Unspecified Book Proceedings Series

On the integer valued SU(3) Casson invariant

Hans U. Boden, Christopher M. Herald, and Paul Kirk

In 1986, Andrew Casson constructed a new invariant $\lambda_{SU(2)}(X)$ for oriented, integral homology 3-spheres X by counting conjugacy classes of irreducible $SU(2)$ representations $\alpha\colon \pi_1 X \to SU(2)$ with sign. The homology restriction on the 3-manifold X guarantees that the set of conjugacy classes of irreducible $SU(2)$ representations is compact, and Casson described a method for perturbing the representation variety to obtain an oriented manifold of dimension zero and used compactness of the irreducible stratum to prove that the resulting algebraic count of points is independent of perturbation.

This article is a survey of the results in [**8, 9**] generalizing the Casson invariant to the $SU(3)$ setting. The main challenge is that the irreducible stratum is no longer compact, and consequently a naive count of conjugacy classes of irreducible representations $\alpha\colon \pi_1 X \to SU(3)$ does not produce a well defined invariant of integral homology 3-spheres.

A similar problem was encountered by Walker in generalizing Casson's original invariant from integral to rational homology 3-spheres [**24**]. The homology condition on a rational homology 3-sphere X does not guarantee compactness of the set of irreducible representations $\alpha\colon \pi_1 X \to SU(2)$. As a consequence a direct count (with sign) of the conjugacy classes of irreducible $SU(2)$ representations ends up depending on the choice of perturbation. Walker solved this problem by introducing a correction term defined entirely in terms of the stratum of reducible (abelian) representations.

In generalizing the Casson invariant from $SU(2)$ to $SU(3)$, the direct count of conjugacy classes of irreducible representations $\alpha\colon \pi_1 X \to SU(3)$ depends on the choice of perturbation, and so again one needs to introduce a correction term defined in terms of the stratum of reducible representations. Correction terms of several different types have been proposed (cf. [**5, 8, 13**]). All three proposals adopt a gauge theoretic point of view, treating the Chern-Simons function or a perturbation of it as a Morse function on the space of connections and viewing the Casson invariant as an Euler characteristic for the space of connections modulo gauge, a framework first developed for $SU(2)$ by Floer and Taubes (cf. [**16, 23**]).

In this article, we describe the complications involved in defining an invariant in the $SU(3)$ gauge theory setting. We discuss several alternatives for resolving the difficulties, following [**5, 8**]. We also present some calculations of the resulting invariant $\tau_{SU(3)}$, following [**9**]. Our approach to the integer valued $SU(3)$ Casson invariant differs from that given in [**8**] in that here, $\tau_{SU(3)}(X)$ is defined as the average of two $SU(3)$ invariants, $\tau_+(X)$ and $\tau_-(X)$. Both of these invariants are a priori integer valued, and they differ only in the choice of basepoints used to determine the correction term.

1. The Casson invariant as an intersection number

We begin with a review of Casson's approach to the transversality and perturbation issues. First, we establish some notation. Given a space X and a Lie group G, let

$$R(X,G) = \mathrm{Hom}(\pi_1 X, G)/\text{conjugation}$$

denote the (real-algebraic) variety of conjugacy classes of representations $\alpha\colon \pi_1 X \to G$. Let $R^*(X,G) \subset R(X,G)$ denote the subset of conjugacy classes of representations $\alpha\colon \pi_1 X \to G$ whose stabilizer $\{g \in G \mid g\alpha(x)g^{-1} = \alpha(x) \text{ for all } x \in \pi_1 X\}$ coincides with the center of G. Such representations are called *irreducible representations*, and the complement $R(X,G) - R^*(X,G)$ is the subvariety of *reducible representations*.

If G is a subgroup of the general linear group $GL(n,\mathbb{C})$ (resp. $GL(n,\mathbb{R})$), then a representation $\alpha : \pi_1 X \to G$ induces a representation of $\pi_1 X$ on the vector space $\mathbb{C}^n$ (resp. $\mathbb{R}^n$). The above definition of reducibility is equivalent to the existence of a proper invariant complex (real) subspace.

A continuous map $f\colon X \to Y$ induces an algebraic map $f^*\colon R(Y,G) \to R(X,G)$. This map is injective if $f_*\colon \pi_1 X \to \pi_1 Y$ is surjective. Notice that if α is a reducible representation, then $f^*(\alpha)$ is also reducible.

Suppose X is a closed 3-manifold, and choose a Heegaard splitting (H_1, H_2, Σ) for X; i.e. H_1 and H_2 are solid handlebodies of genus g with boundary a closed surface Σ and

$$X = H_1 \cup_\Sigma H_2.$$

The inclusions $\pi_1\Sigma \to \pi_1 H_1$ and $\pi_1\Sigma \to \pi_1 H_2$ are surjective, and so the Seifert-Van Kampen theorem implies that $\pi_1 H_1 \to \pi_1 X$ and $\pi_1 H_2 \to \pi_1 X$ are surjective. Thus every morphism in the diagram

$$\begin{tikzcd}
 & R(H_1, SU(2)) \arrow[dr] & \\
R(X, SU(2)) \arrow[ur] \arrow[dr] & & R(\Sigma, SU(2)) \\
 & R(H_2, SU(2)) \arrow[ur] &
\end{tikzcd}$$

is injective. This identifies $R(X,SU(2))$ as the intersection of $R(H_1,SU(2))$ with $R(H_2,SU(2))$ in $R(\Sigma,SU(2))$.

The representation varieties $R(H_1,SU(2)), R(H_2,SU(2))$ and $R(\Sigma,SU(2))$ are singular along the reducible representations (those conjugacy classes whose stabilizer is larger than the center $\mathbb{Z}_2 \subset SU(2)$), and if $g \geq 2$ the irreducible points $R^*(\Sigma,SU(2))$, $R^*(H_1,SU(2))$, and $R^*(H_2,SU(2))$ are smooth (non-compact) manifolds of dimension $6g-6$, $3g-3$, $3g-3$, respectively.

When X is an integral homology 3-sphere, every intersection point except for the trivial representation lies in the intersection of the irreducible strata

$$R^*(X,SU(2)) = R^*(H_1,SU(2)) \cap R^*(H_2,SU(2)).$$

This is because a reducible $SU(2)$ representation has abelian image, but $\pi_1 X$ has no nontrivial abelian representations if $H_1(X;\mathbb{Z}) = 0$. (There are nontrivial reducible representations if one replaces $SU(2)$ by $SU(3)$, or if $H_1(X;\mathbb{Z})$ is non-zero; this is the source of the difficulty in extending Casson's invariant to $SU(3)$ or to rational homology 3-spheres.)

An orientation on X induces orientations on each of the three representation varieties $R^*(H_1, SU(2)), R^*(H_2, SU(2))$ and $R^*(\Sigma, SU(2))$ in a natural way, and Casson defined $\lambda_{SU(2)}(X)$ to be one half times the algebraic intersection number of these two $(3g-3)$-dimensional submanifolds of $R^*(\Sigma, SU(2))$:

$$\lambda_{SU(2)}(X) := \tfrac{1}{2} R^*(H_1, SU(2)) \cdot R^*(H_2, SU(2)).$$

To make sense of this intersection number Casson proved that one can make $R^*(H_1, SU(2))$ and $R^*(H_2, SU(2))$ transverse using a compactly supported isotopy of $R^*(H_2, SU(2))$ in $R^*(\Sigma, SU(2))$.

That one can choose the isotopy to be compactly supported in this context is a crucial point. Even though neither $R^*(H_1, SU(2))$ nor $R^*(H_2, SU(2))$ are themselves compact, their intersection $R^*(X, SU(2)) = R^*(H_1, SU(2)) \cap R^*(H_2, SU(2))$ is compact. This follows from the fact that, up to conjugacy, there is one only reducible representation (the trivial representation), and its conjugacy class is an isolated point in $R(X, SU(2))$.

If X is not a homology 3-sphere, then $R^*(X, SU(2))$ is not generally compact. This leads to problems in trying to use the algebraic intersection number of $R^*(H_1, SU(2))$ and $R^*(H_2, SU(2))$ to produce a well defined invariant. For example, if X is a rational homology 3-sphere but not an integral homology 3-sphere (so $H_1(X;\mathbb{Z})$ is nontrivial and finite), Walker observed that as one perturbs the representation variety $R(H_1, SU(2))$ relative to $R(H_2, SU(2))$ in $R(\Sigma, SU(2))$, an intersection point of $R^*(H_1, SU(2))$ and $R^*(H_2, SU(2))$ can suddenly appear out of (or disappear into) the stratum of reducible representations. Coincident with such a birth (or death) is a simultaneous change in a certain Maslov index determined by the varieties $R(H_1, U(1))$ and $R(H_2, U(1))$ in $R(\Sigma, U(1))$ and the normal bundles of these strata in the full representation space.

Using these Maslov indices, Walker derived a correction term involving the reducible representations which cancels out the effect of these births and deaths when X is a rational homology 3-sphere. Thus Walker was able to extend Casson's invariant to rational homology 3-spheres by adding this correction term to the intersection number of $R^*(H_1, SU(2))$ and $R^*(H_2, SU(2))$ [**24**].

The behavior of Casson's invariant and Walker's extension under various cut-and-paste operations is well understood. The most important such result is the Dehn surgery formula for the Casson-Walker invariant. Motivated by the surgery formula, Lescop gave a combinatorial definition of an invariant for all 3-manifolds which agrees with the Casson-Walker invariant on rational homology 3-spheres [**22**].

We now briefly review the behavior of the Casson-Walker-Lescop invariant $\lambda_{SU(2)}$ under change of orientation, connected sum, and Dehn surgery operations (cf. §1.5, [**22**]). If $-X$ denotes the 3-manifold X with the opposite orientation, then

$$\lambda_{SU(2)}(-X) = (-1)^{b_1(X)+1} \lambda_{SU(2)}(X),$$

where $b_1(X) = \operatorname{rank} H_1(X;\mathbb{Z})$. Furthermore,

$$\lambda_{SU(2)}(X_1 \# X_2) = |H_1(X_2;\mathbb{Z})| \lambda_{SU(2)}(X_1) + |H_1(X_1;\mathbb{Z})| \lambda_{SU(2)}(X_2),$$

where

$$|H_1(X;\mathbb{Z})| = \begin{cases} \text{the order of } H_1(X;\mathbb{Z}) & \text{if } b_1(X) = 0, \\ 0 & \text{if } b_1(X) > 0. \end{cases}$$

Finally, Walker and Lescop have generalized Casson's original surgery formula, which states: Given a knot $K \subset X$ in an integral homology 3-sphere X, if $\Delta_K(t)$ is the (symmetric) Alexander polynomial of K and X_n is the integral homology 3-sphere obtained by performing $1/n$-Dehn surgery on K, then

$$\lambda_{SU(2)}(X_n) = \lambda_{SU(2)}(X) + \frac{n}{2}\Delta_K''(1).$$

There is a similar formula for surgeries on a link which effectively computes the Casson-Walker-Lescop invariant for all 3-manifolds, since every 3-manifold is surgery on some link L in S^3.

In the next section, we describe Taubes' reinterpretation of the Casson invariant as an Euler characteristic [**23**], using gauge theory. This is the approach we take in generalizing the Casson invariant to the $SU(3)$ context.

2. The Casson invariant as an Euler characteristic

Throughout the remainder of the paper, we assume that X is an oriented integral homology 3-sphere. We now recall the gauge theoretic description of the Casson invariant provided by Taubes [**23**]. For the purposes of expediency, we develop notation in the general $SU(n)$ context here even though a treatment of the $SU(2)$ case would suffice for this section.

Let $P_n = X \times SU(n)$ be the trivial(ized) principal $SU(n)$ bundle over X, and denote by $\mathscr{A}_n$ and $\mathscr{G}_n$ the space of connections and gauge transformations on P_n. Since P_n is trivialized, any smooth $SU(n)$ connection can be described as a differential operator $d_A : C^\infty(X, \mathbb{C}^n) \to C^\infty(T^*X \otimes \mathbb{C}^n)$ of the form $d_A = d + A$, where d denotes the de Rham exterior derivative and $A \in \Gamma(T^*X \otimes su(n))$, i.e. A is a 1-form with values in the Lie algebra $su(n)$ (where we view an $su(n)$ element as an endomorphism of $\mathbb{C}^n$). The group $\mathscr{G}_n$ is $C^\infty(X, SU(n))$, the set of functions from X to $SU(n)$. There is a group action of $\mathscr{G}_n$ on $\mathscr{A}_n$ defined by

$$d_{g\cdot A} = g \circ d_A \circ g^{-1}.$$

A connection is called *irreducible* if its stabilizer is just the center of $\mathscr{G}_n$ (which consists precisely of constant functions on X into the center of $SU(n)$). If a connection has a larger stabilizer, it is called *reducible*.

