

# Donaldson = Seiberg–Witten from Mochizuki’s Formula and Instanton Counting

by

Lothar GÖTTSCHE, Hiraku NAKAJIMA and Kōta YOSHIOKA

## Abstract

We propose an explicit formula connecting Donaldson invariants and Seiberg–Witten invariants of a 4-manifold of simple type via Nekrasov’s deformed partition function for the  $N = 2$  SUSY gauge theory with a single fundamental matter. This formula is derived from Mochizuki’s formula, which makes sense and was proved when the 4-manifold is complex projective. Assuming our formula is true for a 4-manifold of simple type, we prove Witten’s conjecture and sum rules for Seiberg–Witten invariants (superconformal simple type condition), conjectured by Mariño, Moore and Peradze.

*2010 Mathematics Subject Classification.* Primary 14D21; Secondary 57R57, 81T13, 81T60.

*Keywords:* Donaldson invariants, Seiberg–Witten invariants, instanton counting.

## Contents

§1. Introduction	308
§2. Preliminaries (I): Donaldson and Seiberg–Witten invariants	311
§3. Preliminaries (II): Instanton counting	316
§4. Mochizuki’s formula and the partition function	319
§5. Blow-up formula for the partition function	327
§6. Seiberg–Witten curves	334
§7. Partition functions at the singular point	340
§8. Computation	348
References	357

---

This is a contribution to the special issue “The golden jubilee of algebraic analysis”.

Communicated by M. Kashiwara. Received January 29, 2010. Revised April 28, 2010.

L. Götsche: International Centre for Theoretical Physics,  
Strada Costiera 11, 34151 Trieste, Italy;  
e-mail: [gottsche@ictp.it](mailto:gottsche@ictp.it)

H. Nakajima: Research Institute for Mathematical Sciences, Kyoto University,  
Kyoto 606-8502, Japan;  
e-mail: [nakajima@kurims.kyoto-u.ac.jp](mailto:nakajima@kurims.kyoto-u.ac.jp)

K. Yoshioka: Department of Mathematics, Faculty of Science, Kobe University,  
Kobe 657-8501, Japan;  
Max-Planck-Institut für Mathematik, Vivatsgasse 7, 53111 Bonn, Germany;  
e-mail: [yoshioka@math.kobe-u.ac.jp](mailto:yoshioka@math.kobe-u.ac.jp)

### §1. Introduction

Let  $X$  be a smooth, compact, connected, and oriented 4-manifold with  $b_1 = 0$  and  $b_+ \geq 3$  odd. We set

$$(K_X^2) := 2\chi(X) + 3\sigma(X), \quad \chi_h(X) := \frac{\chi(X) + \sigma(X)}{4}.$$

When  $X$  is a complex projective surface, these are the self-intersection of the canonical bundle and the holomorphic Euler characteristic respectively, and our notation is consistent.

Let  $\xi \in H^2(X, \mathbb{Z})$ ,  $\alpha \in H_2(X)$  and  $p \in H_0(X)$  be the point class. In [40] Witten explained that the generating function  $\mathcal{D}^\xi(\alpha)$  of Donaldson invariants (see (2.3) for the definition) is related to Seiberg–Witten invariants by

$$(1.1) \quad \begin{aligned} \mathcal{D}^\xi(\alpha) &:= \sum_{n,k} \frac{1}{k!} \left( D^{\xi,n}(\alpha^k) + \frac{1}{2} D^{\xi,n}(\alpha^k p) \right) \\ &= 2^{(K_X^2) - \chi_h(X) + 2} (-1)^{\chi_h(X)} e^{(\alpha^2)/2} \sum_{\mathfrak{s}} \text{SW}(\mathfrak{s}) (-1)^{(\xi, \xi + c_1(\mathfrak{s}))/2} e^{(c_1(\mathfrak{s}), \alpha)}, \end{aligned}$$

where  $(\ , \ )$  is the intersection form,  $(\alpha^2) = (\alpha, \alpha)$ ,  $\text{SW}(\mathfrak{s})$  is the Seiberg–Witten invariant of a  $\text{spin}^c$  structure  $\mathfrak{s}$ , and  $c_1(\mathfrak{s}) = c_1(S^+) \in H^2(X, \mathbb{Z})$  is the first Chern class of the spinor bundle of  $\mathfrak{s}$ . And  $X$  is assumed to be of *SW-simple type*, i.e.,  $c_1(\mathfrak{s})^2 = (K_X^2)$  if  $\text{SW}(\mathfrak{s}) \neq 0$ .

Witten’s argument was based on Seiberg–Witten’s ansatz [37] of  $N = 2$  SUSY gauge theory, which is a physical theory underlying Donaldson invariants [39]. It was not given in a way which mathematicians can justify, so (1.1) has become *Witten’s conjecture* among mathematicians.

Let us explain the main point of Witten’s argument. (See [27, Introduction] for a more detailed exposition for mathematicians.) Seiberg–Witten’s ansatz roughly says that the  $N = 2$  SUSY gauge theory is controlled by a family of elliptic curves (called Seiberg–Witten curves)

$$y^2 = 4x(x^2 + ux + \Lambda^4)$$

parametrized by  $u \in \mathbb{C}$ . Here  $\Lambda$  is a formal variable used to count the dimension of instanton moduli spaces in the prepotential of the theory. (In the Donaldson series, one usually sets  $\Lambda = 1$ .) Witten explained that  $\mathcal{D}^\xi(\alpha)$  is given by an integration over  $u \in \mathbb{C}$ , and the integrand is supported only at points  $u = \pm 2\Lambda^2$ , where the corresponding elliptic curve is singular, when  $b_+ \geq 3$ . Those points contribute as given on the right hand side of (1.1).

In mathematics, the Seiberg–Witten curves appear as elliptic curves for the  $\sigma$ -function in Fintushel–Stern’s blow-up formula [11] for Donaldson invariants, and

the parameter  $u$  corresponds to the point class  $p$ . However, no mathematician succeeded in making Witten’s argument rigorous.

An alternative mathematically rigorous approach was proposed by Pidstrigach and Tyurin [36], and further pursued by Feehan–Leness [8]. It is based on moduli spaces of  $\mathrm{SO}(3)$ -monopoles, which are a higher rank analog of  $\mathrm{U}(1)$ -monopoles used to define Seiberg–Witten invariants. In particular, under a certain technical assumption on a property of  $\mathrm{SO}(3)$ -monopole moduli spaces, Feehan–Leness [8] (see also [6, 7], in particular [9, Th. 3.1]) showed that Donaldson invariants have the form

$$(1.2) \quad D^{\xi, n}(\alpha^k p^l) = \sum_{\mathfrak{s}} f_{k, l}(\chi_h(X), (K_X^2), \mathfrak{s}, \xi, \alpha, \mathfrak{s}_0) \mathrm{SW}(\mathfrak{s}),$$

where the coefficients  $f_{k, l}$  are not explicit, but depend only on the  $\chi_h(X)$ ,  $(K_X^2)$  and various intersection products among  $\mathfrak{s}$ ,  $\xi$ ,  $\alpha$ ,  $\mathfrak{s}_0$ . Here  $\mathfrak{s}_0$  is an auxiliary  $\mathrm{spin}^c$  structure needed for  $\mathrm{SO}(3)$ -monopole moduli spaces. As an application, they proved Witten’s conjecture for  $X$  which satisfies  $(K_X^2) \geq \chi_h(X) - 3$  or is abundant, i.e., the orthogonal complement of Seiberg–Witten classes contains a hyperbolic sublattice [9].

For a complex projective surface  $X$ , Mochizuki, motivated by [36, 8] (and also by [4, 14]), proved a formula expressing Donaldson invariants in the form (1.2), but the coefficients  $f_{k, l}$  are given as residue of an explicit  $\mathbb{C}^*$ -equivariant integral over the product of Hilbert schemes of points on  $X$  (see Theorem 4.1). He obtained the formula by applying the Atiyah–Bott–Berline–Vergne fixed point formula to the algebro-geometric counterpart of  $\mathrm{SO}(3)$ -monopole moduli spaces.

Our first main result (Theorem 4.5) says that Mochizuki’s coefficients are given by leading terms, denoted by  $F_0, H, A, B$ , of Nekrasov’s deformed partition function for the  $N = 2$  SUSY gauge theory with a single fundamental matter, which is the physics counterpart of the  $\mathrm{SO}(3)$ -monopole theory. Thus the coefficients are ‘equivariant  $\mathrm{SO}(3)$ -monopole invariants for  $\mathbb{R}^4$ ’ in some sense.

The proof is almost the same as that of the authors’ wall-crossing formula for Donaldson invariants with  $b_+ = 1$ , expressed in terms of Nekrasov’s partition function for pure gauge theory [15]: By a cobordism argument (due to Ellingsrud–Göttsche–Lehn [5]), it is enough to show it for toric surfaces. Then the integral is given as the product of local contributions from torus fixed points of  $X$ , and the local contribution can be considered as the case  $X = \mathbb{R}^4$ . Thus it is, by its definition, Nekrasov’s partition function.

From this result, we see that Mochizuki’s coefficients depend on the various data in the same way as those of Feehan–Leness. In particular, they make sense also for a smooth 4-manifold  $X$ . (Here  $\mathfrak{s}_0$  is given by the complex structure.) Hoping

that Mochizuki's coefficients are the same as those of Feehan–Leness,<sup>1</sup> we propose a conjecture: our formula remains true for a smooth 4-manifold of SW-simple type (Conjecture 4.6).

Nekrasov's partition functions are defined in a mathematically rigorous way and have explicit combinatorial expressions in terms of Young diagrams [33]. Furthermore, the leading part  $F_0$  is given by certain period integrals over Seiberg–Witten curves [26, 34, 3],  $H$  is explicit, and  $A, B$  are also given in terms of Seiberg–Witten curves [27]. The proofs in [26, 27] were given only for the pure theory, but we extend them for the theory with one matter in this paper using the theory of perverse coherent sheaves [31]. Thus Mochizuki's coefficients are now given by residue of a differential form expressed by Seiberg–Witten curves.

The pole at which we take the residue is at  $u = \infty$ . It is a very deep pole, and a direct computation of the residue looks difficult. Fortunately there is a hint: a similar problem, for certain limits of Donaldson invariants with  $b_+ = 1$ , was analyzed by Göttsche–Zagier [16]. Observing that their differential is defined on  $\mathbb{P}^1$ , holomorphic outside  $\infty, \pm 2\Lambda^2$ , they showed that it is enough to compute the residues at poles  $\pm 2\Lambda^2$  which are simple, and proved an analog of Witten's conjecture. Also this picture is close to Witten's original intuition.<sup>2</sup>

Let us emphasize that the extension of the differential to  $\mathbb{P}^1$  is already a nontrivial assertion. In the original formulation the parameter  $u$  was a *formal* variable used to introduce a generating function of invariants. Therefore it is, a priori, defined only in the formal neighborhood of  $u = \infty$ . The extension is done, so far, by an explicit formula of the differential. Thus the geometric picture of moduli spaces becomes obscure at the points  $\pm 2\Lambda^2$ .

Our situation is similar to one in [16], but slightly different. The Seiberg–Witten curve for the theory with a fundamental matter is

$$y^2 = 4x^2(x + u) + 4m\Lambda^3x + \Lambda^6,$$

and has one more parameter  $m$ , called the *mass* of the matter field. And in our formula, this  $m$  is chosen so that the above curve is singular. Therefore the family of curves is different from what Witten used. We have two features of the new family. First since the curves are singular, the differential is written in terms of elementary functions, and not modular functions as in [16]. This makes our computation much easier. Second, more importantly, we get another pole besides  $\infty, \pm 2\Lambda^2$ , which is

<sup>1</sup>See §4.2 for our heuristic proof of this hope. This paper is motivated by Feehan–Leness' papers, but the proof is independent.

<sup>2</sup>G. Moore pointed out to us that the  $u$ -plane integrand is a total derivative, at least if we take a derivative with respect to the metric. See [22, (11.16)].

called the *superconformal point* in the physics literature. (In the main text, we change the variable from  $u$  to another variable  $\phi$  called the *contact term*.)

The contribution of this point to the gauge theory with one matter was studied by Mariño, Moore and Peradze [20] at a physical level of rigor. They argued that the partition function must be regular at the superconformal point and then this condition leads to sum rules on Seiberg–Witten invariants, i.e.,  $X$  must satisfy the following condition.

**Definition 1.3** ([20]). Suppose that a 4-manifold  $X$  is of SW-simple type. We say  $X$  is of *superconformal simple type* if  $(K_X^2) \geq \chi_h(X) - 3$  or

$$(1.4) \quad \sum_{\mathfrak{s}} (-1)^{(\tilde{w}_2(X), \tilde{w}_2(X) + c_1(\mathfrak{s})) / 2} \text{SW}(\mathfrak{s})(c_1(\mathfrak{s}), \alpha)^n = 0$$

for any integral lift  $\tilde{w}_2(X)$  of  $w_2(X)$  and  $0 \leq n \leq \chi_h(X) - (K_X^2) - 4$ .

Remark that  $(K_X^2) \geq \chi_h(X) - 3$  is the condition which Feehan–Leness [9] assumed to prove Witten’s conjecture. It should be remarked that they also proved that  $X$  is of superconformal simple type if  $X$  is abundant under the same technical assumption as before [10, 9].

We analyze the residue of our differential at the superconformal point and show that 1) the fact that  $\mathcal{D}^\xi(\alpha)$ , up to sign, depends only on  $(\xi \bmod 2)$  implies that  $X$  is of superconformal simple type, and 2) the differential is regular at the superconformal point if  $X$  is of superconformal simple type. Thus the residue vanishes at the superconformal point, and hence we prove Witten’s conjecture for a 4-manifold  $X$  of simple type under Conjecture 4.6, and under no assumption on the complex projective surface  $X$ .

