}_P)$. Thus, the sum of the series defines a C^∞ map of Banach manifolds,

$$V : \mathcal{A} \to L^2_l(X; \Lambda^+ \otimes \mathfrak{su}_P).$$

Definition 3.3. Fix maps q_α and constants C_α as above. We define W to be the Banach space consisting of all sequences $\boldsymbol{\omega} = (\omega_\alpha)_{\alpha\in\mathbb{N}}$ such that the sum

$$\sum_\alpha C_\alpha \|\omega_\alpha\|_{C^l}$$

converges. For each $\boldsymbol{\omega} \in W$, we write $V_{\boldsymbol{\omega}}$ for the sum of the series $\sum_\alpha V_{q_\alpha,\omega_\alpha}$, which is a smooth map

$$V_{\boldsymbol{\omega}} : \mathcal{A} \to L^2_l(X; \Lambda^+ \otimes \mathfrak{su}_P).$$

The dependence of $V_{\boldsymbol{\omega}}$ on $\boldsymbol{\omega}$ is linear, and the map $(\boldsymbol{\omega}, A) \mapsto V_{\boldsymbol{\omega}}(A)$ is a smooth map of Banach manifolds, $W \times \mathcal{A} \to L^2_l(X; \Lambda^+ \otimes \mathfrak{su}_P)$. Given $\boldsymbol{\omega}$ in W, we define the perturbed anti-self-duality equation to be the equation

$$F_A^+ + V_{\boldsymbol{\omega}}(A) = 0 \tag{7}$$

for $A \in \mathcal{A}$. Note that the left-hand side lies in $L^2_{l-1}(X; \Lambda^+ \otimes \mathfrak{su}_P)$. The equation is gauge-invariant, and we define the perturbed moduli space to be the quotient of the set of solutions by the gauge group:

$$M^w_{\kappa,\boldsymbol{\omega}}(X) = \{\, [A] \in \mathcal{B}^w_\kappa(X) \mid \text{equation (7) holds} \,\}.$$

3.2. Regularity and compactness for the perturbed equations. If A is a solution of the perturbed equation (7), we can consider the linearization of the equation at A, as a map

$$L^2_l(X; \Lambda^1 \otimes \mathfrak{su}_P) \to L^2_{l-1}(X; \Lambda^+ \otimes \mathfrak{su}_P). \tag{8}$$

The derivative of the perturbation $V_{\boldsymbol{\omega}}$ defines a bounded operator from L^2_l to L^2_l; so in the topologies of (8), the derivative of $V_{\boldsymbol{\omega}}$ is a compact operator. When we restrict the perturbed equations to the Coulomb slice through A, we therefore have a smooth Fredholm map, just as in the unperturbed case, and we have a Kuranishi theory for the perturbed moduli space. We write the Kuranishi complex at a solution A as

$$L^2_{l+1}(X; \Lambda^0 \otimes \mathfrak{su}_P) \xrightarrow{d_A} L^2_l(X; \Lambda^1 \otimes \mathfrak{su}_P) \xrightarrow{d^+_{A,\boldsymbol{\omega}}} L^2_{l-1}(X; \Lambda^+ \otimes \mathfrak{su}_P),$$

where $d^+_{A,\boldsymbol{\omega}}$ is the linearization of the perturbed equation,

$$d^+_{A,\boldsymbol{\omega}} = d^+_A + DV_{\boldsymbol{\omega}}|_A.$$

We denote the homology groups of this complex by H^0_A, $H^1_{A,\boldsymbol{\omega}}$ and $H^2_{A,\boldsymbol{\omega}}$. If A is irreducible and $H^2_{A,\boldsymbol{\omega}}$ is zero, then the moduli space $M^w_{\kappa,\boldsymbol{\omega}}$ is smooth near $[A]$, with tangent space $H^1_{A,\boldsymbol{\omega}}$.

Understanding the Uhlenbeck compactness theorem for the perturbed equations is more delicate than the Fredholm theory: if the curvature of a sequence of connections A_n concentrates at a point, then effect is seen in $V_{\boldsymbol{\omega}}(A_n)$ throughout the manifold, because the perturbations are non-local. To obtain a compactness result, we must work with weaker norms than L^2_l.

To this end, we start by observing that Proposition 3.1 continues to hold in the L^p_k norms, for any $p \geq 1$ and $k \geq 0$. That is to say, if we let $\mathcal{A}^p_k$ denote the L^p_k completion of the space of smooth connections, the $V_{q,\omega}$ defines a smooth map

$$V_{q,\omega} : \mathcal{A}^p_k \to L^p_k(X; \Lambda^+ \otimes \mathfrak{su}_P),$$

and the derivatives of this map are uniformly bounded in terms of constants which depend only on q and the C^k norm of ω. As long as the constants C_α are chosen appropriately, we can define the Banach space of perturbations W as before, and for $\boldsymbol{\omega} \in W$, we have again a smooth map $V_{\boldsymbol{\omega}}$

$$V_{\boldsymbol{\omega}} : \mathcal{A}^p_k \to L^p_k(X; \Lambda^+ \otimes \mathfrak{su}_P),$$

as long as $k \leq l$. By a diagonalization argument, we may suppose if we wish that the constants C_α are chosen so that $V_{\boldsymbol{\omega}}$ is smooth for *all* p and all $k \leq l$. We can also note that the constants C_α are inevitably bounded below, and this implies that $V_{\boldsymbol{\omega}}(A)$ satisfies a uniform L^∞ bound, independent of A.

Consider now a sequence $[A_n]$ in $M^w_{\kappa,\boldsymbol{\omega}}$. As just observed, the equations imply that the self-dual curvatures $F^+_{A_n}$ are uniformly bounded in L^∞, and F_{A_n} is then uniformly bounded in $L^2(X)$ because of the Chern–Weil formula. After passing to a subsequence, we may therefore assume as usual that there is a finite set of points $\mathbf{x}$ and a cover of $X \setminus \mathbf{x}$ by metric balls Ω_i, such that

$$\int_{\Omega_i} |F_{A_n}|^2 \leq \epsilon$$

for all n and i. Here, ϵ may be specified in advance to be smaller than the constant that appears in Uhlenbeck's gauge-fixing theorem. We may therefore put A_n in Coulomb gauge relative to a trivial connection on the ball Ω_i. For all $p \geq 2$, we have a bound on $F^+_{A_n}$ in L^p, and the connection form A_n in Coulomb gauge on Ω_i is therefore bounded in L^p_1 by a constant depending only on i.

By patching gauge transformations, we arrive at the following situation. There are gauge transformations g_n on X of class $L^2_{l+1,\mathrm{loc}}$, such that the sequence $g_n(A_n)$ is bounded in $L^p_1(K)$ for any compact subset K in $X \setminus \mathbf{x}$. We rename $g_n(A_n)$ as A_n.

Lemma 3.4. *Under the assumptions above, the sequence $V_{\boldsymbol{\omega}}(A_n)$ has a subsequence that is Cauchy in $L^p(X)$.*

Proof. For $r > 0$, let $K_r \subset X \setminus \mathbf{x}$ be the complement of the open metric balls around radius r around the points of $\mathbf{x}$. The uniform bound on the L^p_1 norm of A_n on K_r means that we can pass to a subsequence which converges in $L^p(K_r)$. By diagonalization, we can arrange the subsequence so that convergence occurs in $L^p(K_r)$ for all $r > 0$. Rename this subsequence as A_n. We claim that $V_{\boldsymbol{\omega}}(A_n)$ is Cauchy in $L^p(X)$. To see this, let $\epsilon > 0$ be given. For any $\beta \in \mathbb{N}$, let $V_{\boldsymbol{\omega}(\beta)}$ be the partial sum

$$V_{\boldsymbol{\omega}(\beta)} = \sum_{\alpha=1}^{\beta} V_{q_\alpha,\omega_\alpha}.$$

Choose β so that

$$\sum_{\alpha=\beta+1}^{\infty} \|\omega_\alpha\|_{C^0} \leq \epsilon.$$

It follows that

$$\|V_{\boldsymbol{\omega}}(A) - V_{\boldsymbol{\omega}(\beta)}(A)\|_{L^p} \leq \mathrm{Vol}(X)^{1/p}\epsilon$$

for all A. For each r, let $Z_{r,\alpha} \subset B_\alpha$ be the set

$$Z_{r,\alpha} = \{\, x \in B_\alpha \mid \; q_\alpha(x \times S^1) \text{ is not contained in } K_r \,\},$$

and let $Z_r = \cup_{\alpha \leq \beta} Z_{r,\alpha}$. The volume of Z_r goes to zero as r goes to zero; and because we have a uniform L^∞ bound on $V_{\boldsymbol{\omega}}(A)$, we can find r_0 sufficiently small that

$$\|V_{\boldsymbol{\omega}(\beta)}(A)|_{Z_{r_0}}\|_{L^p(Z_{r_0})} \leq \epsilon$$

for all A. The L^p convergence of A_n on K_{r_0} implies that $V_{\boldsymbol{\omega}(\beta)}(A_n)$ is Cauchy in $X \setminus Z_{r_0}$. So, we can find n_0 such that, for all n_1, n_2 greater than n_0, we have

$$\|V_{\boldsymbol{\omega}(\beta)}(A_{n_1}) - V_{\boldsymbol{\omega}(\beta)}(A_{n_2})\|_{L^p(X \setminus Z_{r_0})} \leq \epsilon.$$

Adding up the inequalities, we have

$$\|V_{\boldsymbol{\omega}}(A_{n_1}) - V_{\boldsymbol{\omega}}(A_{n_2})\|_{L^p(X)} \leq \big(3 + 2\mathrm{Vol}(X)^{1/p}\big)\epsilon$$

for all n_1, n_2 greater than n_0. q.e.d.

Bootstrapping and the removability of singularities now leads us to the following version of Uhlenbeck's theorem for our situation.

Proposition 3.5. *Let A_n be a sequence of connections in $P \to X$ representing points $[A_n]$ in the moduli space $M^w_{\boldsymbol{\omega},\kappa}$. Then, there is a point $\mathbf{x}$ in X^m/S_m and a connection A' in a bundle $P' \to X$ representing a point of a moduli space $M^w_{\boldsymbol{\omega},\kappa-m}$ with the following property. After passing to a subsequence, there are isomorphisms*

$$h_n : P|_{X\setminus\mathbf{x}} \to P'|_{X\setminus\mathbf{x}},$$

such that $(h_n)_*(A_n)$ converges to A' in $L^p_1(K)$, for all compact K contained in $X\setminus\mathbf{x}$. The proof of the lemma above shows that $(h_n)_*(V_{\boldsymbol{\omega}}(A_n))$, extended by zero to all of X, is Cauchy in $L^p(X)$ and has $V_{\boldsymbol{\omega}}(A')$ as its limit. It follows that A' solves the perturbed equations.

Note that, in an Uhlenbeck limit involving bubbles, we cannot expect anything better than L^p convergence of the curvatures on compact subsets disjoint $\mathbf{x}$, nor anything better than L^p_1 convergence of the connection forms. This is in contrast to the unperturbed case, where the convergence can be taken to be C^∞.

3.3. Transversality. Let $\mathcal{A}^*$ again denote the irreducible connections of class L^2_l. Let W be the Banach space parametrizing the perturbations $\boldsymbol{\omega}$, and consider the perturbed equations (7) for an irreducible connection A and perturbation $\boldsymbol{\omega}$ as the vanishing of a map

$$\mathcal{F} : W \times \mathcal{A}^* \to L^2_{l-1}(X; \Lambda^+ \otimes \mathfrak{su}_P), \tag{9}$$

given by

$$\mathcal{F}(\boldsymbol{\omega}, A) = F^+_A + V_{\boldsymbol{\omega}}(A).$$

We want this map to be transverse to zero. To achieve this, we need the family of balls B_α and maps q_α to be sufficiently large:

Condition 3.6. We require the balls B_α and the submersions $q_\alpha : S^1 \times B_\alpha \to X$ to satisfy the following additional condition: for every x in X, the maps

$$\{\, q_\alpha|_{S^1\times\{x\}} \mid \alpha \in \mathbb{N}, x \in \mathrm{int}(B_\alpha) \,\}$$

should be C^1-dense in the space of smooth loops $q : S^1 \to X$ based at x.

Lemma 3.7. *Suppose that the balls B_α and maps q_α are chosen to satisfy the condition above. Then, the map $\mathcal{F}$ is transverse to zero.*

Proof. Let $(\boldsymbol{\omega}, A)$ be in the zero set of $\mathcal{F}$. The derivative of $\mathcal{F}$ at $(\boldsymbol{\omega}, A)$ is the map

$$L : (\boldsymbol{\nu}, a) \mapsto d^+_{A,\boldsymbol{\omega}} a + V_{\boldsymbol{\nu}}(A).$$

We shall consider the restriction of L to the Coulomb slice: we write

$$L'_l : W \times K_l \to L^2_{l-1}(X; \Lambda^+ \otimes \mathfrak{su}_P),$$

where K_l is the Coulomb slice at A in the L^2_l topology. It will suffice to prove that L'_1 is surjective, because the regularity for the operator d^+_A allows us to use a bootstrapping argument to show that, if $L'_1(\boldsymbol{\nu}, a)$ belongs to L^2_{l-1}, then a belongs to L^2_l.

Our Fredholm theory already tells us that the cokernel of L'_1 has finite dimension. So, we need only show the image of L'_1 is dense in L^2. We shall show that, for a given irreducible A, the map $\boldsymbol{\nu} \mapsto V_{\boldsymbol{\nu}}(A)$ has dense image in L^2.

In order to use the hypothesis that A is irreducible, we recall that if $\rho : H \to U(N)$ is a unitary representation of a group H, then the hypothesis of irreducibility of H means that the complex span of the linear transformations $\rho(h)$ is all of $GL(N,\mathbb{C})$. So, if x is any point of X, we find loops γ_i based at x such that the holonomies $\mathrm{Hol}_{\gamma_i}(A)$, regarded as elements of $(\mathfrak{gl}_P)_x$, are a spanning set.

Because of Condition 3.6, we can therefore achieve the following. Given any x in X, we can find an $r > 0$ and N^2 elements α_i ($i = 1,\dots,N^2$), such that $B_{r_0}(x) \subset B_{\alpha_i}$ for all i and such that the holonomies $\mathrm{Hol}_{q_{\alpha_i},x}(\tilde{A})$ span $(\mathfrak{gl}_P)_x$. Decreasing r if necessary, we may assume further that the sections $e_i = \mathrm{Hol}_{q_{\alpha_i}}(\tilde{A})$ of the bundle $\mathfrak{gl}_P$ are a basis for the fiber at every point of the ball $B_r(x)$. Here, we have used the continuity of these sections, which follows from our assumption that L^2_l is contained in C^0.

If ω_i are any complex-valued self-dual forms of class C^l supported in $B_r(x)$, then the section $\sum \pi(\omega_i \otimes e_i)$ of $\lambda^+ \otimes \mathfrak{su}_P$ can be realized as $V_{\boldsymbol{\nu}}(A)$, by taking $\nu_{\alpha_i} = \omega_i$. (Here, π is again the projection $\mathfrak{gl}_P \to \mathfrak{su}_P$.) As the ω_i vary, we obtain a dense subset of the continuous sections supported in the ball $B_r(x)$.

Because x is arbitrary and X is compact, this is enough to show that the sections $V_{\boldsymbol{\nu}}(A)$ are dense in C^0 as $\boldsymbol{\nu}$ runs through W. They are therefore dense in L^2 also. q.e.d.

Corollary 3.8. *Suppose that Condition 3.6 holds. Then, there is a residual subset of W such that for all $\boldsymbol{\omega}$ in this set, and all w and κ, the irreducible part of the moduli space $M^w_{\kappa,\boldsymbol{\omega}}$ is regular, and therefore, a smooth manifold of dimension d given by (3).*

Proof. As usual, the zero set of $\mathcal{F}$ is a C^∞ Banach manifold by the implicit function theorem. The projection of $\mathcal{F}^{-1}(0)$ to W is a smooth Fredholm map of index d between separable Banach manifolds, and the fiber over $\boldsymbol{\omega}$ is the irreducible part of $M^w_{\kappa,\boldsymbol{\omega}}$; so, the result follows from the Sard–Smale theorem. q.e.d.

