Adv. Theor. Math. Phys. **15** (2011) 71–129

# Supersymmetric surface operators, four-manifold theory and invariants in various dimensions

**Meng-Chwan Tan**

Department of Physics, National University of Singapore,
Singapore 119260, Singapore
tan@ias.edu

**Abstract**

We continue our program initiated in [1] to consider supersymmetric surface operators in a topologically twisted $\mathcal{N}=2$ pure $SU(2)$ gauge theory, and apply them to the study of four-manifolds and related invariants. Elegant physical proofs of various seminal theorems in four-manifold theory obtained by Ozsváth and Szabó [2, 3] and Taubes [4], will be furnished. In particular, we will show that Taubes' groundbreaking and difficult result — that the ordinary SW invariants are in fact the Gromov invariants which count pseudo-holomorphic curves embedded in a symplectic four-manifold $X$ — nonetheless lends itself to a simple and concrete physical derivation in the presence of "ordinary" surface operators. As an offshoot, we will be led to several interesting and mathematically novel identities among the Gromov and "ramified" SW invariants of $X$, which in certain cases, also involve the instanton and monopole Floer homologies of its three-submanifold. Via these identities, and a physical formulation of the "ramified" Donaldson invariants of four-manifolds with boundaries, we will uncover completely new and economical ways of deriving and understanding various important

e-print archive: http://lanl.arXiv.org/abs/1006.3313

mathematical results concerning (i) knot homology groups from "ramified" instantons by Kronheimer and Mrowka [5]; and (ii) monopole Floer homology and SW theory on symplectic four-manifolds by Kutluhan–Taubes [4, 6]. Supersymmetry, as well as other physical concepts such as $R$-invariance, electric–magnetic duality, spontaneous gauge symmetry-breaking and localization onto supersymmetric configurations in topologically twisted quantum field theories, play a pivotal role in our story.

## Contents

## 1 Introduction and summary

Supersymmetric surface operators in a topologically twisted $\mathcal{N} = 2$ pure $SO(3)$ or $SU(2)$ gauge theory have recently been analyzed in detail in [1], where, among other things, concrete physical proofs of various seminal theorems in four-dimensional geometric topology obtained by Kronheimer and Mrowka in [7–9], were also furnished. For example, it was shown in [1] that the Kronheimer–Mrowka result of [7] — which identifies the "ramified" Donaldson invariants as the ordinary Donaldson invariants of an "admissible" four-manifold $X$ with $b_2^+ > 1$ — is a direct consequence of a required modular invariance over the $u$-plane in the presence of nontrivially embedded surface operators. It was also shown in [1] that a generalization of the Thom conjecture proved by Kronheimer and Mrowka in [7] — which leads to a minimal genus formula for embedded surfaces of non-negative self-intersection in $X$ — is a direct result of the $R$-invariance of the correlation functions in the microscopic non-abelian gauge theory which correspond to the ("ramified") Donaldson invariants of $X$.

In this paper, we continue the program initiated in [1]; we consider arbitrarily embedded surface operators in a topologically twisted $\mathcal{N} = 2$ pure $SU(2)$ gauge theory, and apply them to the study of four-manifolds and invariants in two, three and four dimensions. The plan and results of our work can be summarized as follows.

In Section 2, we will review various aspects of the topologically twisted $\mathcal{N} = 2$ pure $SU(2)$ gauge theory on $X$ with arbitrarily embedded surface operators, and the corresponding physical interpretations of the "ramified" Donaldson and Seiberg–Witten (SW) invariants and their associated moduli spaces, all of which will be useful and relevant to our arguments and computations in the later sections.

In Section 3, with the aid of key results computed in [1], we will furnish physical proofs of various seminal theorems in four-dimensional geometric topology obtained by Ozsváth–Szabó in [2, 3]; in particular, we will physically demonstrate a minimal genus formula obtained earlier in [2] for embedded surfaces of *negative* self-intersection. $R$-invariance and electric–magnetic duality underlie our proofs in this section.

In Section 4, we will present an elegant physical derivation of Taubes' stunning result in [4], which identifies the SW invariants as the Gromov invariants on a symplectic four-manifold with $b_2^+ > 1$. The crucial ingredients in this derivation are supersymmetry, $R$-invariance, electric–magnetic duality, spontaneous gauge symmetry-breaking and localization onto supersymmetric configurations in topologically twisted quantum field theories. In essence, one can understand Taubes' result to be a consequence of the scale invariance of a particular instanton sector of the topologically twisted gauge theory in the presence of "ordinary" curved surface operators which wrap pseudo-holomorphic curves embedded in the symplectic four-manifold.

In Section 5, we will explore the mathematical implications of the underlying physics. We will compute — using certain intermediate results obtained in Section 3 and Section 4 — various mathematically novel identities involving the Gromov and ("ramified") SW invariants of a symplectic four-manifold with $b_2^+ > 1$. These identities, which one can understand to exist because of $R$-invariance, are also found to be consistent with more general theorems established in the mathematical literature. In addition, for symplectic $X = M \times \mathbf{S}^1$, where $M$ is a closed, oriented three-submanifold, we will show — via a supersymmetric quantum mechanical interpretation of the topological gauge theory with surface operators — that a knot homology conjecture proposed by Kronheimer and Mrowka in [5] ought to hold on purely physical grounds, and that the Gromov invariant of $X$ is given by the Euler characteristic of the instanton Floer homology of $M$. In turn, because the Euler characteristic of the instanton Floer homology of $M$ is given by the Casson–Walker–Lescop invariant of $M$, the Gromov invariant of $X$ is zero if $b_2^+(X) > 3$. We will also derive, amidst other things, an interesting relation between the instanton and monopole Floer homologies of $M$, and a novel identity between the SW invariants of $M$. Last but not least, we will formulate "ramified" generalizations of various formulas presented by Donaldson and Atiyah in [10, 11] relating ordinary Donaldson and Floer theory on four-manifolds with boundaries, in terms of "three-one branes".

In Section 6, we will generalize our computations in Section 4 to involve multiple surface operators which are *disjoint*. This will allow us to physically derive Taubes' result in all generality.

In Section 7, the final section, by further applying our physical insights and results obtained hitherto, we will first provide — via the topological gauge theory with nontrivially embedded surface operators on a general four-manifold with boundaries — a physical derivation of certain key properties of knot homology groups from "ramified" instantons defined and proved by Kronheimer and Mrowka in [5]. Then, via the identities obtained in Section 5, and certain key relations computed in Section 3 and Section 4, we will re-derive various important mathematical results concerning the monopole Floer homology of three-manifolds and SW theory on symplectic four-manifolds.

This paper is dedicated to See–Hong, whose strength, courage and optimism in the face of grave adversity have made this otherwise impossible endeavor, possible.

## 2 Surface operators and the "ramified" Donaldson and SW invariants

In this section, we will present some background material that will be necessary for a coherent, self-contained understanding of the main discussions in this paper. We will be brief in our exposition, although concepts deemed to play a crucial role will be reviewed in greater detail.

## 2.1 Embedded surfaces and the "ramified" Donaldson invariants

Let us first review the mathematical definition of embedded surfaces and the "ramified" Donaldson invariants by Kronheimer and Mrowka (henceforth denoted as KM) in [7]. To this end, let $X$ be a smooth, compact, simply connected, oriented four-manifold with Riemannian metric $\bar{g}$, and let $E \to X$ be an $SO(3)$-bundle over $X$ (i.e., a rank-three real vector bundle with a metric).

*Embedded surfaces*

An embedded surface $D$ is characterized by a two-submanifold of $X$ that is a complex curve of genus $g$ and self-intersection number $D^2$. Consider the case where the second SW class $w_2(E) = 0$; the structure group of $E$ can then be lifted to its $SU(2)$ double-cover. In the neighborhood of $D$, one can choose a decomposition of $E$ as

$$E = L \oplus L^{-1}, \tag{2.1}$$

where $L$ is a complex line bundle over $X$. In the presence of $D$, the connection matrix of $E$ restricted to $X \backslash D$ (in the *real* Lie algebra) will look like

$$A = \alpha d\theta + \cdots , \tag{2.2}$$

where $\alpha$ is a real number valued in the generator

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{2.3}$$

of the Cartan subalgebra $\mathfrak{t}$, $\theta$ is the angular variable of the coordinate $z = r\,\mathrm{e}^{\mathrm{i}\theta}$ of the plane normal to $D$, and the ellipses refer to the ordinary terms that are regular near $D$. Notice that since $d\theta = \mathrm{i}dz/z$, the connection is singular as $z \to 0$, i.e., as one approaches $D$. In any case, the singularity in the connection induces the following gauge-invariant holonomy:

$$\exp(2\pi\alpha) \tag{2.4}$$

around any small circle that links $D$. Hence, if the holonomy is trivial, we are back to considering ordinary connections on $E$. Therefore, $\alpha$ actually takes values in $\mathbb{T}$, the maximal torus of the gauge group with Lie algebra $\mathfrak{t}$. As we shall see shortly, this mathematical definition of embedded surfaces will coincide with our (more general) physical definition of surface operators.

*The "ramified" Donaldson invariants*

In analogy with the original formulation of Donaldson theory [12], KM introduced the notion of "ramified" Donaldson invariants — i.e., Donaldson invariants of $X$ with an embedded surface $D$. According to KM [7, 9], the "ramified" Donaldson polynomials $\mathcal{D}'_E$ can be defined as polynomials on the homology of $X\backslash D$ with real coefficients:[1]

$$\mathcal{D}'_E : H_0(X\backslash D, \mathbb{R}) \oplus H_2(X\backslash D, \mathbb{R}) \to \mathbb{R}. \tag{2.5}$$

Assigning degree 4 to $p \in H_0(X\backslash D, \mathbb{R})$ and 2 to $S \in H_2(X\backslash D, \mathbb{R})$, the degree $s$ polynomial may be expanded as

$$\mathcal{D}'_E(p, S) = \sum_{2m+4t=s} S^m p^t d^{k'}_{m,t}, \tag{2.6}$$

where $s$ is the dimension of the moduli space $\mathcal{M}'$ of gauge-inequivalent classes of anti-self-dual connections on $E$ restricted to $X\backslash D$ with first Pontrjagin number $p_1(E)$ and instanton number $k' = -\int_{X\backslash D} p_1(E)/4$. The numbers $d^{k'}_{m,t}$ — in other words, the "ramified" Donaldson invariants of $X$ — can be defined as in Donaldson theory in terms of intersection theory on the moduli space $\mathcal{M}'$; for maps

$$\begin{aligned} p \in H_0(X\backslash D, \mathbb{R}) &\to \Omega^0(p) \in H^4(\mathcal{M}'),\\ S \in H_2(X\backslash D, \mathbb{R}) &\to \Omega^2(S) \in H^2(\mathcal{M}'), \end{aligned} \tag{2.7}$$

the "ramified" Donaldson invariants can be written as

$$d^{k'}_{m,t} = \int_{\mathcal{M}'} [\Omega^0(p)]^t \wedge \Omega^2(S_{i_1}) \wedge \cdots \wedge \Omega^2(S_{i_m}). \tag{2.8}$$

Moreover, one can also package the "ramified" Donaldson polynomials into a generating function: by summing over all topological types of bundle $E$ with fixed $\xi = w_2(E)$ (where $\xi$ may be non-vanishing in general) but

---

[1]To be precise, KM actually defines the "ramified" Donaldson polynomials to be the map $\mathcal{D}'_E : \mathrm{Sym}[H_0(X\backslash D, \mathbb{R}) \oplus H_2(X\backslash D, \mathbb{R})] \otimes \wedge * H_1(X\backslash D, \mathbb{R}) \to \mathbb{R}$. However, their definition can be truncated as shown, in accordance with Donaldson's original formulation in [12].

varying $k'$, the generating function can be defined as

$$\mathbf{Z}'_{\xi,\bar{g}}(p,S) = \sum_{k'} \sum_{m\geq 0, t\geq 0} \frac{S^m}{m!}\frac{p^t}{t!} d^{k'}_{m,t}. \tag{2.9}$$

Clearly, $\mathbf{Z}'_{\xi,\bar{g}}$ depends on the class $w_2(E)$ but not on the instanton number $k'$ (as this has been summed over). In analogy with the ordinary case, one can also define the "ramified" Donaldson series as

$$\mathscr{D}'_{\xi}(S) = \left(1 + \frac{1}{2}\frac{\partial}{\partial p}\right) \cdot \mathbf{Z}'_{\xi,\bar{g}}(p,S)|_{p=0}. \tag{2.10}$$

*About the moduli space of "ramified" instantons*

Another relevant result by KM is the following. Assuming that there are no reducible connections on $E$ restricted to $X \backslash D$, $\mathcal{M}'$ — which we will hereon refer to as the moduli space of "ramified" instantons — will be a smooth manifold of finite dimension

$$s = 8k - \tfrac{3}{2}(\chi + \sigma) + 4l - 2(g-1) \tag{2.11}$$

for *any* nontrivial value of $\alpha$. Here, $\chi$ and $\sigma$ are the Euler characteristic and signature of $X$, and for $\xi = 0$, the integer $k$ is given by

$$k = -\frac{1}{8\pi^2}\int_X \mathrm{Tr}\, F \wedge F, \tag{2.12}$$

where $F$ is the curvature of the bundle $E$ over $X$, and Tr is the trace in the two-dimensional representation of $SU(2)$. The integer $l$ — called the monopole number by KM — is given by

$$l = -\int_D c_1(L). \tag{2.13}$$

Here, $c_1(L) = -F_L/2\pi$, where $F_L$ is the curvature of $L$; thus, $l$ measures the degree of the reduction of $E$ near $D$.

As shown in [1], $l$ will depend explicitly on $\alpha$ because the singular term in the connection $A$ will result in a singularity proportional to $\alpha$ along $D$ in the field strength (extended over $D$). Likewise, $k$ will also depend explicitly on $\alpha$. Thus, the invariance of $s$ must mean that both $l$ and $k$ will vary with $\alpha$ in such a way as to keep it fixed for any nontrivial value of $\alpha$.

*Topological invariance of* $\mathbf{Z}'_{\xi,\bar{g}}$

Let $b_2^+$ denote the self-dual part of the second Betti number of $X$. According to KM (see Section 7 of [9]), if $b_2^+ > 1$, $\mathbf{Z}'_{\xi,\bar{g}}$ is independent of the metric $\bar{g}$ and hence, just like the generating function of the ordinary Donaldson invariants, defines invariants of the smooth structure of $X$. This is consistent with the fact that for $b_2^+ \geq 3$ (and $b_1 = 0$), the "ramified" Donaldson invariants can be expressed solely in terms of the ordinary Donaldson invariants (see Theorem 5.10 of [7], and its physical proof in Section 8 of [1]).

However, if $b_2^+ = 1$, we run into the phenomenon of chambers; $\mathbf{Z}'_{\xi,\bar{g}}$ will jump as we move across a "wall" in the space of metrics on $X$. (This phenomenon was demonstrated via a purely physical approach in Section 6 of [1].)

## 2.2 Surface operators in pure $SU(2)$ theory with $\mathcal{N} = 2$ supersymmetry

*Supersymmetric surface operators*

We would like to define surface operators along $D$ which are compatible with $\mathcal{N} = 2$ supersymmetry. In other words, they should be characterized by solutions to the supersymmetric field configurations of the underlying gauge theory on $X$ that are singular along $D$.

In order to ascertain what these solutions are, first note that any supersymmetric field configuration of a theory must obey the conditions implied by setting the supersymmetric variations of the fermions to zero. In the original (untwisted) theory without surface operators, this implies that any supersymmetric field configuration must obey $F = 0$ and $\nabla_\mu a = 0$, where $a$ is a scalar field in the $\mathcal{N} = 2$ vector multiplet [13]. Let us assume for simplicity the trivial solution $a = 0$ to the condition $\nabla_\mu a = 0$ (so that the relevant moduli space is non-singular); this means that any supersymmetric field configuration must be consistent with *irreducible* flat connections on $X$ that obey $F = 0$. Consequently, any surface operator along $D$ that is supposed to be supersymmetric and compatible with the underlying $\mathcal{N} = 2$ supersymmetry, ought to correspond to a *flat* irreducible connection on $E$ restricted to $X \backslash D$ which has the required singularity along $D$.[2] Let us for convenience choose the singularity of the connection along an oriented $D$

[2]This prescription of considering connections on the bundle $E$ restricted to $X \backslash D$ whenever one inserts a surface operator that introduces a field singularity along $D$, is just a two-dimensional analog of the prescription one adopts when inserting an 't Hooft loop operator in the theory. See Section 10.1 of [14] for a detailed explanation of the latter.

to be of the form shown in (2.2). Then, since $d(\alpha d\theta) = 2\pi\alpha\delta_D$, where $\delta_D$ is a delta two-form Poincaré dual to $D$ with support in a tubular neighborhood of $D$ [15], our surface operator will equivalently correspond to a *flat* irreducible connection on a bundle $E'$ over $X$ whose field strength is $F' = F - 2\pi\alpha\delta_D$, where $F$ is the field strength of the bundle $E$ over $X$.[3] In other words, a supersymmetric surface operator will correspond to a gauge field solution over $X$ that satisfies

$$F = 2\pi\alpha\delta_D \tag{2.14}$$

along $D$. Indeed, the singular term in $A$ of (2.2) that is associated with the inclusion of an embedded surface, is such a solution. Thus, our physical definition of supersymmetric surface operators coincides with the mathematical definition of embedded surfaces.

Some comments on (2.14) are in order. Note that even though $\alpha$ is formally defined in (2.2) to be $\mathfrak{t}$-valued, we saw that it actually takes values in the maximal torus $\mathbb{T}$. Since $\mathbb{T} = \mathfrak{t}/\Lambda_{\text{cochar}}$, where $\Lambda_{\text{cochar}}$ is the cocharacter lattice of the underlying gauge group, (2.14) appears to be unnatural, since one is free to subject $F$ to a shift by an element of $\Lambda_{\text{cochar}}$. This can be remedied by lifting $\alpha$ in (2.14) from $\mathfrak{t}/\Lambda_{\text{cochar}}$ to $\mathfrak{t}$. Equivalently, this corresponds to a choice of an extension of the bundle $E$ over $D$ — something that was implicit in our preceding discussion.

*The "quantum" parameter $\eta$*

With an extension of the bundle $E$ over $D$, the restriction of the field strength $F$ to $D$ will be $\mathfrak{t}$-valued. Hence, we roughly have an abelian gauge theory in two dimensions along $D$. As such, one can generalize the physical definition of the surface operator, and introduce a two-dimensional theta-like angle $\eta$ as an additional "quantum" parameter which enters in the Euclidean path-integral via the phase

$$\exp\left(2\pi i \mathrm{Tr}\ \eta\mathfrak{m}\right), \tag{2.15}$$

[3]To justify this statement, note that the instanton number $\tilde{k}$ of the bundle $E$ over $X\backslash D$ is (in the mathematical convention) given by $\tilde{k} = k + 2\alpha l - \alpha^2 D \cap D$, where $k$ is the instanton number of the bundle $E$ over $X$ with curvature $F$, and $l$ is the monopole number (cf. equation (1.7) of [8]). On the other hand, the instanton number $k'$ of the bundle $E'$ over $X$ with curvature $F' = F - 2\pi\alpha\delta_D$ is (in the physical convention) given by $k' = -(1/8\pi^2)\int_X \mathrm{Tr} F' \wedge F' = k + 2\alpha l - \alpha^2 D \cap D$. Hence, we find that the expressions for $\tilde{k}$ and $k'$ coincide, reinforcing the notion that the bundle $E$ over $X\backslash D$ can be equivalently interpreted as the bundle $E'$ over $X$. Of course, for $F'$ to qualify as a nontrivial field strength, $D$ must be a homology cycle of $X$, so that $\delta_D$ (like $F$) is in an appropriate cohomology class of $X$.

where $\mathfrak{m} = \int_D F/2\pi$. Since $F$ restricted to $D$ is $\mathfrak{t}$-valued, and since the monopole number $l = \int_D F_L/2\pi$ is an integer, it will mean that $\mathfrak{m}$ must take values in the subset of diagonal, traceless $2 \times 2$ matrices — which generate the maximal torus $\mathbb{T}$ — that have *integer* entries only; i.e., $\mathfrak{m} \in \Lambda_{\text{cochar}}$. Also, values of $\eta$ that correspond to a nontrivial phase must be such that Tr $\eta\mathfrak{m}$ is *non-integral*. Because Tr $\mathfrak{m}'\mathfrak{m}$ is an integer if $\mathfrak{m}' \in \Lambda_{\text{cochar}}$, it will mean that $\eta$ must takes values in $\mathfrak{t}/\Lambda_{\text{cochar}} = \mathbb{T}$. Just like $\alpha$, one can shift $\eta$ by an element of $\Lambda_{\text{cochar}}$ whilst leaving the theory invariant.[4] Note that modular invariance requires that $\eta$ be set to *zero* if the surface operator is *nontrivially-embedded*; this condition is a crucial ingredient in the physical proof of KM's relation between the "ramified" and ordinary Donaldson invariants in [1].