## §2. Preliminaries (I): Donaldson and Seiberg–Witten invariants

### §2.1. Donaldson invariants

Let  $y = (2, \xi, n) \in H^{\text{even}}(X, \mathbb{Z})$ . We take a Riemannian metric  $g$  on  $X$  and consider the moduli space  $M(y)$  of irreducible anti-self-dual connections on the adjoint bundle  $\text{ad}(P)$  of a principal  $U(2)$ -bundle  $P$  with  $c_1(P) = \xi$ ,  $c_2(P) = n$ . For a generic metric  $g$ , this is a manifold of dimension  $8n - 2(\xi^2) - 6\chi_h(X)$ . A choice of an orientation of  $H^+$ , a maximal positive definite subspace of  $H^2(X)$  with respect to the intersection pairing, gives an orientation on  $M(y)$ .

Let  $\mathcal{P} \rightarrow X \times M(y)$  be a universal  $PU(2)$ -bundle and let  $\mu: H_i(X) \rightarrow H^{4-i}(M(y))$  be the  $\mu$ -map defined by  $\mu(\beta) := -\frac{1}{4}p_1(\mathcal{P})/\beta$ . Then the *Donaldson*

invariant of  $X$  is a polynomial on  $H_0(X) \oplus H_2(X)$  defined by

$$(2.1) \quad D^{\xi,n}(\alpha^k p^l) = \int_{M(y)} \mu(\alpha)^k \mu(p)^l,$$

where  $p \in H_0(X)$  is the point class. This is nonzero only when  $k + 2l = 4n - (\xi^2) - 3\chi_h(X)$ . As  $M(y)$  is not compact, this integral must be justified by using the Uhlenbeck compactification of  $M(y)$ . When  $b_+ \geq 3$  as we assumed, the integral is independent of the choice of the Riemannian metric  $g$ . The moduli space does not change on twisting of  $P$  by a line bundle, since the adjoint bundle remains the same. Only the orientation is different. Thus the integral depends only on  $\xi \bmod 2 \in H^2(X, \mathbb{Z}/2)$  up to sign.

We consider the generating function

$$D^\xi(\exp(\alpha z + px)) = \sum_{n,k,l} D^{\xi,n}(\alpha^k p^l) \frac{z^k x^l}{k! l!} \Lambda^{4n - (\xi^2) - 3\chi_h(X)}.$$

Since  $n$  can be read off from  $k, l$  as above, the variable  $\Lambda$  is redundant, and we often put  $\Lambda = 1$ , but it is also useful when we will consider the partition function.

**Definition 2.2.** A 4-manifold  $X$  is of *KM-simple type* if for any  $\xi$  and  $\alpha$ ,

$$\frac{\partial^2}{\partial x^2} D^\xi = 4\Lambda^4 D^\xi.$$

For a 4-manifold of KM-simple type, we define

$$(2.3) \quad \mathcal{D}^\xi(\alpha) := D^\xi\left(\exp(\alpha)\left(1 + \frac{1}{2}p\right)\right) = \sum_{n,k} D^{\xi,n}(\alpha^k) \frac{1}{k!} + \frac{1}{2} \sum_{n,k} D^{\xi,n}(\alpha^k p) \frac{1}{k!}.$$

Kronheimer–Mrowka’s structure theorem [17] says that there is a finite distinguished collection of 2-dimensional cohomology classes  $K_i \in H^2(X, \mathbb{Z})$  and nonzero rational numbers  $\beta_i$  such that

$$\mathcal{D}^\xi(\alpha) = \exp((\alpha^2)/2) \sum_i (-1)^{(\xi, \xi + K_i)/2} \beta_i \exp(K_i, \alpha).$$

Each  $K_i$  is an integral lift of the second Stiefel–Whitney class  $w_2(X)$ .

**§2.2. Complex projective surfaces**

Now suppose  $X$  is a complex projective surface. Take an ample line bundle  $H$  and consider the moduli space  $M_H(y)$  of torsion free  $H$ -semistable sheaves  $E$  with  $c_1(E) = \xi, c_2(E) = n$ . Here we assume  $\xi$  is of type  $(1, 1)$ . We take the orientation on  $H^+$  given by  $c_1(H)$  and the complex orientation on  $H^{0,2}(X)$ .

It is known that Donaldson invariants can be defined using  $M_H(y)$  instead of  $M(y)$  in (2.1) if  $M_H(y)$  is of expected dimension [18, 23]. We define the  $\mu$ -map by using a universal sheaf  $\mathcal{E}$  instead of  $\mathcal{P}$ , as  $\mu(\beta) = (c_2(\mathcal{E}) - c_1(\mathcal{E})^2/4)/\beta$ . The orientation we used above differs from the complex orientation by  $(-1)^{(\xi, \xi + K_X)/2}$ , where  $K_X$  is the canonical class.

If  $M_H(y)$  is not of expected dimension, we consider the blow-up at sufficiently many points  $p_1, \dots, p_N$  disjoint from cycles representing  $\alpha, p$ . Then the moduli becomes of expected dimension on the blow-up if  $N$  is sufficiently large. We then use the blow-up formula as the definition of the integral over  $M_H(y)$ . See [15, §1.1] for detail.

Mochizuki defines the invariants by using obstruction theory on the moduli spaces of pairs of sheaves and their sections with a suitable stability condition. When the vector  $y$  is primitive, the stability is equivalent to the semistability for  $M_H(y)$ , and Mochizuki's moduli is a projective bundle over  $M_H(y)$ . If, furthermore, the moduli space  $M_H(y)$  is of expected dimension, the virtual fundamental class coincides with the ordinary one, and hence Mochizuki's invariants are equal to the usual Donaldson invariants ([21, Lem. 7.3.5]). In order to prove that his invariant coincides with the above invariant for any  $y$ , one needs to prove the blow-up formula for Mochizuki's invariants. It follows a posteriori from our main result that this is true. It should be possible to give a more direct proof by combining the theory of perverse coherent sheaves [29, 30, 31] with Mochizuki's method.

### §2.3. Seiberg–Witten invariants

Let  $\mathfrak{s}$  be a  $\text{spin}^c$  structure and let  $c_1(\mathfrak{s}) = c_1(S^+) \in H^2(X)$  be the first Chern class of its spinor bundle.

Let  $N(\mathfrak{s})$  be the moduli space of the solutions of monopole equations. This is a compact manifold (more precisely, after a perturbation) of dimension  $d(\mathfrak{s}) := (c_1(\mathfrak{s})^2 - (K_X)^2)/4$ . It has the orientation induced from that of  $H^+$  as in the case of Donaldson invariants. Let  $\mathcal{Q}$  be the  $S^1$ -bundle associated with the evaluation homomorphism from the gauge group at a point in  $X$ , and  $c_1(\mathcal{Q})$  be its first Chern class. The *Seiberg–Witten invariant* of  $\mathfrak{s}$  is defined as

$$\text{SW}(\mathfrak{s}) := \int_{N(\mathfrak{s})} c_1(\mathfrak{s})^{d(\mathfrak{s})/2}$$

This is independent of the choice of  $g$  and the perturbation.

We call  $\mathfrak{s}$  (or  $c_1(\mathfrak{s})$ ) a *Seiberg–Witten class* if  $\text{SW}(\mathfrak{s}) \neq 0$ . It is known that there are only finitely many Seiberg–Witten classes.

**Definition 2.4.** A 4-manifold  $X$  is of *SW-simple type* if  $\text{SW}(\mathfrak{s})$  is zero for all  $\mathfrak{s}$  with  $d(\mathfrak{s}) > 0$ .

For  $c \in H^2(X; \mathbb{Z})$  which is a lift of  $w_2(X)$ , we define

$$\text{SW}(c) = \sum_{c_1(\mathfrak{s})=c} \text{SW}(\mathfrak{s}).$$

When  $X$  is a complex projective surface, it is known that all Seiberg–Witten classes are of type  $(1, 1)$ . The moduli space  $N(\mathfrak{s})$  is identified with the moduli space of pairs of a holomorphic line bundle and its section. It is an unperturbed moduli space, and does not have the expected dimension  $d(\mathfrak{s})$  in general, but can be equipped with an obstruction theory to define the invariants [13]. It is also known that  $X$  is of SW-simple type.

We will not use so much results on Seiberg–Witten invariants, except the most basic one:

$$\text{SW}(-\mathfrak{s}) = (-1)^{\chi_h(X)} \text{SW}(\mathfrak{s}),$$

where  $-\mathfrak{s}$  is the complex conjugate of the  $\text{spin}^c$  structure  $\mathfrak{s}$ . (See, e.g., [24, Cor. 6.8.4].)

#### §2.4. Witten’s conjecture

Witten’s conjecture states that if  $X$  is of SW-simple type, it is also of KM-simple type and  $\beta_i, K_i$  are determined by Seiberg–Witten invariants. See (1.1) in Introduction.

**Example 2.5.** Let  $X$  be a K3 surface. The Donaldson series is known [35]:

$$\mathcal{D}^\xi(\alpha) = (-1)^{(\xi, \xi)/2} \exp((\alpha^2)/2).$$

The only Seiberg–Witten class is  $c_1(\mathfrak{s}) = 0$  and  $\text{SW}(\mathfrak{s}) = 1$ .

**Example 2.6.** Let  $X$  be a quintic surface in  $\mathbb{P}^3$ . The Donaldson series was given in [17, Example 2]:

$$\begin{aligned} \mathcal{D}^0(\alpha) &= 8 \exp((\alpha^2)/2) \sinh(K_X, \alpha), \\ \mathcal{D}^{K_X}(\alpha) &= -8 \exp((\alpha^2)/2) \cosh(K_X, \alpha). \end{aligned}$$

We have  $\chi_h = 5$ ,  $(K_X^2) = 5$ . The Seiberg–Witten classes are  $\pm K_X$ , and  $\text{SW}(-K_X) = 1$ ,  $\text{SW}(K_X) = (-1)^{\chi_h} = -1$  by [24, Prop. 7.3.1].

**Example 2.7.** Let  $X$  be an elliptic surface  $X$  without multiple fibers such that  $H^1(X, \mathcal{O}_X) = 0$ . Let  $f$  be the class of a fiber. We have  $K_X = \mathcal{O}_X(df)$  with  $\chi_h(X) = d + 2$  and  $(K_X^2) = 0$ . The Donaldson series is given by Fintushel–Stern [12]:

$$\mathcal{D}^0(\alpha) = \exp((\alpha^2)/2) \sinh^{\chi_h(X)-2}(f, \alpha).$$

The Seiberg–Witten invariants were computed by Friedman–Morgan [13]:

$$\text{SW}((2p - d)f) = (-1)^p \binom{d}{p} \quad \text{for } p = 0, \dots, d, \quad \text{SW}(c) = 0 \quad \text{for other } c.$$

**§2.5. Superconformal simple type**

Let us briefly study the superconformal simple type condition (Definition 1.3) in this subsection. More examples can be found in [20].

If we take another integral lift  $\tilde{w}'_2(X)$  of  $w_2(X)$  in (1.4), we have

$$(-1)^{(\tilde{w}'_2(X), \tilde{w}'_2(X) + c_1(\mathfrak{s}))/2} = (-1)^{(\tilde{w}_2(X), \tilde{w}_2(X) + c_1(\mathfrak{s}))/2} (-1)^{((\tilde{w}'_2(X) - \tilde{w}_2(X))/2)^2}.$$

Therefore it is enough to assume (1.4) for *some* integral lift  $\tilde{w}_2(X)$  of  $w_2(X)$ . We will consider the case when  $X$  is a complex projective surface, and take  $K_X$  as a lift.

If  $X$  is a minimal surface of general type, we have the Noether inequality  $(K_X^2)/2 + 2 \geq \chi_h(X) - 1$ . Together with  $(K_X^2) \geq 1$ , it implies  $(K_X^2) \geq \chi_h(X) - 3$ . Thus  $X$  is of superconformal simple type by definition ([20, §7.1]). In fact, it is known that the Seiberg–Witten classes are  $\pm K_X$ , and  $\text{SW}(-K_X) = 1, \text{SW}(K_X) = (-1)^{\chi_h(X)}$  (see, e.g., [24]). Therefore we cannot have a nontrivial identity like (1.4).

We consider

$$\mathcal{SW}(\alpha) := \sum_{\mathfrak{s}} (-1)^{(K_X, K_X + c_1(\mathfrak{s}))/2} \text{SW}(\mathfrak{s}) \exp(c_1(\mathfrak{s}), \alpha).$$

The condition (1.4) is equivalent to  $\mathcal{SW}(\alpha)$  having zero of order  $\geq \chi_h(X) - (K_X^2) - 3$  at  $\alpha = 0$ . We have

$$\mathcal{SW}(-\alpha) = (-1)^{\chi_h(X) - (K_X^2)} \mathcal{SW}(\alpha)$$

as  $\text{SW}(-c) = (-1)^{\chi_h(X)} \text{SW}(c)$ . Therefore  $\mathcal{SW}$  is an even (resp. odd) function if  $\chi_h(X) - (K_X^2)$  is even (resp. odd). Therefore the order of zero is automatically  $\geq \chi_h(X) - (K_X^2) - 2$  under the above condition.

**Example 2.8** ([20, §7.2]). Let  $X$  be an elliptic surface without multiple fibers such that  $H^1(X, \mathcal{O}_X) = 0$ . Let  $f$  be the class of a fiber. Then  $K_X = \mathcal{O}_X(df)$  with  $\chi_h(X) = d + 2$ . We have  $(K_X^2) = 0$ . The Seiberg–Witten invariants were computed by Friedman–Morgan [13] as in Example 2.7. Therefore

$$\mathcal{SW}(\alpha) = (-2)^{\chi_h(X) - 2} \sinh^{\chi_h(X) - 2}(f, \alpha).$$

This has zero of order  $\chi_h(X) - 2$  at  $\alpha = 0$ . Hence  $X$  is of superconformal simple type. This example can be generalized to the case of elliptic surfaces with multiple fibers.

