Second International Conference on
Geometry, Integrability and Quantization
June 7–15, 2000, Varna, Bulgaria
Ivaïlo M. Mladenov and Gregory L. Naber, Editors
Coral Press, Sofia 2001

A SURVEY OF DONALDSON THEORY

GREGORY L. NABER

Department of Mathematics, California State University
Chico, CA 95929-0525, U.S.A.

Abstract. The purpose of this series of lectures is to provide an informal survey of the background in Donaldson theory required to understand the motivation behind Witten's construction (in 1988) of the first topological quantum field theory.

1. The Moduli Spaces

B will denote a compact, simply connected, oriented, smooth 4-manifold and we will write

$$SU(2) \hookrightarrow P_k \xrightarrow{\pi_k} B$$

for the principal $SU(2)$-bundle over B with Chern class k. $\mathcal{A}(P_k)$ is the set of all smooth (C^∞) connection 1-forms on P_k and $\mathcal{G}(P_k)$ is the group of all automorphisms of the bundle (diffeomorphisms f of P_k onto itself which satisfy $\pi_k \circ f = \pi_k$ and $f(p \cdot g) = f(p) \cdot g$ for all $p \in P_k$ and $g \in SU(2)$). $\mathcal{G}(P_k)$ is called the **gauge group**, its elements are called **(global) gauge transformations** and it acts on $\mathcal{A}(P_k)$ on the right by pullback, i. e., for any $\omega \in \mathcal{A}(P_k)$ and any $f \in \mathcal{G}(P_k)$ we have $\omega \cdot f = f^*\omega \in \mathcal{A}(P_k)$. Two connections ω and ω' are said to be **gauge equivalent** if there exists an $f \in \mathcal{G}(P_k)$ for which $\omega' = \omega \cdot f$. The set of gauge equivalence classes $[\omega]$ for $\omega \in \mathcal{A}(P_k)$ is called the **moduli space** and written $\mathcal{B}(P_k) = \mathcal{A}(P_k)/\mathcal{G}(P_k)$.

It is this moduli space that we wish to study. Unfortunately, it has no reasonable mathematical structure in the smooth context in which we have just introduced it. For this reason we must introduce appropriate Sobolev completions of the objects just described and this requires that some of the definitions be recast in other, but equivalent terms. Let us denote by $\Omega^i(P_k, su(2))$ the vector space of all i-forms on P_k with values in the Lie algebra $su(2)$. $\Omega^i_{\mathrm{ad}}(P_k, su(2))$ denotes the subspace consisting of those elements of $\Omega^i(P_k, su(2))$ that are "tensorial of type ad", i. e., satisfy the following two conditions: Any $\varphi \in \Omega^i_{\text{ad}}(P_k, su(2))$ vanishes if one of its arguments is vertical (i. e., tangent to a fiber in P_k). Moreover, if, for each $g \in SU(2)$, we consider the diffeomorphism $\sigma_g \colon P_k \to P_k$ given by $\sigma_g(p) = p \cdot g$, then $\sigma_g^* \varphi = g^{-1} \varphi g$ for each $\varphi \in \Omega^i_{\text{ad}}(P_k, su(2))$. Finally, we let $\Omega^i(B, \text{ad}\, P_k)$ denote the vector space of all i-forms on B with values in the adjoint bundle $\text{ad}\, P_k$ (the vector bundle associated to P_k by the adjoint (conjugation) action of $SU(2)$ on its Lie algebra $su(2)$). It is easy to show that $\Omega^i_{\text{ad}}(P_k, su(2))$ and $\Omega^i(B, \text{ad}\, P_k)$ are isomorphic. Our interest in this vector space is accounted for by the following Proposition.

Proposition 1.1. *$\mathcal{A}(P_k)$ is an affine space modelled on the vector space $\Omega^1_{\text{ad}}(P_k, su(2)) \cong \Omega^1(B, \text{ad}\, P_k)$.*

The proof is simple. One need only verify that, if ω and ω_0 are two connection 1-forms on P_k, then $\omega - \omega_0$ is tensorial of type ad. Thus, if we fix some ω_0 in $\mathcal{A}(P_k)$ we may write

$$\mathcal{A}(P_k) = \{\omega = \omega_0 + \varphi;\ \varphi \in \Omega^1_{\text{ad}}(P_k, su(2))\} \,. \tag{1.1}$$

Shortly, this description will permit us to define Sobolev completions of $\mathcal{A}(P_k)$. Next we consider the so-called nonlinear adjoint bundle $\text{Ad}\, P_k$. This is the fiber bundle associated to P_k by the adjoint (conjugation) action of $SU(2)$ on itself. Let $\Omega^0(B, \text{Ad}\, P_k)$ be the set of smooth sections of $\text{Ad}\, P_k$. It is a group under pointwise multiplication in the fibers $(SU(2))$ and is easily seen to be isomorphic to the group $\Omega^0_{\text{Ad}}(P_k, su(2))$ of smooth maps $\psi \colon P_k \to SU(2)$ that are equivariant, i. e., satisfy $\psi(p \cdot g) = g^{-1}\psi(p)g$ for all $p \in P_k$ and $g \in SU(2)$. We care about this because of the following Proposition.

Proposition 1.2. $\mathcal{G}(P_k) \cong \Omega^0_{\text{Ad}}(P_k, SU(2)) \cong \Omega^0(B, \text{Ad}\, P_k)$.

The proof is again simple. An automorphism $f \colon P_k \to P_k$ in $\mathcal{G}(P_k)$ preserves the fibers of P_k so, for each $p \in P_k$, there is a unique $\psi(p) \in SU(2)$ such that $f(p) = p \cdot \psi(p)$ and this defines the appropriate $\psi \in \Omega^0_{\text{Ad}}(P_k, SU(2))$. This alternative view of the gauge group will soon permit us to define its Sobolev completions.

We shall briefly recall how one defines Sobolev completions of a space of sections of a vector bundle. Begin with some principal bundle $G \hookrightarrow P \xrightarrow{\pi} B$ on which is defined a connection ω. Let $\mathcal{V}$ be a finite-dimensional vector space with an inner product and let $\rho \colon G \to GL(\mathcal{V})$ be an orthogonal representation of the structure group G on $\mathcal{V}$. Let $E = P \times_\rho \mathcal{V}$ be the associated vector bundle and $\Omega^i(B, E)$, $i = 0, 1, 2, \ldots$, the space of i-forms on B with values in E. In particular, $\Omega^0(B, E)$ is the space of sections of E. There are natural inner products defined on each $\Omega^i(B,E)$ and, moreover, ω induces a covariant exterior differentiation operator d^ω on each.

$$\Omega^0(B,E) \xrightarrow{\mathrm{d}^\omega} \Omega^1(B,E) \xrightarrow{\mathrm{d}^\omega} \Omega^2(B,E) \xrightarrow{\mathrm{d}^\omega} \cdots$$

Remark: Unlike the ordinary exterior derivative this sequence of operators does not form a complex in general. Indeed, when $E = \operatorname{ad} P$, the composition of the first two

$$\mathrm{d}^\omega \circ \mathrm{d}^\omega \colon \Omega^0(B, \operatorname{ad} P) \longrightarrow \Omega^2(B, \operatorname{ad} P)$$

is given by

$$\mathrm{d}^\omega \circ \mathrm{d}^\omega = [F_\omega, \cdot] \ , \tag{1.2}$$

where F_ω is the curvature of ω (just compute in coordinates).

Now, suppose $\xi \in \Omega^0(B,E)$. For each $m = 0,1,2,\ldots$, one defines the **Sobolev m-norm** $\|\xi\|_m$ of ξ by

$$\|\xi\|_m^2 = \sum_{j=0}^m \int_B \Big\| \big(\overbrace{\mathrm{d}^\omega \circ \cdots \circ \mathrm{d}^\omega}^{m} \big)(\xi) \Big\|^2 \mathrm{vol} \ .$$

This is, indeed, a norm on $\Omega^0(B,E)$ and different choices for the connection ω, the inner product on $\mathcal{V}$, etc. give rise to equivalent norms. The completion of $\Omega^0(B,E)$ in this norm is a Hilbert space $L_m^2(E)$. Sobolev embedding theorems guarantee that, by choosing m sufficiently large, one can achieve any desired degree of smoothness for the elements of $L_m^2(E)$. More precisely, if $m > \frac{1}{2}(\dim B) + l$, then $L_m^2(E)$ embeds in the space $C^l(B,E)$ of l-times continuously differentiable sections of E. Finally note that each $\Omega^i(B,E)$ is itself a space of sections of a vector bundle and so has Sobolev completions.

Since $\Omega^i(B, \operatorname{ad} P_k)$ is a space of sections of a vector bundle it has a Sobolev completion $\Omega^i_m(B, \operatorname{ad} P_k)$ for any $m = 0,1,2,\ldots$. For sufficiently large m its elements are all continuous sections so the isomorphism $\Omega^i(B, \operatorname{ad} P_k) \cong \Omega^i_{\mathrm{ad}}(P_k, su(2))$ serves to define $\Omega^i_{\mathrm{ad},m}(P_k, su(2))$. With this and (1.1) and taking $m = 3$ we define the Sobolev space of connections

$$\mathcal{A}_3(P_k) = \Big\{ \omega = \omega_0 + \varphi ; \ \varphi \in \Omega^1_{\mathrm{ad},3}(P_k, su(2)) \Big\} \ , \tag{1.3}$$

where ω_0 is any fixed smooth connection on P_k. Thus, $\mathcal{A}_3(P_k)$ is an affine space and therefore a Hilbert manifold whose tangent space at any ω is

$$T_\omega\left(\mathcal{A}_3(P_k)\right) = \Omega^1_{\mathrm{ad},3}\left(P_k, su(2)\right) \cong \Omega^1_3(B, \operatorname{ad} P_k) \ . \tag{1.4}$$

Sobolev completions of the gauge group $\mathcal{G}(P_k)$ take a bit more work since, by Proposition 1.2, it can be identified with the space of sections of (not a vector bundle, but) a fiber bundle with fiber $SU(2)$. But if we regard $SU(2)$ as a subset of the vector space $M_{2\times 2}(\mathbb{C})$ of 2×2 complex matrices, then $\mathrm{Ad}\, P_k$ embeds in the vector bundle $E = P_k \times_{SU(2)} M_{2\times 2}(\mathbb{C})$, where the action of $SU(2)$ on $M_{2\times 2}(\mathbb{C})$ is by conjugation. But the Sobolev space $L_4^2(E)$ is defined (and its elements are $\mathbb{C}^1$) so we may define $\mathcal{G}_4(P_k) = \{s \in L_4^2(E)\colon s(B) \subseteq \mathrm{Ad}\, P_k\}$. We will abuse the notation slightly and write

$$\mathcal{G}_4(P_k) = \Omega_4^0(B, \mathrm{Ad}\, P_k)\,.$$

This is a group under pointwise multiplication in the fibers of $\mathrm{Ad}\, P_k$. In fact, one can show ([6] or [12]) that $\mathcal{G}_4(P_k)$ is a Hilbert Lie group with Lie algebra (tangent space at the identity $\mathbf{1}$) that can be identified with

$$T_{\mathbf{1}}\left(\mathcal{G}_4(P_k)\right) = \Omega_4^0(B, \mathrm{ad}\, P_k)\,. \tag{1.5}$$

(this is easy enough to believe since the sections in $\Omega_4^0(B, \mathrm{ad}\, P_k)$ can be exponentiated pointwise to give elements of $\Omega_4^0(B, \mathrm{Ad}\, P_k)$).

The action of $\mathcal{G}(P_k)$ on $\mathcal{A}(P_k)$ extends to an action of $\mathcal{G}_4(P_k)$ on $\mathcal{A}_3(P_k)$ (same formulas since the elements of $\mathcal{G}_4(P_k)$ are differentiable and those of $\mathcal{A}_3(P_k)$ are continuous). It can be shown ([6] or [12]) that this action is smooth and that, if $\omega \in \mathcal{A}_3(P_k)$ is fixed, the map of $\mathcal{G}_4(P_k)$ to $\mathcal{A}_3(P_k)$ given by $f \to \omega \cdot f$ has a derivative at $\mathbf{1}$ that can be identified with

$$-\,\mathrm{d}^\omega\colon \Omega_4^0(B, \mathrm{ad}\, P_k) \longrightarrow \Omega_3^1(B, \mathrm{ad}\, P_k)\,.$$

In particular, the tangent space at ω to the orbit $\omega \cdot \mathcal{G}_4(P_k)$ of ω under $\mathcal{G}_4(P_k)$ is given by

$$T_\omega\left(\omega \cdot \mathcal{G}_4(P_k)\right) = \mathrm{Im}(\,\mathrm{d}^\omega) = \,\mathrm{d}^\omega\left(\Omega_4^0(B, \mathrm{ad}\, P_k)\right)\,. \tag{1.6}$$

Remark: Differential operators extend to bounded operators on Sobolev completions and this is the meaning of d^ω here and henceforth.

Given an $\omega \in \mathcal{A}_3(P_k)$ its **stabilizer** is the subgroup $\mathrm{Stab}(\omega)$ of $\mathcal{G}_4(P_k)$ that leaves ω fixed, i. e.,

$$\mathrm{Stab}(\omega) = \{f \in \mathcal{G}_4(P_k);\, \omega \cdot f = \omega\}\,.$$

Any such stabilizer contains the subgroup $\mathbb{Z}_2$ of $\mathcal{G}_4(P_k)$ generated by $\pm\mathbf{1}$ and if this is all it contains, i. e., if $\mathrm{Stab}(\omega) = \mathbb{Z}_2$, we will say that the connection ω is **irreducible**; otherwise, ω is **reducible**. The following characterization of reducibility is proved in [6] and [12].

Theorem 1.3. *The following are equivalent for any* $\omega \in \mathcal{A}_3(P_k)$.

a: ω *is reducible, i. e.,* $\mathrm{Stab}(\omega)/\mathbb{Z}_2$ *is nontrivial.*

b: $\mathrm{Stab}(\omega)/\mathbb{Z}_2 \cong U(1)$.

c: $\mathrm{d}^\omega : \Omega^0_4(B, \mathrm{ad}\, P_k) \longrightarrow \Omega^1_3(B, \mathrm{ad}\, P_k)$ *has nontrivial kernel.*

We will denote by $\hat{\mathcal{A}}_3(P_k)$ the subset of $\mathcal{A}_3(P_k)$ consisting of irreducible connections. It follows from Theorem 1.3 that irreducibility is invariant under gauge transformations.

With these preliminaries out of the way we can introduce the objects of real interest to us. The **moduli space**

$$\mathcal{B}_3(P_k) = \mathcal{A}_3(P_k)/\mathcal{G}_4(P_k)$$

of **connections** on P_k is the set of gauge equivalence classes of elements of $\mathcal{A}_3(P_k)$ modulo the action of $\mathcal{G}_4(P_k)$. Similarly, the orbit of space of $\hat{\mathcal{A}}_3(P_k)$ under the action of $\mathcal{G}_4(P_k)$ is the **moduli space**

$$\hat{\mathcal{B}}_3(P_k) = \hat{\mathcal{A}}_3(P_k)/\mathcal{G}_4(P_k)$$

of **irreducible connections** on P_k. Each of these is assumed to have the quotient topology and with this one can show ([6] and [12]) that $\mathcal{B}_3(P_k)$ is Hausdorff and $\hat{\mathcal{B}}_3(P_k)$ is open in $\mathcal{B}_3(P_k)$.

We investigate the local structure of these moduli spaces near each point $[\omega]$. First suppose $\omega \in \hat{\mathcal{A}}_3(P_k)$. We will produce a "slice" of the $\mathcal{G}_4(P_k)$-action on $\mathcal{A}_3(P_k)$ near ω, i. e., a submanifold $\mathcal{O}$ of $\mathcal{A}_3(P_k)$ such that

$$T_\omega\left(\mathcal{A}_3(P_k)\right) = T_\omega\left(\omega \cdot \mathcal{G}_4(P_k)\right) \oplus T_\omega(\mathcal{O})$$

and such that the restriction to $\mathcal{O}$ of the projection into the moduli space is injective near ω. Then the local structure of the moduli space near $[\omega]$ is the same as that of $\mathcal{O}$ near ω.