3.4. Reducible solutions of the perturbed equations.

When b_2^+ is non-zero, we would like to arrange that $M^w_{\kappa,\boldsymbol{\omega}}$ contains no reducible solutions. For the unperturbed case, the relevant statement is Corollary 2.6 above; but the discussion that led to this corollary breaks down when the equations are perturbed. Using the compactness theorem however, we can obtain a result when the perturbation $\boldsymbol{\omega}$ is small enough.

Proposition 3.9. *Let κ_0 be given, and let g be a metric on X with the property that the unperturbed moduli spaces M^w_κ contain no reducible solutions for any $\kappa \le \kappa_0$. Then, there exists $\epsilon > 0$ such that for all*

$\boldsymbol{\omega} \in W$ with $\|\boldsymbol{\omega}\|_W \le \epsilon$, the perturbed moduli spaces $M^w_{\kappa,\boldsymbol{\omega}}$ also contain no reducibles.

Proof. Suppose the contrary, and let $[A_n]$ be a reducible connection in $M^w_{\kappa,\boldsymbol{\omega}_n}$ for some sequence $\boldsymbol{\omega}_n$ in W with $\|\boldsymbol{\omega}_n\|_W \to 0$ as $n \to \infty$. The proof of Uhlenbeck compactness for the perturbed equations extends to this case, so there is a weak limit $([A], \mathbf{x})$ in $M^w_{\kappa'}$ for some $\kappa' \le \kappa$. The condition of being reducible is a closed condition, so A is reducible and we have a contradiction. q.e.d.

If we combine this proposition with the transversality results, we obtain the following.

Corollary 3.10. *Suppose that b_2^+ is non-zero and $c_1(P)$ is coprime to N. Then given any κ_0, we can find a Riemannian metric g and perturbation $\boldsymbol{\omega}$ such that the moduli spaces $M^w_{\kappa,\boldsymbol{\omega}}$ on (X, g) are regular for all $\kappa \le \kappa_0$ and contain no reducibles.*

4. Orienting the moduli spaces

4.1. Conventions defining orientations. On the irreducible part of the configuration space $\mathcal{B}^w_\kappa$, there is a line bundle Λ, obtained as the quotient by $\mathcal{G}$ of the line bundle on $\mathcal{A}$ given by the determinant of the family of operators

$$D_A = d_A^* \oplus d^+_{A,\boldsymbol{\omega}} \tag{10}$$

acting on $\mathfrak{su}_P$-valued forms. A trivialization of this line bundle provides an orientation of the moduli space $M^w_{\kappa,\boldsymbol{\omega}}$ at all points where the moduli space is regular. The space of perturbations $\boldsymbol{\omega}$ is contractible, so in studying trivializations of Λ, we may as well take $\boldsymbol{\omega} = 0$.

The first result is that Λ is indeed trivial: a proof for $U(N)$ bundles is given in [**6**, Section 5.4.3]. Our moduli spaces are, therefore, orientable manifolds. The trickier task is to isolate what choices need to be made to in order to specify a particular orientation, and to determine how the orientation depends on the choices made. For this task, we follow [**3**].

First, we orient the Lie algebra $\mathfrak{su}(N)$. Let $\mathfrak{s}$ be the subalgebra of $\mathfrak{sl}(N)$ consisting of the traceless upper triangular matrices whose diagonal entries are imaginary. The orthogonal projection $\mathfrak{s} \to \mathfrak{su}(N)$ is a linear isomorphism, so, we can orient $\mathfrak{su}(N)$ by specifying an orientation of $\mathfrak{s}$. We write $\mathfrak{s}$ as the direct sum of the diagonal algebra $\mathfrak{h}$ and a complex vector space, the strictly upper triangular matrices. We take the complex orientation on the latter; so an orientation of $\mathfrak{h}$ now determines an orientation of $\mathfrak{su}(N)$. To orient $\mathfrak{h}$, we write δ_n for the matrix with a single 1 at the n-th spot on the diagonal, and we take the ordered basis $i(\delta_n - \delta_{n+1})$, $n = 1, \dots, N-1$, to be an oriented basis of $\mathfrak{h}$.

Next, we choose a homology orientation for the 4-manifold X: that is, we choose an orientation o_X of the determinant line of the operator

$$D = d^* \oplus d^+ : \Omega^1(X) \to \Omega^0(X) \oplus \Omega^+(X).$$

In the case that $P \to X$ is trivial and A is the trivial connection, the operator D_A is obtained from D by tensor product with the vector space $\mathfrak{su}(N)$. So

$$\begin{aligned} \ker(D_A) &= \ker(D) \otimes \mathfrak{su}(N) \\ \operatorname{coker}(D_A) &= \operatorname{coker}(D) \otimes \mathfrak{su}(N). \end{aligned}$$

In general, if V and W are oriented vector spaces, we orient a tensor product $V \otimes W$ by choosing oriented ordered bases $(v_1, \dots, v_d)$ and $(w_1, \dots, w_e)$ and taking as an oriented basis for the tensor product the basis

$$(v_1 \otimes w_1, \dots, v_1 \otimes w_e, v_2 \otimes w_1, \dots).$$

If the dimension of W is even, this orientation of $V \otimes W$ is independent of the chosen orientation of V. Note also that if W is complex and has the complex orientation, then this convention equips $V \otimes W$ with its complex orientation also. Using this convention, our homology orientation o_X and our preferred orientation of $\mathfrak{su}(N)$ together determine an orientation of the determinant line of D_A at the trivial connection. This trivializes the line bundle Λ in the case of the trivial bundle P. Note that if N is odd, then the dimension of $\mathfrak{su}(N)$ is even; and in this case, the orientation of Λ is independent of the choice of homology orientation o_X.

Moving away from the trivial bundle, we now consider a hermitian rank-N vector bundle $E \to X$ decomposed as

$$E = L_1 \oplus \cdots \oplus L_N, \tag{11}$$

and we take P to be the corresponding principal $U(N)$ bundle. In line with our use of the upper triangular matrices $\mathfrak{s}$, we can decompose the associated bundle $\mathfrak{su}_P$ as the sum of a trivial bundle with fiber $\mathfrak{h}$ and a complex vector bundle

$$\bigoplus_{j>i} \operatorname{Hom}(L_j, L_i).$$

Using our preferred orientation of $\mathfrak{h}$ and a chosen homology orientation o_X, we obtain in this way an orientation

$$O(L_1, \dots, L_N, o_X)$$

for the determinant line $\Lambda \to \mathcal{B}_P$. In the case that all the L_n are trivial, this reproduces the same orientation as we considered in the previous paragraph. Again, this orientation is independent of o_X when N is odd. As a special case, given a line bundle $w \to X$, we define $O(w, o_X)$ to be the orientation just defined, in the case that $L_1 = w$ and L_n is trivial for $n > 1$. The corresponding bundle P has $\kappa = \kappa_0$ given by $\kappa_0 = -((N-1)/(2N))c_1^2(w)$. In this way, for each w, we orient the determinant line over $\mathcal{B}^w_{\kappa_0}$.

For different values of κ, the determinant lines over $\mathcal{B}^w_\kappa$ and $\mathcal{B}^w_{\kappa'}$ can be compared by the "addition of instantons", as in [**3**]. On S^4, there is a standard $SU(N)$ instanton with $\kappa = 1$, obtained from the standard $SU(2)$ instanton by the standard inclusion of $SU(2)$ in $SU(N)$; using a concentrated instanton and excision, an orientation of the determinant line over $\mathcal{B}^w_\kappa$ determines an orientation over $\mathcal{B}^w_{\kappa+1}$. A precise convention can be fixed by requiring that this determination respects the complex orientations in the case of a Kähler manifold (see below). In this way, the orientation $O(w, o_X)$ over $\mathcal{B}^w_{\kappa_0}$ determines an orientation over $\mathcal{B}^w_\kappa$ for all κ.

To summarize, given a homology orientation o_X for the 4-manifold, we have fixed standard orientations for the smooth parts of all the moduli spaces M^w_κ. In the case that N is odd, the choice of o_X is immaterial.

4.2. Comparing orientations. Suppose E is a hermitian rank-N bundle and $E' = E \otimes v$ is another, where v is a complex line bundle. The top exterior powers are related by $w' = w + Nv$. (We use additive notation here for a tensor product of line bundles.) The configuration spaces $\mathcal{B}_P$ and $\mathcal{B}_{P'}$ for the corresponding principal bundles P and P' are canonically identified, as are the moduli spaces, because the adjoint bundles $\mathrm{ad}(P)$ and $\mathrm{ad}(P')$ are the same. While this provides an identification of two moduli spaces M^w_κ and M^{w+Nv}_κ, it is not necessarily the case that this identification respects the orientations of their smooth parts given by $O(w, o_X)$ and $O(w + Nv, o_X)$ as in the previous subsection.

We follow the proof of [**3**, Proposition 3.25] closely to compare the two. The main step is to consider the case that X is Kähler, and compare the orientation $O(w, o_X)$ to the complex orientation, where o_X is the preferred homology orientation of the Kähler manifold defined in [**3**]. For a bundle E decomposed as (11), the operator

$$D_A : \Omega^1\Big(\mathfrak{h} \oplus \bigoplus_{j>i} \mathrm{Hom}(L_j, L_i)\Big) \longrightarrow \big(\Omega^0 \oplus \Omega^+\big)\Big(\mathfrak{h} \oplus \bigoplus_{j>i} \mathrm{Hom}(L_j, L_i)\Big)$$

decomposes as a sum of operators $(-\bar\partial^* \oplus \bar\partial)$ acting on bundles

$$\Omega^{0,1} \otimes_{\mathbb{R}} \mathfrak{h} \longrightarrow \Big((\Omega^0)^{\mathbb{C}} \oplus \Omega^{0,2} \Big) \otimes_{\mathbb{R}} \mathfrak{h},$$

$$\Omega^{0,1} \otimes_{\mathbb{C}} \mathrm{Hom}(L_j, L_i) \longrightarrow \Big((\Omega^0)^{\mathbb{C}} \oplus \Omega^{0,2} \Big) \otimes_{\mathbb{C}} \mathrm{Hom}(L_j, L_i), \qquad (j > i),$$

$$\Omega^{0,1} \otimes_{\mathbb{C}} \mathrm{Hom}(L_i, L_j) \longrightarrow \Big((\Omega^0)^{\mathbb{C}} \oplus \Omega^{0,2} \Big) \otimes_{\mathbb{C}} \mathrm{Hom}(L_i, L_j), \qquad (j > i).$$

We treat these three summands in turn. In each case, the index of the operator has a complex orientation (coming from the complex structure of $\Omega^{0,i}$) which we wish to compare to our preferred orientation. In the first summand, one comparison that arises is between the complex orientation of $V \otimes_{\mathbb{R}} \mathfrak{h}$ and the tensor product orientation of $V \otimes_{\mathbb{C}} \mathfrak{h}$ defined in the previous subsection, when V is a complex vector space. These two differ by the sign

$$(-1)^{(\dim_{\mathbb{C}} V)(N-1)(N-2)/2},$$

irrespective of our convention for orienting $\mathfrak{h}$. This consideration contributes a sign

$$(-1)^{(\mathrm{ind}(-\bar\partial^* \oplus \bar\partial))(N-1)(N-2)/2},$$

to our comparison. As in [3], there is a further consideration in the first term, in that the preferred homology orientation of X uses the opposite orientation to that given by the complex structure of $(-\bar\partial^* \oplus \bar\partial)$, and this contributes another $(\mathrm{ind}(-\bar\partial^* \oplus \bar\partial))(N-1)$. Altogether, the first summand contributes a sign

$$(-1)^{(\mathrm{ind}(-\bar\partial^* \oplus \bar\partial))N(N-1)/2}. \tag{12}$$

In the second summand, our convention for defining $O(L_1, ..., L_N, o_X)$ uses the complex structure of $\mathrm{Hom}(L_j, L_i)$, and this is the same complex structure on the index as arises from the complex structure of $\Omega^{0,i}$. So there is no correction here. In the third summand, our complex structure is opposite to the one inherited from $\Omega^{0,i}$. We therefore, get a correction

$$\prod_{j>i} (-1)^{\mathrm{ind}(-\bar\partial^* \oplus \bar\partial)_{\mathrm{Hom}(L_i, L_j)}}. \tag{13}$$

We now specialize to the case that $L_1 = w$ and the other L_n are trivial, which is the case that we used to define $O(w, o_X)$ above. The bundles $\mathrm{Hom}(L_i, L_j)$ with $j > i$ contribute $N-1$ copies of $\bar w$ and $(N-2)(N-1)/2$ trivial bundles. The formula (13) above (the contribution of the third summand) thus simplifies to

$$(-1)^{\mathrm{ind}(-\bar\partial^* \oplus \bar\partial)_{\bar w}(N-1) + \mathrm{ind}(-\bar\partial^* \oplus \bar\partial)(N-1)(N-2)/2}.$$

Now, we add the contribution (12) from the first summand, and we obtain

$$(-1)^{(N-1)(\mathrm{ind}(-\bar\partial^*\oplus\bar\partial)_{\bar w}-\mathrm{ind}(-\bar\partial^*\oplus\bar\partial)},$$

which is equal to

$$(-1)^{(N-1)(w\cdot w+K\cdot w)/2},$$

where K is the canonical class of X. Thus, we have proved:

Proposition 4.1. *If X is Kähler and o_X is the preferred homology orientation of the Kähler manifold, then the distinguished orientation $O(w, o_X)$ of the moduli space M^w_κ coincides with its complex orientation if N is odd. If N is even, then the two orientations compare by the sign $(-1)^{(w\cdot w+K\cdot w)/2}$.*

Finally, we can use Proposition 4.1 to compare $O(w, o_X)$ with $O(w+Nv, o_X)$ on a Kähler manifold, using the fact that the complex orientations of M^w_κ and M^{w+Nv}_κ coincide. We consider the case that N is even. If we set $w' = w + Nv$, then

$$\big((w\cdot w+K\cdot w)-(w'\cdot w'+K\cdot w')\big)\ /\ 2=(N/2)v\cdot v, \quad (\mathrm{mod}\ 2).$$

Thus, we have:

Proposition 4.2. *The two orientations $O(w, o_X)$ and $O(w+Nv, o_X)$ of the moduli space $M^w_\kappa = M^{w+Nv}_\kappa$ are the same when N is odd or when N is zero mod 4. When N is 2 mod 4, they compare with the sign $(-1)^{v\cdot v}$.*

Proof. We have just dealt with the case that X is Kähler. The general case is reduced to this one by excision, as in [**3**]. q.e.d.

4.3. Dual bundles. The map $\delta : U(N) \to U(N)$ given by $g \mapsto \bar g$ allows us to associate to our principal bundle P a new principal bundle $\bar P$. The corresponding vector bundles E and $\bar E$ obtained from the defining representation of $U(N)$ are dual. To each connection A in the adjoint bundle $\mathfrak{su}_P$, we obtain a connection $\delta(A)$ in $\mathfrak{su}_{\bar P}$; and this gives us a map

$$\delta : M^w_\kappa \to M^{-w}_\kappa.$$

In the case $N = 2$, this is the same as one obtains by tensoring E with $\bar w$; but for other N, it is different.

Proposition 4.3. *Let the smooth parts of M^w_κ and M^{-w}_κ be given the orientations $O(w, o_X)$ and $O(-w, o_X)$, where o_X is a chosen homology-orientation for X. Then, the dualizing map δ respects these orientations in the case that N is odd. In the case that N is even, the map δ preserves or reverses orientation according to the sign $(-1)^{w\cdot w}$.*

Proof. In the case that X is Kähler, the map δ respects the complex orientations of the moduli spaces, for it is a complex-analytic map between them. Using Proposition 4.1 and excision, we can again compare the orientations $O(w, o_X)$ and $O(-w, o_{X)})$. q.e.d.

5. Integer invariants and product formulae

In this section, we use the gauge theory and perturbations that we have developed for $PSU(N)$ to define simple integer invariants of 4-manifolds. As a preliminary step for our later calculations, we also prove a product formula for such invariants, for a 4-manifold X decomposed along a 3-torus.

