Topological Quantum Field Theory

Edward Witten[*]

School of Natural Sciences, Institute for Advanced Study, Olden Lane, Princeton, NJ 08540, USA

Abstract. A twisted version of four dimensional supersymmetric gauge theory is formulated. The model, which refines a nonrelativistic treatment by Atiyah, appears to underlie many recent developments in topology of low dimensional manifolds; the Donaldson polynomial invariants of four manifolds and the Floer groups of three manifolds appear naturally. The model may also be interesting from a physical viewpoint; it is in a sense a generally covariant quantum field theory, albeit one in which general covariance is unbroken, there are no gravitons, and the only excitations are topological.

1. Introduction

One of the dramatic developments in mathematics in recent years has been the program initiated by Donaldson of studying the topology of low dimensional manifolds via nonlinear classical field theory [1, 2]. Donaldson's work uses heavily the self-dual Yang-Mills equations, which were first introduced by physicists [3], and depends on some important results originally obtained by mathematical physicists, e.g. Taubes' theorem on existence of instantons on certain smooth four manifolds [4] (as well as hard analysis of instanton moduli spaces [5]). Thus there have been many conjectures that Donaldson's work may be related to physical ideas in an intimate way. However, such a relation has not been apparent in Donaldson's detailed constructions.

This picture has changed considerably because of the work of Floer on three manifolds [6]. Floer's work involves tunneling amplitudes in $3 + 1$ dimensions, and has been interpreted by Atiyah [7] in terms of a modified version of supersymmetric quantum gauge theory. (Floer theory has also been reviewed in [8].) In this viewpoint, Floer theory can be seen as a generalization to infinite dimensional function space of the supersymmetric approach to Morse theory [9].

The Floer homology groups are then the ground states of a certain Hamiltonian H which is closely related to physical quantum field theories. The Hamiltonian H, which will be described later, contains anticommuting fields of integer spin. Since H also acts in a Hilbert space of positive metric, the spin-statistics theorem implies that the theory must not be Lorentz invariant. It is easy to see that this is so (the anticommuting fields do not form Lorentz multiplets, and there are no anti-commuting gauge invariances).

   Purely from the point of view of Floer theory, which is a theory of three manifolds, a non-relativistic description is adequate. But one of the most beautiful features of Floer theory is its connection with Donaldson's theory of four manifolds. The Donaldson polynomial invariants of four manifolds were origin-ally defined for a four manifold M without boundary. It has turned out that to generalize Donaldson's original definitions to the case in which M has a non-empty boundary B, one must define relative Donaldson invariants of M that take values in the Floer groups of B. This connection between the Floer and Donaldson theories has led Atiyah to conjecture that the "Morse theory" interpretation of Floer homology must be an approximation to a relativistic quantum field theory. That conjecture was the motivation for the present work. We will find a relativistic formulation which turns out to require a not entirely trivial generalization of the nonrelativistic treatment in [7]. This generalization is described in Sect. 2. In Sects. 3 and 4, we describe from this point of view the origin of the Donaldson polynomials and their connection with Floer theory. In Sect. 5, some explicit formulas are worked out. Finally, in Sect. 6, we will discuss the possible physical interpretation of this work.

   There are many results on instantons in the physical and mathematical literature which are important background for the present work. Instantons were used to solve the "$U(1)$ problem" by 't Hooft [10] and were interpreted in terms of tunneling in [11, 12]. The formal theory of deformations of instantons, relevant in Sect. 3 and later, was developed in [13]. In addition, many remarkable properties of instantons in supersymmetric gauge theories have been uncovered in the physics literature. In particular, the ideas of [14, 15] may well be important for future developments in Floer and Donaldson theory, perhaps connected with the role of the reducible connections. Our treatment of Donaldson polynomials in Sect. 3 has a close formal similarity with the arguments given in [16] to determine certain correlation functions in strongly coupled supersymmetric gauge theories. Finally, it should be noted that many arguments in Sect. 3 and later will be quite recognizable to string theorists. This analogy is in fact tantalizing and is further pursued in Sect. 6. Introductions to the relevant string theory are [17–19].

2. Construction of the Lagrangian

Let us first recall the description of Floer theory in a non-relativistic quantum field theory [7]. One begins with gauge fields $A_i^a(x)$ on a three manifold Y. Here $i = 1 \ldots 3$ labels the components of a tangent vector to M, a runs over the generators of a gauge group G, and x labels a point in Y. Y is endowed with a metric tensor g_{ij}. We

wish to consider differential forms on the space $\mathscr{A}$ of all gauge connections on Y.[1] A basis for the one forms would be the $\delta A_i^a(x)$.[2] The $\delta A_i^a(x)$ can be regarded as operators on the differential forms on $\mathscr{A}$ [if ω is a differential form on $\mathscr{A}$, then $\delta A_i^a(x)$ acts on ω by $\omega \to \delta A_i^a(x) \wedge \omega$]. Regarded thus as operators on differential forms, the $\delta A_i^a(x)$ anticommute, $\{\delta A_i^a(x), \delta A_j^b(y)\} = 0$. They thus correspond to second quantized fermi fields. Following physical terminology, we will denote the $\delta A_i^a(x)$ as $\psi_i^a(x)$.

The exterior derivative on $\mathscr{A}$ is

$$d = \int d^3 x \psi_i^a(x) \frac{\delta}{\delta A_i^a(x)} \, . \tag{2.1}$$

Its adjoint is

$$d^* = - \int d^3 x \chi_i^a(x) \frac{\delta}{\delta A_i^a(x)} \, , \tag{2.2}$$

where the $\chi_i^a(x)$ [vector fields on $\mathscr{A}$ dual to the $\psi_i^a(x)$] obey $\{\chi_i^a(x), \chi_j^b(y)\} = 0$, $\{\chi_i^a(x), \psi_j^b(y)\} = g_{ij} \delta^{ab} \delta^3(x - y)$. One then considers the Chern-Simons functional $W = \frac{1}{2} \int_Y \mathrm{Tr}(A \wedge dA + \frac{2}{3} A \wedge A \wedge A)$ as a Morse function on Y. Thus, as in finite dimensions [9], one introduces a real number t and defines $d_t = e^{-tW} d e^{tW}$, $d_t^* = e^{tW} d^* e^{-tW}$ to be the "supersymmetry charges." They obey

$$d_t^2 = 0 \, , \quad d_t^{*2} = 0 \, , \quad d_t d_t^* + d_t^* d_t = 2H \, , \tag{2.3}$$

where H, defined by the last equation, is the Hamiltonian of the nonrelativistic theory. Explicitly,

$$H = \int d^3 x \left[\frac{1}{2} \sum_{i,a} \left(-i \frac{\delta}{\delta A_i^a(x)} \right)^2 + \frac{t^2}{2} \mathrm{Tr} B_i B^i + t \cdot \varepsilon_{ijk} \mathrm{Tr} \psi^i D^j \chi^k \right], \tag{2.4}$$

with $B_i = \frac{1}{2} \varepsilon_{ijk} F^{jk}$, $F_{ij} = \partial_i A_j - \partial_j A_i + [A_i, A_j]$.[3] In the first two terms of (2.4), we recognize the Hamiltonian of conventional (Lorentz-invariant) bosonic Yang-Mills theory. The last term is a Lorentz non-invariant coupling to fermions. (The fermions have spin one, impossible if the coupling were Lorentz invariant.) As described in [7], the ground states of (2.4) are the (rational) Floer groups of Y. These groups are graded by an additive quantum number which we will call U, with $U = 1$ for ψ and $U = -1$ for χ. (In finite dimensions, U would correspond to

---

[1] The relation of differential forms in function space to quantum fields was described for sigma models in [20], which may serve as useful background

[2] That is, for each x, i, and a, we view $A_i^a(x)$ as a function or zero form on $\mathscr{A}$, and $\delta A_i^a(x)$ is a one form which is the exterior derivative of the zero form $A_i^a(x)$. The symbol δ simply denotes the exterior derivative on the function space $\mathscr{A}$

[3] Our gauge theory conventions are that the covariant derivative of a charged field ϕ is $D_i \phi = \partial_i \phi + [A_i, \phi]$. Under an infinitesimal gauge transformation with $\delta \phi = [\varepsilon, \phi]$, the transformation of A_i is $\delta A_i = -D_i \varepsilon$. We regard the generator of the gauge group G as real, skew-symmetric matrices in the adjoint representation, and any field ϕ with values in the adjoint representation, when not described explicitly by components ϕ^a, is such a real skew-symmetric matrix. The symbol "Tr" denotes the positive definite Cartan-Killing form on the Lie algebra of G

the grading of the de Rham complex by dimension.) Because of instantons, U is only conserved modulo a constant; for $SU(2)$, the constant is 8.

2.1. Relativistic Generalization

In trying to find a relativistic version of this picture, our first problem is to decide what the supersymmetry algebra will be. In (2.3), there appears the generator H of time translations. There is a natural notion of time translations as long as we work on $Y \times R^1$ (Y is "space" and R^1 is "time"). However, the Donaldson theory applies to a general (compact, smooth) four manifold. On a general four manifold, there is no natural notion of "time translations," so one must work with a smaller supersymmetry algebra in which H does not appear. This means that we cannot keep both d_t and d_t^* in the intrinsic, four dimensional, theory. We must keep just one supersymmetry generator, say d_t, which we will call Q; it will obey simply $Q^2 = 0$. [We will see how to retrieve the algebra (2.3) if one specializes to a four manifold $Y \times R^1$.] Obeying $Q^2 = 0$, Q will be rather similar to a BRST charge, and we will suppose that it plays a BRST-like role of identifying physical states; physical states ψ will be states obeying $Q\psi = 0$, modulo those of the form $\psi = Q\lambda$.[4] It will turn out that the physical states in that sense are just the Floer groups. In this way, the negative norm states that one might have expected in a Lorentz invariant theory with anticommuting fields of integer spin will disappear.[5] Despite this BRST-like role of the supercharge Q, I have no idea how to obtain the "BRST" invariant Lagrangian considered below by gauge fixing of a gauge invariant Lagrangian. It will be argued in Sect. 6 that such a Lagrangian would have to be a generally covariant one of a new type.

Trying to extend (2.4) to a Lorentz invariant theory, the next problem is to put the fields $A_i^a, \psi_i^a, \chi_i^a$ into Lorentz multiplets. Clearly, the gauge field A_i^a is in a four dimensional picture part of a (Lie algebra valued) one form A_α^a, $\alpha = 1 \ldots 4$. (Tangent indices to a four manifold M will be denoted α, β, γ.) As for ψ and χ, we will take these to be a one form ψ_α^a and a self-dual two form $\chi_{\alpha\beta}^a$ (thus $\chi_{\alpha\beta}^a = -\chi_{\beta\alpha}^a = \frac{1}{2}\varepsilon_{\alpha\beta\gamma\delta}\chi^{\gamma\delta a}$). Also, we will supplement these with a zero form η^a of $U = -1$. The rationale behind these choices is that the $(\eta, \psi_\alpha, \chi_{\alpha\beta})$ multiplet is known to play a role in four dimensional instanton moduli problems analogous to the role of (ψ_i, χ_i) in Floer theory.[6]

Let us try to make a supersymmetric theory from these fields. The only reasonable (scale invariant and U conserving) Lagrangian that we can write is

$$\mathcal{L} = \int_M dx \, \mathrm{Tr} \left[\tfrac{1}{4} F_{\alpha\beta} F^{\alpha\beta} - i\eta D_\alpha \psi^\alpha + i(D_\alpha \psi_\beta)\chi^{\alpha\beta} \right]. \tag{2.5}$$

As for the supersymmetry transformation laws, the only reasonable try is

$$\delta A_\alpha = i\varepsilon\psi_\alpha, \quad \delta\eta = 0, \quad \delta\psi_\alpha = 0, \quad \delta\chi_{\alpha\beta} = \varepsilon(F_{\alpha\beta} + \tfrac{1}{2}\varepsilon_{\alpha\beta\gamma\delta}F^{\gamma\delta}) \tag{2.6}$$

with ε a constant anticommuting parameter.

---

[4] BRST quantization of gauge theories was originated in [21, 22]. The role of the BRST charge in identifying physical states emerged in [23]

[5] A similar phenomenon occurs in string theory in the no-ghost theorem. This aspect of the analogy between Donaldson theory and string theory was suggested by D. Friedan

[6] See [13, 2] for background; a brief sketch appears in Sect. 3 below

One quickly sees that (2.5) is invariant under (2.6) if the gauge group is abelian, but in the non-abelian case there is an uncancelled term of the form $\varepsilon \operatorname{Tr}\eta[\psi_\alpha, \psi^\alpha]$. There is no way to avoid this except by adding more fields. Heuristically, we should expect to have to add more bosons, because we have added new fermions (η and ψ_0) to the supersymmetric non-relativistic theory. (The addition of A_0 to the non-relativistic theory goes in the wrong direction, because it implies a constraint, rather than being a physical propagating field.) To guess what new fields are required, one may note that the propagating modes of A_α have helicities $(1, -1)$, while the propagating modes of (η, ψ, χ) have helicities $(1, -1, 0, 0)$. Since the supersymmetry parameter ε is to carry no spin, supersymmetry will require that the propagating modes of commuting and anticommuting fields should have the same helicities. Thus, we need two helicity zero commuting modes, and to accommodate them we introduce two new spinless fields ϕ and λ (in the adjoint representation of the gauge group).

A little experimentation leads to the Lagrangian

$$
\mathcal{L}_0 = \int_M d^4x \operatorname{Tr}\left[\frac{1}{4}F_{\alpha\beta}F^{\alpha\beta} + \frac{1}{2}\phi D_\alpha D^\alpha\lambda - i\eta D_\alpha\psi^\alpha + iD_\alpha\psi_\beta \cdot \chi^{\alpha\beta} \right.
$$
$$
\left. - \frac{i}{8}\phi[\chi_{\alpha\beta}, \chi^{\alpha\beta}] - \frac{i}{2}\lambda[\psi_\alpha, \psi^\alpha]\right].
\tag{2.7}
$$

This Lagrangian is invariant under the fermionic symmetry

$$
\delta A_\alpha = i\varepsilon\psi_\alpha, \quad \delta\phi = 0, \quad \delta\lambda = 2i\varepsilon\eta, \quad \delta\eta = \tfrac{1}{2}\varepsilon[\phi, \lambda],
$$
$$
\delta\psi_\alpha = -\varepsilon D_\alpha\phi, \quad \delta\chi_{\alpha\beta} = \varepsilon(F_{\alpha\beta} + \tfrac{1}{2}\varepsilon_{\alpha\beta\gamma\delta}F^{\gamma\delta}).
\tag{2.8}
$$

This action is also invariant under global scaling if the scaling dimensions of $(A, \phi, \lambda, \eta, \psi, \chi)$ are $(1, 0, 2, 2, 1, 2)$, and preserves the additive U symmetry if the U assignments are $(0, 2, -2, -2, -1, 1, -1)$.