*A point on nontrivially embedded surface operators*

More can also be said about the "classical" parameter $\alpha$ as follows. In the case when the surface operator is trivially embedded in $X$ — i.e., $X = D' \times D$ and the normal bundle to $D$ is hence trivial — the self-intersection number

$$D \cap D = \int_X \delta_D \wedge \delta_D \tag{2.16}$$

vanishes. On the other hand, for a non-trivially embedded surface operator supported on $D \subset X$, the normal bundle is nontrivial, and the intersection number is non-zero. The surface operator is then defined by the gauge field with singularity in (2.2) in each normal plane.

When the surface operators are non-trivially embedded, there is a condition on the allowed gauge transformations that one can invoke in the physical theory [17]. Let us explain this for when the underlying gauge group is $U(1)$ with gauge bundle $L$. Since there is a singularity of $2\pi\alpha\delta_D$ in the abelian field strength $F_L$ restricted to $D$, we find, using (2.16), that $\int_D F_L/2\pi = \alpha D \cap D \mod \mathbb{Z}$. Since $\int_D F_L/2\pi = l$ is always an integer, we must have

$$\alpha D \cap D \in \mathbb{Z}. \tag{2.17}$$

In fact, underlying the integrality of $l$ is actually the condition $c_1(L) \in H^2(X, \mathbb{Z})$. This implies that for *any* integral homology two-cycle $U \subset X$ (assuming, for simplicity, that $H_2(X, \mathbb{Z})$ is torsion-free), $-c_1(L)[U] =$

[4] This characteristic of $\eta$ is consistent with an $S$-duality in the corresponding, low-energy effective abelian theory, which maps $(\alpha, \eta) \to (\eta, -\alpha)$ [16].

$\int_U F_L/2\pi = \alpha(U \cap D)$ mod $\mathbb{Z}$ is always an integer; in other words, we must have

$$\alpha(U \cap D) \in \mathbb{Z}. \tag{2.18}$$

Now consider a gauge transformation — in the normal plane — by the following $U(1)$-valued function

$$(r, \theta) \to \exp(\theta u), \tag{2.19}$$

where $u \in \mathfrak{u}(1)$; its effect is to shift $\alpha \to \alpha + u$ whilst leaving the holonomy $\exp(2\pi\alpha)$ of the abelian gauge connection around a small circle linking $D$ which underlies the *effective* "ramification" of the theory, *unchanged.* Clearly, the only gauge transformations of this kind that can be globally-defined along $D$, are those whereby the corresponding shifts in $\alpha$ are compatible with (2.18). For effectively nontrivial $\alpha$, since $U \cap D \in \mathbb{Z}$, the relevant gauge transformations are such that $u \notin \mathbb{Z}$; in other words, $\exp(2\pi u) \neq 1$, and the gauge transformations are not single-valued under $\theta \to \theta + 2\pi$. Such twisted gauge transformations can certainly be defined for a non-simply-connected gauge group like $U(1)$.

*The effective field strength in the presence of surface operators*

In any gauge theory, supersymmetric or not, the kinetic term of the gauge field has a positive-definite real part. As such, the Euclidean path-integral (which is what we will eventually be interested in) will be non-zero if and only if the contributions to the kinetic term are strictly non-singular. Therefore, as a result of the singularity (2.14) when one includes a surface operator in the theory, the effective field strength in the Lagrangian that will contribute non-vanishingly to the path-integral must be a shifted version of the field strength $F$. In other words, whenever we have a surface operator along $D$, one ought to study the action with field strength $F' = F - 2\pi\alpha\delta_D$ instead of $F$. This means that the various fields of the theory are necessarily coupled to the gauge field $A'$ with field strength $F'$. This important fact was first pointed out in [17], and further exploited in [16] to prove an $S$-duality in a general, abelian $\mathcal{N} = 2$ theory without matter in the presence of surface operators.

## 2.3 A physical interpretation of the "ramified" Donaldson invariants

*Correlation functions of $\mathcal{Q}$-invariant observables*

Consider a topologically twisted version of a pure $SU(2)$ or $SO(3)$ theory with $\mathcal{N} = 2$ supersymmetry — also known as Donaldson–Witten theory —

in the presence of surface operators. This theory has a nilpotent scalar supercharge $\mathcal{Q}$, and its action can be written as [1]

$$S_E = \frac{\{\mathcal{Q}, V\}}{e^2} + \frac{\mathrm{i}\Theta}{8\pi^2}\int_X \mathrm{Tr}\ F' \wedge F' - \mathrm{i}\int_X \mathrm{Tr}\ \eta\delta_D \wedge F' \tag{2.20}$$

for some fermionic operator $V$ of $R$-charge -1 and scaling dimension 0, and complexified gauge coupling $\tau = 4\pi\mathrm{i}/e^2 + \Theta/2\pi$. The action is thus $\mathcal{Q}$-exact up to purely topological terms.

Now consider the set of $\mathcal{Q}$-invariant observables $\mathscr{O}_i$ and their correlation function

$$\langle \mathscr{O}_1 \cdots \mathscr{O}_n \rangle = \int \mathcal{D}\Phi\ \mathscr{O}_1 \cdots \mathscr{O}_n e^{-S_E}, \tag{2.21}$$

where $\mathcal{D}\Phi$ denotes the total path-integral measure in all fields. Note that one of the central features of the twisted theory is that its stress tensor $T_{\mu\nu}$ is $\mathcal{Q}$-exact, i.e., $T_{\mu\nu} = \{\mathcal{Q}, G_{\mu\nu}\}$ for some fermionic operator $G_{\mu\nu}$. Consequently, a variation of the correlation function with respect to the metric yields $\delta_{\bar{g}}\langle \mathscr{O}_1 \cdots \mathscr{O}_n \rangle = -\langle \mathscr{O}_1 \cdots \mathscr{O}_n \cdot \delta S_E/\delta \bar{g}^{\mu\nu} \rangle = -\langle \mathscr{O}_1 \cdots \mathscr{O}_n \cdot T_{\mu\nu} \rangle = -\langle \mathscr{O}_1 \cdots \mathscr{O}_n \cdot \{\mathcal{Q}, G_{\mu\nu}\} \rangle = -\langle \{\mathcal{Q}, \mathscr{O}_1 \cdots \mathscr{O}_n\, G_{\mu\nu}\} \rangle = 0$, where we have made use of the fact that $\langle \{\mathcal{Q}, \cdots\} \rangle = 0$ since $\mathcal{Q}$ generates a (super)symmetry of the theory. Notice also that a differentiation of the correlation function with respect to the gauge coupling $e$ yields $\partial/\partial e \langle \mathscr{O}_1 \cdots \mathscr{O}_n \rangle = 2/e^3 \langle \mathscr{O}_1 \cdots \mathscr{O}_n \{\mathcal{Q}, V\} \rangle = 2/e^3 \langle \{\mathcal{Q}, \mathscr{O}_1 \cdots \mathscr{O}_n\, V\} \rangle = 0$. In other words, the correlation function of $\mathcal{Q}$-invariant observables is independent of the gauge coupling $e$; as such, the semiclassical approximation to its computation will be exact. In this approximation, one can freely send $e$ to a very small value in the correlation function. Consequently, from (2.20) and (2.21), we see that the non-zero contributions to the correlation function will be centered around classical field configurations — or the zero-modes of the fields — which minimize $\{\mathcal{Q}, V\}$ and therefore the action. Thus, it suffices to consider quadratic fluctuations around these zero-modes.

Let us first consider the fluctuations. Assuming that the operators $\mathscr{O}_i$ can be expressed purely in terms of zero-modes, the path-integral over the fluctuations of the fields in the kinetic terms of the action give rise to determinants of the corresponding kinetic operators. Due to supersymmetry, the determinants resulting from the bose and fermi fields cancel up to a sign. (This point will be important when we physically derive Taubes' result in a later section).

Let us now consider the zero-modes. The bosonic zero-modes obey the constraints obtained by setting the supersymmetric variation of the fermi fields in the twisted theory to zero. We find that these constraints are

$F'_+ = 0$, $\nabla'\phi = 0$ and $[\phi, \phi^\dagger] = 0$, where $\nabla'$ is the gauge-covariant derivative and $\phi$ is a complex bose field [1]. If we assume the trivial solution $\phi = 0$ to the constraints $\nabla'\phi = 0$ and $[\phi, \phi^\dagger] = 0$, it will mean that the zero-modes of $A'$ do *not* correspond to reducible connections, and that there are *no* zero-modes of $\phi$. Altogether, this means that the only bosonic zero-modes come from the gauge field $A'$, and that they correspond to irreducible, anti-self-dual connections which are characterized by the relation

$$F_+ = 2\pi\alpha\delta_D^+. \tag{2.22}$$

Again, recall that the bundle $E'$ over $X$ with curvature $F' = F - 2\pi\alpha\delta_D$ can be equivalently viewed as the bundle $E$ with curvature $F$ restricted to $X\backslash D$. Hence, the constraint (2.22) just defines anti-self-dual connections on the bundle $E$ restricted to $X\backslash D$, whose holonomies around small circles linking $D$ are as given in (2.4). In other words, modulo gauge transformations that leave (2.22) invariant, the expansion coefficients of the zero-modes of $A'$ that appear in the path-integral measure will correspond to the collective coordinates on $\mathcal{M}'$ - the moduli space of "ramified" instantons.

The rest of the fields in the theory are given by the fermions $(\zeta, \chi^+_{\mu\nu}, \psi_\mu)$. Since we have restricted ourselves to connections $A'$ that are irreducible, and moreover, if we assume that they are also regular, it can be argued that $\zeta$ and self-dual $\chi^+$ do not have any zero-modes [1]. Thus, the only fermionic zero-modes whose expansion coefficients contribute to the path-integral measure come from $\psi$.

The number of bosonic zero-modes is, according to our analysis above, given by the dimension $s$ of $\mathcal{M}'$. What about the number of zero-modes of $\psi$? Well, since there are no zero-modes for $\zeta$ and $\chi^+$, the dimension of the kernel of the kinetic operator $\Delta_F$ which acts on $\psi$ in the ("ramified") Lagrangian is equal to the index of $\Delta_F$; in other word, the number of zero-modes of $\psi$ is given by $\dim(\mathrm{Ker}(\Delta_F)) = \mathrm{ind}(\Delta_F)$. This index also counts the number of infinitesimal connections $\delta A'$ where gauge-inequivalent classes of $A' + \delta A'$ satisfy $F'_+ = 0$, i.e., (2.22). Therefore, the number of zero-modes of $\psi$ will also be given by the dimension $s$ of $\mathcal{M}'$. Altogether, this means that after integrating out the non-zero modes, we can write the remaining part of the measure in the expansion coefficients $a'_i$ and $\psi_i$ of the zero-modes of $A'$ and $\psi$ as

$$\prod_{i=1}^{s} da'_i d\psi_i. \tag{2.23}$$

Notice that the $s$ distinct $d\psi_i$'s anti-commute. Hence, (2.23) can be interpreted as a natural measure for the integration of a differential form in $\mathcal{M}'$.

*Correlation functions corresponding to the "ramified" Donaldson invariants*

In the relevant case that $b_1(X) = 0$, one has the following $\mathcal{Q}$-invariant observables

$$I'_0(p) = \frac{1}{8\pi^2}\mathrm{Tr}\langle\phi(p)\rangle^2, \tag{2.24}$$

$$I'_2(S) = -\frac{1}{\sqrt{32}\pi^2}\int_S \mathrm{Tr}\left(\langle\phi\rangle F\right)_0 \tag{2.25}$$

for any $p \in H_0(X\backslash D)$ and $S \in H_2(X\backslash D)$; up to lowest order in $e$ in the semiclassical approximation,

$$\langle\phi(x)\rangle = -\frac{\mathrm{i}}{\sqrt{2}}\int_X d^4y\, G(x-y)[\psi(x), \psi(y)]_0, \tag{2.26}$$

where $G(x-y)$ is the unique solution to the relation $\nabla'^2 G(x-y) = \delta^4(x-y)$; the subscript "0" in (2.25) and (2.26) just denotes their restriction to zero-modes.

Like the assumption made of the operators $\mathscr{O}_i$ in (2.21), $I'_0(p)$ and $I'_2(S)$ are express purely in terms of zero-modes. Moreover, based on our above discussion about (2.23) being a natural measure for the integration of differential forms in $\mathcal{M}'$, we see that $I'_0(p)$ and $I'_2(S)$ (which contain 4 and 2 zero-modes of $\psi$, respectively) can be interpreted as 4-forms and 2-forms in $\mathcal{M}'$.

Let us compute an arbitrary correlation function in the $\mathcal{Q}$-invariant observables $I'_0(p)$ and $I'_2(S)$. For the correlation function to be non-vanishing, the $d\psi_i$'s in the remaining measure (2.23) have to be "soaked up" by the zero-modes of $\psi$ that appear in the combined operator whose correlation function we wish to consider, *exactly*. This just reflects the fact that a non-vanishing correlation function is necessarily $R$-invariant: the field $\psi$ carries a non-zero $R$-charge of 1, and under an $R$-transformation, the integration measure and an appropriately chosen combined operator will transform with weights $-\Delta R$ and $\Delta R$, respectively, where $\Delta R$ is the number $N_\psi$ of zero-modes of $\psi$. In turn, this means that the combined operator ought to correspond to a top-form (of degree $s$) in $\mathcal{M}'$. Therefore, if such a combined operator is given by $[I'_0(p)]^t I'_2(S_{i_1})\dots I'_2(S_{i_m})$,[5] its *topological* correlation

[5]Notice that we are considering a combined operator in which there are $t$ operators $I'_0(p_{i_1}),\dots,I'_0(p_{i_t})$ that coincide at one particular point $p$ in $X$. Such a combined operator can be consistently defined in any physical correlation function; this is because the $I'_0(p_{i_k})$'s consist only of non-interacting zero-modes and moreover, any correlation function of the topological theory is itself independent of the insertion points of the operators.

function — for "*ramified*" instanton number $k' = -\int_X p_1(E')/4$ — can be written as

$$\langle [I'_0(p)]^t I'_2(S_{i_1}) \dots I'_2(S_{i_m}) \rangle_{k'} = \int_{\mathcal{M}'} [I'_0(p)]^t \wedge I'_2(S_{i_1}) \wedge \dots \wedge I'_2(S_{i_m}), \quad (2.27)$$

where $2m + 4t = s$. This coincides with the definition of the "ramified" Donaldson invariants $d^{k'}_{m,t}$ in (2.8). Thus, we have found, in (2.27), a physical interpretation of the "ramified" Donaldson invariants in terms of the correlation functions of $\mathcal{Q}$-invariants observables $I'_0(p)$ and $I'_2(S)$. As a result, the generating function $\mathbf{Z}'_{\xi,\bar{g}}(p,S)$ in (2.9) can also be interpreted in terms of $I'_0$ and $I'_2$ as

$$\mathbf{Z}'_{\xi,\bar{g}}(p,S) = \sum_{k'} \langle e^{pI'_0 + I'_2(S)} \rangle_{k'}. \quad (2.28)$$

The "ramified" Donaldson series in (2.10) is then given by

$$\mathscr{D}'_{\xi}(S) = \sum_{k'} \left(1 + \tfrac{1}{2}\frac{\partial}{\partial p}\right) \cdot \langle e^{pI'_0 + I'_2(S)} \rangle_{k'}|_{p=0}. \quad (2.29)$$

### 2.4 The "ramified" SW equations and invariants

In the topologically twisted version of the corresponding low-energy SW theory with surface operators, the supersymmetric configurations correspond to solutions of the equations [1]

$$(F^d_L)_+ = (\bar{M}M)_+ + 2\pi\alpha_d \delta^+_D \quad (2.30)$$

and

$$\not{D} M = 0, \quad (2.31)$$

where $\not{D}$ is the Dirac operator *coupled* to the effective $U(1)$ photon with field strength $F^{d'}_L = F^d_L - 2\pi\alpha_d\delta_D$; the label "*d*" indicates that the field or parameter is that which is defined in the preferred *dual* "magnetic" frame; and $M$ is a section of the complex vector bundle $S_+ \otimes L'_d$, where $S_+$ and $L'_d$ are a positive spinor bundle and a $U(1)$-bundle with curvature field strength $F^{d'}_L$, respectively. (2.30) and (2.31) together define the "ramified" SW equations, whence the relevant *topological* correlation functions in the case where $b_1(X) = 0$ are

$$\langle [J^d_0(p)]^q \rangle_{\lambda'} = \int_{\mathcal{M}^{\lambda'}_{\rm sw}} [J^d_0(p)]^q = SW_{\lambda'}. \quad (2.32)$$

Here, $q = d^{\lambda'}_{\rm sw}/2$, where $d^{\lambda'}_{\rm sw}$ — the even dimension of the moduli space $\mathcal{M}^{\lambda'}_{\rm sw}$ of the "ramified" SW equations determined by the "ramified" first

Chern class $\lambda' = \frac{1}{2}c_1(L_d'^{\otimes 2})$ of the determinant line bundle $L_d'^{\otimes 2}$ of the Spin$^c$-structure associated with a choice of the complex vector bundle $S_+ \otimes L'_d$ — is given by [18]

$$d_{\text{sw}}^{\lambda'} = -\frac{2\chi + 3\sigma}{4} + (\lambda')^2. \tag{2.33}$$

Also

$$J_0^d(p) = \langle \varphi_d(p) \rangle = a_d, \tag{2.34}$$

where $p \in H_0(X)$,[6] and $\varphi_d$ is a complex scalar in the "magnetic" $\mathcal{N} = 2$ vector multiplet of the SW theory. As required, $J_0^d(p)$ is expressed purely in terms of non-fluctuating zero-modes; it has $R$-charge 2 (associated with an accidental $U(1)_R$ symmetry at low-energy) and consequently, it can be interpreted as a 2-form in $\mathcal{M}_{\text{sw}}^{\lambda'}$. Following [1], let us call $\text{SW}_{\lambda'}$ the "ramified" SW invariant for the basic class $\lambda'$.

*When X is of ("ramified") SW simple-type*

If $X$ is of "ramified" SW simple-type, i.e., $\dim(\mathcal{M}_{\text{sw}}^{\lambda'}) = 0$, we have $q = 0$ in (2.32), and as explained in [18], we have

$$\text{SW}_{\lambda'} = \text{SW}(\lambda') = \sum_{x_i} (-1)^{n_i}, \tag{2.35}$$

where the $x_i$'s are the points that span the zero-dimensional space $\mathcal{M}_{\text{sw}}^{\lambda'}$, and the $n_i$'s are integers which are determined by the corresponding sign of the determinant of an elliptic operator associated with a linearization of the "ramified" SW equations. In other words, $\text{SW}(\lambda')$ counts (with signs) the number of solutions of the "ramified" SW equations determined by $\lambda'$; in particular, it is an *integer*, just like its ordinary counterpart.

Recall that the ordinary limit whence there is effectively no "ramification" along $D$ is achieved when the effective value of $\alpha_d$ approaches an integer. Since the foregoing discussion holds for arbitrary values of $\alpha_d$, a four-manifold of "ramified" SW simple-type is necessarily of ordinary SW simple-type, too.

---

[6] Note that in contrast to the physical definition of the correlation functions that correspond to the "ramified" Donaldson invariants, here, we need not restricted the zero-cycles $p$ to $X \backslash D$. This is because the operator $a_d$ — unlike $F$, $F_L$ or $F_L^d$ — does not contain a singularity along $D$.

# 3 Physical proofs of seminal theorems by Ozsváth and Szabó

In this section, physical proofs of various seminal theorems in four-dimensional geometric topology obtained by Ozsváth and Szabó in [2,3], will be furnished. Our computations in this section will soon prove to be useful when we physically derive Taubes' spectacular result in the next section.

## 3.1 Adjunction inequality for embedded surfaces of negative self-intersection

In Corollary 1.7 of their seminal paper [2], Ozsváth and Szabó demonstrated that embedded surfaces of *negative* self-intersection in four-manifolds actually obey an adjunction inequality which involves the first Chern class of the Spin$^c$-structure. We will now present a physical proof of this mathematical corollary.

*A minimal genus formula from R-invariance*

Firstly, a relevant result from [1] is the following. Consider $X$ with $b_1 = 0$ and odd $b_2^+ > 1$; assume that $g \geq 1$, where $g$ is the genus of the surface operator $D \subset X$; then, for *any* $D \cap D \neq 0$, $R$-invariance of the non-vanishing correlation functions in (2.27) will imply that

$$2D \cap D - (2g-2) \leq 4l \leq (2g-2), \tag{3.1}$$

where $l = \int_D F_L/2\pi$.

Secondly, since $g \geq 1$ and therefore $(2g-2) \geq 0$, we can infer from (3.1) that

$$(2g-2) \geq D \cap D - 2l. \tag{3.2}$$

Moreover, note that due to electric–magnetic duality, $F_L$ is physically *equivalent* to the field strength $F_L^d$ of the low-energy $U(1)$ theory; in turn, $F_L^d$ corresponds to $-\pi c_1(L_d^2)$. In other words, we can identify $-2l$ with $c_1(L_d^2)[D]$. Consequently, we can write (3.2) as

$$(2g-2) \geq D \cap D + c_1(L_d^2)[D]. \tag{3.3}$$

The above formula coincides with Theorem 1.7(b) of [7].