5.1. Simple integer invariants. Let X be a smooth, compact oriented 4-manifold, with $b_2^+(X) \geq 2$. Fix N and let $P \to X$ be a principal $U(N)$ bundle. Let w be the $U(1)$ bundle $\det(P)$, and let κ be the instanton number. If N is even, choose also a homology orientation o_X for X. We impose the following simplifying condition.

Hypothesis 5.1. We shall suppose that $c_1(w)$ is coprime to N in the sense of Definition 2.4. We shall also suppose that the formal dimension d of the moduli space M^w_κ, as given by the formula (3), is zero.

By Corollary 2.6, we can choose a Riemannian metric g on X so that the moduli spaces $M^w_{\kappa'}$ contain no reducibles solutions. More specifically, we want to choose g so that there is no integer class c and no $n < N$ such that $c - (n/N)c_1(w)$ is represented by an anti-self-dual form. Having chosen such a g, we can find $\epsilon(g)$ as in Proposition 3.9 such that for all $\boldsymbol{\omega} \in W$ with $\|\boldsymbol{\omega}\|_W \leq \epsilon(g)$, the perturbed moduli space $M^w_{\kappa'}$ contains no reducibles for any $\kappa' \leq \kappa$. Finally, we use Corollary 2.6 to find a $\boldsymbol{\omega}$ with $\|\boldsymbol{\omega}\|_W < \epsilon(g)$ such that the moduli spaces $M^w_{\kappa'}$ are smooth for all $\kappa' \leq \kappa$. We call any $(g, \boldsymbol{\omega})$ arrived at in this way a *good pair.*

Let $(g, \boldsymbol{\omega})$ be a good pair. The moduli space $M^w_{\kappa,\boldsymbol{\omega}}$ is a smooth 0-manifold, and is compact because the moduli spaces $M^w_{\kappa-m,\boldsymbol{\omega}}$ are empty, being of negative dimension. So $M^w_{\kappa,\boldsymbol{\omega}}$ is a finite set of points. Using $O(w, o_X)$, we can orient this moduli space: the points of the moduli space then acquire signs; and in the usual way, we can count the points, with signs, to obtain an integer.

The integer which we obtain in this way is independent of the choice of good pair $(g, \boldsymbol{\omega})$, by the usual cobordism argument. In a little more detail, suppose $(g_0, \boldsymbol{\omega}_0)$ and $(g_1, \boldsymbol{\omega}_1)$ are good pairs. Because b_2^+ is greater than 1, Corollary 2.5 allows us to choose a smooth path of metrics g_t joining g_0 to g_1, such that the moduli spaces $M^w_\kappa(X, g_t)$ contain no reducibles. We can then find a continuous function *epsilon*(t) such that

$M^w_{\kappa,\boldsymbol{\omega}_t}(X, g_t)$ contains no reducibles whenever $\|\boldsymbol{\omega}_t\|_{W_t} \leq \epsilon(t)$. Here, W_t is the Banach space of perturbations on (X, g_t): we can take it that the maps q_α and balls B_α used in its definition are independent of t. Let $\mathcal{W} \to [0, 1]$ be the fiber bundle with fiber the perturbation space W_t. Over $[0, 1]$, the maps $\mathcal{F}$ of (9) define a total map

$$\mathcal{F} : \mathcal{W} \times \mathcal{A}^* \to \mathcal{L},$$

where $\mathcal{L}$ is the fiber bundle over $[0, 1]$ with fiber $L^2_{l-1}(X; \Lambda^+_t \otimes \mathfrak{su}_P)$. We can choose a section $\boldsymbol{\omega}_t$ of the $\mathcal{W}$ over $[0, 1]$, transverse to $\mathcal{F}$, and always smaller than $\epsilon(t)$. The family of moduli spaces $M^w_{\kappa,\boldsymbol{\omega}_t}(X, g_t)$ will then sweep out an oriented 1-dimensional cobordism between $M^w_{\kappa,\boldsymbol{\omega}_0}(X, g_0)$ and $M^w_{\kappa,\boldsymbol{\omega}_1}(X, g_1)$.

We summarize our discussion.

Definition 5.2. Let X be a closed, oriented smooth 4-manifold with $b^+_2(X)$ at least 2, equipped with a homology orientation o_X. Let w be a line bundle with $c_1(w)$ coprime to N, and suppose there exists a $U(N)$ bundle P, with $\det(P) = w$, such that the corresponding moduli space M^w_κ has formal dimension 0. Then, we define an integer $q^w(X)$ as the signed count of the points in the moduli space $M^w_{\kappa,\boldsymbol{\omega}}(X, g)$, where $(g, \boldsymbol{\omega})$ is any choice of good pair. This integer depends only on X, w and o_X up to diffeomorphism; and if N is odd, it is independent of o_X.

We extend the definition of q^w by declaring it to be zero if there is no P for which the corresponding moduli space is zero-dimensional.

For given w, there may be no corresponding P for which the moduli space M^w_κ has formal dimension zero. Referring to the formula (3) for d and the definition of κ, we see that a necessary and sufficient condition is that the integer

$$(N^2 - 1)(b^+_2(X) - b_1(X) + 1) + 2(N - 1)w \cdot w$$

should be divisible by $4N$. In the case that N is even, this implies in particular that $b^+_2 - b_1$ should be odd, and that $w \cdot w$ has the same parity as $(b^+_2 - b_1 + 1)/2$. From Proposition 4.3, we therefore derive:

Proposition 5.3. *The integer invariant $q^w(X)$ satisfies $q^w(X) = q^{-w}(X)$ if N is odd. If N is even, then $q^w(X)$ and $q^{-w}(X)$ differ by the sign*

$$(-1)^{(b^+_2(X) - b_1(X) + 1)/2}.$$

5.2. Cylindrical ends. Let X be an oriented 4-manifold with boundary a connected 3-manifold Y, and let $P \to X$ be a principal $U(N)$ bundle. Let $\mathcal{R}^w(Y)$ denote the *representation variety* of Y: the quotient of the space of flat connections in $\mathfrak{su}_P \to Y$ by the action of the group of determinant-1 gauge transformations of $P|_Y$. We make the following simplifying assumption, in order to extend the definition of the integer invariant q^w in a straightforward way to such manifolds with boundary.

Hypothesis 5.4. We shall suppose that $\mathcal{R}^w(Y)$ consists of a single point $\alpha = [A_Y]$. We shall further suppose that α is irreducible, and non-degenerate, in the sense that the cohomology group $H^1(Y;\alpha)$ with coefficients in the flat bundle $(\mathfrak{su}_P, \alpha)$ is trivial.

Once a particular representative connection A_Y on $\mathfrak{su}_P|_Y$ is chosen, we can extend A_Y to all of $\mathfrak{su}_P$ to obtain a connection A_X on X. But there is an essential topological choice in this process. The determinant-1 gauge group $\mathcal{G}(Y)$ is not connected: its group of components is infinite cyclic; and the gauge transformations on Y that extend to X are precisely the identity component of $\mathcal{G}(Y)$.

Let $\tilde{\mathcal{B}}_Y$ denote the quotient of the space of connections $\mathcal{A}_Y$ by the identity component $\mathcal{G}_1(Y) \subset \mathcal{G}(Y)$. Let $\tilde{\mathcal{R}}^w(Y)$ denote the subset of $\tilde{\mathcal{B}}_Y$ consisting of flat connections: this is an infinite set acted on transitively by the infinite cyclic group $\mathcal{G}(Y)/\mathcal{G}_1(Y)$. Let $\tilde{\alpha}$ be a choice of a point in $\tilde{\mathcal{R}}^w_Y$. Let A_Y be a connection representing $\tilde{\alpha}$, and let A_X be any extension of A_Y to all of X.

Up to this point, we have made no essential use of a metric on X. We now choose a Riemannian metric on X that is cylindrical in a collar of the boundary. We also suppose that A_X is flat in the collar. Form a cylindrical-end manifold X^+ by attaching a cylinder $[0,\infty) \times Y$ to X, and extend A_X as a flat connection on the cylindrical part. We now introduce a space of connections on X^+,

$$\mathcal{A}(X^+;\tilde{\alpha}) = \{\, A \mid A - A_X \in L^2_{l,A_X}(X^+;\Lambda^1 \otimes \mathfrak{su}_P\,\},$$

a gauge group

$$\mathcal{G}(X^+;\tilde{\alpha}) = \{\, g \mid g - 1 \in L^2_{l+1,A_X}(X^+; SU(P))\,\},$$

and the quotient space $\mathcal{B}(X^+;\tilde{\alpha})$. The gauge group acts freely, and the quotient space is a Banach manifold. Inside this Banach manifold is the moduli space of anti-self-dual connections:

$$M^w(X^+;\tilde{\alpha}) = \{\, [A] \in \mathcal{B}(X^+;\tilde{\alpha}) \mid F^+_A = 0\,\}.$$

We have the usual results concerning moduli spaces on cylindrical-end manifolds. Any finite-energy anti-self-dual connection A in the bundle $P \to X^+$ is gauge-equivalent to a connection in $M^w(X^+;\tilde{\alpha})$, for some unique $\tilde{\alpha} \in \tilde{\mathcal{R}}^w(Y)$. Given $[A]$ in this moduli space, there is an elliptic complex

$$L^2_{l+1,A_X}(X^+;\Lambda^0\otimes\mathfrak{su}_P)\xrightarrow{d_A}L^2_{l,A_X}(X^+;\Lambda^1\otimes\mathfrak{su}_P)\\ \xrightarrow{d_A^+}L^2_{l-1,A_X}(X^+;\Lambda^+\otimes\mathfrak{su}_P).$$

We write H^i_A for its cohomology groups. We always have $H^0_A=0$, and again describe $[A]$ as regular if H^2_A is zero, in which case the moduli space is smooth near $[A]$. The dimension of the moduli space at regular points is the index of the operator

(14)

$$d^*_{A_X}\oplus d^+_{A_X}:L^2_{l,A_X}(X^+;\Lambda^1\otimes\mathfrak{su}_P)\longrightarrow L^2_{l-1,A_X}(X^+;(\Lambda^0\oplus\Lambda^+)\otimes\mathfrak{su}_P).$$

We write $d(X^+;\tilde\alpha)$ for this index. We write the transitive action of $\mathbb{Z}$ on $\tilde{\mathcal{R}}^w(Y)$ using the notation $\tilde\alpha\mapsto\tilde\alpha+k$. The sign convention is fixed so that

$$d(X^+;\tilde\alpha+k)=d(X^+;\tilde\alpha)+4Nk.$$

We also have

$$\kappa(X^+;\tilde\alpha+k)=\kappa(X^+;\tilde\alpha)+k,$$

where κ is defined by the same Chern–Weil integral that would define the characteristic class (2) in the closed case:

$$\begin{aligned}\kappa(X^+;\tilde\alpha)&=-\frac{1}{2N}\int_{X^+}p_1(\mathfrak{su}_P;A)\\ &=\frac{1}{16N\pi^2}\int_{X^+}\operatorname{tr}\big(\operatorname{ad}(F_A)\wedge\operatorname{ad}(F_A)\big).\end{aligned}\tag{15}$$

(The notation $\operatorname{ad}(F_A)$ denotes a 2-form with values in $\operatorname{End}(\mathfrak{su}_P)$, and tr is the trace on $\operatorname{End}(\mathfrak{su}_P)$.)

We can construct a Banach space W parametrizing perturbations of the anti-self-duality equations on X^+. In order to achieve transversality, we again choose a family of balls and maps q_α satisfying the density condition (Condition 3.6). The constants C_α in the definition of W can be chosen so that for all $\boldsymbol{\omega}\in W$, the perturbing term $V_{\boldsymbol{\omega}}(\alpha)$ has rapid decay on the end of X^+: for some constants K_j, and t a function equal to the first coordinate on the cylindrical end, we can require

$$|\nabla^j_{A_X}V_{\boldsymbol{\omega}}(A)|\le K_je^{-t}\|\boldsymbol{\omega}\|_W,$$

for all A in $\mathcal{A}(X^+;\tilde\alpha)$, all $\boldsymbol{\omega}$ in W, and all $j\le l$. With such a condition, the term $V_{\boldsymbol{\omega}}(A)$ will contribute a compact perturbation to the linearized equations, and we have moduli spaces $M^w_{\boldsymbol{\omega}}(X^+;\tilde\alpha)$. For a residual set of perturbations $\boldsymbol{\omega}$ in W, the perturbed moduli space $M^w_{\boldsymbol{\omega}}(X^+;\tilde\alpha)$ will be regular, and therefore a smooth manifold of dimension $d(X^+;\tilde\alpha)$. A pair

$(g, \boldsymbol{\omega})$ consisting of a cylindrical-end metric and perturbation $\boldsymbol{\omega}$ will be called a *good pair* if all the moduli spaces $M^w_{\boldsymbol{\omega}}(X^+; \tilde{\alpha})$ are regular.

The moduli space is orientable, and can be oriented by choosing a trivialization of the determinant line bundle on the connected Banach manifold $\mathcal{B}^w(X^+; \tilde{\alpha})$. We make no effort here to define a canonical orientation for the moduli space.

The proof of the compactness theorem also adapts to the cylindrical end case. If $[A_n]$ is a sequence in $M^w_{\boldsymbol{\omega}}(X^+; \tilde{\alpha})$, then after passing to a subsequence there is an Uhlenbeck limit $([A]; \mathbf{x})$, where $\mathbf{x} \in \mathrm{Sym}^m(X^+)$ and $[A] \in M^w_{\boldsymbol{\omega}}(X^+; \tilde{\alpha} - m')$. The main difference from the case of a closed manifold is that we only have $m' \geq m$, rather than equality, because some energy may be lost on the cylindrical end.

We can now define an integer invariant (with an ambiguous sign)

$$q^w(X; \alpha) \in \mathbb{Z}/\{\pm 1\} \tag{16}$$

whenever Hypothesis 5.4 holds. We define it by picking a good pair $(g, \boldsymbol{\omega})$, choosing an orientation for the determinant line, and then counting with signs the points in a zero-dimensional moduli space $M^w_{\boldsymbol{\omega}}(X^+; \tilde{\alpha})$ if one exists. If there is no $\tilde{\alpha}$ for which the moduli space is zero-dimensional, we define $q^w(X; \alpha)$ to be zero.

Note that there is no hypothesis on $b_2^+(X)$ in this construction, because there are no reducible solutions in the moduli spaces.

5.3. Perturbations with compact support. The following lemma is useful for the gluing formula in the subsection below.

Lemma 5.5. *Let X^+ be a manifold with cylindrical end, as above, and suppose that w satisfies Hypothesis 5.4. Suppose $M^w_\kappa(X^+; \tilde{\alpha})$ has formal dimension 0. Then, we can find a cylindrical-end metric g and a perturbation $\boldsymbol{\omega}$ making the moduli spaces $M^w_{\kappa,\boldsymbol{\omega}}(X^+; \tilde{\alpha} - m)$ regular for all $m \geq 0$, with the additional condition that the sum (6) defining the perturbation has only finitely many non-zero terms: that is, only finitely many of the ω_α are non-zero.*

Proof. We already know that we can find $\boldsymbol{\omega}$ in W such that the perturbed moduli spaces $M^w_{\kappa,\boldsymbol{\omega}}(X^+; \tilde{\alpha} - m)$ are regular. (For m strictly positive, regular means that this moduli space is empty.) We will show that the regularity of these moduli spaces for all $m \geq 0$ is an open condition. This will suffice, because the perturbations given by finite sums are dense.

Suppose then $\boldsymbol{\omega}_n$ is a sequence converging to $\boldsymbol{\omega}$ in W. Because of the Chern–Weil formula, we know that there exists m_0 such that $M^w_{\kappa,\boldsymbol{\omega}_n}(X^+; \tilde{\alpha} - m)$ is empty (and therefore, regular), for all $m \geq m_0$ and all n.