Let us now work out the algebra obeyed by the fermionic symmetry. Let $\delta_\varepsilon(\Phi)$ denote the variation of any field Φ under (2.8). Let $T_\sigma(\Phi)$ denote the variation of Φ in a gauge transformation generated by an infinitesimal parameter σ (the gauge field A_α transforms as $T_\sigma(A_\alpha) = -D_\alpha\sigma$, and charged fields Φ transform as $T_\sigma(\Phi) = [\sigma, \Phi]$). Then one can verify that for all Φ,

$$
(\delta_\varepsilon\delta_{\varepsilon'} - \delta_{\varepsilon'}\delta_\varepsilon)(\Phi) = T_\varrho(\Phi),
\tag{2.9}
$$

where

$$
\varrho^a = -2i\varepsilon\varepsilon' \cdot \phi^a.
\tag{2.10}
$$

So the commutator of two supersymmetry transformations is a gauge transformation with infinitesimal parameter ϱ^a. In verifying (2.9) for $\Phi = A, \phi, \lambda, \eta, \psi$, one need not use the equations of motion, but for $\Phi = \chi$ the equations of motion must be used.

The Lagrangian (2.7) is not quite uniquely determined by its symmetries. The reason for this is as follows. Let us define a linear transformation $\{Q, \ \}$ of the space of all functionals of the field variables as follows. $\{Q, \ \}$ is defined by saying that for any functional $\mathcal{O}$, the variation $\delta\mathcal{O}$ of $\mathcal{O}$ under the fermionic symmetry (2.8) is

$$
\delta\mathcal{O} = -i\varepsilon \cdot \{Q, \mathcal{O}\}.
\tag{2.11}
$$

[Here $\{Q, V\}$ is simply a linear transformation on a suitable space of functionals of $A, \phi, \lambda, \eta, \psi, \chi$. The rationale for writing this transformation as $\{Q, \mathcal{O}\}$ and not merely as $Q(\mathcal{O})$ is simply that in the Hamiltonian framework, this transformation really corresponds to the graded commutator of $\mathcal{O}$ with the supercharge Q defined later.] Insofar as it is true that $Q^2 = 0$, $\{Q, \mathcal{O}\}$ is Q invariant and can be added to the Lagrangian without spoiling the fermionic symmetry. Actually, since the proof that $Q^2 = 0$ uses the χ equation of motion, we are only entitled to add $\{Q, \mathcal{O}\}$ to the Lagrangian if χ does not appear in $\mathcal{O}$. Also, since Q^2 is only zero up to a gauge transformation, we must pick $\mathcal{O}$ to be gauge invariant. In practice, there is one choice of $\mathcal{O}$ that respects all the symmetries, namely $\mathcal{O} = \frac{1}{4} \text{Tr}([\phi, \lambda]\eta)$. This gives us the possibility to add to the Lagrangian a new term

$$\mathcal{L}_1 = s \int d^4x \{Q, \mathcal{O}\} = s \int d^4x \, \text{Tr}\left[\frac{i}{2}\phi[\eta, \eta] + \frac{1}{8}[\phi, \lambda]^2\right] \tag{2.12}$$

with s an arbitrary parameter. Some of our later considerations will be simplified if we add (2.12) with $s = -1$, and thus the Lagrangian we actually use for many purposes will be

$$\mathcal{L} = \int_M d^4x \sqrt{g} \, \text{Tr}\left[\frac{1}{4}F_{\alpha\beta}F^{\alpha\beta} + \frac{1}{2}\phi D_\alpha D^\alpha \lambda - i\eta D_\alpha \psi^\alpha + iD_\alpha \psi_\beta \cdot \chi^{\alpha\beta}\right.$$
$$\left. - \frac{i}{8}\phi[\chi_{\alpha\beta}, \chi^{\alpha\beta}] - \frac{i}{2}\lambda[\psi_\alpha, \psi^\alpha] - \frac{i}{2}\phi[\eta, \eta] - \frac{1}{8}[\phi, \lambda]^2\right]. \tag{2.13}$$

The following remarks (though not strictly necessary for understanding this paper) may be helpful for readers acquainted with conventional supersymmetric gauge theories. The above construction of a four dimensional supersymmetric Lagrangian, starting with the non-relativistic version, adding fields, and adjusting couplings, undoubtedly seems rather ad hoc. There is, however, a simple way to relate the output Lagrangian (2.13) to standard physical constructions. [The argument will explain the form of the Lagrangian (2.13), but does not quite explain why it is supersymmetric on an arbitrary four manifold.] Consider the usual $N = 2$ supersymmetric gauge theory in flat Euclidean space R^4. The rotation group of R^4 is $SU(2)_L \times SU(2)_R$. The $N = 2$ Lagrangian has a global internal symmetry which we will denote as $SU(2)_I \times U(1)_U$. Under $SU(2)_L \times SU(2)_R \times SU(2)_I \times U(1)_U$, the fields of $N = 2$ supersymmetric Yang-Mills theory transform as follows. The gauge fields are

$$(1/2, 1/2, 0)^0, \tag{2.14}$$

the spinless bosons are

$$(0, 0, 0)^2 \oplus (0, 0, 0)^{-2}, \tag{2.15}$$

and the fermions are

$$(1/2, 0, 1/2)^1 \oplus (0, 1/2, 1/2)^{-1}. \tag{2.16}$$

[Here the three numbers in parenthesis denote $SU(2)_L \times SU(2)_R \times SU(2)_I$ representations, and the superscript is the $U(1)$ charge.] Suppose now that while

remaining in flat four dimensional Euclidean space, we consider an exotic action of the four dimensional rotation group, replacing $SU(2)_L \times SU(2)_R$ by $SU(2)_L \times SU(2)'_R$, where $SU(2)'_R$ is the diagonal sum of $SU(2)_R$ and $SU(2)_I$. It is easy to see that under $SU(2)_L \times SU(2)'_R \times U(1)$, the bosons transform as

$$(1/2, 1/2)^0 \oplus (0, 0)^2 \oplus (0, 0)^{-2} \tag{2.17}$$

and the fermions as

$$(1/2, 1/2)^1 \oplus (0, 1)^{-1} \oplus (0, 0)^{-1}. \tag{2.18}$$

These are precisely the fields appearing in (2.13). The fermions ψ_α, $\chi_{\alpha\beta}$, and η transform as $(\frac{1}{2}, \frac{1}{2})^1$, $(0, 1)^{-1}$, and $(0, 0)^{-1}$, respectively, while the bosons A_α, ϕ, and λ transform as $(\frac{1}{2}, \frac{1}{2})^0$, $(0, 0)^2$, and $(0, 0)^{-2}$. As long as the four manifold M is flat Euclidean space, the Lagrangian (2.13) is simply the standard $N = 2$ Lagrangian with an exotic action of the rotation group.

The global supersymmetries of the standard $N = 2$ model transform under $SU(2)_L \times SU(2)_R \times SU(2)_I \times U(1)_U$ as $(\frac{1}{2}, 0, \frac{1}{2})^{-1} \oplus (0, \frac{1}{2}, \frac{1}{2})^{+1}$. So they transform under $SU(2)_l \otimes SU(2)_R$ as $(\frac{1}{2}, \frac{1}{2})^{-1} \oplus (0, 1)^1 \oplus (0, 0)^1$. The Lorentz singlet supercharge that we have considered is simply the $(0, 0)^1$ component.

Thus, from this standpoint, it is obvious that (2.13) is supersymmetric if M is R^4 with flat metric. It is crucial and less obvious that (2.13) is supersymmetric for M an arbitrary orientable Riemannian four manifold. In verifying supersymmetry, one sometimes meets commutators of covariant derivatives, and the Riemann tensor might appear. However, in verifying supersymmetry of (2.13), there is only one point at which one meets the commutator of covariant derivatives. This is in computing $\delta(\mathrm{Tr}(D_\alpha \psi_\beta \cdot \chi^{\alpha\beta})) = \frac{1}{2}\varepsilon\,\mathrm{Tr}([D_\alpha, D_\beta]\phi \cdot \chi^{\alpha\beta})$. Since the commutator of covariant derivatives is here acting on the spin zero field ϕ, the Riemann tensor does not appear, and all is well. I do not know whether twisted versions of other $N \geq 2$ supersymmetric field theories will similarly be supersymmetric on a general four manifold.

2.2. Some Useful Formulas

We conclude this section by working out certain formulas that will be useful in later sections of this paper. First, we would like the formula for the supersymmetry current which, according to Noether's theorem, generates the fermionic symmetry (2.8). The recipe for finding the supersymmetry current is standard. One considers a transformation of the form (2.8) with ε an anticommuting parameter that is *not* necessarily constant. Since the variation of the Lagrangian would be zero if ε is constant, it must be proportional to the derivative of ε, and so has the general form

$$\delta\mathscr{L} = \int_M \partial_\alpha\varepsilon \cdot J^\alpha \tag{2.19}$$

for some J^α. In the case at hand one computes

$$J^\alpha = \mathrm{Tr}\left[(F^{\alpha\beta} + \tilde{F}^{\alpha\beta})\psi_\beta - \eta D^\alpha\phi - D_\beta\phi \cdot \chi^{\alpha\beta} - \frac{1}{2}\psi^\alpha[\lambda, \phi]\right] \tag{2.20}$$

with

$$\tilde{F}^{\alpha\beta} = \frac{1}{2}\varepsilon^{\alpha\beta\gamma\delta}F_{\gamma\delta}. \tag{2.21}$$

Now, the variation of the Lagrangian (2.13) under (2.8) has the form (2.19) regardless of the behavior of the fields $(A, \phi, \lambda, \eta, \psi, \chi)$. If, however, the Euler-Lagrange field equations are obeyed, then $\mathscr{L}$ is stationary under arbitrary (compactly supported) variations of the fields and in particular under (2.19). Thus, the Euler-Lagrange equations imply vanishing of (2.19) for arbitrary compactly supported ε; this must mean that the Euler-Lagrange equations imply that

$$D_\alpha J^\alpha = 0, \tag{2.22}$$

as one can verify directly. This enables us to construct a conserved charge. Given a homology three cycle Y in M, the integral

$$Q_Y = \int_Y d\Sigma_\mu J^\mu \tag{2.23}$$

$\left(\text{or equivalently } Q_Y = \int_Y * J, \text{ with } * J \text{ the closed three form dual to the current } J^\mu\right)$

depends only on the homology class of Y.

Now, we would like to compute the energy-momentum tensor of the theory. This is defined in terms of the variation of the Lagrangian under a change in the metric tensor $g_{\alpha\beta}$ of M. The definition of $T_{\alpha\beta}$ is that under an infinitesimal change of the metric $g^{\alpha\beta} \to g^{\alpha\beta} + \delta g^{\alpha\beta}$, the action changes by

$$\delta\mathscr{L} = \tfrac{1}{2} \int_M \sqrt{g}\,\delta g^{\alpha\beta} T_{\alpha\beta}. \tag{2.24}$$

There is one subtlety that must be noted here. The antisymmetric tensor $\chi_{\alpha\beta}$ is subject to a self-duality constraint

$$\chi_{\alpha\beta} = \tfrac{1}{2}\varepsilon_{\alpha\beta\gamma\delta}g^{\gamma\gamma'}g^{\delta\delta'}\chi_{\gamma'\delta'} \tag{2.25}$$

which must be preserved when computing the variation of the Lagrangian with respect to $g^{\alpha\beta}$. To preserve this condition, an arbitrary change $\delta g^{\alpha\beta}$ in the metric must be accompanied by

$$\delta\chi_{\alpha\beta} = \tfrac{1}{2}\varepsilon_{\alpha\beta\gamma\delta}\delta g^{\gamma\gamma'}g^{\delta\delta'}\chi_{\gamma'\delta'} - \tfrac{1}{8}(\delta g^{\sigma\tau}g_{\sigma\tau})\varepsilon_{\alpha\beta\gamma\delta}g^{\gamma\gamma'}g^{\delta\delta'}\chi_{\gamma'\delta'}. \tag{2.26}$$

It is then straightforward, although slightly tedious, to compute the energy-momentum tensor. One finds

$$T_{\alpha\beta} = \mathrm{Tr}\left[\left(F_{\alpha\sigma}F_\beta{}^\sigma - \frac{1}{4}g_{\alpha\beta}F_{\sigma\tau}F^{\sigma\tau}\right) + \frac{i}{2}\left[(D_\alpha\psi_\sigma - D_\sigma\psi_\alpha)\chi_\beta{}^\sigma + (D_\beta\psi_\sigma - D_\sigma\psi_\beta)\chi_\alpha{}^\sigma\right.\right.$$

$$\left.- \frac{1}{2}g_{\alpha\beta}(D_\lambda\psi_\sigma - D_\sigma\psi_\lambda)\chi^{\lambda\sigma}\right] - \frac{1}{2}(D_\alpha\phi D_\beta\lambda + D_\beta\phi D_\alpha\lambda - g_{\alpha\beta}D_\sigma\phi D^\sigma\lambda)$$

$$- i(D_\alpha\eta\psi_\beta + D_\beta\eta\psi_\alpha - g_{\alpha\beta}D_\sigma\eta\psi^\sigma) - 2i\left(\lambda\psi_\alpha\psi_\beta - \frac{1}{2}g_{\alpha\beta}\lambda\psi_\sigma\psi^\sigma\right)$$

$$\left.+ \frac{i}{2}g_{\alpha\beta}\phi[\eta,\eta] + \frac{1}{8}g_{\alpha\beta}[\phi,\lambda]^2\right]. \tag{2.27}$$

The single most important property of $T_{\alpha\beta}$ is, of course, that it is conserved (in the covariant sense) if the equations of motion are obeyed:

$$D_\alpha T^{\alpha\beta} = 0. \tag{2.28}$$

Equation (2.28) follows by a formal argument similar to that which gives (2.22). Under a coordinate transformation $\delta x^\alpha = u^\alpha$, with x^α coordinates on M and u^α an infinitesimal vector field, the metric changes by $\delta g^{\alpha\beta} = -(D^\alpha u^\beta + D^\beta u^\alpha)$. This coordinate transformation induces some change in $(A, \phi, \lambda, \eta, \psi, \chi)$. If the Euler-Lagrange equations are valid, the action will be invariant to lowest order. The change of the action is in fact $\delta \mathscr{L} = \frac{1}{2} \int \sqrt{g} (\delta g^{\alpha\beta}) T_{\alpha\beta} = -\frac{1}{2} \int \sqrt{g} (D^\alpha u^\beta + D^\beta u^\alpha) T_{\alpha\beta}$. This vanishes for arbitrary compactly supported u^α if and only if (2.28) is obeyed.