To produce $\mathcal{O}$ we choose a Riemannian metric g on B and an invariant inner product on $su(2)$. These give rise to natural inner products on the vector spaces $\Omega^i(B, \mathrm{ad}\, P_k)$ so that the operator $\mathrm{d}^\omega : \Omega^0(B, \mathrm{ad}\, P_k) \longrightarrow \Omega^1(B, \mathrm{ad}\, P_k)$ has a formal adjoint δ^ω.

$$\Omega^0(B, \mathrm{ad}\, P_k) \underset{\delta^\omega}{\overset{\mathrm{d}^\omega}{\leftrightarrows}} \Omega^1(B, \mathrm{ad}\, P_k)$$

(in fact, $\delta^\omega = -{}^*\mathrm{d}^\omega{}^*$, where $*$ is the Hodge dual corresponding to g and the given orientation of B). It turns out that

$$\delta^\omega \circ \mathrm{d}^\omega : \Omega^0(B, \mathrm{ad}\, P_k) \longrightarrow \Omega^0(B, \mathrm{ad}\, P_k)$$

is a (formally self-adjoint) elliptic operator. We use the same symbols for the extensions of these operators to the Sobolev completions $\Omega^0_4(B, \mathrm{ad}\, P_k)$ and $\Omega^1_3(B, \mathrm{ad}\, P_k)$. Elliptic theory (Generalized Hodge Decomposition Theorem) implies that $\ker(\delta^\omega \circ \mathrm{d}^\omega) = \ker(\mathrm{d}^\omega)$ is finite dimensional, $\mathrm{Im}(\mathrm{d}^\omega) = \ker(\delta^\omega)^\perp$, d^ω has closed range and that there is an orthogonal decomposition

$$\Omega^1_3(B, \mathrm{ad}\, P_k) = \mathrm{Im}(\mathrm{d}^\omega) \oplus \ker(\delta^\omega)\,,$$

i. e.,

$$T_\omega\left(\mathcal{A}_3(P_k)\right) = T_\omega\left(\omega \cdot \mathcal{G}_4(P_k)\right) \oplus \ker(\delta^\omega)\,.$$

Now, for each $\varepsilon > 0$, the submanifold

$$\mathcal{O}_{\omega,\varepsilon} = \{\omega + A;\ A \in \ker(\delta^\omega), \|A\|_3 < \varepsilon\} \tag{1.7}$$

clearly satisfies

$$T_\omega\left\{\mathcal{O}_{\omega,\varepsilon}\right) = \ker(\delta^\omega)$$

so

$$T_\omega\left(\mathcal{A}_3(P_k)\right) = T_\omega\left(\omega \cdot \mathcal{G}_4(P_k)\right) \oplus T_\omega\left(\mathcal{O}_{\omega,\varepsilon}\right)\,.$$

We claim that, for $\varepsilon > 0$ sufficiently small, $\mathcal{O}_{\omega,\varepsilon}$ projects injectively into the moduli space.

First note that, since $\omega \in \hat{\mathcal{A}}_3(P_k)$ and $\hat{\mathcal{A}}_3(P_k)$ is open in $\mathcal{A}_3(P_k)$, we may take ε small enough to ensure $\mathcal{O}_{\omega,\varepsilon} \subseteq \hat{\mathcal{A}}_3(P_k)$. Now consider the map

$$\Psi\colon \mathcal{G}_4(P_k) \times \mathcal{O}_{\omega,\varepsilon} \longrightarrow \hat{\mathcal{A}}_3(P_k)$$
$$\Psi\left(f, \omega'\right) = \omega' \cdot f\,.$$

The derivative of Ψ at $(\mathbf{1}, 0)$ is computed to be

$$(\mathrm{d}\Psi)_{(\mathbf{1},0)} : \Omega^0_4(B, \mathrm{ad}\, P_k) \oplus \ker(\delta^\omega) \longrightarrow \mathrm{Im}(\mathrm{d}^\omega) \oplus \ker(\delta^\omega)$$
$$(\mathrm{d}\Psi)_{(\mathbf{1},0)} = (\mathrm{d}^\omega, \mathrm{Id})\,.$$

This is obviously surjective. Moreover, because ω is assumed irreducible, Theorem 1.3(c) implies that it is injective. By the Open Mapping Theorem, $(\mathrm{d}\Psi)_{(\mathbf{1},0)}$ is an isomorphism. Thus, the (infinite dimensional version of the) Inverse Function Theorem implies that Ψ is a local diffeomorphism near $(\mathbf{1}, 0)$. Consequently, for some (perhaps smaller) $\varepsilon > 0$ there is an open neighborhood U_ω of ω in $\hat{\mathcal{A}}_3(P_k)$ and a set $\eta_{\mathbf{1},\varepsilon} = \{f \in \mathcal{G}_4(P_k);\ \|\mathbf{1} - f\|_4 < \varepsilon\}$ such that

$$\Psi\colon \eta_{\mathbf{1},\varepsilon} \times \mathcal{O}_{\omega,\varepsilon} \longrightarrow U_\omega$$

is a diffeomorphism. In particular, no two things in $\mathcal{O}_{\omega,\varepsilon}$ are gauge equivalent by any gauge transformation with ε of $\mathbf{1}$. A "bootstrapping" argument ([6] or [12]) now shows that, for a (possibly) still smaller $\varepsilon > 0$, no two things in $\mathcal{O}_{\omega,\varepsilon}$ are gauge equivalent by any gauge transformation. For such an ε,

$\mathcal{O} = \mathcal{O}_{\omega,\varepsilon}$ projects injectively into $\hat{\mathcal{B}}_3(P_k)$ and so is our slice and provides a local manifold structure for $\hat{\mathcal{B}}_3(P_k)$ near $[\omega]$. In particular, $\hat{\mathcal{B}}_3(P_k)$ has the structure of a smooth Hilbert manifold. Although $\mathcal{G}_4(P_k)$ does not act freely on $\hat{\mathcal{A}}_3(P_k)$, the action of $\tilde{\mathcal{G}}_4(P_k) = \mathcal{G}_4(P_k)/\mathbb{Z}_2$ on $\hat{\mathcal{A}}_3(P_k)$ is free and we have a smooth principal $\tilde{\mathcal{G}}_4(P_k)$-bundle

$$\tilde{\mathcal{G}}_4(P_k) \hookrightarrow \hat{\mathcal{A}}_3(P_k) \longrightarrow \hat{\mathcal{B}}_3(P_k)\,.$$

If $\omega \in \hat{\mathcal{A}}_3(P_k)$ is reducible the analysis is similar, except that to get an injective projection into the moduli space one must first factor out the action of the stabilizer of ω. More precisely, defining $\mathcal{O}_{\omega,\varepsilon}$ as in (1.7) and $\widetilde{\mathrm{Stab}}(\omega) = \mathrm{Stab}(\omega)/\mathbb{Z}_2 \cong U(1)$ one finds that, for sufficiently small ε, the projection

$$\mathcal{O}_{\omega,\varepsilon}/\widetilde{\mathrm{Stab}}(\omega) \longrightarrow \hat{\mathcal{B}}_3(P_k)$$

is a homeomorphism onto an open neighborhood of $[\omega]$ in $\mathcal{B}_3(P_k)$ which, in fact, is a diffeomorphism outside the fixed point set of $\widetilde{\mathrm{Stab}}(\omega)$.

The objects of real interest in Donaldson theory are certain subspaces of $\mathcal{B}_3(P_k)$ and $\hat{\mathcal{B}}_3(P_k)$ which we now introduce. Begin by selecting a Riemannian metric g on B. Together with the orientation of B this gives a Hodge star operation $*$ on forms defined on B and we wish to consider connections ω on P_k that are anti-self-dual (ASD) with respect to g in the sense that the curvature form F_ω satisfies ${}^*F_\omega = -F_\omega$ (we will emphasize the dependence on g by calling such a connection g-ASD). We note, however, that such connections can exist only if the Chern class k is non-negative. Indeed,

$$k = c_2(P_k) = \frac{1}{8\pi^2}\int_B \mathrm{Tr}\,(F_\omega \wedge F_\omega) = \frac{1}{8\pi^2}\int_B \left(|F_\omega^-|^2 - |F_\omega^+|^2\right)\mathrm{vol}\,,$$

where $F_\omega^\pm$ are the self-dual and anti-self-dual parts of F_ω. Consequently, if $k < 0$ one must have $F_\omega^+ \neq 0$. Moreover, if $k = 0$, any ASD connection is necessarily flat. We will therefore restrict our attention henceforth to the bundles with $k > 0$. For such k we define

$$\mathrm{Asd}_3(P_k, g) = \left\{\omega \in \mathcal{A}_3(P_k);\ \omega \ \text{ is } g\text{-ASD}\right\}$$
$$\widehat{\mathrm{Asd}}_3(P_k, g) = \left\{\omega \in \hat{\mathcal{A}}_3(P_k);\ \omega \ \text{ is } g\text{-ASD}\right\}$$

and the **moduli spaces**

$$\mathcal{M}(P_k, g) = \mathrm{Asd}_3(P_k, g)/\mathcal{G}_4(P_k)$$
$$\hat{\mathcal{M}}(P_k, g) = \widehat{\mathrm{Asd}}_3(P_k, g)/\mathcal{G}_4(P_k)$$

of **g-ASD** and **irreducible g-ASD** connections.

The curvature operator $F\colon \mathcal{A}(P_k) \to \Omega^2(B, \operatorname{ad} P_k)$ extends to a smooth map on the Sobolev completions ([6] or [12]) and, by continuity, the Hodge star is defined on $\Omega^2_2(B, \operatorname{ad} P_k)$ so our definitions make sense. Also note that the Sobolev indices have been dropped on $\mathcal{M}(P_k, g)$ and $\hat{\mathcal{M}}(P_k, g)$. The reason is that, for any $\omega \in \mathrm{Asd}_3(P_k, g)$, elliptic regularity implies that there is an $f \in \mathcal{G}_4(P_k, g)$ such that $\omega \cdot f$ is a smooth connection (see Section 5 of [17]). Thus, these moduli spaces do not depend on the choice of (sufficiently large) Sobolev index.

For a given B, g and k, $\mathrm{Asd}_3(P_k, g)$ (and therefore $\mathcal{M}(P_k, g)$) might well be empty. This is the case, for example, when B is either $\mathbb{CP}^2$ or $\mathbb{S}^2 \times \mathbb{S}^2$, $k = 1$ and g is the standard metric (Fubini–Study, in the case of $\mathbb{CP}^2$). Before pursuing the general analysis any further we pause to describe some examples in which the moduli space of ASD connections is non-empty.

Examples:

1. Let B be an arbitrary compact, simply connected, oriented, smooth 4-manifold and g any Riemannian metric on B. The $k = 0$ bundle $SU(2) \hookrightarrow P_0 \to B$ is trivial and, as we observed earlier, any ASD connection on P_0 is necessarily flat. Conversely, a flat connection is certainly ASD ($F_\omega = 0 \Rightarrow F_\omega^+ = 0$). Since flat connections exist on any trivial bundle, the moduli space $\mathcal{M}(P_0, g)$ is non-empty. Since B is simply connected, any two flat connections on P_0 are gauge equivalent (Proposition 2.2.3 of [5]) so $\mathcal{M}(P_0, g)$ is, in fact, a single point.

2. A much more interesting example is obtained when $B = \mathbb{S}^4$ with its standard metric and orientation and $k = 1$. It will be convenient to describe the results in terms of quaternions so we identify $\mathbb{R}^4 = \mathbb{H}$ and $SU(2) = Sp(1)$ (the group of unit quaternions). Let

$$\mathbb{S}^7 = \left\{ (q^1, q^2) \in \mathbb{H}^2 ;\ |q^1|^2 + |q^2|^2 = 1 \right\}$$

and define a right action of $Sp(1)$ on $\mathbb{S}^7$ by

$$(p, a) \in \mathbb{S}^7 \times Sp(1) \longrightarrow p \cdot a = (q^1, q^2) \cdot a = (q^1 a, q^2 a)\,.$$

The orbits $p \cdot Sp(1)$ are submanifolds of $\mathbb{S}^7$ diffeomorphic to $Sp(1) \cong \mathbb{S}^3$. The orbit space is, by definition, $\mathbb{HP}^1$ which is diffeomorphic to $\mathbb{S}^4$. Choosing a natural identification of $\mathbb{HP}^1$ with $\mathbb{S}^4$ we obtain the **quaternionic Hopf bundle**

$$Sp(1) \hookrightarrow \mathbb{S}^7 \longrightarrow \mathbb{S}^4\,.$$

This bundle admits a natural connection which, as an $sp(1) = \operatorname{Im} \mathbb{H}$-valued 1-form ω on $\mathbb{S}^7$ is the restriction to $\mathbb{S}^7$ of the 1-form

$$\operatorname{Im}\left(\bar{q}^1\, \mathrm{d}q^1 + \bar{q}^2\, \mathrm{d}q^2\right) \qquad \text{on } \mathbb{H}^2.$$

Remark: This connection is "natural" in the sense that, at each point of $\mathbb{S}^7$, its horizontal subspace (kernel) is that part of the $\mathbb{R}^8$-orthogonal complement of the tangent space to the fiber that lies in the tangent space to $\mathbb{S}^7$ (see [13]).

Trivializing the bundle over $\mathbb{S}^4 \backslash \{\text{north pole}\}$ and identifying this with $\mathbb{H}$ via stereographic projection one obtains a corresponding section (gauge) s and a coordinate expression for the gauge potential

$$\mathcal{A} = s^* \omega = \operatorname{Im}\left(\frac{\bar{q}}{1 + |q|^2} \, \mathrm{d}q \right) .$$

If $\Omega = \mathrm{d}\omega + \omega \wedge \omega$ is the curvature of ω, then the corresponding gauge field strength $\mathcal{F}$ is given by

$$\mathcal{F} = s^* \Omega = \mathrm{d}\mathcal{A} + \mathcal{A} \wedge \mathcal{A} = \frac{1}{\left(1 + |q|^2\right)^2} \, \mathrm{d}\bar{q} \wedge \mathrm{d}q \, .$$

This uniquely determines F_ω (since the trivialization covers all but one point of $\mathbb{S}^4$) and is ASD because

$$\begin{aligned} \mathrm{d}\bar{q} \wedge \mathrm{d}q = 2 \big[\left(\mathrm{d}x^0 \wedge \mathrm{d}x^1 - \mathrm{d}x^2 \wedge \mathrm{d}x^3 \right) \mathbf{i} + \left(\mathrm{d}x^0 \wedge \mathrm{d}x^2 + \mathrm{d}x^1 \wedge \mathrm{d}x^3 \right) \mathbf{j} \\ + \left(\mathrm{d}x^0 \wedge \mathrm{d}x^3 - \mathrm{d}x^1 \wedge \mathrm{d}x^2 \right) \mathbf{k} \big] \end{aligned}$$

and these components form a basis for the ASD 2-forms on $\mathbb{R}^4$. Since stereographic projection from the north pole is an orientation preserving diffeomorphism one can calculate the Chern number of the bundle by integrating $\operatorname{Tr}(\mathcal{F} \wedge \mathcal{F})$ over $\mathbb{R}^4$. The result is 1 (see [13]) so the quaternionic Hopf bundle is, indeed, the $k = 1$ bundle over $\mathbb{S}^4$ and we have one point $[\omega]$ in the moduli space $\mathcal{M}(P_1, g)$.

Remark: One can show that ω is irreducible. In fact, a general result asserts that, if B is a spin manifold (e. g., $\mathbb{S}^4$), then there are no reducible ASD connections on any P_k, $k > 0$, so $\mathcal{M}(P_k, g) = \hat{\mathcal{M}}(P_k, g)$ for any g.

One can generate new connection 1-forms on $Sp(1) \hookrightarrow \mathbb{S}^7 \to \mathbb{S}^4$ by pulling ω back by diffeomorphisms of $\mathbb{S}^7$ onto $\mathbb{S}^7$ that respect the $Sp(1)$-action. By a judicious choice of these diffeomorphisms one can produce, for each $\lambda > 0$ and each $n \in \mathbb{H}$, a connection $\omega_{\lambda,n}$ whose gauge potential and field strength are given by

$$\mathcal{A}_{\lambda,n} = \operatorname{Im}\left(\frac{\bar{q} - \bar{n}}{\lambda^2 + |q - n|^2} \, \mathrm{d}q \right)$$

and

$$\mathcal{F}_{\lambda,n} = \frac{\lambda^2}{\left(\lambda^2 + |q - n|^2\right)^2} \, \mathrm{d}\bar{q} \wedge \mathrm{d}q \, .$$

All of these are ASD (for the same reason that $\omega = \omega_{1,0}$ is ASD). A calculation (see [13]) shows that

$$\|\mathcal{F}_{\lambda,n}(q)\|^2 = \frac{48\lambda^2}{(\lambda^2 + |q-n|^2)^2}$$

which has a maximum of $\frac{48}{\lambda^2}$ at n. Notice that, for a fixed n, the curvature $\mathcal{F}_{\lambda,n}$ becomes more and more concentrated at n as $\lambda \to 0$. Moreover, since $\|\mathcal{F}_{\lambda,n}(q)\|$ is a gauge invariant, each $\omega_{\lambda,n}$ gives rise to a different point $[\omega_{\lambda,n}]$ in the moduli space.

Remark: n is called the **center** and λ the **scale** of $\omega_{\lambda,n}$. The connections $\omega_{\lambda,n}$ (or, more often, the potentials $\mathcal{A}_{\lambda,n}$) are called **BPST instantons** and were first discovered in the physics literature (see [2]).