Suppose that m_0 is greater than 1, and consider the moduli spaces

$$M^w_{\kappa,\boldsymbol{\omega}_n}(X^+;\tilde\alpha - m_0 + 1).$$

If these are non-empty for infinitely many n, then we have a contradiction, arising from Uhlenbeck's compactness theorem and the emptiness of $M^w_{\kappa,\boldsymbol{\omega}}(X^+;\tilde\alpha - m)$ for m positive. In this way, we prove inductively that $M^w_{\kappa,\boldsymbol{\omega}_n}(X^+;\tilde\alpha - m)$ is empty for all $n \geq n_0$ and all $m \geq 1$.

Now, consider the moduli spaces $M^w_{\kappa,\boldsymbol{\omega}_n}(X^+;\tilde\alpha)$, which have formal dimension 0. Suppose these are non-regular for infinitely many n. After passing to a subsequence, we may suppose there is an irregular solution $[A_n]$ in $M^w_{\kappa,\boldsymbol{\omega}_n}(X^+;\tilde\alpha)$ for all n, and that these converge in Uhlenbeck's sense. The fact that the lower moduli spaces are empty means that this is in fact, strong convergence of the connections A_n after gauge transformation. Irregularity is a closed condition under strong limits, so the limit is an irregular point $[A]$ in $M^w_{\kappa,\boldsymbol{\omega}}(X^+;\tilde\alpha)$, which contradicts our hypothesis. q.e.d.

Remark. One should expect to prove that irregularity is a closed condition under Uhlenbeck limits, and so extend the lemma to higher-dimensional moduli spaces. Because of the non-local nature of our perturbations, this would require some more work.

5.4. Gluing. Suppose X is a closed 4-manifold with $b_2^+ \geq 2$, so that the integer invariants $q^w(X)$ are defined. Suppose X contains a connected 3-manifold Y that separates X into two manifolds with common boundary, X_1 and X_2, and suppose that the representation variety $\mathcal{R}^w(Y)$ satisfies Hypothesis 5.4. Let w_i be the restriction of w to X_i, so that we have integer invariants (with ambiguous sign) $q^{w_i}(X_i;\alpha)$ for $i = 1, 2$.

The sign-ambiguity can be partly resolved as follows. Let Λ be the orientation line bundle on $\mathcal{B}^w(X)$, and let Λ_i be the orientation bundles on $\mathcal{B}^{w_i}(X^+;\tilde\alpha)$. Then, there is a preferred isomorphism

$$\Lambda = \Lambda_1 \otimes \Lambda_2.$$

Suppose we choose a homology orientation o_X for X, and so trivialize Λ using $O(w, o_X)$. Then, an orientation for Λ_1 determines an orientation for Λ_2, using the above product rule. Thus, a choice of sign for $q^{w_1}(X_1;\alpha)$ determines a choice of sign for $q^{w_2}(X_2;\alpha)$. In the next proposition, we assume that the signs are resolved in this way.

Proposition 5.6. *In the above situation, we have a product law,*

$$q^w(X) = q^{w_1}(X_1;\alpha)q^{w_2}(X_2;\alpha).$$

Proof. First, choose perturbation $\boldsymbol{\omega}_1$ and $\boldsymbol{\omega}_2$ for X_1^+ and X_2^+ satisfying the condition in the conclusion of Lemma 5.5. This condition means that there is a compact subset $K_i \subset X_i^+$ such that, for all A on X_i^+, the perturbing term $V_{\boldsymbol{\omega}_i}(A)$ is supported in K_i and depends only on the restriction of A to K_i. Equip X with a metric g_R containing a long cylindrical neck $[-R, R] \times Y$, in the usual way. For R sufficiently large, X contains isometric copies of K_1 and K_2, so we can regard $\boldsymbol{\omega}_1$ and $\boldsymbol{\omega}_2$ as defining a perturbation $V_{\boldsymbol{\omega}}$ of the equations on (X, g_R). Since this perturbation is supported away from the neck region, it does not interfere with the standard approaches to gluing anti-self-dual connections. The conclusion is that, for R sufficiently large, the moduli space $M^w_{\boldsymbol{\omega}}(X, g_R)$ is regular and is the product of the two moduli spaces $M^{w_i}_{\boldsymbol{\omega}_i}(X_i^+; \tilde{\alpha})$, provided there exists a lift $\tilde{\alpha}$ such that these are zero-dimensional. q.e.d.

5.5. Other cases. We now consider a slightly more general setting, in which the representation variety $\mathcal{R}^w(Y)$ is no longer required to be a single point. We still ask that $c_1(w)$ is coprime to N on Y, but we suppose now that $\mathcal{R}^w(Y)$ consists of n irreducible elements α_i $(i = 1, \dots, n)$, all of which are non-degenerate. Rather than develop a complete Floer homology theory, we continue to make some simplifying assumptions, which we now lay out.

If $\gamma : [0,1] \to \mathcal{B}^w(Y)$ is a path joining α_i to α_j, we can associate to γ two quantities, both of which are additive along composite paths. The first is the spectral flow of the family of operators

$$\begin{bmatrix} 0 & -d^*_{A(t)} \\ -d_{A(t)} & *d_{A(t)} \end{bmatrix}$$

acting on $(\Omega^0 \oplus \Omega^1)(Y; \mathfrak{su}_P)$, where $A(t)$ is a lift of the path γ to $\mathcal{A}(Y)$. We call this integer $d(\gamma)$. The second quantity is the Chern–Weil integral of the same kind as (15). We regard $A(t)$ as defining a connection A on $[0,1] \times Y$ in temporal gauge, and define

$$\kappa(\gamma) = \frac{1}{16N\pi^2} \int_{[0,1]\times Y} \operatorname{tr}\bigl(\operatorname{ad}(F_A) \wedge \operatorname{ad}(F_A)\bigr).$$

Modulo $\mathbb{Z}$, we can interpret $\kappa(\gamma)$ as the drop in the suitably-normalized Chern–Simons invariant from α_i to α_j. The normalization is such that, for a suitably-oriented closed loop in $\mathcal{B}^w(Y)$ representing the generator of first homology, the value of $\kappa(\gamma)$ is 1. For the same loop, $d(\gamma)$ is $4N$.

If $A(t)$ is a path of connections lifting γ, and A is the corresponding 4-dimensional connection as above, we may extend A to a connection A_Z on the infinite cylinder $Z = \mathbb{R} \times Y$, just as we did for the cylindrical-end case. The connection A_Z is flat on both ends of the cylinder. We can then define a space of connections

$$\mathcal{A} = \{\, A_Z + a \mid a \in L^2_{l,a_Z}(Z; \mathfrak{su}_P) \,\}$$

and a quotient space $\mathcal{B} = \mathcal{B}^w(Z;\gamma)$. Inside the configuration space is the moduli space $M^w(Z;\gamma)$, and $d(\gamma)$ is its formal dimension.

Hypothesis 5.7. Let γ denote a path from α_i to α_j as above. We impose two conditions:

1) $d(\gamma)$ is not 1 for any such γ;
2) if $d(\gamma) \leq 0$, then $\kappa(\gamma) \leq 0$ also.

Assume these two conditions hold, and consider a moduli space $M^w_{\boldsymbol{\omega}}(X^+;\tilde{\alpha}_k)$ on a cylindrical-end manifold X^+. Suppose $\boldsymbol{\omega}$ is chosen so that all moduli spaces are regular, and that this particular moduli space is zero-dimensional. The two conditions imply that if $M^w_{\boldsymbol{\omega}}(X^+;\tilde{\alpha}_j)$ is another moduli space, with $\kappa(X^+;\tilde{\alpha}_j) \leq \kappa(X^+;\tilde{\alpha}_k)$, then the formal dimension of this moduli space is -2 or less. This moduli space is therefore empty, and remains empty for a generic path of metrics an perturbations. It follows that there is a well-defined integer invariant

$$q^w(X;\alpha_k) \in \mathbb{Z}/\pm 1$$

counting points in this zero-dimensional moduli space. Once again, we extend the definition by declaring it to be zero if there is no lift $\tilde{\alpha}_k$ for which the moduli space is zero-dimensional.

The product law from Proposition 5.6 then extends to this more general situation. In the situation considered in the proposition, we now have:

$$q^w(X) = \sum_{k=1}^{n} q^{w_1}(X_1;\alpha_k)\, q^{w_2}(X_2;\alpha_k). \tag{17}$$

6. Calculations

6.1. The $K3$ surface. Let X be a *K3* surface (with the complex orientation) and let $P \to X$ be a $U(N)$ bundle with

$$\langle c_2(P), [X] \rangle = N(N^2-1)$$
$$\langle c_1(P)^2, [X] \rangle = 2(N+1)^2(N-1).$$

More specifically, we suppose these conditions are achieved by having

$$c_1(P) = \binom{N+1}{1} h,$$
$$c_2(P) = \binom{N+1}{2} h^2$$

where h is a primitive class of square $2(N-1)$. These conditions ensure in particular, that $c_1(P)$ is coprime to N. The bundle P has $\kappa = N-(1/N)$, and the formal dimension of the corresponding moduli space $M^w_{\kappa,\boldsymbol{\omega}}$ is zero. We choose the homology orientation o_X arising from the Kähler structure, and Proposition 4.1 tells us that the moduli space has the same orientation as the one it obtains as a complex-analytic space. In this situation, there is an integer invariant $q^w(X)$ as in Definition 5.2, and our aim is to calculate it.

Proposition 6.1. *The above invariant $q^w(X)$ for the $K3$ surface X is* 1.

Proof. We use a familiar circle of ideas to evaluate the invariant: see [**4**] for the calculation in the case $N = 2$, on which our argument is based. We assume $N \geq 3$ for the rest of this proof.

We take X to be an algebraic surface embedded in $\mathbb{CP}^N$ by the complete linear system of a line bundle with first Chern class h. This can be done in such a way that h generates the Picard group of X: all holomorphic line bundles on X are powers of the hyperplane bundle H.

For the inherited metric g, the corresponding moduli space M^w_κ can be identified with the moduli space of poly-stable holomorphic bundles with the topology of P. The condition on $c_1(P)$ ensures that there are no reducible bundles, so poly-stable means stable in this case. As usual on a $K3$, irreducible stable bundles are regular, so the moduli space M^w_κ is smooth and 0-dimensional; and the moduli spaces $M^w_{\kappa-m}$ are empty for m positive. We can therefore use the stable bundles to calculate the integer invariant. By the remark about orientations above, the integer q^w is simply the number of stable bundles on X with the correct topology.

Mukai's argument [**16**], shows that the moduli space is connected, so to prove the proposition, we just have to exhibit one stable bundle. The Chern classes of P have been specified so that the restriction to X of the holomorphic tangent bundle $T\mathbb{CP}^N$ as the correct topology. We will show that the holomorphic vector bundle

$$\mathcal{E} = T\mathbb{CP}^N|_X$$

is stable, under the hypothesis that H generates the Picard group. To do this, we recall that $\mathcal{E}$ has a description as a quotient

$$0 \to \mathcal{O} \xrightarrow{s} V \otimes H \xrightarrow{q} \mathcal{E} \to 0,$$

where V is a vector space of dimension $N+1$. The bundle $\mathcal{E}$ has rank N and determinant $(N+1)H$; so if it is unstable then there is subsheaf $\mathcal{F}$ with rank n and determinant at least $(n+1)H$, where n lies in the range $1 \le n \le N$. We can suppose that $\mathcal{E}/\mathcal{F}$ is torsion-free. Let $\tilde{\mathcal{F}}$ be the inverse image of $\mathcal{F}$ under the quotient map q. Then, $\tilde{\mathcal{F}}$ has rank $n+1$ and determinant at least $(n+1)H$; and the quotient of $V \otimes H$ by $\tilde{\mathcal{F}}$ is torsion-free. This forces

$$\tilde{\mathcal{F}} = U \otimes H,$$

where U is a non-trivial proper subspace of V. However, the image of the section s in the above sequence is not contained in $U \otimes H$, for any proper subspace U. (Geometrically, this is the statement that X is not contained in any hyperplane in $\mathbb{CP}^N$.) Since $\tilde{\mathcal{F}}$ contains $s(\mathcal{O})$ by construction, we have a contradiction. q.e.d.

6.2. Flat connections on the 3-torus. We consider the 3-torus T^3 as the quotient $\mathbb{R}^3/\mathbb{Z}^3$, and take x, y and z to be the simple closed curves which are the images of the three coordinate axes in $\mathbb{R}^3$. We regard these as oriented curves, defining elements of the fundamental group $\pi_1(T^3, z_0)$, where z_0 is the base point. We take $w \to T^3$ to be a line bundle with $c_1(w)$ Poincaré dual to z, and we take P to be a principal $U(N)$ bundle having w as determinant.

Lemma 6.2. *For the* 3*-manifold $T^3 = \partial Z_K$, the representation variety $\mathcal{R}^w(T^3)$ consists of N points $\alpha_0, \dots, \alpha_{N-1}$. Each of these points is irreducible and non-degenerate.*

Proof. Let $z' \subset T^3$ be a parallel copy of z, disjoint from the curves x and y and meeting the 2-torus $T^2 \times \{0\}$ spanned by x and y transversely in a single point. Let η be a 2-form supported in a tubular neighborhood of z', representing the dual class to z. The support of η should be disjoint from x, y and z. Equip the line bundle w with a connection θ whose curvature is $-(2\pi i)\eta$ and which is trivial on the complement of the tubular neighborhood. Use θ to trivialize w away from z'. This trivialization of the determinant reduces the structure group of P to $SU(N)$ on the complement of the tubular neighborhood of z'.

Using θ, we identify $\mathcal{R}^w(T^3)$ with the moduli space of flat $SU(N)$ connections $\tilde{A}$ on $T^3 \setminus z'$ such that the corresponding homomorphism

$$\rho : \pi_1\big(T^3 \setminus z', z_0\big) \to SU(N)$$

satisfies

$$[\rho(x), \rho(y)] = e^{2\pi i/N}\mathbf{1}_N.$$

The last condition implies that, up to similarity $\rho(x)$ and $\rho(y)$ can be taken to be the pair

$$\rho(x) = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & \zeta & 0 & \cdots & 0 \\ 0 & 0 & \zeta^2 & \cdots & 0 \\ & & & \ddots & 0 \\ 0 & 0 & 0 & 0 & \zeta^{N-1} \end{bmatrix} \quad \rho(y) = \begin{bmatrix} 0 & 0 & 0 & \cdots & \pm 1 \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ & & & \ddots & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix},$$

where $\zeta = e^{2\pi i/N}$ and the sign in the top row of $\rho(y)$ is negative if N is even (so that $\rho(y)$ has determinant 1). As an element of $\pi_1(T^3 \setminus z', z_0)$, the class of z commutes with x and y, and this tells us that

$$\rho(z) = \zeta^k \mathbf{1}_N$$

for some k with $0 \le k \le N-1$. The elements x, y and z generate the fundamental group, so there are exactly N elements in the representation variety, as claimed. Let ρ_k be the representation with $\rho(z) = \zeta^k \mathbf{1}_N$, and let α_k be the corresponding point of $\mathcal{R}^w(T^3)$.

We have already remarked that the only elements commuting with $\rho_k(x)$ and $\rho_k(y)$ are the central elements; so each ρ_k is irreducible. Each ρ_k determines a representation $\bar{\rho}_k$ of $\pi_1(T^3, z_0)$ in the adjoint group $PSU(N)$; and all these are equal. To see that they are non-degenerate, we look at the cohomology $H^1(T^3; (\mathfrak{su}_P, \bar{\rho}_k)$ with coefficients in the adjoint bundle. Any element of this group is represented by a covariant-constant $\mathfrak{su}_P$-valued form, because the torus is flat; and there are no non-zero covariant-constant forms because the bundle is irreducible.

q.e.d.