Now we wish to discuss scale and conformal invariances. For the trace of the energy-momentum tensor one finds

$$g^{\alpha\beta} T_{\alpha\beta} = \mathrm{Tr} \left[D_\sigma \phi \cdot D^\sigma \lambda - 2i D_\sigma \eta \cdot \psi^\sigma + 2i\lambda [\psi_\sigma, \psi^\sigma] + 2i\phi [\eta, \eta] + \frac{1}{2} [\phi, \lambda]^2 \right] .$$

$$(2.29)$$

This does not vanish, so the Lagrangian (2.13) is not conformally invariant – it is not invariant under $\delta g^{\alpha\beta} = w(x) g^{\alpha\beta}$ with $w(x)$ an arbitrary real function on M. However, the trace of the energy-momentum tensor can (using the equations of motion) be written as a total divergence,

$$g^{\alpha\beta} T_{\alpha\beta} = D_\alpha R^\alpha , \qquad (2.30)$$

with

$$R^\alpha = \mathrm{Tr} \left[\lambda D^\alpha \phi - 2i\eta \psi^\alpha \right] . \qquad (2.31)$$

The fact that the trace of the stress-tensor is a total divergence means that the Lagrangian is invariant under a global rescaling of the metric; that is, if $\delta g^{\alpha\beta} = w g^{\alpha\beta}$, with w constant, then $\delta \mathscr{L} = \frac{1}{2} \int_M \sqrt{g} \delta g^{\alpha\beta} T_{\alpha\beta} = \frac{w}{2} \int_M \sqrt{g} g^{\alpha\beta} T_{\alpha\beta} = \frac{w}{2} \int_M \sqrt{g} D_\alpha R^\alpha = 0$. A closely related statement is as follows. Take M to be R^4 with the usual flat metric, and with Euclidean coordinates x^α. Then the scaling transformation $\delta x^\alpha = w x^\alpha$ of Euclidean space (w an infinitesimal parameter) is generated by the conserved current

$$S^\alpha = T^{\alpha\beta} x_\beta - R^\alpha . \qquad (2.32)$$

Conservation of S^α follows immediately from (2.30). S^α generates the scale invariance of (2.13) that is visible to the naked eye, with scaling dimensions $(1, 0, 2, 2, 1, 2)$ for $(A, \phi, \lambda, \eta, \psi, \chi)$.[7]

Now we come to a point which – as we will see in the next section – is of utmost importance in understanding Donaldson theory. An operator which can be written as $\{Q, \mathcal{O}\}$ for some $\mathcal{O}$ is known in string theory as a "BRST commutator." [The operator $\{Q, \ \}$ was introduced in (2.11).] The correlation functions of BRST commutators are subject to powerful restrictions [17], which we will review in Sect. 3. In understanding Donaldson theory, it will turn out that one of the most important facts is that the energy-momentum tensor is a BRST commutator. One finds

$$T_{\alpha\beta} = \{Q, \lambda_{\alpha\beta}\} \qquad (2.33)$$

---

[7] Note, in particular, that (2.13) is a counterexample to some frequent mis-statements of the relation between scale and conformal invariance. For an elucidation of these issues see [24]

with

$$\lambda_{\alpha\beta} = \tfrac{1}{2} \mathrm{Tr}(F_{\alpha\sigma}\chi_\beta{}^\sigma + F_{\beta\sigma}\chi_\alpha{}^\sigma - \tfrac{1}{2}g_{\alpha\beta}F_{\sigma\tau}\chi^{\sigma\tau})$$
$$+ \tfrac{1}{2} \mathrm{Tr}(\psi_\alpha D_\beta\lambda + \psi_\beta D_\alpha\lambda - g_{\alpha\beta}\psi_\sigma D^\sigma\lambda) + \tfrac{1}{4}g_{\alpha\beta}\mathrm{Tr}(\eta[\phi,\lambda]). \qquad (2.34)$$

Now, one might guess from (2.28) and (2.33) that $D_\alpha\lambda^{\alpha\beta}$ would vanish. Rather one finds

$$D_\alpha(\lambda^{\alpha\beta} + U^{\alpha\beta}) = 0, \qquad (2.35)$$

with $U^{\alpha\beta} = -U^{\beta\alpha}$ an antisymmetric tensor defined by

$$U^{\alpha\beta} = -\tfrac{1}{2}\mathrm{Tr}[(F^{\alpha\beta} - \tilde{F}^{\alpha\beta})\eta] + \tfrac{1}{2}\varepsilon^{\alpha\beta\gamma\delta}\mathrm{Tr}\psi_\gamma D_\delta\lambda + \tfrac{1}{4}\mathrm{Tr}([\phi,\lambda]\chi^{\alpha\beta}). \qquad (2.36)$$

Equations (2.28), (2.35), and (2.36) together imply that

$$0 = D_\alpha T^{\alpha\beta} = D_\alpha(\{Q, \lambda^{\alpha\beta}\}) = \{Q, D_\alpha\lambda^{\alpha\beta}\}$$
$$= -\{Q, D_\alpha U^{\alpha\beta}\} = -D_\alpha(\{Q, U^{\alpha\beta}\}), \qquad (2.37)$$

so $\{Q, U^{\alpha\beta}\}$ must be conserved, even though $U^{\alpha\beta}$ is not. It is easy to check this; in fact

$$\{Q, U^{\alpha\beta}\} = \tfrac{1}{2}\varepsilon^{\alpha\beta\gamma\delta}D_\gamma R_\delta \qquad (2.38)$$

with R_δ defined in (2.31). [I do not know why precisely the same object R_δ appears in both (2.30) and (2.38).] From (2.38) it follows that $D_\alpha(\{Q, U^{\alpha\beta}\}) = 0$, as expected.

As preparation for the next section, we will require one more formula of a similar nature. Let

$$V = \tfrac{1}{4}\mathrm{Tr}F_{\alpha\beta}\chi^{\alpha\beta} + \tfrac{1}{2}\mathrm{Tr}\psi_\alpha D^\alpha\lambda - \tfrac{1}{4}\mathrm{Tr}(\eta[\phi,\lambda]). \qquad (2.39)$$

Then one computes (with the aid of the χ equation of motion) that

$$\{Q, V\} = \mathcal{L}', \qquad (2.40)$$

where

$$\mathcal{L}' = \mathcal{L} + \tfrac{1}{4}\int_M \sqrt{g}\,\mathrm{Tr}F_{\alpha\beta}\tilde{F}^{\alpha\beta}. \qquad (2.41)$$

Here $\mathcal{L}$ is the Lagrangian of Eq. (2.13), and the second term (which equals $\tfrac{1}{8}\int_M \mathrm{Tr}F \wedge F$) is a topological invariant, which measures the first Pontryagin class of the vector bundle on which A_α is a connection and $F_{\alpha\beta}$ a curvature tensor. Adding this topological invariant to the Lagrangian would not disturb any of our previous considerations, since the new term, being a topological invariant, is certainly invariant under all of the infinitesimal transformations that we considered earlier. Because of Eq. (2.40), $\mathcal{L}'$ is in many ways a more convenient choice of Lagrangian than $\mathcal{L}$, as we will see in the next section.

Before tackling the quantum theory (and the Donaldson polynomials), a few more points should be noted. The bosonic part of the Lagrangian contains the Yang-Mills action and certain scalar couplings. Although the Yang-Mills action is positive definite, the (ϕ, λ) kinetic energy is indefinite, so the $\mathrm{Tr}[\phi, \lambda]^2$ term in (2.13) has the wrong sign. These facts might appear to make problematical the Feynman

path integral formalism which we will use in the next section. There are several points of view one might adopt.

One point of view is to go back to the form (2.7) of the Lagrangian, perhaps with the addition of the topological charge term. [Recall that (2.7) and (2.13) are equivalent, differing only by a BRST commutator. We have introduced (2.13) only because of its more obvious relation to $N=2$ supersymmetry and the higher symmetry it has when $M = Y \times R^1$; see Sect. 2.4.] In this version, the scalars only appear quadratically; they can be integrated out by Gaussian integration, and the indefiniteness of the kinetic energy does no harm.

Another possible viewpoint is that instead of regarding ϕ and λ as independent real fields, one can view λ as a complex field and set $\phi = -\lambda^*$. Then the (ϕ, λ) terms in (2.13) become positive definite. This is what one would get if one takes twisted $N=2$ supersymmetry literally. All of the formulas given above still go through in this viewpoint.[8] The drawback of this approach is that if λ is complex and $\phi = -\lambda^*$, the Lagrangian is not real, and it is not obvious that the Donaldson invariants will come out to be real numbers.

In Floer and Donaldson theory, reducible connections (that is, gauge fields which are invariant under a non-trivial subgroup of the gauge group) are well known to cause many difficulties. In the present framework, these problems show up in zero modes of the (ϕ, λ) system for reducible connections. [In other words, for reducible connections, the Laplacian $-D_\alpha D^\alpha$, which is the kinetic operator for the (ϕ, λ) system in the linearized approximation, has a non-trivial kernel.] The proper treatment of the (ϕ, λ) system, which is not yet clear, is bound to interact in a non-trivial way with the proper treatment of reducible connections in Floer and Donaldson theory.

3. Path Integral Representation of Donaldson Polynomials

In the last section, we formulated a version of supersymmetric Yang-Mills theory that possesses some fermionic symmetry on an arbitrary smooth orientable four manifold M. This makes it possible to use techniques of quantum field theory to describe invariants of four manifolds. As we will see, a natural description of the Donaldson invariants will emerge.

In this section, we will see what can be obtained by formal manipulations of Feynman path integrals. Of course, a rigorous framework for four dimensional quantum gauge theory has not yet been developed to a sufficient extent to justify all of our considerations. Perhaps the connection we will uncover between quantum field theory and Donaldson theory may serve to broaden the interest in constructive field theory, or even stimulate the development of new approaches to that subject. Though our considerations in this section and the next one will be purely formal, we will see in Sect. 2.5 that even without having a rigorous construction of the quantum field theory, one can extract from it concrete and

---

[8] It is not necessary to worry about whether the variation (2.8) is compatible with $\phi = -\lambda^*$. The supercurrent (2.20) is still conserved by virtue of the equations of motion if λ is complex and $\phi = -\lambda^*$, and this is what counts

rigorous formulas that are relevant to Donaldson theory, together with a recipe for proving some of the main properties of those formulas.

Proceeding formally, we will find path integral representations for certain topological invariants.[8a] The integrals considered will be integrals over all the fields $(A, \phi, \lambda, \eta, \psi, \chi)$ considered in Sect. 2. The integration measure $\mathscr{D}A \cdot \mathscr{D}\phi \cdot \mathscr{D}\lambda \cdot \mathscr{D}\eta \cdot \mathscr{D}\psi \cdot \mathscr{D}\chi$ will be abbreviated as $(\mathscr{D}X)$.[9] The integrals we consider will be of the form

$$Z(W) = \int (\mathscr{D}X) \exp(-\mathscr{L}'/e^2) \cdot W, \tag{3.1}$$

where $\mathscr{L}'$ is the Lagrangian of Eq. (2.41) (with a topological term included), e is a real number usually known as the "gauge coupling constant," and W will usually be a polynomial in the integration variables. The integral in (3.1) is known as the (unnormalized) expectation value of W, and denoted $\langle W \rangle$.

Recall from Sect. 2 that the variation of any field $\mathcal{O}$ (by a "field" we simply mean a functional of the integration variables $A, \phi, \lambda, \eta, \psi, \chi$) under the fermionic transformation (2.8) is denoted as $\{Q, \mathcal{O}\}$. The most important property of Feynman path integration in the supersymmetric theory under discussion here is that $\langle \{Q, \mathcal{O}\} \rangle = 0$ for any $\mathcal{O}$. This holds for the following reason. Because the integration measure $(\mathscr{D}X)$ is invariant under supersymmetry, the integral

$$Z_\varepsilon(0) = \int (\mathscr{D}X) \exp(\varepsilon Q) \cdot \{\exp(-\mathscr{L}'/e^2) \cdot \mathcal{O}\} \tag{3.2}$$

is independent of the infinitesimal parameter ε. Expanding this out, and using the fact that the action is supersymmetric ($\{Q, \mathscr{L}'\} = 0$), we see that

$$Z_\varepsilon(\mathcal{O}) = \int (\mathscr{D}X) \exp(-\mathscr{L}'/e^2)(\mathcal{O} + \varepsilon\{Q, \mathcal{O}\}). \tag{3.3}$$

The assertion that $Z_\varepsilon(\mathcal{O})$ is independent of ε thus means that

$$0 = \langle \{Q, \mathcal{O}\} \rangle = \int (\mathscr{D}X) \exp(-\mathscr{L}'/e^2) \cdot \{Q, \mathcal{O}\}. \tag{3.4}$$

A corollary which perhaps should be stated explicitly is this. If $\{Q, A\} = 0$, then $\langle A\{Q, B\} \rangle = 0$ for any B, since $\{Q, A\} = 0$ implies $A\{Q, B\} = \{Q, AB\}$ so that

$$\langle A\{Q, B\} \rangle = \langle \{Q, AB\} \rangle = 0. \tag{3.5}$$

Now we are ready to define topological invariants. We pick a smooth four manifold M, a compact gauge group G, and a G bundle E on which the gauge field A_α^a is to be a connection. Our simplest topological invariant is simply the partition function:

$$Z = \int (\mathscr{D}X) \exp(-\mathscr{L}'/e^2). \tag{3.6}$$

Why is Z a topological invariant? To define the Lagrangian $\mathscr{L}'$, one needs to pick a Riemannian metric $g_{\alpha\beta}$ on M. To show that (3.6) is a topological invariant, it is necessary and sufficient to show that Z is invariant under an infinitesimal change in the metric. We recall from Sect. 2 that the change of $\mathscr{L}'$ under a change in g is by definition the energy-momentum tensor:

$$\delta\mathscr{L}' = \tfrac{1}{2} \int_M \sqrt{g}\, \delta g^{\alpha\beta} T_{\alpha\beta}. \tag{3.7}$$

[8a] What we will loosely call topological invariants are really "smooth invariants", that is, they depend on the smooth structure but not on the metric of M

[9] As is usual in gauge theories, we really intend here an integration over orbits of the gauge group – that is, over fields modulo gauge transformations

Also, recall that the energy-momentum tensor is a BRST commutator:

$$T_{\alpha\beta} = \{Q, \lambda_{\alpha\beta}\} \tag{3.8}$$

with $\lambda_{\alpha\beta}$ defined in Eq. (2.34). Therefore, the change in Z under a change in metric is

$$\delta Z = \int (\mathcal{D}X)[\exp(-\mathcal{L}'/e^2)] \cdot \left(-\frac{\delta\mathcal{L}'}{e^2} \right)$$

$$= -\frac{1}{e^2} \int (\mathcal{D}X) \exp(-\mathcal{L}'/e^2) \cdot \left\{ Q, \int_M \sqrt{g}\,\delta g^{\alpha\beta} \lambda_{\alpha\beta} \right\}$$

$$= -\frac{1}{e^2} \left\langle \left\{ Q, \int_M \sqrt{g}\,\delta g^{\alpha\beta} \lambda_{\alpha\beta} \right\} \right\rangle = 0. \tag{3.9}$$

This shows that Z is a topological invariant.[10]