*And the proof*

Now, let us consider a surface operator $D = \Sigma$ with $\Sigma \cap \Sigma \leq 0$ and genus $g \geq 1$. Since (3.1) is valid for any value of $D \cap D$, we can write

$$2\Sigma \cap \Sigma - (2g-2) \leq 4l \leq (2g-2). \tag{3.4}$$

Since $(2g-2) \geq 0$ and $\Sigma \cap \Sigma \leq 0$, we obtain from (3.4) the following inequality:

$$(2g-2) \geq \Sigma \cap \Sigma - c_1(L_d^2)[\Sigma], \tag{3.5}$$

after identifying $-2l$ with $c_1(L_d^2)[\Sigma]$.

As (3.3) is also valid for arbitrary values of $D \cap D$, it will mean that

$$(2g-2) \geq \Sigma \cap \Sigma + c_1(L_d^2)[\Sigma]. \tag{3.6}$$

Thus, if (3.5) and (3.6) are to hold simultaneously, it will mean that

$$|c_1(L_d^2)[\Sigma]| + \Sigma \cap \Sigma \leq (2g-2). \tag{3.7}$$

This is just Corollary 1.7 of [2]. In fact, our physical proof asserts that (3.7) should also hold for $\Sigma$ with $\Sigma \cap \Sigma = 0$, and not just for $\Sigma$ with $\Sigma \cap \Sigma < 0$ (as stipulated in Corollary 1.7 of [2]); this physical assertion is indeed consistent with Theorem 1.1 of [3] for $X$ of SW simple-type (for which (3.7) is also valid).

### 3.2 A relation among the ordinary SW invariants

Ozsváth and Szabó also showed in Theorem 1.3 of [2] and Theorem 1.6 of [3], that there exists relations among the ordinary SW invariants which arise from the above embedded surfaces with negative self-intersection in $X$ with $b_2^+(X) > 1$. We will now present the physical proofs of these mathematical theorems.

*The "magic" formula*

First, note that for $X$ with $b_1 = 0$ and $b_2^+ > 1$, the "magic" formula which expresses the generating function $Z'_{\text{D}}$ of the "ramified" Donaldson invariants in terms of the ("ramified") SW invariants when $\Sigma \cap \Sigma \neq 0$, is (via (7.20) and (7.15) of [1])

$$
\begin{aligned}
Z'_{\rm D} = \sum_{\bar\lambda} \frac{SW_{\bar\lambda}}{16} \cdot {\rm e}^{2{\rm i}\pi(\lambda_0\cdot\bar\lambda+\lambda_0^2)} \cdot {\rm e}^{2\bar\lambda[\widetilde\Sigma]} \\
\cdot \,{\rm Res}_{q_M=0}\left[\frac{dq_M}{q_M} q_M^{-\bar\lambda^2/2} \frac{\vartheta_2^{8+\sigma}}{a_d h_M}\left(2i\frac{a_d}{h_M^2}\right)^{(\chi+\sigma)/4}\right. \\
\left.\times \;\exp\left[2pu_M - {\rm i}(\bar\lambda, S)/h_M + S^2 T_S^M\right]\right] \\
+ \, i^{\{(\chi+\sigma)/4 - w_2(E)^2\}} \sum_{\lambda'} \frac{SW_{\lambda'}}{16} \cdot {\rm e}^{2{\rm i}\pi(\lambda_0\cdot\lambda+\lambda_0^2+\alpha^2\Sigma^2/2)} \cdot {\rm e}^{2\lambda[\widetilde\Sigma]} \\
\cdot \,{\rm Res}_{q_M=0}\left[\frac{dq_M}{q_M} q_M^{-(\lambda')^2/2} \frac{\vartheta_2^{8+\sigma}}{a_d h_M}\left(2{\rm i}\frac{a_d}{h_M^2}\right)^{(\chi+\sigma)/4}\right. \\
\left.\times \exp\left[-2pu_M + {\rm i}(\lambda, {\rm i}S)/h_M - S^2 T_S^M - 4\widetilde\Sigma^2 T_{\widetilde\Sigma}^M\right]\right], \qquad (3.8)
\end{aligned}
$$

where $\lambda' = \lambda - \alpha\delta_\Sigma$ is a "*ramified*" (first Chern class of the) Spin$^c$-structure for effectively nontrivial values of $\alpha$; $\vartheta_2^{8+\sigma}(\tau)$ is a certain Jacobi theta function in $q_M = {\rm e}^{2\pi{\rm i}\tau}$, while $a_d$, $u_M$, $h_M$, $T_S^M$ and $T_{\widetilde\Sigma}^M$ are polynomial functions in $q_M$ (see appendix A and Section 4.2 of [1] for their explicit expansions); $\widetilde\Sigma = {\rm i}\pi\alpha\Sigma/2$; $2\lambda_0$ is an integral lift of $w_2(E)$; and $\bar\lambda$ is an *ordinary* (first Chern class of the) Spin$^c$-structure.

Second, let us specialize to the case where the microscopic gauge group is $SU(2)$; i.e., $\xi = w_2(E) = 0$. In this case, $\lambda_0$ can be set to zero in the "ramified" and ordinary theories [1, 19]. Moreover, if we assume $X$ to be such that the values of $\bar\lambda^2/2$ and $(\lambda')^2/2$ of the first and second residues in (3.8), respectively, are both given by $(\chi+\sigma)/4 + \sigma/8$, the computation of (3.8) will simplify considerably; only the leading terms in the $q_M$-expansion contribute non-vanishingly. Using $u_M = 1 + \cdots$, $T_S^M = 1/2 + \cdots$, $T_{\widetilde\Sigma}^M = 1/4 + \cdots$, $h_M = 1/(2i) + \cdots$ and $a_d = 16iq_M + \cdots$, one will compute (3.8) to be

$$
\begin{aligned}
Z'_{\rm D} = 2^{1+\frac{7\chi}{4}+\frac{11\sigma}{4}} \left\{\sum_{\bar\lambda} SW_{\bar\lambda}\, {\rm e}^{2p+S^2/2}\, {\rm e}^{2(S+\widetilde\Sigma,\bar\lambda)} + {\rm i}^{(\chi+\sigma)/4}\right. \\
\left. \times \sum_{\lambda'} SW_{\lambda'}\,(-1)^{\alpha^2\Sigma^2}\, {\rm e}^{-2p-S^2/2-\widetilde\Sigma^2}\, {\rm e}^{-2{\rm i}(S+i\widetilde\Sigma,\lambda)}\right\}. \qquad (3.9)
\end{aligned}
$$

*On to the proofs*

Before we proceed further, note that since our objective is to provide a physical proof of a mathematical result, one needs to express (3.9) in the mathematical convention. One can do so by replacing $\alpha$ with $i\alpha$ and the relevant $U(1)$ field strengths $F$ with $\mathrm{i}F$, throughout; the gauge fields are then valued in the complex Lie algebra, as desired.

Coming back to our main discussion, let us send the effective value of $\alpha$ to $\pm 1$. Then, the condition $(\lambda - \mathrm{i}\alpha\delta_\Sigma)^2/2 = (\chi+\sigma)/4 + \sigma/8$ at the dyon point (i.e., the second contribution in $Z'_{\mathrm{D}}$) implies that we have $d_{L_d^2} = \lambda^2 - (2\chi + 3\sigma)/4 = \mp c_1(L_d^2)[\Sigma] + \Sigma \cap \Sigma$. As there is effectively no "ramification" in the $U(1)$ theory at the dyon point when $\alpha = \pm 1$, we can, via (2.32) and (2.34), write $\mathrm{SW}_{\lambda'}$ in (3.9) as

$$SW_{\lambda'} = \int_{\mathcal{M}_{\mathrm{sw}}^{\lambda'}} (a_d)^{d_{L_d^2}/2}, \tag{3.10}$$

where $\mathcal{M}_{\mathrm{sw}}^{\lambda'}$ is the moduli space of the ordinary SW equations whose even dimension is therefore $d_{L_d^2}$. As explained in footnote 6, $a_d$ can be defined at *any* point in $X$; thus, let us, for later convenience, define $a_d$ at some point $z \in \Sigma$.

When $\alpha = \pm 1$, there is also no "ramification" in the microscopic $SU(2)$ theory associated with $Z'_{\mathrm{D}}$ on the left-hand side (LHS) of (3.9); moreover, recall from Section 2.2 that modular-invariance requires that the phase term (2.15) be equal to 1 whenever the surface operators are nontrivially-embedded; as such, one can express $Z'_{\mathrm{D}}$ on the LHS of (3.9) as equation (7.25) of [1] via Witten's ordinary magic formula. Furthermore, since the condition $\bar{\lambda}^2/2 = (\chi+\sigma)/4 + \sigma/8$ implies that $X$ is of SW simple-type, one can denote $\mathrm{SW}_{\bar{\lambda}}$ as $\mathrm{SW}(\bar{\lambda})$ on the right-hand side (RHS) of (3.9). Last but not least, note that one can appeal to a regular gauge transformation of the kind in (2.19) with $u = \mp 1$ — *which leaves the gauge-invariant $\lambda'$ unchanged* — to shift $\alpha$ and therefore $\widetilde{\Sigma}$ to zero on the RHS of (3.9). Altogether, this means that one can also express (3.9) as

$$\begin{aligned}\sum_{\hat{\lambda}} &\left\{ SW(\hat{\lambda})\, \mathrm{e}^{2p+S^2/2}\, \mathrm{e}^{2(S,\hat{\lambda})} + \mathrm{i}^{(\chi+\sigma)/4}\, SW(\hat{\lambda})\, \mathrm{e}^{-2p-S^2/2}\, \mathrm{e}^{-2i(S,\hat{\lambda})} \right\} \\ &= \sum_{\lambda} \left\{ SW(\lambda)\, \mathrm{e}^{2p+S^2/2}\, \mathrm{e}^{2(S,\lambda)} + \mathrm{i}^{(\chi+\sigma)/4}\, SW_{\lambda'}\, \mathrm{e}^{-2p-S^2/2}\, \mathrm{e}^{-2\mathrm{i}(S,\lambda)} \right\}.\end{aligned} \tag{3.11}$$

In (3.11), $\hat{\lambda}$ is an ordinary Spin$^c$-structure; $\lambda' = \lambda \mp i\delta_\Sigma$; and $\lambda = -F_L^d/2\pi$, where $F_L^d$ is an ordinary $U(1)$ field strength, i.e., the holonomy of its gauge field around a small circle that links $\Sigma$ is trivial. Via a term-by-term comparison of (3.11), we conclude that we have an equivalence

$$SW_{\mathfrak{s}'} = SW(\mathfrak{s}) \tag{3.12}$$

of *ordinary* SW invariants, where $\mathfrak{s}' = -i\lambda'$ and $\mathfrak{s} = -i\lambda$ are the respective (first Chern class of the) Spin$^c$-structures expressed in the mathematical convention.

A few observations are in order. First, notice that if $d_{L_d^2} = |c_1(L_d^2)[\Sigma]| + \Sigma \cap \Sigma$, then $\mathfrak{s}' = \mathfrak{s} + \epsilon\delta_\Sigma$, where $\epsilon = \pm 1$ is the sign of $c_1(L_d^2)[\Sigma]$. Second, since the scalar variable $a_d$ has $R$-charge 2, it will represent a class in $H^2(\mathcal{M}_{\text{sw}}^{\mathfrak{s}'})$. Third, because we have assumed that $b_1(X) = 0$, it will mean that $H_1(X, \mathbb{Z})$ is empty. Fourth, as $d_{L_d^2} \geq 0$, it will mean that $|c_1(L_d^2)[\Sigma]| + \Sigma \cap \Sigma \geq 0$; together with (3.7), this implies that $g > 0$. With these four points in mind, it is thus clear that (3.10) and (3.12) are *precisely* Theorem 1.3 of [2]; here, $a_d$ and $d_{L_d^2}/2$ can be identified with $U$ and $(m + g)$ in Theorem 1.3 of [2], respectively, while $a$ in Theorem 1.3 of [2] can be set to 1 since $X$ is of SW simple-type.

*When $X$ is not of SW simple-type*

What if $X$ is not of SW simple-type? The analysis is similar: one just substitutes $\lambda^2/2$ and $(\lambda')^2/2$ as $(\chi + \sigma)/4 + \sigma/8 + p$ — where $p$ is some fixed positive integer — in the first and second residues of (3.8), respectively, and proceed as above. The only difference now is that the effective value of $d_{L_d^2}$ will be shifted by $2p$, and instead of (3.12), we will have

$$\text{SW}_{\mathfrak{s}'} = \text{SW}_{\mathfrak{s}}, \tag{3.13}$$

where

$$\text{SW}_{\mathfrak{s}'} = \int_{\mathcal{M}_{\text{sw}}^{\mathfrak{s}'}} (a_d)^{d_{L_d^2}/2} (a_d)^p \quad \text{and} \quad \text{SW}_{\mathfrak{s}} = \int_{\mathcal{M}_{\text{sw}}^{\mathfrak{s}}} (a_d)^p. \tag{3.14}$$

In this case, $a$ of [2] is no longer equal to 1 but rather, it is $U^p \in \mathbb{A}(X)$, where $\mathbb{A}(X)$ is the polynomial algebra $\mathbb{Z}[U]$. It is also clear from (3.14) that $2p$ must be equal to the dimension of the moduli space $\mathcal{M}_{\text{sw}}^{\mathfrak{s}}$ of the SW equations with Spin$^c$-structure $\mathfrak{s}$.

*When $b_2^+(X) = 1$*

And what if $b_2^+ = 1$? In this case, as explained in Section 6.3 of [1], the monopole and dyon point contributions to $Z'_D$ — which depend on $\mathrm{SW}_\lambda$ and $\mathrm{SW}_{\lambda'}$, respectively — will jump as we cross certain "walls" in the forward light cone $V_+ = \{\omega_+ \in H^{2,+}(X;\mathbb{R}) : (\omega_+)^2 > 0\}$. In particular, $\mathrm{SW}_\lambda$ will jump if we cross the "wall" defined by

$$(\omega_+, \lambda) = \frac{\mathrm{i}\alpha_M}{2}(\omega_+, \Sigma), \tag{3.15}$$

where $\alpha_M = -\eta$, while $\mathrm{SW}_{\lambda'}$ will jump if we cross the "wall" defined by

$$(\omega_+, \lambda) = \frac{\mathrm{i}\alpha_D}{2}(\omega_+, \Sigma), \tag{3.16}$$

where $\alpha_D = 2\alpha$.

Notice that when $\alpha = \pm 1$, we can rewrite (3.16) as

$$(\omega_+, \mathfrak{s} + \epsilon\delta_\Sigma) = 0, \tag{3.17}$$

where $\epsilon = \mp$.

Recall also that since modular-invariance requires $\eta$ to vanish whenever we have a nontrivially embedded surface operator such as $\Sigma$, the RHS of (3.15) is identically zero. This is tantamount to setting

$$(\omega_+, \Sigma) = 0. \tag{3.18}$$

Consequently, one can rewrite (3.15) as

$$(\omega_+, \mathfrak{s}) = 0. \tag{3.19}$$

Therefore, for (3.12) or (3.13) to continue to hold unambiguously when $b_2^+ = 1$, $\omega_+$ must not lie anywhere along the "walls" defined by (3.17) and (3.19) where the values of the LHS and RHS of (3.12) or (3.13), respectively, will jump; in addition, $\omega_+$ must also satisfy (3.18). This observation matches *exactly* the claim in Theorem 1.3 of [2] for four-manifolds with $b_2^+ = 1$. This completes our physical proof of Theorem 1.3 of [2].

*When $\Sigma$ has arbitrary self-intersection number*

Finally, note that (3.12) and its generalization (3.13) also hold for $\Sigma$ with *arbitrary* (as opposed to just negative) self-intersection number; this has been proved mathematically as Theorem 1.6 of [3] by Ozsváth and Szabó. Once more, one can furnish a physical proof of this theorem.

To this end, let us consider the case where $\alpha = -\epsilon = 1$. Then, $d_{L_d^2} = -c_1(L_d^2)[\Sigma] + \Sigma \cap \Sigma$. Since $d_{L_d^2} \geq 0$, we have

$$-c_1(L_d^2)[\Sigma] + \Sigma \cap \Sigma \geq 0. \tag{3.20}$$

On the other hand, since $d_{L_d^2}$ can be identified with $2(m+g)$ in [2] (and therefore [3]), as $2m \geq 0$, it will mean that $d_{L_d^2} - 2g \geq 0$. In turn, this implies that

$$-c_1(L_d^2)[\Sigma] + \Sigma \cap \Sigma + 2p > 2g - 2, \tag{3.21}$$

since $p$ is a positive integer.

As we have not appealed to (3.7) in deducing the above inequalities, the self-intersection number of $\Sigma$ is allowed to be arbitrary in (3.20) and (3.21). Consequently, after noting that $\mathfrak{s}' = \mathfrak{s} - \delta_\Sigma$ because $\epsilon = -1$, we find that (3.20) and (3.21), with (3.13) and (3.14), is nothing but Theorem 1.6 of [3]; here, $a_d$, $p$ and $(d_{L_d^2}/2 + p)$ can be consistently identified as $U$, $d$ and $d'$ in Theorem 1.6 of [3], respectively. When $b_2^+(X) = 1$, these relations will continue to hold as long as the above-stated conditions on $\omega_+$ are satisfied.

## 4 A physical derivation of Taubes' groundbreaking result

In a series of four long papers collected in [4], Taubes showed that on any compact, oriented symplectic four-manifold $X$ with $b_2^+ > 1$, the ordinary SW invariants are (up to a sign) equal to what is now known as the Gromov–Taubes invariants which count (with signs) the number of pseudo-holomorphic (complex) curves which can be embedded in $X$. This astonishing result, as formidable as its mathematical proof may be, nonetheless lends itself to a simple and concrete physical derivation, as we shall now demonstrate.

*Pseudo-holomorphic curves in a symplectic four-manifold*

Let $\omega_{\text{sp}}$ be a self-dual symplectic two-form on $X$ that is compatible with an almost-complex structure $J$. Let $K$ be the canonical line bundle on $X$. If a surface operator $\Sigma$ is a *pseudo-holomorphic* curve embedded in $X$ (in the sense of Gromov [20]), $J$ will map the tangent space of $\Sigma$ to itself. Moreover, we will have $\int_\Sigma \omega_{\rm sp} > 0$; in other words, $\Sigma$ will be homologically nontrivial so that the Poincaré dual $\delta_\Sigma$ of its fundamental class lies in $H^2(X, \mathbb{Z})$.

A connected $\Sigma$ is also known to satisfy the adjunction formula

$$2 - 2g + \Sigma \cap \Sigma = -c_1(K)[\Sigma]. \tag{4.1}$$

This implies that a flat torus with zero self-intersection number can potentially have multiplicity greater than 1; consequently, counting of such curves can be a delicate issue [4]. Therefore, for simplicity, let us choose $\Sigma$ to be curved with a *non-zero* self-intersection number. Such a choice is guaranteed by the fact that one can always find a basis of homology two-cycles $\{U_i\}_{i=1,\dots,b_2}$ in $X$ that has a purely diagonal, unimodular intersection matrix, whence $\Sigma$ can a priori be selected from the $b_2$ number of $U_i$'s with $\Sigma \cap \Sigma \neq 0$.

However, since being compatible with $J$ implies that $\omega_{\rm sp} \in H^{2,+}(X, \mathbb{R})$, together with $(\omega_{\rm sp}, \Sigma) > 0$, it will mean that there can be at most $b_2^+$ choices of $\Sigma$ among the $U_i$'s such that

$$\delta_\Sigma = \delta_\Sigma^+ \tag{4.2}$$

for some connected, non-multiply-covered, pseudo-holomorphic curve $\Sigma \subset X$ which obeys $\Sigma \cap \Sigma > 0$. (Exceptional spheres which may be multiply covered are also being automatically excluded here since they have negative self-intersections [21].)

*A particular instanton sector*

As in the previous section, let us now consider the case where $E$ can be lifted to an $SU(2)$-bundle, i.e., $w_2(E) = \xi = 0$. At low energies, the $SU(2)$ gauge symmetry is spontaneously-broken to a $U(1)$ gauge symmetry in the underlying physical theory. Mathematically, this means that $E$ can be expressed over all of $X$ as

$$E = L \oplus L^{-1} \tag{4.3}$$

at macroscopic scales, where $L$ is the complex line bundle corresponding to the unbroken $U(1)$ gauge symmetry. This implies that

$$c_2(E) = -c_1(L)^2. \tag{4.4}$$

However, since we have an equivalence of characteristic classes which are themselves topological invariants, it will mean that (4.4) will also hold in the microscopic $SU(2)$ theory; in particular, since $k$ and $l$ are given by $\int_X c_2(E)$ and $-\int_\Sigma c_1(L)$, respectively, the values of $k$ and $l$ will be correlated for any particular choice of $X$ and surface operator $\Sigma$.