Thus far we have a point $[\omega_{\lambda,n}]$ in the moduli space $\mathcal{M}(P_1, g)$ for each $\lambda > 0$ and each $n \in \mathbb{R}^4$. Atiyah, Hitchin and Singer [1] have proved that, in fact, every ASD connection on $Sp(1) \hookrightarrow \mathbb{S}^7 \to \mathbb{S}^4$ is gauge equivalent to some $\omega_{\lambda,n}$ so that the map

$$(\lambda, n) \in (0,\infty) \times \mathbb{R}^4 \longrightarrow [\omega_{\lambda,n}] \in \mathcal{M}(P_1, g)$$

is a bijection. It is, in fact, a homeomorphism (intuitively, nearby points in $(0,\infty) \times \mathbb{R}^4$ give rise to connections that are close in the sense that their field strengths are centered at nearby points in $\mathbb{R}^4$ and have approximately the same scale). Since any two differentiable structures on $(0,\infty) \times \mathbb{R}^4 \cong \mathbb{R}^5$ are diffeomorphic we identify $\mathcal{M}(P_1, g)$ with $(0,\infty) \times \mathbb{R}^4 \cong \mathbb{R}^5$.

There is a more instructive picture of $\mathcal{M}(P_1, g)$, however. Note that there is an orientation preserving conformal diffeomorphism of $(0,\infty) \times \mathbb{R}^4 \cong \mathbb{R}^5$ onto the open 5-dimensional ball B^5. One can parametrize the points of $\mathcal{M}(P_1, g)$ by the elements of B^5 in such a way that $[\omega] = [\omega_{1,0}]$ corresponds to the center of B^5 and with the property that moving radially out from $[\omega] = [\omega_{1,0}]$ toward a point $n \in \mathbb{S}^4 = \partial B^5$ one encounters the connections which (in coordinates obtained by stereographically projecting from the antipodal point $-n \in \mathbb{S}^4$) are the BPST instantons whose curvature concentrates at n ([13]). Notice that $\mathcal{M}(P_1, g)$ has a natural compactification $\bar{\mathcal{M}}(P_1, g)$ (namely, the closed ball $\bar{B}^5$) in which $\bar{\mathcal{M}}(P_1, g) - \mathcal{M}(P_1, g)$ is a copy of the base manifold $\mathbb{S}^4$ and that the points in this copy of $\mathbb{S}^4$ can be intuitively identified with instantons that are concentrated entirely at a single point of $\mathbb{R}^4$ (δ-function potentials). We shall see that this is quite a general phenomenon.

3. Let $\mathbb{CP}^2$ be the complex projective plane with its natural orientation as a complex manifold and $\overline{\mathbb{CP}}^2$ the same manifold with the opposite orientation. In both cases we take g to be the standard (Fubini–Study) metric. We have already mentioned that the $k = 1$ moduli space for $\mathbb{CP}^2$ is empty (the $k = 2$

moduli space is non-empty, however). For $\overline{\mathbb{CP}}^2$ on the other hand one can write out explicit formulas for representatives of each gauge equivalence class of ASD connections on the $k = 1$ bundle very much as we did for $\mathbb{S}^4$ above. The moduli space $\hat{\mathcal{M}}(P_1, g)$ turns out to be an open cone over $\overline{\mathbb{CP}}^2$. The base of the cone is a copy of $\overline{\mathbb{CP}}^2$ each point of which corresponds, as for $\mathbb{S}^4$, to a sequence of irreducible ASD connections becoming ever more concentrated at that point. The vertex of the cone corresponds to a sequence of irreducible ASD connections approaching a reducible connection. We shall find that reducible connections always give rise to such a cone structure in the moduli space. By adding the base and the vertex one again obtains a natural compactification of $\hat{\mathcal{M}}(P_1, g)$.

4. Taubes [21] has used the $k = 1$ instantons on $\mathbb{S}^4$ described in Example 2 to show that, if $b_2^+(B) = 0$, then the $k = 1$ bundle over B always admits g-ASD connections for any Riemannian metric g on B.

Remark: We briefly recall the definition of $b_2^+(B)$. In a compact, oriented, smooth 4-manifold B any element of $H_2(B, \mathbb{Z})$ can be represented by a smoothly embedded, oriented surface (2-manifold). The intersection form

$$q_B\colon H_2(B, \mathbb{Z}) \times H_2(B, \mathbb{Z}) \longrightarrow \mathbb{Z}$$

is defined as follows: Represent $\alpha, \beta \in H_2(B, \mathbb{Z})$ by surfaces Σ_α and Σ_β that intersect transversally. Thus, $\Sigma_\alpha \cap \Sigma_\beta$ is a finite set of points at each of which the tangent space to Σ_α and the tangent space to Σ_β together span the tangent space to B. An intersection point b is assigned the value $+1$ if an oriented basis for $T_b(\Sigma_\alpha)$ together with an oriented basis for $T_b(\Sigma_\beta)$ give an oriented basis for $T_b(B)$, and -1 otherwise. Then $q_B(\alpha, \beta)$ is the sum of these values over all intersection points. q_B is a unimodular, symmetric, bilinear form on $H_2(B, \mathbb{Z})$. If $b_2(B)$ is the rank of $H_2(B, \mathbb{Z})$, then $b_2(B) = b_2^+(B) + b_2^-(B)$, where $b_2^+(B)$ ($b_2^-(B)$) is the maximal dimension of a subspace of $H_2(B, \mathbb{Z})$ on which q_B is positive (negative) definite. One can show that $b_2^+(B)$ ($b_2^-(B)$) is also the dimension of the space of self-dual (anti-self-dual) harmonic 2-forms on B and this accounts for the role it plays in the study of ASD connections (we will say a bit more about this shortly).

Taubes' proof consists of an ingenious "grafting" procedure which can be very roughly described as follows: Let B denote a compact, simply connected, oriented, smooth 4-manifold with $b_2^+(B) = 0$. Any $SU(2)$-bundle over B can be constructed as the pullback of an $SU(2)$-bundle over $\mathbb{S}^4$ by some map $\Phi\colon B \to \mathbb{S}^4$. If the degree of Φ is 1, then the pullback of the $k = 1$ bundle over $\mathbb{S}^4$ is the $k = 1$ bundle over B. Taubes constructs such maps explicitly and, for sufficiently small $\lambda > 0$, pulls the instanton $\omega_{\lambda,n}$ back to B. These pullbacks acquire a "small" self-dual part and Taubes shows that a "small" perturbation

$(\omega \to \omega + \phi,\ \phi \in \Omega^1_{\mathrm{ad},3}(P_1, su(2)))$ of these connections kills the self-dual part. During the perturbation part of the proof Taubes requires a certain eigenvalue estimate (Theorem 7.16 of [6]). Assuming that the desired inequality is not satisfied he is able to produce a non-zero, harmonic, self-dual 2-form on B, thus contradicting $b_2^+(B) = 0$.

Now we return to the general study of the moduli spaces $\mathcal{M}(P_k, g)$ and $\hat{\mathcal{M}}(P_k, g)$ for $k > 0$. For this we define a smooth map

$$\mathrm{Pr}_+ \circ F \colon \mathcal{A}_3(P_k) \longrightarrow \Omega^2_{+,2}(B, \mathrm{ad}\, P_k)\,,$$

where F is the curvature map and Pr_+ projects onto the self-dual part. Thus,

$$(\mathrm{Pr}_+ \circ F)(\omega) = F_\omega^+$$

and

$$\mathrm{Asd}_3(P_k, g) = (\mathrm{Pr}_+ \circ F)^{-1}(0)\,. \tag{1.8}$$

At any $\omega \in \mathrm{Asd}_3(P_k, g)$ the derivative of this map can be identified with

$$\mathrm{d}_+^\omega = \mathrm{Pr}_+ \circ \mathrm{d}^\omega \colon \Omega^1_3(B, \mathrm{ad}\, P_k) \longrightarrow \Omega^2_{+,2}(B, \mathrm{ad}\, P_k)$$

(see page 54 of [6]). Now, we have already observed that, in general, $\mathrm{d}^\omega \circ \mathrm{d}^\omega$ is not zero, but is given by (1.2). However, if ω is ASD,

$$\mathrm{d}_+^\omega \circ \mathrm{d}^\omega = [F_\omega^+, \cdot] = [0, \cdot] = 0$$

so

$$\mathrm{Im}(\mathrm{d}^\omega) = \mathrm{d}^\omega\left(\Omega^0_4(B, \mathrm{ad}\, P_k)\right) \subseteq \ker\left(\mathrm{d}_+^\omega\right). \tag{1.9}$$

Consequently, we have associated with ω a complex $\mathcal{E}(\omega)$

$$0 \;\to\; \Omega^0_4(B, \mathrm{ad}\, P_k) \;\underset{\delta^\omega}{\overset{\mathrm{d}^\omega}{\leftrightarrows}}\; \Omega^1_3(B, \mathrm{ad}\, P_k) \;\underset{\delta_+^\omega}{\overset{\mathrm{d}_+^\omega}{\leftrightarrows}}\; \Omega^2_{+,2}(B, \mathrm{ad}\, P_k) \;\to\; 0$$

where we have included the adjoints δ^ω and δ_+^ω of d^ω and d_+^ω. This complex is, in fact, elliptic and our entire analysis of the local structure of the moduli space near $[\omega]$ is based on properties of $\mathcal{E}(\omega)$ (the so-called **fundamental elliptic complex** associated with $\omega \in \mathrm{Asd}_3(P_k, g)$). We begin by simply enumerating some consequences of the generalized Hodge Decomposition Theorem for elliptic complexes:

1. The Laplacians

$$\begin{aligned}
\Delta_0^\omega &= \delta^\omega \circ \mathrm{d}^\omega \\
\Delta_1^\omega &= \mathrm{d}^\omega \circ \delta^\omega + \delta_+^\omega \circ \mathrm{d}_+^\omega \\
\Delta_2^\omega &= \mathrm{d}_+^\omega \circ \delta_+^\omega
\end{aligned}$$

are all self-adjoint elliptic operators.

2. The spaces $\ker(\Delta_k^\omega)$, $k = 0, 1, 2$, of harmonic forms are finite dimensional and consist of smooth forms (elliptic regularity).

3. Each of the cohomology groups

$$
\begin{aligned}
H^0(\omega) &= \ker(\,\mathrm{d}^\omega) \\
H^1(\omega) &= \ker(\,\mathrm{d}_+^\omega)/\,\mathrm{Im}(\,\mathrm{d}^\omega) \\
H^2(\omega) &= \Omega^2_{+,2}(B, \mathrm{ad}\, P_k)/\,\mathrm{Im}(\,\mathrm{d}_+^\omega)
\end{aligned}
$$

associated with $\mathcal{E}(\omega)$ contains a unique harmonic representative. In particular,

$$
H^k(\omega) \cong \ker(\Delta_k^\omega)\,, \qquad k = 0, 1, 2,
$$

so all of these cohomology groups are finite-dimensional and we may define the **index** of $\mathcal{E}(\omega)$ by

$$
\begin{aligned}
\mathrm{Ind}(\mathcal{E}(\omega)) &= \dim\left(H^0(\omega)\right) - \dim\left(H^1(\omega)\right) + \dim\left(H^2(\omega)\right) \\
&= \dim(\ker(\Delta_0^\omega)) - \dim(\ker(\Delta_1^\omega)) + \dim(\ker(\Delta_2^\omega))\,.
\end{aligned}
$$

4. There are orthogonal decompositions

$$
\begin{aligned}
\Omega^0_4(B, \mathrm{ad}\, P_k) &= \mathrm{Im}(\Delta_0^\omega) \oplus \ker(\Delta_0^\omega) = \mathrm{Im}(\delta^\omega) \oplus \ker(\,\mathrm{d}^\omega) \\
\Omega^1_3(B, \mathrm{ad}\, P_k) &= \mathrm{Im}(\Delta_1^\omega) \oplus \ker(\Delta_1^\omega) = \mathrm{Im}(\,\mathrm{d}^\omega) \oplus \ker(\delta^\omega) \\
\Omega^2_{+,2}(B, \mathrm{ad}\, P_k) &= \mathrm{Im}(\Delta_2^\omega) \oplus \ker(\Delta_2^\omega) = \mathrm{Im}(\,\mathrm{d}_+^\omega) \oplus \ker(\delta_+^\omega)\,.
\end{aligned}
$$

We put this information to use in the following way. Fix $\omega \in \mathrm{Asd}_3(P_k, g)$ and restrict the map $\mathrm{Pr}_+ \circ F$ to a slice $\mathcal{O}_{\omega,\epsilon}$ through ω.

$$
\mathrm{Pr}_+ \circ F \mid \mathcal{O}_{\omega,\epsilon} \colon \mathcal{O}_{\omega,\epsilon} \longrightarrow \Omega^2_{+,2}(B, P_k)\,.
$$

Then

$$
(\mathrm{Pr}_+ \circ F \mid \mathcal{O}_{\omega,\epsilon})^{-1}(0) = \mathrm{Asd}_3(P_k, g) \cap \mathcal{O}_{\omega,\epsilon} \tag{1.10}
$$

and the derivative at ω is

$$
\mathrm{d}_+^\omega | \ker(\delta^\omega) \colon\ \ker(\delta^\omega) \longrightarrow \Omega^2_{+,2}(B, P_k)\,. \tag{1.11}
$$

We will show that this map is always Fredholm and determine conditions under which it is surjective, thus setting up an application of the (infinite dimensional version of the) Implicit Function Theorem to obtain a smooth manifold structure for $\mathrm{Asd}_3(P_k, g) \cap \mathcal{O}_{\omega,\epsilon}$ near ω. If, in addition, the projection of $\mathrm{Asd}_3(P_k, g) \cap \mathcal{O}_{\omega,\epsilon}$ into the moduli space is injective (i. e., if ω is irreducible) this will give a smooth manifold structure for the moduli space near $[\omega]$.

Remark: Recall that a bounded linear map $T\colon H_1 \to H_2$ between two Hilbert spaces is **Fredholm** if either of the following two equivalent conditions is satisfied:

1. $\dim(\ker T) < \infty$, $\dim(\ker T^*) < \infty$ and $\operatorname{Im} T$ is closed.
2. $H_1 = \ker T \oplus \operatorname{Im} T^*$ and $H_2 = \ker T^* \oplus \operatorname{Im} T$.

The version of the **Implicit Function Theorem** to which we will appeal asserts the following: Let X and Y be Hilbert manifolds, $F\colon X \to Y$ a smooth map and $x_0 \in X$ a point at which the derivative $Df_{x_0}\colon T_{x_0}(X) \to T_{F(x_0)}(Y)$ is a surjective Fredholm map. Then there exists an open neighbourhood of x_0 in $F^{-1}(F(x_0))$ that is a smooth manifold of dimension $\dim(\ker Df_{x_0})$.

To see that $\mathrm{d}_+^\omega | \ker(\delta^\omega)$ is Fredholm we reason as follows: First, $\ker(\mathrm{d}_+^\omega | \ker(\delta^\omega)) = \ker(\mathrm{d}_+^\omega)/\operatorname{Im}(\mathrm{d}^\omega) = H^1(\omega)$ which is finite-dimensional by consequence #3 of the Hodge theorem above. Next observe that $(\mathrm{d}_+^\omega | \ker(\delta^\omega))^* = \delta_+^\omega | \operatorname{Im}(\mathrm{d}_+^\omega | \ker(\delta^\omega)) = \delta_+^\omega | \operatorname{Im}(\mathrm{d}_+^\omega)$ so

$$\ker((\mathrm{d}_+^\omega | \ker(\delta^\omega))^*) = \ker(\delta_+^\omega | \operatorname{Im}(\mathrm{d}_+^\omega))$$

which is finite dimensional because $\ker(\delta_+^\omega) = H^2(\omega)$ is finite dimensional. Finally, $\operatorname{Im}(\mathrm{d}_+^\omega | \ker(\delta^\omega)) = \operatorname{Im}(\mathrm{d}_+^\omega) = (\ker(\delta_+^\omega))^\perp$ is closed.