Lemma 6.3. *For each k, there is a path γ joining α_k to α_{k+1} in $\mathcal{B}^w(T^3)$ such that the spectral flow $d(\gamma)$ is 4. For the same path, the energy $\kappa(\gamma)$ is $1/N$.*

Proof. The Chern–Simons invariant of α_k in $\mathbb{R}/\mathbb{Z}$ is calculated in [**1**], where it is shown to be $-k/N$ mod $\mathbb{Z}$. It follows that there is a path γ_0 joining α_0 to α_1 with $\kappa(\gamma_0) = 1/N$. The connections represented by α_k and α_{k+1} are gauge-equivalent under the larger gauge group consisting of all $PSU(N)$ gauge transformations of $\mathrm{ad}(P)$. It follows that we can lift γ_0 to a path of connections $A : [0,1] \to \mathcal{A}^w(T^3)$ with $A(1) = g(A(0))$ for some automorphism g of $\mathrm{ad}(P)$. Applying g to this whole path, we obtain a path of connections $A : [1,2] \to \mathcal{A}^w(T^3)$ joining $A(2)$ to a representative $A(3)$ for some other point of $\mathcal{R}^w(T^3)$, which must be α_3 on the grounds of its Chern–Simons invariant.

In this way, we construct paths γ_k from α_k to α_{k+1}, each of which has the same energy and the same spectral flow. The composite path is a loop δ based at α_0 with $\kappa(\delta) = 1$. It follows that $d(\delta)$ is $4N$. So each γ_k has $d(\gamma_k) = 4$. q.e.d.

6.3. Knot complements. Let K be a knot in S^3, and let M be the knot complement, which we take to be a 3-manifold with 2-torus boundary carrying a distinguished pair of oriented closed curves m and l, the meridian and longitude, on its boundary.

We write Z_K for the 4-manifold $S^1 \times M$. The boundary of Z_K is a 3-torus. We identify m and l with curves $1 \times m$ and $1 \times l$ on ∂Z_K, and write s for the curve $S^1 \times z_0$ on ∂Z_K, where z_0 is the point of intersection of m and l, which we take as base-point in Z_K. The three classes

$$[s], [m], [l] \in H_1(\partial Z_K)$$

generate the first homology of the boundary. We orient K, so as to orient the longitude l. We suppose m is oriented so that the oriented tangent vectors of s, m and l (in that order) are an oriented basis for the tangent space to ∂Z_K at z_0.

The first homology of Z_K itself is generated by $[s]$ and $[m]$, as the class l is the boundary of an oriented Seifert surface for K. We take Σ to be such a Seifert surface, regarded as a 2-dimensional submanifold with boundary in $(Z_K, \partial Z_K)$, and we take $w \to Z_K$ to be the line bundle with $c_1(w)$ dual to $[\Sigma, \partial\Sigma] \in H_2(Z_K, \partial Z_K)$.

We identify ∂Z_K with the standard 3-torus T^3 in such a way that s, m and l are identified with x, y and z (the curves considered in the previous subsection). From the discussion above, we have connections α_k in $\mathcal{R}^w(\partial Z_K)$ for $k = 0, \dots, N-1$, with spectral flow 4 along paths from each one to the next. Because of the lemmas above, the conditions of Hypothesis 5.7 hold for the 3-torus ∂Z_K, so, we can consider the integer-valued invariants $q^w(Z_K; \alpha_k)$, as defined in section 5.5, for the manifold-with-boundary Z_K. Here is the result:

Proposition 6.4. *When N is odd, the integer invariant $q^w(Z_K; \alpha_0)$ for $Z_K = S^1 \times M$ can be expressed in terms of the Alexander polynomial $\Delta(t)$ of the knot K by the formula*

$$q^w(Z_K; \alpha_0) = \pm \prod_{k=1}^{N-1} \Delta(e^{2\pi i k/N}),$$

provided this quantity is non-zero. For k non-zero, the invariant $q^w(Z_K; \alpha_k)$ is zero, because there is no moduli space $M^w(Z_K^+; \tilde{\alpha}_k)$ of formal dimension 0.

Proof. We prove the proposition in a series of lemmas. The main idea is that we can calculate this invariant by counting certain flat connections on Z_K, so reducing the problem to one about representations of the fundamental group, rather than having to understand non-flat solutions to the anti-self-duality equations.

Lemma 6.5. *There is a lift $\tilde{\alpha}_0 \in \tilde{\mathcal{R}}^w(T^3)$ of $\alpha_0 \in \mathcal{R}^w(T^3)$ such that the moduli space $M^w(Z_K^+; \tilde{\alpha}_0)$ on the cylindrical-end manifold Z_K^+ consists only of flat connections and has formal dimension* 0.

Proof. The moduli space $M^w(Z_K^+; \tilde{\alpha}_0)$ will consist entirely of flat connections if $\kappa(Z_K^+; \tilde{\alpha}_0)$ is zero, which in turn holds if there is at least one flat connection in the moduli space. The $PSU(N)$ representation $\bar{\rho}_0$ of $\pi_1(\partial Z_K; z_0)$ which corresponds to α_0 sends l to 1, so it extends as an abelian representation of the fundamental group of Z_K. This representation defines a $PSU(N)$ connection on the non-trivial bundle $\mathrm{ad}(P)$, so it defines a lift $\tilde{\alpha}_0$ of α_0 of the sort required.

By excision, the formal dimension of the moduli space $M^w(Z_K^+; \tilde{\alpha}_0)$ with $\kappa = 0$ is independent of the knot K, so we may consider just the unknot, in which case the formal dimension is easily seen to be zero (for example by using the product formula for the integer invariants, as in the section below). q.e.d.

Lemma 6.6. *The number of flat connections in $M^w(Z_K^+; \tilde{\alpha}_0)$ can be expressed in terms of the Alexander polynomial $\Delta(t)$ by the formula*

$$\left| \prod_{k=1}^{N-1} \Delta(e^{2\pi i k/N}) \right|, \tag{18}$$

provided that this quantity is non-zero.

Proof. Because of the previous lemma, we can regard $M^w(Z_K^+; \tilde{\alpha}_0)$ as a moduli space of flat connections on the compact space Z_K. More exactly, it is the space of flat $PSU(N)$ connections in $\mathrm{ad}(P) \to Z_K$, divided by the determinant-1 gauge group. (Any such flat connection must define the representation α_0 at the boundary, rather than α_k for non-zero k, because the latter have non-zero Chern–Simons invariant.)

We proceed as we did for the 3-torus T^3 previously. Let l' be a parallel copy of l, just as z' was a parallel copy of z in our earlier language. We take l' to be the boundary of Σ', which is a parallel copy of Σ in the 4-manifold $Z_K = S^1 \times M$. The surface Σ' should lie in $p \times M$, where p is distinct from the basepoint $1 \in S^1$. We equip $w \to Z_K$ with a connection θ that is trivial outside the tubular neighborhood of Σ' and whose curvature is $(-2\pi i)\eta$, where η represents the Poincaré dual class to Σ' and is supported in the tubular neighborhood. Using θ, we can lift a flat connection A in $\mathrm{ad}(P)$ lifts to a $U(N)$ connection $\tilde{A}$ in P that is flat outside the neighborhood of Σ'. The trivialization of w provided by θ reduces the structure group of P to $SU(N)$ away from the tubular neighborhood, and $\tilde{A}$ becomes a flat $SU(N)$ connection there.

Suppose, then that A and $\tilde{A}$ are such connections. Let

$$\rho : \pi_1(Z_K \setminus \Sigma'; z_0) \to SU(N)$$

be the representation defined by the holonomy of $\tilde{A}$, and let

$$\bar{\rho} : \pi_1(Z_K; z_0) \to PSU(N) \tag{19}$$

be defined by the holonomy of A. We have already seen, from our discussion of the boundary T^3, that after change of basis, we must have (20)

$$\rho(s) = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & \zeta & 0 & \cdots & 0 \\ 0 & 0 & \zeta^2 & \cdots & 0 \\ & & & \ddots & 0 \\ 0 & 0 & 0 & 0 & \zeta^{N-1} \end{bmatrix} \quad \rho(m) = \begin{bmatrix} 0 & 0 & 0 & \cdots & \pm 1 \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ & & & \ddots & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix},$$

and

$$\rho(l) = \mathbf{1}_N.$$

In $\pi_1(Z_K; z_0)$ the class of the circle s is central. So the image of $\bar{\rho}$ lies in the centralizer of $\bar{\rho(s)}$. Let $\bar{J}$ denote this centralizer, and let J be its inverse image in $SU(N)$. We can describe the group J as a semi-direct product

$$1 \to H \to J \to V \to 1,$$

where H is the standard maximal torus of $SU(N)$ and V is the cyclic group of order N generated by the matrix $\rho(m)$ above.

We identify the 3-manifold M with the submanifold $1 \times M$ in the product $Z_K = S^1 \times M$. Then, by restriction, ρ determines a homomorphism

$$\sigma : \pi_1(M, z_0) \to J,$$

with

$$\sigma(m) = \begin{bmatrix} 0 & 0 & 0 & \cdots & \pm 1 \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ & & & \ddots & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \tag{21}$$

and $\sigma(l) = \mathbf{1}_N$. Conversely, σ determines ρ and hence $\bar{\rho}$. We are therefore, left with the task of enumerating such homomorphisms σ.

Let π stand as an abbreviation for the group $\pi_1(M, z_0)$, let $\pi^{(1)}$ be the commutator subgroup, and $\pi^{(2)} = [\pi^{(1)}, \pi^{(1)}]$. The group J is a two-step solvable group, with the torus H as commutator subgroup. So, we must have

$$\sigma(\pi^{(1)}) \subset H,$$

and $\sigma(\pi^{(2)}) = \{1\}$.

Because M is a knot complement, the quotient $\pi/\pi^{(1)}$ is infinite cyclic, generated by the coset represented by the meridian m. The abelian group $\pi^{(1)}/\pi^{(2)}$ becomes a $\mathbb{Z}[t, t^{-1}]$-module when we let the action of t on $\pi^{(1)}/\pi^{(2)}$ be defined by the action of conjugation, $x \mapsto mxm^{-1}$, on $\pi^{(1)}$. In a similar way, the abelian group H becomes a $\mathbb{Z}[t, t^{-1}]$-module if we define the action of t to be given by conjugation by the element $v \in SU(N)$ that appears in (21). In this way, σ determines (and is determined by) a homomorphism of $\mathbb{Z}[t, t^{-1}]$-modules,

$$\sigma' : \frac{\pi^{(1)}}{\pi^{(2)}} \to H.$$

We are left with the task of enumerating such $\mathbb{Z}[t, t^{-1}]$-module homomorphisms. The structure of $\pi^{(1)}/\pi^{(2)}$ as a $\mathbb{Z}[t, t^{-1}]$ module is described by the Alexander polynomial Δ: there is an isomorphism

$$\frac{\pi^{(1)}}{\pi^{(2)}} \cong \mathbb{Z}[t, t^{-1}] \;/\; \Delta(t). \tag{22}$$

(See [**22**] for example.) So, the homomorphism σ' is entirely determined by the element

$$h = \sigma'(1) \in H.$$

We switch to additive notation for the torus H. We write τ_v for the automorphism of H given by conjugation by v. In the ring of endomorphisms of H, we can consider the element $\Delta(\tau_v)$. The description (22) of $\pi^{(1)}/\pi^{(2)}$ means that the element $h = \sigma'(1)$ must lie in the kernel of $\Delta(\tau_v)$.

Thus, we have shown that the moduli space $M^w(Z_K^+; \tilde{\alpha}_0)$ is in one-to-one correspondence with the elements h in the abelian group

$$\ker(\Delta(\tau_v)) \subset H.$$

Let $\tilde{H}$ be the universal cover of H (or the Lie algebra of this torus), and let $\tilde{\tau}_v$ be the lift of τ_v to this vector space. Then, the order of the kernel of $\Delta(\tau_v)$ is equal to the absolute value of the determinant

$$\det\bigl(\Delta(\tilde{\tau}_v)\bigr)$$

of the linear transformation $\Delta(\tilde{\tau}_v)$, provided that this determinant is non-zero. If the determinant is zero, then the kernel is infinite, for it is a union of tori, whose dimension is equal to the dimension of the null-space of $\Delta(\tilde{\tau}_v)$. As a linear transformation of the $(N-1)$-dimensional real vector space $\tilde{H}$, the operator $\tilde{\tau}_v$ is diagonalizable after complexification, and its eigenvalues on $\tilde{H} \otimes \mathbb{C}$ are the non-trivial N-th roots of unity, ζ^k for $k = 1, \dots, N-1$. The eigenvalues of $\Delta(\tilde{\tau}_v)$ are therefore, the complex numbers $\Delta(\zeta^k)$, so

$$\det\bigl(\Delta(\tilde{\tau}_v)\bigr) = \prod_{k=1}^{N-1} \Delta(\zeta^k).$$

This completes the proof of the lemma. q.e.d.

Remark. The formula (18) has another, closely-related interpretation. It is the order of the first homology group $H_1(Y_N;\mathbb{Z})$ of the 3-manifold Y_N obtained as the N-fold cyclic cover of S^3 branched over the knot. By Poincaré duality, this group is the same as $H^2(Y_N;\mathbb{Z})$, which (when finite) classifies flat complex line bundles on Y_N. Let $p : M' \to M$ be the N-fold cyclic cover of the knot complement. If we take a flat complex line bundle on Y_N and restrict it to M', we can then push it forward by the map p to obtain a flat rank N bundle on M. In the case that N is odd, this construction produces a flat $SU(N)$ bundle on M for each element of $H^2(Y_N;\mathbb{Z})$. In another language, we are constructing N-dimensional representations of $\pi_1(M)$ as induced representations, starting from 1-dimensional representations of the subgroup $\pi_1(M')$. (In the case N even, a slight adjustment is needed, corresponding to the negative sign in the top-right entry of $\rho(m)$ above.)

Lemma 6.7. *If the quantity in Lemma 6.6 is non-zero, then the moduli space $M^w(Z_K^+;\tilde{\alpha}_0)$ is regular.*

Proof. Let A be a flat connection in $\mathrm{ad}(P)$ representing a point of this moduli space. The formal dimension of the moduli space is zero, and A is irreducible. So, to establish regularity, it is sufficient to show that H^1_A is zero, or equivalently that $\ker d_A^+$ is equal to $\mathrm{im} d_A$ in $L^2_{l,A}(Z_K^+;\mathfrak{su}_P)$. Harmonic theory on the cylindrical end manifold means that we can rephrase this as saying that the kernel of $d_A^+ \oplus d_A^*$ is zero. Any element of the kernel of this Fredholm operator is exponentially decaying on the cylindrical end, and we can integrate by parts (using the fact that A is flat), to conclude that an element a in the kernel must have $d_A^- a = 0$. Thus, a is in the kernel of the operator $d_A \oplus d_A^*$.

The kernel of $d_A \oplus d_A^*$ represents the first cohomology group with coefficients in the flat bundle $\mathfrak{su}_P$ with connection A:

$$H^1(Z_K;(\mathfrak{su}_P, A)).$$

There is no difference here between the absolute or relative group, because the connection A is irreducible on the boundary of Z_K. We must, therefore, only show that this cohomology group is zero.

Under the adjoint action of the matrix $\rho(s)$ from (20), the complexified Lie algebra $\mathfrak{sl}(N, \mathbb{C})$ decomposes as a sum

$$\mathfrak{sl}(N, \mathbb{C}) = E_0 \oplus \cdots \oplus E_{N-1},$$

where E_k is the ζ^k-eigenspace of $\mathrm{Ad}(\rho(s))$, which is the span of the elementary matrices e_{ij} with a 1 in row i and column j, with $i - j = k \pmod{N}$. There is a corresponding decomposition of the bundle $\mathfrak{su}_P$ which is respected by the flat connection, because s is central in the fundamental group.

So, we look now at the individual summands $H^1(Z_K; (E_k, A))$. The manifold Z_K is a product, and s is the element of π_1 corresponding to the circle fiber. On the circle $H^1(S^1, \lambda)$ is zero if λ is a local system with fiber $\mathbb{C}$ and non-trivial holonomy. By the Kunneth theorem, it follows that $H^1(Z_K; (E_k, A))$ is zero for k non-zero.

We are left with the summand $H^1(Z_K; (E_0, A))$. The holonomy of A on E_0 is trivial in the circle factor of Z_K; so, we may as well consider $H^1(M; (E_0, A))$. On the knot complement M, the bundle E_0 has rank $N - 1$ and the holonomy of A factors through the abelianization of the fundamental group, $H_1(M; \mathbb{Z})$, which is generated by m. Under the action of $\mathrm{Ad}(\rho(m))$, the bundle E_0 decomposes into $N - 1$ flat line bundles L_k $(k = 1, \ldots N)$; and the holonomy of L_k along the meridian is ζ^k.