**Example 2.9** ([20, §7.3]). Consider a one-point blow-up  $\widehat{X} \rightarrow X$ . Let  $C$  be the exceptional divisor. We have  $(K_{\widehat{X}}^2) = (K_X^2) - 1$  and  $\chi_h(\widehat{X}) = \chi_h(X)$ . Let us add the subscript  $X$  and  $\widehat{X}$  to the Seiberg–Witten invariants  $\text{SW}$  (and  $\mathcal{SW}$ ) in order to clarify which surface we consider. Then we have  $\text{SW}_{\widehat{X}}(c \pm C) = \text{SW}_X(c)$  for  $c \in H^2(X)$  and the other  $\text{SW}_{\widehat{X}}(c + nC)$  vanish. Therefore

$$\mathcal{SW}_{\widehat{X}}(\alpha + zC) = -2\mathcal{SW}_X(\alpha) \sinh(z).$$

Thus  $\mathcal{SW}_X(\alpha)$  has a zero of order  $\geq \chi_h(X) - (K_X^2) - 3$  at  $\alpha = 0$  if and only if  $\mathcal{SW}_{\widehat{X}}$  has a zero of order  $\geq \chi_h(X) - (K_X^2) - 2 = \chi_h(\widehat{X}) - (K_{\widehat{X}}^2) - 3$  at  $(\alpha, z) = 0$ . Thus  $X$  is of superconformal simple type if and only if so is  $\widehat{X}$ .

From these two examples and the classification of complex surfaces, we conclude that all complex projective surfaces with  $p_g > 0$ ,  $b_1 = 0$  are of superconformal simple type. (See [20, §7.3].)

### §3. Preliminaries (II): Instanton counting

#### §3.1. Framed moduli spaces of torsion free sheaves

We briefly recall the framed moduli spaces of torsion free sheaves on  $\mathbb{P}^2$ . See [25, Chap. 2] and [27, §3] for more detail.

Let  $\ell_\infty$  be the line at infinity of  $\mathbb{P}^2$ . A *framed sheaf*  $(E, \varphi)$  on  $\mathbb{P}^2$  is a pair consisting of

- a coherent sheaf  $E$ , which is locally free in a neighborhood of  $\ell_\infty$ , and
- an isomorphism  $\varphi: E|_{\ell_\infty} \rightarrow \mathcal{O}_{\ell_\infty}^{\oplus r}$ , where  $r$  is the rank of  $E$ .

Let  $M(r, n)$  be the moduli space of framed sheaves  $(E, \varphi)$  of rank  $r$  and  $c_2(E) = n$ . This is a nonsingular quasi-projective variety of dimension  $2rn$ . It has an ADHM type description.

Let  $M_0(r, n)$  be the corresponding Uhlenbeck partial compactification. There is a projective morphism  $\pi: M(r, n) \rightarrow M_0(r, n)$ . This is a crepant resolution of  $M_0(r, n)$ .

Let  $\mathbb{C}^* \times \mathbb{C}^*$  act on  $\mathbb{P}^2$  by  $[z_0 : z_1 : z_2] \mapsto [z_0 : t_1 z_1 : t_2 z_2]$ , where the line  $\ell_\infty$  at infinity is  $z_0 = 0$ . Let  $T$  be the maximal torus of  $\text{SL}_r(\mathbb{C})$  consisting of the diagonal matrices and let  $\widetilde{T} = \mathbb{C}^* \times \mathbb{C}^* \times T$ . It acts on  $M(r, n)$  as follows: the first factor  $\mathbb{C}^* \times \mathbb{C}^*$  acts by pull-backs of sheaves  $E$ , and  $T$  acts by the change of the framing  $\varphi$ . It also acts on  $M_0(r, n)$  and  $\pi$  is equivariant. We consider the equivariant homology groups  $H_*^{\widetilde{T}}(M(r, n))$  and  $H_*^{\widetilde{T}}(M_0(r, n))$ . Let  $[M(r, n)]$  and  $[M_0(r, n)]$  be the fundamental classes.

The fixed points  $M(r, n)^{\tilde{T}}$  are parametrized by  $r$ -tuples of Young diagrams  $\vec{Y} = (Y_1, \dots, Y_r)$ . Each  $Y_\alpha$  corresponds to a monomial ideal  $I_\alpha$  of the polynomial ring  $\mathbb{C}[x, y]$ , and gives a framed rank 1 sheaf. The direct sum  $I_1 \oplus \dots \oplus I_r$  is a torus fixed point. The equivariant Euler class  $\text{Eu}(T_{\vec{Y}}M(r, n))$  of the tangent space of  $M(r, n)$  at  $\vec{Y}$  is given by a certain combinatorial formula (see [27, §§3,4]). Its explicit form is used only implicitly in this paper when we quote Maple calculations. On the other hand,  $M_0(r, n)$  has a unique fixed point: the rank  $r$  trivial sheaf together with a singularity concentrated at the origin.

Let  $\varepsilon_1, \varepsilon_2, a_1, \dots, a_r$  (with  $a_1 + \dots + a_r = 0$ ) be the coordinates of the Lie algebra of  $\tilde{T}$ . We also use the notation  $\vec{a} = (a_1, \dots, a_r)$ . The equivariant cohomology  $H_{\tilde{T}}^*(\text{pt})$  of a single point is naturally identified with the polynomial ring  $S(\tilde{T}) := \mathbb{C}[\varepsilon_1, \varepsilon_2, a_2, \dots, a_r]$ . Let  $\mathfrak{S}(\tilde{T})$  be its quotient field. The localization theorem for the equivariant homology group says that the push-forward homomorphism  $\iota_{0*}$  of the inclusion  $M_0(r, n)^{\tilde{T}} \rightarrow M_0(r, n)$  induces an isomorphism of equivariant homology groups after tensoring by  $\mathfrak{S}(T)$ . Since  $M_0(r, n)^{\tilde{T}}$  is a single point, as we remarked above, we have

$$\iota_{0*} : H_{\tilde{T}}^*(M_0(r, n)^{\tilde{T}}) \otimes_{S(\tilde{T})} \mathfrak{S}(T) = \mathfrak{S}(T) \xrightarrow{\cong} H_{\tilde{T}}^*(M_0(r, n)) \otimes_{S(\tilde{T})} \mathfrak{S}(T).$$

Let  $\iota_{0*}^{-1}$  be the inverse of  $\iota_{0*}$ .

We also have an isomorphism

$$\iota_* : H_{\tilde{T}}^*(M(r, n)^{\tilde{T}}) \otimes_{S(\tilde{T})} \mathfrak{S}(T) = \mathfrak{S}(T)^{\oplus \#\{\vec{Y}\}} \xrightarrow{\cong} H_{\tilde{T}}^*(M(r, n)) \otimes_{S(\tilde{T})} \mathfrak{S}(T),$$

where  $\iota : M(r, n)^{\tilde{T}} \rightarrow M(r, n)$ . By the functoriality of push-forward homomorphisms, we have

$$(3.1) \quad \sum_{\vec{Y}} \iota_*^{-1} = \iota_{0*}^{-1} \circ \pi_*,$$

where  $\sum_{\vec{Y}}$  is the map  $\mathfrak{S}(T)^{\oplus \#\{\vec{Y}\}} \rightarrow \mathfrak{S}(T)$  defined by taking the sum of components.

Since  $M(r, n)$  is smooth,  $\iota_*^{-1}$  is given by

$$\frac{\iota^*(\bullet)}{\text{Eu}(T_{\vec{Y}}M(r, n))},$$

where  $\iota^*$  is the pull-back homomorphism of equivariant cohomology groups, considered as a map between equivariant homology groups via Poincaré duality.

Nekrasov’s deformed partition function for *pure* gauge theory is defined as the generating function of  $\iota_{0*}^{-1} \pi_* [M(r, n)]$ , where we let  $n$  vary. By the discussion above, it is the generating function of  $1/\text{Eu}(T_{\vec{Y}}M(r, n))$  for all  $\vec{Y}$ . It was introduced in [33] and studied in [34, 26, 27].

**§3.2. The partition function for the theory with fundamental matters**

We need a variant of the partition function. It is called the partition function for the theory with fundamental matters in the physics literature.

Over the moduli space  $M(r, n)$ , we have a natural vector bundle  $\mathcal{V}$ , whose fiber at  $(E, \varphi)$  is  $H^1(E(-\ell_\infty))$ . It has rank  $n$ . If  $\mathcal{E}$  denotes the universal sheaf on  $\mathbb{P}^2 \times M(r, n)$ , we have  $\mathcal{V} = R^1 q_{2*}(\mathcal{E} \otimes q_1^*(\mathcal{O}(-\ell_\infty)))$ , where  $q_1, q_2$  are the projections from  $\mathbb{P}^2 \times M(r, n)$  to the first and second factors respectively.

In fact, a computation shows that it is more natural to replace  $H^1(E(-\ell_\infty))$  by the  $L^2$ -kernel of the Dirac operator  $D_A: E \otimes S^- \rightarrow E \otimes S^+$ , where  $A$  is the instanton corresponding to  $E$  (assuming it is locally free). This is simply given by tensoring the half canonical bundle  $K_{\mathbb{C}^2}^{1/2}$  of  $\mathbb{C}^2$ , i.e. a trivial line bundle with weight  $e^{-(\varepsilon_1 + \varepsilon_2)/2}$ . This makes sense even if  $E$  is not locally free, so we can consider it as a definition of the kernel of the Dirac operator. Thus we consider  $R^1 q_{2*}(\mathcal{E} \otimes q_1^*(\mathcal{O}(-\ell_\infty) \otimes K_{\mathbb{C}^2}^{1/2}))$ . As  $K_{\mathbb{C}^2}^{1/2}$  is a trivial bundle, this factors as  $\mathcal{V} \otimes K_{\mathbb{C}^2}^{1/2}$ .

For a positive integer  $N_f$ , we consider a vector space  $M = \mathbb{C}^{N_f}$ , called the *flavor space*. The group  $GL(M)$  naturally acts on  $M$ . Let  $T_M$  be the diagonal subgroup. Let  $\vec{m} = (m_1, \dots, m_{N_f})$  denote an element in  $\text{Lie } T_M$ . We consider the equivariant class  $\text{Eu}(\mathcal{V} \otimes K_{\mathbb{C}^2}^{1/2} \otimes M) \cap [M(r, n)]$ . This has the degree (or virtual dimension)  $(2r - N_f)n$ . The theory is called *conformal* when  $N_f = 2r$  and *asymptotically free* when  $N_f < 2r$ . We assume  $N_f < 2r$  hereafter. We set  $\gamma := 2r - N_f$ .

We define the instanton part of the partition function by

$$(3.2) \quad Z^{\text{inst}}(\varepsilon_1, \varepsilon_2, \vec{a}, \vec{m}; \Lambda) = \sum_{n=0}^{\infty} \Lambda^{\gamma n} \iota_{0*}^{-1} \pi_* (\text{Eu}(\mathcal{V} \otimes K_{\mathbb{C}^2}^{1/2} \otimes M) \cap [M(r, n)]),$$

where  $\Lambda$  is a formal variable.

Using (3.1) we can replace  $\iota_{0*}^{-1} \pi_*$  by  $\sum_{\vec{Y}} \iota_*^{-1}$ . Then we get

$$(3.3) \quad Z^{\text{inst}}(\varepsilon_1, \varepsilon_2, \vec{a}, \vec{m}; \Lambda) = \sum_{n=0}^{\infty} \sum_{|\vec{Y}|=n} \Lambda^{\gamma|\vec{Y}|} \frac{\text{Eu}(\mathcal{V}|_{\vec{Y}} \otimes K_{\mathbb{C}^2}^{1/2} \otimes M)}{\text{Eu}(T_{\vec{Y}} M(r, n))},$$

where  $\mathcal{V}|_{\vec{Y}}$  is the fiber of  $\mathcal{V}$  at the fixed point  $\vec{Y}$ , and  $|\vec{Y}|$  is the sum of the numbers of boxes in the Young diagrams  $Y_\alpha$ . The right hand side has a combinatorial expression, which we do not present since we use it only implicitly.

It is known that

- (1)  $\varepsilon_1 \varepsilon_2 \log Z^{\text{inst}}(\varepsilon_1, \varepsilon_2, \vec{a}, \vec{m}; \Lambda)$  is regular at  $\varepsilon_1, \varepsilon_2 = 0$ , and hence has the expansion

$$(3.4) \quad F_0^{\text{inst}}(\vec{a}, \vec{m}; \Lambda) + (\varepsilon_1 + \varepsilon_2) H^{\text{inst}}(\vec{a}, \vec{m}; \Lambda) + \varepsilon_1 \varepsilon_2 A^{\text{inst}}(\vec{a}, \vec{m}; \Lambda) + \frac{\varepsilon_1^2 + \varepsilon_2^2}{3} B^{\text{inst}}(\vec{a}, \vec{m}; \Lambda) + \dots$$

- (2) The leading term  $F_0^{\text{inst}}(\vec{a}, \vec{m}; \Lambda)$  is the instanton part of the Seiberg–Witten prepotential.

For the pure theory (i.e.,  $N_f = 0$ ) these were proved by the second and third-named authors [26], Nekrasov–Okounkov [34], and Braverman–Etingof [3] independently. The proof in [34] works also for theories with matters. See [34, §7.1]. We also need to know the next three terms  $H, A, B$ . These were computed in [27] for the pure theory. The corresponding results for our case  $r = 2, N_f = 1$  along the argument in [26, 27] will be explained below (§6). We need to use the theory which we have developed in [29, 30, 31]. In particular, we have  $H^{\text{inst}} \equiv 0$ , which means that the partition function is ‘topological’:  $H^{\text{inst}}$  is coupled with  $\varepsilon_1 + \varepsilon_2 = -c_1(K_{\mathbb{C}^2})$ , which depends on the complex structure, but it vanishes.