Now consider the sector of the $SU(2)$ theory where $N_\psi$ is *zero*; this is the sector where[7]

$$k' = \tfrac{3}{16}(\chi + \sigma). \tag{4.5}$$

If $\mathbf{Z}^{p'}_{0,\bar{g}}(p, S)$ is the $p'$-instanton sector of the generating function (2.28) of the "ramified" $SU(2)$ Donaldson invariants of $X$, then

$$\mathbf{Z}^{k'}_{0,\bar{g}}(p, S) = \mathbf{Z}^{k'}_{0,\bar{g}}(0, 0) = \langle 1 \rangle_{k'}. \tag{4.6}$$

Let $\mathscr{D}^{p'}_0(S)$ be the $p'$-instanton sector of the "ramified" $SU(2)$ Donaldson series $\mathscr{D}'_0(S)$ in (2.29); then, from (4.6) and (2.29), we have

$$\mathscr{D}^{k'}_0(S) = \mathscr{D}^{k'}_0(0) = \langle 1 \rangle_{k'}. \tag{4.7}$$

At any rate, note that if $X$ is of ("ramified") SW simple-type, from (3.9), we have

$$\sum_{p'} \mathscr{D}^{p'}_0(S) = \sum_{\bar{\lambda}} SW(\bar{\lambda})\, \mathrm{e}^{2(S+\widetilde{\Sigma},\bar{\lambda}) + S^2/2 + f(\chi+\sigma)}, \tag{4.8}$$

where $f(\chi + \sigma)$ is a real-valued function in $\chi$ and $\sigma$. Two conclusions can be drawn from (4.8) at this point. First, since there is, as explained in Section 3.1, a one-to-one correspondence between $l$ and $-\int_\Sigma F^d_L/2\pi$ due to electric–magnetic duality in the low-energy $U(1)$ theory, it will mean — via the relation $k' = k + 2\alpha l - \alpha^2 \Sigma \cap \Sigma$,[8] and the correlation between $k$ and $l$ for any particular choice of $X$, $\Sigma \cap \Sigma$ and $\alpha$ — that there is also a one-to-one correspondence between $k'$ and a *certain* basic class $\lambda$.[9] Hence, $\mathscr{D}^{k'}_0(0)$ on the LHS of (4.8) will correspond to the $\lambda$-term on the RHS (4.8). Second, as our notation indicates, $\mathscr{D}^{k'}_0(0)$ is independent of $S$; also, in a supersymmetric topological quantum field theory whence the semiclassical approximation is exact, (4.7) will mean that the topological invariant $\mathscr{D}^{k'}_0(0)$ — like SW($\lambda$) of the $\lambda$-term on the RHS of (4.8) — is necessarily an *integer* (a fact that will be elucidated shortly). Consequently, the exponential factor in (4.8)

[7]Based on our discussions in Section 2.2, the expression for the index of the kinetic operator $\Delta_F$ of $\psi$ that counts the number $N_\psi$ of $\psi$ zero-modes, is the expression for the index in the ordinary case but with gauge bundle $E'$. In other words, $N_\psi = 8k' - \frac{3}{2}(\chi + \sigma)$, where the "ramified" instanton number $k' = -1/8\pi^2 \int_X \mathrm{Tr} F' \wedge F'$.

[8]Recall here that $k' = -1/8\pi^2 \int_X \mathrm{Tr} F' \wedge F' = k + \mathrm{Tr}\alpha\mathfrak{m} - (1/2)\,\mathrm{Tr}\alpha^2 \Sigma \cap \Sigma$, and since $\mathfrak{t}$ is generated by (2.3), it will mean that $k' = k + 2\alpha l - \alpha^2 \Sigma \cap \Sigma$.

[9]Recall here that $\lambda = -F^d_L/2\pi$. In fact, there is a one-to-one correspondence between all values of $p'$ and $\bar{\lambda}$ in (4.8), although the RHS of (4.8) is known to consist of a finite number of terms only because some of the SW($\bar{\lambda}$)'s are zero: cancellations can occur in the SW($\bar{\lambda}$)'s since they take the form given in (2.35).

will imply that $2(S,\lambda) + S^2/2 + f(\chi+\sigma) = 0$ and $(\widetilde{\Sigma}, 2\lambda) = \mathrm{i}\pi\mathbb{Z}$; the former condition will hold as long as the operator $I'_2(S)$ in (2.25) is normalized correctly[10] (a physical requirement that was implicit in our discussions hitherto), and the latter condition just reflects the fact that one is free to appeal to a "ramification"-preserving, twisted $U(1)$-valued gauge transformation (2.19) which shifts $\alpha$ in a way compatible with (2.18).[11] Altogether, it will mean that

$$\mathrm{SW}(\mathfrak{s}) = \langle 1 \rangle_{k'} \tag{4.9}$$

up to a sign, where $\mathfrak{s} = -i\lambda$ is the corresponding *ordinary* Spin$^c$-structure.

*The SW invariants are the Gromov–Taubes invariants*

What we would like to do now is to determine $\langle 1 \rangle_{k'}$ of (4.9) explicitly. To this end, first note that the parameter $\eta$ in $S_E$ of (2.20) must be set to zero since we are considering non-trivially embedded surface operators $\Sigma$; then, via a chiral rotation of the massless fermions in the theory which inconsequentially shifts $\Theta$ in (2.20) to a convenient value, we can write

$$\begin{aligned} S_E = \frac{1}{e^2}\int_X d^4x\,\sqrt{g}\,\mathrm{Tr}\bigg[&\frac{1}{4}F'_{\mu\nu}F^{\mu\nu\,\prime} + \frac{1}{2}\phi\nabla'_\mu\nabla^{\mu\,\prime}\phi^\dagger - \mathrm{i}\zeta\nabla'_\mu\psi^\mu - \mathrm{i}\chi_+^{\mu\nu}\nabla'_\mu\psi_\nu \\ &- \frac{\mathrm{i}}{8}\phi[\chi^+_{\mu\nu}, \chi_+^{\mu\nu}] - \frac{\mathrm{i}}{2}\phi^\dagger[\psi_\mu, \psi^\mu] - \frac{\mathrm{i}}{2}\phi[\zeta,\zeta] - \tfrac{1}{8}[\phi,\phi^\dagger]^2\bigg]. \end{aligned} \tag{4.10}$$

Second, let $\Phi = (A', \phi, \phi^\dagger)$ and $\Psi = (\zeta, \chi^+, \psi)$ represent the bosonic and fermionic fields of the theory, respectively. Since the semiclassical approximation is exact, one can expand $S_E$ to lowest order in $e$ when terms beyond quadratic order in the non-zero modes $\tilde{\Phi}$ and $\tilde{\Psi}$ need *not* be considered; as such, because there are, for $N_\psi = 0$, no zero-modes of $\psi$ and $(\phi, \phi^\dagger, \zeta, \chi^+)$ (as explained in Section 2.3), one can ignore the non-kinetic terms in (4.10)

---

[10] For any $\bar{S} \in H_2(X\backslash\Sigma, \mathbb{R})$, let $r$ be the correct normalization factor (of the classical zero-mode wavefunction) of the operator $I'_2(\bar{S})$; from (2.25), it will mean that $I'_2(S)$ — where $S = r\cdot\bar{S}$ — is a correctly normalized version of $I'_2(\bar{S})$; the condition $2(S,\lambda) + S^2/2 + f(\chi+\sigma) = 0$ can then be written as $ar^2 + br + c = 0$, where $a = \bar{S}^2$, $b = 4(\bar{S},\lambda)$ and $c = 2f(\chi+\sigma)$ are real constants for any particular choice of $X$ and $\bar{S}$ such that a solution of $r$ can always be found — in other words, the condition $2(S,\lambda) + S^2/2 + f(\chi + \sigma) = 0$ will hold if the operator $I'_2(S)$ is normalized correctly, and vice-versa.

[11] Since $2\lambda \in H^2(X,\mathbb{Z})$, we can write $\alpha(\Sigma, 2\lambda) = \alpha\sum_i(\Sigma, U_i)$, where $U_i \in H_2(X,\mathbb{Z})$ and $(\Sigma, U_i) \in \mathbb{Z}$. Via (2.19), one can shift $\alpha$ to satisfy (2.18); in particular, one can regard $\alpha(\Sigma, 2\lambda)$ as an even integer so that $(\widetilde{\Sigma}, 2\lambda) = \mathrm{i}\pi\alpha(\Sigma,2\lambda)/2 = \mathrm{i}\pi\mathbb{Z}$ under such a gauge symmetry.

which are beyond quadratic order in $\tilde{\Phi}$ and $\tilde{\Psi}$; thus, we can write

$$S_E = \int_X d^4x \sqrt{g} \left( \tilde{\Phi} \Delta_B \tilde{\Phi} + \tilde{\Psi} \Delta_F \tilde{\Psi} \right), \tag{4.11}$$

where $\Delta_B$ and $\Delta_F$ are certain second and first-order elliptic operators, respectively. Hence, the Gaussian integrals over $\tilde{\Phi}$ and $\tilde{\Psi}$ will be given by

$$\frac{\det(\Delta_F)}{\sqrt{\det(\Delta_B)}}. \tag{4.12}$$

Note at this point that due to supersymmetry, there is a pairing of the excitations of the fields $\Phi$ and $\Psi$ at every non-zero energy level. Moreover, it is a fact that $\det(\Delta_F) = \det(\Delta_B^{1/2})$ (after one fixes a sign ambiguity by specifying an orientation of the underlying moduli space $\mathcal{M}'$ of "ramified" instantons). Consequently, we have

$$\Delta_F \tilde{\Psi}_n = \xi_n \tilde{\Psi}_n \tag{4.13}$$

and

$$\Delta_B \tilde{\Phi}_n = \xi_n^2 \tilde{\Phi}_n, \tag{4.14}$$

where the subscript "$n$" refers to the $n$th energy level with corresponding real eigenvalue $\xi_n \neq 0$. Therefore, one can compute (4.12) to be

$$\prod_n \frac{\xi_n}{\sqrt{|\xi_n|^2}} = \pm 1 = \text{sign}(\det \Delta_F). \tag{4.15}$$

Third, recall from our discussions in Section 2.3 that the non-vanishing contributions (4.15) to $\langle 1 \rangle_{k'}$ localize around "ramified" instantons which satisfy (2.22). Moreover, according to our discussions in Section 2.3, since there are no $\zeta$ and $\chi^+$ zero-modes, we have $\text{ind}(\Delta_F) = \dim(\mathcal{M}')$ and $\ker(\Delta_F) = T\mathcal{M}'$.

Let us now send the effective value of $\alpha$ to $+1$; i.e., we now have an "ordinary" surface operator along $\Sigma$ whence the relation (4.9) is exact.[12] Then, from the above three points, the fact that $N_\psi = \dim(\mathcal{M}') = 0$, and

[12]When $\alpha = +1$, one can (as was done earlier in computing (3.11)) set $\widetilde{\Sigma} = i\pi\alpha\Sigma/2$ to zero in the sign $e^{(\widetilde{\Sigma}, 2\lambda)}$ of (4.9) via the gauge transformation (2.19) where $u = -1$.

the relations (4.15), (2.22) and (4.2), one can conclude that in this case,

$$\langle 1 \rangle_{k'} = \sum_x \mathrm{sign}(\det \mathcal{D}), \tag{4.16}$$

where $\mathcal{D}$ is a certain first-order elliptic operator whose kernel is the tangent space to the space $H$ of solutions to the relation

$$c_1(\mathcal{E}) = \delta_\Sigma. \tag{4.17}$$

In (4.16), the $x$'s are just the points which span the space $H$ of dimension *zero*; in (4.17), $\mathcal{E}$ is some nontrivial complex line bundle with a self-dual $\mathfrak{u}(1)$-valued connection $A_\mathcal{E}$ and curvature $c_1(\mathcal{E})$: recall from our discussion in Section 2.2 that a choice of an extension of $E$ over $\Sigma$ results in $\alpha$ being $\mathfrak{t}$-valued, and since the maximal torus $\mathbb{T}$ of $SU(2)$ is actually $U(1)$, $\alpha$ and therefore $F_+$ (i.e., $2\pi c_1(\mathcal{E})$) are actually $\mathfrak{u}(1)$-valued in (2.22). Such a complex line bundle $\mathcal{E}$ — where $c_1(\mathcal{E}) \cdot c_1(\mathcal{E}) > 0$ — can always be found, as $b_2^+ > 1$. Since $\mathrm{sign}(\det \mathcal{D}) = \pm 1$, (4.16) will imply that $\langle 1 \rangle_{k'}$ is an integer, consistent with (4.9).

Notice that the relation (4.17) means that one can interpret $H$ as the space of pseudo-holomorphic curves in $X$ whose Poincaré dual is $c_1(\mathcal{E})$; hence, with the above description of the kernel of the first-order elliptic operator $\mathcal{D}$, one can further conclude that

$$\sum_x \mathrm{sign}(\det \mathcal{D}) = \mathrm{Gr}(c_1(\mathcal{E})), \tag{4.18}$$

where $\mathrm{Gr}(c_1(\mathcal{E}))$ is the Gromov–Taubes invariant defined in [22] for a connected, non-multiply-covered, pseudo-holomorphic curve in $X$ whose fundamental class is Poincaré dual to $c_1(\mathcal{E})$. Since $\mathrm{Gr}(c_1(\mathcal{E}))$ depends only on the homology class of $\Sigma$, it will be invariant under smooth deformations of the metric and complex structure on $\Sigma$; i.e., $\mathrm{Gr}(c_1(\mathcal{E}))$ is also a two-dimensional topological invariant of $\Sigma$.

On the other hand, because (4.9) is valid for $X$ of SW simple-type, i.e., $\lambda^2 - (2\chi + 3\sigma)/4 = d_{L_d^2} = -c_1(L_d^2)[\Sigma] + \Sigma \cap \Sigma = 0$, we find that (3.12), (3.10) and $\mathfrak{s}' = \mathfrak{s} - \delta_\Sigma$ will imply that the LHS of (4.9) is $\mathrm{SW}(\mathfrak{s} - \delta_\Sigma)$. In fact, since the non-vanishing contributions to the RHS of (4.9) localize around supersymmetric field configurations which obey (4.17), the LHS of (4.9) can actually be written as $\mathrm{SW}(\mathfrak{s} - c_1(\mathcal{E}))$; as a result, by (4.18), (4.16) and (4.9), we have

$$\mathrm{SW}(\mathfrak{s} - c_1(\mathcal{E})) = \mathrm{Gr}(c_1(\mathcal{E})). \tag{4.19}$$

In any case, because $\Sigma$ is such that $\Sigma \cap \Sigma > 0$, it must satisfy (cf. (3.3))

$$2 - 2g + \Sigma \cap \Sigma \leq -c_1(L_d^2)[\Sigma], \tag{4.20}$$

in addition to (4.1); consequently, we necessarily have $\mathfrak{s} = \frac{1}{2}c_1(L_d^2) = \frac{1}{2}c_1(K)$. By noting that as $b_2^+ > 1$, the ordinary SW invariants satisfy $\mathrm{SW}(\bar{\mathfrak{s}}) = \pm\mathrm{SW}(-\bar{\mathfrak{s}})$ for any ordinary Spin$^c$-structure $\bar{\mathfrak{s}}$ [23], we can also write (4.19) as

$$\mathrm{SW}(\hat{\mathfrak{s}}) = \pm\mathrm{Gr}(c_1(\mathcal{E})), \tag{4.21}$$

where

$$\hat{\mathfrak{s}} = \tfrac{1}{2}c_1(\mathscr{L}), \tag{4.22}$$

and

$$\mathscr{L} = K^{-1} \otimes \mathcal{E}^2. \tag{4.23}$$

Moreover, since $c_1(L_d^2) = c_1(K)$, the condition $d_{L_d^2} = 0$ can also be expressed as

$$-c_1(K) \cdot c_1(\mathcal{E}) + c_1(\mathcal{E}) \cdot c_1(\mathcal{E}) = 0. \tag{4.24}$$

Because the LHS of (4.24) is the dimension of $H$ as defined mathematically in [22], we see that (4.24) is indeed consistent with the fact that $H$ is zero-dimensional as implied by $N_\psi = 0$. Moreover, (4.24) and (4.1) together imply that the genus of the pseudo-holomorphic curve represented by $c_1(\mathcal{E})$ will be given by

$$g = 1 + c_1(\mathcal{E}) \cdot c_1(\mathcal{E}). \tag{4.25}$$

Finally, note that (4.21)–(4.25) are *precisely* Theorem 4.1 and Propositions 4.2–4.3 of [24] which summarizes the results collected in [4]! This completes our physical derivation of Taubes' groundbreaking result that the ordinary SW invariants are (up to a sign) equal to the Gromov–Taubes invariants on any compact, oriented symplectic four-manifold with $b_2^+ > 1$.

## 5 Mathematical implications of the underlying physics

Now that we have physically re-derived the above mathematically established theorems by Ozsváth–Szabó and Taubes, one might wonder if the physics can, in turn, offer any new and interesting mathematical insights. Indeed it can, as we shall now elucidate.

### 5.1 The Gromov–Taubes and "ramified" SW invariants

Assume that $X$ is a compact, oriented, symplectic four-manifold which contains at least one trivially embedded curve that is connected; assume also that $b_1(X) = 0$ and $b_2^+(X) > 1$. Then, from (4.21), and (9.9) of [1], we find that

$$SW(\hat{\mathfrak{s}}_{\mathrm{r}} + \alpha_{\mathrm{r}}\delta_{\mathfrak{D}}) = \pm \mathrm{Gr}(c_1(\mathcal{E})). \tag{5.1}$$

Here, $\mathrm{SW}(\hat{\mathfrak{s}}_{\mathrm{r}} + \alpha_{\mathrm{r}}\delta_{\mathfrak{D}})$ is a "ramified" SW invariant; in addition, the gauge field underlying $\hat{\mathfrak{s}}_{\mathrm{r}}$ picks up a *non-trivial* holonomy — parameterized by a non-integer $\alpha_{\mathrm{r}}$ — as one traverses a closed loop linking a connected curve $\mathfrak{D}$ that is *trivially-embedded* in $X$. In other words, the Gromov–Taubes invariants — which count the connected, non-multiply covered, pseudo-holomorphic curves with positive self-intersection and fundamental class Poincare dual to $c_1(\mathcal{E})$ — are (up to a sign) equal to the "ramified" SW invariants of $X$!

*A rigorous mathematical proof?*

Let us now attempt to explain why (5.1) ought to be amenable to a rigorous mathematical proof. First, note that the presence of a bona-fide "ramification" along $\mathfrak{D}$ implies that the LHS of (5.1) counts (with signs) the number of solutions to the "ramified" SW equations which are defined (in the mathematical convention) by

$$F_+ = (\bar{M}M)_+ - \mu \tag{5.2}$$

and

$$\not{D}M = 0. \tag{5.3}$$

Here, $F$ is an imaginary-valued, curvature two-form given by $F = -\mathrm{i}\pi c_1$ $(\mathscr{L}_F)$, where $\mathscr{L}_F = \mathscr{L} \otimes \mathscr{L}_{\mathfrak{D}}^{2\alpha_{\mathrm{r}}}$; the complex line bundle $\mathscr{L}_{\mathfrak{D}}$ is such that $c_1(\mathscr{L}_{\mathfrak{D}})$ is the Poincaré dual $[\mathfrak{D}]$ of $\mathfrak{D}$; $\mu = i\epsilon\delta_{\mathfrak{D}}^+$, where $\epsilon$ is a positive real constant, is a fixed, imaginary-valued, self-dual two-form on $X$ that cannot be set to zero; and $M$ is a section of the complex vector bundle $S_+ \otimes \mathscr{L}^{1/2}$, where $\mathscr{L}$ is as given in (4.23).

Second, notice that (5.2) is the same as

$$F_+ = (\bar{M}'M')_+ - \mu, \tag{5.4}$$

where $\bar{M}' = \bar{M}\,\mathrm{e}^{\mathrm{i}\alpha_{\mathrm{r}}\theta}$ and $M' = \mathrm{e}^{-\mathrm{i}\alpha_{\mathrm{r}}\theta}M$ are gauge-transformed versions of $\bar{M}$ and $M$, respectively. Notice also that (5.3) is the same as

$$\not{D}'M' = 0, \tag{5.5}$$

where (assuming a small but non-zero $\alpha_r$) $\not{D}' M' = \mathrm{e}^{-\mathrm{i}\alpha_\mathrm{r}\theta} \not{D} M$ such that $M'$ is a section of the complex vector bundle $S_+ \otimes \mathscr{L}_F^{1/2}$. Altogether, this means that one can interpret the "ramified" SW equations of (5.2)–(5.3) as the ordinary, *perturbed* SW equations of (5.4)–(5.5) with perturbation two-form $\mu$ and Spin$^c$-structure $\mathfrak{s}_\mathrm{o} = \frac{1}{2} c_1(\mathscr{L}_F)$. Moreover, via (4.23), one can also write $\mathscr{L}_F = K^{-1} \otimes \mathcal{E}_F^2$, where $\mathcal{E}_F = \mathcal{E} \otimes \mathscr{L}_{\mathfrak{D}}^{\alpha_\mathrm{r}}$.

Third, note that since $\mathfrak{D} \cap \mathfrak{D} = 0$, it will mean that $\mathfrak{D} \neq \Sigma$, where $\Sigma \subset X$ is the connected, psuedo-holomorphic curve introduced at the start of Section 4 with *positive* self-intersection. In particular, $[\mathfrak{D}]$ and $[\Sigma]$ are necessarily distinct. Consequently, since $c_1(\mathcal{E}) = [\Sigma]$, it must be that $\mathscr{L}_{\mathfrak{D}} \neq \mathcal{E}$.