For analytical reasons, one must use the completions of the space of connections and gauge transformations with respect to certain Sobolev norms: $\mathscr{A}_n$ is the Banach manifold obtained by completing the space of smooth connections on P_n in the L^2_1 norm, and $\mathscr{G}_n$ is the Banach Lie group obtained by completing the group of smooth bundle automorphisms of P_n in the L^2_2 norm. We will not stress this point here.

We adopt the usual notational convention $d_A \leftrightarrow A$ which identifies connections with their associated connection 1-form and denote by θ the canonical trivial connection (i.e. $d \leftrightarrow \theta$). In this notation the $\mathscr{G}_n$ action on 1-forms is written

$$g \cdot A = gAg^{-1} - (dg)g^{-1}.$$

We say A and $g \cdot A$ are *gauge equivalent*.

Lett $\Omega(X, x_0)$ denote the monoid of piecewise smooth loops in X based at x_0. Associated to each smooth connection A is its *holonomy representation*

$$hol_A \colon \Omega(X, x_0) \longrightarrow SU(n)$$

obtained from parallel translation with respect to A. The connection is called *flat* if its curvature 2-form $F(A) = dA + A \wedge A$ vanishes. By an elliptic regularity argument, flat connections are (up to gauge) smooth, and for a flat smooth connection the holonomy depends only on the homotopy class of a loop. Thus the holonomy representation of a flat connection A factors through the fundamental group $\pi_1(X)$. Gauge equivalent flat connections have conjugate holonomy representations. This establishes a homeomorphism between the space of gauge equivalence classes of flat $SU(n)$ connections on X and conjugacy classes of $SU(n)$ representations of $\pi_1 X$. This correspondence preserves the notions of reducibility.

The Chern-Simons function $cs \colon \mathscr{A}_n \longrightarrow \mathbb{R}$ is given by the formula (viewing a connection as a matrix valued 1-form):

$$cs(A) = \frac{1}{8\pi^2} \int_X tr\left(A \wedge dA + \tfrac{2}{3} A \wedge A \wedge A\right).$$

This is a smooth function with critical point set equal to the set of flat connections in $\mathscr{A}_n$. As we next explain, this suggests that the Casson invariant, which is an algebraic count of flat connections up to gauge, is the analogue of the Euler characteristic (for the quotient $\mathscr{A}_n/\mathscr{G}_n$), assuming that cs satisfies a Morse condition adapted to this equivariant, infinite dimensional setting.

To explain this properly, we first recall the analogous results in finite dimensions, namely the Poincaré-Hopf theorem. Suppose M is a smooth compact manifold and $f \colon M \to \mathbb{R}$ a smooth function. A point $p \in M$ is called a *critical point* of f if $df_p = 0$.

The *Hessian* of f at a critical point p is the symmetric bilinear form

$$\operatorname{Hess} f_p : T_pM \times T_pM \to \mathbb{R}$$

defined by $\operatorname{Hess} f_p(X, Y)) = \tilde{X}(\tilde{Y}(f))$, where $\tilde{X}$ and $\tilde{Y}$ are vector fields which extend X and Y, and vector fields act on functions in the usual way, i.e. $X(f) = df(X)$. A function is called *Morse* if its Hessian is nondegenerate at each critical point.

If we endow M with a Riemannian metric, the gradient vector field of f is defined by the condition

$$X(f) = \langle \operatorname{grad} f_p, X \rangle_p$$

for all tangent vectors $X \in T_pM$. In other words $\operatorname{grad} f$ is dual to the differential df with respect to the Riemannian metric. A point $p \in M$ is a critical point of f if $\operatorname{grad} f_p = 0$, and the function f is Morse if $\operatorname{grad} f$ is transverse to the zero section.

At a critical point $p \in M$ of f, the *Morse index* $\mu(p)$ is defined to be the dimension of the negative eigenspace of the Hessian of f at p. Then the Poincaré-Hopf theorem states that the number of critical points, counted with sign determined by the Morse index, equals the Euler characteristic:

$$\sum_{p \in \operatorname{Crit}(f)} (-1)^{\mu(p)} = \chi(M). \tag{2.1}$$

There is a very useful way to describe the difference in Morse indices at critical points which will be used to extend the discussion to the infinite dimensional case. In general the Hessian of f is only defined at a critical point. The following construction provides a suitable extension of the Hessian to all points in M.

Let $\nabla : C^\infty(TM) \times C^\infty(TM) \to C^\infty(TM)$ denote the Levi-Civita connection on M. Then for *any* point $p \in M$, the linear map

$$H_p : T_pM \to T_pM \tag{2.2}$$

defined by $H_p(X) = \nabla_X(\operatorname{grad} f)_p$ satisfies

$$\langle H_p(X), Y\rangle = \langle X, H_p(Y)\rangle$$

for $X, Y \in T_pM$, hence is self-adjoint with respect to the Riemannian metric. Moreover, at a critical point p, H_p is related to the Hessian by the equation

$$\operatorname{Hess}(f)_p(X, Y) = \langle H_p(X), Y\rangle_p.$$

Thus, if p_0 and p_1 are two critical points of f, the difference $\mu(p_1) - \mu(p_0)$ equals the spectral flow of the family $-H_{p_t}$ for any path p_t in M from p_0 to p_1. (Given a continuous path $D_t,\ t \in [0,1]$ of self-adjoint matrices, or more generally of self-adjoint Fredholm operators, the spectral flow of the path $SF(D_t)$ is the difference between the number of eigenvalues that change from negative to non-negative minus the number that change from non-negative to negative for the path of operators D_t. This is the so-called $(-\varepsilon, -\varepsilon)$ *convention*. Spectral flow is a homotopy invariant of the path of operators, rel endpoints; in addition, with this convention, spectral flow is additive under composition of paths. A careful construction of the spectral flow can be found in [**11**].)

Thus the symmetric linear map H_p extends the notion of the Hessian to all points in M. We remark that the preceding discussion applies with no change to circle valued functions $f : M \to S^1$.

We now explain how Taubes adapted these ideas to the infinite dimensional setting of the Chern-Simons function $cs : \mathscr{A}_n \to \mathbb{R}$. The infinite dimensional manifold $\mathscr{A}_n$ is an affine space modeled on $\Omega^1_X \otimes su(n)$. We choose a Riemannian metric on X, and then this space of forms inherits an L^2 inner product defined by

$$\langle a, b\rangle_{L^2} = -\int_X tr(a \wedge *b)$$

for $a, b \in T_A\mathscr{A}_n = \Omega^1_X \otimes su(n)$, where $* : \Omega^p \to \Omega^{3-p}$ denotes the Hodge star operator. This gives an inner product on the tangent space $T_A\mathscr{A}_n = L^2_1(\Omega^1_X \otimes su(n))$ for each $A \in \mathscr{A}_n$. Alternatively, we can view it as a Riemannian metric on $\mathscr{A}_n$. Computing the derivative of the Chern-Simons function, we see that if $A \in \mathscr{A}_n$ and $a \in \Omega^1(X; su(n))$, then

$$\begin{aligned}
\left.\frac{d}{dt}\right|_{t=0} cs(A + ta) &= \frac{1}{8\pi^2}\int_X tr(A \wedge da + a \wedge dA + 2a \wedge A \wedge A) \\
&= \frac{1}{8\pi^2}\int_X tr(-dA \wedge a + a \wedge dA + 2a \wedge A \wedge A) \\
&= \frac{1}{4\pi^2}\int_X tr(a \wedge (dA + A \wedge A)) \\
&= -\frac{1}{4\pi^2}\langle a, *F(A)\rangle_{L^2},
\end{aligned}$$

where $F(A) = dA + A \wedge A \in \Omega^2 \otimes su(n)$ denotes the curvature of the connection A. Hence, the gradient of the Chern-Simons function is given by

$$\operatorname{grad} cs_A = -\frac{1}{4\pi^2} * F(A),$$

and the critical set of the Chern-Simons function consists of all flat connections, namely all $A \in \mathscr{A}_n$ for which $F(A) \equiv 0$.

In order to extend the constructions described above in the finite-dimensional case to this new setting, one must deal with several issues. First, the critical set of the Chern-Simons function is infinite dimensional. This is because $F(g \cdot A) = ad_g(F(A))$. Hence if A is a critical point so is $g \cdot A$ for any g in the infinite dimensional group $\mathscr{G}_n$.

This difficulty is resolved by passing to the quotient

$$\mathscr{B}_n := \mathscr{A}_n / \mathscr{G}_n.$$

As is typical in the context of group actions, one describes the tangent space of an orbit space using the slice theorem. In the present context, the tangent space to $\mathscr{G}_n$ at the identity is just $\Omega^0_X \otimes su(n)$. For fixed $A \in \mathscr{A}_n$, the differential of the action map

$$\mathscr{G}_n \to \mathscr{A}_n, \; g \mapsto g \cdot A$$

at the identity is given by the covariant derivative:

$$d_A : \Omega^0_X \otimes su(n) = T_1\mathscr{G}_n \to \Omega^1_X \otimes su(n) = T_A\mathscr{A}_n.$$

The action of $\mathscr{G}_n$ is not free, but as long as one restricts to the open subset $\mathscr{A}^*_n$ of irreducible connections, $\mathscr{G}_n$ acts freely (up to its center) and so $\mathscr{B}^*_n = \mathscr{A}^*_n / \mathscr{G}_n$ is a (Banach) manifold. The slice theorem then identifies the tangent space of $\mathscr{B}^*_n$ at $[A]$ with the cokernel of $d_A : \Omega^0_X \otimes su(n) \to \Omega^1_X \otimes su(n)$. The L^2 inner product then identifies this cokernel with the kernel of $d^*_A : \Omega^1_X \otimes su(n) \to \Omega^0_X \otimes su(n)$.

The gauge invariance property of cs (invariance under the action of the identity component of $\mathscr{G}$) implies that $\operatorname{grad} cs$ is orthogonal to the orbit tangent space, and may be viewed either as a $\mathscr{G}$ equivariant vector field on $\mathscr{A}$ or as a tangent vector field on the quotient (at least on $\mathscr{B}^*$, where this makes sense).

In particular, let $\mathscr{V}_n \to \mathscr{A}^*_n$ be the subbundle of the tangent bundle $T\mathscr{A}^*_n$ whose fiber at $A \in \mathscr{A}^*_n$ is $\ker d^*_A$. Using the trivial connection on $T\mathscr{A}^*_n \to \mathscr{A}^*_n$ and the fiberwise L^2 projection onto the subbundle $\mathscr{V}_n$, we construct a connection ∇ on $T\mathscr{B}^*_n \to \mathscr{B}^*_n$ defined for vector fields $X, Y : \mathscr{B}^*_n \to T\mathscr{B}^*_n$ as follows. First lift X and Y to equivariant sections $\tilde{X}, \tilde{Y} : \mathscr{A}^*_n \to \mathscr{V}_n$. Then let $\nabla_X(Y)$ be the vector field on $\mathscr{B}^*_n$ whose lift to $\mathscr{V}_n$ at A equals $\operatorname{proj}_{\ker d^*_A}(d_X(Y))$, where d_X denotes the trivial connection in the product bundle $T\mathscr{A}^*_n \to \mathscr{A}^*_n$. It is typical to (and we will) abuse notation and blur the distinction between sections of $T\mathscr{B}^*_n \to \mathscr{B}^*_n$ and equivariant sections of $\mathscr{V}_n \to \mathscr{A}^*_n$.

Having described the tangent space to $\mathscr{B}^*_n$, we continue our discussion of the Morse theory of the Chern-Simons function. The Chern-Simons function is not invariant under the action of $\mathscr{G}_n$. In fact, $cs(g \cdot A) = cs(A) + \deg g$. (The *degree* $\deg g$ of a gauge transformation $g : X \to SU(n)$ is defined by the equation $g_*([X]) = \deg g \cdot [SU(2)]$, where $[X] \in H_3(X; \mathbb{Z})$ denotes the fundamental class of X, and $[SU(2)] \in H_3(SU(n); \mathbb{Z}) \cong \mathbb{Z}$ denotes the class represented by the inclusion $SU(2) \subset SU(n)$.) Since $\deg \colon \pi_0\mathscr{G}_n \to \mathbb{Z}$ is an isomorphism, the Chern-Simons function descends to a circle valued function:

$$cs : \mathscr{B}_n \to \mathbb{R}/\mathbb{Z}.$$

The gradient of the Chern-Simons function on $\mathscr{B}_n^*$ (or rather, its lift to $\mathscr{V}_n$) again satisfies $\operatorname{grad} cs_A = -\frac{1}{4\pi^2} * F(A)$. Thus the critical set is the set of flat connections modulo gauge transformations, a space homeomorphic (via the holonomy map) to $R(X, SU(n))$, the finite dimensional real-analytic variety of conjugation classes of $SU(n)$ representations of $\pi_1 X$.