Since we will only consider the case  $r = 2, N_f = 1$ , we denote  $a_2, m_1$  simply by  $a, m$  respectively. In application to Mochizuki’s formula below, we need to specialize  $a = m$ . This is well-defined: Setting  $a = m$  means that we restrict the acting group from  $\tilde{T} \times T_M$  to a smaller subgroup. But the smaller subgroup still has the same fixed points (as  $T_M$  acts trivially), and the fixed point formula can be specialized.

In view of Conjecture 4.6 below, it is desirable to have a direct definition of the partition function in terms of the Uhlenbeck compactification  $M_0(r, n)$ , not appealing to the algebro-geometric object  $M(r, n)$ . Since  $\mathcal{V}$  is not a pull-back from  $M_0(r, n)$ , this is a nontrivial problem. If we consider  $M_0(r, n)$  as an affine algebraic variety, then  $\pi_*(\text{Eu}(\mathcal{V} \otimes K_{\mathbb{C}^2}^{1/2} \otimes M) \cap [M(r, n)])$  is a limit of the formal  $\tilde{T} \times T_M$ -character of the space of sections of certain virtual sheaves on  $M_0(r, n)$  as in [26, §4]. It should be possible to replace this virtual sheaf by a complex of vector bundles.

#### §4. Mochizuki’s formula and the partition function

As we mentioned in the Introduction, we will use Mochizuki’s formula relating Donaldson invariants and Seiberg–Witten invariants. Before stating his formula, let us briefly explain the idea behind its proof. The reader can safely jump to §4.1 if he/she accepts Mochizuki’s formula. But the authors encourage the reader to learn Mochizuki’s beautiful ideas. Of course he/she should read the book [21] for more detail.

When  $X$  is a complex projective surface, Mochizuki first developed the obstruction theory for moduli spaces of pairs of sheaves and sections and related spaces. Then he obtained a general machinery to write down the difference of invariants for two moduli spaces defined with different stability conditions. A point is to introduce a  $\mathbb{C}^*$ -equivariant obstruction theory on the ‘master space’ containing

two moduli spaces as  $\mathbb{C}^*$ -fixed point loci. He integrated the class  $\exp(\mu(\alpha z + px)) \cup a$  over the master space, where  $a$  is the generator of the equivariant cohomology group  $H_{\mathbb{C}^*}^*(\text{pt})$  of a single point. Since the integral vanishes at the nonequivariant limit  $a = 0$ , the sum of the residues of the fixed point loci contributions is zero by the Atiyah–Bott–Berline–Vergne fixed point formula. This gives the difference of the invariants as the sum of the residues of ‘exceptional’ fixed points loci contributions. The exceptional fixed points are products of lower rank sheaves and pairs. Up to this point, the framework is essentially the same as the  $\text{SO}(3)$ -monopole cobordism program, except for a systematic use of obstruction theory. But a crucial difference is that Mochizuki’s obstruction theory enables him to treat moduli spaces as if they were *smooth*. In particular, his ‘residues’ are given explicitly in terms of equivariant Euler classes of virtual normal bundles.

He applied this theory to the case of moduli spaces of rank 2 pairs. When a stability condition is suitably chosen, moduli spaces of pairs are projective bundles over moduli of genuine sheaves, thus the invariants are reduced to Donaldson invariants. On the other hand, for another stability condition, moduli spaces become empty sets. The difference of the invariants, which is just Donaldson invariants, is given by the sum of the residues of equivariant integrals over other ‘exceptional’ fixed point loci, which are moduli spaces of pairs of rank 1 sheaves with sections of one factor. These exceptional contributions can be identified with the product of the Seiberg–Witten invariant and an equivariant integral over the product  $X^{[n_1]} \times X^{[n_2]}$  of Hilbert schemes of points in  $X$ . This is because rank 1 sheaves are just ideal sheaves twisted by line bundles. The class  $Q$  appearing in the formula below is the equivariant Euler class of the normal bundle mentioned above.

#### §4.1. Mochizuki’s formula

Let  $y = (2, \xi, n)$ ,  $\alpha, p, z, x$  be as in the definition of Donaldson invariants (§2.1). Suppose that we have decompositions  $\xi = \xi_1 + \xi_2$  and  $n - (\xi_1, \xi_2) = n_1 + n_2$ . We denote by  $e^{\xi_i}$  the holomorphic line bundle whose first Chern class is  $\xi_i$ . Let  $\mathcal{I}_i$  (resp.  $\mathcal{O}_{Z_i}$ ) denote the universal ideal sheaf (resp. subscheme) over  $X \times X^{[n_i]}$ . Their pull-backs to  $X \times X^{[n_1]} \times X^{[n_2]}$  are denoted by the same notation. Let  $q_2: X \times X^{[n_1]} \times X^{[n_2]} \rightarrow X^{[n_1]} \times X^{[n_2]}$  be the projection.

Let  $\mathbb{C}^*$  act trivially on  $X^{[n_1]} \times X^{[n_2]}$  and consider the equivariant cohomology group  $H_{\mathbb{C}^*}^*(X^{[n_1]} \times X^{[n_2]}) \cong H^*(X^{[n_1]} \times X^{[n_2]})[a]$ , where  $a$  is the variable for  $H_{\mathbb{C}^*}^*(\text{pt})$ , i.e.,  $H_{\mathbb{C}^*}^*(\text{pt}) = \mathbb{C}[a]$ . We consider the following equivariant cohomology classes on  $X^{[n_1]} \times X^{[n_2]}$ :

$$P(\mathcal{I}_1 e^{\xi_1 - a} \oplus \mathcal{I}_2 e^{\xi_2 + a}) := \exp(-\text{ch}_2(\mathcal{I}_1 e^{\xi_1 - a - \xi/2} \oplus \mathcal{I}_2 e^{\xi_2 + a - \xi/2}) / (\alpha z + px)),$$

$$Q(\mathcal{I}_1 e^{\xi_1 - a} \oplus \mathcal{I}_2 e^{\xi_2 + a}) := \text{Eu}(-\text{Ext}_{q_2}^*(\mathcal{I}_1 e^{\xi_1 - a}, \mathcal{I}_2 e^{\xi_2 + a})) \text{Eu}(-\text{Ext}_{q_2}^*(\mathcal{I}_2 e^{\xi_2 + a}, \mathcal{I}_1 e^{\xi_1 - a})),$$

where  $\text{Ext}_{q_2}^*$  is the alternating sum  $\text{Ext}_{q_2}^0 - \text{Ext}_{q_2}^1 + \text{Ext}_{q_2}^2$ , and  $\text{Ext}_{q_2}^\bullet$  is the derived functor of the composite  $q_{2*} \circ \mathcal{H}om$ .

Roughly speaking,  $Q$  is the equivariant Euler class of the virtual normal bundle of  $X^{[n_1]} \times X^{[n_2]}$  in  $M_H(y)$ . Here one should consider that the embedding is given by  $(I_1, I_2) \mapsto e^{\xi_1} I_1 \oplus e^{\xi_2} I_2$ . And  $P$  is the restriction of the integrand appearing in Donaldson invariants. But the precise formulation requires the master space, and is omitted in this paper.

Note that  $Q$  is invertible in  $H^*(X^{[n_1]} \times X^{[n_2]}][a, a^{-1}]$  as it has the form  $Q(\mathcal{I}_1 e^{\xi_1 - a} \oplus \mathcal{I}_2 e^{\xi_2 + a}) = a^N + (\text{lower degree in } a)$  for some  $N$ . We consider the following class in  $H^*(X^{[n_1]} \times X^{[n_2]}][a, a^{-1}]$ :

$$\begin{aligned} &\tilde{\Psi}(\xi_1, \xi_2, n_1, n_2; a) \\ &:= \frac{P(\mathcal{I}_1 e^{\xi_1 - a} \oplus \mathcal{I}_2 e^{\xi_2 + a}) \text{Eu}(H^*((\mathcal{O}/\mathcal{I}_1)e^{\xi_1})) \text{Eu}(H^*((\mathcal{O}/\mathcal{I}_2)e^{\xi_2 + 2a}))}{Q(\mathcal{I}_1 e^{\xi_1 - a} \oplus \mathcal{I}_2 e^{\xi_2 + a}) (2a)^{n_1 + n_2 - p_g}}, \end{aligned}$$

where  $H^*(\mathcal{O}/\mathcal{I}_i)$  is the alternating sum of the higher direct image sheaves  $R^\bullet q_{2*}(\mathcal{O}/\mathcal{I}_i)$ . This is the same as Mochizuki’s  $\Psi$  ([21, §1.4.2]), except that we do not take the residue with respect to  $a$ . Therefore we add tildes in the notation.

We set

$$\tilde{\mathcal{A}}(\xi_1, y; a) = 2^{1-\chi(y)} \sum_{n_1 + n_2 = n - (\xi_1, \xi_2)} \int_{X^{[n_1]} \times X^{[n_2]}} \tilde{\Psi}(\xi_1, \xi_2, n_1, n_2; a),$$

where  $\chi(y)$  is the Euler characteristic of the class  $y$ . By Riemann–Roch, we have

$$\chi(y) = \frac{(\xi, \xi - K_X)}{2} + 2\chi_h(X) - n.$$

**Theorem 4.1** ([21, Th. 1.4.6]). *Assume that  $\chi(y) > 0$ ,  $(\xi, H)/2 > (K_X, H)$  and  $(\xi, H) > (c_1(\mathfrak{s}) + K_X, H)$  for any Seiberg–Witten class  $\mathfrak{s}$ . Then*

$$(4.2) \quad \frac{1}{2} \int_{M_H(y)} \exp(\mu(\alpha z + px)) = \sum_{\xi_1} \text{SW}(\tilde{\xi}_1) \text{Res}_{a=\infty} \tilde{\mathcal{A}}(\xi_1, y; a) da,$$

where  $\tilde{\xi}_1 := 2\xi_1 - K_X$ .

Let us give several remarks.

**Remarks 4.3.** (1) The left hand side is Mochizuki’s definition of the invariant using the obstruction theory. It is equal to the usual Donaldson invariant if  $y$  is

primitive and  $M_H(y)$  is of expected dimension. This is not an essential assumption, as we explained in §2.2.

(2) Mochizuki took the residue at  $a = 0$ , instead of  $a = \infty$ . But ours is just the negative of Mochizuki's, as  $\tilde{\mathcal{A}}(\xi_1, y; a)$  is in  $\mathbb{C}[a, a^{-1}]$ .

(3) The factor  $1/2$  on the left hand side comes from Mochizuki's convention. He considered integration over the moduli space of *oriented* sheaves. There is a natural étale proper morphism from the oriented moduli space to the usual one of degree  $(\text{rank})^{-1} = 1/2$ .

(4) The assumption is satisfied if we replace  $y$  by  $ye^{kH}$  for sufficiently large  $k$ . But it is not clear, a priori, that the right hand side is independent of  $k$ . This will become important for our later analysis of the residue of  $\tilde{\mathcal{A}}$ .

(5) Mochizuki denoted the usual Seiberg–Witten invariant by  $\widetilde{\text{SW}}$  and set  $\text{SW}(\xi_1) = \widetilde{\text{SW}}(2\xi_1 - K_X)$ . We keep  $\text{SW}$  for the notation of the usual Seiberg–Witten invariant. On the other hand, the Seiberg–Witten class  $2\xi_1 - K_X$  will naturally appear in Witten's formula (1.1). Therefore we have denoted it by  $\tilde{\xi}_1$ . Thus our  $\text{SW}(\tilde{\xi}_1)$  is Mochizuki's  $\text{SW}(\xi_1)$ .

(6) Since the expected dimension  $\dim M_H(y)$  is  $4n - (\xi^2) - 3\chi_h(X)$ , we have

$$(4.4) \quad 4\chi(y) = ((\xi - K_X)^2) - (K_X^2) - \dim M_H(y) + 5\chi_h(X).$$

**§4.2. Formula in terms of the partition function**

Now we prove our first main result. Recall  $y = (2, \xi, n)$ . Let us introduce the generating function of the  $\tilde{\mathcal{A}}(\xi_1, y; a)$ :

$$\mathcal{B}(\xi_1, \xi; a) := \sum_n \Lambda^{4n - (\xi^2) - 3\chi_h(X)} \tilde{\mathcal{A}}(\xi_1, (2, \xi, n); a).$$

**Theorem 4.5.** *We have*

$$\begin{aligned} \mathcal{B}(\xi_1, \xi; a) da &= \frac{da}{a} (-1)^{(\xi, \xi + K_X)/2 + (K_X, K_X + \tilde{\xi}_1)/2 + \chi_h(X)} 2^{-2\chi_h(X) - (\xi - K_X - \tilde{\xi}_1, \xi - K_X)/2} \\ &\times \left(\frac{2a}{\Lambda}\right)^{((\xi - K_X)^2) + (K_X^2) + 3\chi_h(X) - 2(\xi - K_X, \tilde{\xi}_1)} \exp(-(\xi - K_X - \tilde{\xi}_1, \alpha)az - a^2x) \\ &\times \exp\left[\frac{1}{3} \frac{\partial F_0^{\text{inst}}}{\partial \log \Lambda} x + \left(\frac{1}{8} \frac{\partial^2 F_0^{\text{inst}}}{\partial a^2} + \frac{1}{4} \frac{\partial^2 F_0^{\text{inst}}}{\partial a \partial m} + \frac{1}{8} \frac{\partial^2 F_0^{\text{inst}}}{\partial m^2}\right)((\xi - K_X)^2) \right. \\ &\quad \left. - \frac{1}{4} \left(\frac{\partial^2 F_0^{\text{inst}}}{\partial a \partial m} + \frac{\partial^2 F_0^{\text{inst}}}{\partial a^2}\right)(\xi - K_X, \tilde{\xi}_1) \right. \\ &\quad \left. + \frac{1}{6} \left(\frac{\partial^2 F_0^{\text{inst}}}{\partial a \partial \log \Lambda} + \frac{\partial^2 F_0^{\text{inst}}}{\partial m \partial \log \Lambda}\right)(\xi - K_X, \alpha)z - \frac{1}{6} \frac{\partial^2 F_0^{\text{inst}}}{\partial a \partial \log \Lambda}(\tilde{\xi}_1, \alpha)z \right. \\ &\quad \left. + \frac{1}{18} \frac{\partial^2 F_0^{\text{inst}}}{(\partial \log \Lambda)^2}(\alpha^2)z^2 + \chi_h(X)(12A^{\text{inst}} - 8B^{\text{inst}}) + (K_X^2) \left(B^{\text{inst}} - A^{\text{inst}} + \frac{1}{8} \frac{\partial^2 F_0^{\text{inst}}}{\partial a^2}\right)\right], \end{aligned}$$

where  $\tilde{\xi}_1 = 2\xi_1 - K_X$  as above, and the derivatives of  $F_0^{\text{inst}}$ ,  $A^{\text{inst}}$  and  $B^{\text{inst}}$  are evaluated at  $(a, m, \Lambda) = (a, a, \Lambda^{4/3}a^{-1/3})$ .