Fourth, notice that for a fixed $\mathfrak{D}$ and $\alpha_\mathrm{r}$, the map $\Sigma \to [\Sigma]$ is potentially many-to-one while the map $\mathfrak{D} \to [\mathfrak{D}]$ is necessarily one-to-one. As such, there is a one-to-one correspondence between (pseudo-holomorphic) curves in $X$ whose Poincaré duals are $c_1(\mathcal{E}_F) = [\Sigma] + \alpha_\mathrm{r}[\mathfrak{D}]$ and $c_1(\mathcal{E}) = [\Sigma]$.

Last but not least, note that $\mu = \mathrm{i}\epsilon\delta_{\mathfrak{D}}^+$ in (5.4) is singular: this is because $\delta_{\mathfrak{D}}$ is actually a delta two-form. Hence, according to Taubes' analysis in [4], each solution to (5.4)–(5.5) ought to determine a pseudo-holomorphic curve in $X$ that is Poincaré dual to $c_1(\mathcal{E}_F)$. Consequently, the above-observed one-to-one correspondence between pseudo-holomorphic curves in $X$ whose Poincaré duals are $c_1(\mathcal{E}_F)$ and $c_1(\mathcal{E})$, will imply that each solution to (5.4)–(5.5) ought to correspond to a pseudo-holomorphic curve in $X$ that is Poincaré dual to $c_1(\mathcal{E})$. This conclusion is indeed consistent with (5.1).

### 5.2 Certain identities among the Gromov–Taubes invariants

Assume that $X$ is a compact, oriented, symplectic four-manifold with $b_1 = 0$ and $b_2^+ > 1$. Then, from (4.9), (4.16) and (4.18), we have

$$\mathrm{SW}(\mathfrak{s}) = \mathrm{Gr}(c_1(\mathcal{E})), \tag{5.6}$$

where as explained in Section 4, we necessarily have

$$\mathfrak{s} = \tfrac{1}{2} c_1(K^{-1} \otimes K^2) \tag{5.7}$$

if $c_1(\mathcal{E})$ is the Poincaré dual of the pseudo-holomorphic curve $\Sigma$.

Via (5.6), (5.7) and (4.21)–(4.24), we find that

$$\mathrm{Gr}(c_1(K)) = \pm \mathrm{Gr}(c_1(\mathcal{E})). \tag{5.8}$$

Note that the dimension of the space of pseudo-holomorphic curves associated with the LHS of (5.8) is given by (4.24), albeit with $\mathcal{E}$ replaced by $K$; in particular, it is zero, just like the dimension of the space of pseudo-holomorphic curves associated with the RHS of (5.8). In this sense, (5.8) can be viewed as a consistent relation. But can we say more? Most certainly.

First, it is clear that (5.8) implies that there exists pseudo-holomorphic curves in $X$ which are Poincaré dual to $c_1(K)$. Since pseudo-holomorphic curves (in a symplectic four-manifold) are automatically symplectic [4], it will mean that the Poincaré dual of $c_1(K)$ can be represented by a fundamental class of an embedded symplectic curve in $X$; this conclusion is just Theorem 0.2 in article 1 of [4]. Moreover, this conclusion also implies via (4.25) that if there are no embedded spheres in $X$ with self-intersection $-1$, then $c_1(K) \cdot c_1(K) \geq 0$; this observation agrees with Proposition 4.2 of [24].

Second, (5.8) also implies that $c_1(\mathcal{E})$ is represented by at least one pseudo-holomorphic curve in $X$ — a fact that is well-established in the mathematical literature [25].

In any event, the above mathematical assertions depend squarely on the non-vanishing of $\mathrm{Gr}(c_1(\mathcal{E}))$; thus, one can understand them to be a consequence of $R$-invariance: $R$-invariance of the topological partition function $\langle 1 \rangle_{k'}$ of the $k'$-instanton sector asserts that it will not vanish, and from (4.16) and (4.18), we see that $\mathrm{Gr}(c_1(\mathcal{E}))$ will not vanish either.

Notice also that since $\mathrm{Gr}(0) = 1$ by definition [4], we have $\mathrm{SW}(\mathfrak{s}) = 1$ from (4.19). Consequently, (5.6) will mean that

$$\mathrm{Gr}(c_1(\mathcal{E})) = +1. \tag{5.9}$$

In other words, the number of points in the zero-dimensional space $H$ of pseudo-holomorphic curves $\Sigma \subset X$ which are positively oriented is greater than the number which are negatively-oriented by *one*. In fact, (5.9) is consistent with the relation $\mathrm{Gr}_0(A) = \pm 1$ proved in Proposition 3.18 of [21], while (5.8) — in light of (5.9) — is consistent with the relation $|\mathrm{Gr}(K)| = 1$ proved in Theorem 3.10 of [21].

Last but not least, since $\mathrm{SW}(\bar{\mathfrak{s}}) = \pm SW(-\bar{\mathfrak{s}})$ for any ordinary Spin$^c$-structure $\bar{\mathfrak{s}}$, from (4.21)–(4.23) and (5.8), we find that

$$\mathrm{Gr}(c_1(K)) = \pm \mathrm{Gr}(c_1(K) - c_1(\mathcal{E})). \tag{5.10}$$

Note that (5.10) is distinct from the widely-known result of Serre–Taubes duality for pseudo-holomorphic curves in $X$ [25] (which one can nevertheless obtain from (5.10) by making the substitution (5.8)).

In summary, for *connected*, *non-multiply covered*, pseudo-holomorphic curves in $X$ that have *positive* self-intersection and fundamental class Poincaré dual to $c_1(\mathcal{E})$, the underlying physics suggests that the relations (5.8), (5.9) and (5.10) ought to hold in addition to those which have already been established in the mathematical literature.

### 5.3 Affirming a knot homology conjecture by Kronheimer and Mrowka

Assume that *general* $X = M \times \mathbf{S}^1$, where $M$ is a compact, oriented three-manifold, and $b_1(M) = b_2^+(X) > 1$. Recall that the effective Lagrangian of the topological $\mathcal{N} = 2$ pure $SU(2)$ gauge theory with an arbitrarily embedded surface operator $\Sigma$, is just the Lagrangian of the ordinary Donaldson–Witten theory with gauge field $A'$ and field strength $F' = F - 2\pi i \alpha \delta_\Sigma$. As such, one can conclude from the analysis in [26] that up to $\mathcal{Q}$-exact terms which are thus irrelevant, $S_E$ in (4.10) is the action[13] for a supersymmetric quantum mechanical sigma model with target manifold $\mathcal{A}'/\mathcal{G}'$ — the space of all gauge-inequivalent classes of "ramified" $SU(2)$-connections $\mathscr{A}'$ on $M$, and potential $h = \frac{1}{2}\int_M \mathrm{Tr}\,(\mathscr{A}' \wedge d\mathscr{A}' + \frac{2}{3}\mathscr{A}' \wedge \mathscr{A}' \wedge \mathscr{A}')$ — the Chern–Simons functional of $\mathscr{A}'$. Specifically, $\mathscr{A}'$ can be regarded a gauge connection of an $SU(2)$-bundle over $M\backslash\Sigma_M$ — where $\Sigma_M \subset \Sigma$ is the component of $\Sigma$ embedded in $M$ — whose holonomy around a small circle linking $\Sigma_M$ in $M$ is $\exp(2\pi i\alpha)$. Moroever, we now have in the supersymmetry algebra a Hamiltonian operator $H$ which generates translations in the "time" direction along $\mathbf{S}^1$, and a second nilpotent supercharge $\bar{\mathcal{Q}}$. In particular, they obey $[H, \mathcal{Q}] = [H, \bar{\mathcal{Q}}] = 0$ and $\{\mathcal{Q}, \bar{\mathcal{Q}}\} = 2H$; consequently, one can easily show that the ground states of the theory are supersymmetric, i.e., they must be annihilated by both $\mathcal{Q}$ and $\bar{\mathcal{Q}}$, and that they are in the $\mathcal{Q}$-cohomology. (See Chapter 10 of [27] for an excellent review of this and other assertions to be made momentarily.)

What we would like to do now is to compute the partition function $\langle 1 \rangle_{k'}$ of the theory via the supersymmetric quantum mechanical sigma model on $\mathcal{A}'/\mathcal{G}'$. Since the presence of $\mathbf{S}^1$ in $X$ enforces a periodic boundary condition on the (fermi) fields of the theory, the path-integral of the sigma model without operator insertions, i.e., $\langle 1 \rangle_{k'}$, will be given by the Witten index $\mathrm{Tr}(-1)^F$, where $F$ is the fermion number. In turn, $\mathrm{Tr}(-1)^F$ is given by

[13] Recall from Section 2.2 that unless the surface operator is nontrivially embedded, there is *no* restriction on the effective values that its $\eta$-parameter can take in order to preserve modular invariance in the corresponding low-energy SW theory. As such, let us for simplicity, take $\eta$ to be zero in our following analysis; then, $S_E$ in (4.10) will be the relevant action regardless of the embedding of the surface operator in $X$.

the Euler characteristic of the $\mathcal{Q}$-complex generated by the $\mathcal{Q}$-cohomology groups, i.e., the supersymmetric ground states.

As first pointed out by Atiyah in [11], in the case that one has an ordinary $SU(2)$ connection $\mathscr{A}$ on (a homology three-sphere) $Y$, the ground states of the corresponding Hamiltonian are, purely formally, the instanton Floer homology groups $HF_*(Y)$ of $Y$ defined by Floer [28]. Analogously, as first suggested in [29], one can formally identify the ground states of $H$ as the "ramified" instanton Floer homology groups $HF_*(M;\Sigma_M;\alpha)$ of $M$; $\langle 1\rangle_{k'}$ will then be given by the Euler characteristic of the "ramified" instanton Floer homology of $M$. In other words, we have

$$\langle 1\rangle_{k'} = \chi(HF_*(M;\Sigma_M;\alpha)). \tag{5.11}$$

Moreover, it was also verified in [1] that the $R$-symmetry under which $\mathcal{Q}$ has charge 1 is only conserved mod 8; as a result, the "ramified" instanton Floer complex, like its ordinary counterpart, has a mod 8 grading under this $R$-symmetry. Nevertheless, unlike its ordinary counterpart, its relative grading is defined mod 4 instead of mod 8.[14]

That (5.11) is a consistent relation can be seen as follows. First, note that $\chi(M \times \mathbf{S}^1) = \chi(M)\chi(\mathbf{S}^1) = 0$; similarly, as $b_2^+(X) = b_2^-(X)$, we have $\sigma(M \times \mathbf{S}^1) = 0$; as such, from (4.5), it will mean that $\langle 1\rangle_{k'}$ is the partition function of the topologically trivial sector where $k' = 0$. Second, note that for $k' = 0$, the dimension of the moduli space of flat "ramified" $SU(2)$-connections on $X$ is given by $-3\chi(M \times \mathbf{S}^1) = 0$; in other words, there are a discrete number of flat solutions of $A'$; consequently, as $X = M \times \mathbf{S}^1$ is a trivial product of two spaces, each such flat solution of $A'$ on $X$ will correspond to a flat solution of $\mathscr{A}'$ on $M$; hence, the dimension $\dim(\mathcal{M}'_f)$ of the moduli space $\mathcal{M}'_f$ of flat "ramified" $SU(2)$-connections $\mathscr{A}'_f$ on $M$, is zero. Third, note that it is well established that the number of supersymmetric ground states of the sigma model is invariant under rescalings of the potential $h$; therefore, one can rescale $h \to \gamma h$, where $\gamma \gg 1$, and the Witten index $\mathrm{Tr}(-1)^F$ — which counts the difference in the number of bosonic and fermionic ground states — will not change; this means that one can compute $\chi(HF_*(M;\Sigma_M;\alpha))$ after such a rescaling of $h$, and still get the correct result. Last but not least, note that when $\gamma \gg 1$, the contributions to $\chi(HF_*(M;\Sigma_M;\alpha))$ will localize

[14]The relative grading, as defined mathematically for the ordinary instanton Floer complex, depends on the index which computes the dimension of the moduli space $\mathcal{M}$ of $SU(2)$-instantons; in particular, since $\dim(\mathcal{M}) = 8k - \frac{3}{2}(\chi + \sigma)$, where $k$ takes different integer values in different topological sectors, the relative grading is defined mod 8. In this sense, since $\dim(\mathcal{M}') = 4(2k + l) - \frac{3}{2}(\chi + \sigma) - 2(g - 1)$, where $k$ and $l$ take different integer values in different topological sectors, the relative grading of the "ramified" instanton Floer complex will be defined mod 4.

onto the critical point set of $h$, i.e., $\mathcal{M}'_f$. Thus, since $\dim(\mathcal{M}'_f) = 0$, i.e., $\mathcal{M}'_f$ consists of zero-dimensional points only, we have

$$\chi(HF_*(M; \Sigma_M; \alpha)) = \sum_x \text{sign}(h''(x)) = \sum_x \pm 1, \tag{5.12}$$

where $h''(x)$ is the Hessian of $h$ at the point $x \in \mathcal{M}'_f$.[15] In particular, the topological invariant $\chi(HF_*(M; \Sigma_M; \alpha))$ is — like $\langle 1 \rangle_{k'}$ computed using (4.15) — a sum of signed points; an *integer*. It is in this sense that (5.11) is deemed to be a consistent relation.

*Implications for a knot homology group from "ramified" instantons*

In fact, one can say more if $\Sigma_M$ is a knot $K \subset M$. In this case, $\chi(HF_*(M; K; \alpha))$ in (5.12) counts (with signs) the number of flat $SU(2)$-connections $\mathscr{A}'_f$ on $M \backslash K$ with holonomy $\exp(2\pi i \alpha)$ around a circle linking $K$ in $M$. Also, $\mathscr{A}'_f$ only picks up nontrivial contributions to the holonomy along a path that lies in the plane normal to (the singularity along) $K$, i.e., along the $\theta$-direction; therefore, if $K$ is a non-trivial knot — i.e., if there are crossings that cannot be undone by any orientation-preserving homeomorphism of $M$ to itself — the holonomy of $\mathscr{A}'_f$ along the longitude of $K$ will always be nontrivial (for some judicious choice of the $\alpha$-parameter of the surface operator). Therefore, one can also interpret $\chi(HF_*(M; K; \alpha))$ as an algebraic count of the number of conjugacy classes of homomorphisms

$$\rho : \pi_1(M \backslash K) \to SU(2) \tag{5.13}$$

which satisfy the constraint that $\rho$ maps — via the holonomy of $\mathscr{A}'_f$ — the longitude of $K$ to a *non-identity* element of $SU(2)$. In turn, this implies that the groups $HF_*(M; K; \alpha)$ are in one-to-one correspondence with the conjugacy classes $\rho$ with the stated constraint. Thus, we can identify $HF_*(M; K; \alpha)$ with the knot homology groups $LI_*(M, K)$ from "ramified" $SU(2)$-instantons defined by Kronheimer and Mrowka in Section 4.4 of [5]; in particular, we have $\chi(HF_*(M; K; \alpha)) = \chi(LI_*(M, K)) \neq 0$.

Note at this point that since the partition function $\langle 1 \rangle_{k'}$ is invariant under deformations of the metric on $X$, the relation (5.11) would imply that

[15] Since $b_1(M) \neq 0$, one might encounter a situation whereby some of the points in $\mathcal{M}'_f$ are degenerate. Nevertheless, for an appropriate nontrivial restriction of the $SU(2)$-bundle to $M$, one can — without altering the Witten index $\text{Tr}(-1)^F$ and therefore, $\chi(HF_*(M; \Sigma_M; \alpha))$ — perturb $h$ so that its critical point set will consist of a finite number of isolated, non-degenerate and irreducible points which we can then interpret as the $x$'s in (5.12) (cf. Prop. 3.12 of [5]).

$\chi(HF_*(M;K;\alpha))$ and hence $\chi(LI_*(M,K))$ are invariant under homeomorphisms of $M$ to itself. Consequently, if $K_0$ is an unknot which thus bounds a (twisted) disk in $M$, one can always — via a suitable orientation-preserving homeomorphism of $M$ to itself — deform $K_0$ to a trivial unknot $\tilde{K}_0$ (i.e., a geometrically round circle) such that $\chi(LI_*(M;K_0)) = \chi(LI_*(M;\tilde{K}_0))$. As the holonomy of $\mathscr{A}'_f$ along the longitude of the trivial unknot $\tilde{K}_0$ can only be the identity element of $SU(2)$ (according to our explanations in the last paragraph), the set of constrained maps $\rho$ in question will be empty for $\tilde{K}_0$; i.e., $LI_*(M,\tilde{K}_0)$ and therefore $\chi(LI_*(M;K_0))$ are zero.

In summary, we find that $\chi(LI_*(M,K))$ is zero only if $K$ is an unknot. Therefore, our above analysis physically affirms a mathematical conjecture proposed by Kronheimer and Mrowka in Section 4.4 of [5], which asserts that $\chi(LI_*(M,K))$ vanishes if the symmetrized Alexander polynomial of the knot $K$ is trivial, i.e., if $K$ is an unknot.

## 5.4 The Gromov–Taubes invariant, instanton floer homology, and the Casson–Walker–Lescop Invariant

*The Gromov–Taubes invariant of $M \times \mathbf{S}^1$ and the instanton floer homology of $M$*

Assume that *symplectic* $X = M \times \mathbf{S}^1$, where $M$ is a compact, oriented three-manifold, and $b_1(M) = b_2^+(X) > 1$. Now, let us consider the surface operator $\Sigma$ to be a pseudo-holomorphic curve in $X$ whose characteristics are as described at the beginning of Section 4. Let us also send the effective value of $\alpha$ to $+1$. Then, according to our analysis in Section 4, the LHS of (5.11) will be given by the Gromov–Taubes invariant $\mathrm{Gr}(c_1(\mathcal{E}))$, where $c_1(\mathcal{E})$ is Poincaré dual to $\Sigma$ with positive self-intersection, and $\mathcal{E}$ is a complex line bundle with self-dual connection $A_{\mathcal{E}}$.

On the other hand, when $\alpha = +1$, the holonomy $\exp(2\pi i\alpha)$ of the $SU(2)$ gauge connection $\mathscr{A}'$ around a small circle which links $\Sigma_M$ in $M$, is trivial; in other words, $\mathscr{A}'$ can, in this case, be regarded as an ordinary $SU(2)$ connection on $M$. In turn, this means that one can, in such a situation, replace $HF_*(M;\Sigma_M;\alpha)$ on the RHS of (5.11) with the *ordinary* instanton Floer homology groups $HF_*(M)$ of $M$.

From the preceding two points, one can therefore conclude that

$$\mathrm{Gr}(c_1(\mathcal{E})) = \chi(HF_*(M)). \tag{5.14}$$

In other words, the Gromov–Taubes invariant which algebraically counts *connected*, *non-multiply covered*, pseudo-holomorphic curves in $M \times \mathbf{S}^1$ with *positive* self-intersection, is equal to the Euler characteristic of the instanton Floer homology of $M$ with $b_1(M) > 1$!

One can immediately validate (5.14) for $M = \mathbf{T}^3$, since the relevant mathematical results exist. In this case, $X = \mathbf{T}^3 \times \mathbf{S}^1$ is symplectic Kähler with $b_2^+(X) = b_1(M) = 3$, and according to [30], $\chi(HF_*(\mathbf{T}^3)) = +1$.[16] What about $\mathrm{Gr}(c_1(\mathcal{E}))$? Well, although $b_2^+(X) > 1$, because $b_1(X) > 0$, one cannot read off from our result in (5.9) (which is defined for $b_1(X) = 0$ and $b_2^+(X) > 1$). However, from the relation $\mathrm{Gr}_0(A) = \pm 1$ proved in Proposition 3.18 of [21], and the fact that on any Kähler manifold, the almost complex structure $J$ is necessarily integrable and thus, all points in the space of pseudo-holomorphic curves have positive orientation, i.e., all points contribute as $+1$ in the computation of $\mathrm{Gr}(c_1(\mathcal{E}))$ [21], we can conclude that $\mathrm{Gr}(c_1(\mathcal{E})) = +1$ on $X$, too. Therefore, we have $\chi(HF_*(\mathbf{T}^3)) = \mathrm{Gr}(c_1(\mathcal{E})) = +1$, which certainly agrees with (5.14).

In fact, one can validate (5.14) for *any* $M = \Sigma_g \times \mathbf{S}^1$, where $\Sigma_g$ is a compact Riemann surface of genus $g > 1$. To this end, first note that $X$ is, in this case, a minimal symplectic manifold with $b_2^+(X) = b_1(M) = 1 + 2g$. Other than the $g = 1$ example above, its canonical bundle $K$ is nontrivial; however, since $g > 1$, $X$ is an elliptic surface of Kodaira dimension 1 — i.e., $K^2 = 0$ [31]. Consequently, because $c_1(\mathcal{E})^2 > 0$, Theorem 3.10 (*iv*) of [21] will imply that $\mathrm{Gr}(c_1(\mathcal{E})) = 0$. At the same time, we have $\chi(HF_*(M)) = 0$ for $g > 1$ [30]. In summary, for the elliptic surfaces $X = \Sigma_g \times \mathbf{T}^2$ where $g \geq 1$, (5.14) is found to be consistent with all known mathematical results.