The second issue concerns the Morse index $\mu(p)$. Mimicking the construction in the finite dimensional case, we define the Hessian, or rather its substitute H (as in equation (2.2)), using the connection ∇ on $T\mathscr{B}_n^*$. (For convenience, we adjust by a factor of $4\pi^2$.) Thus we define

$$H_A : T_A\mathscr{B}_n^* \to T_A\mathscr{B}_n^*, \ b \mapsto 4\pi^2 \nabla_b(\operatorname{grad} cs)_A,$$

which we compute using the definition of ∇:

$$\begin{aligned} H_A(b) &= 4\pi^2 \nabla_b(\operatorname{grad} cs)_A \\ &= -\operatorname{proj}_{\ker d_A^*} \left(\tfrac{d}{dt}\big|_{t=0} * F(A + tb) \right) \\ &= -\operatorname{proj}_{\ker d_A^*} (*d_A(b)). \end{aligned} \tag{2.3}$$

The operator H_A is closed, unbounded, and self-adjoint. It has infinitely many positive and negative eigenvalues and so the traditional definition of the Morse index $\mu(A)$ at a critical (i.e. flat) connection does not make sense. Taubes overcame this obstacle by using the spectral flow in place of the Morse index (as we described above in the finite dimensional context). This gives a relative sign to all critical points, and, assuming cs is Morse, determines an invariant up to an overall sign. Specifying the sign then amounts to deciding on a basepoint from which to measure spectral flow. We will return to the basepoint question shortly.

An interesting wrinkle that appears is that the spectral flow of $-H_{A_t}$ changes by a multiple of $4n$ if we replace A_0 or A_1 by gauge equivalent connections, but in any case the parity of $SF(-H_{A_t})$ is well defined, and this is enough to make sense of equation (2.1) in this context assuming the Chern-Simons function is Morse. We will sometimes be sloppy and write $SF(A_0, A_1)$ for $SF(-H_{A_t})$.

If $cs : \mathscr{B}_n^* \to \mathbb{R}/\mathbb{Z}$ is not Morse, we wish to perturb it by adding a gauge invariant function $h \colon \mathscr{A}_n \to \mathbb{R}$ so that $cs + h$ has only isolated nondegenerate critical points. For the remainder of the section, we restrict to the case $n = 2$. Taubes and Floer described a class of functions $\mathscr{F}$ satisfying:

(i) $\operatorname{grad}(cs + h)$ is a compact perturbation of $\operatorname{grad}(cs)$ for all $h \in \mathscr{F}$, and
(ii) $cs + h$ is a Morse function for generic $h \in \mathscr{F}$.

Let $\mathscr{M}_h^* \subset \mathscr{B}_2^*$ denote the irreducible gauge orbits in the critical set of the perturbed Chern-Simons function $cs + h$.

The main result in [**23**] states that for a generic small perturbation h,

$$\lambda_{SU(2)}(X) = \tfrac{1}{2} \sum_{[A] \in \mathscr{M}_h^*} (-1)^{SF(\theta, A)}. \tag{2.4}$$

This is the infinite dimensional analogue of the Poincaré-Hopf Theorem (2.1) for the manifold $\mathscr{B}_2^*$ and the Chern-Simons function.

The right hand side of equation (2.4) has the virtue that it gives an invariant of homology 3-spheres without making use of a Heegaard splitting. Indeed, Taubes's proof that the right hand side of equation (2.4) is well defined avoids a sticky issue that comes up in Casson's original proof, namely stablilization of Heegaard diagrams. Thus in trying to define an SU(3) version of Casson's invariant one is motivated to generalize Taubes's construction, rather than Casson's, and this is what has been done in the articles [**5, 8, 13**].

We end this section with a more detailed discussion of exactly what is meant by "$SF(\theta, A),\ [A] \in \mathscr{M}_h$" in equation (2.4). There are two issues: The first is that we described H_A above as the Hessian of the Chern-Simons function, but one must extend this to the perturbed Chern-Simons function $cs+h$. The second issue is that the base point θ (the trivial connection) is not irreducible and therefore its gauge orbit $[\theta]$ is a singular point of $\mathscr{B}_2$. So H_θ is not well-defined. We will replace the self-adjoint operator H_A with an operator that makes sense over reducible connections as well as irreducibles.

We first describe the space $\mathscr{F}$ of admissible perturbations. This space consists of gauge invariant functions $h : \mathscr{A}_2 \to \mathbb{R}$ constructed as follows. Fix a collection of thickened loops $\{\gamma_i \colon S^1 \times D^2 \to X \mid i = 1, \ldots, n\}$ whose cores generate $\pi_1 X$ and with a common normal disk at one point, i.e. $\gamma_i(s_0, x) = \gamma_j(s_0, x)$ for all i, j. For any connection A, denote by $hol_i(x, A)$ the holonomy of A around the ith loop $\gamma_i(S^1 \times \{x\})$ from some fixed basepoint. Choose $\eta(x)$ to be a bump function on the 2-disk and define the space $\mathscr{F}$ of admissible perturbations $h \colon \mathscr{A}_2 \to \mathbb{R}$ to be functions of the form

$$h(A) = \int_{D_2} f(hol_1(x, A), \ldots, hol_n(x, A))\eta(x) d^2 x,$$

where $f \colon SU(2)^n \to \mathbb{R}$ is a C^3 function invariant under the adjoint action. For example, for every word $\omega = \omega(g_1, \ldots, g_n)$ in the free group, one obtains an admissible perturbation by taking $f = tr(\omega)$.

The gradient $\operatorname{grad}(cs+h)_A = -\frac{1}{4\pi^2} * F(A) + \operatorname{grad} h_A$ can be viewed either as an equivariant vector field on $\mathscr{A}^*$ or as a vector field on $\mathscr{B}^*$. A connection A is called *h-perturbed flat* if $[A]$ is a critical point of $cs+h$ (i.e. if $\big(*F(A) - 4\pi^2 \operatorname{grad}(h)_A\big) = 0$).

Arguing as before we define the self-adjoint operator $H_{A,h} : \ker d_A^* \to \ker d_A^*$ for any A and h by

$$H_{A,h}(b) = 4\pi^2 \nabla_b(\operatorname{grad}(cs+h)) = \operatorname{proj}_{\ker d_A^*}\left(- * d_A(b) + 4\pi^2 d \operatorname{grad}(h)_A(b)\right).$$

Abusing notation slightly, write

$$\text{Hess } h_A(b) = d \operatorname{grad}(h)_A(b).$$

If A_0, A_1 are two h-perturbed flat irreducible connections, $SF(A_0, A_1)$ means the spectral flow of the path of operators $-H_{A_t,h}$, where A_t is a path of irreducible connections joining A_0 to A_1.

In order to simplify and extend the definition of $SF(A_0, A_1)$ to reducible connections, Taubes introduced the following trick. For notational convenience write Ω^i for $\Omega^i_X \otimes su(2)$. Given an arbitrary connection $A \in \mathscr{A}_2$ and an admissible perturbation h, write

$$d_{A,h} = d_A - *4\pi^2(\text{Hess } h)_A : \Omega^1 \to \Omega^2$$

and consider the operator

$$K_{A,h} : \Omega^0 \oplus \Omega^1 \to \Omega^0 \oplus \Omega^1$$

defined by $K_{A,h}(\xi,a) = (d_A^* a, d_A\xi + *d_{A,h}(a))$. Then $K_{A,h}$ is self-adjoint (with respect to the L^2 inner product). Using the Hodge decomposition

$$\Omega^0 \oplus \Omega^1 = \Omega^0 \oplus \text{image } d_A \oplus \ker d_A^*,$$

one can check that if A is an irreducible connection, the difference

$$K_{A,h} - \begin{pmatrix} 0 & d_A^* & 0 \\ d_A & 0 & 0 \\ 0 & 0 & -H_{A,h} \end{pmatrix}$$

is a compact operator. (The compactness properties of the perturbation enter at this point.)

Moreover, if A is h-perturbed flat then the two maps $d_A \circ d_{A,h} : \Omega^1 \to \Omega^3$ and $d_{A,h} \circ d_A : \Omega^0 \to \Omega^2$ are both zero. It follows that for A an h-perturbed flat connection,

$$K_{A,h} = \begin{pmatrix} 0 & d_A^* & 0 \\ d_A & 0 & 0 \\ 0 & 0 & -H_{A,h} \end{pmatrix}.$$

Thus at an h-perturbed flat connection A, the $K_{A,h}$ is obtained from $-H_{A,h}$ by adding an operator with symmetric spectrum.

These facts, together with the invariance of spectral flow with respect to homotopy of paths of self-adjoint operators rel endpoints, imply that if A_0, A_1 are any two h-perturbed irreducible flat connections, and A_t is any path of irreducible connections joining them, the spectral flow of the family $K_{A_t,h}$ equals the spectral flow of the family $-H_{A_t,h}$.

The advantage is that $K_{A,h}$ is a smooth family of self-adjoint operators *for any perturbation h and any connection A*. Thus we can define the spectral flow $SF((A_0,h_0),(A_1,h_1))$ for any pair of connections and admissible perturbations to be the spectral flow of the family K_{A_t,h_t} for any path A_t joining A_0 to A_1 and any path h_t joining h_0 to h_1. This is well defined since the space of connections and the space of perturbations is contractible. It agrees with the previous definition of $SF(A_0,A_1)$ in terms of the Hessian $H_{A,h}$ when A_0 and A_1 are irreducible h-perturbed flat connections for a fixed perturbation h.

If A is an h-perturbed flat connection, we use the shorthand $SF(\theta,A)$ for $SF((\theta,0),(A,h))$. This, finally, is the meaning of this term in equation (2.4). We call the operator $K_{A,h}$ the *perturbed odd signature operator*, since when A is a flat connection, $K_{A,0}$ is exactly the odd signature operator coupled to A as described in [**3**]. The reader will note that this discussion applies equally well to $SU(n)$ connections. In this general case we have $SF(A_0, g\cdot A_1) = SF(A_0,A_1) + 4n\ \deg g$.

3. Reducible connections, singularities and bifurcations

As explained in the previous section, Taubes interpreted the Casson invariant as an infinite dimensional Morse theoretic Euler characteristic for the space $\mathscr{B}_2^*$ of $SU(2)$ connections modulo gauge. This quotient space is singular along the reducible connections, but since X is a homology 3-sphere the critical set of the function cs avoids the singularities (except for the orbit of the trivial connection, which is isolated from other critical orbits), and so the singularities do not cause any serious difficulties.

In the setting of $SU(3)$ gauge theory, a reducible flat connection need not be gauge equivalent to the trivial connection, even on a homology 3-sphere. Restricting a gauge transformation to a fiber of the principal bundle determines an isomorphism from the stabilizer subgroup of a connection A to a subgroup of $SU(3)$. Up to conjugation, there are five possibilities for this subgroup:

(1) The center $\mathbb{Z}_3$ (so A is irreducible),
(2) The 1-dimensional torus $\{\mathrm{diag}(\lambda, \lambda, \lambda^{-2}) \mid \lambda \in U(1)\}$,
(3) The maximal torus $\{\mathrm{diag}(\lambda_1, \lambda_2, (\lambda_1\lambda_2)^{-1}) \mid \lambda_i \in U(1)\}$,
(4) $S(U(2) \times U(1)) = \left\{ \begin{pmatrix} A & 0 \\ 0 & \det A^{-1} \end{pmatrix} \middle| A \in U(2) \right\}$, and
(5) $SU(3)$.

Since $H_1(X;\mathbb{Z}) = 0$, all abelian representations of $\pi_1 X$ are trivial, so the third and fourth possibilities do not occur as stabilizers of representations (or of flat connections). The three types of flat connections on a homology 3-sphere, then, are the following:

(1) Irreducible connections A, with $\mathrm{stab}(A) = \mathbb{Z}_3$.
(2) Connections A with $\mathrm{stab}(A)$ equal to a 1-dimensional torus. For these connections, the image of the holonomy representation $hol_A \colon \pi_1 X \to SU(3)$ lies in $SU(2) \subset S(U(2) \times U(1))$ up to conjugation.
(3) Connections which are gauge equivalent to the trivial connection. For these connections, $\mathrm{stab}(A) \cong SU(3)$.