Observe that our formula does not depend on the complex structure of  $X$  when we consider the canonical class  $K_X$  as a choice of a  $\text{spin}^c$  structure. Therefore, the above expression makes sense for a smooth 4-manifold  $X$ . Further observe that  $K_X$  appears in the above expression only as either  $(K_X^2)$  or the combination  $\xi - K_X$ , except in the sign factor. If we ignore the sign, the Donaldson invariants depend only on  $(\xi \bmod 2)$ , so we can consider  $\xi - K_X$  as auxiliary cohomology class. The only requirement is that it is equal to  $(\xi \bmod 2) + w_2(X)$  in  $H^2(X, \mathbb{Z}/2)$ .

Therefore we pose the following:

**Conjecture 4.6.** Mochizuki’s result (Theorem 4.1) holds for a smooth 4-manifold  $X$  with  $b_1 = 0$ ,  $b_+ \geq 3$  odd, up to sign, if we replace  $\tilde{\mathcal{A}}(\xi_1, y; a)$  by coefficients of  $\mathcal{B}(\xi_1, \xi; a)$  in Theorem 4.5.

We have the conditions  $(\xi, H)/2 > (K_X, H)$ ,  $(\xi, H) > (c_1(\mathfrak{s}) + K_X, H)$ , which we do not know how to interpret for a smooth 4-manifold  $X$ . Therefore we just ignore this condition and conjecture that Mochizuki’s result holds without it.

This conjecture is compatible with Feehan–Leness’ result (1.2). Our formula in Theorem 4.5 involves only the intersection pairings among  $\tilde{\xi}_1$ ,  $\xi - K_X$  and  $\alpha$ . Their formula involves an auxiliary cohomology class, denoted by  $\Lambda$  in [9, Th. 3.1], and equal to  $\Lambda = c_1(\mathfrak{s}_0) + \xi$  for a chosen  $\text{spin}^c$  structure  $\mathfrak{s}_0$ . We take the canonical  $\text{spin}^c$  structure of the complex surface  $X$  as  $\mathfrak{s}_0$ , so their  $\Lambda$  should be identified with our  $\xi - K_X$ . In fact,  $\Lambda$  satisfies the same condition which we have assumed for  $\xi - K_X$ . It is required to satisfy the same condition as  $\chi(y) > 0$  thanks to (4.4) (written as ‘ $I(\Lambda) > \delta$ ’ [9]). We also remark that the exponent  $((\xi - K_X)^2) + (K_X^2) + 3\chi_h(X) - 2(\xi - K_X, \tilde{\xi}_1)$  of  $2a/\Lambda$  is equal to  $-r(\Lambda, c_1(\mathfrak{s}))$  in [7, (1.12)] if we take  $\Lambda = \xi - K_X$ ,  $\mathfrak{s} = \tilde{\xi}_1$  and replace  $(K_X^2)$  by  $(c_1(\mathfrak{s})^2)$ .

Thus our conjecture follows immediately if the coefficients  $f_{k,l}$  appearing in Feehan–Leness’ formula (1.2) are the same as ours. This does not directly follow from Feehan–Leness’ statement itself, as Seiberg–Witten invariants satisfy nontrivial relations, namely superconformal simple type conditions, therefore the coefficients are not uniquely determined.

The authors’ heuristic proof is the following: there is a morphism from Mochizuki’s master space to the moduli space of  $\text{SO}(3)$ -monopole when  $X$  is complex projective. Then by the functoriality of push-forward homomorphisms as used in (3.1), the contributions of Seiberg–Witten invariants are the same for Mochizuki’s and Feehan–Leness’ formulas. Since Feehan–Leness’ formula is universal, it is enough to calculate them for complex projective  $X$ , and our calculation gives the answer.

This proof works only for  $X$  of simple type and for which there is a complex projective surface  $X_0$  with  $\chi(X) = \chi(X_0)$ ,  $\sigma(X) = \sigma(X_0)$ . To generalize it to hypothetical  $X$  of non-simple type, we need to connect Feehan–Leness’ coefficients with Nekrasov’s partition function more directly.

*Proof of Theorem 4.5.* The proof is similar to that of the wall-crossing formula for  $b_+ = 1$  in [15]. When the argument is really the same, we just refer to the corresponding argument there.

We denote  $a$  in Mochizuki’s  $\tilde{\mathcal{A}}$  by  $s$  for a moment.

We first write  $\mathcal{B}$  as a product of the ‘perturbative term’, i.e., an expression independent of  $n_1, n_2$  and the ‘instanton part’, which is 1 if  $n_1 = n_2 = 0$ . For the term  $P$ , we have

$$P(\mathcal{I}_1 e^{\xi_1 - s} \oplus \mathcal{I}_2 e^{\xi_2 + s}) = \exp(-(\xi_2 - \xi_1, \alpha)sz - s^2x) \exp([c_2(\mathcal{I}_1) + c_2(\mathcal{I}_2)]/(\alpha z + px)).$$

Thus the perturbative term is  $\exp(-(\xi_2 - \xi_1, \alpha)sz - s^2x)$ . For  $Q$ , the perturbative term is

$$\begin{aligned} & \text{Eu}(H^*(\mathcal{O}_X(\xi_1 - \xi_2))e^{-2s}) \text{Eu}(H^*(\mathcal{O}_X(\xi_2 - \xi_1))e^{2s}) \\ &= (-2s)^{\chi(\mathcal{O}_X(\xi_1 - \xi_2))} (2s)^{\chi(\mathcal{O}_X(\xi_2 - \xi_1))} \\ &= (-1)^{(\xi_1 - \xi_2, \xi_1 - \xi_2 - K_X)/2 + \chi_h(X)} (2s)^{((\xi_1 - \xi_2)^2) + 2\chi_h(X)} \\ &= (-1)^{(\xi, \xi - K_X)/2 + (K_X, \xi_2) + \chi_h(X)} (2s)^{((\xi_1 - \xi_2)^2) + 2\chi_h(X)}. \end{aligned}$$

We also have  $2^{1-\chi(y)} (2s)^{p_g - n_1 - n_2}$  whose perturbative part is

$$2^{1 - (\xi, \xi - K_X)/2 - 2\chi_h(X) + (\xi_1, \xi_2)} (2s)^{\chi_h(X) - 1}.$$

For the power of  $\Lambda$ , the perturbative part is

$$\Lambda^{4(\xi_1, \xi_2) - (\xi^2) - 3\chi_h(X)} = \Lambda^{-((\xi_1 - \xi_2)^2) - 3\chi_h(X)}.$$

Combining all these terms, we find that the perturbative part of  $\mathcal{B}$  is

$$(4.7) \quad (-1)^{(\xi, \xi - K_X)/2 + (K_X, \xi_2) + \chi_h(X)} \times \frac{1}{s} \left( \frac{2s}{\Lambda} \right)^{((\xi_1 - \xi_2)^2) + 3\chi_h(X)} e^{-(\xi_2 - \xi_1, \alpha)sz - s^2x} 2^{-2\chi_h(X) - (\xi, \xi - K_X)/2 + (\xi_1, \xi_2)}.$$

By the argument in [15, §5], it is enough to compute the instanton part for a toric surface  $X$ .

We write  $\chi := \chi(X)$  for brevity. Let  $p_1, \dots, p_\chi$  be the torus fixed points,  $x_i, y_i$  the torus equivariant coordinates at  $p_i$ , and  $w(x_i), w(y_i)$  the weights of the torus action. As in [15, §3.2] we apply the Atiyah–Bott–Berline–Vergne fixed point

formula to  $\tilde{\mathcal{A}}(\xi_1, y; s)$ . At torus fixed points, the ideal sheaves  $\mathcal{I}_1, \mathcal{I}_2$  are the intersection of ideal sheaves supported at  $p_i$ . Accordingly the cohomology groups in  $Q$  and the matter factor  $\text{Eu}(H^*((\mathcal{O}/\mathcal{I}_1)e^{\xi_1})) \text{Eu}(H^*((\mathcal{O}/\mathcal{I}_2)e^{\xi_2+2s}))$  decompose as products of local contributions at  $p_i$ . As in [15, §3.2] we will identify these local contributions with factors in the partition function  $Z^{\text{inst}}$ .

Let us first study how variables appearing in  $Z^{\text{inst}}$  will be identified with expressions in the local contribution of  $\tilde{\mathcal{A}}$  at  $p_i$ . The variables  $\varepsilon_1, \varepsilon_2$  in  $Z^{\text{inst}}$  are identified with  $w(x_i), w(y_i)$ . In order to identify  $a_1, a_2$ , consider the factor  $Q$ . In the definition of the partition function the first Chern class of the universal sheaf is normalized to be 0 as  $a_1 + a_2 = 0$ . In view of [15, Lemma 3.4], this normalization must be performed for  $Q$  as

$$\text{Ext}_{q_2}^*(\mathcal{I}_1 e^{\xi_1-s}, \mathcal{I}_2 e^{\xi_2+s}) = \text{Ext}_{q_2}^*(\mathcal{I}_1 e^{\xi_1-s-\xi/2}, \mathcal{I}_2 e^{\xi_2+s-\xi/2}).$$

Thus we get the same expression appearing in  $P$ , and we will identify variables as

$$\begin{aligned} a_1 &= -s + \iota_{p_i}^*(\xi_1 - \xi/2) = -s - \iota_{p_i}^*(\xi_2 - \xi_1)/2, \\ a_2 &= s + \iota_{p_i}^*(\xi_2 - \xi/2) = s + \iota_{p_i}^*(\xi_2 - \xi_1)/2 \end{aligned}$$

in  $Z^{\text{inst}}$  and  $\tilde{\mathcal{A}}$ . Here  $\iota_{p_i}^*$  is the pull-back homomorphism associated with the inclusion  $\iota_{p_i}$  of the fixed point  $p_i$  into  $X$ .

Accordingly we normalize the matter factor as

$$\begin{aligned} &\text{Eu}(H^*((\mathcal{O}/\mathcal{I}_1)e^{\xi_1})) \text{Eu}(H^*((\mathcal{O}/\mathcal{I}_2)e^{\xi_2+2s})) \\ &= \text{Eu}(H^*((\mathcal{O}/\mathcal{I}_1)e^{\xi_1-s-\xi/2+s+\xi/2})) \text{Eu}(H^*((\mathcal{O}/\mathcal{I}_2)e^{\xi_2+s-\xi/2+s+\xi/2})). \end{aligned}$$

Recalling that we put  $K_{\mathbb{C}^2}^{1/2}$  in the partition function, we identify the variable  $m$  for the matter with  $s + \iota_{p_i}^*(\xi - K_X)/2$ , as  $\xi_\alpha - s - \xi/2$  is  $a_\alpha$  for  $\alpha = 1, 2$ .

Next we consider the variable  $\Lambda$ . After removing the perturbative part as above, we consider  $\Lambda^{4(n_1+n_2)}$  in  $\mathcal{B}$ . On the other hand, we use  $\Lambda^{3(n_1+n_2)}$  in the definition of the partition function. We combine this with  $s^{n_1+n_2}$  in  $\tilde{\Psi}$ , which we then absorb into the variable  $\Lambda$  in the partition function (3.2). Therefore  $\Lambda$  in (3.2) will be replaced by  $\Lambda^{4/3}s^{-1/3}$ .

Now we use the argument in [15, §3.2] to write the instanton part of  $\mathcal{B}$  in terms of the partition function:

$$\lim_{\varepsilon_1, \varepsilon_2 \rightarrow 0} \prod_{i=1}^X Z^{\text{inst}} \left( w(x_i), w(y_i), \iota_{p_i}^* \left( \frac{\xi_2 - \xi_1}{2} \right) + s, \iota_{p_i}^* \left( \frac{\xi - K_X}{2} \right) + s; \frac{\Lambda^{4/3}}{s^{1/3}} e^{\iota_{p_i}^* \left( \frac{\alpha z + px}{3} \right)} \right).$$

We need to explain the last expression  $e^{\iota_{p_i}^* \left( \frac{\alpha z + px}{3} \right)}$ . This comes from  $\exp([c_2(\mathcal{I}_1) + c_2(\mathcal{I}_2)]/(\alpha z + px))$ , which is the instanton part of  $P$ . We use the same argument as in [15, Cor. 3.18], which was based on [27, §4.5]. Let us briefly

recall the point of the argument: We can put more variables  $\vec{\tau} = (\tau_\rho)_{\rho \geq 1}$  into the partition function  $Z^{\text{inst}}$  as in [15, (1.4)], [27, §4.5]. But we only need  $\tau_1$  since we only use  $\text{ch}_2$  and not higher Chern classes in the Donaldson invariants. Then  $\tau_1$  can be absorbed into the variable  $\Lambda$  as  $\text{ch}_2$  is determined by  $n$  of  $M(r, n)$ . We identify  $\tau_1 = -i_{p_i}^*(\alpha z + px)$  as in [15]. In fact, the absorption of  $\tau_1$  into  $\Lambda$  is simpler than in [27, §4.5], as we do not put the perturbative term in the partition function. We just need to note that it is a multiplication of  $e^{i_{p_i}^*((\alpha z + px)/3)}$  instead of  $e^{i_{p_i}^*((\alpha z + px)/4)}$ , because we use  $\Lambda^{\gamma n} = \Lambda^{3n}$  instead of  $\Lambda^{4n}$  in the definition of the partition function. The partition function  $Z^{\text{inst}}$  with the extra variable  $\tau_1$  does not have the property that its constant part in  $\Lambda$  is 1. However, by the first displayed formula in [27, §4.5], when  $\tau_1$  is absorbed into  $\Lambda$ ,  $Z^{\text{inst}}$  is multiplied by a function independent of  $\Lambda$ , so that afterwards its constant part in  $\Lambda$  is 1, matching the instanton part of  $\mathcal{B}$ .