To complete the proof, we have to see that $H^1(M; (L_k, A))$ is zero unless $\Delta(\zeta^k) = 0$. This is a standard result, and can be proved as follows. If the cohomology group is non-zero, let α be a non-zero element, and let α' be the pull-back to the N-fold cyclic cover $p : M' \to M$. This pull-back is an element of $H^1(M'; \mathbb{C})$ and has trivial pairing with the loop $m' = p^{-1}(m)$. It therefore, pairs non-trivially with a loop $\gamma \subset M'$ that lifts to the $\mathbb{Z}$-covering $\tilde{M}$. Thus, the pull-back of α to the $\mathbb{Z}$-covering is a non-zero class $\tilde{\alpha} \in H^1(\tilde{M}; \mathbb{C})$. By construction, $\tilde{\alpha}$ lies in the ζ^k-eigenspace of the action of the covering transformation. Up to units in the ring $\mathbb{C}[t, t^{-1}]$, the characteristic polynomial of the covering transformation acting on $H^1(\tilde{M}; \mathbb{C})$ is $\Delta(t)$. So, $\Delta(\zeta^k)$ is zero. q.e.d.

Lemma 6.8. *Suppose N is odd. Under the same hypothesis as Lemma 6.7, all the points of the moduli space $M^w(Z_K^+; \tilde{\alpha}_0)$ have the same sign.*

Proof. We wish to compare the orientations of two points $[A_0]$ and $[A_1]$ in the moduli space $M^w(Z_K^+;\tilde{\alpha}_0)$. Let $[A_t]$ be a 1-parameter family of connections in $\mathcal{A}(Z_K^+;\tilde{\alpha}_0)$ joining A_0 to A_1. Let D_t be the corresponding Fredholm operators on the cylindrical-end manifold Z_K^+, as in (14). The operators D_0 and D_1 are invertible, and so, there are canonical trivializations of the determinant lines $\det(D_0)$ and $\det(D_1)$. The lemma asserts that these canonical trivializations can be extended to a trivialization of the determinant line $\det(D_\bullet)$ on $[0,1]$.

Let $p : M' \to M$ be again the N-fold cyclic cover of the knot complement, and let $p : Z'_K \to Z_K$ be the corresponding cyclic cover of $Z_K = S^1 \times M$. Let ${Z'}_K^+$ be the corresponding cylindrical-end manifold. Set

$$A'_t = p^*(A_t).$$

These are connections on ${Z'}_K^+$, asymptotic to $p^*(\alpha_0)$ on the cylindrical end. There are corresponding Fredholm operators $\det(D'_t)$, for $t \in [0,1]$; we use a weighted Sobolev space with a small exponential weight to define the domain and codomain of D'_t, because the pull-back of $\tilde{\alpha}_0$ under the covering map p is not isolated in the representation variety. The operators D'_0 and D'_1 are invertible, as one can see by a calculation similar to that in the proof of Lemma 6.7 above.

The covering group $\mathbb{Z}/(N\mathbb{Z})$ acts on the domain and codomain of D'_t, which therefore decompose into isotypical components according to the representations of this cyclic group. The restriction of D'_t to the component on which $\mathbb{Z}/(N\mathbb{Z})$ acts trivially is just D_t (with the inconsequential change that D_t is acting now on a weighted Sobolev space). The nontrivial representations of $\mathbb{Z}/(N\mathbb{Z})$ are all complex, because N is odd, and their contribution to the determinant is therefore trivial. We can therefore replace D_t by D'_t without changing the question: we want to know whether the canonical trivializations of $\det(D'_0)$ and $\det(D'_1)$ can be extended to a trivialization of $\det(D'_\bullet)$ on the interval $[0,1]$.

Let $\tilde{A}'_t$ be the $U(N)$ connection with determinant $p^*(\theta)$, corresponding to the $PSU(N)$ connection A'_t. The description of the representations $\bar{\rho}$ corresponding to the flat connections in the moduli space $M^w(Z_K^+;\tilde{\alpha}_0)$ show that $\tilde{A}'_0$ and $\tilde{A}'_1$ are each compatible with a decomposition of the $U(N)$ bundle $P' = p^*(P)$ into a direct sum of $U(1)$ bundles (though a different decomposition in the two cases). That is, we have decompositions of the associated rank-N vector bundle E as

$$E = L_{i,1} \oplus \cdots \oplus L_{i,N}$$

for $i = 0$ and 1. Since $\tilde{A}'_i$ is projectively flat, the first Chern class of each $L_{i,m}$ is the same over the reals, independent of i and m: the common first Chern class of all these line bundles is $(1/N)c_1(p^*(w))$. We can include Z'_K in a closed 4-manifold W, and since all the $L_{i,m}$ are the same on the boundary, we can suppose they all extend to W in such a way that their rational first Chern classes are equal. The comparison of the orientations determined by D'_0 and D'_1 is then the same as the comparison between orientations

$$O(L_{i,1}, \ldots, L_{i,n}, o_W), \qquad (i = 0, 1)$$

on W. The analysis of section 4.1 shows that these two orientations depend only on the real or rational first Chern classes of the line bundles, so the two orientations agree. q.e.d.

The proof of Proposition 6.4 is now complete, because the definition of the invariant is the signed count of the points of the moduli space. q.e.d.

Remarks. When N is even, the proof of Lemma 6.8 breaks down: in the decomposition of the domain and codomain of D'_t, there is now a component on which the generator of $\mathbb{Z}/N\mathbb{Z}$ acts as -1; and we know no more about the determinant line of D'_t on this component than we did about the determinant line of the original D_t.

One should expect to prove Proposition 6.4 also in the case that the expression that appears there is zero. In this case, the moduli space is a union of circles or higher-dimensional tori, and one should try to use a holonomy perturbation to make the perturbed moduli space empty.

6.4. The Fintushel-Stern construction. Let X be a closed oriented 4-manifold with $b_2^+(X) \geq 2$. Let w be a line bundle with $c_1(w)$ coprime to N, and suppose that the integer invariant $q^w(X)$ is non-zero.

Let T be an embedded torus in X, with trivial normal bundle, and suppose that the pairing of $c_1(w)$ with $[T]$ is 1. The torus T has a closed tubular neighborhood $N \subset X$ of the form $T \times D^2$. If U is the unknot in S^3, then the 4-manifold Z_U obtained from the unknot is also $T \times D^2$. We choose an identification ϕ of Z_U with N, in such a way that the longitudinal curve l on the boundary of Z_U is matched by ϕ with the curve (point) $\times\, \partial D^2$ that links the embedded torus. Using the standard curves l, m and s, the boundary of Z_U is canonically identified with the boundary of Z_K, for any other knot K in S^3. Thus, the choice of ϕ gives us preferred diffeomorphisms

$$\psi : \partial Z_K \to \partial N$$

for all K. We now form a closed 4-manifold X_K by removing $N = Z_U$ and replacing it with Z_K:

$$X_K = Z_K \cup_\psi (X \setminus N).$$

This is Fintushel and Stern's construction, from [**8**]. The Mayer–Vietoris sequence tells us that the cohomology rings of X and X_K are canonically isomorphic.

The line bundle w on $X \backslash N$ extends to X_K. The extension w is unique if we demand the additional property that $c_1(w)^2[X_K]$ and $c_1(w)^2[X]$ are equal. The moduli spaces $M^w_\kappa(X_K)$ and $M^w_\kappa(X_K)$ have the same formal dimension, and there is a potentially non-zero integer invariant $q^w(X_K)$. We can calculate this invariant for odd N:

Proposition 6.9. *When N is odd, the integer invariant $q^w(X_K)$ can be expressed in terms of $q^w(X)$ and the Alexander polynomial Δ of K as*

$$q^w(X_K) = q^w(X) \times \prod_{k=1}^{N-1} \Delta(\zeta^k),$$

where $\zeta = e^{2\pi i/N}$.

Proof. Up to an overall ambiguity in the sign, this proposition is now a formal consequence of the calculation of the invariants $q^w(Z_K; \alpha_k)$ and $q^w(Z_U; \alpha_k)$ which are provided by Proposition 6.4, and the product law (17). These tell us that

$$q^w(X_K) = \pm\, n \times \prod_{k=1}^{N-1} \Delta(\zeta^k),$$

where $n = q^w(X \setminus N^\circ; \alpha_0)$ is a quantity independent of K. Because X_U is X, and the Alexander polynomial of the unknot is 1, it follows that $n = \pm q^w(X)$.

It remains to check the overall sign. Certainly, the sign is correct if K is the unknot. Note also that since the complex numbers $\Delta(\zeta^k)$ come in conjugate pairs when N is odd, their product is positive; so our task is to check that the sign of the invariant is independent of K. To this end, we observe that among the flat connections on the cylindrical-end manifold Z_K^+ that are enumerated in Lemma 6.6, there is a unique connection, say $[A_*]$, distinguished by the fact that the corresponding representation $\bar{\rho}$ of the fundamental group (see (19)) factors through the abelianization $H_1(Z_K; \mathbb{Z})$. Let $[B]$ be an isolated solution in the moduli space $M^w_{\boldsymbol{\omega}}((X \setminus N^\circ)^+; \tilde{\alpha}_0)$, and let $[C]$ be the connection on X_K obtained as the result of gluing $[B]$ to $[A_*]$ with a long neck. It is not important that C is actually a solution of the equations: we can instead just use cut-off functions to patch together the connections A_* and B; but it is important that the operator D_C is invertible (10), so that we can ask about the sign of the point $[C]$.

To complete the proof, we will show that the sign of $[C]$ is independent of the knot K. If K is any knot and K_0 is the unknot, then we can transform $Z_K = S^1 \times M$ into Z_{K_0} by a sequence of surgeries along tori

$$T = S^1 \times \delta \subset S^1 \times M,$$

where in each surgery, the curve δ is a null-homologous circle in M (compare [**8**]). The connection A_* (and therefore also C) is flat on such a torus and lifts to an abelian representation ρ of $\pi_1(T)$, sending the S^1 generator to the first matrix in (20) and the generator δ to 1. Thus, C is compatible with a decomposition of the associated rank-N vector bundle as a sum of flat line bundles. It now follows from excision and the material of section 4.1 that the surgery does not alter the sign of $[C]$. q.e.d.

As a particular case, we can take X to be a $K3$ surface. With w as in Section 6.1, we calculated the invariant $q^w(X)$ to be 1, for the preferred homology-orientation of X. The value of $c_1(w)$ in our calculation was $(N+1)h$; but we could equally have taken $c_1(w)$ to be just the primitive class h (with square $2(N-1)$). In this case, the invariant $q^w(X)$ would have again been 1, by Proposition 4.2. A primitive cohomology class on a $K3$ surface has pairing 1 with some embedded torus T. So, choose T with $h[T] = 1$. We are then in a position to apply the proposition above, to obtain

$$q^w(X_K) = \prod_{k=1}^{N-1} \Delta(\zeta^k)$$

for odd N. As Fintushel and Stern observe in [**8**], the 4-manifold X_K is a homotopy $K3$-surface if we take T to be a standardly embedded torus (so that $X \setminus T$ is simply connected).

7. Polynomial invariants

We now generalize the integer-valued invariant $q^w(X)$, to define polynomial invariants, as Donaldson did in [**4**]. Our approach follows [**4**], [**6**] and [**14**] closely.

7.1. Irreducibility on open sets. Let $\Omega \subset X$ be a connected, non-empty open set, chosen so that $\pi_1(\Omega; x_0) \to \pi_1(X; x_0)$ is surjective, for some $x_0 \in \Omega$. The unique continuation argument from [**6**] shows that if an anti-self-dual connection $[A] \in M^w_\kappa(X)$ is irreducible, then its restriction to Ω is also irreducible. For the perturbed equations, it may be that this result fails. The next lemma provides a suitable substitute for our purposes. It is only a very slight modification of Proposition 3.9, and we therefore omit the proof.

Lemma 7.1. *Let $\Omega \subset X$ be as above. Let κ_0 be given, and let g be a metric on X with the property that the unperturbed moduli spaces M^w_κ contain no reducible solutions for any $\kappa \leq \kappa_0$. Then, there exists $\epsilon > 0$ such that for all $\boldsymbol{\omega} \in W$ with $\|\boldsymbol{\omega}\|_W \leq \epsilon$, and all $[A]$ in $M^w_{\kappa,\boldsymbol{\omega}}(X)$, the restriction of A to Ω is irreducible.*

7.2. Cohomology classes and submanifolds. Fix as usual a $U(N)$ bundle $P \to X$, and let $\mathcal{B}^* = \mathcal{B}^*(X)$ denote the irreducible part of the usual configuration space. There is a universal family of connections, carried by a bundle

$$\mathrm{ad}(\mathbb{P}) \to X \times \mathcal{B}^*.$$

This universal bundle can be constructed as the quotient of $\mathrm{ad}(P) \times \mathcal{A}^*$ by the action of $\mathcal{G}$. Note, however, that we cannot form a universal $U(N)$ bundle $\mathbb{P}$ in this way, and our notation is not meant to imply that such a $U(N)$ bundle exists. We define a 4-dimensional cohomology class on $X \times \mathcal{B}^*$ by taking the first Pontryagin class of the adjoint bundle, with the now-familiar normalization:

$$\begin{aligned}\mathbf{c} &= -(1/2N)p_1(\mathfrak{su}_{\mathbb{P}}) \\ &= (1/2N)c_2(\mathfrak{sl}_{\mathbb{P}}) \\ &\in H^4(X \times \mathcal{B}^*; \mathbb{Q}).\end{aligned}$$

Using the slant product, we now define

$$\mu : H_i(X;\mathbb{Q}) \to H^{4-i}(\mathcal{B}^*;\mathbb{Q}) \tag{23}$$

by the formula $\mu(\alpha) = \mathbf{c}/\alpha$. We will concern ourselves here only with even-dimensional classes on X, so we define $\mathbb{A}(X)$ to be the polynomial algebra

$$\mathbb{A}(X) = \mathrm{Sym}(H_{\mathrm{even}}(X;\mathbb{Q})),$$

and we extend μ to a ring homomorphism

$$\mu : \mathbb{A}(X) \to H^*(\mathcal{B}^*;\mathbb{Q}).$$

We regard $\mathbb{A}(X)$ as a graded algebra, defining the grading of $H_i(X;\mathbb{Q})$ to be $4 - i$ so that μ respects the grading.

Now, let Σ be an oriented embedded surface in X representing a 2-dimensional homology class $[\Sigma]$. Although $\mu([\Sigma])$ is not necessarily an integral cohomology class, the definition of μ makes evident that $(2N)\mu([\Sigma])$ is integral. There is therefore a line bundle

$$\mathcal{L} \to B^*$$

with $c_1(\mathcal{L}) = (2N)\mu([\Sigma])$. We can realize this line bundle as a determinant line bundle: if $\bar\partial$ denotes the $\bar\partial$-operator on Σ, then

$$\mathcal{L} = \det(\bar\partial^*_{\mathfrak{su}(\mathbb{P})}),$$

where the notation means that we couple $\bar{\partial}^*$ to the $PSU(N)$ connections in the family over $\mathcal{B}^*$. The following lemma and Corollary 7.3 below play the same role as Lemma (5.2.9) of [**6**]. There is a slight extra complication in our present setup: we need to deal with the weaker L^p_1 norms, because of the way our perturbations behave under Uhlenbeck limits. (See Proposition 3.5.)