The map $\mathrm{d}_+^\omega | \ker(\delta^\omega)$ is not always surjective. To determine when it is surjective we proceed as follows: Since d_+^ω acts on $\Omega_3^1(B, \operatorname{ad} P_k) = \operatorname{Im}(\mathrm{d}^\omega) \oplus \ker(\delta^\omega)$ and, since d_+^ω vanishes identically on $\operatorname{Im}(\mathrm{d}^\omega)$, $\mathrm{d}_+^\omega | \ker(\delta^\omega)$ is surjective if and only if d_+^ω is surjective, i. e., if and only if $\operatorname{Im}(\mathrm{d}_+^\omega) = \Omega_{+,2}^2(B, \operatorname{ad} P_k)$. But this is the case if and only if $H^2(\omega)$ is trivial. Noting that $H^0(\omega) = \ker(\mathrm{d}^\omega)$ is trivial if and only if ω is irreducible (Theorem 1.3(c)) we arrive at the following interpretations of the cohomology groups of $\mathcal{E}(\omega)$:

$$H^0(\omega) = 0 \iff \omega \quad \text{irreducible} \tag{1.12}$$

$$H^1(\omega) = \ker\left(\mathrm{d}_+^\omega | \ker(\delta^\omega)\right) = \begin{array}{r}\text{kernel of the derivative of} \\ \operatorname{Pr}_+ \circ F \mid \mathcal{O}_{\omega,\epsilon} \ \text{ at } \omega\end{array} \tag{1.13}$$

$$H^2(\omega) = 0 \iff \mathrm{d}_+^\omega | \ker(\delta^\omega) \ \text{ is surjective.} \tag{1.14}$$

Recalling that when ω is irreducible the projection into the moduli space is injective we can summarize all of this in the following Theorem.

Theorem 1.4. *At* $\omega \in \operatorname{Asd}_3(P_k, g)$ *the map*

$$\operatorname{Pr}_+ \circ F \mid \mathcal{O}_{\omega,\epsilon}\colon \mathcal{O}_{\omega,\epsilon} \longrightarrow \Omega_{+,2}^2(B, \operatorname{ad} P_k)$$

has derivative

$$\mathrm{d}_+^\omega | \ker(\delta^\omega)\colon \ \ker(\delta^\omega) \longrightarrow \Omega_{+,2}^2(B, \operatorname{ad} P_k)$$

that is Fredholm. The derivative is surjective if and only if $H^2(\omega) = 0$ *and, in this case,*

$$\left(\mathrm{Pr}_+ \circ F \mid \mathcal{O}_{\omega,\epsilon}\right)^{-1}(0) = \mathrm{Asd}_3(P_k, g) \cap \mathcal{O}_{\omega,\epsilon}$$

is a smooth manifold of dimension

$$\dim\left(\ker\left(\mathrm{d}_+^\omega \,|\, \ker(\delta^\omega)\right)\right) = \dim(H^1(\omega))$$

near ω. *If, in addition,* $H^0(\omega) = 0$, *then the projection into the moduli space* $\mathcal{M}(P_k, g)$ *gives a chart of dimension* $\dim(H^1(\omega))$ *near* $[\omega]$.

Notice that if $H^0(\omega)$ and $H^2(\omega)$ are both trivial, then $\mathrm{Ind}(\mathcal{E}(\omega)) = -\dim(H^1(\omega))$ so, if we could calculate the index of the elliptic complex $\mathcal{E}(\omega)$, we would have (minus) the dimension of the moduli space near $[\omega]$. The Atiyah–Singer Index Theorem gives the index of $\mathcal{E}(\omega)$ in terms of topological data on B and $SU(2) \hookrightarrow P_k \to B$. In our present context the result is

$$\mathrm{Ind}\left(\mathcal{E}(\omega)\right) = -8k + 3\left(1 + b_2^+(B)\right) \tag{1.15}$$

(see pages 267–271 of [5]). Note, in particular, that the result is independent of ω so we obtain the following consequence of Theorem 1.4.

Corollary 1.5. *If* $H^0(\omega) = 0$ *and* $H^2(\omega) = 0$ *for every* $\omega \in \mathrm{Asd}_3(P_k, g)$, *then* $\mathcal{M}(P_k, g) = \hat{\mathcal{M}}(P_k, g)$ *is a smooth manifold of dimension* $8k - 3(1 + b_2^+(B))$.

For a given choice of the Riemannian metric g it may or may not be the case that $H^0(\omega) = 0$ and $H^2(\omega) = 0$ for every $\omega \in \mathrm{Asd}_3(P_k, g)$. We will see shortly, however, that for "almost all" choices of g, $H^2(\omega)$ will be trivial for all $\omega \in \mathrm{Asd}_3(P_k, g)$ and, with one additional restriction on the topology of B, the same is true $H^0(\omega)$. First, however, we note that if $\omega \in \mathrm{Asd}_3(P_k, g)$ and $H^2(\omega) = 0$, then the slice $\mathrm{Asd}_3(P_k, g) \cap \mathcal{O}_{\omega,\epsilon}$ still has a local manifold structure near ω, but if $H^0(\omega) \neq 0$ one can only obtain a one-to-one projection into the moduli space by first factoring out the action of the stabilizer of ω. The consequence is that, near $[\omega]$, $\mathcal{M}(P_k, g)$ is not smooth but has a neighborhood homeomorphic to this quotient, which turns out to be a cone over $\mathbb{CP}^{4k-2}$ with $[\omega]$ at the vertex. Reducible connections, when they exist, give rise to cone-like "singularities" in the moduli space.

We conclude this section with a brief discussion of various "generic metrics" theorems which assert that, under certain circumstances, "almost all" choices for the Riemannian metric g give rise to "nice" moduli spaces of ASD connections. Begin by considering the space $\mathcal{R}$ of all Riemannian metrics on B. This is a space of sections of a fiber bundle over B and can be given the structure of a (pathwise connected) Hilbert manifold. With this structure one can show that

1. There is a dense G_δ-set in $\mathcal{R}$ such that, for every g in this set, any g-ASD connection ω on P_k, $k > 0$, satisfies $H^2(\omega) = 0$.
2. If $b_2^+(B) > 0$, then there is a dense G_δ-set in $\mathcal{R}$ such that, for every g in this set, any g-ASD connection ω on P_k, $k > 0$, satisfies $H^2(\omega) = 0$ and $H^0(\omega) = 0$.

In short, for "generic" g, $\mathcal{M}(P_k, g) = \hat{\mathcal{M}}(P_k, g)$ is (either empty or) a smooth manifold of dimension $8k - 3(1 + b_2^+(B))$.

Very roughly, here is how one might go about showing that $\hat{\mathcal{M}}(P_k, g)$ is smooth for a generic choice of g. Consider the so-called parametrized moduli space

$$\hat{\mathcal{M}}(P_k, \mathcal{R}) = \left\{([\omega], g) \in \hat{\mathcal{B}}_3(P_k) \times \mathcal{R};\ \omega \text{ is } g\text{-ASD}\right\}.$$

This is an infinite dimensional smooth submanifold of $\hat{\mathcal{B}}(P_k) \times \mathcal{R}$. One shows that the projection map

$$\hat{\mathcal{M}}(P_k, \mathcal{R}) \longrightarrow \mathcal{R}$$

is smooth with Fredholm derivative at each point. The Sard–Smale Theorem (infinite dimensional version of Sard's Theorem) implies that the set of regular values of the projection is a dense G_δ-set in $\mathcal{R}$. But then, for any g in this set, the inverse image

$$\left(\hat{\mathcal{B}}_3(P_k) \times \{g\}\right) \cap \hat{\mathcal{M}}(P_k, \mathcal{R}) = \hat{\mathcal{M}}(P_k, g)$$

is a smooth submanifold of $\hat{\mathcal{M}}(P_k, \mathcal{R})$.

Remark: The restriction on $b_2^+(B)$ arises because the subset of $\mathcal{R}$ consisting of those g for which reducible g-ASD connections on P_k exist is a countable union of smooth submanifolds of codimension $b_2^+(B)$. If $b_2^+(B) = 0$, then reducibles are generically unavoidable. As we shall see in Section 3 this is a good, not a bad thing as it leads to Donaldson's theorem on definite intersection forms.

Crudely put, the idea behind defining the Donaldson polynomial invariants (to which we turn in Section 4) is to regard the moduli spaces $\hat{\mathcal{M}}(P_k, g)$ as cycles over which to integrate certain carefully selected differential forms. In order to carry out such a program these moduli spaces must be orientable and, if the result is to be a differential topological invariant, the integrals must be independent of the choice of (generic) g. We now record two results that guarantee this.

Theorem 1.6. *Suppose g is a Riemannian metric on B for which $H^2(\omega) = 0$ for every g-ASD connection ω on P_k, $k > 0$. Then the moduli space $\hat{\mathcal{M}}(P_k, g)$*

is orientable. An orientation for $\hat{\mathcal{M}}(P_k, g)$ *can be uniquely specified by choosing an orientation for* B *and an orientation for the vector space* $H^2_+(B, \mathbb{R})$ *of self-dual 2-forms on* B.

The proof amounts to constructing an explicit model for the determinant line bundle (top exterior power) of $\hat{\mathcal{M}}(P_k, g)$ from a family of differential operators on B and exhibiting a non-zero section (see Sections 5.4.1 and 7.1.6 of [5]).

To state the final result of this section we consider two metrics g_0 and g_1 in the dense G_δ-set in $\mathcal{R}$ on which $H^2(\omega) = 0$ for all g-ASD connections. Since $\mathcal{R}$ is pathwise connected we can join them with a path $\{g_t : 0 \le t \le 1\}$ in $\mathcal{R}$. Define the parametrized moduli space

$$\hat{\mathcal{M}}\left(P_k, \{g_t\}\right) = \left\{([\omega], t) \in \hat{\mathcal{B}}_3(P_k) \times [0,1];\ [\omega] \in \hat{\mathcal{M}}(P_k, g_t)\right\}.$$

Theorem 1.7. *If* g_0 *and* g_1 *are in the dense* G_δ*-set in* $\mathcal{R}$ *on which* $H^2(\omega) = 0$ *for all g-ASD connections* ω*, then, for a generic path* $\{g_t : 0 \le t \le 1\}$ *joining them,* $\hat{\mathcal{M}}(P_k, \{g_t\})$ *is a smooth, orientable submanifold of* $\hat{\mathcal{B}}_3(P_k) \times [0,1]$ *with boundary. A choice of orientation* μ *for* $H^2_+(B, \mathbb{R})$ *determines an orientation for* $\hat{\mathcal{M}}(P_k, \{g_t\})$*. Moreover, the oriented boundary of* $\hat{\mathcal{M}}(P_k, \{g_t\})$ *is the disjoint union of* $\hat{\mathcal{M}}(P_k, g_1)$ *with the orientation induced by* μ *and* $\hat{\mathcal{M}}(P_k, g_0)$ *with the orientation opposite to that induced by* μ.

In short, a generic variation of g varies $\hat{\mathcal{M}}(P_k, g)$ within a single homology class. Even if $\hat{\mathcal{M}}(P_k, g_0)$ and $\hat{\mathcal{M}}(P_k, g_1)$ are smooth manifolds one cannot, in general, arrange that the intermediate moduli spaces $\hat{\mathcal{M}}(P_k, g_t)$ are all smooth. There may be finitely many values of t for which one encounters reducible connections. Even this can be avoided, however, if one is willing to assume of B that $b_2^+(B) > 1$. In this case a generic path $\{g_t\}$ from g_0 to g_1 has the property that each $\hat{\mathcal{M}}(P_k, g_t)$ is a smooth, orientable manifold. These results are discussed in more detail in [5].

2. The Uhlenbeck Compactification

We mentioned in the last section that, to the 0-th approximation, Donaldson invariants are obtained by integrating certain differential forms over the moduli spaces $\hat{\mathcal{M}}(P_k, g)$. This would be straightforward enough if these moduli spaces were compact, but they generally are not (see the Examples in Section 1). However, we will now outline the construction of a natural compactification of these moduli spaces and it is this that one actually integrates over to obtain the invariants. Throughout this section we will assume that the Riemannian metric g is "generic" in the following sense:

1. $H^2(\omega) = 0$ for every g-ASD connection on P_k, $k > 0$. Thus,
   a. $\hat{\mathcal{M}}(P_k, g)$ is either empty or a smooth manifold of dimension $8k-3(1+b_2^+(B))$.
   b. If ω is reducible, then $[\omega]$ has a neighborhood in $\mathcal{M}(P_k, g)$ homeomorphic to a cone over $\mathbb{CP}^{4k-2}$ with $[\omega]$ at the vertex.
2. If $b_2^+(B) > 0$, then $\mathcal{M}(P_k, g) = \hat{\mathcal{M}}(P_k, g)$, i. e., there are no reducible g-ASD connections on any P_k, $k > 0$.

To compactify $\hat{\mathcal{M}}(P_k, g)$ we must understand how a sequence of points in $\hat{\mathcal{M}}(P_k, g)$ can fail to converge in $\hat{\mathcal{M}}(P_k, g)$. In essence, a sequence $\{\omega_n\}$ of irreducible ASD connections can give rise to a sequence $\{[\omega_n]\}$ of points in $\hat{\mathcal{M}}(P_k, g)$ without a convergent subsequence in only two ways:

1. $\{[\omega_n]\}$ has a subsequence that converges to a reducible ASD connection. In this case the compactification will contain a point of $\mathcal{M}(P_k, g)$ with its local cone structure.

2. Deep analytical work of Karen Uhlenbeck shows that the only other possibility is that the curvatures $\{F_{\omega_n}\}$ have pointwise norms $\|F_{\omega_n}(b)\|$ that, like those of the BPST instantons, become increasingly concentrated at some point (or finite set of points) in B. Intuitively, these "converge to connections with curvature concentrated at a point in B" which, because of their singular nature, do not appear in the moduli space. The compactification adds these and this essentially amounts to adding a copy of B (the points of concentration).

To describe the compactification process in somewhat more detail we will record a result which summarizes the analytical work referred to above.

Remark: The most crucial element of this result is the so-called "Removable Singularities Theorem" of Karen Uhlenbeck. To gain some appreciation of the content of this theorem we will pause momentarily to describe what it has to say in the more familiar context of ASD connections on the trivial bundle over $\mathbb{R}^4$ (with its usual metric). As we have seen in our discussion of BPST instantons, these are just ASD gauge potentials $\mathcal{A}$ on $\mathbb{R}^4$ (i. e., $su(2)$-valued 1-forms on $\mathbb{R}^4$). The BPST potentials $\mathcal{A}_{\lambda,n}$ are examples and all of these have "finite action" in the sense that

$$\int_{\mathbb{R}^4} \|\mathcal{F}_{\lambda,n}(q)\|^2 \,\mathrm{vol} < \infty \,.$$

In fact, all of these integrals are $8\pi^2$ because, except for a factor of $\dfrac{1}{8\pi^2}$, they all compute the Chern number of the quaternionic Hopf bundle (which is 1). Furthermore, the potentials $\mathcal{A}_{\lambda,n}$ all "extend to $\mathbb{S}^4$" in the sense that there is an $SU(2)$-bundle over $\mathbb{S}^4$ (namely, the $k = 1$ bundle) and a connection $\omega_{\lambda,n}$ on this bundle whose restriction to $\mathbb{S}^4 \setminus \{\text{north pole}\}$ is, when pulled back by the inverse of stereographic projection, equal to $\mathcal{A}_{\lambda,n}$. Uhlenbeck's theorem asserts that "finite action" and "extendibility to $\mathbb{S}^4$" are the same thing for ASD connections on $\mathbb{R}^4$ (think of $\mathbb{S}^4$ as the 1-point compactification of $\mathbb{R}^4$, $\mathcal{A}$ as a connection on $\mathbb{R}^4$ that is "singular at the point at infinity" and "finite action" as a sufficient condition to "remove" the singularity). More precisely, the special case of Uhlenbeck's Theorem of interest to us at the moment asserts the following: Let $\mathcal{A}$ be an $su(2)$-valued 1-form on $\mathbb{R}^4$ that is ASD and for which

$$\int_{\mathbb{R}^4} \|\mathcal{F}(q)\|^2 \,\mathrm{vol} < \infty \,. \tag{2.1}$$

Then there exists a unique $SU(2)$-bundle $SU(2) \hookrightarrow P \xrightarrow{\pi} \mathbb{S}^4$ over $\mathbb{S}^4$ and a unique connection 1-form ω on P such that $s^*\omega = \mathcal{A}$ for some section s defined on $\mathbb{S}^4\backslash\{\text{north pole}\}$. Furthermore, the integral in (2.1), except for a factor of $\frac{1}{8\pi^2}$, is the Chern number of P (so that the asymptotic behavior of the field strength $\mathcal{F}$ is directly encoded in the topology of the bundle to which $\mathcal{A}$ extends).

Here, then, is the analytical result upon which the construction of the compactification is based.