We now use the expansion (3.4) together with  $H^{\text{inst}} = 0$ . As in [15, proof of Th. 4.2], we have

$$\begin{aligned} & \prod_{i=1}^X Z^{\text{inst}} \left( w(x_i), w(y_i), i_{p_i}^* \left( \frac{\xi_2 - \xi_1}{2} \right) + s, i_{p_i}^* \left( \frac{\xi - K_X}{2} \right) + s; \Lambda \left( \frac{\Lambda}{s} \right)^{1/3} e^{i_{p_i}^*((\alpha z + px)/3)} \right) \\ &= \exp \left[ \sum_i \frac{1}{w(x_i)w(y_i)} \left( F_0^{\text{inst}} + \frac{\partial F_0^{\text{inst}}}{\partial a} i_{p_i}^* \left( \frac{\xi_2 - \xi_1}{2} \right) + \frac{\partial F_0^{\text{inst}}}{\partial m} i_{p_i}^* \left( \frac{\xi - K_X}{2} \right) \right. \right. \\ & \quad + \frac{\partial F_0^{\text{inst}}}{\partial \log \Lambda} i_{p_i}^* \left( \frac{\alpha z + px}{3} \right) + \frac{1}{2} \frac{\partial^2 F_0^{\text{inst}}}{\partial a^2} i_{p_i}^* \left( \frac{\xi_2 - \xi_1}{2} \right)^2 \\ & \quad + \frac{\partial^2 F_0^{\text{inst}}}{\partial a \partial m} i_{p_i}^* \left( \frac{\xi_2 - \xi_1}{2} \right) i_{p_i}^* \left( \frac{\xi - K_X}{2} \right) \\ & \quad + \frac{1}{2} \frac{\partial^2 F_0^{\text{inst}}}{\partial m^2} i_{p_i}^* \left( \frac{\xi - K_X}{2} \right)^2 + \frac{\partial^2 F_0^{\text{inst}}}{\partial a \partial \log \Lambda} i_{p_i}^* \left( \frac{\xi_2 - \xi_1}{2} \right) i_{p_i}^* \left( \frac{\alpha z + px}{3} \right) \\ & \quad + \frac{\partial^2 F_0^{\text{inst}}}{\partial m \partial \log \Lambda} i_{p_i}^* \left( \frac{\xi - K_X}{2} \right) i_{p_i}^* \left( \frac{\alpha z + px}{3} \right) + \frac{1}{2} \frac{\partial^2 F_0^{\text{inst}}}{\partial \log \Lambda^2} i_{p_i}^* \left( \frac{\alpha z + px}{3} \right)^2 \\ & \quad \left. \left. + w(x_i)w(y_i)A^{\text{inst}} + \frac{w(x_i)^2 + w(y_i)^2}{3} B^{\text{inst}} \right) \right] + O(\varepsilon_1, \varepsilon_2) \\ &= \exp \left[ \frac{1}{3} \frac{\partial F_0^{\text{inst}}}{\partial \log \Lambda} x + \frac{1}{8} \frac{\partial^2 F_0^{\text{inst}}}{\partial a^2} ((\xi_2 - \xi_1)^2) + \frac{1}{4} \frac{\partial^2 F_0^{\text{inst}}}{\partial a \partial m} (\xi_2 - \xi_1, \xi - K_X) \right. \\ & \quad + \frac{1}{8} \frac{\partial^2 F_0^{\text{inst}}}{\partial m^2} ((\xi - K_X)^2) + \frac{1}{6} \frac{\partial^2 F_0^{\text{inst}}}{\partial a \partial \log \Lambda} (\xi_2 - \xi_1, \alpha)z + \frac{1}{6} \frac{\partial^2 F_0^{\text{inst}}}{\partial m \partial \log \Lambda} (\xi - K_X, \alpha)z \\ & \quad \left. + \frac{1}{18} \frac{\partial^2 F_0^{\text{inst}}}{\partial \log \Lambda^2} (\alpha^2)z^2 + \chi(X)A^{\text{inst}} + \sigma(X)B^{\text{inst}} \right] + O(\varepsilon_1, \varepsilon_2), \end{aligned}$$

where we evaluate the derivatives of  $F_0^{\text{inst}}$ ,  $A^{\text{inst}}$ ,  $B^{\text{inst}}$  at  $a_2 = a = s$ ,  $m = s$ ,  $\Lambda = \Lambda^{4/3}/s^{1/3}$ . We now safely change  $s$  back to  $a$ .

We use  $\chi(X) = 12\chi_h(X) - (K_X)^2$ ,  $\sigma(X) = (K_X^2) - 8\chi_h(X)$ ,  $\xi = \xi_1 + \xi_2$ ,  $\tilde{\xi}_1 = 2\xi_1 - K_X$  and  $(\tilde{\xi}_1^2) = (K_X^2)$  to get the assertion, where the last equality is nothing but the SW-simple type condition.  $\square$

**§5. Blow-up formula for the partition function**

We start to analyze the partition function in this section. Our technique is the same as in [26, 27, 28]: we study the blow-up formula of the partition function.

**§5.1. Partition function on the blow-up**

Let  $p: \widehat{\mathbb{P}^2} \rightarrow \mathbb{P}^2$  be the blow-up of  $\mathbb{P}^2$  at the origin  $[1 : 0 : 0]$ . Let  $C = p^{-1}([1 : 0 : 0])$  be the exceptional divisor. Let  $\widehat{M}(r, k, n)$  be the moduli space of framed sheaves  $(E, \Phi)$  on  $\widehat{\mathbb{P}^2}$  with rank  $r$ ,  $c_1(E) = kC$ ,  $(c_2(E) - (r - 1)c_1(E)^2 / (2r), [\widehat{\mathbb{P}^2}]) = n$ , where the framing is defined on  $p^{-1}(\ell_\infty)$ . (See [26, §3] or [27, §3.2].) This is nonsingular quasi-projective of dimension  $2rn$ . We normalize as  $0 \leq k < r$ . This is always possible by twisting by a power of  $\mathcal{O}(C)$ . There is a projective morphism  $\widehat{\pi}: \widehat{M}(r, k, n) \rightarrow M_0(r, n - k(r - k) / (2r))$ .

We pull-back the  $\mathbb{C}^* \times \mathbb{C}^*$ -action on  $\mathbb{P}^2$  to  $\widehat{\mathbb{P}^2}$ . Then we have an action of  $\widetilde{T}$  on  $\widehat{M}(r, k, n)$  as in the case of  $M(r, n)$ . The action is lifted to the universal sheaf  $\mathcal{E}$  on  $\widehat{\mathbb{P}^2} \times \widehat{M}(r, k, n)$ . The morphism  $\widehat{\pi}$  is  $\widetilde{T}$ -equivariant.

We define  $\mu(C)$  as appeared in the definition of Donaldson’s invariants:

$$\mu(C) = \left( c_2(\mathcal{E}) - \frac{r - 1}{2r} c_1(\mathcal{E})^2 \right) / [C] \in H_T^2(\widehat{M}(r, k, n)).$$

Over  $\widehat{M}(r, k, n)$  we have two natural vector bundles, which correspond to  $\mathcal{V}$ :

$$\mathcal{V}_0 := R^1 q_{2*}(\mathcal{E} \otimes q_1^* \mathcal{O}(-\ell_\infty)), \quad \mathcal{V}_1 := R^1 q_{2*}(\mathcal{E} \otimes q_1^* \mathcal{O}(C - \ell_\infty)).$$

These are vector bundles of rank  $n + k^2 / (2r) - k/2$  and  $n + k^2 / (2r) + k/2$  respectively thanks to the vanishing of other higher direct image sheaves, and play a fundamental role in the ADHM type description of  $\widehat{M}(r, k, n)$  (see, e.g., [29]).

Therefore we have two possible choices of *matters* on blow-up. Here we take  $\mathcal{V}_0$  since the  $\mathcal{V}_1$  version can be reduced to the  $\mathcal{V}_0$  one after twisting by the line bundle  $\mathcal{O}(C)$ . We define the partition function (or rather the correlation function since we put the operator  $\mu(C)$ ) as in (3.2) by

$$\begin{aligned} & \widehat{Z}_{c_1=kC}^{\text{inst}}(\varepsilon_1, \varepsilon_2, \vec{a}, \vec{m}; t; \Lambda) \\ & := \Lambda^{N_f k(r-k)/(2r)} \sum_{n=0}^{\infty} \Lambda^{\gamma n} \iota_{0*}^{-1} \widehat{\pi}_* (e^{t\mu(C)} \cap \text{Eu}(\mathcal{V}_0 \otimes p^*(K_{\mathbb{C}^2}^{1/2}) \otimes M) \cap [\widehat{M}(r, k, n)]). \end{aligned}$$

Here  $p^*(K_{\mathbb{C}^2}^{1/2})$  is the trivial line bundle with weight  $e^{-(\varepsilon_1+\varepsilon_2)/2}$ . It looks a little artificial, but is necessary as in the case of  $\mathbb{C}^2$ . The square root  $K_{\widehat{\mathbb{C}^2}}^{1/2}$  does not make sense since  $\widehat{\mathbb{C}^2}$  is not spin.

As in the case of the original partition function  $Z^{\text{inst}}(\varepsilon_1, \varepsilon_2, \vec{a}, \vec{m}; \Lambda)$ , this one also has a combinatorial expression like (3.3). We do not write it down here, we only explain the parameter set for the fixed points. Similar to the case of  $M(r, n)^{\widehat{T}}$ , it is the set of triples  $(\vec{k}, \vec{Y}^1, \vec{Y}^2)$  of an  $r$ -tuple of integers  $\vec{k} = (k_1, \dots, k_r)$  and the pair of  $r$ -tuples of Young diagrams  $\vec{Y}^1 = (Y_1^1, \dots, Y_r^1)$ ,  $\vec{Y}^2 = (Y_1^2, \dots, Y_r^2)$ . The corresponding framed sheaf is  $I_1(k_1 C) \oplus \dots \oplus I_r(k_r C)$ , where  $I_\alpha$  is an ideal sheaf fixed by the  $\mathbb{C}^* \times \mathbb{C}^*$ -action. The blow-up  $\widehat{\mathbb{C}^2}$  has two fixed points  $p_1, p_2$ , and  $I_\alpha$  is given by two monomial ideals with respect to toric coordinates at  $p_1$  and  $p_2$ . In this way,  $I_\alpha$  is parametrized by a pair of Young diagrams  $(Y_\alpha^1, Y_\alpha^2)$ .

From this combinatorial description of the fixed point set, we can write down the correlation function  $\widehat{Z}_{c_1=kC}^{\text{inst}}$  as sum over the lattice for  $\{\vec{k}\}$  of products of two  $Z^{\text{inst}}$ 's for  $p_1, p_2$ , and contribution from line bundles  $\mathcal{O}(k_\alpha C)$ . We postpone to write down the explicit formula until we introduce the perturbative term in the next subsection.

**§5.2. Perturbative term**

The partition function defined above does not behave well in many aspects. It is more natural to add what is called the *perturbative term*, which is an explicit function. We recall its definition in this subsection. We return back to arbitrary  $r, N_f$ .

Let  $\gamma_{\varepsilon_1, \varepsilon_2}(x; \Lambda)$  be the function used to define the perturbative part of the partition function in [27, §E]:

$$(5.1) \quad \gamma_{\varepsilon_1, \varepsilon_2}(x; \Lambda) = \frac{1}{\varepsilon_1 \varepsilon_2} \left\{ -\frac{1}{2} x^2 \log\left(\frac{x}{\Lambda}\right) + \frac{3}{4} x^2 \right\} + \frac{\varepsilon_1 + \varepsilon_2}{2 \varepsilon_1 \varepsilon_2} \left\{ -x \log\left(\frac{x}{\Lambda}\right) + x \right\} - \frac{\varepsilon_1^2 + \varepsilon_2^2 + 3 \varepsilon_1 \varepsilon_2}{12 \varepsilon_1 \varepsilon_2} \log\left(\frac{x}{\Lambda}\right) + \sum_{n=3}^{\infty} (n-3)! c_n (-x)^{2-n},$$

where  $c_n$  is defined by

$$\frac{1}{(1 - e^{-\varepsilon_1 t})(1 - e^{-\varepsilon_2 t})} = \sum_{n=0}^{\infty} c_n t^{n-2}.$$

If we consider the equivariant cohomology group  $H_{T^2}^*(\mathbb{C}^2)$  of  $\mathbb{C}^2$  with respect to the two-dimensional torus action, we have

$$c_n = \int_{\mathbb{C}^2} \text{Todd}_n(\mathbb{C}^2),$$

where  $\text{Todd}_n$  is the degree  $n$  part of the Todd genus, and  $\int_{\mathbb{C}^2}$  is defined by the localization formula applied to  $\mathbb{C}^2: \iota_0^*(\bullet)/\text{Eu}(T_0\mathbb{C}^2)$ . Here  $0$  is the unique fixed point and  $\iota_0$  is the inclusion  $\{0\} \rightarrow \mathbb{C}^2$ .