Before trying to evaluate this invariant, let us observe that it is for similar reasons independent of the gauge coupling e. Indeed, the variation of Z with respect to e^2 is

$$\delta Z = \delta \int (\mathcal{D}X) \exp(-\mathcal{L}'/e^2) = \delta\left(-\frac{1}{e^2} \right) \int (\mathcal{D}X) \exp -(\mathcal{L}'/e^2) \cdot \mathcal{L}'$$

$$= \delta\left(-\frac{1}{e^2} \right) \int \mathcal{D}X \exp -(\mathcal{L}'/e^2) \cdot \{Q, V\} = \delta\left(-\frac{1}{e^2} \right) \cdot \langle \{Q, V\} \rangle = 0, \tag{3.10}$$

where we are borrowing from Eq. (2.40) the fact that $\mathcal{L}' = \{Q, V\}$. This shows that Z is independent of e^2 as long as $e^2 \neq 0$.[11]

Therefore, we can evaluate Z by going to the limit of very small e^2, whereupon the path integral is dominated by classical minima. To find these minima, note that the gauge field terms in $\mathcal{L}'$ are

$$\mathcal{L}'_A = \tfrac{1}{4} \int_M \sqrt{g}\, \mathrm{Tr}(F_{\alpha\beta}F^{\alpha\beta} + F_{\alpha\beta}\tilde{F}^{\alpha\beta})$$

$$= \tfrac{1}{8} \int_M \sqrt{g}\, \mathrm{Tr}(F_{\alpha\beta} + \tilde{F}_{\alpha\beta})(F^{\alpha\beta} + \tilde{F}^{\alpha\beta}). \tag{3.11}$$

This is positive semidefinite, and vanishes if and only if

$$F_{\alpha\beta} = -\tilde{F}_{\alpha\beta}, \tag{3.12}$$

that is if and only if the gauge field is anti-self-dual. Therefore, the evaluation of Z depends on expansion around solutions of (3.12), known as instantons. (It would

---

[10] We ignored in (45) a possible dependence of the measure $(\mathcal{D}X)$ on the metric g. Taking account of this dependence is really the problem of showing that the crucial equation $T_{\alpha\beta} = \{Q, \lambda_{\alpha\beta}\}$ is true quantum mechanically. Making this completely rigorous is one of the tasks of constructive quantum field theory. In this paper we will restrict ourselves to essentially classical considerations

[11] An attempt to go to $e^2 = 0$, by writing the whole factor $\exp(-\mathcal{L}'/e^2)$ as a BRST commutator, fails for two reasons. First, without the $\exp(-\mathcal{L}'/e^2)$ convergence factor, the integration by parts in function space used to prove that $\langle \{Q, \mathcal{O}\} \rangle = 0$ for any $\mathcal{O}$ would not be valid. Second, the verification that $\mathcal{L}' = \{Q, V\}$ used the χ equation of motion, and this can be justified in correlation functions only with use of the $\exp(-\mathcal{L}'/e^2)$ factor

be tiresome to call them anti-instantons; solutions of the opposite equation $F_{\alpha\beta} = + \tilde{F}_{\alpha\beta}$ will play no role in this section.)

Let us therefore make a few remarks about instantons. Depending on the choice of the manifold M and the bundle E, instantons may or may not exist. If they do exist, then (for a generic choice of metric g) the instantons have a moduli space $\mathcal{M}$ (smooth except for some relatively mild singularities) whose "formal" dimension $d(\mathcal{M})$ is given by a certain topological formula [13, 2].[12] If the gauge group G is $SU(2)$, this formula is

$$d(\mathcal{M}) = 8p_1(E) - \tfrac{3}{2}(\chi(M) + \sigma(M)), \qquad (3.13)$$

where $p_1(E)$ is the first Pontryagin number of the bundle E, and $\chi(M)$ and $\sigma(M)$ are the Euler characteristic and signature of M.

If we do manage to find an instanton gauge field A, we can look for a nearby instanton $A + \delta A$. The condition for $A + \delta A$ to obey (3.12) is

$$0 = D_\alpha \delta A_\beta - D_\beta \delta A_\alpha + \varepsilon_{\alpha\beta\gamma\delta} D^\gamma \delta A^\delta. \qquad (3.14)$$

In addition, we are interested in requiring δA to be orthogonal to the variations in A that can be obtained purely by a gauge transformation. This is conveniently achieved by imposing the gauge condition

$$0 = D_\alpha \delta A^\alpha. \qquad (3.15)$$

Let n be the number of solutions of (3.14), (3.15). These solutions describe infinitesimal instanton moduli, so at a generic point in moduli space, $n = d(\mathcal{M})$ (at least if conditions are such that the formal dimension equals the actual dimension).

Now let us look at fermion zero modes in the instanton field. The χ equation gives

$$D_\alpha \psi_\beta - D_\beta \psi_\alpha + \varepsilon_{\alpha\beta\gamma\delta} D^\gamma \psi^\delta = 0, \qquad (3.16)$$

while the η equation gives

$$D_\alpha \psi^\alpha = 0. \qquad (3.17)$$

These are the equations we have just seen, so the number of ψ zero modes is the number we have called n. (This relation between the fermion equations and the instanton moduli problem was the motivation for introducing precisely this collection of fermions in Sect. 2.)

For a generic $SU(2)$ instanton, there are no (η, χ) zero modes. This is so precisely when $n = d(\mathcal{M})$. The general statement, governed by an index theorem, is that the number of ψ zero modes minus the number of (η, χ) zero modes is equal to $d(\mathcal{M})$.

Recall that the Lagrangian (2.13) has a global symmetry U at the classical level. ψ has $U = +1$, and η, χ have $U = -1$. As in [10], the number of ψ zero modes minus the number of η, χ zero modes is the net violation of U by the instanton at the quantum level; we will call this number ΔU. Thus

$$\Delta U = d(\mathcal{M}). \qquad (3.18)$$

---

[12] The formal dimension equals the actual dimension under certain conditions noted later

(The meaning of this statement is that the integration measure $\mathscr{D}X$ is not invariant under U but transforms with a definite weight $-\Delta U$.) Equation (3.18) holds for any gauge group G; of course, $d(\mathscr{M})$ must be computed using the appropriate generalization of (3.13). It must be borne in mind, though, that the appropriate $d(\mathscr{M})$ is the "formal" dimension of the instanton moduli space, which equals the actual dimension only if G and E are such that the generic instanton is not invariant under any subgroup of G; if G is larger than $SU(2)$, this is only so under certain restrictions on E.

Let us now go back to our problem of computing Z. Z vanishes unless M, G, and E are such that $d(\mathscr{M})=0$. Otherwise, $\Delta U \neq 0$, and the partition function vanishes because of the fermion zero modes.

To further simplify the discussion, we will suppose that in addition to the formal dimension $d(\mathscr{M})$ of the instanton moduli space vanishing, the actual dimension also vanishes. In fact, we will assume that the moduli space consists of discrete, isolated instantons. (Standard quantum field theory methods could, however, be used to deal with a more general situation.) In expanding around an isolated instanton, it is enough in the weak coupling limit to keep only quadratic terms in the bose fields $\Phi=(A, \phi, \lambda)$ and fermi fields $\Psi=(\eta, \psi, \chi)$. (The weak coupling limit is adequate because we have seen that Z is independent of coupling.) The quadratic terms are of the general form,

$$\mathscr{L}_{(2)} = \int_M \sqrt{g} (\Phi \Delta_B \Phi + i\Psi D_F \Psi), \tag{3.19}$$

where Δ_B and D_F are certain second and first order operators, respectively.[13] The operator D_F is a real, skew-symmetric operator. The Gaussian integral over Δ_B and D_F gives

$$\frac{\text{Pfaff}(D_F)}{\sqrt{\det(\Delta_B)}}. \tag{3.20}$$

Here Pfaff denotes the Pfaffian of the real, skew-symmetric operator D_F. (Recall that, up to sign, the Pfaffian is the same as the square root of the determinant.)

The important point now is that D_F and Δ_B are related by supersymmetry. A look back to Eq. (2.8) shows that a classical field configuration in which $F_{\alpha\beta} + \tilde{F}_{\alpha\beta} = 0$ and $\phi, \lambda, \eta, \psi, \chi$ vanish is invariant under supersymmetry (the requirement $F_{\alpha\beta} + \tilde{F}_{\alpha\beta} = 0$ is needed to ensure $\delta\chi_{\alpha\beta} = 0$). Therefore, supersymmetry relates the bosonic and fermionic excitations about such a field configuration. To be precise, for every eigenvalue of D_F

$$iD_F \Psi = \lambda \Psi \tag{3.21}$$

with $\lambda \neq 0$, there is a corresponding eigenvalue of Δ_B,

$$\Delta_B \Phi = \lambda^2 \Phi. \tag{3.22}$$

(More exactly, as D_F is a skew symmetric operator, its eigenvalues occur in complex conjugate pairs. Each such pair corresponds to an eigenvalue of Δ_B.) At least for $M = S^4$, this relation between bosonic and fermionic eigenvalues is a

---

[13] Δ_B is an elliptic operator acting in the directions in field space transverse to the gauge orbits

standard result in the theory of instantons [25]. For the particular supersymmetric theory we are considering, the argument goes through for general M in the same way (since we have arranged so that supersymmetry holds for general M).

The ratio of determinants in (3.20) is thus formally

$$\frac{\text{Pfaff } D_F}{\sqrt{\det \Delta_B}} = \pm \prod_i \frac{\lambda_i}{\sqrt{|\lambda_i|^2}} \tag{3.23}$$

with the product running over all non-zero eigenvalues of Δ_B (or equivalently, over all non-zero eigenvalue pairs of D_F). The reason for the uncertain sign on the right-hand side of (3.23) is that although up to sign the Pfaffian is the square root of the determinant, the sign depends on a choice of orientation, which we must now discuss.

In fact, for any given gauge field A, there is no natural way to determine the sign of Pfaff D_F, since there is no natural way to pick an orientation (or equivalently, to fix the sign of the fermion measure). One may as well pick a particular gauge field $A = A_0$, and declare Pfaff $D_F(A = A_0) > 0$. Once this is done, there is a natural way to determine the sign of Pfaff $D_F(A = A')$ for any other gauge field A' (with A_0 and A' being connections on the same bundle). One simply interpolates from A_0 to A', via (say) the one parameter family of gauge fields $A_t = tA_0 + (1-t)A'$, $0 \leq t \leq 1$, and requires that the sign of Pfaff $D_F(A_t)$ changes sign whenever D_F has a zero eigenvalue for $A = A_t$. This uniquely determines the sign of Pfaff $D_F(A)$, but one must still ask whether the assignment is consistent – whether the sign that one will obtain this way depends on the choice of an interpolation from A_0 to A'. It is equivalent to ask whether Pfaff D_F will change sign when followed continuously around a non-contractible loop in the space $\mathscr{A}/\mathscr{G}$ of gauge fields modulo gauge transformations. Physically, this is the question of whether the theory we are trying to discuss has a global anomaly, in the sense of [26]. If so, the theory under investigation is inconsistent, in the usual sense of quantum field theory, and one cannot expect to learn anything of interest by studying it.

A priori, the Pfaffian of D_F must be regarded not as a function on $\mathscr{A}/\mathscr{G}$ but as a section of a certain real line bundle λ, which we may call the Pfaffian line bundle. The issue is whether the Pfaffian line bundle is trivial (orientable). Precisely this question has arisen in Donaldson's work. For Donaldson, it was important to know whether instanton moduli space $\mathscr{M}$ is orientable. If we denote the highest exterior power of the tangent bundle of $\mathscr{M}$ as ε, then for Donaldson the issue was whether the real line bundle ε was orientable. The two questions are related because, thinking of $\mathscr{M}$ as a subspace of $\mathscr{A}/\mathscr{G}$, the restriction of λ to $\mathscr{M}$ is canonically isomorphic to ε, at least if conditions are such that the formal and actual dimensions of $\mathscr{M}$ are equal. (This is so because the kernel of D_F corresponds, under such conditions, to the tangent space of moduli space.) Donaldson actually proved orientability of $\mathscr{M}$ by using index theory and certain topological arguments to prove that λ is always orientable, and thus his results show that there is never a global anomaly that would prevent a consistent determination of the sign of Pfaff D_F.

In our problem, we simply pick one instanton and declare that (3.23) is $+1$ for this instanton. Once this is done, there is a well defined way to evaluate (3.23) for

any other instanton; for the ith instanton it equals $(-1)^{n_i}$, where $n_i = 0$ or 1 according to the outcome of the process sketched above. The contribution of the ith instanton to Z being $(-1)^{n_i}$, we have finally

$$Z = \sum_i (-1)^{n_i}. \tag{3.24}$$

This is a familiar formula, originally introduced by Donaldson (who motivated the definition of the n_i in a slightly different but equivalent way). Donaldson showed on topological grounds that if M, G, and E are such that $d(\mathcal{M}) = 0$, then the right-hand side of (3.24) is a topological invariant. We have argued for the same conclusion by using the equation $T_{\alpha\beta} = \{Q, \lambda_{\alpha\beta}\}$ to prove that Z is a topological invariant, and then evaluating Z to arrive at (3.24).

Equation (3.24) is only the first of Donaldson's invariants. More generally, when $d(\mathcal{M}) > 0$, Donaldson defines certain more subtle analogues of (3.24), which have had rather dramatic implications for the study of smooth four manifolds. We would like to bring these within the framework of quantum field theory.

When $d(\mathcal{M}) > 0$, the non-vanishing path integrals will be of the form

$$Z(\mathcal{O}) = \int (\mathcal{D}X) \exp(-\mathscr{L}'/e^2) \cdot \mathcal{O}, \tag{3.25}$$

where $\mathcal{O}$ must carry a U quantum number equal to $d(\mathcal{M})$, so as to absorb the fermion zero modes. Let us determine the conditions on $\mathcal{O}$ for (3.25) to be a topological invariant.

The variation of (3.25) under a change in the metric is

$$\delta Z(\mathcal{O}) = \int (\mathcal{D}X) \exp(-\mathscr{L}'/e^2) \cdot \left(-\frac{\delta\mathscr{L}'}{e^2} \cdot \mathcal{O} + \delta_g \mathcal{O} \right)$$

$$= \int (\mathcal{D}X) e^{-I} \cdot \left(-\frac{1}{2e^2} \{Q, \int \sqrt{g} \delta g^{\alpha\beta} \lambda_{\alpha\beta}\} \cdot \mathcal{O} + \delta_g \mathcal{O} \right), \tag{3.26}$$

where $\delta_g \mathcal{O}$ is the variation of $\mathcal{O}$ with respect to $g_{\alpha\beta}$ (if $g_{\alpha\beta}$ appears explicitly in the definition of $\mathcal{O}$), and $I = \mathscr{L}'/e^2$.