The presence of $S(U(2) \times U(1))$ flat connections means that, unlike the $SU(2)$ case, the $SU(3)$ situation requires analysis of reducible connections to define a perturbation independent quantity, because the singular stratum is not disjoint from the critical set of the Chern-Simons function. (This is similar to what Walker faced in extending the $SU(2)$ invariant to rational homology 3-spheres; in that case there were also three strata to consider.) As in the $SU(2)$ setting, the gauge orbit of the trivial connection is isolated in the space of flat connections modulo gauge. Thus the difficulties are focused on the singular stratum in the quotient $\mathscr{B}_3 = \mathscr{A}_3/\mathscr{G}_3$ consisting of connections whose stabilizer is a 1-dimensional torus. Notice that the reducible flat $SU(3)$ connections may be viewed as irreducible flat $SU(2)$ connections, i.e. those considered by Taubes.

The irreducible portion $\mathscr{B}_3^* \subset \mathscr{B}_3$ is a manifold. In addition, the stratum

$$\mathscr{B}_3^r = \{A \in \mathscr{A}_3 \mid \mathrm{stab}(A) \cong U(1)\}/\mathscr{G}_3$$

is also smooth (and in fact has a cone bundle neighborhood in $\mathscr{B}_3$). The strata of $\mathscr{B}_3$ corresponding to the third and fourth orbit type listed above are disjoint from the critical set of of the Chern-Simons function, so we will not dwell on them.

Again, $\mathrm{grad}\, cs$ and $\mathrm{grad}(cs + h)$ may be viewed as either equivariant vector fields on $\mathscr{A}_3$ or as tangent vector fields on each stratum. (Invariance implies that at a reducible connection A, the gradient is tangent to the space of reducible connections.) Before describing the $SU(3)$ invariants, we illustrate the difficulties in defining an invariant by counting critical points (zeros of a gradient vector field) in such a stratified context.

For the purpose of illustration, we consider a finite dimensional manifold M with a semifree $U(1)$ action. Recall that an action of a group G on a space M is called *semifree* if the stabilizer subgroup of every point in M is either the trivial subgroup or G.

Let $L \subset M$ denote the set of fixed points of the group action. Like $\mathscr{B}^* \cup \mathscr{B}^r$, the quotient space $M/U(1)$ consists of two strata, $L/U(1) \cong L$ and $(M - L)/U(1)$, and the fixed point stratum has a normal bundle similar to that of $\mathscr{B}^r$ in $\mathscr{B}$. Namely, the normal bundle of $L/U(1)$ has fiber homeomorphic to $c(\mathbb{C}P^k)$, whereas the normal bundle of $\mathscr{B}^r$ has fiber $c(\mathbb{C}P^\infty)$. In other words, in this illustration, the non-fixed points play the role of the irreducible $SU(3)$ connections, and the fixed points play the role of the (nontrivial) reducible connections.

Let $f\colon M \to \mathbb{R}$ be a $U(1)$ invariant smooth function. If $m \in M$ is a critical point for f, then so is every point in the orbit of m. Thus the critical set of f consists of a collection of fixed points together with a union of circle orbits of critical points in $M - L$. Fix an invariant Riemannian metric on M so that we can identify the critical set of f with the zero set of the $U(1)$-equivariant gradient vector field $\operatorname{grad} f$. Because $U(1)$ acts with no nonzero fixed vectors on the normal bundle to L, if $p \in L$ then $\operatorname{grad} f(p) \in T_pL$. It follows that the critical fixed points are exactly the critical points of $f|_L$.

On a compact manifold M without group actions, standard techniques in differential topology show that generic functions are Morse (i.e. have gradient vector fields transverse to the zero section of TM). This nondegeneracy condition implies that the critical set of the function (the set of zeros of the vector field) is compact. Moreover, for a generic path of functions connecting two Morse functions, the only topological changes in the critical set are births and deaths of pairs of critical points of Morse indices differing by one. This fact can be used to give a Cerf theoretic proof that the signed count of critical points is independent of the Morse function, for the algebraic contribution from such a pair will be zero. (In fact, as mentioned in Section 2, the signed count can be identified with the Euler characteristic, so there are other ways of seeing that it is an invariant, too.)

Things are more complicated in the equivariant setting. We may view $M/U(1)$ as a union of two manifolds, $(M-L)/U(1)$ and $L/U(1) \cong L$. The singular nature of the quotient space has to do with the normal bundle structure of the second in the first. A theorem of Wasserman [**25**] asserts in this setting that a generic invariant function on M will give a Morse function with finite critical set on each of these strata. Since $(M-L)/U(1)$ is not compact, finiteness here is less straightforward. But Wasserman's theorem is stronger than the statement that the function restricted to each stratum is Morse. In addition, at a critical point $p \in L$, $\operatorname{Hess}(f)_p$ is also nondegenerate on the normal bundle fiber to N_pL. This prevents critical orbits in $M - L$ from accumulating at p.

For a generic path of invariant functions on M connecting two such generic functions, three types of bifurcations can occur in the critical set. The first two correspond to standard Morse births and deaths of cancelling pairs of critical points, in L, and in the (smooth) quotient space $(M-L)/U(1)$. In other words, a pair of critical points in L of Morse index difference one can be born or can annihilate one another, and similarly for a pair of circle orbits of critical points of index difference one.

The third type of bifurcation involves a critical circle orbit popping out of a critical fixed point (or this process in reverse). That is to say, there is a critical point in L, with no other critical points nearby for $t \le t_0$, and then for $t > t_0$ there is in addition a circle orbit of critical points nearby. Figure 1 illustrates the topology of the critical set (in a neighborhood nearby). This phenomenon is illustrated with the path of functions $f_t\colon \mathbb{C} \to \mathbb{R}$, $f_t(z) = -t|z|^2 + |z|^4$.

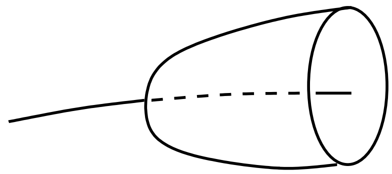

Figure 1. A picture of the parameterized critical set near t_0. For $t \leq t_0$, the critical set consists of only one fixed point, but for $t > t_0$, it also includes a circle orbit.

As a consequence, for a generic path f_t of functions joining two generic functions, the *parameterized critical set*

$$\{([m], t) \in M/U(1) \times [0,1] \mid m \text{ is a critical point of } f_t \}$$

has a description as a compact 1-dimensional singular bordism with 'T'-intersections from the critical set of f_0 to the critical set of f_1. More precisely, it is the union of

$$C_{\text{fixed}} := \{([p], t) \in L/U(1) \times [0,1] \mid p \in L \text{ is a critical fixed point of } f_t \}$$

and

$$C_{\text{free}} := \{([m], t) \in M/U(1) \times [0,1] \mid m \in M - L \text{ is a critical point of } f_t \}.$$

C_{fixed} is a compact 1-dimensional manifold with boundary, and the boundary lies in the "ends" $L/U(1) \times \{0,1\}$. On the other hand, C_{free} is a 1-manifold (generally not compact). All of its endpoints lie in $M/U(1) \times \{0,1\}$, but its noncompact ends limit to points in the interior of C_{fixed}. An example of a possible parameterized critical set is illustrated in Figure 2.

To obtain an analogue of the Euler characteristic in this $U(1)$ manifold setting, we must modify the usual signed count of critical points (or critical orbits) to obtain a formula which is invariant under all three types of bifurcations. The crucial observation is that, simultaneous with a bifurcation of the third type described above, there is a change in the normal Morse index of the function at the critical fixed point.

To describe the change in detail, we must refine our notion of Morse index at the critical points in L. The fact that the Hessian of an invariant function at a

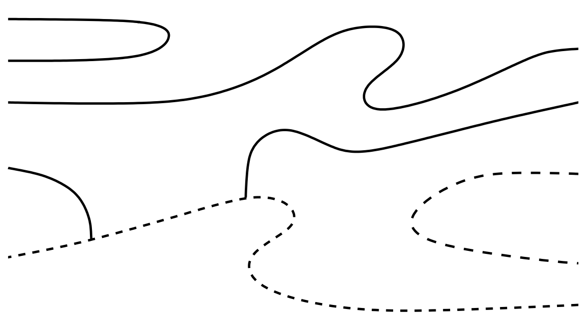

Figure 2. A cobordism with 'T'-intersections. The dashed curves represent the varying critical set in L, and the solid curves represent the varying critical set (modulo $U(1)$) in $M - L$.

critical point $p \in L$ is invariant under the (linearized) group action on the tangent space T_pM implies that the Hessian decomposes into a direct sum of operators on the two summands $T_pL \oplus N_pL \cong T_pM$. We define the tangential and normal components of the Morse index to be the number of negative eigenvalues of these summands, and denote them by $\mu_t(p)$ and $\mu_n(p)$. On N_pL, the Hessian commutes with the $U(1)$ action, so its eigenspaces are $U(1)$ invariant subspaces. The $U(1)$ action on N_pL has no trivial subrepresentations, therefore $\mu_n(p)$ is always even.

A close examination of the relationship between the Morse indices of the critical points involved in the bifurcations demonstrates that the quantity

$$\sum_{q \in ((M-L)\cap crit(f))/U(1)} (-1)^{\mu(q)} - \sum_{p \in L \cap crit(f)} (-1)^{\mu_t(p)} \left(\frac{\mu_n(p)}{2} \right) \tag{3.1}$$

is invariant under all three bifurcations and hence is independent of the invariant Morse function. In fact, one can see by choosing the function appropriately that this invariant of the $U(1)$ manifold M is the relative Euler characteristic of the quotient space relative to the singular set (i.e. the Euler characteristic of the relative cohomology).

4. Correction terms in the SU(3) setting

In this section, we consider the $SU(3)$ gauge theory setting (and so we now drop the subscript on $\mathscr{A}$). There is a classification of the types of bifurcations that occur in the critical set for a generic one parameter family of admissible perturbations of the Chern-Simons function analogous to that given in the previous section. That is to say, the bifurcations involve standard Morse births/deaths in the irreducible stratum of $\mathscr{B}$, standard Morse births/deaths in the reducible (i.e. $S(U(2)\times U(1))$) stratum of $\mathscr{B}$, and irreducible orbits of critical points popping out of reducible critical points.

Motivated by the finite dimensional model discussed in the previous section and Taubes' gauge theoretic description of the Casson invariant, it is natural to try to define an $SU(3)$ Casson invariant by replacing the Morse indices in (3.1) with the spectral flow analogues. We begin by describing the spectral flow analogue of the decomposition of Morse index $\mu(p) = \mu_t(p) + \mu_n(p)$ at critical points with circle stabilizer.

We first identify the tangent space of the space $\mathscr{A}$ of $SU(3)$ connections with $\Omega^1(X) \otimes su(3)$, the space of differential 1-forms on X with values in the Lie algebra $su(3)$. Every reducible connection is gauge equivalent to an $S(U(2)\times U(1))$ connection. Under the adjoint action of $S(U(2)\times U(1))$ on the Lie algebra $su(3)$, we have the decomposition

$$su(3) = s(u(2) \times u(1)) \oplus \mathbb{C}^2, \tag{4.1}$$

where an element $A \oplus (\det A)^{-1} \in S(U(2)\times U(1))$ acts by the adjoint action on the first factor and by matrix multiplication by the $U(2)$ matrix $\det(A)A$ on the second factor. Any reducible $SU(3)$ connection can be gauge transformed into the path connected space of $S(U(2)\times U(1))$ connections. Thus one can decompose the spectral flow of the odd signature operator K_{A_t} along a path A_t of reducible connections into $s(u(2)\times u(1))$ and $\mathbb{C}^2$ components according to the decomposition of equation (4.1). We denote the spectral flow in the $s(u(2)\times u(1))$ (resp. $\mathbb{C}^2$)

subspace suggestively by SF_t (resp. SF_n), where 't' stands for tangential and 'n' for normal.

The observation which underlies the construction of the $SU(3)$ Casson invariant is the following: *At a bifurcation where a new irreducible critical point pops out of a reducible one, an eigenvalue of multiplicity two in the $\mathbb{C}^2$ component crosses zero.* More precisely, if $(A_t, h_t)_{t\in(-\varepsilon,\varepsilon)}$ is a generic path of reducible connections such that A_t is h_t-perturbed flat, and so that a bifurcation occurs at $t = 0$, then $SF_n(A_t) = \pm 2$.