If  $\gamma_0(x; \Lambda) = -\frac{1}{2}x^2 \log(x/\Lambda) + \frac{3}{4}x^2$  denotes the leading part of  $\gamma_{\varepsilon_1, \varepsilon_2}(x; \Lambda)$  ('genus 0 part'), we have

$$\gamma_{\varepsilon_1, \varepsilon_2}(x; \Lambda) = \sum_{n=0}^{\infty} \int_{\mathbb{C}^2} \text{Todd}_n(\mathbb{C}^2) \gamma_0^{(n)}(x).$$

We introduce the function for the matter contribution as

$$\begin{aligned} \delta_{\varepsilon_1, \varepsilon_2}(x; \Lambda) &:= \gamma_{\varepsilon_1, \varepsilon_2}\left(x - \frac{\varepsilon_1 + \varepsilon_2}{2}; \Lambda\right) \\ &= \frac{1}{\varepsilon_1 \varepsilon_2} \left\{ -\frac{1}{2}x^2 \log\left(\frac{x}{\Lambda}\right) + \frac{3}{4}x^2 \right\} + \frac{\varepsilon_1^2 + \varepsilon_2^2}{24} \log\left(\frac{x}{\Lambda}\right) + \dots \end{aligned}$$

The shift  $-(\varepsilon_1 + \varepsilon_2)/2$  is identified with  $K_{\mathbb{C}^2}/2$ , and is compatible with our shift for the instanton partition function.

We define the full partition function as

$$\begin{aligned} Z(\varepsilon_1, \varepsilon_2, \vec{a}, \vec{m}; \Lambda) &:= \exp\left[-\sum_{\vec{\alpha} \in \Delta} \gamma_{\varepsilon_1, \varepsilon_2}(\langle \vec{a}, \vec{\alpha} \rangle; \Lambda) + \sum_{f, \alpha} \delta_{\varepsilon_1, \varepsilon_2}(a_\alpha + m_f; \Lambda)\right] \\ &\times Z^{\text{inst}}(\varepsilon_1, \varepsilon_2, \vec{a}, \vec{m}; \Lambda), \end{aligned}$$

where  $\Delta$  is the set of type  $A$  roots, i.e.,  $\langle \vec{a}, \vec{\alpha} \rangle = a_\alpha - a_\beta$  and the summation runs over all  $\alpha \neq \beta$ .

**Remark 5.2.** We do not make precise to which ring the full partition functions belong, as functions in  $\Lambda$ . We just use them formally to make the blow-up equation (5.4) shorter. The rigorous form of the blow-up equation is given by rewriting it as an equation for the instanton part. This was given in [26, Cor. 6.12], for the Nekrasov partition function without matters, and the version with matters is similar. But we do not give it here since it is not enlightening. This remark applies to all formulas below until they (more precisely their leading coefficients) are identified with those defined via Seiberg–Witten curves, which are really functions defined over an appropriate open set in  $\Lambda$ .

Let us expand the perturbative part as

$$\begin{aligned} &-\sum_{\vec{\alpha} \in \Delta} \gamma_{\varepsilon_1, \varepsilon_2}(\langle \vec{a}, \vec{\alpha} \rangle; \Lambda) + \sum_{f, \alpha} \delta_{\varepsilon_1, \varepsilon_2}(a_\alpha + m_f; \Lambda) \\ &= \frac{1}{\varepsilon_1 \varepsilon_2} \left( F_0^{\text{pert}} + (\varepsilon_1 + \varepsilon_2)H^{\text{pert}} + \varepsilon_1 \varepsilon_2 A^{\text{pert}} + \frac{\varepsilon_1^2 + \varepsilon_2^2}{3} B^{\text{pert}} + \dots \right) \end{aligned}$$

as in the instanton part.

For future reference, we give explicit formulas for some terms when  $r = 2$ ,  $N_f = 1$ :

$$\begin{aligned}
(5.3) \quad & H^{\text{pert}} = \pi\sqrt{-1}a, \\
& \frac{\partial F_0^{\text{pert}}}{\partial \log \Lambda} = -3a^2 + m^2, \\
& \frac{\partial^2 F_0^{\text{pert}}}{\partial (\log \Lambda)^2} = 0, \quad -\frac{1}{\gamma} \frac{\partial^2 F_0^{\text{pert}}}{\partial \log \Lambda \partial a} = 2a, \\
& \frac{\partial^2 F_0^{\text{pert}}}{\partial a^2} = 8 \log \frac{-2\sqrt{-1}a}{\Lambda} - \log \frac{(a+m)(-a+m)}{\Lambda^2}, \\
& \frac{\partial^2 F_0^{\text{pert}}}{\partial a \partial m} = \log \left( \frac{-a+m}{\Lambda} \right) - \log \left( \frac{a+m}{\Lambda} \right), \\
& \frac{\partial^2 F_0^{\text{pert}}}{\partial m^2} = -\log \left( \frac{-a+m}{\Lambda} \right) - \log \left( \frac{a+m}{\Lambda} \right), \\
& \frac{\partial^2 F_0^{\text{pert}}}{\partial m \partial \log \Lambda} = 2m, \\
& A^{\text{pert}} = \frac{1}{2} \log \left( \frac{-2\sqrt{-1}a}{\Lambda} \right), \\
& B^{\text{pert}} = \frac{1}{2} \log \left( \frac{-2\sqrt{-1}a}{\Lambda} \right) + \frac{1}{8} \log \left( \frac{(m-a)(m+a)}{\Lambda^2} \right).
\end{aligned}$$

### §5.3. Blow-up formula

Similarly we define the perturbative part to the correlation function on the blow-up as

$$\begin{aligned}
& \widehat{Z}_{c_1=kC}(\varepsilon_1, \varepsilon_2, \vec{a}, \vec{m}; t; \Lambda) \\
& := \exp \left[ - \sum_{\vec{a} \in \Delta} \gamma_{\varepsilon_1, \varepsilon_2}(\langle \vec{a}, \vec{a} \rangle; \Lambda) + \sum_{f, \alpha} \delta_{\varepsilon_1, \varepsilon_2}(a_\alpha + m_f; \Lambda) \right] \widehat{Z}_{c_1=kC}^{\text{inst}}(\varepsilon_1, \varepsilon_2, \vec{a}, \vec{m}; t; \Lambda).
\end{aligned}$$

As in [27, §4.4], we get

$$\begin{aligned}
(5.4) \quad & \widehat{Z}_{c_1=kC}(\varepsilon_1, \varepsilon_2, \vec{a}, \vec{m}; t; \Lambda) \\
& = \exp \left[ \frac{t}{\gamma} \left( \left( \frac{r}{12}(2r + N_f - 2) + \frac{N_f k^2}{2r} \right) (\varepsilon_1 + \varepsilon_2) + \left( \frac{r}{2} - k \right) \sum_f m_f \right) \right] \\
& \quad \times \sum_{\vec{k}} Z \left( \varepsilon_1, \varepsilon_2 - \varepsilon_1, \vec{a} + \varepsilon_1 \vec{k}, \vec{m} + \left( \frac{k}{r} - \frac{1}{2} \right) \varepsilon_1 \vec{e}; \Lambda e^{t\varepsilon_1/\gamma} \right) \\
& \quad \times Z \left( \varepsilon_1 - \varepsilon_2, \varepsilon_2, \vec{a} + \varepsilon_2 \vec{k}, \vec{m} + \left( \frac{k}{r} - \frac{1}{2} \right) \varepsilon_2 \vec{e}; \Lambda e^{t\varepsilon_2/\gamma} \right)
\end{aligned}$$

by analyzing the fixed points in  $\widehat{M}(r, k, n)$  and then using a difference equation satisfied by the perturbative term. Here  $\vec{e} = (1, \dots, 1) \in \mathbb{Z}^r$ , and  $\vec{k}$  runs over

$$\left\{ \vec{k} = (k_1, \dots, k_r) \in \mathbb{Q}^r : \sum k_\alpha = 0, k_\alpha \equiv -\frac{k}{r} \pmod{\mathbb{Z}} \right\}.$$

This is slightly different from the  $\vec{k}$  which appeared in the parametrization of the fixed point set  $\widehat{M}(r, k, n)^{\vec{T}}$ : we subtract  $k/r$  from each factor so that the sum of entries becomes 0.

The complete proof will be given in [32], but is a straightforward modification of the original one.

In [31, Th. 2.1] we proved the following vanishing theorem:

$$(5.5) \quad \frac{\widehat{Z}_{c_1=0}(\varepsilon_1, \varepsilon_2, \vec{a}, \vec{m}; t; \Lambda)}{Z(\varepsilon_1, \varepsilon_2, \vec{a}, \vec{m}; \Lambda)} = 1 + O(t^{\max(r+1, 2r-N_f)}).$$

This is a generalization of the vanishing theorem for the pure theory ( $N_f = 0$ ), which was proved by the dimension counting argument in [26]. The proof of this generalization requires the theory of perverse coherent sheaves in [29, 30, 31], but has a similar flavor to the original one. In particular, the exponent  $2r - N_f$ , which is written  $\gamma$  here, comes from the formula for

$$\deg(\text{Eu}(\mathcal{V} \otimes K_{\mathbb{C}^2}^{1/2} \otimes M) \cap [M(r, n)]) = (2r - N_f)n = \gamma n.$$

From (5.4) together with (5.5), we can prove

- (1)  $\varepsilon_1 \varepsilon_2 \log Z(\varepsilon_1, \varepsilon_2, \vec{a}, \vec{m}; \Lambda)$  is regular at  $\varepsilon_1, \varepsilon_2 = 0$ .
- (2) The instanton part satisfies  $Z^{\text{inst}}(\varepsilon_1, -2\varepsilon_1, \vec{a}, \vec{m}; \Lambda) = Z^{\text{inst}}(2\varepsilon_1, -\varepsilon_1, \vec{a}, \vec{m}; \Lambda)$ .

The proofs of these assertions are exactly as in [27, §5.2] and [26, Lem. 7.1] respectively. They will be reproduced in [32] for this version, and are not repeated here.

We expand the partition function as in (3.4):

$$\begin{aligned} &\varepsilon_1 \varepsilon_2 \log Z(\varepsilon_1, \varepsilon_2, \vec{a}, \vec{m}; \Lambda) \\ &= F_0(\vec{a}, \vec{m}; \Lambda) + (\varepsilon_1 + \varepsilon_2)H(\vec{a}, \vec{m}; \Lambda) + \varepsilon_1 \varepsilon_2 A(\vec{a}, \vec{m}; \Lambda) + \frac{\varepsilon_1^2 + \varepsilon_2^2}{3} B(\vec{a}, \vec{m}; \Lambda) + \dots \end{aligned}$$

From the symmetry property (2) of  $Z$ , we see that  $H$  comes only from the perturbative part. This is already explained above. As in [27, §5.3] (which has a sign mistake) we have

$$H(\vec{a}, \vec{m}; \Lambda) = -\pi \sqrt{-1} \langle \vec{a}, \rho \rangle,$$

where  $\rho$  is half the sum of the positive roots.

As in [27, §6] we can take the limit of (5.4) to get

$$\begin{aligned}
 (5.6) \quad & \lim_{\varepsilon_1, \varepsilon_2 \rightarrow 0} \frac{\widehat{Z}_{c_1=kC}(\varepsilon_1, \varepsilon_2, \vec{a}, \vec{m}; t; \Lambda)}{Z(\varepsilon_1, \varepsilon_2, \vec{a}, \vec{m}; \Lambda)} \\
 &= \exp \left[ -\frac{1}{2} \sum_{f, f'} \frac{\partial^2 F_0}{\partial m_f \partial m_{f'}} \left( \frac{k}{r} - \frac{1}{2} \right)^2 + A - B \right. \\
 &\quad \left. - \frac{t}{\gamma} \left\{ \sum_f \left( \frac{k}{r} - \frac{1}{2} \right) \left( \frac{\partial^2 F_0}{\partial \log \Lambda \partial m_f} - m_f r \right) \right\} - \frac{1}{\gamma^2} \frac{\partial^2 F_0}{\partial (\log \Lambda)^2} \frac{t^2}{2} \right] \\
 &\quad \times \Theta_{E_k} \left( -\frac{1}{2\pi\sqrt{-1}} \frac{\partial^2 F_0}{\partial \vec{a} \partial m_f} \left( \frac{k}{r} - \frac{1}{2} \right) - \frac{t}{\gamma} \frac{1}{2\pi\sqrt{-1}} \frac{\partial^2 F_0}{\partial \vec{a} \partial \log \Lambda} \mid \tau \right),
 \end{aligned}$$

where  $\Theta_{E_k}$  is the Riemann theta function with the characteristic  $E_k$  as in [27, §B]. The period matrix  $\tau$  is given by

$$\tau_{kl} = -\frac{1}{2\pi\sqrt{-1}} \frac{\partial^2 F_0}{\partial a^k \partial a^l}.$$

Here we change the coordinate from  $(a_2, \dots, a_r)$  to the root system coordinate defined as  $\vec{a} = \sum a^i \alpha_i^\vee$  by simple coroots  $\alpha_i^\vee = (0, \dots, 0, \overset{i}{1}, \overset{i+1}{-1}, 0, \dots, 0)$ ,  $i = 1, \dots, r$ .

In the  $r = 2$  case, we have  $a^1 = a_1 = -a_2 = -a$ . Therefore we need to note  $\partial/\partial \vec{a} = -\partial/\partial a$  when we use (5.6).