The first term on the right-hand side of (3.26) vanishes if $\{Q, \mathcal{O}\} = 0$, for then

$$\int (\mathcal{D}X) e^{-I} \left(-\frac{1}{e^2} \{Q, \int \sqrt{g} g^{\alpha\beta} \lambda_{\alpha\beta}\} \right) \cdot \mathcal{O}$$

$$= -\frac{1}{e^2} \langle \{Q, (\int \sqrt{g} g^{\alpha\beta} \lambda_{\alpha\beta}) \cdot \mathcal{O}\} \rangle$$

$$= 0. \tag{3.27}$$

[Notice that we are using (3.5).] The second term in (3.26) vanishes if $\mathcal{O}$ has no explicit dependence on $g_{\alpha\beta}$, or more generally if $\delta\mathcal{O} = \{Q, \varrho\}$ for some ϱ.

If $\mathcal{O}$ obeys these conditions, then $Z(\mathcal{O})$ is a topological invariant. However, $Z(\mathcal{O})$ will vanish if $\mathcal{O} = \{Q, \varrho\}$ for some ϱ, since then

$$Z(\mathcal{O}) = \langle \{Q, \varrho\} \rangle = 0. \tag{3.28}$$

Thus, topological invariants will come from operators $\mathcal{O}$ such that $\{Q, \mathcal{O}\} = 0$, modulo those of the form $\mathcal{O} = \{Q, \varrho\}$, and with $\mathcal{O}$ obeying the extra condition

$\delta_g \mathcal{O} = \{Q, \varrho\}$ ($\delta_g \mathcal{O}$ being the change in $\mathcal{O}$ under a change in $g_{\alpha\beta}$). In our actual examples, we will always have $\delta_g \mathcal{O} = 0$.

Looking back to the supersymmetry variations in (2.8), it is easy to find operators that obey these criteria. The spin zero field ϕ is BRST invariant, does not depend explicitly on the metric, and (being the only local field of scaling dimension zero) cannot be written as $\{Q, \varrho\}$. Of course, ϕ itself is not gauge invariant, but invariant polynomials in ϕ such as $\mathrm{Tr}\,\phi^2$, $\mathrm{Tr}\,\phi^4$, etc., are gauge invariant as well as obeying our other criteria. The number of independent invariant polynomials is equal to the rank of G; they correspond to the independent Casimir operators. For $G = SU(2)$, there is only one, which we may take to be

$$W_0(P) = \tfrac{1}{2}\,\mathrm{Tr}\,\phi^2(P). \tag{3.29}$$

Here P denotes a point in M; we are emphasizing by the notation that W_0 is a local operator that depends on the choice of a point P. Note that W_0 has $U = 4$.

We can now define some new topological invariants. Let the manifold M and the bundle E be such that $d(\mathcal{M}) = 4k$. Pick k points $P_1, \ldots, P_k$ on M, and define

$$Z(k) = \int (\mathcal{D}X) e^{-I} \cdot \prod_{i=1}^{k} W_0(P_i) = \langle W_0(P_1) \ldots W_0(P_k) \rangle. \tag{3.30}$$

Then $Z(k)$ is independent of the choice of metric on M by virtue of the discussion above. It is also independent of the choice of points $P_1, \ldots, P_k$, since the choice of k points has no intrinsic significance independent of a choice of metric.

While this argument shows that $Z(k)$ is a topological invariant, it is very illuminating to check more explicitly that $Z(k)$ is independent of the choice of points $P_1, \ldots, P_k$. To this aim we differentiate $W_0(P)$ with respect to the coordinates x^α of P, and find

$$\frac{\partial}{\partial x^\alpha} W_0 = \frac{\partial}{\partial x^\alpha}\left(\frac{1}{2}\,\mathrm{Tr}\,\phi^2\right) = \mathrm{Tr}\,\phi D_\alpha \phi = i\{Q, \mathrm{Tr}\,\phi\psi_\alpha\}. \tag{3.31}$$

Thus, although W_0 is not a BRST commutator, its derivative is. It follows from (3.31) that

$$W_0(P) - W_0(P') = \int_{P'}^{P} \frac{\partial W_0}{\partial x^\alpha} dx^\alpha = i\left\{Q, \int_{P'}^{P} W_1\right\}, \tag{3.32}$$

where W_1 is the operator valued 1 form $W_1 = \mathrm{Tr}(\phi\psi_\alpha) \cdot dx^\alpha$. We now see that

$$\left\langle (W_0(P) - W_0(P')) \cdot \prod_j W_0(P_j) \right\rangle = \left\langle \left\{Q, i\int_{P'}^{P} W_1 \cdot \prod_j W_0(P_j)\right\} \right\rangle = 0. \tag{3.33}$$

Here, of course, we have used the fact that $\{Q, W_0\} = 0$, and we have again used (3.5).

The key equations so far have been

$$0 = i\{Q, W_0\}, \qquad dW_0 = i\{Q, W_1\} \tag{3.34}$$

with W_0 and W_1 being a zero form and one form on M, respectively. This process has an important generalization. One finds recursively

$$dW_1 = i\{Q, W_2\}, \qquad dW_2 = i\{Q, W_3\},$$
$$dW_3 = i\{Q, W_4\}, \qquad dW_4 = 0 \tag{3.35}$$

with

$$W_2 = \text{Tr}(\tfrac{1}{2}\psi \wedge \psi + i\phi \wedge F),$$
$$W_3 = i\,\text{Tr}(\psi \wedge F), \qquad W_4 = -\tfrac{1}{2}\text{Tr}(F \wedge F). \tag{3.36}$$

In these formulas, ϕ, ψ, and F are regarded as zero, one, and two forms on M. W_k for $0 \leq k \leq 4$ is a k form. Notice that W_k has $U = 4 - k$.

If γ is a k dimensional homology cycle on M, consider the integral

$$I(\gamma) = \int_\gamma W_k. \tag{3.37}$$

This integral is BRST invariant, since

$$\{Q, I\} = \int_\gamma \{Q, W_k\} = -i\int_\gamma dW_{k-1} = 0. \tag{3.38}$$

In addition, up to a BRST commutator, I depends only on the homology class of γ. For if γ is a boundary, say $\gamma = \partial\beta$, then

$$I(\gamma) = \int_\gamma W_k = \int_\beta dW_k = i\int_\beta \{Q, W_{k+1}\} = i\left\{Q, \int_\beta W_{k+1}\right\}. \tag{3.39}$$

This formula is to be seen as the generalization of (3.32) from zero cycles (points P and P') to k cycles. It says that if γ is trivial in homology, then $I(\gamma)$ is trivial in the BRST sense.

Now we are ready to propose quantum field theory formulas for the general Donaldson invariants. Let M, G, and E be such that $d(\mathcal{M}) \geq 0$. Pick homology cycles $\gamma_1 \ldots \gamma_r$ of dimensions $k_1 \ldots k_r$, such that

$$\sum_{i=1}^{r} (4 - k_r) = d(\mathcal{M}).$$

This formula ensures that $\prod_{i=1}^{r} W_i$ has $U = d(\mathcal{M})$. Then let

$$Z(\gamma_1, \ldots, \gamma_r) = \int (\mathscr{D}X) \exp(-\mathscr{L}'/e^2) \cdot \prod_{i=1}^{r} \int_{\gamma_i} W_{k_i} = \left\langle \prod_{i=1}^{r} \int_{\gamma_i} W_{k_i} \right\rangle. \tag{3.40}$$

This integral is a topological invariant by virtue of our standard arguments [including a use of (3.5), (3.38), and (3.40) to show that (3.40) depends only on the homology classes of the γ_i]. Of course, if the group G is other than $SU(2)$, we can write similar formulas beginning with $W_0' = \text{Tr}\,\phi^4$ or other invariant polynomials in ϕ. This corresponds precisely to the fact that a vector bundle with a rank r gauge group has r essentially independent characteristic classes, each of which can be used in principle in constructing Donaldson invariants, though so far the interesting applications have come from the second Chern class. In Sect. 5 we will extract from the quantum field theory viewpoint some explicit formulas for the Donaldson invariants as integrals over the instanton moduli space.

4. Hamiltonian Treatment and Floer Theory

In the last section, we worked on an arbitrary four manifold M, with no preferred "time" direction. As a result, there was no natural Hamiltonian formalism, and we

have used Feynman path integral techniques, manipulating the BRST charge Q in a way that is familiar in string theory. In this section we will specialize to the case $M = Y \times R^1$, with Y a three manifold and R^1 corresponding to "time." In this situation, we will discuss the Hamiltonian formalism. We will in the process see how to recover the results anticipated in the nonrelativistic treatment of [7].[14]

In the Hamiltonian formalism, one constructs a Hilbert space H, a Hamiltonian H, and a fermionic charge Q obeying $Q^2 = 0$, $[Q, H] = 0$. One of the main interests is to study the Q cohomology groups, that is the kernel of Q modulo its image (a Q cohomology class is an equivalence class of states ψ such that $Q\psi = 0$, the equivalence relation being that $\psi \simeq \psi + Q\lambda$ for any λ). The cohomology groups are the quantum ground states and are precisely the Floer groups. These groups are graded by the global quantum number U introduced in Sect. 2. We recall that $[U, Q] = +Q$ and that U is conserved modulo a constant; for $SU(2)$ this constant is 8, corresponding to the mod 8 grading of the Floer groups.

The assertion that the Q cohomology groups are finite dimensional and consist only of the quantum ground states may at first surprise string theorists. In string theory, the cohomology of the more or less analogous operator Q_{BRST} consists of the whole infinite spectrum of physical states. The difference is that in string theory Q_{BRST} acts nonlinearly on the fields. [In fact, Q_{BRST} is cubic in oscillators; if c, b are the conformal ghosts and X the matter field then $Q_{\mathrm{BRST}} \sim c\partial cb + c(\partial X)^2$.] On the other hand, in the twisted supersymmetric theory that we are considering,[15] Q acts linearly on the fields (i.e., it acts in a nondegenerate fashion even in a linearized approximation). At non-zero momentum every field has a superpartner, as is usual in supersymmetric theories, and they cancel out upon constructing the cohomology, leaving only the quantum ground states.

One way to prove that the cohomology groups correspond to the ground states involves an analogue of "Hodge theory." For $M = Y \times R^1$, we will find, in addition to the supercharge Q, a second operator $\bar{Q}$ obeying

$$\{Q, \bar{Q}\} = 2H. \tag{4.1}$$

$\bar{Q}$ will also obey $(\bar{Q})^2 = 0$. From (4.1), a standard argument, which we will discuss later, shows that the cohomology consists of the ground states.

The key to finding $\bar{Q}$ is the equation that played a central role in Sect. 3, namely

$$T_{\alpha\beta} = \{Q, \lambda_{\alpha\beta}\}. \tag{4.2}$$

For $M = Y \times R^1$, the Hamiltonian is defined as

$$H = \int d^3 x\, T_{00}. \tag{4.3}$$

---

[14] In discussing the Hamiltonian formalism, we will use the form (2.13) of the Lagrangian. It is true that adding $\int \mathrm{Tr}\, F \wedge F$ with a very precise coefficient [as in (2.41)], so as to cancel the classical instanton action, makes the path integral treatment of the last section more elegant. However, this term with the coefficient in (2.41) would be very embarrassing in a Hamiltonian treatment (it corresponds in physical terms to an imaginary θ angle, $\theta \sim 1/e^2$). Floer theory definitely seems to correspond to quantization of (2.13), not (2.41)

[15] Recall from Sect. 2 that for $M = R^4$, Q reduces to one of the eight supercharges in an $N = 2$ supersymmetric gauge theory

Evidently then, we can find an operator $\bar{Q}$ obeying $\{Q, \bar{Q}\} = 2H$ simply by choosing

$$\bar{Q} = 2\int d^3x\lambda^{00}. \tag{4.4}$$

Let us now show that $[H, \bar{Q}] = 0$. We recall from Sect. 2 that $D_\alpha\lambda^{\alpha\beta} = -D_\alpha(U)^{\alpha\beta}$ with $(U)^{\alpha\beta} = -(U)^{\beta\alpha}$. So

$$-i[H, \bar{Q}] = \frac{d\bar{Q}}{dt} = 2\int_Y d^3x\frac{\partial\lambda^{00}}{\partial t} = -2\int_Y d^3xD_i(\lambda^{0i} - U^{0i}) = 0. \tag{4.5}$$

Here we are using the fact that $U^{00} = 0$ (since $U^{\alpha\beta} = -U^{\beta\alpha}$) and that the integral of a total divergence over the compact manifold Y is zero.

With $H = \frac{1}{2}\{Q, \bar{Q}\}$, the fact that $[H, \bar{Q}] = 0$ means that $[Q, \bar{Q}^2] = 0$. It is actually true in the case at hand that $\bar{Q}^2 = 0$. To see this in the most transparent way, let us write out the Lagrangian of Eq. (2.13) (without the topological term) in a $3 + 1$ dimensional language. Thus, with $M = Y \times R^1$, we have

$$\mathcal{L} = \int dt \int_M d^3x\sqrt{g}\,\mathrm{Tr}\left[\frac{1}{2}F_{0i}F_{0i} + \frac{1}{2}\tilde{F}_{0i}\tilde{F}_{0i} - \frac{1}{2}D_0\phi D_0\lambda - \frac{1}{2}D_i\phi D_i\lambda + i\varepsilon^{ijk}(D_j\psi_k)\chi_i\right.$$

$$+ i(D_0\psi_i)\chi_i + i\psi_0 D_i\chi^i - i\eta D_0\psi_0 - i\eta D_i\psi^i$$

$$\left. - \frac{i}{2}\phi[\chi_i, \chi_i] - \frac{i}{2}\lambda[\psi_i, \psi_i] - \frac{i}{2}\lambda[\psi_0, \psi_0] - \frac{i}{2}\phi[\eta, \eta] - \frac{1}{8}[\phi, \lambda]^2\right]. \tag{4.6}$$

Here t (parametrizing R^1) is "time," "0" denotes the time direction, $i, j, k = 1, 2, 3$ run over a basis of the tangent space to Y, $\chi^i = \chi^{0i}$, and $\tilde{F}_{0i} = \frac{1}{2}\varepsilon_{ijk}F^{jk}$. We have taken $Y \times R^1$ with signature $(+ + + +)$ in writing the above. It is easy to see that (4.6) has a symmetry under $t \to -t$ together with

$$\phi \to \lambda, \quad \lambda \to \phi, \quad \psi_i \to \chi_i,$$
$$\chi_i \to -\psi_i, \quad \eta \to -\psi_0, \quad \psi_0 \to \eta. \tag{4.7}$$

Let us denote this operation as T. It is easy to see that $T^2 = (-1)^F$ (the latter being the operation that changes the sign of all anticommuting fields).

Since T (mapping $t \to -t$) is a time reversal symmetry, it will be realized in quantum field theory as an anti-unitary operation. This means that the Floer groups, rather than just being complex vector spaces, have a real structure. (Of course, they actually have an integral structure, but this is not evident from the quantum field theory viewpoint that we are developing here.)