Theorem 2.1. *Let $\{\omega_n\}$ be a sequence in $\mathrm{Asd}_3(P_k, g)$, $k > 0$, without a convergent subsequence. After passing to a subsequence, also denoted $\{\omega_n\}$, one can find*

1. *A principal $SU(2)$-bundle $SU(2) \hookrightarrow P_{k'} \longrightarrow B$ with $0 \leq k' \leq k$,*
2. *Points $\{b_1, \ldots, b_t\}$ in B, where $t \leq k$,*
3. *Positive integers $\{m_1, \ldots, m_t\}$,*
4. *A g-ASD connection ω' on $P_{k'}$,*
5. *Bundle isomorphisms $f_n \colon P_{k'} \mid B - \{b_1, \ldots, b_t\} \longrightarrow P_k \mid B - \{b_1, \ldots, b_t\}$ for each n which are L^2_4 on each compact set $K \subseteq B - \{b_1, \ldots, b_t\}$,*

such that

a. *The connections $f_n^*(\omega_n|K)$ converge in L^2_3 to $\omega' \mid K$ for every compact set $K \subseteq B - \{b_1, \ldots, b_t\}$, and*

b. *The functions $\|F_{\omega_n}(b)\|^2$ on B converge weakly to the measure*

$$\|F_{\omega'}(b)\|^2 \,\mathrm{vol}_g + \sum_{i=1}^{t} 8\pi^2 m_i \delta_{b_i} \,,$$

where δ_{b_i} is the point measure of mass 1 concentrated at b_i.
More precisely, for any continuous real-valued function φ on B

$$\int_B \varphi \,\|F_{\omega_n}\|^2 \,\mathrm{vol}_g \longrightarrow \int_B \varphi \,\|F_{\omega'}\|^2 \,\mathrm{vol}_g + \sum_{i=1}^{t} 8\pi^2 m_i \varphi(b_i) \,.$$

Example: Consider the family $\omega_{\lambda,0}$, $\lambda > 0$, of BPST instantons centered at $0 \in \mathbb{H} = \mathbb{R}^4$. For $n = 1, 2, 3, \dots$ let $\omega_n = \omega_{1/n,0}$. Then $\{\omega_n\}$ is a sequence of ASD connections on the $k = 1$ bundle over $\mathbb{S}^4$. Pulling back to $\mathbb{H}$ by the inverse of stereographic projection from the north pole gives gauge potentials

$$\mathcal{A}_n = \mathrm{Im}\left(\frac{\bar{q}}{1/n^2 + |q|^2}\right) \mathrm{d}q$$

and field strengths

$$\mathcal{F}_n = \frac{1/n^2}{(1/n^2 + |q|^2)^2}\, \mathrm{d}\bar{q} \wedge \mathrm{d}q\,.$$

The functions

$$\|\mathcal{F}_n(q)\|^2 = \frac{48/n^2}{(1/n^2 + |q|^2)^2}$$

have a maximum of $48n^2$ which approaches infinity as $n \to \infty$, but the integrals $\int_{\mathbb{R}^4} \|\mathcal{F}_n(q)\|^2 \,\mathrm{vol}$ remain constant at $8\pi^2$. Thus, $\{\omega_n\}$ "converges" to the flat connection away from the south pole S of $\mathbb{S}^4$ and the curvature densities $\|\mathcal{F}_n\|^2$ "converge" to the point measure $8\pi^2\delta_S$ so, in Uhlenbeck's Theorem, $k' = 0$, $P_{k'} = P_0$ is the trivial bundle over $\mathbb{S}^4$, $t = 1$, $b_1 = S$, $m_1 = 1$ and ω' is the flat connection.

Remark: Analogous "k-instantons" on bundles over $\mathbb{S}^4$ with larger Chern class behave like superpositions of BPST instantons and so may concentrate at several points, or several of them may concentrate at the same point, thus giving rise to $b_1, \dots, b_t$ and $m_1, \dots, m_t$ ($m_i\delta_{b_i}$ in Uhlenbeck's Theorem means $\overbrace{\delta_{b_i} + \cdots + \delta_{b_i}}^{m_i}$).

With this result as motivation we introduce the following definitions. An **ideal** (or **generalized**, or **virtual**) **ASD connection** of **Chern class** k is a pair ($[\omega]$, $(b_1, \dots, b_t)$), where $0 \le t \le k$, $[\omega]$ is a point in $\mathcal{M}(P_{k-t}, g)$ and $(b_1, \dots, b_t)$), is a point in the t-th symmetric product $\mathbb{S}^t(B)$ of B.

Remark: $\mathbb{S}^t(B)$ is the space of unordered, and possibly non-distinct, points of B, i. e., it is the quotient of $\overbrace{B \times \cdots \times B}^{t}$ by the action of the symmetric group on t letters that permutes the coordinates.

The connection ω is called the **background connection** for ($[\omega]$, $(b_1, \dots, b_t)$). The **curvature density** of ($[\omega]$, $(b_1, \dots, b_t)$) is the measure

$$\|F_\omega\|^2 \,\mathrm{vol}_g + \sum_{i=1}^{t} 8\pi^2\delta_{b_i}\,.$$

Notice that the integral of this measure over B is $8\pi^2(C_2(P_{k-t}) + t) = 8\pi^2 k$.

If $\{\omega_n\}$ is a sequence of g-ASD connections on $SU(2) \hookrightarrow P_k \to B$, then $\{[\omega_n]\}$ is said to **converge** (**weakly**) to the ideal ASD connection $([\omega],\ (b_1, \ldots, b_t))$ if

1. $\displaystyle\int_B \varphi \left\|F_{\omega_n}\right\|^2 \mathrm{vol}_g \longrightarrow \int_B \varphi \left\|F_\omega\right\|^2 \mathrm{vol}_g + \sum_{i=1}^{t} 8\pi^2 \varphi(b_i)$ for every continuous real-valued function φ on B.
2. There exist bundle isomorphisms $f_n\colon P_{k-t} \mid B - \{b_1, \ldots, b_t\} \to P_k \mid B - \{b_1, \ldots, b_t\}$ which are L^2_4 on every compact set $K \subseteq B - \{b_1, \ldots, b_t\}$ and such that, on every such compact set, $f_n^*(\omega_n|K)$ converges to $\omega|K$ in L^2_3.

Similarly, one defines what it means for a sequence $\{([\omega_n],\ (b_1^n, \ldots, b_{t(n)}^n))\}$ of ideal ASD connections to converge (weakly) to an ideal ASD connection $([\omega],\ (b_1, \ldots, b_t))$, i. e., the curvature densities

$$\left\|F_{\omega_n}\right\|^2 \mathrm{vol}_g + \sum_{i=1}^{t(n)} 8\pi^2 \delta_{b_i^n}$$

converge weakly to $\|F_\omega\|^2 \mathrm{vol}_g + \sum_{i=1}^{t} 8\pi^2 \delta_{b_i}$ and, on compact sets in $B - \{b_1, \ldots, b_t\}$, the restrictions of the ω_n converge, up to gauge transformations, to the restriction of ω in L^2_3. This notion of sequential convergence defines a topology on the disjoint union

$$\begin{aligned} \mathcal{M}(P_k, g) \cup \Big(\mathcal{M}(P_{k-1}, g) \times \mathbb{S}^1(B)\Big) \cup \Big(\mathcal{M}(P_{k-2}, g) \times \mathbb{S}^2(B)\Big) \cup \cdots \\ \cup \Big(\mathcal{M}(P_1, g) \times \mathbb{S}^{k-1}(B)\Big) \cup \Big(\mathcal{M}(P_0, g) \times \mathbb{S}^k(B)\Big). \end{aligned}$$

The topology is, in fact, metrizable, each "stratum"

$$\mathcal{M}(P_{k-t}, g) \times \mathbb{S}^t(B)$$

inherits its usual topology and $\mathcal{M}(P_k, g)$ is an open subspace. Recall that, with our hypotheses, $\mathcal{M}(P_0, g)$ consists of a single point so that $\mathcal{M}(P_0, g) \times \mathbb{S}^k(B)$, which consists of those ideal ASD connections with trivial background connection, is really just a copy of $\mathbb{S}^k(B)$.

Uhlenbeck's Theorem 2.1 now implies that this space is compact so that the closure $\bar{\mathcal{M}}(P_k, g)$ of $\mathcal{M}(P_k, g)$ in it is a compactification of $\mathcal{M}(P_k, g)$ called its **Uhlenbeck compactification**. Since $\hat{\mathcal{M}}(P_k, g)$ is dense in $\mathcal{M}(P_k, g)$, $\bar{\mathcal{M}}(P_k, g)$ is also a compactification of $\hat{\mathcal{M}}(P_k, g)$. Even when there are no reducible ASD connections, however, so that $\mathcal{M}(P_k, g) = \hat{\mathcal{M}}(P_k, g)$ is a smooth manifold, the compactification $\bar{\mathcal{M}}(P_k, g)$ is not generally a manifold. It is, rather, what is known as a stratified space. In the next section we will look more closely at a particular case ($k = 1$ and $b_2^+(B) = 0$) and see how it leads to a famous theorem of Donaldson on definite intersection forms.

3. Donaldson's Theorem on Definite Intersection Forms

Let us consider a compact, simply connected, oriented, smooth 4-manifold B with $b_2^+(B) = 0$ (i. e., indefinite intersection form) and the $k = 1$ bundle

$$SU(2) \hookrightarrow P_1 \longrightarrow B$$

over B. Choose a generic Riemannian metric g on B. We have already mentioned (Example 4, of Section 1) that Taubes has proved that $\mathcal{M}(P_1, g)$ is non-empty. Now, since $k = 1$, the Uhlenbeck compactification has only two strata

$$\mathcal{M}(P_1, g) \cup (\mathcal{M}(P_0, g) \times \mathbb{S}^1(B)) \cong \mathcal{M}(P_1, g) \times B\,.$$

Since $b_2^+(B) = 0$, P_1 will generally have reducible as well as irreducible g-ASD connections. Thus, $\hat{\mathcal{M}}(P_k, g)$, which we know to be a smooth manifold of dimension

$$8k - 3\,(1 + b_2^+(B)) = 8 \cdot 1 - 3(1 + 0) = 5\,,$$

has two types of "ends" in the Uhlenbeck compactification $\bar{\mathcal{M}}(P_k, g)$:

1. Points in $\mathcal{M}(P_1, g) - \hat{\mathcal{M}}(P_1, g)$ corresponding to sequences of irreducible connections converging to reducible connections. These all have neighborhoods in $\mathcal{M}(P_1, g) \subseteq \bar{\mathcal{M}}(P_1, g)$ that are homeomorphic to a cone over $\mathbb{CP}^{4k-2} = \mathbb{CP}^2$.

   Remark: One can show that the number of equivalence classes of reducible g-ASD connections is half the number of $\alpha \in H_2(B, \mathbb{Z})$ for which $q_B(\alpha, \alpha) = -1$ and so, in particular, is finite (see [6] or [12]).
2. Points in $\mathcal{M}(P_0, g) \times \mathbb{S}^1(B) \cong B$ which correspond to sequences in $\hat{\mathcal{M}}(P_1, g)$ that do not converge in $\mathcal{M}(P_1, g)$ because the corresponding connections "concentrate" at some point of B. Since Taubes' "grafting" procedure can produce such a sequence at each point of B the Uhlenbeck compactification is, in this case, all of $\mathcal{M}(P_1, g) \times B$.

$$\bar{\mathcal{M}}(P_1, g) \cong \mathcal{M}(P_1, g) \cup B\,.$$

Taubes' analysis actually yields much more detailed information about $\bar{\mathcal{M}}(P_1, g)$ near the ideal points represented by $B \subseteq \bar{\mathcal{M}}(P_1, g)$. Specifically, he shows that there is a neighborhood of B in $\bar{\mathcal{M}}(P_1, g)$ homeomorphic to a cylinder $B \times [0, \lambda_0)$. Intuitively, each $\{b\} \times [0, \lambda_0)$ corresponds to a family of grafted instantons parametrized by λ in $[0, \lambda_0)$ that concentrates at b as $\lambda \to 0$. The picture that emerges of $\bar{\mathcal{M}}(P_1, g)$ is therefore as follows: There are finitely many points $p_1, \ldots, p_m$ corresponding to equivalence classes of reducible g-ASD connections on P_1, each with a neighborhood in $\bar{\mathcal{M}}(P_1, g)$ homeomorphic to a cone over $\mathbb{CP}^2$. $\bar{\mathcal{M}}(P_1, g) - \{p_1, \ldots, p_m\}$ is a smooth 5-manifold with boundary and the boundary is just the copy $B \times \{0\} \subseteq B \times [0, \lambda_0)$ of B.

We shall now briefly indicate how Donaldson [3] used this picture of $\bar{\mathcal{M}}(P_1, g)$ to prove a remarkable result on smooth 4-manifolds with definite intersection forms. From $\bar{\mathcal{M}}(P_1, g)$ we construct a new space $\mathcal{M}_0(P_1, g)$ by deleting the (open) top half of the conical neighborhood of each point $p_1, \ldots, p_m$. Then $\mathcal{M}_0(P_1, g)$ is a compact 5-manifold with boundary and the boundary consists of a copy of B and the disjoint union of m copies of $\mathbb{CP}^2$:

$$\partial\mathcal{M}_0(P_1, g) = B \sqcup \bigsqcup \mathbb{CP}^2 \,.$$

Here $\bigsqcup \mathbb{CP}^2$ ranges over all pairs $\{\pm\alpha\}$ in $H_2(B, \mathbb{Z})$ for which $q_B(\alpha, \alpha) = -1$. One can show that $\mathcal{M}_0(P_1, g)$ is naturally oriented and that the induced orientation on B is its given orientation (see [6]). The copies of $\mathbb{CP}^2$ may inherit their usual orientation (as a complex manifold) or the opposite so we write

$$\partial\mathcal{M}_0(P_1, g) = B \sqcup p\mathbb{CP}^2 \sqcup q\overline{\mathbb{CP}}^2 \,, \tag{3.1}$$

where $p + q = m$. Thus, $p + q$ is the number of pairs $\{\pm\alpha\}$ in $H_2(B, \mathbb{Z})$ for which $q_B(\alpha, \alpha) = -1$.

The crucial observation is that, because of (3.1), $\mathcal{M}_0(P_1, g)$ is a cobordism from B to $p\mathbb{CP}^2 \sqcup q\overline{\mathbb{CP}}^2$ and that the signature $\sigma(B) = b_2^+(B) - b_2^-(B)$ of the intersection form is a cobordism invariant (see [20]). Since the signature of $p\mathbb{CP}^2 \sqcup q\overline{\mathbb{CP}}^2$ is $p - q$ and $b_2^+(B)$ is assumed to be 0 we have $-b_2^-(B) = p - q$ and so

$$b_2(B) = b_2^+(B) + b_2^-(B) = -p + q \,. \tag{3.2}$$

In particular,

$$b_2(B) \leq p + q \,. \tag{3.3}$$

We now show that $b_2(B) \geq p + q$ as well so that, in fact, $b_2(B) = p + q$. The argument is as follows: If $p + q = 0$ the result is obvious so suppose $p + q > 0$. Then we can select an $\alpha_1 \in H_2(B, \mathbb{Z})$ with $q_B(\alpha_1, \alpha_1) = -1$. Then there is an orthogonal (with respect to q_B) decomposition

$$H_2(B, \mathbb{Z}) = \mathbb{Z}\alpha_1 \oplus G_1$$

obtained by writing any $\beta \in H_2(B, \mathbb{Z})$ as

$$\beta = (q_B(\beta, \alpha_1))\,\alpha_1 + (\beta - (q_B(\beta, \alpha_1))\,\alpha_1) \,.$$

Now, for any $\alpha_2 \in H_2(B, \mathbb{Z})$ with $q_B(\alpha_2, \alpha_2) = -1$ and $\alpha_2 \neq \pm\alpha_1$, the Schwartz Inequality gives

$$(q_B(\alpha_1, \alpha_2))^2 < q_B(\alpha_1, \alpha_1) q_B(\alpha_2, \alpha_2) = 1 \,.$$

But $q_B(\alpha_1, \alpha_2)$ is an integer so $q_B(\alpha_1, \alpha_2) = 0$, i. e.,

$$\alpha_2 \in G_1 \,.$$

Now repeat the argument inside G_1 and continue until you run out of α for which $q_B(\alpha, \alpha) = -1$. The result is an orthogonal decomposition

$$H_2(B, \mathbb{Z}) = \mathbb{Z}\alpha_1 \oplus \cdots \oplus \mathbb{Z}\alpha_m \oplus G \,,$$

where m is the number of pairs $\{\pm\alpha\}$ in $H_2(B, \mathbb{Z})$ for which $q_B(\alpha, \alpha) = -1$ and G is either empty or the orthogonal complement of $\mathbb{Z}\alpha_1 \oplus \cdots \oplus \mathbb{Z}\alpha_{p+q}$. In particular, $p + q \leq b_2(B)$ and this, together with (3.3) gives $b_2(B) = p + q$. Consequently, G must be empty and we have

$$H_2(B, \mathbb{Z}) = \mathbb{Z}\alpha_1 \oplus \cdots \oplus \mathbb{Z}\alpha_{b_2(B)} \,,$$

where

$$q_B(\alpha_i, \alpha_i) = -1, \qquad i = 1, \ldots, b_2(B) \,.$$

This concludes our sketch of the proof of Donaldson's Theorem.