Let $\mathscr{M}_h \subset \mathscr{A}/\mathscr{G}$ be the h-perturbed flat moduli space, which consists of the gauge orbits of critical points of $cs + h$ for an admissible perturbation h. We can split $\mathscr{M}_h = \mathscr{M}_h^* \cup \mathscr{M}_h^{red} \cup \{[\theta]\}$ as a disjoint union of irreducible and reducible orbits and the trivial orbit.

Replacing the Morse indices of formula (3.1) with the corresponding spectral flows, we obtain, when $\mathscr{M}_h$ is generic,

$$\sum_{[A]\in\mathscr{M}_h^*} (-1)^{SF(\theta,A)} - \sum_{[B]\in\mathscr{M}_h^{red}} (-1)^{SF_t(\theta,B)} \left(\frac{SF_n(\theta,B)}{2}\right). \tag{4.2}$$

Equation (4.2) is the expression which generalizes to the gauge theory setting the discussion we carried out above in the finite-dimensional setting, and which presents the natural candidate for the definition of an $SU(3)$ Casson invariant following the approach of Taubes.

But there is a hitch, arising from the fact that the spectral flow is not well defined on gauge orbits. For the exponents of (-1), this is not a problem, since $SF(\theta, [A])$ is well defined modulo 12 and $SF_t(\theta, [B])$ is well defined modulo 8. But the gauge ambiguity of the normal spectral flow, which is only well defined modulo 4, makes formula (4.2) dependent on the gauge representatives chosen for the reducible perturbed flat orbits.

Several approaches have been proposed to correct this gauge ambiguity. In [**5**], the perturbation h is taken to be small and the correction term is defined by replacing $\frac{1}{2}SF_n(\theta, B)$ with $\frac{1}{2}(SF_n(\theta, B) - 4cs(\widehat{B}) + 2)$ in the second sum of (4.2). Here, $\widehat{B}$ is a genuine flat connection near B. This quantity does not depend on the choice of gauge representative B for the orbit $[B]$ because

$$SF_n(\theta, gB) - 4cs(g\widehat{B}) = SF_n(\theta, B) - 4cs(\widehat{B}).$$

This choice for the $SU(3)$ correction term is motivated by Walker's correction term for rational homology 3-spheres.

The resulting real valued invariant,

$$\lambda_{SU(3)} := \sum_{[A]\in\mathscr{M}_h^*} (-1)^{SF(\theta,A)} - \tfrac{1}{2} \sum_{[B]\in\mathscr{M}_h^{red}} (-1)^{SF_t(\theta,B)} \left(SF_n(\theta, B) - 4cs(\widehat{B}) + 2\right),$$

was shown in [**6**] to satisfy:

$$\begin{aligned} \lambda_{SU(3)}(-X) &= \lambda_{SU(3)}(X) \\ \lambda_{SU(3)}(X\#Y) &= \lambda_{SU(3)}(X) + \lambda_{SU(3)}(Y) + 16\lambda_{SU(2)}(X)\lambda_{SU(2)}(Y). \end{aligned} \tag{4.3}$$

The behavior of $\lambda_{SU(3)}$ under Dehn surgery appears to be very complicated. A rather delicate analysis was used to compute the values of $\lambda_{SU(3)}$ for surgeries on $(2, q)$-torus knots in [**7**], using analytical surgery methods to compute spectral flow. The data obtained from these calculations did not suggest any obvious surgery formula for $\lambda_{SU(3)}$. Moreover, the calculations show that $\lambda_{SU(3)}$ is not a finite-type invariant.

The value of $\lambda_{SU(3)}$ for Seifert fibered homology spheres is rational since the Chern-Simons invariant of flat connections on Seifert fibered spaces are known to be rational. But it is not generally known whether the value of $\lambda_{SU(3)}$ on homology 3-spheres is always a rational number. This is closely related to the conjectured rationality of the Chern-Simons invariants, which remains an open problem.

Cappell, Lee, and Miller presented an alternative method for correcting the gauge ambiguities and derived a new $SU(3)$ Casson invariant [**13**]. Their result is based on the observation that one can correct for the gauge ambiguities of (4.2) using the tangential spectral flow instead of the Chern-Simons invariant: the quantity $SF_n(\theta, B) - SF_t(\theta, B)/2$ is independent of the choice of gauge representative of $[B]$. Thus Cappell, Lee, and Miller solved the problem of gauge ambiguities by replacing $SF_n(\theta, B)$ in the second sum in (4.2) by $SF_n(\theta, B) - SF_t(\theta, B)/2 + 5/4$. This restores gauge invariance but destroys the independence of perturbation: one still needs to correct for births and deaths of pairs of *reducible* critical orbits. To handle this phenomenon, they add to

$$\sum_{[A]\in\mathscr{M}_h^*} (-1)^{SF(\theta,A)} - \tfrac{1}{2} \sum_{[B]\in\mathscr{M}_h^{red}} (-1)^{SF_t(\theta,B)} \left(SF_n(\theta, B) - SF_t(\theta, B)/2 + 5/4\right)$$

a third correction term involving the ranks of the boundary operators for the Floer instanton homology complex (on $S(U(2) \times U(1))$ connections). That the resulting expression is an invariant follows from the fact that Floer homology is itself independent of the perturbation, and hence a birth and death of a pair of reducibles coincides with a change (of ± 1) in the rank of some boundary operator.

The Cappell-Lee-Miller invariant λ_{CLM} takes values in $\frac{1}{4}\mathbb{Z}$ and also satisfies $\lambda_{CLM}(-X) = \lambda_{CLM}(X)$, but its behavior under connected sum is obscured by complications arising from the third correction term. These complications are closely related to the subtle problem of determining the Floer homology for connected sums.

In [**8**], the present authors developed a different method for defining a correction term, which we now describe in detail. Let $\mathscr{G}_0$ denote the identity component of $\mathscr{G}$, which consists of degree zero gauge transformations. Then $\mathscr{A}/\mathscr{G}_0 \longrightarrow \mathscr{A}/\mathscr{G}$ is a nontrivial $\mathbb{Z}$ covering and is classified by the Chern-Simons function. Thus spectral flow is well defined on $\mathscr{A}/\mathscr{G}_0$. This has the consequence that if $[A]$ and $[B]$ are close enough in $\mathscr{A}/\mathscr{G}$ then they have close lifts in $\mathscr{A}/\mathscr{G}_0$ and the spectral flow between these two lifts is well defined and independent of the lift (provided they are close). If $[A]$ and $[B]$ happen to be reducible and close, then we can also define the normal spectral flow between them and it is also independent of the choice of close lifts. In fact, for any connected subset $\mathscr{U} \subset \mathscr{A}^{red}/\mathscr{G}$ over which $\widetilde{\mathscr{U}} \longrightarrow \mathscr{U}$ is the trivial $\mathbb{Z}$ cover, we define the normal spectral flow $SF_n([A], [B])$ between any two gauge orbits $[A], [B] \in \mathscr{U}$ by simply choosing lifts of $[A]$ and $[B]$ and a path between them in the same connected component of $\widetilde{\mathscr{U}}$. For the remainder of this article, we adopt this convention for defining the normal spectral flow between reducible gauge orbits.

We apply this to the path components of $\mathscr{M}^{red}$. Let $\widetilde{\mathscr{M}}^{red} \subset \mathscr{A}/\mathscr{G}_0$ be the lift (i.e. inverse image) of the moduli space $\mathscr{M}^{red}$ of reducible flat $SU(3)$ connections. Then since

(i) $cs(g \cdot A) = cs(A) + \deg(g)$, and
(ii) $cs\colon \mathscr{A} \longrightarrow \mathbb{R}$ is constant on components of flat connections,

the cover $\widetilde{\mathscr{M}}^{red} \longrightarrow \mathscr{M}^{red}$ is trivial and so the normal spectral flow is well defined between gauge orbits lying in the same connected component of $\widetilde{\mathscr{M}}^{red}$ and, more generally, the normal spectral flow is defined between gauge orbits lying in a small enough neighborhood of a path component of $\mathscr{M}^{red}$.

The new correction term is defined in terms of $SF_n([B_i^{\pm}],[A])$ for fixed basepoints $[B_i^{\pm}]$ chosen as follows. Let X be a homology 3-sphere and index the components of the reducible flat moduli space $C_1, \ldots, C_n$. Each component C_i is compact, hence the normal spectral flow is bounded and one can choose orbits $[B_i^-]$ and $[B_i^+]$ in C_i for which the normal spectral flow $SF_n([B_i^-],[B_i^+])$ is maximal. If h is a small perturbation, then every reducible h-perturbed flat connection $[B]$ is close to one of the components C_i, and so there is an unambiguous notion of normal spectral flow from the basepoints $[B_i^{\pm}]$ to $[B]$. Although the basepoints need not be unique, the normal spectral flows are well defined since different choices $[\widehat{B}_i^{\pm}]$ satisfy

$$SF_n([B_i^-],[\widehat{B}_i^-]) = 0 = SF_n([B_i^+],[\widehat{B}_i^+]).$$

Given a small, nondegenerate perturbation h, we define $\tau_{\pm}(X)$ by

$$\tau_+(X) = \sum_{[A]} (-1)^{SF(\theta,A)} - \tfrac{1}{2} \sum_{i=1}^{n} \sum_{[B]} (-1)^{SF_t(\theta,B)} SF_n([B_i^+],[B])$$

$$\tau_-(X) = \sum_{[A]} (-1)^{SF(\theta,A)} - \tfrac{1}{2} \sum_{i=1}^{n} \sum_{[B]} (-1)^{SF_t(\theta,B)} \left(SF_n([B_i^-],[B]) + \dim H^1_{B_i^-}(X;\mathbb{C}^2) \right),$$

where the first and second sums in $\tau_{\pm}(X)$ are over $[A] \in \mathscr{M}_h^*$ and $[B] \in \mathscr{M}_h^{red} \cap \mathscr{U}_i$, and $\mathscr{U}_i \subset \mathscr{B}^{red}$ is a small open neighborhood of C_i.

Note that the normal spectral flow $SF_n([B_i^{\pm}],[B])$ is always even because it counts the (real) eigenvalues of a complex operator. This observation shows that both $\tau_+(X)$ and $\tau_-(X)$ are integer valued. Under change in orientations, one can easily verify that $\tau_+(-X) = \tau_-(X)$. (The appearance of the $\dim H^1$ term in the

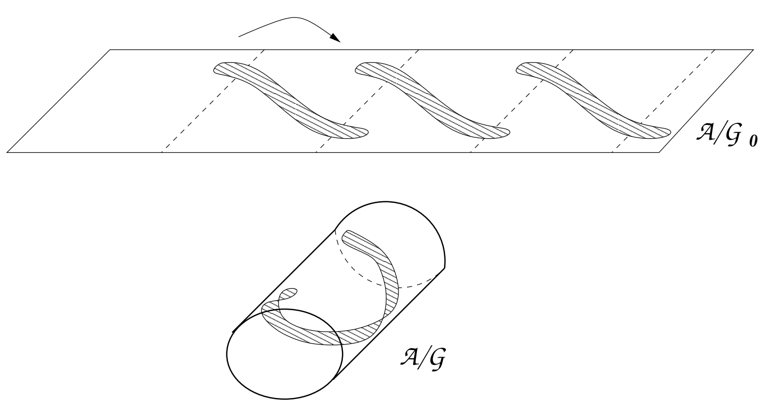

Figure 3. A set $\mathscr{U}$ with trivial cover. Given $[A],[B] \in \mathscr{U}$, compute the spectral flow by choosing lifts in the same component of $\widetilde{\mathscr{U}}$.

definition of τ_- comes from our choice of convention for spectral flow.) Hence it is natural to define an orientation-independent invariant by averaging them. Thus we set

$$\begin{aligned}\tau_{SU(3)}(X) =& \tfrac{1}{2}(\tau_+(X)+\tau_-(X)) \\ =& \sum_{[A]}(-1)^{SF(\theta,A)} - \tfrac{1}{4}\sum_{i=1}^{n}\sum_{[B]}(-1)^{SF_t(\theta,B)}\Big(SF_n([B_i^+],[B]) \\ & \qquad + SF_n([B_i^-],[B]) + \dim H^1_{B_i^-}(X;\mathbb{C}^2)\Big),\end{aligned}$$

where the sums are defined as before.

We claim that $\tau_{SU(3)}(X)$ is integer valued. To prove this claim, it is sufficient to show that, for any $B \in M_h^r$, the quantity

$$SF_n([B_i^+],[B]) + SF_n([B_i^-],[B]) + \dim H^1_{B_i^-}(X;\mathbb{C}^2)$$

is divisible by 4.