Lemma 7.2. *Fix an even integer $p > 2$. Then, there is a complex Banach space E and a continuous linear map $s : E \to \Gamma(\mathcal{B}^*; \mathcal{L})$ with the following properties.*

1) *For each $e \in E$, the section $s(e)$ is smooth, and furthermore the map*
$$s^\dagger : E \times \mathcal{B}^* \to \mathcal{L}$$
obtained by evaluating s is a smooth map of Banach manifolds.
2) *For each $[A] \in \mathcal{B}^*$, there exists $e \in E$ such that $s(e)([A])$ is non-zero.*
3) *For each $e \in E$, the section $s(e)$ is continuous in the L^p_1 topology on $\mathcal{B}^*$.*

Proof. Let $[A]$ belong to $\mathcal{B}^*$, and let $S \subset \mathcal{A}$ be the Coulomb slice through A. Using the definition of $\mathcal{L}$ as a determinant line bundle, construct a non-vanishing local section t for $\mathcal{L}$ on some neighborhood U of A in S. The section t can be constructed to be smooth in the L^q_1 topology for all q; so, we can take it that U is open in the L^p_1 topology. Because p is an even integer, the L^p_1 norm (or an equivalent norm) has the property that $\| \|^p_{L^p_1}$ is a smooth function. We can therefore construct a smooth cut-off β function on the L^p_1 completion of S, supported in (the completion) of U. Then, the section $s = \beta t$ can be extended as a smooth section of $\mathcal{L}$ on the L^p_1 configuration space $\mathcal{B}^*_{L^p_1}$.

Now, choose a countable section collection of points $[A_i]$ in $\mathcal{B}^*$ and sections s_i constructed as above, such that the support of the s_i covers $\mathcal{B}^*$. Fix a collection of positive real numbers c_i, and let E be the Banach space of all sequences $\mathbf{z} = \{z_i\}_{i\in\mathbb{N}}$ such that the sum
$$\sum c_i |z_i|$$
is finite. If the sequence c_i is sufficiently rapidly increasing, then for all $\mathbf{z}$ in E, the sum
$$\sum z_i s_i$$
converges to a smooth section of $\mathcal{L}$ on $\mathcal{B}^*$, and the resulting linear map $s : E \to \Gamma(\mathcal{B}^*; \mathcal{L})$ satisfies the conditions of the lemma. q.e.d.

Corollary 7.3. *Let M be a finite-dimensional manifold and $f : M \to \mathcal{B}^*$ a smooth map. Then, there is a section s of $\mathcal{L} \to \mathcal{B}^*$ such that the section $s \circ f$ of the line bundle $f^*(\mathcal{L})$ on M is transverse to zero.*

The above corollary can be applied directly to the situation that M is a regular moduli space $M^w_{\kappa,\boldsymbol{\omega}}$ of perturbed anti-self-dual connections, containing no reducibles. In this case, the zero set of s on M is a cooriented submanifold of codimension 2 whose dual class is $(2N)\mu([\Sigma])$. As with Donaldson's original construction however, we must modify the construction to obtain a codimension-2 submanifold which behaves a little better with respect to the Uhlenbeck compactification.

Let Σ be an embedded surface as above, and let $\nu(\Sigma)$ be a closed neighborhood that is also a 4-dimensional submanifold with boundary in X. By adding 1-handles to $\nu(\Sigma)$, we may suppose that this neighborhood is connected and that $\pi_1(\nu(\Sigma)) \to \pi_1(X)$ is surjective. Let $\mathcal{B}^*(\nu(\Sigma)$ be the configuration space of irreducible connections on $\nu(\Sigma)$. We suppose w and g are such that the moduli spaces $M^w_{\kappa'}$ contain no reducible solutions for all $\kappa' \leq \kappa$. Lemma 7.1 tells us that we may choose a small regular perturbation $\boldsymbol{\omega}$ so that there are well-defined restriction maps

$$f : M^w_{\boldsymbol{\omega},\kappa'} \to \mathcal{B}^*(\nu(\Sigma)).$$

The determinant line bundle $\mathcal{L}$ is defined on $\mathcal{B}^*(\nu(\Sigma))$. As in [**6**], we choose a section s of the $\mathcal{L} \to \mathcal{B}^*(\nu(\Sigma))$ such that the section $s \circ f$ is transverse to zero. This can be done by repeating the proof of the lemma above with $\mathcal{B}^*(\nu(\Sigma))$ replacing $\mathcal{B}^*(X)$. We can even choose s so that transversality holds for all $\kappa' \leq \kappa$ simultaneously.

As is customary, we write $V_\Sigma \subset \mathcal{B}^*(\nu(\Sigma))$ for the zero-set of a section s of this sort; and we write

$$M^w_{\boldsymbol{\omega},\kappa} \cap V_\Sigma$$

for the zero set in the moduli space, suppressing mention of the restriction map f.

Lemma 7.4. *Let V_Σ be constructed as above, and let $[A_n]$ be a sequence in $M^w_{\boldsymbol{\omega},\kappa} \cap V_\Sigma$, converging in the sense of Proposition 3.5 to an ideal solution $([A'], \mathbf{x})$, with $[A']$ in $M^w_{\boldsymbol{\omega},\kappa-m}$. Then, either:*

1) *the limit $[A']$ belongs to $M^w_{\boldsymbol{\omega},\kappa-m} \cap V_\Sigma$; or*
2) *one of the points of the multi-set $\mathbf{x}$ lies in the closed neighborhood $\nu(\Sigma)$.*

Proof. If $\mathbf{x}$ is disjoint from $\nu(\Sigma)$, then the A_n converge to A' in the L^p_1 topology on $\nu(\Sigma)$ after gauge transformation. The set V_Σ is closed in $\mathcal{B}^*(\nu(\Sigma))$ in the L^p_1 topology, because of the last condition in Lemma 7.2, so $[A']$ belongs to V_Σ. q.e.d.

We now wish to represent $(2N)\mu([x])$ (x is a point in X) by a codimension-4 intersection $M^w_{\boldsymbol{\omega},\kappa} \cap V_x$ in the same manner. As in [**14**], we must now take V_x to be a stratified space with smooth manifolds for strata, rather than simply a smooth submanifold.

To obtain a suitable definition of V_x, we observe that $(2N)\mu([x])$ is represented on $\mathcal{B}^*$ by $c_2(\mathfrak{sl}_{\mathbb{P}_{\frown}})$, where $\mathbb{P}_x \to \mathcal{B}^*$ is the base-point fibration: the restriction of $\mathbb{P}$ to $\{x\} \times \mathcal{B}^*$. To represent c_2 geometrically, we take $N^2 - 2$ sections $s_1, \ldots s_{N^2-2}$ of the complex vector bundle $\mathfrak{sl}_{\mathbb{P}_x}$ and examine the locus on which they are linearly dependent. We must require a transversality condition: the following lemma and corollary are essentially the same as Lemma 7.2 and Corollary 7.3 above.

Lemma 7.5. *Fix an even integer $p > 2$, and write $\mathcal{V} \to \mathcal{B}^*$ for the bundle* $\mathrm{Hom}(\mathbb{C}^{N^2-2}, \mathfrak{sl}_{\mathbb{P}_x})$. *Then, there is a complex Banach space E and a continuous linear map $s : E \to \Gamma(\mathcal{B}^*; \mathcal{V})$ with the following properties.*

1) *For each $e \in E$, the section $s(e)$ is smooth, and furthermore, the map*
$$s^\dagger : E \times \mathcal{B}^* \to \mathcal{L}$$
obtained by evaluating s is a smooth map of Banach manifolds.
2) *For each $[A] \in \mathcal{B}^*$, the map $E \to \mathcal{V}_{[A]}$ obtained by evaluating s is surjective.*
3) *For each $e \in E$, the section $s(e)$ is continuous in the L^p_1 topology on $\mathcal{B}^*$.*

Corollary 7.6. *Let M be a finite-dimensional manifold and $f : M \to \mathcal{B}^*$ a smooth map. Then, there is a section s of* $\mathrm{Hom}(\mathbb{C}^{N^2-2}, \mathfrak{sl}_{\mathbb{P}_x}) \to \mathcal{B}^*$ *such that the section $s \circ f$ of the pull-back $f^*(\mathcal{L})$ on M is transverse to the stratification of* $\mathrm{Hom}(\mathbb{C}^{N^2-2}, \mathfrak{sl}_{\mathbb{P}_x})$ *by rank.*

As with the case of 2-dimensional classes, we now take a closed neighborhood $\nu(x)$ that is also a manifold with boundary. We require that $\pi_1(\nu(x), x) \to \pi_1(X, x)$ is surjective. We can again suppose that for $\kappa' \le \kappa$, the restriction of every solution in $M^w_{\boldsymbol{\omega},\kappa'}$ to $\nu(x)$ is irreducible, so that we have maps
$$f : M^w_{\boldsymbol{\omega},\kappa'} \to \mathcal{B}^*(\nu(x)).$$
We take a section s satisfying the transversality condition of the above corollary, taking M to be the union of the $M^w_{\boldsymbol{\omega},\kappa'}$ and replacing $\mathcal{B}^*$ by $\mathcal{B}^*(\nu(x))$ in the statement. Write
$$V_x \subset \mathcal{B}^*(\nu(x))$$
for the locus where s does not have full rank. Then, the intersection
$$M^w_{\boldsymbol{\omega},\kappa} \cap V_x$$

is a stratified subset, with each stratum a smooth, cooriented manifold, and top stratum of real codimension 4 in $M^w_{\boldsymbol{\omega},\kappa}$. The dual class to V_x is $(2N)\mu([x])$. Lemma 7.4 continues to hold with V_x in place of V_Σ.

7.3. Polynomials in the coprime case. We now have the necessary definitions to construct the $PSU(N)$ generalization of Donaldson's polynomial invariants, in the case that $w = c_1(P)$ is coprime to N.

We suppose as usual that $b^+_2(X)$ is at least 2. Because of the coprime condition, we can choose a Riemannian metric g on X such that the moduli spaces M^w_κ contain no reducible solutions. Given some κ_0, we may find ϵ so that for all perturbations $\boldsymbol{\omega}$ with $\|\boldsymbol{\omega}\|_W \leq \epsilon$, the moduli spaces $M^w_{\boldsymbol{\omega},\kappa}$ contain no reducibles for $\kappa \leq \kappa_0$. Amongst such small perturbations, we may choose one so that the moduli spaces are regular. If N is even, choose a homology orientation for X so as to orient the moduli spaces.

Let d be the dimension of $M^w_{\boldsymbol{\omega},\kappa}$, and suppose that d is even. (This parity condition holds if N is odd, or if $b^+_2(X) - b_1(X)$ is odd.) Consider an element of $\mathbb{A}(X)$, given by a monomial

$$z = [x]^r[\Sigma_1][\Sigma_2]\dots[\Sigma_t]. \tag{24}$$

Suppose that $4r + 2t = d$, so that z is of degree d. We may assume that the surfaces Σ_i intersect only in pairs, and we can represent the class $[x]$ by r distinct points $x_1, \dots, x_r$ all disjoint for the surfaces. Choose neighborhoods

$$\nu(x_1), \dots, \nu(x_r), \nu(\Sigma_1), \dots, \nu(\Sigma_t)$$

satisfying the surjectivity condition on π_1 as above. We may assume that these satisfy the same intersection conditions: the only non-empty intersections are possible intersections of pairs $\nu(\Sigma_i) \cap \nu(\Sigma_j)$. Construct submanifolds (or stratified subspaces) V_{x_i} and V_{Σ_j} as above. By a straightforward extension of the transversality argument, we may arrange that all multiple intersections

$$M^w_{\boldsymbol{\omega},\kappa'} \cap V_{x_{i_1}} \cap \dots \cap V_{x_{i_p}} \cap V_{\Sigma_{j_1}} \cap \dots \cap V_{\Sigma_{j_q}}$$

are transverse. In the case of the V_{x_i}, transversality means transversality for each stratum. The intersection

$$M^w_{\boldsymbol{\omega},\kappa} \cap V_{x_1} \cap \dots \cap V_{x_r} \cap V_{\Sigma_1} \cap \dots \cap V_{\Sigma_t} \tag{25}$$

has dimension zero: it is an oriented set of points, contained entirely in the top stratum of the V_{x_i}. The next lemma is proved by the same dimension-counting that is used for the case $N = 2$:

Lemma 7.7. *Under the stated transversality assumptions, the zero-dimensional intersection* (25) *is compact.*

Finally, we can state:

Proposition 7.8. *Fix N, and let X be a closed, oriented smooth 4-manifold with $b_2^+(X) \geq 2$, equipped with a homology orientation o_X. Let w be a line bundle with $c_1(w)$ coprime to N, and let P be a $U(N)$ bundle on X with $\det(P) = w$, such that the corresponding moduli space M_κ^w has formal dimension an even integer d. Let V_{x_i} and V_{Σ_j} be as above. Then, the signed count of the points in the intersection (25) has the following properties:*

1) *It is independent of the choice of Riemannian metric g, perturbation $\boldsymbol{\omega}$, the neighborhoods $\nu(x_i)$ and $\nu(\Sigma_j)$, and the sections defining the submanifolds V_{x_i} and V_{Σ_j}, subject to the conditions laid out in the construction above.*
2) *It depends on the Σ_j only through their homology classes $[\Sigma_j]$.*
3) *It is linear in the homology class of each Σ_j.*

Proof. The proof can be modeled on the arguments given in [**4**] and [**6**]. q.e.d.

Staying with our assumptions that $b_2^+(X) \geq 2$, that $c_1(w)$ is coprime to N and that a homology orientation o_X is given, we now define the polynomial invariant

$$q^w(X) : \mathbb{A}(X) \to \mathbb{Q}$$

by declaring its value on the monomial z given in (24) to be given by

$$q^w(X)(z) = (2N)^{-r-t} \#(M^w_{\boldsymbol{\omega},\kappa} \cap V_{x_1} \cap \cdots \cap V_{x_r} \cap V_{\Sigma_1} \cap \cdots \cap V_{\Sigma_t}),$$

where κ is chosen (if possible) to make the moduli space d-dimensional. If there is no such κ, we define $q^w(X)(z)$ to be zero for this monomial of degree d. The factor $(2N)^{-r-t}$ is there because the subvarieties V_Σ and V_μ are dual to $(2N)\mu([\Sigma])$ and $(2N)\mu([x])$ respectively: so, if the moduli space were actually compact, then the $q^w(X)(z)$ would be the pairing

$$q^w(X)(z) = \langle \mu(z), [M^w_{\boldsymbol{\omega},\kappa}] \rangle.$$

The definition of $q^w(X)$ is set up so that the integer invariant that we defined earlier can be recovered from this more general invariant as the value $q^w(X)(1)$.

7.4. The non-coprime case. Thus far, we have defined $q^w(X)$ only in the case that $c_1(w)$ is coprime to N, because this condition allows us to avoid the difficulties arising from reducibles. In the case that $c_1(w)$ is *not* coprime to N, however, we can still give a useful definition of $q^w(X)$ by a formal device previously used in [**14**].

Given X and w (with $b^+(X) \geq 2$ as usual), we consider the blow-up

$$\tilde{X} = X \# \bar{\mathbb{CP}}^2.$$

A maximal positive-definite subspace of $H^2(X;\mathbb{R})$ is also a maximal positive-definite subspace of $H^2(\tilde{X};\mathbb{R})$; so a homology-orientation o_X determines a homology-orientation $o_{\tilde{X}}$ for the blow-up. We set

$$\tilde{w} = w + e,$$

where $e = \mathrm{PD}(E)$ is the Poincaré dual of a chosen generator of E in $H_2(\bar{\mathbb{CP}}^2;\mathbb{Z})$. Notice that $\tilde{w}$ is coprime to N on $\tilde{X}$, so $q^{\tilde{w}}(\tilde{X})$ is already well-defined. We now *define* $q^w(X)$ in terms of $q^{\tilde{w}}(\tilde{X})$ by the formula

$$q^w(X)(z) = q^{\tilde{w}}(\tilde{X})(E^{N-1}z), \tag{26}$$

for $z \in \mathbb{A}(X) \subset \mathbb{A}(\tilde{X})$. (The result is independent of the choice of generator E.)

The rationale for this definition, as in [**14**], is that the formula (26) is an *identity* in the case that $c_1(w)$ is already coprime to N (so that the left-hand side is defined by our previous construction). This is a blow-up formula, which in the case $N = 2$ was first proved in [**13**]. We now outline the proof of this blow-up formula, for general N.