Now, the explicit formulas for Q and $\bar{Q}$ may be determined from $Q = \int_Y J^0$, $\bar{Q} = 2\int_Y \lambda^{00}$ (with J^μ the conserved supercurrent found in Sect. 2). One finds

$$Q = \int_Y \mathrm{Tr}[(F_{0i} + \tilde{F}_{0i})\psi_i - \eta D_0\phi - D_i\phi\chi_i - \psi_0[\lambda, \phi]/2],$$
$$\bar{Q} = \int_Y \mathrm{Tr}[(F_{0i} - \tilde{F}_{0i})\chi_i + \psi_0 D_0\lambda - \psi_i D_i\lambda + \eta[\phi, \lambda]/2]. \tag{4.8}$$

We see that under T,

$$\bar{Q} \to Q, \quad Q \to -\bar{Q}. \tag{4.9}$$

Therefore, the fact that $Q^2 = 0$ implies that also $\bar{Q}^2 = 0$.

Before discussing the Q cohomology, a few comments are in order. In the best of worlds, a quantum field theory has a Hamiltonian that is bounded below and a Hilbert space with a Lorentz invariant and positive metric. In the case at hand these properties do not hold. The indefiniteness of the (ϕ, λ) kinetic energy means that the Hamiltonian will not be bounded below. The Hilbert space inner product will be indefinite because of the indefiniteness of the $\eta D_0\psi_0$ and $D_0\psi_i \cdot \chi_i$ terms[16]. The former of these problems was already discussed at the end of Sect. 2, where it was pointed out that it can be avoided if one sets $\phi = -\lambda^*$, and for the present purposes we will accept this.

Let us temporarily postpone worrying about the positivity of the norm and review the standard Hodge theory argument relating the cohomology to the ground states of the Hamiltonian (which is positive semi-definite since we are setting $\phi = -\lambda^*$). Since $[Q, H] = 0$, we can represent the cohomology classes by H eigenstates. Given such a class ($Q\psi = 0$), if $H\psi = \lambda\psi$ with $\lambda \neq 0$, then since $H = \frac{1}{2}\{Q, \bar{Q}\}$ we get $\psi = Q\left(\frac{1}{2\lambda}\bar{Q}\psi\right)$ so ψ is trivial in cohomology. Hence the cohomology classes are zero eigenstates of H. Conversely, if $H\psi = 0$, then $0 = \frac{1}{2}\langle\psi|H|\psi\rangle = \frac{1}{2}|Q|\psi\rangle|^2 + \frac{1}{2}|\bar{Q}|\psi\rangle|^2$, so $Q|\psi\rangle = \bar{Q}|\psi\rangle = 0$. In particular, ψ represents a Q cohomology class. And if $\psi \neq 0$, this class is not zero. For as $[H, Q] = 0$, if ψ can be written as $Q\alpha$ we can assume $H\alpha = 0$, but as we have seen $H\alpha = 0$ implies $Q\alpha = 0$ and so $\psi = 0$.

Clearly, the proof that cohomology classes give quantum ground states does not depend on positivity of the scalar product, but the proof that quantum ground states are annihilated by Q and $\bar{Q}$ does. Since in the theory of interest, the natural Lorentz invariant scalar product is not positive definite, some discussion of the validity of the above argument is required. Let us first describe the computation of the space of quantum ground states for small coupling.

For small coupling, the quantum Hilbert space is straightforwardly constructed by expanding around the classical minima of the potential. Because of a term $\text{Tr} F_{ij}F_{ij}$ in the energy, a classical minimum corresponds to $F_{ij} = 0$, that is, to a flat connection. Once we pick a flat connection, we must choose ϕ and λ so $D_i\phi = D_i\lambda = [\phi, \lambda] = 0$, to set the scalar contribution to the energy to zero.

If we pick a flat connection that is "irreducible" (there is no subgroup of the gauge group G that leaves it invariant), the conditions in the last sentence require $\phi = \lambda = 0$. If, in addition, the flat connection is "isolated" (no zero modes of A_i^a), life is very simple. There being no bosonic zero modes (and, by supersymmetry, no fermionic ones), the quantization will give rise to a unique quantum ground state for each isolated, irreducible flat connection. The value of U for this state (and thus its "dimension" in Floer theory) must be determined by computing the fermion normal ordering constant. Perturbative corrections cannot give this mode a non-zero energy, since this would violate the invariance of the Euler characteristic [or $\text{Tr}(-1)^F$]. Instanton corrections lead precisely to Floer's considerations.

---

[16] The reason is as follows. Consider a general Lagrangian with real anticommuting fields α_i whose time derivatives enter the Lagrangian only via a term $\mathscr{L}_F = \int dt i M_{ij}\alpha_i D_0\alpha_j$, with M_{ij} a constant symmetric matrix. Quantization will give the anticommutation relations $\{\alpha_i, \alpha_j\} = (M^{-1})_{ij}$, and this permits the α_i to be Hermitian operators in a Hilbert space of positive metric only if the matrix M_{ij} is positive definite

For connections that are not isolated and irreducible there are bosonic (and fermionic) zero modes, and it will be a more subtle problem to determine the quantum ground states. This is the counterpart in the present framework of the well known problems in Floer theory with reducible and non-isolated flat connections. Roughly speaking, for flat connections that are irreducible but not isolated, there are zero modes of A_i but not of ϕ and λ. The flat connections will then form a moduli space of positive dimension, and just as in finite dimensional degenerate Morse theory, the evaluation of the Floer groups will involve the cohomology of the space of flat connections. But for reducible connections, there are ϕ and λ zero modes, and one will meet new phenomena, perhaps of a subtle quantum field theoretic nature.

The important question is now whether the quantum ground states are really annihilated by Q, as would follow from the Hodge theory argument if the scalar product on the quantum Hilbert space were positive definite. In fact, it is quite straightforward to see in perturbation theory that this is so. An isolated flat connection is annihilated by Q classically [since the right-hand side of (2.8) is zero if the connection is flat and all other fields are zero]. Expanding around an isolated flat connection, Q is quadratic (plus higher orders) in oscillators, and certainly annihilates the ground state. The structure is just as predicted by Hodge theory, even though the Lorentz invariant inner product is not positive definite.

A natural way to explain this seems to be that one can define a modified inner product which is positive definite but not Lorentz invariant and which perhaps is the appropriate one to consider in the "Hodge theory" argument.[17] If $(\ |\ )_L$ is the Lorentz invariant scalar product on the quantum field theory Hilbert space $\mathscr{H}$, one can define a new inner product $(\ |\ )_+$ by saying that for $u, v \in \mathscr{H}$,

$$(u|v)_+ = (u|Tv)_L, \tag{4.10}$$

where T is the time reversal operation which was introduced in Eq. (4.7). The idea behind (4.10) is that while η, for example, is self-adjoint in the sense of $(\ |\ )_L$, its adjoint in the sense of $(\ |\ )_+$ is $-\psi_0 = T\eta$. Quantization of the Lagrangian (2.13) shows that the canonical conjugate of η is $-\psi_0$, so that a positive metric on $\mathscr{H}$ must be one in which $-\psi_0$ is the adjoint of η. Notice that in the sense of $(\ |\ )_+$, Q and $\bar{Q}$ are adjoints of one another, as the Hodge theory argument requires [in the sense of $(\ |\ )_L$, Q, and $\bar{Q}$ are each self-adjoint]. Thus, it may well be that $(\ |\ )_+$ is the proper structure for use in the Hodge theory argument. It might even be appropriate to turn this argument around in the following sense. In showing that the nonlinear quantum field theory under study really does exist, the positive definite inner product may be the right one to use. One would then introduce the Lorentz invariant one via (4.10) at the end of the construction in order to achieve Lorentz invariance.

4.1. Relation of Donaldson and Floer Theory

The next issue that we should discuss is the connection of Floer and Donaldson theory. According to [7], to define Donaldson invariants of a four manifold M

---

[17] Such a situation has also arisen recently in work by D. Olive on a Hodge theoretic interpretation of the "no ghost" theorem of string theory

with boundary B one must specify a state in the Floer homology of B. (As explained in the introduction, this fact was really the motivation for the present paper.) In the context of quantum field theory, the relation of Donaldson invariants to Floer homology has the following interpretation (which was anticipated by Atiyah).

In quantum field theory on a closed four manifold M, the nicest path integrals are of the form

$$Z(\mathcal{O}) = \int (\mathcal{D}X) e^{-I} \cdot \mathcal{O} \qquad (4.11)$$

with I the action, and $\mathcal{O}$ a product of local fields (usually polynomials). "X" is an abbreviation for the whole collection of integration variables. It is well known, though, that if M has a non-empty boundary B, the path integral requires a "boundary condition" on B. Such a boundary condition may consist simply of specifying the values of the field on B. More generally, one picks an arbitrary state in the Hilbert space $\mathcal{H}$ of the quantum theory formulated on B (or more exactly, the theory formulated on $B \times R^1$ as in our above discussion). If $X|_B$ represents the restriction of the whole collection of integration variables to B, then $\mathcal{H}$ is a certain space of functionals of the $X|_B$, and a state in $\mathcal{H}$ corresponds to a functional $\Psi(X_B)$. The path integral "with boundary conditions determined by Ψ" is just

$$Z(\mathcal{O}, \Phi) = \int (\mathcal{D}X) \exp(-\mathcal{L}'/e^2) \cdot \mathcal{O} \cdot \Psi(X_B). \qquad (4.12)$$

Now we can ask, For what $\mathcal{O}$ and Ψ is (4.12) a topological invariant? The arguments of Sect. 3 show that we need $Q\mathcal{O} = 0 = Q\Psi$. Thus Ψ represents a Floer cohomology class. Moreover, with $Q\mathcal{O} = 0$, the arguments in Sect. 3 show that (4.12) is zero if $\Psi = Q\Lambda$ for some Λ; therefore, (4.12) depends only on the Floer cohomology class represented by Ψ. On the other hand, the interesting choices for $\mathcal{O}$ are precisely the ones that we considered in the case that M had no boundary, namely

$$\mathcal{O} = \prod_i \int_{\gamma_i} W_{k_i}. \qquad (4.13)$$

Here γ_i are certain cohomology classes on M, and the W_{k_i} were constructed in Sect. 3. Thus, in (4.12) we obtain Donaldson polynomials with values in (the dual of) the Floer groups of B.

As an example (described to me by Atiyah and part of the inspiration behind the present paper), let B consist of several connected components B_i. Choose the Floer classes on the B_i so that one may take $\mathcal{O} = 1$. The connected components of B can be considered roughly as incoming and outgoing "three-branes" (higher dimensional generalizations of strings). In this situation, (4.12) can thus be considered roughly as a "three-brane scattering amplitude." Further thoughts along these lines are one route to certain speculations about the physical interpretation of Donaldson theory which can be found in Sect. 6.

There is a slight modification of (4.12) which is also significant (and related to a recent axiomatization of conformal field theory [27]). Let us group the connected components of B into "incoming three-branes" B_i and "outgoing three-branes" $\tilde{B}_j$. Suppose we are given a functional $\Psi(X|_{B_i})$ of the boundary values of the fields on the B_i. Then the path integral can be used to compute a functional of the fields on the $\tilde{B}_j$, by

$$\tilde{\Psi}(X') = \int_{(X|_{\tilde{B}_j} = X')} (\mathcal{D}X) \exp(-\mathcal{L}'/e^2) \Psi(X|_{B_i}). \qquad (4.14)$$

[This formula requires some explanation. The integral in (4.14) is carried out over all fields X whose restriction to the $\tilde{B}_j$ is equal to some given field X'. The dependence of the integral on X' gives a functional of X' which we are calling $\tilde{\Psi}$.] Moreover, $Q\tilde{\Psi}=0$ if $Q\Psi=0$; in fact, if $Q\Psi=0$, then everything on the right-hand side of (4.14) is Q-invariant. Thus, $\Psi \to \tilde{\Psi}$ is a morphism of the tensor product of the Floer groups of the B_i to that of the $\tilde{B}_j$.

A specialization of this gives what is perhaps the nicest way, within the framework of the present paper, to show that the Floer groups are topological invariants (and depend, that is, on the three manifold Y but not on a choice of metric). Let $M=Y\times R$, with R denoting the real line, parametrized by a "time" variable t, with $-\infty < t < +\infty$. Pick on M a metric of the form

$$ds^2 = dt^2 + g_{ij}(x^k, t)\, dx^i dx^j, \tag{4.15}$$

with x^k being coordinates on Y.

If we are given a metric g on Y, let $HF^*(Y; g)$ denote the Floer groups of Y computed with this particular metric. Suppose we are given two metrics $g^{(1)}$ and $g^{(2)}$, and we wish to compare the corresponding Floer groups $HF^*(Y; g^{(1)})$ and $HF^*(Y; g^{(2)})$. To do so, pick on $M=Y\times R$ a metric of the form (4.15), with the additional requirement that $g_{ij}(x^k, t)$ goes over to $g^{(1)}$ for $t \ll 0$ and to $g^{(2)}$ for $t \gg 0$. For every Ψ in $HF^*(Y; g^{(1)})$, (4.14) computes a corresponding $\tilde{\Psi}$ in $HF^*(Y; g^{(2)})$. Let us denote this linear transformation from $HF^*(Y; g^{(1)})$ to $HF^*(Y; g^{(2)})$ as $W(g^{(2)}, g^{(1)})$. Our standard arguments (using the fact that $T_{\alpha\beta} = \{Q, \lambda_{\alpha\beta}\}$ for some λ) show that W is independent of the detailed choice of the t dependence of the metric in (4.15).

The W's may readily be seen to obey some formal properties which imply that they are isomorphisms. If $g^{(1)} = g^{(2)}$, we have

$$W(g^{(2)}, g^{(1)}) = 1. \tag{4.16}$$

For if $g^{(2)} = g^{(1)}$, then W can be computed by picking a time independent metric in (4.15). W is then equal to e^{-Ht}, with $t \to \infty$. But we know that H annihilates the Floer groups, so $W = 1$ in cohomology.

And given three metrics $g^{(1)}$, $g^{(2)}$, and $g^{(3)}$, the product $W(g^{(3)}, g^{(2)}) \cdot W(g^{(2)}, g^{(1)})$ can be evaluated simply by modifying the t dependence of the metric in (4.15). One simply chooses $g_{ij}(x^k, t)$ in (4.15) to equal $g^{(1)}$ for $t \to -\infty$, $g^{(2)}$ for $t \sim 0$, and $g^{(3)}$ for $t \to +\infty$. This gives a path integral recipe for computing the two step transition $W(g^{(3)}, g^{(2)}) \cdot W(g^{(2)}, g^{(1)})$, but it also is clearly the definition of $W(g^{(3)}, g^{(1)})$. We therefore have

$$W(g^{(3)}, g^{(1)}) = W(g^{(3)}, g^{(2)}) \cdot W(g^{(2)}, g^{(1)}). \tag{4.17}$$

Equations (4.16) and (4.17) imply [on picking $g^{(3)} = g^{(1)}$] that $1 = W(g^{(1)}, g^{(2)}) \cdot W(g^{(2)}, g^{(1)})$, so that the W's are invertible and therefore are isomorphisms. Equation (4.17) means then that the Floer groups are independent of metric up to the canonical isomorphisms given by the W's.