Identifying $\mathbb{C}^2$ with the quaternions and $SU(2)$ with the unit quaternions, the natural action of $SU(2)$ on $\mathbb{C}^2$ becomes left quaternionic multiplication. Since B_i^- is a flat $SU(2)$ connection, right quaternionic multiplication endows the cohomology group $H^1_{B_i^-}(X;\mathbb{C}^2)$ with a quaternionic structure, which implies that the dimension of this cohomology is a multiple of 4. Furthermore, for any $SU(2)$ connection, the twisted signature operator on $\mathbb{C}^2$ valued forms commutes with this right multiplication, so its eigenspaces are all quaternionic.

This leaves the two normal spectral flow terms. Additivity of the spectral flow under composition of paths gives that

$$SF_n([B_i^+],[B]) + SF_n([B_i^-],[B]) = 2SF_n([B_i^+],[B]) + SF_n([B_i^-],[B_i^+]).$$

The complex structure shows that $SF_n([B_i^+],[B])$ is even. Furthermore, since both B_i^+ and B_i^- are flat $SU(2)$ connections, they can be connected by a path of $SU(2)$ connections. As noted above, the corresponding path of twisted signature operators will have quaternionic eigenspaces, which implies that $SF_n([B_i^-],[B_i^+])$ is divisible by 4. This completes the proof of the integrality claim.

Thus $\tau_{SU(3)}$ is integer valued, invariant under change of orientation, and moreover Theorem 4 of [**8**] shows that

$$\tau_{SU(3)}(X\#Y) = \tau_{SU(3)}(X) + \tau_{SU(3)}(Y) + 16\lambda_{SU(2)}(X)\lambda_{SU(2)}(Y).$$

(The difference in coefficients here and in [**8**] is the result of our normalizing $\lambda_{SU(2)}(X)$ as Casson originally did here, whereas it was normalized according to Walker's convention [**24**] in [**8**].)

The invariant $\tau_{SU(3)}$ is vastly easier to compute than either $\lambda_{SU(3)}$ or λ_{CLM}. To see why, consider a homology 3-sphere X whose cohomology $H^1_\alpha(X;\mathbb{C}^2)$ vanishes for every $SU(2)$ representation $\alpha\colon X \to SU(2)$. Since the twisted signature operator at each reducible flat connection has no zero modes on $\mathbb{C}^2$ valued forms, all normal spectral flows $SF_n(B_i^\pm, B)$ vanish for all sufficiently small perturbations. This implies that the correction term for $\tau_{SU(3)}(X)$ vanishes for small perturbations, and thus $\tau_{SU(3)}(X)$ simply equals the quantity

$$\sum_{[A]\in\mathscr{M}_h^*}(-1)^{SF(\theta,A)}$$

obtained by perturbing the $SU(3)$ moduli space and counting only irreducible, perturbed flat gauge orbits. In this sense, the advantage of $\tau_{SU(3)}$ over $\lambda_{SU(3)}$ and λ_{CLM} is that the correction term for $\tau_{SU(3)}$ is nontrivial only when absolutely necessary. For example, the correction term vanishes for Brieskorn spheres of the form $\Sigma(2,p,q)$. This was used in the computations of $\tau_{SU(3)}(\Sigma(2,p,q))$ in [**8**]. In Section 5, we present computations of $\tau_{SU(3)}$ in cases where the correction term does not vanish.

5. Computations for Brieskorn spheres $\Sigma(p,q,r)$

In this section, we outline a method for computing $\tau_{SU(3)}$ on Brieskorn homology 3-spheres [**9**].

5.1. SU(3) representations of Brieskorn 3-spheres. Fix pairwise relatively prime numbers p,q,r and consider the Brieskorn homology 3-sphere

$$\Sigma(p,q,r) = \{(x,y,z) \in \mathbb{C}^3 \mid x^p + y^q + z^r = 0\} \cap S^5_\epsilon.$$

Taking a,b,c with

$$aqr + bpr + cpq = 1,$$

the fundamental group of $\Sigma(p,q,r)$ has a presentation

$$\pi_1(\Sigma(p,q,r)) = \langle x,y,z,h \mid h \text{ central}, \, x^p = h^a, y^q = h^b, z^r = h^c, \, xyz = 1\rangle.$$

The next result is not difficult to prove; details can be found in [**9**].

Proposition 5.1. *If $\alpha\colon \pi_1\Sigma(p,q,r) \to SU(3)$ is irreducible, then $\alpha(h) = e^{2\pi ik/3}I$ is central. If $\alpha\colon \pi_1\Sigma(p,q,r) \to SU(3)$ is reducible, then up to conjugation* $\operatorname{image}(\alpha) \subset SU(2)\times\{1\}$ *and*

$$\alpha(h) = \begin{bmatrix} \pm 1 & 0 & 0 \\ 0 & \pm 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Path components of $R(\Sigma(p,q,r), SU(3))$ are indexed by fixing $\alpha(h)$ and choosing $\alpha(x), \alpha(y)$ and $\alpha(z)$ to be p-th, q-th and r-th roots of $\alpha(h^a)$, $\alpha(h^b)$ and $\alpha(h^c)$, respectively.

Using the Seifert fibration $\Sigma(p,q,r) \to S^2$, one can interpret $R(\Sigma(p,q,r), SU(3))$ in terms of moduli spaces of parabolic bundles over S^2 with 3 marked points. This

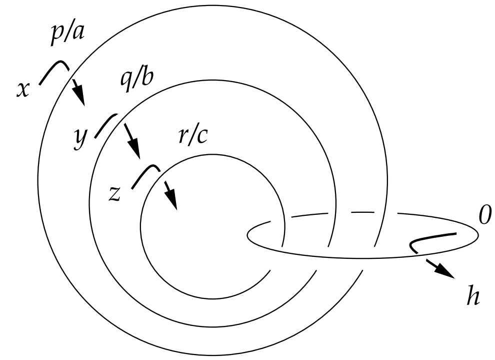

Figure 4. The Brieskorn sphere $\Sigma(p,q,r)$ as surgery on a link.

Figure 5. Components of types III and IV in $R(\Sigma, SU(3))$.

allows us to make use of the description of these moduli spaces given in [**4**] and [**19**]. In these articles it is shown that each path component of $R(\Sigma(p,q,r), SU(3))$ is homeomorphic to either an isolated point or a 2-sphere. These can be further broken down into 4 types:

(1) *Type I* consist of isolated irreducible $SU(3)$ representations,
(2) *Type II* are isolated reducible $SU(3)$ representations,
(3) *Type III* are smooth 2-spheres consisting entirely of irreducible $SU(3)$ representations, and
(4) *Type IV* are 2-spheres which have one non-smooth point: the smooth points are irreducible $SU(3)$ representations and the non-smooth point is a reducible representation. These components cause the most trouble: we call them *pointed 2-spheres*.

When $p = 2$, there are no 2-sphere components. In this case, $\tau_{SU(3)}(\Sigma(p,q,r))$ is simply a count of the Type I points; the Type II points do not contribute to $\tau_{SU(3)}$ since $H^1_\alpha(\Sigma; \mathbb{C}^2) = 0$ at these points (see the last paragraph of Section 4). The count of the Type I points is carried out in [**4**], where it is shown that each Type I point contributes positively to $\tau_{SU(3)}$. Table 1 gives some calculations of $\tau_{SU(3)}(\Sigma(2,q,r))$ for various q and r.

Σ	$\tau_{SU(3)}(\Sigma)$
$\Sigma(2,3,6k \pm 1)$	$3k^2 \pm k$
$\Sigma(2,5,10k \pm 1)$	$33k^2 \pm 9k$
$\Sigma(2,5,10k \pm 3)$	$33k^2 \pm 19k + 2$
$\Sigma(2,7,14k \pm 1)$	$138k^2 \pm 26k$
$\Sigma(2,7,14k \pm 3)$	$138k^2 \pm 62k + 4$
$\Sigma(2,7,14k \pm 5)$	$138k^2 \pm 102k + 16$
$\Sigma(2,9,18k \pm 1)$	$390k^2 \pm 58k$
$\Sigma(2,9,18k \pm 5)$	$390k^2 \pm 210k + 24$
$\Sigma(2,9,18k \pm 7)$	$390k^2 \pm 298k + 52$

Table 1. Calculations of the integer valued SU(3) Casson invariant for Brieskorn spheres $\Sigma(2,q,r)$.

5.2. Twisting perturbations. If $p,q,r > 2$, then $R(\Sigma(p,q,r), SU(3))$ always contains pointed 2-spheres. Thus perturbations are needed to resolve them. In [**9**] we use a preliminary step to improve the moduli space $\mathscr{M}$: we show that a certain class of *twisting* perturbations resolve each Type IV component into the union of one Type II point and one Type III smooth 2-sphere. Rather than to introduce the twisting perturbations explicitly here, we describe their effect.

The starting point is that pertubations alter the flatness equations in the neighborhood of a knot. A perturbation function $h\colon \mathscr{A} \to \mathbb{R}$ takes a connection A to a number determined by the holonomies of the A along curves $S^1 \times \{z\}$ in a solid torus $S^1 \times D^2$ contained in the homology 3-sphere. Floer proved that for such a perturbation h, an h-perturbed flat connection is in fact flat outside the solid torus [**16**].

Let $Y = S^1 \times D^2$ be a neighborhood of the singular r-fiber in $\Sigma = \Sigma(p,q,r)$ and set $Z = \Sigma - Y$. Then

$$\Sigma = Y \cup_T Z.$$

We perturb the flatness equations in the solid torus Y and study the effect on a pointed 2-sphere. To better visualize this situation, we regard $\mathscr{M}(\Sigma)$ and $\mathscr{M}_h(\Sigma)$ as the pullback of an intersection in the 4-dimensional representation space of the torus

$$\mathscr{M}(T) = T^2 \times T^2/S_3,$$

where S_3 acts diagonally. Let h denote the twisting perturbation and let h_t be a path from 0 to h. Consider the diagram

$$\begin{tikzcd} & \mathscr{M}_{h_t}(Y) \arrow[dr, "i^*"] & \\ \mathscr{M}_{h_t}(\Sigma) \arrow[ur] \arrow[dr] & & \mathscr{M}(T) \\ & \mathscr{M}(Z) \arrow[ur, "j^*"'] & \end{tikzcd}$$

The flat moduli space of Z near an intersection giving rise to a Type IV component has the form of a union of a 2-dimensional reducible stratum and a 2-parameter family of 2-spheres. This second piece consists of all irreducible points except for a single point on each of the spheres along a curve in the 2-parameter family, and this arc of reducible points is identified with a curve in the 2-dimensional reducible stratum. The restriction map $j^* : \mathscr{M}(Z) \to \mathscr{M}(T)$ collapses 2-spheres to points. See Figure 6.

When $t = 0$, $\mathscr{M}_0(\Sigma)$ contains a pointed 2-sphere, corresponding to an intersection point of $i^*\mathscr{M}_0(Y)$ and $j^*\mathscr{M}(Z)$. The important property of the twisting perturbation is that as t increases, this intersection point moves transversely to the image of $\mathscr{M}^*(Z)$ under j^*. Thus, for $t > 0$, $j^*|_{\mathscr{M}^*(Z)}$ is a submersion with nondegenerate 2-sphere fibers over the intersection point of $i^*\mathscr{M}_{h_t}(Y)$ and $j^*\mathscr{M}(Z)$.

Suppose $[B_0]$ is the reducible flat connection on the pointed 2-sphere. We claim that this point persists under the twisting perturbation. To see why, consider the intersection of $\mathscr{M}_{h_t}(Y)$ and $\mathscr{M}^{red}(Z)$ in $\mathscr{M}(T)$. At $t = 0$, $\mathscr{M}_0(Y) \cap \mathscr{M}^{red}(Z) = \{[B_0]\}$, but the map $\mathscr{M}^{red}(Z) \to \mathscr{M}(T)$ is transverse to $\mathscr{M}_0(Y)$ at $[B_0]$ along the reducible stratum. The situation is illustrated in Figure 6. Thus, after pertubing, the intersection $\mathscr{M}_h(\Sigma) = \mathscr{M}(Z) \cap \mathscr{M}_h(Y)$ near the pointed sphere consists of one

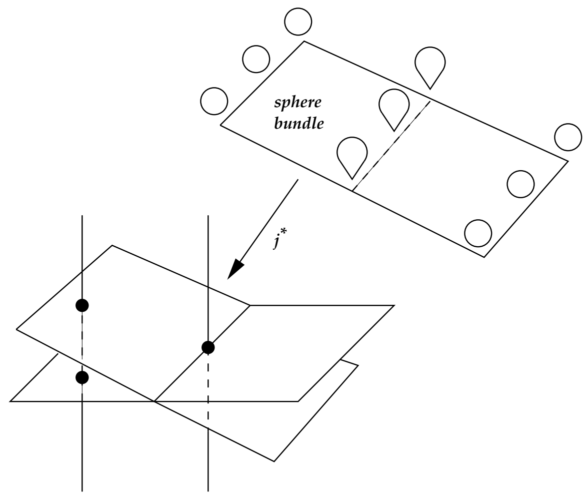

FIGURE 6. THE EFFECT OF A TWISTING PERTURBATION. The two intersecting 2-planes here represent the images of $\mathscr{M}^{red}(Z)$ and $\mathscr{M}^*(Z)$ in $\mathscr{M}(T)$. The two vertical lines, from left to right, represent the images of $\mathscr{M}_h(Y)$ and $\mathscr{M}(Y)$, which are 2-dimensional. Also depicted is the 2-sphere bundle $\mathscr{M}^*(Z)$ being mapped under j^* to the slanted 2-plane.

isolated reducible orbit and a smooth nondegenerate 2-sphere of irreducible orbits (see Figure 7).