Theorem 3.1. *Let B be a compact, simply connected, oriented, smooth 4-manifold with negative definite intersection form q_B (i. e., $b_2^+(B) = 0$). Then q_B is diagonalizable over $\mathbb{Z}$, i. e., there is a basis $\alpha_1, \ldots, \alpha_{b_2(B)}$ for $H_2(B, \mathbb{Z})$ over $\mathbb{Z}$ such that $q_B(\alpha_i, \alpha_i) = -1$ for $i = 1, \ldots, b_2(B)$ and so the matrix of q_B is $(q_B(\alpha_i, \alpha_i)) = -\,\mathrm{Id}$.*

Remark: Some sources (e. g., [6] and [12]) consider self-dual connections on the $k = -1$ bundle to prove that if q_B is positive definite, then, relative to some basis, the matrix of q_B is Id. By reversing orientations, the two results are equivalent.

One cannot appreciate just how remarkable Donaldson's Theorem is without contrasting it with Freedman's classification of compact, simply connected, topological 4-manifolds. The intersection form can be defined also in the topological category and Freedman proved that every integer-valued, unimodular, symmetric, bilinear form can be realized as the intersection form of at least one (and at most two) compact, simply connected topological 4-manifold(s). Since there are vast numbers of such forms that are definite, but not diagonalizable over $\mathbb{Z}$, there must also be vast numbers of topological 4-manifolds that admit no smooth structure. A more subtle combination of the work of Donaldson and Freedman yields the existence of "fake $\mathbb{R}^4$s", i. e., smooth 4-manifolds that are homeomorphic, but not diffeomorphic to $\mathbb{R}^4$ (see Chapter 6 of [7]).

4. The Donaldson Polynomial Invariants

We now turn our attention to 4-manifolds B with $b_2^+(B) > 0$ so that we can select a generic metric g for B such that, for $k > 0$, $\mathcal{M}(P_k, g) = \hat{\mathcal{M}}(P_k, g)$ is either empty or a smooth manifold of dimension $8k - 3(1 + b_2^+(B))$. The case $b_2^+(B) = 1$ presents technical difficulties because, as we mentioned in Section 1, a generic variation of g can, in this case, introduce reducible connections and therefore singularities in the moduli spaces. We intend to leap over these difficulties by assuming $b_2^+(B) > 1$. Moreover, for the constructions we have in mind it will be necessary that the dimension of the moduli space be even (because we want to integrate certain distinguished even rank forms over it). Since $8k - 3(1 + b_2^+(B))$ is even only when $b_2^+(B)$ is odd **we shall henceforth assume that $b_2^+(B) > 1$ and odd**.

In order to get some sense of what such manifolds "look like" we will briefly record what Freedman's classification theorem says about their topological (i. e., homeomorphism) type. We first note that $b_2^+(B) > 1$ and odd implies that q_B must be indefinite and that indefinite, unimodular, symmetric, integer-valued bilinear forms are classified up to equivalence as follows (two such forms are equivalent if there are bases relative to which they have the same matrix and we use $\sim$ to denote this relation):

$$\text{Type I} \quad (q_B \text{ even}) : q_B \sim n(1) \oplus m(-1)\,, \qquad n > 0\,,\ m > 0$$

$$\text{Type II} \quad (q_B \text{ odd}) : q_B \sim n\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \oplus mE_8\,, \qquad n > 0\,,\ m \geq 0$$

where E_8 is the Cartan matrix of the exceptional Lie algebra with the same name, i. e.,

$$E_8 = \begin{pmatrix} 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 2 \end{pmatrix}.$$

Freedman's theorem gives the corresponding topological types for smooth manifolds with such intersection forms ($\cong_h$ means "is homeomorphic to"):

$$\text{Type I} \quad (B \text{ not spin}) : B \cong_h n\mathbb{CP}^2 \# m\overline{\mathbb{CP}}^2\,, \qquad n > 0\,,\ m > 0$$

$$\text{Type II} \quad (B \text{ spin}) : B \cong_h n(\mathbb{S}^2 \times \mathbb{S}^2) \# m|E_8|\,, \qquad n > 0\,,\ m \geq 0$$

where $|E_8|$ is the unique compact, simply connected, topological 4-manifold with intersection form E_8 (guaranteed to exist by Freedman's result). We emphasize that this describes the topological, not the smooth type of B. Indeed, not all of these topological manifolds admit smooth structures.

Remark: Any topological manifold of the form $n\mathbb{CP}^2 \# m\overline{\mathbb{CP}}^2$ admits a smooth structure and the same is true of $n(\mathbb{S}^2 \times \mathbb{S}^2) \# m|E_8|$ provided $m \leq 6n$. The current conjecture is that $m \leq 6n$ is a necessary as well as sufficient condition for the smoothability of $n(\mathbb{S}^2 \times \mathbb{S}^2) \# m|E_8|$. This inequality is equivalent to

$$\frac{b_2(B)}{|\sigma(B)|} \geq \frac{11}{8}$$

so this is called the $\frac{11}{8}$**-conjecture**.

Example: The Kummer surface $K3$ may be defined as the submanifold of $\mathbb{CP}^3$ with homogeneous coordinates z^1, z^2, z^3, z^4 satisfying

$$\left(z^1\right)^4 + \left(z^2\right)^4 + \left(z^3\right)^4 + \left(z^4\right)^4 = 0\,.$$

It is a compact, simply connected 4-manifold with a natural orientation as a complex manifold. As it happens, $b_2(K3) = 22$, $b_2^+(K3) = 3$ and $q_{K3} \sim 3\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \oplus 2E_8$ so

$$K3 \cong_h 3\left(\mathbb{S}^2 \times \mathbb{S}^2\right) \# 2|E_8|\,.$$

We turn now to the construction of the Donaldson polynomial invariants for B. For this we define, for each $k > 0$, an integer d_k by

$$d_k = 4k - \frac{3}{2}\left(1 + b_2^+(B)\right)\,.$$

Note that this is half the formal dimension of the moduli space $\mathcal{M}(P_k, g) = \hat{\mathcal{M}}(P_k, g)$ and is an integer because $b_2^+(B)$ is odd. The corresponding Donaldson invariant γ_{d_k} is to be a symmetric multilinear $\mathbb{Z}$-valued function on $H_2(B, \mathbb{Z})$:

$$\gamma_{d_k}(B)\colon \overbrace{H_2(B,\mathbb{Z}) \times \cdots \times H_2(B,\mathbb{Z})}^{d_k} \longrightarrow \mathbb{Z}$$

(or, equivalently, a homogeneous polynomial of degree d_k on $H_2(B, \mathbb{Z})$). Very roughly, the idea behind the definition is as follows: For each $k > 0$ we will define a map $\mu\colon H_2(B,\mathbb{Z}) \to H^2(\hat{\mathcal{M}}(P_k, g), \mathbb{Z})$ which associates cohomology classes in the moduli space with homology classes in B (intuitively, 2-forms on $\hat{\mathcal{M}}(P_k, g)$ with surfaces in B). We will find that this map extends to the Uhlenbeck compactification $\bar{\mu}\colon H_2(B,\mathbb{Z}) \to H^2(\bar{\mathcal{M}}(P_k, g), \mathbb{Z})$. Next we observe that, with certain restrictions on k, $\bar{\mathcal{M}}(P_k, g)$ carries a fundamental class $[\bar{\mathcal{M}}(P_k, g)] \in H_{2d_k}(\bar{\mathcal{M}}(P_k, g), \mathbb{Z})$ which can then be paired with $2d_k$-rank cohomology classes on $\bar{\mathcal{M}}(P_k, g)$ (intuitively, a $2d_k$-form can be integrated over $\bar{\mathcal{M}}(P_k, g)$. Then, for $x_1, \ldots, x_{d_k} \in H_2(B, \mathbb{Z})$ we have $\bar{\mu}(x_1), \ldots, \bar{\mu}(x_{d_k}) \in H^2(\bar{\mathcal{M}}(P_k, g), \mathbb{Z})$ so we can define

$$\begin{aligned}\gamma_{d_k}(x_1, \ldots, x_{d_k}) &= \langle \bar{\mu}(x_1) \smile \cdots \smile \bar{\mu}(x_{d_k}), [\bar{\mathcal{M}}(P_k, g)] \rangle \\ &= \int_{\bar{\mathcal{M}}(P_k, g)} \bar{\mu}(x_1) \wedge \cdots \wedge \bar{\mu}(x_{d_k})\,.\end{aligned}$$

The definition of the μ-map depends on the construction of a certain $SO(3)$-bundle which we shall now sketch. For convenience, we fix a $k > 0$ and a generic metric g on B and write $P_k = P$, $\hat{\mathcal{A}}_3(P_k) = \hat{\mathcal{A}}$, $\mathcal{G}_4(P_k) = \mathcal{G}$, $\hat{\mathcal{B}}_3(P_k) = \hat{\mathcal{B}}$ and $\hat{\mathcal{M}}(P_k, g) = \hat{\mathcal{M}}$. Now, $\mathcal{G}$ acts on $\hat{\mathcal{A}} \times P$ on the right by $(\omega, p) \cdot f = (f^*\omega, f^{-1}(p))$ and the action is free because the elements of $\hat{\mathcal{A}}$ are irreducible $((f^*\omega, f^{-1}(p)) = (\omega, p) \Rightarrow f^*\omega = \omega \Rightarrow f = \pm\,\mathrm{Id}$, but then $f^{-1}(p) = p$ gives $f = \mathrm{Id})$. The orbit space $\hat{\mathcal{A}} \times_{\mathcal{G}} P$ is a Hilbert manifold and

$$\mathcal{G} \hookrightarrow \hat{\mathcal{A}} \times P \longrightarrow \hat{\mathcal{A}} \times_{\mathcal{G}} P$$

is a principal $\mathcal{G}$-bundle, where the last map is the canonical projection $(\omega, p) \to [\omega, p]$.

Notice that there is a natural map of $\hat{\mathcal{A}} \times_{\mathcal{G}} P$ to $\hat{\mathcal{B}} \times B$ given by $[\omega, p] \to ([\omega], \pi(b))$, where π is the projection map of the bundle $SU(2) \hookrightarrow P \xrightarrow{\pi} B$. There is an action of $SU(2)$ on $\hat{\mathcal{A}} \times P$ given by $(\omega, p) \cdot g = (\omega, p \cdot g)$ and this commutes with the action of $\mathcal{G}$ on $\hat{\mathcal{A}} \times P$ because

$$\begin{aligned}((\omega, p) \cdot g) \cdot f = (\omega, p \cdot g) \cdot f &= (f^*\omega, f^{-1}(p \cdot g)) \\ &= (f^*\omega, f^{-1}(p) \cdot g) = (f^*\omega, f^{-1}(p)) \cdot g \\ &= ((\omega, p) \cdot f) \cdot g\,.\end{aligned}$$

Consequently, the $SU(2)$-action on $\hat{\mathcal{A}} \times P$ descends to the quotient $\hat{\mathcal{A}} \times_{\mathcal{G}} P$ to give an action of $SU(2)$ on $\hat{\mathcal{A}} \times_{\mathcal{G}} P$:

$$[\omega, p] \cdot g = [\omega, p \cdot g]\,.$$

This action is not free, but, because the elements of $\hat{\mathcal{A}}$ are irreducible, one finds that $[\omega, p] \cdot g = [\omega, p]$ if and only if $g = \pm\,\mathrm{Id} \in SU(2)$. Thus, there is a free $SU(2)/\pm\,\mathrm{Id} = SO(3)$-action on $\hat{\mathcal{A}} \times_{\mathcal{G}} P$ and we have a principal $SO(3)$-bundle

$$\begin{array}{c} SO(3) \hookrightarrow \hat{\mathcal{A}} \times \mathcal{G}P \\ \Big\downarrow \quad [\omega, p] \longrightarrow ([\omega], \pi(p)) \\ \hat{\mathcal{B}} \times B \end{array}$$

called the **Poincaré bundle** and denoted $\mathcal{P}$.

Being an $SO(3)$-bundle, $\mathcal{P}$ has a 1st Pontryagin class $p_1(\mathcal{P}) \in H^4(\hat{\mathcal{B}} \times B, \mathbb{Z})$. One can show (see [8]) that $p_1(\mathcal{P})$ is divisible by 4 so we may consider the cohomology class

$$-\frac{1}{4}p_1(\mathcal{P}) \in H^4(\hat{\mathcal{B}} \times B, \mathbb{Z}) .$$

Remark: The reason for the $-\frac{1}{4}$ is as follows: When an $SO(3)$-bundle $\mathcal{P}$ lifts to an $SU(2)$-bundle $\mathcal{P}'$ the relationship between the 1st Pontryagin class of $\mathcal{P}$ and the 2nd Chern class of $\mathcal{P}'$ is given by $c_2(\mathcal{P}') = -\frac{1}{4}p_1(\mathcal{P})$. In fact, such a lift actually exists for the Poincaré bundle under our assumptions (B simply connected and $b_2^+(B)$ odd) so we could actually proceed with our construction using $c_2(\mathcal{P}')$ instead of $-\frac{1}{4}p_1(\mathcal{P})$.

Now, there is a general operation in algebraic topology, called the slant product, which maps

$$\begin{aligned} H^{p+q}(X \times Y) \times H_p(Y) &\longrightarrow H^q(X) \\ (\gamma, \alpha) &\longrightarrow \gamma/\alpha. \end{aligned}$$

Fixing a $\gamma_0 \in H^{p+q}(X \times Y)$ then gives a map

$$\begin{aligned} H_p(Y) &\longrightarrow H^2(X) \\ \alpha &\longrightarrow \gamma_0/\alpha . \end{aligned}$$

With $X = \hat{\mathcal{B}}$, $Y = B$, $p = q = 2$ and $\gamma_0 = -\frac{1}{4}p_1(\mathcal{P})$ this gives a map

$$\begin{aligned} H_2(B, \mathbb{Z}) &\longrightarrow H^2(\hat{\mathcal{B}}, \mathbb{Z}) \\ x &\longrightarrow -\frac{1}{4}p_1(\mathcal{P})/x . \end{aligned}$$

Restricting the cohomology classes $-\frac{1}{4}p_1(\mathcal{P})/x$ to $\hat{\mathcal{M}} \subseteq \hat{\mathcal{B}}$ gives the μ-map:

$$\begin{aligned} \mu \colon H_2(B, \mathbb{Z}) &\longrightarrow H^2(\hat{\mathcal{M}}, \mathbb{Z}) \\ \mu(x) &= -\frac{1}{4}p_1(\mathcal{P})/x . \end{aligned}$$

Rather than describing the slant product operation in detail (see [18]) we will show, more informally and in the language of de Rham cohomology, how μ associates a cohomology class of $\hat{\mathcal{M}}$ with a homology class of B. For convenience, we write $H^q_{\text{de Rham}}(X, \mathbb{R}) = H^q(X)$ and $H_p(Y, \mathbb{R}) = H_p(Y)$. Recall that the Künneth formula implies that

$$H^4(\hat{\mathcal{M}} \times B) = \bigoplus_{p+q=4} H^p(\hat{\mathcal{M}}) \otimes H^q(B)$$

so $-{}^1\!/\!_4 p_1(\mathcal{P}) \in H^4(\hat{\mathcal{M}} \times B)$ has a (2, 2)-summand $\alpha = \alpha_1 \otimes \alpha_2 \in H^2(\hat{\mathcal{M}}) \otimes H^2(B)$. But since $H^2(B) \cong (H_2(B))^*$,

$$\begin{aligned} H^2(\hat{\mathcal{M}}) \otimes H^2(B) &\cong H^2(B) \otimes H^2(\hat{\mathcal{M}}) \cong (H_2(B))^* \otimes H^2(\hat{\mathcal{M}}) \\ &\cong \mathrm{Hom}\left(H_2(B), H^2(\hat{\mathcal{M}})\right) \end{aligned}$$

so α corresponds to an element of $\mathrm{Hom}(H_2(B), H^2(\hat{\mathcal{M}}))$ and this is the μ-map. One can describe this more explicitly as follows: An element of the 2nd homology of B can be represented by a smoothly embedded, oriented, surface Σ in B and, with $\alpha = \alpha_1 \otimes \alpha_2$ as above,

$$\mu(\Sigma) = \left(\int_\Sigma \alpha_2\right) \alpha_1 \, .$$

The μ-maps $\mu\colon H_2(B, \mathbb{Z}) \to H^2(\hat{\mathcal{M}}(P_k, g), \mathbb{Z})$ can be shown to extend to the Uhlenbeck compactification $\bar{\mathcal{M}}(P_k, g)$:

$$\bar{\mu}\colon H_2(B, \mathbb{Z}) \longrightarrow H^2(\bar{\mathcal{M}}(P_k, g), \mathbb{Z}) \, .$$

The construction of the extension proceeds one stratum $\mathcal{M}(P_{k-t}, g) \times \mathbb{S}^t(B)$ at a time and requires a detailed understanding of the Taubes glueing procedure for piecing these strata together in $\bar{\mathcal{M}}(P_k, g)$. We simply refer to [5] or [8] for the details.