Another nontrivial check on the validity of (5.14) is as follows. Let $\Sigma = K_0 \times \mathbf{S}^1$, where $K_0 \subset M$ is an unknot whence $\Sigma$ is homeomorphic to a genus one curve in $X$. Recall from our discussion at the beginning of Section 4 that in our case, the topological invariant $\mathrm{Gr}(c_1(\mathcal{E}))$ does *not* count curves of genus one in $X$; in other words, $\mathrm{Gr}(c_1(\mathcal{E})) = 0$ for such a $\Sigma$. At the same time, for such a $\Sigma$, we have $\chi(HF_*(M)) = \chi(HF_*(M; K_0; 1)) = 0$ from our discussion in the previous subsection. Again, this observation agrees with (5.14).

[16] Note that this mathematical result of [30] is actually valid for $G = SO(3)$. However, $\alpha$ continues to take values in $\mathfrak{u}(1)$ when $G = SO(3)$ instead of $SU(2)$; consequently, when $G = SO(3)$, our computations will still lead us to (5.14) — i.e., (5.14) also holds for $G = SO(3)$. Hence, we can still check against this mathematical result.

*The Gromov–Taubes and the Casson–Walker–Lescop invariants*

Let us also mention that it was argued in [32] that $\chi(HF_*(M)) = \lambda_{\rm CWL}(M)$, where $\lambda_{\rm CWL}(M)$ is the Casson–Walker–Lescop invariant of $M$ [33]; for example, one has $\chi(HF_*(\mathbf{T}^3)) = \lambda_{\rm CWL}(\mathbf{T}^3) = +1$. Hence, (5.14) will imply that for symplectic $X = M \times \mathbf{S}^1$,

$$\mathrm{Gr}(c_1(\mathcal{E})) = \lambda_{\rm CWL}(M). \tag{5.15}$$

Consequently, since $\lambda_{\rm CWL}(M) = 0$ if $b_1(M) > 3$, it will mean that

$$\mathrm{Gr}(c_1(\mathcal{E})) = 0, \quad \text{if} \quad b_2^+(X) > 3. \tag{5.16}$$

Notice that (5.15) and (5.16) indeed agree with our analysis of $X = \Sigma_g \times \mathbf{T}^2$ above.

## 5.5 The monopole Floer homology and SW invariants of three-manifolds

*A relation between the instanton and monopole Floer homologies of M*

Again, let us consider $X = M \times \mathbf{S}^1$ to be *symplectic*, where $M$ is a compact, oriented, three-manifold with $b_1(M) = b_2^+(X) > 1$. In this case, the relation (5.14) also leads to an important implication for a SW or monopole Floer homology group $HM_*(Y, \mathfrak{s}_Y)$ of a general three-manifold $Y$ with Spin$^c$-structure $\mathfrak{s}_Y$ described by Kronheimer in [34].[17] This can be understood as follows.

First, let us denote $\mathrm{SW}(X, \mathfrak{s})$ as the SW invariant of $X$ determined by a Spin$^c$-structure $\mathfrak{s}$ on $X$; let us also denote $\mathrm{SW}(M, \mathfrak{s}_M)$ as the SW invariant of $M$ (which "counts" the number of solutions of the three-dimensional SW equations on $M$ obtained by dimensional reduction along $\mathbf{S}^1$ of the original four-dimensional SW equations on $X$) determined by a Spin$^c$-structure $\mathfrak{s}_M$ on $M$; then, one can prove that $\mathrm{SW}(X, \pi^{-1}(\mathfrak{s}_M)) = \mathrm{SW}(M, \mathfrak{s}_M)$, where $\pi : M \times \mathbf{S}^1 \to M$ [34]. Second, note that $\chi(HM_*(M, \mathfrak{s}_M)) = \mathrm{SW}(M, \mathfrak{s}_M)$ [36]. Third, recall that as explained in footnote 16, (5.14) is also valid for when the gauge group underlying $HF_*(M)$ is $SO(3)$. Altogether, this means that

[17] Other variants of this monopole homology group were subsequently defined and constructed by Kronheimer–Mrowka in [35]; they were also studied in detail by Kutluhan–Taubes in [6] for when $Y = M$, $b_1(M) > 1$ and $X = M \times \mathbf{S}^1$ is symplectic — in other words, our case at hand.

(5.14), in light of (4.21), will imply that up to a sign, we have an equivalence

$$\chi(HM_*(M, \pi(\hat{\mathfrak{s}}))) = \chi(HF^w_*(M)) \tag{5.17}$$

between the Euler characteristics of the monopole and instanton Floer homologies of $M$! Here, $w = w_2(E_M)$ is the second SW class of the $SO(3)$ gauge bundle $E_M$ over $M$, and according to the Main Theorem of [6] and Section 40.1 of [35], the first Chern class of the projection $\pi(\hat{\mathfrak{s}})$ to $M$ of the Spin$^c$-structure $\hat{\mathfrak{s}}$ on $X$ is necessarily non-torsion. (Recall that $c_1(\hat{\mathfrak{s}}) = c_1(\mathcal{E}) - \frac{1}{2}c_1(K)$, where $K$ is the canonical line bundle of $X$, and $c_1(\mathcal{E})$ is the Poincaré-dual of the fundamental class of the *connected*, *non-multiply covered* pseudo-holomorphic curve $\Sigma$ with *positive* self-intersection.)

From (5.17), it is also clear that if the monopole Floer homology $HM_*(M, \hat{\mathfrak{s}}_M)$ is nontrivial for some Spin$^c$-structure $\hat{\mathfrak{s}}_M$ whose first Chern class is non-torsion, then the instanton Floer homology $HF^w_*(M)$ is nontrivial too. In the case that $M = \Sigma_g \times \mathbf{S}^1$ with $g \geq 1$, this result is just a generalization of Conjecture 6.3 in [34] proposed by Kronheimer for $g = 0$.

In fact, for $M = \Sigma_g \times \mathbf{S}^1$ with $g \geq 1$, it was proved in [37] that $HF^w_*(M)$ is isomorphic to the quantum cohomology $QH^*(\mathcal{M}_{\Sigma_g})$ of the moduli space $\mathcal{M}_{\Sigma_g}$ of flat $SO(3)$-connections on $\Sigma_g$ with nontrivial second SW class $w$; it was also proved in [38] that $HM_*(M, \pi(\hat{\mathfrak{s}}))$ is isomorphic to the quantum cohomology $QH^*(s^r(\Sigma_g))$ of the $r$-symmetric product $s^r(\Sigma_g)$ of the Riemann surface $\Sigma_g$, where the integer $r$ is related to the choice of the Spin$^c$-structure $\pi(\hat{\mathfrak{s}})$; since it is shown in [39] that the space $\mathcal{M}_{\Sigma_g}$ can be smoothly linked to the space $s^r(\Sigma_g)$, one ought to be able to identify $HF^w_*(M)$ with $HM_*(M, \pi(\hat{\mathfrak{s}}))$ such that (5.17) will hold. In short, for $M = \Sigma_g \times \mathbf{S}^1$ with $g \geq 1$, (5.17) is found to be consistent with expectations from existing mathematical results.

*Topology of the moduli space of flat $SU(2)$-connections on $M$*

Another implication of (5.17) can be understood as follows. First, note that $\mathrm{SW}(X_4, \hat{\mathfrak{s}})$ vanishes identically if the four-manifold $X_4$ has positive scalar curvature and $b^+_2(X_4) > 1$ [23]; in our case, this will mean that $\mathrm{SW}(X, \hat{\mathfrak{s}}) = \mathrm{SW}(M, \pi(\hat{\mathfrak{s}})) = \chi(HM_*(M, \pi(\hat{\mathfrak{s}}))) = 0$ identically if $M$ has positive scalar curvature. Second, note that $\chi(HF_*(M; \Sigma_M; \alpha)) = \chi(HF^{w=0}_*(M))$ if we send the effective value of $\alpha$ to $+1$; hence, from (5.12), we have $\chi(HF^{w=0}_*(M)) = \sum_x \pm 1$, where the $x$'s are the isolated points that span the zero-dimensional space $\mathcal{M}_f$ of flat (ordinary) $SU(2)$-connections on $M$; this just reflects the established fact that $\chi(HF^{w=0}_*(M))$ can also be interpreted as the Euler number $\chi(\mathcal{M}_f)$ [40]. By these two points, (5.17) will then mean that $\chi(\mathcal{M}_f)$ is zero if $M$ has positive scalar curvature — in other words, if $M$ has *positive* scalar curvature, $\mathcal{M}_f$ is either empty or spanned by an *equal* number of positively and negatively oriented points.

*Implications for the SW invariants of M*

Another useful thing to note is that it was pointed out in [32] that $\lambda_{\mathrm{CWL}}(M)$ in (5.15) can be expressed as a sum of all coefficients of the Reidemeister–Milnor torsion; in turn, by the work of Meng–Taubes in [41], this sum is given by a certain combination of SW invariants of $M$ called the SW series $\mathrm{SW}(t_i)$, where the $t_i$'s are variables whose details will not be important. Consequently, by (5.15) and (4.21), we have

$$\mathrm{SW}(M,\pi(\hat{\mathfrak{s}})) = \sum_{x\in H} \sum_{\mathfrak{s}_M|\bar{c}_1(\mathfrak{s}_M)=x} \mathrm{SW}(M,\mathfrak{s}_M) \tag{5.18}$$

up to a sign, where $H = H^2(M,\mathbb{Z})/\mathrm{Tor}(H^2(M,\mathbb{Z}))$ is the torsion-free part of the second integral cohomology of $M$, $\bar{c}_1(\mathfrak{s}_M)$ is the projection of $c_1(\mathfrak{s}_M)$ to $H$, and $c_1(\pi(\hat{\mathfrak{s}})) \in H$.

Once again, we can validate (5.18) for $M = \Sigma_g \times \mathbf{S}^1$ (or $X = \Sigma_g \times \mathbf{T}^2$) with $g \geq 1$. As we saw in Section 5.4, the magnitude of $\mathrm{Gr}(c_1(\mathcal{E}))$ and thus that of $\mathrm{SW}(X,\hat{\mathfrak{s}}) = SW(M,\pi(\hat{\mathfrak{s}}))$ on the LHS of (5.18) is equal to 1 and 0 for $g = 1$ and $g > 1$, respectively. At the same time, it is known that the SW series and hence the RHS of (5.18) is given by SW(1), where $\mathrm{SW}(t) = (t^{1/2} - t^{-1/2})^{2g-2}$; in other words, the magnitude of the RHS of (5.18) is 1 and 0 for $g = 1$ and $g > 1$, too. Therefore, for $M = \Sigma_g \times \mathbf{S}^1$ with $g \geq 1$, (5.18) is found to be consistent with all known mathematical results.

*A Non-vanishing theorem for the monopole Floer homology of M*

Now consider $X = M \times \mathbf{S}^1$ to be *general* with $b_2^+(X) = b_1(M) > 1$, where $M$ is a compact, oriented three-manifold. Let $\Sigma$ be an oriented two-surface of genus $g > 0$ that is smoothly embedded in $M$; then, the normal bundle of $\Sigma$ in $X$ is trivial,[18] i.e., $\Sigma \cap \Sigma = 0$. As such, by $\mathrm{SW}(X,\pi^{-1}(\mathfrak{s}_M)) = \mathrm{SW}(M,\mathfrak{s}_M)$, and by Theorem 1.3 of [3] — which generalizes (3.7) to $X$ with

[18]Consider the restriction $TM|_\Sigma$ to $\Sigma$ of the tangent bundle $TM$ of $M$; it splits as $TM|_\Sigma = T\Sigma \oplus TN$, where $T\Sigma$ and $TN$ are the tangent and normal bundles of $\Sigma$ in $M$, respectively. Note that $TM$ is oriented and so is $T\Sigma$; hence, $TN$ is also orientable. However, an orientable real line bundle such as $TN$ must be trivial; therefore, the normal bundle of $\Sigma$ in $X$ is also trivial, and it is given by $\Sigma \times \mathbf{R}^2$.

$b_1 \neq 0$ — we have

$$|\langle c_1(\mathfrak{s}_M), [\Sigma]\rangle| \leq (2g-2) \tag{5.19}$$

if $\mathrm{SW}(M, \mathfrak{s}_M) \neq 0$. Hence, since $\chi(HM_*(M, \mathfrak{s}_M)) = SW(M, \mathfrak{s}_M)$, it will mean that

$$HM_*(M, \mathfrak{s}_M) \neq 0 \tag{5.20}$$

as long as (5.19) holds. For $c_1(\mathfrak{s}_M)$ non-torsion, this claim is just Corollary 40.1.2 of [35].

### 5.6 "Ramified" generalizations of various relations between Donaldson and Floer theory

We shall now formulate, purely physically, "ramified" generalizations of various formulas presented by Donaldson and Atiyah in [10, 11] that relate ordinary Donaldson and Floer theory on four-manifolds with boundaries.

*"Ramified" Donaldson invariants with values in knot homology groups from "ramified" instantons*

To this end, let general $X = B \times \mathbf{R}_{\geq 0}$, where $B$ can be interpreted as the boundary of $X$, and the half real-line $\mathbf{R}_{\geq 0}$ can be interpreted as the "time" direction. Let the surface operator $\Sigma = K_B \times \mathbf{R}_{\geq 0}$, where $K_B$ is an arbitrary knot embedded in $B$. In such a case, there exists in the supersymmetry algebra a Hamiltonian $H$ which generates translations along $\mathbf{R}_{\geq 0}$, and by replacing $\mathbf{S}^1$ with $\mathbf{R}_{\geq 0}$ in our earlier explanation, we find that the "ramified" Donaldson–Witten theory can be interpreted as a supersymmetric quantum mechanical sigma-model with worldline $\mathbf{R}_{\geq 0}$, target manifold $\mathscr{A}'_B/\mathcal{G}'_B$ — the space of all gauge-inequivalent classes of "ramified" $SU(2)$-connections $\mathscr{A}'_B$ on $B$, and potential $h_B = \frac{1}{2}\int_B \mathrm{Tr}\,(\mathscr{A}'_B \wedge d\mathscr{A}'_B + \frac{2}{3}\mathscr{A}'_B \wedge \mathscr{A}'_B \wedge \mathscr{A}'_B)$ — the Chern–Simons functional of $\mathscr{A}'_B$. In particular, $\mathscr{A}'_B$ can be regarded a gauge connection of an $SU(2)$-bundle over $B\backslash K_B$ whose holonomy around the meridian of $K_B$ is given by $\exp(2\pi i\alpha)$, while the $\mathcal{Q}$-cohomology of the sigma-model — which is furnished by the supersymmetric ground states that correspond to the critical points of $h_B$ — can, in fact, be identified with the "ramified" instanton Floer homology $HF_*(B; K_B; \alpha)$.

According to the general ideas of quantum field theory, when the theory is formulated on such an $X$, one must specify the boundary values of the path-integral fields along $B$. Let us denote $\Phi_B$ to be the restriction of these fields to $B$; then, in the space $\mathscr{H}$ of functionals of the $\Phi_B$, specifying a set of boundary values for the fields on $B$ is tantamount to selecting a functional $\Psi(\Phi_B) \in \mathscr{H}$. Since the $\mathcal{Q}$-cohomology of the sigma-model is annihilated by $H$, i.e., it is time-invariant, one can take an arbitrary time-slice in $X$ and study the quantum theory formulated on $B$ instead; in this way, $\Psi(\Phi_B) \in \mathscr{H}$ can be interpreted as a state in the Hilbert space $\mathscr{H}$ of the quantum theory on $B$. As a result, via a state-operator mapping of the topological field theory, the correlation function "with boundary values of the fields determined by $\Psi$" will be given by

$$\langle \mathscr{O}_1 \dots \mathscr{O}_n \rangle_{\Psi(\Phi_B)} = \int \mathcal{D}\Phi \ e^{-S_E} \ \mathscr{O}_1 \dots \mathscr{O}_n \cdot \Psi(\Phi_B). \tag{5.21}$$

Since the theory ought to remain topological in the presence of a boundary $B$, according to our discussion in Section 2.3, it must be that $\{\mathcal{Q}, \mathscr{O}_i] = 0 = \{\mathcal{Q}, \Psi]$. Moreover, if $\Psi = \{\mathcal{Q}, \dots]$, the fact that $\{\mathcal{Q}, \mathscr{O}_i] = 0$ implies that (5.21) will also be zero. Thus, (5.21) depends on $\Psi$ via its interpretation as a $\mathcal{Q}$-cohomology class only, and since $\Psi$ is associated with the quantum theory on $B$, we can identify $\Psi$ as a class in the "ramified" instanton Floer homology $HF_*(B; K_B; \alpha)$. Altogether, since the $\mathscr{O}_i$'s represent either the operators $I'_0(p)$ or $I'_2(S)$ in (2.24)-(2.25), by (2.27), we find that (5.21) will represent a "ramified" Donaldson invariant with values in $HF_*(B; K_B; \alpha)$ — a knot homology group from "ramified" instantons. This is just a "ramified" generalization of the ordinary relation between Donaldson and Floer theory on $X$ described by Donaldson in [10].

*Interpretation as a scattering amplitude of "three-one branes"*

Now, let us assume that the total boundary $\partial X$ of $X$ consists not of a single boundary $B$, but a disjoint union of boundaries $B_j$, $j = 1, \dots, r$; i.e.,

$$\partial X = \bigsqcup_{j=1}^{r} B_j. \tag{5.22}$$

Let the surface operator $\Sigma = K_{\partial X} \times (X \backslash \partial X)$. If one is to choose the $\Psi(\Phi_{B_j})$'s appropriately such that one can replace all the $\mathscr{O}_i$'s with the identity operator 1 and yet have a non-vanishing path-integral, the resulting correlation function

$$\langle 1 \rangle_{\Psi(\Phi_{B_1}); \dots; \Psi(\Phi_{B_r})} = \int \mathcal{D}\Phi \ e^{-S_E} \ \Psi(\Phi_{B_1}) \cdots \Psi(\Phi_{B_r}) \tag{5.23}$$

can be interpreted as a scattering amplitude of incoming and outgoing "three-one branes" (the knot $K_{B_j}$ being the one-brane with "magnetic" charge $\alpha$ that is embedded in the three-brane $B_j$).

At any rate, according to the general ideas of quantum field theory, one can also write

$$\langle 1 \rangle_{\Psi(\Phi_{B_1});\ldots;\Psi(\Phi_{B_r})} = \int \mathcal{D}\Phi \, \mathrm{e}^{-S_E} \ \Psi(\Phi_{\partial X}). \tag{5.24}$$

As such, (5.23) can be expressed as

$$\int \mathcal{D}\Phi \, \mathrm{e}^{-S_E} \ \Psi(\Phi_{\partial X}) = \int \mathcal{D}\Phi \, \mathrm{e}^{-S_E} \ \Psi(\Phi_{B_1}) \cdots \Psi(\Phi_{B_r}). \tag{5.25}$$

In turn, this implies the relation

$$\begin{aligned} HF_*(\partial X; K_{\partial X}; \alpha) = {} & HF_*(B_1; K_{B_1}; \alpha) \otimes HF_*(B_2; K_{B_2}; \alpha) \\ & \otimes \cdots \otimes HF_*(B_r; K_{B_r}; \alpha) \end{aligned} \tag{5.26}$$

for knot homology groups from "ramified" instantons — which can be interpreted as a "ramified" generalization of equation (6.5) of [10] — that describes a scattering amplitude of "three-one branes".

*A Poincaré duality map of knot homology groups from "ramified" instantons*

Note that it is known [42] that one can always decompose a general $X$ along a homology three-sphere $Y$ into two parts $X^+$ and $X^-$, as shown in figure 1 below. Let $\Sigma^{\pm}$ be the parts of the surface operator $\Sigma$ which are embedded in $X^{\pm}$, and let $\Sigma_K^+ = K_Y$ and $\Sigma_K^- = K_{\bar{Y}}$ be their corresponding knot components embedded in $Y$ and $\bar{Y}$, respectively, where $\bar{Y}$ and $K_{\bar{Y}}$ are oppositely oriented copies of $Y$ and $K_Y$.