Note that the perturbed Chern-Simons function is not generic, since there are critical 2-spheres. However, the 2-sphere is a Bott-Morse critical submanifold, and it is a general fact that such submanifolds contribute their Euler characteristic to the $SU(3)$ Casson invariant, up to sign. Precisely, a Bott-Morse smooth critical submanifold $M \subset \mathscr{M}_h^*$ contributes $\chi(M) \cdot (-1)^{SF(\theta,A)}$ where A is any connection whose gauge orbit lies on M. In the present case, $\chi(S^2) = 2$ and $SF(\theta, A)$ was shown to be even in [**4**]. Thus this 2-sphere contributes 2 to $\tau_{SU(3)}(\Sigma)$. A similar comment applies to compute the contribution from the Type III smooth 2-sphere components, which remain smooth 2-spheres for small enough perturbations. The Type I isolated irreducible points remain isolated irreducible for sufficiently small perturbations.

The results of [**4**] then apply to calculate the contribution of the Type I, Type III, and the resolved smooth 2-spheres coming from the Type IV components. The Type II components do not contribute to $\tau_{SU(3)}(\Sigma)$. To complete the calculation it remains to compute the contribution of the reducible point in $\mathscr{M}_h(\Sigma)$ to the second term of $\tau_{SU(3)}(\Sigma)$ coming from the resolution of the Type IV components.

5.3. Normal spectral flow along the path of twisted perturbed flat reducible connections. For $0 \leq t \leq \varepsilon$, let B_t be the path of h_t-perturbed flat reducible connections near a pointed 2-sphere, i.e. B_t corresponds to the intersection of $\mathscr{M}_{h_t}(Y)$ with $\mathscr{M}(Z)$ in $\mathscr{M}(T)$ near the singular point of the pointed 2-sphere. In light of the definition of the second term of $\tau_{SU(3)}$, we need to calculate the normal

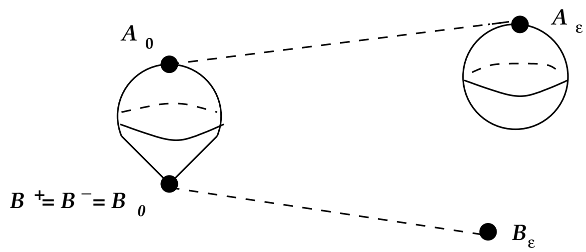

FIGURE 7. A POINTED 2-SPHERE RESOLVES INTO A SMOOTH 2-SPHERE AND A REDUCIBLE POINT.

spectral flow $SF_n(B^{\pm}, B_\varepsilon)$. Notice that B_0 is isolated in the reducible flat moduli space, and so we take $B^+ = B^- = B_0$.

To calculate $SF_n(B_0, B_\epsilon)$, we split the spectral flow according to the manifold decomposition

$$\Sigma = Y \cup_T Z.$$

We use Atiyah-Patodi-Singer boundary conditions on Y and Z to get self-adjoint operators. Consider first a connection A on Y whose restriction to a collar of the boundary is flat, and let $a = A|_T$. The tangential operator S_a of the odd signature operator K_A is

$$S_a \colon \Omega^{0+1+2}(T;\mathbb{C}^2) \to \Omega^{0+1+2}(T;\mathbb{C}^2)$$
$$S_a(\alpha,\beta,\gamma) = (*d_a\beta, -*d_a\alpha - d_a * \gamma, d_a * \beta)\,.$$

(In this formula, $d_a \colon \Omega^i(T;\mathbb{C}^2) \to \Omega^{i+1}(T;\mathbb{C}^2)$ is the exterior derivative associated to the connection a.) After a gauge transformation, if needed, the formula $K_A = \sigma(du)(S_a + \frac{\partial}{\partial u})$ holds in a collar $T \times I$ of the boundary, where σ denotes the symbol of K_A and u is the inward normal coordinate.

In this context Atiyah, Patodi and Singer in [**2**] showed that the operator K_A acting on those forms in $\Omega^{0+1}(Y;\mathbb{C}^2)$ whose restriction to the boundary lie in the positive eigenspan P_a^+ of S_a is Fredholm, and self-adjoint if in addition $\ker S_a = 0$. When a is flat, $\ker S_a$ is isomorphic via the Hodge theorem to $H_a^{0+1+2}(T;\mathbb{C}^2)$. This cohomology group vanishes for nontrivial flat connections a.

Applying this to the path B_t restricted to Y gives a family of self-adjoint operators whose (normal) spectral flow is denoted by $SF_n(B_t;P^+;Y;0\le t\le\varepsilon)$. Similar considerations on Z (taking into account the change in orientation of the inward normal) gives the spectral flow $SF_n(B_t;P^-;Z;0\le t\le\varepsilon)$. A theorem of Bunke ([**12**], see [**14**] for general splitting theorems for spectral flow) implies that

$$SF_n(B_t;\Sigma;0\le t\le\varepsilon) = SF_n(B_t;P^+;Y;0\le t\le\varepsilon) + SF_n(B_t;P^-;Z;0\le t\le\varepsilon).$$

THEOREM 5.2. *Let B_t be the path of reducible perturbed flat connections near a pointed 2-sphere described above. Then*

(i) $SF_n(B_t,P^+;Y;0\le t\le\varepsilon) = 0,$
(ii) $SF_n(B_t,P^-;Z;0\le t\le\varepsilon) = -2.$

Hence each pointed 2-sphere contributes $+2$ to $\tau_{SU(3)}(\Sigma(p,q,r))$.

Sketch of proof. Part (i) follows from vanishing of cohomology $H_{B_0}^{0+1}(Y;\mathbb{C}^2)$ and is left to the reader. Instead, we outline the argument for part (ii), which uses additivity of the spectral flow under composition of paths.

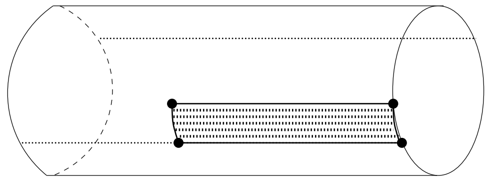

Figure 8. When p and q are both odd, the components of reducible $SU(3)$ representations of $\pi_1 Z$ are homeomorphic to cylinders $S^1 \times [0,1]$.

The space of conjugacy classes of $SU(2)$ representations of $\pi_1 Z$ is a union of arcs $\rho_s, s \in [0,1]$. The endpoints of these arcs are abelian $SU(2)$ representations. (These facts are proved in [**21**].) Since $H_1(Z;\mathbb{Z}) = \mathbb{Z}$, we can twist each $SU(2)$ representation ρ_s to get a representation $\pi_1 Z \to U(2)$ by simply multiplying ρ_s by the (unique) abelian representation $\chi_t \colon \pi_1 Z \to U(1)$ with $\chi_t(\gamma) = e^{2\pi i t}$ for γ generator of $H_1(Z;\mathbb{Z})$.

Identifying $U(2)$ with the subgroup $S(U(2)\times U(1)) \subset SU(3)$, the family $\alpha_{s,t} = \chi_t \rho_s$, parameterizes a family of reducible $SU(3)$ representations where $(s,t) \in [0,1] \times (-\epsilon, \epsilon)$ and $\epsilon > 0$ is chosen small.

The restriction to Z of the path B_t is a path of flat connections whose holonomy is essentially of the form $\alpha_{s_0,t}$ for some fixed $s_0 \in (0,1)$. Thus there is a two parameter family $B_{s,t}$ of flat connections on Z defined for $(s,t) \in [0,1] \times [0,\varepsilon]$ with $B_{s_0,t} = B_t$ and $B_{1,t}$ abelian.

Using the invariance of spectral flow with respect to homotopy rel endpoints and the additivity with respect to composition of paths, we have:

$$\begin{aligned} SF(B_t; P^-; Z; 0 \le t \le \varepsilon) &= SF(B_{s_0,t}; P^-; Z; 0 \le t \le \varepsilon), \\ &= SF(B_{s,0}; P^-; Z; s_0 \le s \le 1) + SF(B_{1,t}; P^-; Z; 0 \le t \le \varepsilon) \\ &\qquad - SF(B_{s,1}; P^-; Z; s_0 \le s \le 1), \\ &= SF(B_{1,t}; P^-; Z; 0 \le t \le \varepsilon). \end{aligned}$$

To see the last step, note that $\dim H^1_{B_{s,0}}(Z;\mathbb{C}^2) = 4$ and $\dim H^1_{B_{s,1}}(Z;\mathbb{C}^2) = 0$ are both constant in s, hence $SF(B_{s,0}; P^-; Z; s_0 \le s \le 1)$ and $SF(B_{s,1}; P^-; Z; s_0 \le s \le 1)$ vanish.

Now $\{[B_{1,t}] \mid 0 \le t \le \varepsilon\} \subset \mathscr{M}^{ab}(Z)$ is a path of connections with abelian holonomy. The operator $K_{B_{1,t}}$ on Z with P^- boundary conditions has four zero modes at $t = 0$ and no zero modes for $0 < t \le \varepsilon$. Working with abelian connections is simpler and one can show (using for example Levine-Tristam signatures, as in [**20**]) that as t increases from $t = 0$, two of the zero modes go up and the other two go down, hence (with our convention for spectral flow)

$$SF_n(B_t; Z; P^-; 0 \le t \le \varepsilon) = SF(B_{1,t}; P^-; Z; 0 \le t \le \varepsilon) = -2.$$

This establishes (ii).

Using Bunke's theorem, the fact that $B^+ = B^- = B_0$ in this case, the computation that $\dim H^1_{B_0}(\Sigma;\mathbb{C}^2) = 4$, and the definition of the correction term for

$\tau_{SU(3)}$, it follows that the perturbed flat reducible gauge orbit $[B_\varepsilon]$ contributes

$$-\tfrac{1}{4}(-1)^{SF(\theta,B_\varepsilon)}\big(2SF_n(B_0,B_\varepsilon)+\dim H^1_{B_0}(\Sigma;\mathbb{C}^2)\big)$$
$$= -\tfrac{1}{4}(-1)^{2SF(\theta,B_\varepsilon)}\big(2(-2)+4\big)=0$$

to $\tau_{SU(3)}(\Sigma)$. Thus the only contribution of a pointed 2-sphere to $\tau_{SU(3)}(\Sigma)$ comes from the smooth 2-sphere of irreducibles obtained after perturbing which contributes $\chi(S^2)=2$ to (the first term in the definition of) $\tau_{SU(3)}(\Sigma)$. □

Σ	$\tau_{SU(3)}(\Sigma)$
$\Sigma(3,4,12k\pm 1)$	$105k^2 \pm 21k$
$\Sigma(3,4,12k\pm 5)$	$105k^2 \pm 87k + 16$
$\Sigma(3,5,15k\pm 1)$	$276k^2 \pm 40k$
$\Sigma(3,5,15k\pm 2)$	$276k^2 \pm 74k + 2$
$\Sigma(3,5,15k\pm 4)$	$276k^2 \pm 148k + 16$
$\Sigma(3,5,15k\pm 7)$	$276k^2 \pm 254k + 56$

TABLE 2. CALCULATIONS OF THE INTEGER VALUED SU(3) CASSON INVARIANT FOR BRIESKORN SPHERES $\Sigma(p,q,r)$ WITH $p>2$.

Table 2 gives some sample computations of the integer valued Casson invariant. Let $K_{p,q}$ = the (p,q) torus knot and set $X_n = 1/n$ Dehn surgery on $K_{p,q}$. Then $X_n = \pm\Sigma(p,q,r)$ for $r=|pqn-1|$. Table 3 gives the value of $\tau_{SU(3)}(X_n)$ for various p,q. These calculations suggest several conjectures. First of all, in all these computations, $\tau_{SU(3)}(\Sigma)$ is even, which lends evidence to the following conjecture of [**8**].