5. Differential Forms on Instanton Moduli Space

In Sect. 3, we saw how to represent the Donaldson polynomial invariants of a four manifold M as correlation functions in quantum field theory. At the end of Sect. 4,

we observed that this picture has a generalization when M has a non-trivial boundary B. To make the quantum field theory viewpoint mathematically rigorous is of course a formidable undertaking. However, we will see in this section that one can extract from the quantum field theory viewpoint concrete formulas for differential forms on the moduli space of instantons whose integrals over instanton moduli space are the Donaldson invariants. Though inspired by quantum field theory, these are perfectly rigorous, classical formulas, whose key properties can be verified by classical manipulations. (One can extract from the quantum field theory a prescription for the requisite manipulations; I will indicate this but not carry it out in detail.) From the topological point of view, the formulas presented here are presumably just a way of reexpressing Donaldson's ideas. From an analytical point of view, however, they may be useful in overcoming some of the technical problems in Donaldson theory. In this discussion, we will ignore the singularities of instanton moduli space. Thus, our discussion will not be complete. In general, the singularities make additional contributions to the BRST invariant correlators, and quantum field theory should give a prescription for evaluating these.

We choose on M a vector bundle E such that $d(\mathcal{M})$, the formal dimension of the instanton moduli space $\mathcal{M}$, is positive. We write $n = d(\mathcal{M})$. For simplicity, we will consider only the case that at a generic point in instanton moduli space the instantons have no symmetries, and the formal dimension of instanton moduli space equals the actual dimension. These assumptions [which are readily obeyed for $G = SU(2)$] mean that there are no (ϕ, λ) zero modes; the only zero modes of the gauge field A_α are the tangents to $\mathcal{M}$; and the only fermion zero modes are the ψ_α zero modes which (related by supersymmetry to zero modes of A_α) also represent tangents to $\mathcal{M}$.

Now, because of the ψ_α zero modes, the partition function

$$Z = \int (\mathcal{D}X) \exp(-\mathcal{L}'/e^2) \tag{5.1}$$

is zero. Non-zero path integrals are of the form

$$Z(\mathcal{O}) = \int (\mathcal{D}X) \exp(-\mathcal{L}'/e^2) \cdot \mathcal{O}, \tag{5.2}$$

where $\mathcal{O}$ has $U = d(\mathcal{M})$, so as to be able to absorb the fermion zero modes.

In the weak coupling approximation (adequate because we are evaluating topological invariants), $(\mathcal{D}X)$ reduces (after integrating out the non-zero modes) to

$$d\mu = da_1 \ldots da_n d\psi_1 \ldots d\psi_n, \tag{5.3}$$

where $a_i, \psi_j, i, j = 1 \ldots n$ are the boson and fermion zero modes. It is important to note that there is a canonical measure $d\mu$ because $d\psi_i$ transforms oppositely to da_i under any change of basis. In the weak coupling limit, $\exp(-\mathcal{L}'/e^2)$ reduces to one, and the functional integral over the non-zero modes equals ± 1, as we noted in Sect. 3; it can be discarded if $\mathcal{M}$ is connected.

We now must study the functional $\mathcal{O}$. In general, non-zero modes as well as zero modes may be present in $\mathcal{O}$. If non-zero modes are present, we must "integrate them out" to get an effective functional $\mathcal{O}'$ only depending on the zero modes. It is of the general form

$$\mathcal{O}' = \Phi_{i_1 \ldots i_n}(a^k) \cdot \psi^{i_1} \ldots \psi^{i_n}. \tag{5.4}$$

Φ is an antisymmetric tensor with n indices – otherwise known as an n-form – on the n-dimensional manifold $\mathcal{M}$. Replacing $\mathcal{O}$ by $\mathcal{O}'$ and inserting (5.3) and (5.4) in (5.2), we get

$$Z(\mathcal{O}) = \int da_i \dots da_n d\psi_1 \dots d\psi_n \Phi_{i_1 \dots i_n} \psi^{i_1} \dots \psi^{i_n} = \int_{\mathcal{M}} \Phi. \qquad (5.5)$$

Thus, computing a correlation function $Z(\mathcal{O})$ in the weak coupling limit amounts to integrating the non-zero modes out of $\mathcal{O}$ so as to get an n-form Φ on instanton moduli space.

Now suppose that $\mathcal{O}$ is a product,

$$\mathcal{O} = \mathcal{O}_1 \cdot \mathcal{O}_2 \cdot \dots \cdot \mathcal{O}_k \qquad (5.6)$$

with $\mathcal{O}_k$ having $U = n_k$, and $\sum_k n_k = n$. By integrating out the non-zero modes from any of the $\mathcal{O}_r$, one gets an object

$$\mathcal{O}'_r = \Phi^{(r)}_{i_1 \dots i_{n_r}} \psi^{i_1} \dots \psi^{i_{n_r}}. \qquad (5.7)$$

Here $\Phi^{(r)}_{i_1 \dots i_{n_r}}$ can be interpreted as an n_r form on $\mathcal{M}$. The process leading from $\mathcal{O}_r$ to $\mathcal{O}'_r$ is just analogous to that leading from $\mathcal{O}$ to $\mathcal{O}'$. Naively, we might hope that

$$\mathcal{O}' = \mathcal{O}'_1 \cdot \mathcal{O}'_2 \cdot \dots \cdot \mathcal{O}'_k. \qquad (5.8)$$

In general, there is no reason for (5.8) to be true, because in integrating the non-zero modes out of the product $\mathcal{O}_1 \cdot \mathcal{O}_2 \cdot \dots \cdot \mathcal{O}_k$, one might need to make Wick contractions between $\mathcal{O}_i$ and $\mathcal{O}_j$ for $i \neq j$. However, it will often be the case that (5.8) is valid to lowest order in e^2, and this is all we need topologically. In the situation under study here, fortune smiles and (5.8) is valid.

Equation (5.8) is equivalent to the statement that the differential forms Φ and $\Phi^{(i)}$ described above are related by

$$\Phi = \Phi^{(1)} \wedge \Phi^{(2)} \wedge \dots \wedge \Phi^{(k)}. \qquad (5.9)$$

When (5.8) is valid for all products of interesting operators $\mathcal{O}_\alpha$, a particularly simple prescription can be given for computing integrals

$$Z(\mathcal{O}_{\alpha_1} \cdot \mathcal{O}_{\alpha_2} \dots \mathcal{O}_{\alpha_s}) = \int (\mathcal{D}X) \exp(-\mathcal{L}'/e^2) \mathcal{O}_{\alpha_1} \cdot \mathcal{O}_{\alpha_2} \dots \mathcal{O}_{\alpha_s}. \qquad (5.10)$$

One extracts from each $\mathcal{O}_\alpha$ (by integrating out the non-zero modes) a differential form $\Phi^{(\alpha)}$ (of appropriate degree) on $\mathcal{M}$. Then

$$Z(\mathcal{O}_{\alpha_1} \dots \mathcal{O}_{\alpha_n}) = \int_{\mathcal{M}} \Phi^{(\alpha_1)} \wedge \Phi^{(\alpha_2)} \wedge \dots \wedge \Phi^{(\alpha_n)}. \qquad (5.11)$$

In our study of Donaldson theory in Sect. 3, the interesting operators were

$$\mathcal{O}^{(\gamma)} = \int_\gamma W_{k_\gamma}. \qquad (5.12)$$

Here γ is a homology cycle on M, of dimension k_γ, and the W_{k_γ} for $k_\gamma = 0, \dots, 4$ are differential forms on M of degree k_γ defined as follows:

$$W_0 = \tfrac{1}{2} \operatorname{Tr} \phi^2,$$

$$W_1 = \operatorname{Tr}(\phi \wedge \psi),$$

$$W_2 = \operatorname{Tr}(\tfrac{1}{2} \psi \wedge \psi + i\phi \wedge F), \qquad (5.13)$$

$$W_3 = i \operatorname{Tr}(\psi \wedge F),$$

$$W_4 = -\tfrac{1}{2} \operatorname{Tr}(F \wedge F).$$

W_k has $U = 4 - k$, so with each $\mathcal{O}^{(\gamma)}$ we should associate a $4 - k_\gamma$ form $\Phi^{(\gamma)}$ on $\mathcal{M}$. Equation (5.8) is valid for arbitrary products of the $\mathcal{O}^{(\gamma)}$, so the Donaldson invariants are simply

$$Z(\mathcal{O}^{(\gamma_1)} \dots \mathcal{O}^{(\gamma_s)}) = \int_{\mathcal{M}} \Phi^{(\gamma_1)} \wedge \dots \wedge \Phi^{(\gamma_s)}. \tag{5.14}$$

To obtain "explicit formulas for Donaldson theory," we need only integrate the non-zero modes out of the W's.

This is easily done. Whenever F appears in (5.13), we may, to lowest order in e^2, simply replace it by the classical instanton field. Whenever ψ appears in (5.13), we simply replace it by zero mode wave functions.[18] Therefore, all that needs to be done is to integrate out ϕ from the W's.

In doing this, the relevant terms in the action are

$$\frac{\mathcal{L}'}{e^2} = \int_M d^4x \sqrt{g} \left[\frac{1}{2e^2} \operatorname{Tr} \phi D_\alpha D^\alpha \lambda - \frac{i}{2e^2} \lambda[\psi_\alpha, \psi^\alpha] + \dots \right]. \tag{5.15}$$

"Integrating out ϕ" means computing the integral

$$\langle \phi^a(x) \rangle = \int \mathcal{D}\phi \mathcal{D}\lambda \exp -(\mathcal{L}'/e^2) \cdot \phi^a(x)$$

$$= \int \mathcal{D}\phi \mathcal{D}\lambda \exp \left[-\frac{1}{2e^2} \int \sqrt{g} \, \phi D_\alpha D^\alpha \lambda \right]$$

$$\times \phi^a(x) \cdot \int_M \frac{i}{2e^2} \operatorname{Tr} \lambda[\psi_\alpha, \psi^\alpha] + \dots. \tag{5.16}$$

(We have expanded $\exp[-(i/2e^2) \int \operatorname{Tr} \lambda[\psi, \psi]]$ to extract the term linear in λ, which is the only one that survives after integrating over ϕ and λ. The $+ \dots$ on the right of (5.16) is irrelevant to lowest order in e^2.) The ϕ and λ dependence reduces to the Gaussian integral

$$\int \mathcal{D}\phi \mathcal{D}\lambda \exp \left[-\frac{1}{2e^2} \int \sqrt{g} \, \phi D_\alpha D^\alpha \lambda \right] \cdot \phi^a(x) \lambda^b(y) \tag{5.17}$$

which we will denote as $\langle \phi^a(x) \lambda^b(y) \rangle$.[19] According to the rules of Gaussian integration,

$$\langle \phi^a(x) \lambda^b(y) \rangle = -2e^2 G^{ab}(x, y), \tag{5.18}$$

where $G^{ab}(x, y)$ is the Green's function of the Laplacian $\Delta = D_\alpha D^\alpha$. G is defined as the unique solution of

$$\Delta G^{ab}(x, y) = \delta^{ab} \delta^4(x - y). \tag{5.19}$$

---

[18] To be more precise, pick a basis $u_{(1)\alpha}{}^a(x) \dots u_{(n)\alpha}{}^a(x)$ of classical zero mode wave functions. Then let $\psi_\alpha{}^a(x) = \sum_i \psi^i u_{(i)\alpha}{}^a(x)$, where ψ^i are the fermion zero mode coordinates that appear in (5.3). The same substitution $\psi_\alpha{}^a(x) = \sum_i \psi^i u_{(i)\alpha}{}^a(x)$ is to be understood in subsequent formulas. The choice of the $u_{(i)}$ does not matter because (using the same basis of A_α zero modes) it cancels out of $d\mu = da_1 \dots da_n d\psi_1 \dots d\psi_n$

[19] We ignore the (ϕ, λ) determinant, which we know will cancel other determinants of fields not written in (5.17)

Substituting (5.18) in (5.16), we learn that

$$\langle \phi^a(x) \rangle = -i \int_M d^4 y \sqrt{g} \, G^{ab}(x,y) \, [\psi_\alpha(y), \psi^\alpha(y)]_b \,. \tag{5.20}$$

Note that the factors of e^2 have cancelled out, a crucial test. Equation (5.20), with ψ_α replaced by its zero modes, is the required formula expressing ϕ^a in terms of zero modes.

Replacing ϕ in (5.13) by $\langle \phi \rangle$ whenever it appears, we get our desired formulas for differential forms $\Phi^{(\gamma)}$ on $\mathcal{M}$ corresponding to homology cycles γ on M. If γ is a zero cycle, say a point P, then

$$\Phi^{(\gamma)} = \tfrac{1}{2} \operatorname{Tr} \langle \phi(P) \rangle^2 \,. \tag{5.21}$$

If γ is a one cycle, then

$$\Phi^{(\gamma)} = \int_\gamma \operatorname{Tr}(\langle \phi \rangle \wedge \psi) \,. \tag{5.22}$$

For a two cycle

$$\Phi^{(\gamma)} = \int_\gamma \operatorname{Tr}(\tfrac{1}{2} \psi \wedge \psi + i \langle \phi \rangle \wedge F) \,. \tag{5.23}$$

For a three cycle,

$$\Phi^{(\gamma)} = i \int_\gamma \operatorname{Tr}(\psi \wedge F) \,. \tag{5.24}$$

And if γ is a four cycle, say a multiple s of the fundamental class $[M]$ of M, Φ^γ is just a constant function (that is, a closed zero form) on $\mathcal{M}$. The constant is equal to $-s/2 \cdot \int_M \operatorname{Tr} F \wedge F$. In general a k cycle on M gives an operator of $U = 4 - k$, corresponding to Donaldson's map $H_k(M) \to H^{4-k}(\mathcal{M})$.

Some readers may have considered the reasoning in Sects. 3 and 4 to lack the precision of mathematics. However, we have arrived at perfectly concrete formulas for differential forms $\Phi^{(\gamma)}$ on instanton moduli space.

At this point, one may wonder what is involved in proving that the $\Phi^{(\gamma)}$ have the necessary properties so that the

$$Z(\gamma_1, \ldots, \gamma_r) = \int_{\mathcal{M}} \Phi^{(\gamma_1)} \wedge \Phi^{(\gamma_2)} \wedge \ldots \wedge \Phi^{(\gamma_r)} \tag{5.25}$$

are topological invariants. There are really four steps.

(a) One must show that the $\Phi^{(\gamma)}$ are closed. This is not self-evident, but it can be checked by a completely classical computation; there is no need for input from quantum field theory.

(b) We must show that $\Phi^{(\gamma)}$ changes by an exact form if we change γ in its homology class. It is enough to show that if γ is a boundary, say $\gamma = \partial \beta$, then $\Phi^{(\gamma)} = dt^{(\beta)}$ for some differential form $t^{(\beta)}$ on $\mathcal{M}$.