Now recall that we wish to regard $\bar{\mathcal{M}}(P_k, g)$ as a cycle and pair it with cohomology classes of the form $\bar{\mu}(x_1) \smile \cdots \smile \bar{\mu}(x_{d_k})$ to define the Donaldson invariants. For this to succeed, $\bar{\mathcal{M}}(P_k, g)$ must carry a fundamental class

$$[\bar{\mathcal{M}}(P_k, g)] \in H_{2d_k}\left(\bar{\mathcal{M}}(P_k, g), \mathbb{Z}\right)$$

(intuitively, it must be something we can integrate over). $\bar{\mathcal{M}}(P_k, g)$ is generally not a manifold (for which a fundamental class is guaranteed), but only a stratified space and the algebraic topology of such spaces guarantees a fundamental class only when all of the strata except $\mathcal{M}(P_k, g)$ have codimension at least 2. Writing out the various dimensions shows that this is the case only if k is in the so-called **stable range**

$$k > \frac{3}{4}\left(1 + b_2^+(B)\right) \, , \tag{4.1}$$

or, equivalently,

$$d_k > \frac{3}{2}\left(1 + b_2^+(B)\right) \, . \tag{4.2}$$

Example: For $B = K3$, $b_2^+(K3) = 3$ so the stable range is $k > 3$, i. e., $d_k > 6$.

For such k we are now in a position to define a Donaldson invariant.

Remark: We shall see later that there are various devices for extending the definition outside the stable range.

Here then is the procedure: Let B be a compact, simply connected, oriented, smooth 4-manifold with $b_2^+(B) > 1$ and odd. Choose a generic Riemannian metric g on B so that, for each $k > 0$, $\mathcal{M}(P_k, g) = \hat{\mathcal{M}}(P_k, g)$ is either empty or a smooth manifold of dimension $2d_k = 8k - 3(1 + b_2^+(B))$. Fix an orientation of $H_+^2(B, \mathbb{R})$. Assume k is in the stable range and let $[\bar{\mathcal{M}}(P_k, g)]$ be the fundamental class of $\bar{\mathcal{M}}(P_k, g)$ (if $\mathcal{M}(P_k, g) \neq \emptyset$). Now define the **Donaldson polynomial invariant**

$$\gamma_{d_k}(B)\colon \overbrace{H_2(B,\mathbb{Z}) \times \cdots \times H_2(B,\mathbb{Z})}^{d_k} \longrightarrow \mathbb{Z}$$

by

$$\begin{aligned}\gamma_{d_k}(B)(x_1, \ldots, x_{d_k}) &= \langle \bar{\mu}(x_1) \smile \cdots \smile \bar{\mu}(x_{d_k}), [\bar{\mathcal{M}}(P_k, g)] \rangle \\ &= \int_{\bar{\mathcal{M}}(P_k, g)} \bar{\mu}(x_1) \wedge \cdots \wedge \bar{\mu}(x_{d_k})\end{aligned}$$

if $\mathcal{M}(P_k, g) \neq \emptyset$ and $\gamma_{d_k}(B) = 0$ if $\mathcal{M}(P_k, g) = \emptyset$.

Remarks:

1. That $\gamma_{d_k}(B)$ does not depend on the choice of the generic metric g follows from the fact (Section 1) that a generic variation of g encounters no reducible connections and varies the moduli space within a single homology class in $\hat{\mathcal{B}}(P_k)$.

2. The $\gamma_{d_k}(B)$ are invariant under orientation preserving diffeomorphisms of B that also preserve the orientation of $H_+^2(B, \mathbb{R})$. More precisely, let β denote an orientation for $H_+^2(B, \mathbb{R})$ and temporarily write $\gamma_{d_k}(B, \beta)$ for the corresponding Donaldson invariant. Then, if $f\colon B_1 \to B_2$ is an orientation preserving diffeomorphism,

$$\gamma_{d_k}(B_1, f^*\beta) = \gamma_{d_k}(B_2, \beta) \circ \hat{f}_*,$$

where $\hat{f}_*$ is the multilinear extension of the map f_* induced in homology. See [5] or [8] for details.

3. Reversing the orientation of $H_+^2(B, \mathbb{R})$ reverses the sign of $\gamma_{d_k}(B)$, but, for most examples (we will see one shortly), the invariants vanish for one of the two possible orientations of B.

To describe examples it is generally convenient to identify the multilinear functions $\gamma_{d_k}(B)$, via polarization, with the corresponding homogeneous polynomial, for which we will use the same symbol:

$$\gamma_{d_k}(B)\colon H_2(B,\mathbb{Z}) \longrightarrow \mathbb{Z}$$

$$\gamma_{d_k}(B)(x) = \gamma_{d_k}(B)(\overbrace{x,\ldots,x}^{d_k}).$$

Example: Recall that the **Kummer surface** $K3$ is a compact, simply connected smooth 4-manifold with a natural orientation as a complex manifold and $b_2^+(K3) = 3$ (which is greater than 1 and odd). Moreover, $d_k = 4k - \frac{3}{2}(1 + b_2^+(B)) = 4k - 6$ is even and the stable range is given by $k > 3$ (i. e., $d_k > 6$). For one of the two possible orientations of $H_+^2(B,\mathbb{R})$ it has been shown ([8] or [16]) that

$$\gamma_{d_k}(K3)\colon H_2(K3,\mathbb{Z}) \longrightarrow \mathbb{Z}$$

is given by

$$\gamma_{d_k}(K3)(x) = \frac{(2n)!}{2^n n!}\left(q_{K3}(x,x)\right)^n .$$

If the orientation of $K3$ is reversed the invariants all vanish identically.

There are various devices for extending the definition of the Donaldson invariants outside the stable range and we will sketch one such shortly. First, however, we record two classical theorems of Donaldson and use them to illustrate a typical application of the theory as outlined thus far.

Theorem 4.1. (Connected Sum Theorem) *Let X and Y be compact, simply connected, oriented, smooth 4-manifolds each with $b_2^+ > 0$. Then each $\gamma_{d_k}(X \# Y)$ is identically zero.*

Remark: In particular, taking the connected sum with even one copy of $\mathbb{S}^2 \times \mathbb{S}^2$ kills the Donaldson invariants. For details on the proof of Theorem 4.1 (and Theorem 4.2 below) see [5].

For the next result we recall that a **complex surface** is a complex manifold of complex dimension 2, e. g., $K3$.

Theorem 4.2. (Nonvanishing Theorem) *Let B be a compact, simply connected, complex surface with $b_2^+(B) > 1$. Then, for sufficiently large k, $\gamma_{d_k}(B)$ is not identically zero.*

Remark: Again, the $K3$ surface is an example of a 4-manifold to which the Nonvanishing Theorem applies, but, in this case, we have already noted that the Donaldson invariants are non-zero if $k > 3$. An old question in differential topology that remained open until the appearance of Donaldson's work is the following: The intersection form of $K3$ is $q_{K3} = 3\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \oplus 2E_8$ and $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ is the intersection form of $\mathbb{S}^2 \times \mathbb{S}^2$. Furthermore, the intersection form of a connect sum is the direct sum of the intersection forms, i. e., $q_{B_1 \# B_2} = q_{B_1} \oplus q_{B_2}$. Does $K3$ smoothly split into a connect sum of one, two, or three copies of $\mathbb{S}^2 \times \mathbb{S}^2$ reflecting this algebraic decomposition of q_{K3}? Note that Freedman's Theorem states that, in the topological category, the answer is yes. Indeed, $K3 \cong_h 3(\mathbb{S}^2 \times \mathbb{S}^2)\#2|E_8|$. However, there can be no such smooth decomposition since Theorem 4.1 would then imply the vanishing of all Donaldson invariants for $K3$ and this contradicts Theorem 4.2 (or the explicit calculation of $\gamma_{d_k}(K3)$ for $k > 3$). We now prove a more general result.

Corollary 4.3. *Let B be a compact, simply connected, complex surface with $b_2^+(B) > 1$ and even intersection form. If B is orientation preserving diffeomorphic to a connected sum $X\#Y$, then one of X or Y must be a homotopy 4-sphere.*

Proof: The Nonvanishing Theorem implies that, for sufficiently large k, $\gamma_{d_k}(B)$ is not identically zero. Moreover, since B is simply connected, so are X and Y so the Connected Sum Theorem implies that not both of $b_2^+(X)$ and $b_2^+(Y)$ can be positive. Assume without loss of generality that $b_2^+(X) = 0$. Then Donaldson's Theorem 3.1 implies that either $b_2(X) = 0$ or there exist $\alpha \in H_2(X, \mathbb{Z})$ with $q_X(\alpha, \alpha) = -1$. But q_B is even, so q_X is also even and the latter is impossible. Thus, $b_2(X) = 0$, i. e., $H_2(X, \mathbb{Z}) = 0$, so X is a homotopy 4-sphere. □

We now turn to the problem of extending the definition of the Donaldson invariants in various directions (e. g., outside the stable range). The basic reference for this material is [8]. The first order of business is to define another μ-map (for each $k > 0$) that carries the 0-dimensional homology of B to the 4-dimensional cohomology of $\hat{\mathcal{M}}(P_k, g)$:

$$\mu \colon H_0(B, \mathbb{Z}) \longrightarrow H^4(\hat{\mathcal{M}}(P_k, g), \mathbb{Z}).$$

The procedure is the same as for our previous μ-maps. We let $p_1(\mathcal{P})$ be the 1st Pontryagin class of the Poincaré bundle $\mathcal{P}$:

$$SO(3) \hookrightarrow \hat{\mathcal{A}}_3(P_k) \times_{\mathcal{G}_4(P_k)} P \longrightarrow \hat{\mathcal{B}}_3(P_k) \times B.$$

Since B is connected, $H_0(B, \mathbb{Z}) \cong \mathbb{Z}$ is generated by $1 = [b]$ for any $b \in B$ so it will be enough to define $\mu(1)$. Now,

$$-\frac{1}{4}p_1(\mathcal{P}) \in H^4(\hat{\mathcal{B}}_3(P_k) \times B, \mathbb{Z}) \cong \bigoplus_{p+q=4} H^p(\hat{\mathcal{B}}_3(P_k), \mathbb{Z}) \otimes H^q(B, \mathbb{Z})$$

so $-\frac{1}{4}p_1(\mathcal{P})$ has a (4, 0)-summand $\beta = \beta_1 \otimes \beta_2$ and we define $\mu(1)$ to be the restriction to $\hat{\mathcal{M}}(P_k, g)$ of the cohomology class

$$-\frac{1}{4}p_1(\mathcal{P})/1 = (\beta_2(b))\beta_1 .$$

Now, recall that $\mu\colon H_2(B,\mathbb{Z}) \to H^2(\hat{\mathcal{M}}(P_k, g), \mathbb{Z})$ has an extension $\bar{\mu}$ to the Uhlenbeck compactification $\bar{\mathcal{M}}(P_k, g)$. Unfortunately, it is not the case that $\mu(1) \in H^4(\hat{\mathcal{M}}(P_k, g), \mathbb{Z})$ extends over the entire compactification and this introduces technical difficulties that are discussed at length in [8]. The class $\mu(1)$ does admit an extension $\bar{\mu}(1)$ to the complement of the set of points in $\bar{\mathcal{M}}(P_k, g)$ with trivial background connection and this is enough, under certain restrictions, to extend our definition of the Donaldson invariants to include the 4-dimensional class. Basically, the restriction is that there must be "enough" 2-dimensional classes as well. Specifically, if $d_k = 4k - \frac{3}{2}(1 + b_2^+(B))$ and if a and b are non-negative integers with

$$d_k = a + 2b$$

and

$$a > \frac{3}{4}(1 + b_2^+(B)) ,$$

then, for any $x_1, \ldots, x_a \in H_2(B, \mathbb{Z})$, the class

$$\bar{\mu}(x_1) \smile \cdots \bar{\mu}(x_a) \smile \bar{\mu}(1)^b$$

can be paired with the fundamental class $[\bar{\mathcal{M}}(P_k, g)]$ and we can define

$$\gamma_{d_k}(B)\colon \overbrace{H_2(B,\mathbb{Z}) \times \cdots \times H_2(B,\mathbb{Z})}^{a} \times \overbrace{H_0(B,\mathbb{Z}) \times \cdots \times H_0(B,\mathbb{Z})}^{b} \longrightarrow \mathbb{Z}$$

by

$$\begin{aligned}
&\gamma_{d_k}(B)(x_1, \ldots, x_a, n_1 1, \ldots, n_b 1) \\
&\quad = n_1 \cdots n_b \left\langle \bar{\mu}(x_1) \smile \cdots \smile \bar{\mu}(x_a) \smile \bar{\mu}(1)^b, [\bar{\mathcal{M}}(P_k, g)] \right\rangle \\
&\quad = n_1 \cdots n_b \int_{\bar{\mathcal{M}}(P_k, g)} \bar{\mu}(x_1) \wedge \cdots \wedge \bar{\mu}(x_a) \wedge \bar{\mu}(1)^b .
\end{aligned}$$

(provided $\mathcal{M}(P_k, g) \neq 0$; otherwise, $\gamma_{d_k}(B) = 0$). This definition is, again, independent of the choice of (generic) g and defines an invariant under orientation preserving diffeomorphisms.

Note: The notation for these invariants should really include some reference to a and b, but we have chosen to avoid the additional clutter this would introduce.

At this point we have defined a Donaldson invariant for any $d \equiv -\frac{3}{2}(1+b_2^+(B))$ mod 4 that satisfies $d > \frac{3}{2}(1+b_2^+(B))$ and $d = a+2b$ with $a, b \geq 0$ and $a > \frac{3}{4}(1+b_2^+(B))$. The congruence is unavoidable since $2d$ must be the dimension $8k - 3(1+b_2^+(B))$ of some moduli space, but we now indicate how to circumvent the stable range restrictions. We call an element

$$s = (x_1, \ldots, x_a, n_1 1, \ldots, n_b 1)$$

of $\overbrace{H_2(B,\mathbb{Z}) \times \cdots \times H_2(B,\mathbb{Z})}^{a} \times \overbrace{H_0(B,\mathbb{Z}) \times \cdots \times H_0(B,\mathbb{Z})}^{b}$ **k-stable** for B if $a+2b = d_k = 4k - \frac{3}{2}(1+b_2^+(B)) > \frac{3}{2}(1+b_2^+(B))$, $a, b \geq 0$ and $a > \frac{3}{4}(1+b_2^+(B))$. Our procedure is based on a "blow-up formula" due to Donaldson.

Remark: A **blow-up** of the 4-manifold B is the 4-manifold $B\#\overline{\mathbb{CP}}^2$ obtained by forming the connected sum of B an $\overline{\mathbb{CP}}^2$. We will denote the **n-fold blow-up** $B\#\overbrace{\overline{\mathbb{CP}}^2\#\cdots\#\overline{\mathbb{CP}}^2}^{n}$ of B simply $B\#n\overline{\mathbb{CP}}^2$.

In preparation for Donaldson's results we recall a few basic facts about $\overline{\mathbb{CP}}^2$ and connected sums. First, $H_2(\overline{\mathbb{CP}}^2, \mathbb{Z}) \cong \mathbb{Z}$ and is generated by any $\overline{\mathbb{CP}}^1 \cong \mathbb{S}^2$ in $\overline{\mathbb{CP}}^2$. The intersection form is given, relative to such a basis, by $q_{\overline{\mathbb{CP}}^2} = (-1)$ so $b_2^+(\overline{\mathbb{CP}}^2) = 0$. Now, in general, $H_2(B_1\#B_2, \mathbb{Z}) \cong H_2(B_1, \mathbb{Z}) \oplus H_2(B_2, \mathbb{Z})$ and $q_{B_1\#B_2} = q_{B_1} \oplus q_{B_2}$ so $b_2^+(B_1\#B_2) = b_2^+(B_1) + b_2^+(B_2)$. In particular

$$b_2^+\left(B\#n\overline{\mathbb{CP}}^2\right) = b_2^+(B)$$

for any $n \geq 0$ so the stable ranges for B and $B\#n\overline{\mathbb{CP}}^2$ are the same. Finally, since $q_{\overline{\mathbb{CP}}^2}$ is negative definite, there are no self-dual harmonic 2- forms on $\overline{\mathbb{CP}}^2$ (Section 1) and this permits a natural identification of $H_+^2(B, \mathbb{R})$ and $H_+^2(B\#n\overline{\mathbb{CP}}^2, \mathbb{R})$. In particular, orienting $H_+^2(B, \mathbb{R})$ will also orient any $H_+^2(B\#n\overline{\mathbb{CP}}^2, \mathbb{R})$ and consequently all of the corresponding moduli spaces.