CONJECTURE 5.3. *$\tau_{SU(3)}(X)$ is even for all integral homology 3-spheres.*

One reason for believing Conjecture 5.3 (apart from the empirical evidence) is the existence of a natural involution on $\mathscr{M}_{SU(3)}$ induced by complex conjugation on $SU(3)$ and $su(3)$. For example, one can use the involution to prove the conjecture under the hypothesis that the flat moduli space $\mathscr{M}_{SU(3)}$ is regular, which is the condition that $H^1_\alpha(X;su(3))=0$ for every nontrivial representation $\alpha:\pi_1 X\to SU(3)$).

A second pattern observed in the data is quadratic growth in the surgery coefficient for successive surgeries on a fixed knot.

CONJECTURE 5.4. *If $K\subset X$ is a knot in a homology 3-sphere and X_n is the result of $1/n$ Dehn surgery on K, then the limit*

$$\lim_{n\to\infty}\tau_{SU(3)}(X_n)/n^2 = A(K)$$

exists and depends only on the knot K.

$p=2$	$\tau_{SU(3)}(X_n)$	$p=3$	$\tau_{SU(3)}(X_n)$	$p=4$	$\tau_{SU(3)}(X_n)$
$K_{2,3}$	$3n^2-n$	$K_{3,4}$	$105n^2-21n$	$K_{4,5}$	$1011n^2-111n$
$K_{2,5}$	$33n^2-9n$	$K_{3,5}$	$276n^2-40n$	$K_{4,7}$	$4110n^2-320n$
$K_{2,7}$	$138n^2-26n$	$K_{3,7}$	$1128n^2-124n$	$K_{4,9}$	$11\,490n^2-712n$
$K_{2,9}$	$390n^2-58n$	$K_{3,8}$	$1953n^2-179n$	$K_{4,11}$	$25\,935n^2-1297n$
$K_{2,11}$	$885n^2-107n$	$K_{3,10}$	$4851n^2-367n$	$K_{4,13}$	$50\,925n^2-2171n$
$K_{2,13}$	$1743n^2-179n$	$K_{3,11}$	$7140n^2-476n$	$K_{4,15}$	$90\,636n^2-3320n$
$K_{2,15}$	$3108n^2-276n$	$K_{3,13}$	$14\,028n^2-812n$	$K_{4,17}$	$149\,940n^2-4888n$
$K_{2,17}$	$5148n^2-404n$	$K_{3,14}$	$18\,915n^2-993n$	$K_{4,19}$	$234\,405n^2-6789n$
$K_{2,19}$	$8055n^2-565n$	$K_{3,16}$	$32\,385n^2-1517n$	$K_{4,21}$	$350\,295n^2-9231n$
$K_{2,21}$	$12\,045n^2-765n$	$K_{3,17}$	$41\,328n^2-1788n$	$K_{4,23}$	$504\,570n^2-12\,072n$
$K_{2,23}$	$17\,358n^2-1006n$	$K_{3,19}$	$64\,620n^2-2544n$	$K_{4,25}$	$704\,886n^2-15\,600n$
$K_{2,25}$	$24\,258n^2-1294n$	$K_{3,20}$	$79\,401n^2-2923n$	$K_{4,27}$	$959\,595n^2-19\,569n$
$K_{2,27}$	$33\,033n^2-1631n$	$K_{3,22}$	$116\,403n^2-3951n$	$K_{4,29}$	$1\,277\,745n^2-24\,363n$

Table 3. The integer valued SU(3) Casson invariant for homology 3-spheres X_n obtained by $1/n$ surgery on $K_{p,q}$

The data from Table 3 even gives a precise (conjectural) formula for $A(K)$ for torus knots:

$$A(K_{p,q}) = \frac{(p^2-1)(q^2-1)(2p^2q^2-3p^2-3q^2-3)}{240},$$

which is consistent with the more general formula (also conjectural) that

$$A(K) = 6c_4(K) + 3c_2(K)^2, \tag{5.1}$$

where $\Delta_K(z) = \sum_{i\geq 0} c_{2i}(K) z^{2i}$ is the Conway polynomial of K. Based on Frohman's work [**17**], formula (5.1) is essentially equivalent to the conjecture that $A(K)$ equals is six times the Frohman-Nicas $SU(3)$ knot invariant, at least for fibered knots (cf. [**18, 10**]).

Table 3 suggests a stronger result, namely that

$$\tau_{SU(3)}(X_n) = A(K)n^2 - B(K)n + \tau_{SU(3)}(X) \tag{5.2}$$

is a quadratic polynomial. Here, the assumptions are as before: $K \subset X$ is a knot in a homology 3-sphere and X_n is the result of performing $1/n$ Dehn surgery on K. For example, interpolating the data from Table 3,we get the formula

$$B(K_{2,q}) = \begin{cases} \frac{1}{12}(q^3-4q+3) & \text{if } q \equiv 1 \mod 4, \\ \frac{1}{12}(q^3-4q-3) & \text{if } q \equiv 3 \mod 4, \end{cases}$$

for $p=2$. Further, for $p=3$ and $p=4$, we get the formulas

$$B(K_{3,q}) = \begin{cases} \frac{1}{54}(20q^3+3q^2-48q+25) & \text{if } q \equiv 1 \mod 6, \\ \frac{1}{54}(20q^3-3q^2-48q+2) & \text{if } q \equiv 2 \mod 6, \\ \frac{1}{54}(20q^3+3q^2-48q-2) & \text{if } q \equiv 4 \mod 6, \\ \frac{1}{54}(20q^3-3q^2-48q-25) & \text{if } q \equiv 5 \mod 6, \end{cases}$$

and

$$B(K_{4,q}) = \begin{cases} \frac{1}{16}(16q^3 + q^2 - 42q + 25) & \text{if } q \equiv 1 \mod 8, \\ \frac{1}{16}(16q^3 - q^2 - 42q + 39) & \text{if } q \equiv 3 \mod 8, \\ \frac{1}{16}(16q^3 + q^2 - 42q - 39) & \text{if } q \equiv 5 \mod 8, \\ \frac{1}{16}(16q^3 - q^2 - 42q - 25) & \text{if } q \equiv 7 \mod 8. \end{cases}$$

The complexity of these formulas makes it difficult to guess a general formula for $B(K)$ in terms of classical invariants of the knot.

Notice that the conjectured formula (5.1) for $A(K)$ is not consistent with the conjecture (5.2) that $\tau_{SU(3)}(X_n)$ is a quadratic polynomial in the surgery coefficient. We explain this point using the well known fact that, for any amphicheiral knot K in S^3, $X_{-n} = -X_n$. Since $\tau_{SU(3)}(-X) = \tau_{SU(3)}(X)$, the conjectured formula (5.2) would imply that $B(K) = 0$ and that $\tau_{SU(3)}(X_n) = A(K)n^2$ for amphicheiral knots in S^3.

The figure eight knot K is amphicheiral, and ± 1 surgery on K gives the Brieskorn sphere $\pm\Sigma(2,3,7)$. Hence $\tau_{SU(3)}(X_{\pm 1}) = 4$. Assuming Conjecture 5.4 is true (so there is a quadratic growth coefficient $A(K)$), the formula (5.1) predicts that $A(K) = 3$ (because the Conway polynomial of K equals $1 - z^2$). But $B(K) = 0$, $\tau_{SU(3)}(S^3) = 0$ and the value of $\tau_{SU(3)}(X_{\pm 1})$ for ± 1 surgery on K equals 4, hence $A(K) = 4$. So if $\tau_{SU(3)}(X_n)$ is a quadratic polynomial as in (5.2), then formula (5.1) for $A(K)$ is incorrect.

Finding a direct method for computing $\tau_{SU(3)}$ for surgeries on the figure eight knot is an important step that promises to help determine the correct surgery formula for the integer valued $SU(3)$ Casson invariant.

Acknowledgements The authors gratefully acknowledge support from the National Science Foundation grants DMS-9971578, DMS-0072564 and DMS-9971020.

References

[1] S. Akbulut and J. McCarthy, Casson's invariant for oriented homology 3-spheres, an exposition, Math. Notes, Vol. **36**, Princeton University Press, Princeton, NJ, 1990.

[2] M. F. Atiyah, V. K. Patodi, and I. M. Singer, *Spectral asymmetry and Riemannian geometry. I.* Math. Proc. Cambridge Philos. Soc. **77** (1975), no. 1 43–69.

[3] M. F. Atiyah, V. K. Patodi, and I. M. Singer, *Spectral asymmetry and Riemannian geometry. II.* Math. Proc. Cambridge Philos. Soc. **78** (1975), no. 3 405–432.

[4] H. U. Boden, *Unitary representations of Brieskorn spheres,* Duke J. Math. **75** (1994) 193–220.

[5] H. U. Boden and C. M. Herald, *The SU(3) Casson invariant for integral homology 3-spheres,* J. Differential Geom. **50** (1998) 147–206.

[6] H. U. Boden and C. M. Herald, *A connected sum formula for the SU(3) Casson invariant,* J. Differential Geom. **53** (1999) 443–465.

[7] H. U. Boden, C. M. Herald, P. Kirk, and E. P. Klassen, *Gauge theoretic invariants of Dehn surgeries on knots*, Geom. Topology **5** (2001) 143–226.

[8] H. U. Boden, C. M. Herald and P. Kirk, *An integer valued $SU(3)$ Casson invariant*, Math. Research Letters **8** (2001) 589–602.

[9] H. U. Boden, C. M. Herald and P. Kirk, *The integer valued $SU(3)$ Casson invariant for Brieskorn homology spheres*, in preparation.

[10] H. U. Boden and A. Nicas, *Universal formulae for SU(n) Casson invariants of knots,* Trans. Amer. Math. Soc. **352** (2000) 3149–3187.

[11] B. Booss-Bavnbek, M. Lesch, and J. Phillips, *Unbounded Fredholm Operators and Spectral Flow,* 2001 preprint, math.FA/0108014.

[12] U. Bunke, *On the gluing problem for the η-invariant,* J. Differential Geom. **41** (1995), no. 2, 397–448.
[13] S. Cappell, R. Lee, and E. Miller, *A perturbative $SU(3)$ Casson invariant,* math.DG/0006018, to appear in Comment. Math. Helv.
[14] A. M. Daniel and P. Kirk, *A general splitting formula for the spectral flow.* With an appendix by K. P. Wojciechowski. Michigan Math. J. **46** (1999), no. 3, 589–617.
[15] S. Donaldson and P. Kronheimer, *The Geometry of Four-Manifolds,* Oxford Mathematical Monographs, Clarendon Press, Oxford, 1990.
[16] A. Floer, *An instanton invariant for 3-manifolds,* Comm. Math. Phys. **118** (1989) 215–240.
[17] C. Frohman, *Unitary representations of knot groups*, Topology **32** (1993) 121–144.
[18] C. Frohman and A. Nicas, *An intersection homology invariant for knots in a rational homology 3-sphere*, Topology **33** (1994) 123–158.
[19] M. Hayashi, *The moduli space of SU(3)-Flat connections and the fusion rules,* Proc. Amer. Math. Soc., **127**, no. 5 (1999), 1545–1555.
[20] P. Kirk, E. Klassen, and D. Ruberman, *Splitting the spectral flow and the Alexander matrix,* Comment. Math. Helv. **69** (1994), no. 3, 375–416.
[21] E. Klassen, *Representations of knot groups in* SU(2)*,* Trans. Amer. Math. Soc. **326** (1991), no. 2, 795–828.
[22] C. Lescop, Global surgery formula for the Casson-Walker invariant, Annals of Math Studies **140**, Princeton University Press, Princeton, NJ, 1996.
[23] C. Taubes, *Casson's invariant and gauge theory,* J. Differential Geom. **31** (1990) 547–599.
[24] K. Walker, An extension of Casson's invariant, Annals of Math Studies **126**, Princeton University Press, Princeton, NJ, 1992.
[25] A. Wasserman, *Equivariant differential topology,* Topology **8** (1969) 127–150.

Department of Mathematics & Statistics, McMaster University, Hamilton, Ontario L8S-4K1, Canada
E-mail address: boden@mcmaster.ca

Department of Mathematics/084, University of Nevada, Reno NV 89557-0045, USA
E-mail address: herald@unr.edu

Department of Mathematics, Indiana University, Bloomington IN 47405, USA
E-mail address: pkirk@indiana.edu