On the manifold $\bar{\mathbb{CP}}^2$, let Q be the $U(N)$ bundle with $c_1(Q) = e$ and $c_2(Q) = 0$. The characteristic number $\kappa(Q)$ for this bundle is $(N-1)/2N$, so the formal dimension d of the moduli space is $2(N-1)-(N^2-1)$. If we pick a base-point $y_0 \in \bar{\mathbb{CP}}^2$ and consider the *framed* moduli space (the quotient of the space of solutions by the based gauge group), then the formal dimension is $2(N-1)$. The fact that $\bar{\mathbb{CP}}^2$ has positive scalar curvature and anti-self-dual Weyl curvature means that the framed moduli space is regular; and since its dimension is less than N^2-1, the moduli space consists only of reducibles. The framed moduli space therefore consists of a single $\mathbb{CP}^{N-1}$, which is the gauge orbit of a connection $[A]$ compatible with the splitting of Q as a sum

$$Q = l \oplus Q'.$$

(The first is the line bundle with $c_1(l) = e$, and the second summand is a trivial $U(N-1)$ bundle.) The stabilizer Γ_A in the gauge group $\mathcal{G}$ is $S(U(1) \times U(N-1))$. Let us write N for this framed moduli space. There is a universal bundle

$$\mathrm{ad}(\mathbb{Q}) \to \bar{\mathbb{CP}}^2 \times N$$

and a corresponding slant-product map

$$\mu : H_i(\bar{\mathbb{CP}}^2) \to H^{4-i}(N).$$

We can describe $\mathbb{Q}$ as the quotient

$$\mathrm{ad}(\mathbb{Q}) = \frac{\mathrm{ad}(Q) \times SU(N)}{S(U(1) \times U(N-1))},$$

which makes it the adjoint bundle of the $U(N)$ bundle

$$\mathbb{Q} = \frac{Q \times U(N)}{U(1) \times U(N-1)}.$$

The decomposition of Q gives rise to a decomposition

$$\mathbb{Q} = (w \boxtimes \tau) \oplus (\mathbb{C} \boxtimes \tau^{\perp}),$$

where $\tau \to N = \mathbb{CP}^{N-1}$ is the tautological line bundle and $\tau^{\perp}$ is its complement. On the submanifold $\mathbb{CP}^1 \times \mathbb{CP}^1 \subset \bar{\mathbb{CP}}^2 \times N$, the bundle $\mathbb{Q}$ has $c_1^2 = 0$ and c_2 is a generator. So $\mu(E)$ is an integral generator for $H^2(N) = \mathbb{Z}$. In particular,

$$\left\langle \mu(E)^{N-1}, [N] \right\rangle = \pm 1. \tag{27}$$

Now, we consider the blow-up formula in the case that w is coprime to N on X and that the moduli space $M^w_{\boldsymbol{\omega},\kappa}(X)$ is zero-dimensional, leading to an integer invariant $q^w(X)$. Let x_0 be a base-point in X. Using Lemma 5.5 and the conformal invariance of the anti-self-duality equations, we can find a good pair $(g, \boldsymbol{\omega})$ with the property that the support of the perturbation is disjoint from a neighborhood of x_0. We can then form a connected sum $\tilde{X} = X \# \bar{\mathbb{CP}}^2$, attaching the $\bar{\mathbb{CP}}^2$ summand to X in a neighborhood of x_0. We equip the connected sum with a metric g_R containing a neck of length R, in the usual way. On $\tilde{X}$, we form the bundle $\tilde{P}$ as the sum of the bundles $P \to X$ and $Q \to \bar{\mathbb{CP}}^2$. As in 5.4, we may regard $\boldsymbol{\omega}$ as defining a perturbation on $\tilde{X}$, with support disjoint from the neck and from the $\bar{\mathbb{CP}}^2$ summand. The formal dimension of the moduli space on $\tilde{X}$ is $2(N-1)$, and the moduli space is compact once R is sufficiently large. Standard gluing techniques apply, to show that the moduli space for large R is a product of the zero-dimensional moduli space $M^w_{\boldsymbol{\omega},\kappa}(X)$ and the framed moduli space $N = \mathbb{CP}^{N-1}$ for $\bar{\mathbb{CP}}^2$. This is a diffeomorphism of oriented manifolds, up to an overall sign. From (27), we deduce that

$$q^w(X)(1) = \pm q^{\tilde{w}}(\tilde{X})(E^{N-1}),$$

which is the desired formula (26), up to sign, for the case $z = 1$. To prove the general case, we take z to be a monomial as in (24), and we replace the moduli space $M^w_{\boldsymbol{\omega},\kappa}(X)$ by the cut-down moduli space (25), having chosen the neighborhoods $\nu(\Sigma_j)$ etc. to be disjoint from the region in which the connected sum is formed. The argument is then essentially the same as the case $z = 1$. To identify the correct sign for the blow-up formula, we can again make a comparison with the Kähler case.

Having dealt in this way with the non-coprime case, we are allowed now to take w trivial, so as to arrive at a satisfactory definition of the $SU(N)$ Donaldson invariant corresponding to a bundle P with $c_1 = 0$.

8. Further thoughts

8.1. Higher Pontryagin classes. We have defined μ in (23) using the 4-dimensional characteristic class $\mathbf{c}$, taking our lead from the usual definition of the $N = 2$ invariants. But there are other possibilities. The easiest generalizations is of the 4-dimensional class $\mu([x])$, where $[x]$ is the generator of $H_0(X)$. Recall that the cohomology class $\mu([x])$ is a rational multiple of the first Pontryagin class, of the base-point bundle $\mathrm{ad}(\mathbb{P}_x) \to \mathcal{B}^*$. Let us drop the factor $-(1/2N)$ in from the definition and write

$$\nu_4 = p_1(\mathrm{ad}(\mathbb{P}_x)).$$

We can consider more generally now the class

$$\nu_{4j} = p_j(\mathrm{ad}(\mathbb{P}_x))$$

in $H^{4j}(\mathcal{B}^*)$. On the moduli spaces $M^w_{\boldsymbol{\omega},\kappa}$, this can be represented as the dual of a subvariety

$$M^w_{\boldsymbol{\omega},\kappa} \cap V_{j,x},$$

where $V_{j,x}$ is again pulled back from a neighborhood of x. We restrict as before to the coprime case, with $b_2^+(X) \geq 2$, in which case, we consider a zero-dimensional intersection

$$M^w_{\boldsymbol{\omega},\kappa} \cap V_{j_1,x_1} \cap \cdots \cap V_{j_r,x_r}. \tag{28}$$

If $j_i \leq N - 1$ for all i, and if the V_{j_i,x_i} are chosen so that all relevant intersections are transverse, then this zero-dimensional intersection will be compact. So, we can define Donaldson-type invariants using the characteristic class ν_j of the base-point fibration, as long as $j \leq N - 1$.

For $j \geq N$, the dimension-counting fails. For example, suppose that $M^w_{\boldsymbol{\omega},\kappa}$ is $4N$-dimensional, so that $M^w_{\boldsymbol{\omega},\kappa-1}$ is 0-dimensional, and suppose that w is coprime to N and that the moduli spaces are regular and without reducibles. We can try taking $j = N$, and considering the 0-dimensional intersection

$$M^w_{\boldsymbol{\omega},\kappa} \cap V_{N,x},$$

so as to arrive at an evaluation of p_N of the base-point fibration on the (potentially non-compact) moduli space. The problem is that the above zero-dimensional intersection need not be compact. We cannot rule out the possibility that there is a sequence $[A_n]$ of solutions belonging to this intersection, converging to an ideal connection $([A'], x')$, with $[A'] \in M^w_{\boldsymbol{\omega},\kappa-1}$ and x' in the neighborhood $\nu(x)$.

8.2. Other Lie groups. The approach we have taken to define the $SU(N)$ polynomial invariants (even in the case of a bundle with $c_1 = 0$) makes essential use of a $U(N)$ bundle on the blow-up, in order to avoid reducible solutions. This approach does not seem to be available in general: we cannot define E_8 Donaldson invariants this way, for example. It is possible that by taking the instanton number to be large, one could sidestep the difficulties that arise from reducible solutions, and so arrive at an alternative construction. This was Donaldson's original approach, in [**4**]. It seems likely that this approach, if successful, would be technically more difficult than what we have done here for $SU(N)$.

8.3. Mahler measure. The large-N behaviour of the $SU(N)$ Donaldson invariants is briefly considered in [**15**]. In the particular examples we have studied in this paper, the leading term of the invariant for large N is easily visible, and turns out to have a rather interesting interpretation.

We turn again to the manifold Z_K with torus boundary, obtained by taking the product of the circle and a knot-complement, and we look again at the integer-valued invariant $q^w(Z_K; \alpha_0)$. The formula for this invariant in Proposition 6.4 relates it to the geometric mean of the Alexander polynomial Δ_K on the unit circle: as long as there are no zeros of Δ_K on the unit circle, we have

$$q^w(Z_K; \alpha_0)_N \sim \alpha^N,$$

as N increases through odd integers, where

$$\alpha = \exp \int_0^1 \log |\Delta_K(e^{2\pi i t})| \, dt.$$

This asymptotic result is very easily obtained by considering the expression for

$$(1/N) \log |q^w(Z_K; \alpha_0)_N|$$

arising from Proposition 6.4: it is a Riemann-sum approximation to the integral of the smooth function $|\Delta_K(z)|$ on the circle. The remaining point is then that the Riemann sum approximates the integral of a periodic function with an error which is at most $O(1/N^3)$ when the integrand is C^2.

If Δ_K has zeros on the unit circle, the situation is much more delicate. We have only the weaker statement

$$\lim_{N\to\infty} (1/N) \log q^w(Z_K; \alpha_0)_N = \log \alpha, \tag{29}$$

as N runs through those odd integers for which $q^w(Z_K; \alpha_0)_N$ is non-zero. The proof of (29) is given in [**11**] and in [**21**], in answer to a question raised by Gordon in [**12**]. A different proof and an extension of the result are contained in [**23**].

The quantity α is known as the *Mahler measure* of the polynomial. It can also be expressed as

$$\alpha = |a| \prod_{\lambda} \max(1, |\lambda|),$$

where a is the leading coefficient of the polynomial and λ runs through the roots in the complex plane. The Mahler measure of the Alexander polynomial, as well as the Mahler measures of the Jones and A-polynomials, have appeared elsewhere in the literature, in connection with interesting quantities such as the hyperbolic volume of the knot complement. See [**24**, **17**] and [**2**], for example.

9. Acknowledgement

This paper owes a lot to the author's continuing collaboration with Tom Mrowka. In particular, the discussion of holonomy perturbations in Section 3 is based in part on earlier, unpublished joint work.

References

[1] A. Borel, R. Friedman, & J.W. Morgan, *Almost commuting elements in compact Lie groups*, Mem. Amer. Math. Soc. **157**(**747**) (2002) x+136, MR 1895253, Zbl 0993.22002.

[2] D.W. Boyd, *Mahler's measure and invariants of hyperbolic manifolds*, in 'Number theory for the millennium, I' (Urbana, IL, 2000), A K Peters, Natick, MA, 2002, 127–143, MR 1956222, Zbl 1030.11055.

[3] S.K. Donaldson, *The orientation of Yang-Mills moduli spaces and* 4*-manifold topology*, J. Differential Geom. **26**(**3**) (1987) 397–428, MR 910015, Zbl 0683.57005.

[4] S.K. Donaldson, *Polynomial invariants for smooth four-manifolds*, Topology **29**(**3**) (1990) 257–315, MR 1066174, Zbl 0715.57007.

[5] S.K. Donaldson, *Floer homology groups in Yang-Mills theory*, Cambridge Tracts in Mathematics, **147**, Cambridge University Press, Cambridge, 2002, With the assistance of M. Furuta and D. Kotschick, MR 1883043, Zbl 0998.53057.

[6] S.K. Donaldson & P.B. Kronheimer, *The geometry of four-manifolds*, Oxford Mathematical Monographs, The Clarendon Press Oxford University Press, New York, 1990, Oxford Science Publications, MR 1079726, Zbl 0820.57002.

[7] P.M.N. Feehan & T.G. Leness, PU(2) *monopoles.* I. *Regularity, Uhlenbeck compactness, and transversality*, J. Differential Geom. **49**(**2**) (1998) 265–410, MR 1664908, Zbl 0998.57057.

[8] R. Fintushel & R.J. Stern, *Knots, links, and* 4*-manifolds*, Invent. Math. **134**(**2**) (1998) 363–400, MR 1650308, Zbl 0914.57015.

[9] A. Floer, *An instanton-invariant for* 3*-manifolds*, Comm. Math. Phys. **118**(**2**) (1988) 215–240, MR 956166, Zbl 0684.53027.

[10] D.S. Freed & K.K. Uhlenbeck, *Instantons and four-manifolds*, second ed., Mathematical Sciences Research Institute Publications, **1**, Springer–Verlag, New York, 1991, MR 1081321, Zbl 0559.57001.

[11] F. González-Acuña & H. Short, *Cyclic branched coverings of knots and homology spheres*, Rev. Mat. Univ. Complut. Madrid **4**(**1**) (1991) 97–120, Zbl 0756.5700.

[12] C.McA. Gordon, *Knots whose branched cyclic coverings have periodic homology*, Trans. Amer. Math. Soc. **168** (1972) 357–370, MR 0295327, Zbl 0238.55001.

[13] D. Kotschick, SO(3)*-invariants for* 4*-manifolds with* $b_2^+ = 1$, Proc. London Math. Soc. (3) **63**(**2**) (1991) 426–448, MR 1114516, Zbl 0699.53036.

[14] P.B. Kronheimer & T.S. Mrowka, *Embedded surfaces and the structure of Donaldson's polynomial invariants*, J. Differential Geom. **41**(**3**) (1995) 573–734, MR 1338483, Zbl 0842.57022.

[15] M. Mariño & G. Moore, *The Donaldson-Witten function for gauge groups of rank larger than one*, Comm. Math. Phys. **199**(**1**) (1998) 25–69, MR 1660219, Zbl 0921.58080.

[16] S. Mukai, *Symplectic structure of the moduli space of sheaves on an abelian or* $K3$ *surface*, Invent. Math. **77**(**1**) (1984) 101–116, MR 751133, Zbl 0565.14002.

[17] H. Murakami, *Mahler measure of the colored Jones polynomial and the volume conjecture*, Sūrikaisekikenkyūsho Kōkyūroku **1279** (2002) 86–99, Volume conjecture and its related topics (Japanese) (Kyoto, 2002), MR 1953834, Zbl 1024.57503.

[18] H. Nakajima & K. Yoshioka, *Instanton counting on blowup*. I. 4*-dimensional pure gauge theory*, preprint, arXiv:math.AG/0306198.

[19] N. Nekrasov & Okounkov A., *Seiberg-Witten Theory and Random Partitions*, preprint, arXiv:hep-th/0306238.

[20] N.A. Nekrasov, *Seiberg-Witten prepotential from instanton counting*, Adv. Theor. Math. Phys. **7**(**5**) (2003) 831–864, MR 2045303, Zbl 1056.81068.

[21] R. Riley, *Growth of order of homology of cyclic branched covers of knots*, Bull. London Math. Soc. **22**(**3**) (1990) 287–297, MR 1041145, Zbl 0727.57002.

[22] D. Rolfsen, *Knots and links*, Publish or Perish Inc., Berkeley, Calif., 1976, Mathematics Lecture Series, **7**, MR 1277811, Zbl 0854.57002.

[23] D.S. Silver & S.G. Williams, *Mahler measure, links and homology growth*, Topology **41**(**5**) (2002) 979–991, MR 1923995, Zbl 1024.57007.

[24] D.S. Silver & S.G. Williams, *Mahler measure of Alexander polynomials*, J. London Math. Soc. (2) **69**(**3**) (2004) 767–782, MR 2050045, Zbl 1055.57017.

[25] C.H. Taubes, *Casson's invariant and gauge theory*, J. Differential Geom. **31**(**2**) (1990) 547–599, MR 1037415, Zbl 0702.53017.

[26] E. Witten, *Monopoles and four-manifolds*, Math. Res. Lett. **1**(**6**) (1994) 769–796, MR 1306021, Zbl 0867.57029.

Harvard University
Cambridge MA 02138