Now consider the manifold $B\#\overline{\mathbb{CP}}^2$. Select a generator e for $H_2(\overline{\mathbb{CP}}^2, \mathbb{Z}) \subseteq H_2(B\#\overline{\mathbb{CP}}^2, \mathbb{Z}) \cong H_2(B, \mathbb{Z}) \oplus H_2(\overline{\mathbb{CP}}^2, \mathbb{Z})$. Donaldson has proved each of the following (see [8]):

1. If s is k-stable for B (and therefore also for $B\#\overline{\mathbb{CP}}^2$), then

$$\gamma_{d_k}\left(B\#\overline{\mathbb{CP}}^2\right)(s) = \gamma_{d_k}(B)(s)\,.$$

2. Suppose $i = 1, 2,$ or 3 and s is not k-stable for B, but $(s, \overbrace{e, \ldots, e}^{i})$ is k-stable for $B\#\overline{\mathbb{CP}}^2$. Then

$$\gamma_{d_k}\left(B\#\overline{\mathbb{CP}}^2\right)(s, \overbrace{e, \ldots, e}^{i}) = 0\,.$$

3. If s is k-stable for B, then (s,e,e,e,e) is $(k+1)$-stable for $B\#\overline{\mathbb{CP}}^2$ and

$$\gamma_{d_{k+1}}\left(B\#\overline{\mathbb{CP}}^2\right)(s,e,e,e,e) = -2\gamma_{d_k}(B)(s)\,.$$

From these one obtains the following theorem.

Theorem 4.4. *Let n be a positive integer and, for each $i=1,\ldots,n$, let e_i be a generator of*

$$H_2(\overline{\mathbb{CP}}^2_i,\mathbb{Z}) \subseteq H_2(B\#n\overline{\mathbb{CP}}^2,\mathbb{Z}) \cong H_2(B,\mathbb{Z}) \oplus \bigoplus_{i=1}^{n} H_2(\overline{\mathbb{CP}}^2_i,\mathbb{Z})\,.$$

If s is k-stable for B, then $(s,e_1,e_1,e_1,e_1,\ldots,e_n,e_n,e_n,e_n)$ is $(k+n)$-stable for $B\#\overline{\mathbb{CP}}^2$ and

$$\begin{aligned} &\gamma_{d_k}(B)(s) \\ &= \left(-\frac{1}{2}\right)^n \gamma_{d_{k+n}}(B\#n\overline{\mathbb{CP}}^2)(s,e_1,e_1,e_1,e_1,\ldots,e_n,e_n,e_n,e_n)\,. \end{aligned} \tag{4.3}$$

The point here is this: (4.3) is known to be true whenever s is known to be k-stable for B, but if s is only assumed to have total degree $2d_k$, then $(s,e_1,e_1,e_1,e_1,\ldots,e_n,e_n,e_n,e_n)$ will be $(k+n)$-stable for $B\#\overline{\mathbb{CP}}^2$ provided only that n is sufficiently large. Moreover, by #3 above, the right-hand side of (4.3) takes the same value for all such sufficiently large n. Thus, the right-hand side of (4.3) allows us to *define* the left-hand side of (4.3) for any s of total degree $2d_k$ for any $k>0$ (note that if k gives $d_k<0$ then the moduli space is generically empty and we again take $\gamma_{d_k}(B)=0$).

Because of the factor $(-\frac{1}{2})^n$ in (4.3) these extended Donaldson invariants $\gamma_{d_k}(B)$, $k>0$, generally take values in $\mathbb{Z}[\frac{1}{2}] \subseteq \mathbb{Q}$ rather than $\mathbb{Z}$. They are still, of course, orientation preserving diffeomorphism invariants of B and agree with the previous definition on the stable elements.

And so, after all of this effort we have arrived only at the definition of the objects to be studied in Donaldson theory. These invariants are, of course, enormously difficult to compute in practice (see [19]). Prior to 1994, much labor was devoted to this task and the rewards were substantial. In 1994, however, two events took place, the first of which we now describe (the second is discussed in the article [15] in these Proceedings).

There are infinitely many Donaldson invariants and it is by no means apparent that there are any relations among them that might diminish the task of computing the invariants for a given 4-manifold B. However, in early 1994, Kronheimer and Mrowka [10] announced a remarkable structure theorem for the Donaldson polynomials which showed that all of the information contained in these invariants could be retrieved from a finite set of data, at least for a very large class of 4-manifolds. In order to describe their result we will need to recast our definition of the Donaldson invariants in somewhat more algebraic terms. Specifically, we consider the graded symmetric algebra

$$\mathbf{A}(B) = \mathrm{Sym}^*(H_0(B,\mathbb{Z}) \oplus H_2(B,\mathbb{Z}))$$

of the $\mathbb{Z}$-module $H_0(B,\mathbb{Z}) \oplus H_2(B,\mathbb{Z})$. Since $H_0(B,\mathbb{Z}) \oplus H_2(B,\mathbb{Z})$ is free and of finite rank (for manifolds B of the type we are considering), $\mathbf{A}(B)$ can be identified with the (commutative) polynomial algebra $\mathbb{Z}[x, h_1, \ldots, h_{b_2(B)}]$, where $h_1, \ldots, h_{b_2(B)}$ is a basis for $H_2(B,\mathbb{Z})$ over $\mathbb{Z}$ and we use x to denote the positive integral generator for $H_0(B,\mathbb{Z}) \cong \mathbb{Z}$ corresponding to the orientation of B (i. e., x is the Poincaré dual of the normalized volume form of B). We assume $\mathbf{A}(B)$ is graded in such a way that the elements of $H_0(B,\mathbb{Z})$ have degree 4 and the elements of $H_2(B,\mathbb{Z})$ have degree 2 and we will reserve the symbol 1 for the unit element of degree 0 in $\mathbf{A}(B)$. Thus, a typical element of $\mathbf{A}(B)$ is a $\mathbb{Z}$-linear combination of elements of the form

$$h_{i_1} \cdots h_{i_a} x^b$$

and the degree of such an element is $2(a+2b)$.

Now, if $d \geq 0$ is an integer and $d \equiv -\frac{3}{2}(1+b_2^+(B)) \mod 4$, then we have defined a Donaldson invariant $\gamma_d(B)$ which, for any $a, b \geq 0$ with $d = a + 2b$, gives a symmetric multilinear map from

$$\overbrace{H_2(B,\mathbb{Z}) \times \cdots \times H_2(B,\mathbb{Z})}^{a} \times \overbrace{H_0(B,\mathbb{Z}) \times \cdots \times H_0(B,\mathbb{Z})}^{b}$$

to $\mathbb{Z}[\frac{1}{2}]$. By the universal property of the tensor product, multilinearity then gives a unique linear map from

$$\overbrace{H_2(B,\mathbb{Z}) \otimes \cdots \otimes H_2(B,\mathbb{Z})}^{a} \otimes \overbrace{H_0(B,\mathbb{Z}) \otimes \cdots \otimes H_0(B,\mathbb{Z})}^{b}$$

to $\mathbb{Z}[\frac{1}{2}]$. By symmetry this then gives a map which associates with every element of $\mathrm{Sym}^*(H_0(B,\mathbb{Z}) \oplus H_2(B,\mathbb{Z}))$ of the form $h_{i_1} \cdots h_{i_a} x^b$ with $a + 2b \equiv -\frac{3}{2}(1+b_2^+(B))) \mod 4$ an element of $\mathbb{Z}[\frac{1}{2}]$. If we denote this value $D_B(h_{i_1} \cdots h_{i_a} x^b)$ and then take $D_B(h_{i_1} \cdots h_{i_a} x^b) = 0$ whenever $a + 2b \not\equiv -\frac{3}{2}(1+b_2^+(B)) \mod 4$ and, finally, extend by linearity, we obtain a map

$$D_B \colon \mathbf{A}(B) \longrightarrow \mathbb{Z}\left[\frac{1}{2}\right].$$

This map, which contains all of the information in the Donaldson invariants, is non-zero only on elements of total degree $2d$ where $d \equiv -\frac{3}{2}(1+b_2^+(B)) \mod 4$.

Now we come to a crucial observation of Kronheimer and Mrowka. Suppose $d \geq 0$ satisfies $d \equiv -\frac{3}{2}(1+b_2^+(B)) \mod 2$, but $d \not\equiv -\frac{3}{2}(1+b_2^+(B)) \mod 4$. Then $D_B(h_{i_1} \cdots h_{i_a}) = 0$. However, since

$$d = -\frac{3}{2}\left(1 + b_2^+(B)\right) + 2k \qquad (k \text{ odd})$$

we have

$$d + 2 = -\frac{3}{2}\left(1 + b_2^+(B)\right) + 4\left(\frac{k+1}{2}\right) = d_{(k+1)/2}\,.$$

Now, $\gamma_{d_{(k+1)/2}}(B)$ is defined and acts on classes $(h_{i_1}, \ldots, h_{i_d}, x)$ since $d + 2(1) = d_{(k+1)/2}$. Thus, whereas $D_B(h_{i_1} \cdots h_{i_d})$ is necessarily zero,

$$D_B(h_{i_1} \cdots h_{i_d} x)$$

may be non-zero. With this as motivation we define a linear map

$$\hat{D}_B \colon \operatorname{Sym}^*(H_2(B,\mathbb{Z})) \longrightarrow \mathbb{Z}\left[\frac{1}{2}\right]$$

by

$$\hat{D}_B(u) = D_B\left(\left(1 + \frac{x}{2}\right)u\right) = D_B(u) + \frac{1}{2}D_B(xu) \tag{4.4}$$

for each u in $\operatorname{Sym}^*(H_2(B,\mathbb{Z}))$. Because of the presence of the second term, $\hat{D}_B$ can be non-zero in degrees $2d$ with $d \equiv -\frac{3}{2}(1 + b_2^+(B)) \mod 2$.

Remark: The $\frac{1}{2}$ in (4.4) is to ensure consistency with the blow-up formulas described earlier.

Finally, following Kronheimer and Mrowka, we define a formal power series $\mathcal{D}_B$ on $H_2(B,\mathbb{Z})$, called the **Donaldson series** of B, by

$$\begin{aligned}\mathcal{D}_B(h) &= \hat{D}_B\left(e^h\right) = D_B\left(\left(1 + \frac{x}{2}\right)e^h\right) = D_B\left(e^h + \frac{1}{2}x\,e^h\right)\\ &= \sum_{d=0}^{\infty}\frac{D_B(h^d)}{d!} + \frac{1}{2}\sum_{d=0}^{\infty}\frac{D_B(xh^d)}{d!}\,.\end{aligned}$$

Notice that, since D_B is non-zero only on elements of degree $2d$, where $d \equiv -\frac{3}{2}(1+b_2^+(B)) \mod 4$, the non-zero terms in these two series involve different values of d (the degree of xh^d is $2d + 4 = 2(d + 2)$ and $d + 2 \not\equiv d \mod 4$). Thus, if one could somehow independently determine the series $\mathcal{D}_B(h)$, then one could extract from it $D_B(h^a)$ and $D_B(xh^a)$ for all $h \in H_2(B,\mathbb{Z})$ and all $a \geq 0$. We now introduce an assumption on B that will, in turn, allow us to extract, from a known $\mathcal{D}_B(h)$, $D_B(x^b h^a)$ for any $b \geq 0$. B is said to be of **simple type** if, for all $z \in \mathbf{A}(B)$

$$D_B(x^2 z) = 4D_B(z)\,. \tag{4.5}$$

Here the 4 is again for consistency with the blow-up formulas. This simple type assumption seems quite special, but there are no known counterexamples to the conjecture that all compact, simply connected, smooth 4-manifolds with b_2^+ greater than 1 and odd are of simple type.

Now, if B is of simple type and if $\mathcal{D}_B(h)$ is known, then one reads $D_B(h^a)$ and $D_B(xh^a)$ off from the series and uses (4.5) to inductively determine $D_B(x^b h^a)$ for any $b \geq 0$. Polarization then gives $D_B(x^b h_{i_1} \cdots h_{i_a})$ for all $h_{i_1}, \ldots, h_{i_a} \in H_2(B, \mathbb{Z})$. This, in particular, determines all of the Donaldson invariants. We conclude that, if B is of simple type, then all of the information contained in the Donaldson invariants is contained in the Donaldson series.

All of this would be rather uninteresting, of course, if it were not possible to obtain information about $\mathcal{D}_B(h)$ independently of actually calculating the Donaldson invariants. The structure theorem of Kronheimer and Mrowka (proved in [11]) asserts that, in fact, the entire Donaldson series is completely determined by a finite set of data.

Theorem 4.5. *Let B be a compact, simply connected, oriented, smooth 4-manifold with $b_2^+(B) > 1$ and odd. Suppose that B is of simple type and that some Donaldson invariant of B is nontrivial. Then there exist finitely many cohomology classes $K_1, \ldots, K_s \in H^2(B, \mathbb{Z})$, called **basic classes**, and non-zero rational numbers $a_1, \ldots, a_s$, called **coefficients**, such that*

$$\mathcal{D}_B(h) = \mathrm{e}^{q_B(h)/2} \sum_{i=1}^{s} a_i\, \mathrm{e}^{K_i(h)}\,.$$

With the appearance of this result in early 1994 Donaldson theory seemed to have turned a corner. Enormously complicated calculations of an apparently infinite set of independent invariants were suddenly replaced by the (certainly not trivial, but at least finite) problem of determining the basic classes and coefficients. As fate would have it, however, the Fall of 1994 witnessed another event which rendered this triumph of Kronheimer and Mrowka moot. Edward Witten, at the end of a (physics) lecture at M.I.T., made a conjecture which, within weeks, brought about the demise of Donaldson theory and initiated an entirely new approach to the study of smooth 4-manifolds. The story of this event and the ensuing frenzy is told by one who was there in [22]. An informal account of the background behind Witten's conjecture is available in [15].

References

[1] Atiyah M. F., N. J. Hitchin and I. M. Singer, *Self-duality in Four Dimensional Riemannian Geometry*, Proc. Roy. Soc. Lond., Series A **362** (1978) 425–461.

[2] Belavin A., A. Polyakov, A. Schwartz and Y. Tyupkin, *Pseudoparticle Solutions of the Yang-Mills Equations*, Phys. Lett. **59B** (1975) 85–87.

[3] Donaldson S. K., *An Application of Gauge Theory to 4-Dimensional Topology*, J. Diff. Geom. **18** (1983) 279–315.

[4] Donaldson S. K., *Polynomial Invariants for Smooth 4-Manifolds, Topology*, **29** (1990) 257–315.

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

[6] Freed D. S. and K. K. Uhlenbeck, *Instantons and Four-Manifolds*, MSRI Publications, Springer-Verlag, New York, Berlin, 1984.

[7] Freedman M. and F. Luo, *Selected Applications of Geometry to Low-Dimensional Topology*, AMS Publications, Providence, RI, 1989.

[8] Friedman R. and J. Morgan, *Smooth Four-Manifolds and Complex Surfaces*, Springer-Verlag, New York, Berlin, 1994.

[9] Kobayashi S. and K. Nomizu, *Foundations of Differential Geometry*, Volumes I and II, Wiley Classics Library, Wiley-Interscience, New York, 1996.

[10] Kronheimer P. B. and T. S. Mrowka, *Recurrence Relations and Asymptotics for Four-manifold Invariants*, Bull. Amer. Math. Soc. **30** (1994) 215–221.

[11] Kronheimer P. B. and T. S. Mrowka, *Embedded Surfaces and the Structure of Donaldson's Polynomial Invariants*, J. Diff. Geom. **3** (1995) 573–734.

[12] Lawson H. B., *The Theory of Gauge Fields in Four Dimensions*, AMS Publications, Providence, RI, 1985.

[13] Naber G. L., *Topology, Geometry and Gauge Fields: Foundations*, Springer-Verlag, New York, Berlin, 1997.

[14] Naber G. L., *Topology, Geometry and Gauge Fields: Interactions*, Springer-Verlag, New York, Berlin, 2000.

[15] Naber G. L., *The Witten Conjecture*, (this issue).

[16] O'Grady K. G., *Donaldson's Polynomials for $K3$ Surfaces*, J. Diff. Geom. **35** (1992) 415–427.

[17] Parker T. H., *Gauge Theories on Four Dimensional Riemannian Manifolds*, Comm. Math. Phys. **85** (1982) 563–602.

[18] Spanier E. H., *Algebraic Topology*, McGraw-Hill, New York 1966.

[19] Stern R. J., *Computing Donaldson Invariants*, In: Gauge Theory and the Topology of Four-Manifolds, AMS Publications, Providence, RI, 1998.

[20] Stong R. E., *Notes on Cobordism Theory*, Princeton Math. Notes, Princeton University Press, Princeton, NJ, 1958.

[21] Taubes C. H., *Self-dual Connections on Non-self-dual 4-Manifolds*, J. Diff. Geom. **17** (1982) 139–170.

[22] Taubes C. H., The Seiberg–Witten Invariants (Videotape) AMS Selected Lectures in Mathematics, 1995.