Here we actually get some useful insight from quantum field theory. If $t^{(\beta)}$ such that $\Phi^{(\gamma)} = dt^{(\beta)}$ exists, there are many such $t^{(\beta)}$, and we would like a "best" choice, so as to be able to push Donaldson theory to its limits and deal with the singularities and non-compactness of $\mathcal{M}$ to the extent possible. Quantum field theory actually

gives a canonical but not entirely obvious formula for $t^{(\beta)}$. If $\gamma = \partial\beta$, then

$$\mathcal{O}^{(\gamma)} = \int_\gamma W_{k_\gamma} = \int_\beta dW_{k_\gamma} = i\left\{Q, \int_\beta W_{k_\gamma+1}\right\}. \tag{5.26}$$

In the small e^2 limit, $\mathcal{O}^{(\gamma)}$ reduces to our differential form $\Phi^{(\gamma)}$, Q reduces to the exterior derivative d on $\mathcal{M}$, and $W_{k_\gamma+1}$ reduces (using the same formulas given above) to a differential form $t^{(\beta)}$ on $\mathcal{M}$. Equation (5.26) is the desired formula $\mathcal{O}^{(\gamma)} = dt^{(\beta)}$.

(c) One must show that under a change in the metric $g_{\alpha\beta}$ of M, the $Z(\gamma_1, \ldots, \gamma_r)$ are invariant. Here again quantum field theory gives canonical formulas that are not completely obvious. Before writing formulas, let us express the problem precisely. Consider a family of metrics on M parametrized by a parameter space S. Let $X = M \times S$ be the total space of the family. Denote the fiber of X above a point $s \in S$ as M_s. Each M_s has a metric g_s. The differential forms $\Phi^{(\gamma)}$ are defined on each fiber; let us denote them as $\Phi^{(\gamma)}(s)$. To show that the $Z(\gamma_1, \ldots, \gamma_r)$ are independent of metric, we must exhibit closed differential forms $\hat{\Phi}^{(\gamma)}$ on X whose restrictions to the M_s coincide with the $\Phi^{(\gamma)}(s)$. This can immediately be done in a canonical way using our standard formulas. Under a displacement of $s \in S$, the metric of M changes. Let $\delta g_{\alpha\beta}(P)$ denote the change of the components $g_{\alpha\beta}(P)$ of the metric of M at a point $P \in M$ and in some basis of the tangent space to M at P. The $\delta g_{\alpha\beta}(P)$ are closed one forms on S.

From our general formulas, under a change in the metric of M, the change in $\Phi^{(\gamma)}$ is

$$\delta\Phi^{(\gamma)} = \left\langle \int_\gamma W_{k_\gamma} \cdot \frac{1}{2e^2} \int_M \sqrt{g}\, \delta g^{\alpha\beta} T_{\alpha\beta} \right\rangle, \tag{5.27}$$

where $\langle \ \rangle$ is an instruction to integrate out the non-zero modes. Using our favorite formula $T_{\alpha\beta} = \{Q, \lambda_{\alpha\beta}\}$, we can rewrite (5.27) in the form

$$\delta\Phi^{(\gamma)} = \left\{Q, \left\langle \int_\gamma W_{k_\gamma} \cdot \frac{1}{2e^2} \int_M \sqrt{g}\, \delta g^{\alpha\beta} \lambda_{\alpha\beta} \right\rangle\right\}. \tag{5.28}$$

Somewhat naively, (5.28) asserts that $\delta\Phi^{(\gamma)}$, which is the change of $\Phi^{(\gamma)}$ under a change in metric, is a BRST commutator, corresponding to a closed differential form on moduli space which will not contribute in (5.25). This is the conclusion we want, and the argument is essentially correct, but the above description is not a very canonical way to express things, because as the metric changes the moduli space also varies (that is, there is no canonical connection on the fiber bundle X). A better description is as follows. On $X = M \times S$, an n form can be decomposed as a sum of what we might call $(k, n-k)$ forms (a k form on M times an $n-k$ form on S). $\Phi^{(\gamma)}$ is in this sense an $(n, 0)$ form on X, for some n. On the other hand,

$$P = \left\langle \int_\gamma W_{k_\gamma} \cdot \frac{1}{2e^2} \int_M \sqrt{g}\, \delta g^{\alpha\beta} \lambda_{\alpha\beta} \right\rangle \tag{5.29}$$

can be understood as an $(n-1, 1)$ form. [Recall that δg is a one form on S or in other words a $(0, 1)$ form.] The exterior derivative in the fibers of $X = M \times S$ is what we have denoted $-iQ$, while δ is the exterior derivative from the base S. The

exterior derivative on the total space X is $d = \delta - iQ$. Equation (5.28), together with $\{Q, \Phi^{(\gamma)}\} = 0$ and $\delta P = 0$ (which can be verified using the form of λ given in Sect. 2), means that the differential form $\hat{\Phi}^{(\gamma)} = \Phi^{(\gamma)} - iP$ is annihilated by d, and this is the desired closed form on X whose restriction to the fibers gives back $\Phi^{(\gamma)}$.

   (d) Finally, of course, to make rigorous assertions about integrals

$$\int_{\mathcal{M}} \Phi^{(\gamma_1)} \wedge \Phi^{(\gamma_2)} \wedge \ldots \wedge \Phi^{(\gamma_r)} \tag{5.30}$$

one must know that the instanton moduli space exists, has singularities that are not too bad, and behaves not too badly under a change in metric. Such questions involve hard analysis [4, 5]. On these questions the viewpoint of this section offers no new insight, except the hope that analysis of the above formulas near singularities of $\mathcal{M}$ may give a new insight about the necessary criterion in "not too bad." Such a hope is supported by experience of physicists with instantons.

   In principle, in Sect. 3 we gave formulas for Donaldson invariants – as correlation functions in quantum field theory – which are valid regardless of whether instanton moduli space exists and what properties it has. From that more fundamental point of view, the considerations of this section are just a recipe to evaluate the correlation functions under favorable circumstances. But to make that fundamental point of view rigorous will indeed require considerable progress in constructive quantum field theory.

6. Physical Interpretation

In this concluding section, I would like to discuss the possible physical meaning of the present work. Here lie many of the most intriguing questions.

   The fermionic symmetry that we have used is very reminiscent of BRST symmetry. Its use is quite similar to the use of BRST symmetry in string theory. So it is natural to think that in a suitable framework, this symmetry arises upon BRST gauge fixing of an underlying gauge invariant theory.

   If so, that theory is of a very unusual kind. The single most important equation of this paper is the assertion that the stress tensor is a "BRST" commutator

$$T_{\alpha\beta} = \{Q, \lambda_{\alpha\beta}\}. \tag{6.1}$$

This assertion does not hold in ordinary BRST gauge fixing of ordinary gauge theories. It does hold in BRST gauge fixing of string theory, because in that case the starting point is the Nambu action, which is generally covariant in a two dimensional sense. This general covariance leads directly to (6.1). The lesson here is that (6.1) is a signal of general covariance. The "topological quantum field theory" that has been studied in this paper is very similar to a BRST gauge fixed version of an underlying generally covariant quantum field theory. To be more blunt about it, what we have been discussing should be viewed as a renormalizable, unitary (in the BRST sense) generally covariant quantum field theory in four dimensions.

   The reader may be surprised at attributing the property of general covariance to a theory without gravitons. The following comments may thus be illuminating. In general relativity, one begins with the gravitational field $g_{\alpha\beta}$ and a generally

covariant action. Then one expands around some classical field. One meets the fact that any metric (except $g_{\alpha\beta} = 0$, usually considered unlawful) is not invariant under diffeomorphisms. One usually expands around $g_{\alpha\beta} = \eta_{\alpha\beta}$ (the Minkowski metric), and this breaks general covariance down to the Poincaré symmetry. Morally speaking (not in any technical sense), the massless graviton should then be viewed as a Goldstone boson of spontaneously broken (local) general covariance. In a similar spirit, massless gauge bosons in gauge theories reflect the breaking of the local gauge invariance down to a global symmetry. If the local symmetry is unbroken, as in QCD, there are no massless gauge mesons.

Since the emergence of the concept of color confinement in QCD, it has been natural to wonder if in general relativity there could be an analogous phase in which general covariance is confined and unbroken.[20] At first sight, this concept seems paradoxical. For unbroken general covariance, there must be no metric tensor (or at least, it must have zero expectation value). Without a metric, one does not know how signals should propagate, so it seems that there can be no physics.

It seems that we have here stumbled upon a resolution of these paradoxes. With unbroken general covariance, there can indeed be no signal propagation and no local physics. Therefore, in quantization on a three manifold Y, the physical Hilbert spaces (in the BRST sense) come out to be global topological objects, the Floer groups, and the only observables that can be computed are global topological invariants, the Donaldson invariants (discussed in Sect. 3), and the relative Donaldson invariants (discussed in Sect. 4). (Recall from the end of Sect. 4 that the latter even include "three-brane scattering.")

Once it is accepted that the theory discussed in this paper is generally covariant, it is clear that there may be other more or less similar theories. It is indeed possible to find a version with explicit gravitational fields [29]. Perhaps there exists a version with spontaneous breaking of general covariance and dynamical gravitons.

One of the real mysteries is how to exhibit a manifestly generally covariant theory whose BRST gauge fixing (at least in some approximation) gives the "topological quantum field theory" we have considered. This is reminiscent of the situation in string theory, where the origin of space-time general covariance is rather obscure. It is rather clear that the theory considered here will not arise in BRST gauge fixing of a conventional generally covariant field theory with a finite number of fields. It would not be too surprising if it arises instead as a low energy approximation to some version of string field theory, in a phase in which general covariance is unbroken. (This possibility is supported to some extent by the existence [30] of $1 + 1$ dimensional sigma models with a BRST-like fermionic symmetry, and with the property that the graviton vertex operator is a BRST commutator. These sigma models may correspond to a realization of unbroken general covariance in string theory.) The proper incorporation of general covariance in string theory may have some unusual features that are reflected in the proper implementation of general covariance even in the low energy world, and

---

[20] This question has been considered over the years by many physicists, though there do not seem to be many published references. The possibility of unbroken general covariance in the context of string field theory has been considered in [28]

these features may be relevant to observable physics, perhaps to the vanishing of the cosmological constant.

References

1. Donaldson, S.: An application of gauge theory to the topology of four manifolds. J. Differ. Geom. **18**, 269 (1983); The orientation of Yang-Mills moduli spaces and 4-manifold topology. J. Differ. Geom. **26**, 397 (1987); Polynomial invariants for smooth four-manifolds. Oxford preprint
2. Freed, D., Uhlenbeck, K.: Instantons and four manifolds. Berlin, Heidelberg, New York: Springer 1984
3. Belavin, A., Polyakov, A., Schwartz, A., Tyupkin, Y.: Phys. Lett. B **59**, 85 (1975)
4. Taubes, C.: Self-dual Yang-Mills connections on non-self-dual 4-manifolds. J. Differ. Geom. **17**, 139 (1982)
5. Uhlenbeck, K.: Connections with L^p bounds on curvature. Commun. Math. Phys. **83**, 31 (1982). Removable singularities in Yang-Mills fields. Commun. Math. Phys. **83**, 11 (1982)
6. Floer, A.: An instanton invariant for three manifolds. Courant Institute preprint (1987); Morse theory for fixed points of symplectic diffeomorphisms. Bull. AMS **16**, 279 (1987)
7. Atiyah, M.F.: New invariants of three and four dimensional manifolds. In: The Symposium on the Mathematical Heritage of Hermann Weyl, Wells, R. et al. (eds.). (Univ. of North Carolina, May, 1987)
8. Braam, P.J.: Floer homology groups for homology three spheres. University of Utrecht Mathematics preprint 484, November, 1987
9. Witten, E.: Supersymmetry and morse theory. J. Differ. Geom. **17**, 661 (1982)
10. 't Hooft, G.: Computation of the quantum effects due to a four dimensional pseudoparticle. Phys. Rev. D **14**, 3432 (1976)
11. Jackiw, R., Rebbi, C.: Phys. Rev. Lett. **37**, 172 (1976)
12. Callan, C.G., Dashen, R., Gross, D.J.: Phys. Lett. **63** B, 334 (1976)
13. Atiyah, M.F., Hitchin, N., Singer, I.: Self-duality in Riemannian geometry. Proc. Roy. Soc. London A **362**, 425 (1978)
14. Affleck, I., Dine, M., Seiberg, N.: Dynamical supersymmetry breaking in supersymmetric QCD. Nucl. Phys. B **241**, 493 (1984); Dynamical supersymmetry breaking in four dimensions and its phenomenological implications. Nucl. Phys. B **256**, 557 (1985)
15. Seiberg, N.: IAS preprint (to appear)
16. Novikov, V.A., Shifman, M.A., Vainshtein, A.I., Zakharov, V.I.: Nucl. Phys. B **229**, 407 (1983) Amati, D., Konishi, K., Meurice, Y., Rossi, G.C., Veneziano, G.: Non-perturbative aspects in supersymmetric gauge theories. Physics Reports (to appear)
17. Friedan, D., Martinec, E., Shenker, S.: Nucl. Phys. B **271**, 93 (1986)
18. Peskin, M.: Introduction to string and superstring theory. SLAC-PUB-4251 (1987)
19. Green, M.B., Schwarz, J.H., Witten, E.: Superstring theory. Cambridge: Cambridge University Press 1987
20. Witten, E.: Global anomalies in string theory. In: Symposium on anomalies, geometry, and topology. White, A., Bardeen, W. (eds.), especially pp. 90–95. Singapore: World Scientific 1985
21. Becchi, C., Rouet, A., Stora, R.: The abelian Higgs-Kibble model, unitarity of the S-operator. Phys. Lett. **69** B, 309 (1974); Renormalization of gauge theories. Ann. Phys. **98**, 287 (1976)
22. Tyupin, I.V.: Gauge invariance in field theory and in statistical physics in the operator formalism. Lebedev preprint FIAN No. 39 (1975), unpublished
23. Kugo, T., Ojima, I.: Manifestly covariant canonical formulation of Yang-Mills theories. Phys. Lett. **73** B, 459 (1978); Local covariant operator formalism of non-abelian gauge theories and quark confinement problem. Supp. Prog. Theor. Phys. **66**, 1 (1979)

24. Polchinski, J.: Scale and conformal invariance in quantum field theory. Univ. of Texas preprint UTTG-22-87
25. D'Adda, A., DiVecchia, P.: Supersymmetry and instantons. Phys. Lett. **73** B, 162 (1978)
26. Witten, E.: An $SU(2)$ anomaly. Phys. Lett. **117** B, 432 (1982)
27. Segal, G.: Oxford preprint (to appear)
28. Horowitz, G.T., Lykken, J., Rohm, R., Strominger, A.: Phys. Rev. Lett. **57**, 283 (1986)
29. Witten, E.: Topological gravity. IAS preprint, February, 1988
30. Witten, E.: Topological sigma models. Commun. Math. Phys. (to appear)