Now, consider the sector of the theory with "ramified" instanton number $k'$ given in (4.5), i.e., the $N_\psi = 0$ sector. The relevant non-vanishing observable is then the partition function $\langle 1 \rangle_{k'}$. Since $X = X^+ \cup_Y X^-$, according

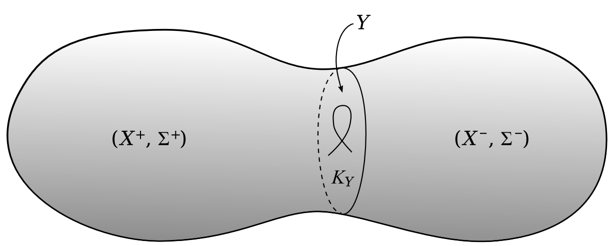


Figure 1: $X = X^+ \cup_Y X^-$.

to the general ideas of quantum field theory, one can evaluate $\langle 1 \rangle_{k'}$ as a path-integral over $X^+$ *towards* $Y$ followed by a second path-integral over $X^-$ *away* from $\bar{Y}$. Specifically, an independent path-integral over $X^+$ towards $Y$ will determine a state $\langle + |$ in the Hilbert space $\mathscr{H}_Y$ of the quantum theory on $Y$, while an independent path-integral over $X^-$ away from $\bar{Y}$ will determine a state $| - \rangle$ in the Hilbert space $\mathscr{H}_{\bar{Y}}$ of the quantum theory on $\bar{Y}$. As $\mathscr{H}_{\bar{Y}}$ is canonically the dual of $\mathscr{H}_Y$, we have

$$\langle + | - \rangle = \langle 1 \rangle_{k'}. \tag{5.27}$$

According to our above discussions, the independent path-integral over $X^+$ will be given by

$$\langle + | = \int \mathcal{D}\Phi \, \mathrm{e}^{-S_E} \; \Psi(\Phi_Y), \tag{5.28}$$

while the independent path-integral over $X^-$ will be given by

$$| - \rangle = \int \mathcal{D}\Phi \, \mathrm{e}^{-S_E} \; \Psi(\Phi_{\bar{Y}}), \tag{5.29}$$

where $\Psi(\Phi_Y)$ and $\Psi(\Phi_{\bar{Y}})$ can be interpreted as classes in $HF_*(Y; K_Y; \alpha)$ and $HF_*(\bar{Y}; K_{\bar{Y}}; \alpha)$, respectively. At the same time, as explained in Section 4, we have $\langle 1 \rangle_{k'} \in \mathbb{Z}$. Altogether, via (5.28) and (5.29), one can interpret (5.27) as a Poincaré duality map

$$HF_*(Y; K_Y; \alpha) \otimes HF_*(\bar{Y}; K_{\bar{Y}}; \alpha) \rightarrow \mathbb{Z} \tag{5.30}$$

of knot homology groups from "ramified" instantons. Note that (5.30) can be interpreted as a "ramified" generalization of a Poincare duality map of ordinary instanton Floer homology groups described by Atiyah in [11].

## 6 Generalization involving multiple surface operators

Notice that our physical derivation of Taubes' result in Section 4 involved only a single, connected surface operator $\Sigma$ that is nontrivially embedded in $X$ with $\Sigma \cap \Sigma > 0$. Let us now revisit that section and consider the case where one has multiple surface operators, which are nevertheless similar to $\Sigma$.

*A disconnected pseudo-holomorphic curve*

Let us start by considering the "total" surface operator

$$e = \sum_m \Sigma_m, \tag{6.1}$$

where $\Sigma_m$ — like $\Sigma$ characterized earlier by (4.1) and (4.2) — is selected from the $b_2^+$ number of homology cycles in the basis $\{U_i\}_{i=1,\ldots,b_2}$ which has a purely diagonal, unimodular intersection matrix. In particular, we have

$$\Sigma_i \cap \Sigma_j = 0, \quad \text{when} \quad i \neq j, \tag{6.2}$$

and therefore, the "total" surface operator $e$ consists of disjoint, non-multiply covered components furnished by the "member" surface operators $\Sigma_m$, which are themselves connected pseudo-holomorphic curves in $X$. Consequently, $e$ is also a pseudo-holomorphic curve, albeit a disconnected one, and since the Poincaré duals of the $\Sigma_m$'s are such that

$$\delta_{\Sigma_m} = \delta^+_{\Sigma_m} \tag{6.3}$$

for all $m$, it will mean that

$$e \cap e > 0. \tag{6.4}$$

*The corresponding moduli space of "ramified" $SU(2)$-instantons*

With the insertion of multiple disjoint surface operators as represented by the "total" surface operator $e$, the path-integral of the topological $SU(2)$ gauge theory localizes onto supersymmetric configurations which satisfy (cf. (6.3))

$$F_e^+ = 2\pi i \sum_m \alpha_m \delta_{\Sigma_m}, \tag{6.5}$$

where $\alpha_m$ is the "classical" parameter of the corresponding "member" surface operator $\Sigma_m$, and $F_e^+$ can be interpreted as an imaginary-valued curvature two-form of some complex line bundle with a self-dual $\mathfrak{u}(1)$-valued connection $A_e$.

Since the surface operators are disjoint, the holonomies of the gauge field around small circles linking the various $\Sigma_m$'s will not "mix" with one another. As such, one can rewrite (6.5) as a set of relations

$$F_m^+ = 2\pi \mathrm{i} \alpha_m \delta_{\Sigma_m}, \quad \text{for} \quad m = 1, 2, \ldots, \tag{6.6}$$

where $F_m^+$ can be interpreted as an imaginary-valued curvature two-form of some complex line bundle with a self-dual $\mathfrak{u}(1)$-valued connection $A_m$.

In other words, the path-integral localizes onto the moduli space $\mathcal{M}'$ of "ramified" $SU(2)$-instantons spanned by field configurations which satisfy the relation (6.5); the equivalent relations in (6.6) then imply that $\mathcal{M}'$ can actually be expressed as

$$\mathcal{M}' = \bigotimes_m \mathcal{M}'_m, \tag{6.7}$$

where $\mathcal{M}'_m$ is the moduli space spanned by field configurations which satisfy the $m$th relation in (6.6).

Similar to the case of a single surface operator, $N_\psi$ is given by the expression

$$N_\psi = 8k' - \tfrac{3}{2}(\chi + \sigma), \tag{6.8}$$

although now, the "ramified" instanton number $k'$ is given by

$$k' = k + 2\sum_m \alpha_m l_m - \sum_m \alpha_m^2 (\Sigma_m \cap \Sigma_m). \tag{6.9}$$

In particular, $k = \int_X c_2(E)$ and $l_m = -\int_{\Sigma_m} c_1(L)$, and according to our discussion surrounding (4.4), the value of $k$ — for any particular choice of $X$ and set of surface operators with parameters $\{\alpha_m\}$ and positive self-intersection numbers $\{\Sigma_m \cap \Sigma_m\}$ — determines the values of all the $l_m$'s, and vice versa. Thus, from (6.9), we find that the value of $k'$ is in one-to-one correspondence with the set $\{l_m\}$.

*The $N_\psi = 0$ sector*

Let us now consider the sector of the $SU(2)$ theory where $k'$ is as given in (4.5); i.e., the sector where $N_\psi = \dim(\mathcal{M}') = 0$. Since electric–magnetic duality in the low-energy $U(1)$ theory implies that there is a one-to-one correspondence between $c_1(L)$ and $\lambda$, there is, according to our preceding discussion, a one-to-one correspondence between $k'$ and $\mathfrak{s} = -\mathrm{i}\lambda$ for any particular choice of $X$ and set of surface operators. Then, by sending the effective value of $\alpha_m$ to $+1$ for every $m$ whence all the surface operators become "ordinary", and by repeating the arguments behind (4.6)–(4.17) whilst noting the fact that if $\mathcal{M}'$ of (6.7) is zero-dimensional, so are the spaces $\mathcal{M}'_m$, we get

$$SW(\mathfrak{s}) = \sum_x q(x), \tag{6.10}$$

where

$$q(x) = \prod_m \operatorname{sign}(\det \mathcal{D}_m). \tag{6.11}$$

Here, the $x$'s are the points which span the zero-dimensional space $H_e$ of solutions to the relation

$$c_1(\mathcal{E}_e) = \delta_e, \tag{6.12}$$

where $\mathcal{E}_e$ is a non-trivial complex line bundle with a self-dual $\mathfrak{u}(1)$-valued connection $A_e$ whence $c_1(\mathcal{E}_e) \cdot c_1(\mathcal{E}_e) > 0$; $\mathcal{D}_m$ for all $m$ is a first-order elliptic operator whose kernel is the tangent space to the space $H_m$ of solutions to the relation

$$c_1(\mathcal{E}_m) = \delta_{\Sigma_m}, \tag{6.13}$$

where $\mathcal{E}_m$ is a non-trivial complex line bundle with a self-dual $\mathfrak{u}(1)$-valued connection $A_m$ whence $c_1(\mathcal{E}_m) \cdot c_1(\mathcal{E}_m) > 0$; and

$$\mathcal{E}_e = \otimes_m \mathcal{E}_m. \tag{6.14}$$

Note that (6.11)–(6.14) mean that one can rewrite (6.10) as

$$\mathrm{SW}(\mathfrak{s}) = \mathrm{Gr}(c_1(\mathcal{E}_e)), \tag{6.15}$$

where $\mathrm{Gr}(c_1(\mathcal{E}_e))$ is the Gromov–Taubes invariant defined in [4] for a disconnected, non-multiply covered, pseudo-holomorphic curve $e$ whose fundamental class is Poincaré dual to $c_1(\mathcal{E}_e)$.

*Arriving at Taubes' result*

Before we proceed further, note that the analysis carried out in Section 3.2 can be generalized to the present case with multiple disjoint surface operators: one simply replaces "$\alpha\delta_\Sigma$" in the relevant analysis therein with "$\sum_m \alpha_m \delta_{\Sigma_m}$". With this in mind, note that since (6.15) is valid for $X$ of SW simple-type, i.e., $\lambda^2 - (2\chi + 3\sigma)/4 = d_e = -c_1(L_d^2)[e] + e \cap e = 0$, we find that the generalizations of (3.12), (3.10) and $\mathfrak{s}' = \mathfrak{s} - \delta_e$ will imply that the LHS of (6.15) is $\mathrm{SW}(\mathfrak{s} - \delta_e)$. In fact, since the non-vanishing contributions to the RHS of (6.15) localize around supersymmetric field configurations which obey (6.12), the LHS of (6.15) can actually be written as $\mathrm{SW}(\mathfrak{s} - c_1(\mathcal{E}_e))$; as a result, we have

$$SW(\mathfrak{s} - c_1(\mathcal{E}_e)) = \mathrm{Gr}(c_1(\mathcal{E}_e)). \tag{6.16}$$

As each $\Sigma_m$ is a connected pseudo-holomorphic curve with positive self-intersection, it must satisfy (4.20) and (4.1) simultaneously. This implies that we necessarily have $\mathfrak{s} = \frac{1}{2}c_1(L_d^2) = \frac{1}{2}c_1(K)$, as in the case of a single surface operator. By noting that as $b_2^+ > 1$, the ordinary SW invariants satisfy $\mathrm{SW}(\bar{\mathfrak{s}}) = \pm SW(-\bar{\mathfrak{s}})$ for any ordinary Spin$^c$-structure $\bar{\mathfrak{s}}$ [23], we can also write (6.16) as

$$SW(\hat{\mathfrak{s}}_e) = \pm \mathrm{Gr}(c_1(\mathcal{E}_e)), \tag{6.17}$$

where

$$\hat{\mathfrak{s}}_e = \tfrac{1}{2} c_1(\mathscr{L}_e), \tag{6.18}$$

and

$$\mathscr{L}_e = K^{-1} \otimes \mathcal{E}_e^2. \tag{6.19}$$

Moreover, since $c_1(L_d^2) = c_1(K)$, by (6.2) and (6.14), we find that the condition $d_e = 0$ can also be expressed as

$$d_e = \sum_m d_m = 0, \tag{6.20}$$

where

$$d_e = -c_1(K) \cdot c_1(\mathcal{E}_e) + c_1(\mathcal{E}_e) \cdot c_1(\mathcal{E}_e), \tag{6.21}$$

and

$$d_m = -c_1(K) \cdot c_1(\mathcal{E}_m) + c_1(\mathcal{E}_m) \cdot c_1(\mathcal{E}_m). \tag{6.22}$$

Because (6.21) and (6.22) are the *non-negative* dimensions of $H_e$ and $H_m$ as defined mathematically in [22], we see that (6.20) is indeed consistent with the fact that $H_e$ and therefore all the $H_m$'s are zero-dimensional as implied by $N_\psi = 0$. In turn, the fact that $d_m = 0$ implies, via (6.22) and (4.1), that the genus $g_m$ of the "member" pseudo-holomorphic curve represented by $c_1(\mathcal{E}_m)$ will be given by

$$g_m = 1 + c_1(\mathcal{E}_m) \cdot c_1(\mathcal{E}_m). \tag{6.23}$$

Finally, note that (6.17)–(6.23) are *precisely* Taubes' theorem [4] equating the ordinary SW invariants to the Gromov–Taubes invariants for disconnected curves in $X$! This implies that the novel mathematical identities obtained in the previous section can be generalized to hold for *disconnected*, non-multiply-covered, pseudo-holomorphic curves in $X$ with positive self-intersection, too. Nevertheless, in favor of brevity, we will not verify this explicitly.

## 7 Further application of our physical insights and results

Let us now, in this final section, apply some of our physical insights and results obtained hitherto to 1). elucidate certain key properties of the knot homology groups from "ramified" instantons discussed in Section 5.3 and

Section 5.6; 2). tell us more about the monopole Floer homology groups discussed in Section 5.5; 3). tell us more about the ordinary SW invariants of a compact, oriented, symplectic four-manifold with $b_1 = 0$ and $b_2^+ > 1$.

## 7.1 Properties of knot homology groups from "ramified" instantons

*Metric-independence*

Let us consider a decomposition of a general $X$ along two disjoint, compact, connected, oriented three-manifolds $Y_0$ and $Y_1$ into three parts $X^+$, $X^-$ and $X'$, as shown in figure 2 below. Let $\Sigma^{\pm}$ and $\Sigma'$ be the parts of the surface operator $\Sigma$ which are embedded in $X^{\pm}$ and $X'$, and let $\Sigma_K^+ = K_{Y_0}$, $\Sigma_K^- = K_{\bar{Y}_0} \cup K_{Y_1}$ and $\Sigma'_K = K_{Y_1}$ be their corresponding knot components embedded in $Y_0$, $\bar{Y}_0$ and $Y_1$, respectively, where $K_B$ indicates the oriented knot embedded in the oriented three-manifold $B$ such that the holonomy of the gauge field around its meridian is $\exp(2\pi i \alpha)$, and $K_{\bar{B}}$ and $\bar{B}$ are just oppositely oriented copies of $K_B$ and $B$.

Let us now, as was done for a similar case in Section 5.6, compute the path-integral over the middle segment labeled by $X^-$. In this case, there will be two sets of boundary values of the fields: one at $\bar{Y}_0$, and the other at $Y_1$. As such, according to the general ideas in quantum field theory, the path-integral will be given by

$$\Psi(\Phi') = \int_{\Phi_{Y_1} = \Phi'} \mathcal{D}\Phi \; e^{-S_E} \; \Psi(\Phi_{\bar{Y}_0}). \tag{7.1}$$

An explanation of the above formula is in order. First, $\Phi_B$ indicates the restriction of the path-integral fields to $B$; correspondingly, $\Psi(\Phi_B)$ is a functional of $\Phi_B$ which determines the boundary values of the fields on $B$. Second, the path-integral is computed over all fields $\Phi$ which when restricted

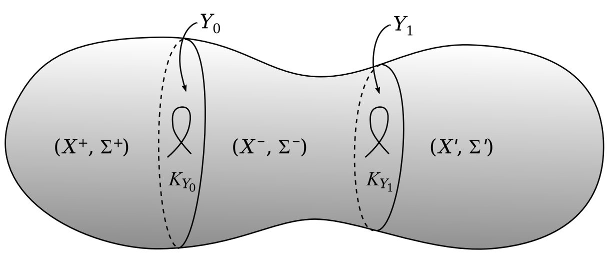


Figure 2: $X = X^+ \cup_{Y_0} X^- \cup_{Y_1} X'$.

to $Y_1$, have values $\Phi'$; this takes care of the boundary values on $Y_1$; according to our discussions in Section 5.6, the insertion of the operator $\Psi(\Phi_{\bar{Y}_0})$ will then take care of the boundary values on $\bar{Y}_0$. Third, we have assumed the boundary values of the fields on $\bar{Y}_0$ and therefore $\Psi(\Phi_{\bar{Y}_0})$, to be a priori determined, while on the other hand, we have assumed the boundary values of the fields on $Y_1$ and therefore $\Phi'$, to be a priori *undetermined*. As such, the path-integral will depend on $\Phi'$, and therefore, it can also be interpreted as a functional $\Psi(\Phi')$ of $\Phi'$, as written in (7.1).

Note at this point that the integration measure $\mathcal{D}\Phi$ is invariant under supersymmetry; in other words, it is $\mathcal{Q}$-closed. Recall also from (2.20) that $S_E = \{\mathcal{Q}, \dots\}$ (since, as explained in footnote 13, we are considering surface operators with $\eta = 0$, while the $\Theta$-angle can always be set to zero via an irrelevant chiral rotation of the massless fermions). Altogether, since $\mathcal{Q}^2 = 0$, it will mean that if $\{\mathcal{Q}, \Psi(\Phi_{\bar{Y}_0})] = 0$, then $\{\mathcal{Q}, \Psi(\Phi')] = 0$, and if $\Psi(\Phi_{\bar{Y}_0}) = \{\mathcal{Q}, \dots]$, then $\Psi(\Phi') = \{\mathcal{Q}, \dots]$, too. Therefore, (7.1) represents a map $\mathcal{H} : \Psi(\Phi_{\bar{Y}_0}) \to \Psi(\Phi')$ of $\mathcal{Q}$-cohomology classes. In addition, as explained in Section 2.3, due to the stress-tensor $T_{\mu\nu}$ of the underlying physical theory being $\mathcal{Q}$-exact, $\mathcal{H}$ is necessarily invariant under metric deformations of $X$.

Now, let us decompose $X$ along three disjoint, compact, connected, oriented three-manifolds $Y_0$, $Y_1$ and $Y_2$ into four parts $X^+$, $X^-$, $X'$ and $X''$, as shown in figure 3 below. Let $\Sigma^\pm$, $\Sigma'$ and $\Sigma''$ be the components of the surface operator $\Sigma$ which are embedded in $X^\pm$, $X'$ and $X''$, and let $\Sigma^+_K = K_{Y_0}$, $\Sigma^-_K = K_{\bar{Y}_0} \cup K_{Y_1}$, $\Sigma'_K = K_{\bar{Y}_1} \cup K_{Y_2}$ and $\Sigma''_K = K_{\bar{Y}_2}$ be their corresponding knot components embedded in $Y_0$, $Y_1$, $Y_2$ and their oppositely oriented copies, respectively. What we would like to do next is to compute the path-integral over the region $X^- \cup_{Y_1} X'$. If we assume the boundary values of the fields on $Y_2$ — like those on $\bar{Y}_0$ — to be a priori determined, the path-integral will be given by

$$Z(X^- \cup_{Y_1} X') = \int_{\Phi_{Y_1} = \Phi'} \mathcal{D}\Phi \, \mathrm{e}^{-S_E} \; \Psi(\Phi_{Y_2}) \cdot \Psi(\Phi_{\bar{Y}_0}). \tag{7.2}$$

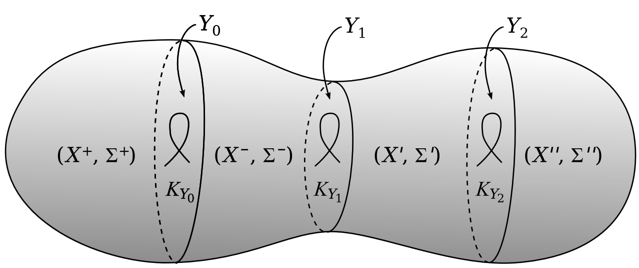


Figure 3: $X = X^+ \cup_{Y_0} X^- \cup_{Y_1} X' \cup_{Y_2} X''$.

That being said, according to the general ideas of quantum field theory, one can also compute $Z(X^- \cup_{Y_1} X')$ as a path-integral over $X^-$ — away from $\bar{Y}_0$ and towards $Y_1$ — followed by a second path-integral over $X'$ — away from $\bar{Y}_1$ and towards $Y_2$. Thus, from our above discussion leading to (7.1), and (7.2), we can write

$$\begin{aligned} &\int_{\Phi_{Y_1}=\Phi'} \mathcal{D}\Phi \, \mathrm{e}^{-S_E} \; \Psi(\Phi_{Y_2}) \cdot \int_{\Phi_{Y_1}=\Phi'} \mathcal{D}\Phi \, \mathrm{e}^{-S_E} \; \Psi(\Phi_{\bar{Y}_0}) \\ &\quad = \int_{\Phi_{Y_1}=\Phi'} \mathcal{D}\Phi \, \mathrm{e}^{-S_E} \; \Psi(\Phi_{Y_2}) \cdot \Psi(\Phi_{\bar{Y}_0}), \end{aligned} \tag{7.3}$$

where we have made use of the fact that specifying the a priori undetermined boundary values of the fields on $\bar{Y}_1$ is equivalent to specifying those on its mirror $Y_1$. Notice that (7.3) means that

$$\mathcal{H}(\Psi(\Phi_{Y_2})) \cdot \mathcal{H}(\Psi(\Phi_{\bar{Y}_0})) = \mathcal{H}(\Psi(\Phi_{Y_2}) \cdot \Psi(\Phi_{\bar{Y}_0})), \tag{7.4}$$

i.e., the map $\mathcal{H}$ is a *homomorphism.*

As per our discussions in Section 5.6, we find that $\Psi(\Phi_{\bar{Y}_0})$, $\Psi(\Phi_{Y_1})$ and $\Psi(\Phi_{Y_2})$ will correspond to classes in $HF_*(\bar{Y}_0; K_{\bar{Y}_0}; \alpha)$, $HF_*(Y_1; K_{Y_1}; \alpha)$ and $HF_*(Y_2; K_{Y_2}; \alpha)$, respectively. Also, according to our discussions in Section 5.6, the state $\Psi(\Phi_{\bar{Y}_0}) \in \mathscr{H}_{\bar{Y}_0}$ is in fact *dual* to the state $\Psi(\Phi_{Y_0}) \in \mathscr{H}_{Y_0}$ (where $\mathscr{H}_B$ refers to the Hilbert space of the quantum theory on $B$), i.e., we can identify $HF_*(\bar{Y}_0; K_{\bar{Y}_0}; \alpha)$ with $HF_*(Y_0; K_{Y_0}; \alpha)$. Hence, the map (7.1) can also be interpreted as the following homomorphism

$$\mathcal{H} : HF_*(Y_0; K_{Y_0}; \alpha) \to HF_*(Y_1; K_{Y_1}; \alpha) \tag{7.5}$$

on knot homology groups from "ramified" instantons. Moreover, $X^-$ is a connected, oriented manifold-with-boundary, and it contains a properly embedded oriented surface-with-boundary $\Sigma^-$, whence we have an orientation-preserving diffeomorphism of pairs

$$r : (\bar{Y}_0, K_{\bar{Y}_0}) \cup (Y_1, K_{Y_1}) \to (\partial X^-, \partial \Sigma^-). \tag{7.6}$$

In other words, we have a cobordism from $(Y_0, K_{Y_0})$ to $(Y_1, K_{Y_1})$ — that underlies the definition of the path-integral over $X^-$ — which gives rise to the homomorphism $\mathcal{H}$ of (7.5); since $\mathcal{H}$ is invariant under metric deformations of $X$, it will depend only on the diffeomorphism class of the cobordism, albeit up to a sign; this sign is determined by a choice of the zero-modes of $(A'_\mu, \psi_\mu, \chi^+_{\mu\nu})$ in the integration measure of (7.1), i.e., a choice of orientation for the line $\Lambda^{\max} H^1(X^-; \mathbb{R}) \otimes \Lambda^{\max} H^{2,+}(X^-; \mathbb{R}) \otimes \Lambda^{\max} H^1(Y_1; \mathbb{R})$.

This result has also been proved via a distinct mathematical approach by Kronheimer and Mrowka as Proposition 3.27 in [5]. In turn, this means that $HF_*(M;K;\alpha)$ for some compact, connected, oriented three-manifold $M$ with an oriented knot $K$ embedded in it, will be independent of the metric on $M$.

*The identity map*

Let us consider the setup in figure 2 again. If $Y_1 = Y_0$ and $K_{Y_1} = K_{Y_0}$, the path-integral over $X^-$, that is (7.1), will, in this case, be given by

$$\Psi(\Phi_{Y_0}) = \int_{\Phi_{Y_0}} \mathcal{D}\Phi \, \mathrm{e}^{-S_E} \; \Psi(\Phi_{\bar{Y}_0}) = \int_{\Phi_{Y_0}} \mathcal{D}\Phi \, \mathrm{e}^{-S_E} \tag{7.7}$$

(since as mentioned above, specifying the boundary values of the fields on $\bar{Y}_0$ is equivalent to specifying those on its mirror $Y_0$). Because (7.7) is a path-integral without operator insertions that, as explained earlier, is also invariant under metric deformations of $X$, we can compute it as $\mathrm{e}^{-Ht}$ [27] in the limit $t \to \infty$, where $t$ is the (stretched) interval of $X^-$. However, since $H$ acts as zero on the $\mathcal{Q}$-cohomology classes, (7.7) is always equal to 1 in our context; in other words, if we label $\mathcal{H}$ in (7.5) as $\mathcal{H}(Y_1; K_{Y_1}, Y_0; K_{Y_0})$, we have from (7.7)

$$\mathcal{H}(Y_0; K_{Y_0}, Y_0; K_{Y_0}) = 1; \tag{7.8}$$

a relation which can be thought to arise from a trivial cobordism from $(Y_0, K_{Y_0})$ to $(Y_0, K_{Y_0})$ furnished by $(X^-, \Sigma^-)$. This result is also part of Proposition 3.27 in [5].

*Composition of cobordisms and maps*

Let us consider the setup in figure 3 again, but now, with the boundary values of the fields on $Y_1$ *determined*. What we would like to do next is to compute the path integral over the segments spanned by $X^-$ and $X'$. Note that from the general ideas of quantum field theory, one can either compute this as a single path-integral starting from $\bar{Y}_0$ and ending at $Y_2$, or as a path-integral over $X^-$ — starting at $\bar{Y}_0$ and ending at $Y_1$ — followed by another path-integral over $X'$ — starting at $\bar{Y}_1$ and ending at $Y_2$. Consequently, we can write

$$\int_{\Phi_{Y_2}} \mathcal{D}\Phi \, \mathrm{e}^{-S_E} \; \Psi(\Phi_{\bar{Y}_1}) \cdot \int_{\Phi_{Y_1}} \mathcal{D}\Phi \, \mathrm{e}^{-S_E} \; \Psi(\Phi_{\bar{Y}_0}) = \int_{\Phi_{Y_2}} \mathcal{D}\Phi \, \mathrm{e}^{-S_E} \; \Psi(\Phi_{\bar{Y}_0}). \tag{7.9}$$

By the fact that the Hilbert spaces $\mathscr{H}_{\bar{B}}$ and $\mathscr{H}_B$ are canonically dual to each other whence one can identify $\Psi(\Phi_{\bar{B}})$ with $\Psi(\Phi_B)$, (7.9) will then mean that

$$\mathcal{H}(Y_2; K_{Y_2}, Y_1; K_{Y_1}) \cdot \mathcal{H}(Y_1; K_{Y_1}, Y_0; K_{Y_0}) = \mathcal{H}(Y_2; K_{Y_2}, Y_0; K_{Y_0}); \quad (7.10)$$

a relation that can be thought to arise from a composite cobordism from $(Y_0, K_{Y_0})$ to $(Y_1, K_{Y_1})$ to $(Y_2, K_{Y_2})$ furnished by $(X^-, \Sigma^-)$ and $(X', \Sigma')$, respectively. This result is also part of Proposition 3.27 in [5].

If we let $(Y_2, K_{Y_2}) = (Y_0, K_{Y_0})$, then (7.8) and (7.10) will imply that

$$\mathcal{H}(Y_1; K_{Y_1}, Y_0; K_{Y_0}) = \mathcal{H}^{-1}(Y_0; K_{Y_0}, Y_1; K_{Y_1}), \quad (7.11)$$

i.e., $\mathcal{H}$ is invertible and therefore, it is also an *isomorphism.*

## 7.2 A vanishing theorem for the monopole Floer homology of three-manifolds

Consider $X = M \times \mathbf{S}^1$ to be *symplectic* with $b_2^+(X) = b_1(M) > 1$, where $M$ is a compact, oriented three-manifold. Recall from our discussions in Section 4 that in our case, $\mathrm{Gr}(c_1(\mathcal{E}))$ only counts (with signs) pseudo-holomorphic curves $\Sigma$ — with Poincaré-dual $c_1(\mathcal{E})$ — which are nontrivially embedded in $X$ such that $c_1(\mathcal{E}) \cdot [\omega_{\mathrm{sp}}] > 0$, where $[\omega_{\mathrm{sp}}]$ is the Poincaré-dual of the symplectic two-form $\omega_{\mathrm{sp}}$ on $X$. Consequently, $\mathrm{Gr}(c_1(\mathcal{E})) = 0$ *identically* if $c_1(\mathcal{E}) \cdot [\omega_{\mathrm{sp}}] \leq 0$, and by (5.14) and (5.17), it will mean that the monopole Floer homology groups $HM_*(M, \pi(\hat{\mathfrak{s}}))$ ought to vanish if $c_1(\mathcal{E}) \cdot [\omega_{\mathrm{sp}}] \leq 0$; here, $\pi : M \times \mathbf{S}^1 \to M$, and the first Chern class of the Spin$^c$-structure $\hat{\mathfrak{s}}$ on $X$ is given by $2c_1(\hat{\mathfrak{s}}) = 2c_1(\mathcal{E}) - c_1(K)$, where $K$ is the canonical line bundle on $X$. Note that this easy-to-reach but nevertheless important conclusion about $HM_*(M, \pi(\hat{\mathfrak{s}}))$ has also been derived via a distinct and highly-involved mathematical approach in the Main Theorem of [6], where "$e$" and "$\mathfrak{s}_e$" therein correspond to $\pi(c_1(\mathcal{E}))$ and $\pi(\hat{\mathfrak{s}})$ herein.

*Mathematical versus physical computation*

In the mathematical proof of the Main Theorem in [6] by Kutluhan and Taubes, the above conclusion about the vanishing of $HM_*(M, \pi(\hat{\mathfrak{s}}))$ was obtained via a head-on analysis of the three-dimensional SW equations on $M$. In particular, the equations were checked for the presence or absence of sensible solutions (which directly generate $HM_*(M, \pi(\hat{\mathfrak{s}}))$) under various conditions; no reference to other related invariants of $M$ or $X$ were made at all.

On the other hand, in our computation leading to the above conclusion, we relied solely on the physically derived relations (5.14) and (5.17) — which connect the topological invariants in various dimensions to one another — without appealing to the three-dimensional SW equations on $M$. Thus, our physical computation provides, in this manner, a completely new way of deriving and understanding the vanishing of $HM_*(M, \pi(\hat{\mathfrak{s}}))$ when $c_1(\mathcal{E}) \cdot [\omega_{\rm sp}] \leq 0$.

### 7.3 SW invariants determined by the canonical basic class

Let $X$ be a compact, oriented, symplectic four-manifold with $b_1 = 0$ and $b_2^+ > 1$. Recall from (3.12) and (3.10) that for $\mathfrak{s} = \frac{1}{2}c_1(K)$, we have

$$\mathrm{SW}(\mathfrak{s} - \delta_{\mathcal{C}}) = \mathrm{SW}(\mathfrak{s}) \tag{7.12}$$

if and only if $d_K = -c_1(K)[\mathcal{C}] + \mathcal{C} \cap \mathcal{C} = 0$, and according to (4.24), $\mathcal{C}$ can be a pseudo-holomorphic curve in $X$.

As explained earlier, (5.8) implies that $c_1(K)$ is the Poincaré dual of some pseudo-holomorphic curve in $X$. Let $\mathcal{C}$ be such a curve, i.e., $\delta_{\mathcal{C}} = \frac{1}{2}c_1(K^2)$; as required, $d_K = 0$. Substituting this in (7.12), we find that

$$\mathrm{SW}(\mathfrak{s}_c) = \mathrm{SW}(\mathfrak{s}), \tag{7.13}$$

where $\mathfrak{s}_c = \frac{1}{2}c_1(K^{-1})$ is the canonical Spin$^c$-structure.

Hence, from (5.6), (5.9) and (7.13), we conclude that

$$\mathrm{SW}(\mathfrak{s}_c) = +1 \tag{7.14}$$

on $X$. This easy-to-reach but nevertheless important conclusion about $\mathrm{SW}(\mathfrak{s}_c)$ has also been proved via a highly involved and distinct mathematical approach in Proposition 2.1 of article 4 in [4].

*Mathematical versus physical computation*

In the mathematical computation of (7.14) in Proposition 2.1 of article 4 in [4], the magnitude of $\mathrm{SW}(\mathfrak{s}_c)$ is determined to be unity because the SW equations are shown to have a unique solution; the positive sign arises because the kernel of an elliptic operator associated with the linearization of the equations, is trivial.

On the other hand, our physical computation of (7.14) depends on the following: first, (7.13), which leads to (7.14), is a consequence of (5.8), which implies that there is a pseudo-holomorphic curve in $X$ that is Poincar'e dual to $c_1(K)$; second, the positive sign in (7.14) can be seen to originate from (5.9), i.e., the fact that the number of connected, non-multiply covered, pseudo-holomorphic curves $\Sigma \subset X$ of positive self-intersection which are "positively oriented" is greater than the number which are "negatively oriented" by one. Thus, our physical computation provides, in this manner, a completely new way of deriving and understanding (7.14).

### 7.4 About the SW invariants of Kähler manifolds

What if $X$ in Section 7.3 is Kähler? Then, one can say the following. First, since every compact, oriented, Kähler manifold is necessarily symplectic, (7.14) will apply to $X$ as well. This observation is just Theorem 3.3.2 of [43]. Second, on any Kähler manifold such as $X$ where the almost complex structure $J$ is necessarily integrable, all points in the space $H$ of pseudo-holomorphic curves have positive orientation; i.e., all points in $H$ contribute as $+1$ in the computation of $\mathrm{Gr}(c_1(\mathcal{E}))$ [21]. In light of (5.9) — i.e., $\mathrm{Gr}(c_1(\mathcal{E})) = +1$ — this means that $H$ consists of a *single* point only. Third, note that via (7.13) and (5.6), we have $\mathrm{SW}(\mathfrak{s}_c) = \mathrm{Gr}(c_1(\mathcal{E}))$. Therefore, this means that from each solution of the ordinary SW equations on $X$ determined by $\mathfrak{s}_c$, one can derive a pseudo-holomorphic curve in $X$ whose fundamental class is Poincaré dual to $c_1(\mathcal{E})$. In other words, the number of points in the moduli space $\mathcal{M}_{\mathfrak{s}_c}$ of solutions of the ordinary SW equations determined by $\mathfrak{s}_c$ equals the number of points in $H$. Thus, according to the second statement above, $\mathcal{M}_{\mathfrak{s}_c}$ consists of a *single* point only. This easy-to-reach but nevertheless important conclusion about $\mathcal{M}_{\mathfrak{s}_c}$ has also been proved via a highly-involved and distinct mathematical approach in Proposition 3.3.1 of [43].

*Mathematical versus physical computation*

In the mathematical computation of (7.14) for Kähler manifolds in Theorem 3.3.2 of [43], the magnitude of $\mathrm{SW}(\mathfrak{s}_c)$ is again determined to be unity because the SW equations are shown in Proposition 3.3.1 of [43] to have a unique solution; the positive sign arises because a relevant map between vector spaces defined by a certain "resonance operator" is orientation-preserving.

On the other hand, the basis of our physical computation of (7.14) for symplectic and thus Kähler manifolds, is as described at the end of the previous subsection. Moreover, in the case where $X$ is Kähler, one can, from our physical computation, understand the uniqueness of the solution of the SW equations to be a consequence of the fact that there is just *one*, connected, non-multiply covered pseudo-holomorphic curve of positive self-intersection in $X$ that is nontrivial in homology. Thus, our physical computation provides, in this manner, a completely new way of deriving and understanding the SW invariants of "admissible" Kähler four-manifolds.

## Acknowledgments

I would like to thank A.J. Berrick, O. Collin, F. Han and Y.L. Wong for mathematical consultations. This work is supported in part by a start-up grant from the National University of Singapore.

## References

[1] M. C. Tan, *Integration over the u-plane in Donaldson theory with surface operators*, J. High Energy Phys. **05** (2011), 007, arXiv:0912.4261.

[2] P. Ozsváth and Z. Szabó, *The symplectic Thom conjecture*, Ann. Math. (2) **151** (1) (2000), 93–124 [arXiv:math/9811087].

[3] P. Ozsváth and Z. Szabó, *Higher type adjunction inequalities in SW theory*, J. Differ. Geom. **55** (3) (2000), 385–440 [arXiv:math/0005268].

[4] C. H. Taubes, *SW and Gromov invariants for symplectic 4-manifolds*, Int. Press, Somerville, MA, 2000.

[5] P. B. Kronheimer and T. S. Mrowka, *Knot homology groups from instantons*, arXiv:0806.1053.

[6] C. Kutluhan and C. H. Taubes, *SW Floer homology and symplectic forms on* $S^1 \times M$, Geom. Topol. **13** (2009), 493–525 [arXiv:0804.1371].

[7] P. B. Kronheimer and T. S. Mrowka, *Embedded surfaces and the structure of Donaldson's polynomial invariants*, J. Differ. Geom. **41** (3) (1995), 573–734.

[8] P. B. Kronheimer and T. S. Mrowka, *Gauge theory for embedded surfaces: I*, Topology **32** (1993), 773–826.

[9] P. B. Kronheimer and T. S. Mrowka, *Gauge theory for embedded surfaces: II*, Topology **34** (1995), 37–97.

[10] S. K. Donaldson, *Floer homology groups in Yang–Mills theory*, Cambridge Tracts in Mathematics vol. 147, 2002.

[11] M. F. Atiyah, *New invariants of 3- and 4-dimensional manifolds.* The mathematical heritage of Hermann Weyl (Durham, NC, 1987), 285–299, Proc. Symp. Pure Math., vol. 48, Amer. Math. Soc., Providence, RI, 1988.

[12] S. Donaldson, *Polynomial invariants for smooth four-manifolds,* Topology **29** (1990), 257.

[13] J. Labastida and M. Marino, *Topological quantum field theory and four-manifolds*, Mathematical Physics Studies, vol. 25, Springer.

[14] P. Deligne, P. Etingof, D. S. Freed, L. Jeffrey, D. Kazhdan, J. Morgan, D. R. Morrison and E. Witten, eds., *Quantum Fields and Strings, A Course for Mathematicians*, vol. 2, AMS IAS, 1999.

[15] R. Bott and L. W. Tu, *Differential forms in algebraic topology*, Graduate Texts in Mathematics, vol. 82, Springer.

[16] M. C. Tan, *Surface operators in $N = 2$ abelian gauge theory,* J. high Energy Phys. 09, 047, 2009. [arXiv:0906.2413].

[17] S. Gukov and E. Witten, *Gauge theory, ramification, and the geometric langlands program*, Current developments in mathematics vol. **2006** (2008), 35–180 [arXiv:hep-th/0612073].

[18] M. C. Tan, *Notes on the ramified SW equations and invariants*, J. High Energy Phys. (in press), arXiv:hep-th/0912.1891.

[19] G. Moore and E. Witten, *Integration over the u-plane in Donaldson theory*, Adv. Theor. Math. Phys. **1** (1998), 298–387 [arXiv:hep-th/9709193].

[20] M. Gromov, *Pseudo-holomorphic curves in symplectic manifolds*, Invent. Math. **82** (1985), 307–347.

[21] D. McDuff, *Lectures on Gromov invariants for symplectic 4-manifolds*, (1996), arXiv:dg-ga/9606005.

[22] C. H. Taubes, *Counting pseudo-holomorphic submanifolds in dimension 4*, J. Differ. Geom. **44** (1996), 818–893.

[23] E. Witten, *Monopoles and four-manifolds*, Math. Res. Lett. **1** (1994), 769–796 [arXiv:hep-th/9411102].

[24] C. H. Taubes, *The seiberg-witten and gromov invariants*, Math. Res. Lett. **2** (1995), 221–238.

[25] A. Scorpan, *The wild world of 4-manifolds*, AMS.

[26] M. Blau and G. Thompson, *Topological Gauge theories from supersymmetric quantum mechanics on spaces of connections*, Int. J. Mod. Phys. A **8** (1993), 573–586 [arXiv:hep-th/9112064].

[27] K. Hori et al, *Mirror symmetry*, Clay mathematics monographs, vol. 1, 2003.

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

[29] S. Gukov, *Surface operators and knot homologies*, arXiv:0706.2369.

[30] V. Munõz, *Ring structure of the Floer cohomology of* $\Sigma \times \mathbf{S}^1$, Topology **38** (1999), 517–528 [arXiv:dg-ga/9710029].

[31] R. Friedman, *Algebraic surfaces and holomorphic vector bundles*, Springer-Verlag.

[32] M. Marino and G. Moore, *3-Manifold topology and the Donaldson–Witten partition function*, Nucl. Phys. B **547** (1999), 569–598 [arXiv:hep-th/9811214].

[33] C. Lescop, *Global surgery formula for the Casson–Walker invariant*, Annals of Mathematical Studies, Princeton University Press, Princeton, NJ, 1996.

[34] P. B. Kronheimer, *Embedded surfaces and Gauge theory in three and four dimensions*, available at http://www.math.harvard.edu/∼kronheim/jdg96.pdf.

[35] P. B. Kronheimer and T. S. Mrowka, *Monopoles and three-manifolds*, New mathematical monographs, vol. 10, Cambridge University Press.

[36] M. Marcolli, *SW Gauge theory*, Text and reading in mathematics vol. 17, Hindustan Book Agency.

[37] V. Muñoz, *Quantum cohomology of the moduli space of stable bundles over a Riemann surface*, arXiv:alg-geom/9711030.

[38] V. Muñoz and B.-L. Wang, *SW–Floer homology of a surface times a circle for non-torsion spin-c structures*, arXiv:math/9905050.

[39] M. Thaddeus, *Stable pairs, linear systems and the Verlinde formula*, arXiv:alg-geom/9210007.

[40] M. Blau and G. Thompson, *$N = 2$ topological gauge theory, the Euler characteristic of moduli spaces, the Casson invariant*, Comm. Math. Phys. **152** (1993), 41–72 [arXiv:hep-th/9112012].

[41] G. Meng and C. H. Taubes, *SW = Milnor torsion*, Math. Res. Lett. **3** (1996), 661–674.

[42] M. Freedman and L. Taylor, *Λ-splitting 4-manifolds*, Topology **16** (1977), 181–184.

[43] L. I. Nicolaescu, *Notes on SW theory.* Graduate Studies in Mathematics, vol. 28, AMS.