For a later purpose, we need another vanishing for  $c_1 \neq 0$ :

$$(5.7) \quad \widehat{Z}_{c_1=kC}(\varepsilon_1, \varepsilon_2, \vec{a}, \vec{m}; t; \Lambda) = O(t^{k(r-k)})$$

for  $0 < k < r$ . See [31, Th. 2.5]. This is again proved by a version of the dimension counting argument, and  $k(r - k)$  appears as the dimension of the Grassmannian of  $k$ -planes in  $\mathbb{C}^r$ .

**§5.4. Lower terms**

We assume  $r = 2$ ,  $N_f = 1$  hereafter. Therefore  $\gamma = 3$ .

Let us define a function  $u$  by

$$(5.8) \quad u := -\frac{1}{\gamma} \left( \frac{\partial F_0}{\partial \log \Lambda} - m^2 \right) = a^2 - \frac{1}{\gamma} \frac{\partial F_0^{\text{inst}}}{\partial \log \Lambda}.$$

In the formula of Theorem 4.5, this appears as the coefficient of  $x$ . Note that  $x$  is a variable for the  $\mu$ -class of the point. Its gauge-theoretic interpretation is already implicitly used in the proof of Theorem 4.5, but becomes clear if we look again

the partition function as follows: Consider

$$(5.9) \quad \frac{\sum_{n=0}^{\infty} \Lambda^{\gamma n} \iota_{0*}^{-1} \pi_* (\text{ch}_2(\mathcal{E})/[0] \cap \text{Eu}(\mathcal{V} \otimes K_{\mathbb{C}^2}^{1/2} \otimes M) \cap [M(2, n)])}{\sum_{n=0}^{\infty} \Lambda^{\gamma n} \iota_{0*}^{-1} \pi_* (\text{Eu}(\mathcal{V} \otimes K_{\mathbb{C}^2}^{1/2} \otimes M) \cap [M(2, n)])},$$

where  $[0]$  is the equivariant homology class of the origin. The denominator is nothing but  $Z^{\text{inst}}(\varepsilon_1, \varepsilon_2, a, m; \Lambda)$ , and we have  $\text{ch}_2(\mathcal{E})/[0] = a^2 - n\varepsilon_1\varepsilon_2$ . Therefore this is equal to

$$a^2 - \frac{\varepsilon_1\varepsilon_2}{\gamma} \frac{\partial}{\partial \log \Lambda} \log Z^{\text{inst}}(\varepsilon_1, \varepsilon_2, a, m; \Lambda).$$

From the expansion (3.4), this converges to (5.8) at  $\varepsilon_1, \varepsilon_2 = 0$ . In other words, the function  $u$  is the limit of (5.9) at  $\varepsilon_1, \varepsilon_2 = 0$ .

This can be generalized as follows. Let us replace  $\text{ch}_2(\mathcal{E})/[0]$  by its  $p^{\text{th}}$  power in (5.9). We have

$$Z^{\text{inst}}(\varepsilon_1, \varepsilon_2, a, m; \Lambda)^{-1} \sum_{k=0}^p \binom{p}{k} a^{2k} \left( -\frac{\varepsilon_1\varepsilon_2}{\gamma} \frac{\partial}{\partial \log \Lambda} \right)^{p-k} Z^{\text{inst}}(\varepsilon_1, \varepsilon_2, a, m; \Lambda).$$

If we set  $\mathcal{F} := \varepsilon_1\varepsilon_2 \log Z^{\text{inst}}$ , we also get higher derivatives of  $\mathcal{F}$ . But they always come with  $\varepsilon_1\varepsilon_2$  and disappear in the limit  $\varepsilon_1, \varepsilon_2 \rightarrow 0$ . Therefore we differentiate  $\mathcal{F}$  only once, and get

$$\sum_{k=0}^p \binom{p}{k} a^{2k} \left( -\frac{1}{\gamma} \frac{\partial F_0^{\text{inst}}}{\partial \log \Lambda} \right)^{p-k} = \left( a^2 - \frac{1}{\gamma} \frac{\partial F_0^{\text{inst}}}{\partial \log \Lambda} \right)^p$$

at  $\varepsilon_1, \varepsilon_2 = 0$ . This is equal to  $u^p$ .

In [31, Th. 2.6] a general structural result of the blow-up formula was proved. The integral

$$\iota_{0*}^{-1} \widehat{\pi}_* (e^{t\mu(C)} \cap \text{Eu}(\mathcal{V}_0 \otimes p^*(K_{\mathbb{C}^2}^{1/2}) \otimes M) \cap [\widehat{M}(2, k, n)]),$$

appearing in the correlation function on the blow-up, can be written as a linear combination of

$$\iota_{0*}^{-1} \widehat{\pi}_* ((\text{ch}_2(\mathcal{E})/[0])^p \cap \text{Eu}(\mathcal{V} \otimes K_{\mathbb{C}^2}^{1/2} \otimes M) \cap [M(2, n - k(2 - k)/4 - j)])$$

for various  $p, j \geq 0$ , where the coefficients are in  $\mathbb{C}[m, \varepsilon_1, \varepsilon_2][[t]]$ . (In higher rank cases, we also need higher Chern classes.) Moreover, the coefficients depend on  $p, j$  (and  $k$ ), but not on  $n$ . Therefore the ratio

$$\frac{\widehat{Z}_{c_1=C}(\varepsilon_1, \varepsilon_2, a, m; t; \Lambda)}{Z(\varepsilon_1, \varepsilon_2, a, m; \Lambda)}$$

is a formal power series in  $t$  with coefficients in  $\mathbb{C}[m, \varepsilon_1, \varepsilon_2, u, \Lambda]$ . Here the finiteness as power series in  $u, \Lambda$  comes from the cohomological degree reason.

In particular, when we expand the ratio in  $t$ , we only get finitely many powers of  $\Lambda$ , and the coefficients can be computed from the integrals over finitely many moduli spaces. By using the combinatorial expressions of the partition and correlation functions, these are really possible to compute. We use a Maple program to get

$$(5.10) \quad \lim_{\varepsilon_1, \varepsilon_2 \rightarrow 0} \frac{\widehat{Z}_{c_1=C}(\varepsilon_1, \varepsilon_2, a, m; t; \Lambda)}{Z(\varepsilon_1, \varepsilon_2, a, m; \Lambda)} \\ = -\Lambda t - \frac{t^3}{3!} \Lambda u - \frac{t^5}{5!} \Lambda(u^2 + 2m\Lambda^3) - \frac{t^7}{7!} \Lambda(u^3 + 6um\Lambda^3 + 6\Lambda^6) + O(t^9).$$

In fact, we have computed the ratio before taking  $\lim_{\varepsilon_1, \varepsilon_2 \rightarrow 0}$ , but imposing  $\varepsilon_1 + \varepsilon_2 = 0$  instead. Otherwise, the program runs very slow.

Let us check the cohomological degree, which we briefly mentioned above. We have  $\deg \Lambda = \deg m = 1$ ,  $\deg u = 2$ . Then the coefficient of  $t^n$  has degree  $n$ .

## §6. Seiberg–Witten curves

In this section we determine the coefficients  $F_0$ ,  $A$ ,  $B$  of  $Z$  in terms of certain ‘periods’ of a family of elliptic curves, called the *Seiberg–Witten curves*. Our derivation of the Seiberg–Witten curves is analogous to Fintushel–Stern’s method [11]: They described (in fact, before Seiberg–Witten’s work) that the blow-up formula for Donaldson invariants is given by elliptic integrals, associated with cubic curves of Weierstrass form. And the moduli parameter  $u$  for the cubics is coupled to the  $\mu$ -class of the point. We define  $u$ , and derive cubic curves in the same way by using the partition function  $Z$  instead of Donaldson invariants. The cubic curves are the Seiberg–Witten curves for the theory with one fundamental matter. In fact, our derivation is much simpler, as we already see the theta function in the blow-up formula.<sup>3</sup>

A similar discussion of the relation between the  $\sigma$ -function and the blow-up formula is given in [27, §6.3] for the partition function without matters. Some relevant formulas for elliptic functions are collected there. (See also [27, App. B] for notation.)

### §6.1. Elliptic curve

As before, we set

$$(6.1) \quad \tau := -\frac{1}{2\pi\sqrt{-1}} \frac{\partial^2 F_0}{\partial a^2}$$

<sup>3</sup>In higher rank cases, the story becomes much more complicated, as we need to show that the theta function is associated with a hyper-elliptic curve. See [32].

and consider the corresponding elliptic curve  $E_\tau$  with period  $\tau$ . We put

$$q = \exp(\pi\sqrt{-1}\tau) = \exp\left(-\frac{1}{2}\frac{\partial^2 F_0}{\partial a^2}\right).$$

Its expansion is given by

$$q = \frac{\Lambda^3(m^2 - a^2)^{1/2}}{(2a)^4}(1 + O(\Lambda^3))$$

from (5.3). Note that the instanton part is a formal power series in  $\Lambda^3$ .

We have defined  $u$  in (5.8). Since

$$u = a^2 + O(\Lambda^3),$$

we can take  $u$  as a variable instead of  $a^2$  since  $\Lambda$  is a formal variable. This viewpoint will be taken later since the curve  $E_\tau$  will be explicitly given as a cubic curve so that its coefficients are polynomials in  $u$ . This  $u$  is the coordinate of what Seiberg–Witten called the  $u$ -plane, a family of vacuum states.

We realize the elliptic curve  $E_\tau$  as  $\mathbb{C}/(\mathbb{Z}\omega + \mathbb{Z}\omega') = \mathbb{C}/(\mathbb{Z}\omega + \mathbb{Z}\omega\tau)$ , where

$$(6.2) \quad \omega := -2\pi\sqrt{-1}\left(\frac{\partial u}{\partial a}\right)^{-1} = \left(\frac{1}{2\pi\sqrt{-1}}\frac{1}{\gamma}\frac{\partial^2 F_0}{\partial a\partial \log \Lambda}\right)^{-1} = -\frac{\pi\sqrt{-1}}{a} + O(\Lambda^3).$$

Using the Weierstrass  $\wp$ -function associated with  $\mathbb{Z}\omega + \mathbb{Z}\omega'$ , we can realize  $E_\tau$  in the Weierstrass form:

$$y^2 = 4x^3 - g_2x - g_3.$$

Since  $g_2, g_3$  are given by (see, e.g., [2, §13.20])

$$g_2 = \frac{2}{3}\left(\frac{\pi}{\omega}\right)^4 (\theta_{00}^4 + \theta_{01}^4 + \theta_{10}^4), \quad g_3 = \frac{4}{27}\left(\frac{\pi}{\omega}\right)^6 (\theta_{00}^4 + \theta_{10}^4)(\theta_{00}^4 + \theta_{01}^4)(\theta_{01}^4 - \theta_{10}^4),$$

$E_\tau$  is defined over the formal power series ring in  $\Lambda$ .

Then the blow-up formula for the  $c_1 = C$  case (5.6) can be rewritten in terms of the  $\sigma$ -function:

$$\begin{aligned} \lim_{\varepsilon_1, \varepsilon_2 \rightarrow 0} \frac{\widehat{Z}_{c_1=C}(\varepsilon_1, \varepsilon_2, a, m; t; \Lambda)}{Z(\varepsilon_1, \varepsilon_2, a, m; \Lambda)} \\ = -\exp\left[A - B - t^2\left\{\frac{1}{2\gamma^2}\frac{\partial^2 F_0}{\partial \log \Lambda^2} + \frac{\pi^2}{6\omega^2}E_2(\tau)\right\}\right]\sigma(t)\frac{\theta'_{11}(0)}{\omega}. \end{aligned}$$

We compare the expansion

$$e^{-Tt^2}\sigma(t) = t - Tt^3 + \left(\frac{T^2}{2} - \frac{g_2}{2 \cdot 5!}\right)t^5 + \left(-\frac{T^3}{3!} + \frac{Tg_2}{2 \cdot 5!} - \frac{6g_3}{7!}\right)t^7 + \dots$$

with our computation of lower terms of the blow-up formula (5.10). We get

$$(6.3) \quad \exp(A - B) \frac{\theta'_{11}(0)}{\omega} = \Lambda,$$

$$(6.4) \quad \frac{1}{2\gamma^2} \frac{\partial^2 F_0}{\partial \log \Lambda^2} + \frac{\pi^2}{6\omega^2} E_2(\tau) = -\frac{u}{6},$$

$$(6.5) \quad g_2 = \frac{4}{3}u^2 - 4m\Lambda^3,$$

$$(6.6) \quad g_3 = -\frac{8}{27}u^3 + \frac{4}{3}um\Lambda^3 - \Lambda^6.$$

In particular, the curve  $E_\tau$  has the Weierstrass form

$$y^2 = 4x^3 - \left(\frac{4}{3}u^2 - 4m\Lambda^3\right)x + \frac{8}{27}u^3 - \frac{4}{3}um\Lambda^3 + \Lambda^6.$$

Replacing  $x$  by  $x + u/3$ , we get

$$(6.7) \quad y^2 = 4x^2(x + u) + 4m\Lambda^3x + \Lambda^6.$$

This is none other than the Seiberg–Witten curve for the theory with one fundamental matter, first determined in [38]. There is a vast literature on this curve. For example, [1] was useful for the authors.

The discriminant  $\Delta = g_2^3 - 27g_3^2$  is given by

$$(6.8) \quad \Delta = -\Lambda^6(16u^3 - 16u^2m^2 - 72um\Lambda^3 + 64m^3\Lambda^3 + 27\Lambda^6).$$

Let  $e_1 - u/3, e_2 - u/3, e_3 - u/3$  be the solutions of the right hand side of (6.7) = 0. We number them as in [2, p. 361]:

$$(6.9) \quad \begin{aligned} e_1 &= \frac{1}{3} \left(\frac{\pi}{\omega}\right)^2 (\theta_{00}^4 + \theta_{01}^4), & e_2 &= \frac{1}{3} \left(\frac{\pi}{\omega}\right)^2 (\theta_{10}^4 - \theta_{01}^4), \\ e_3 &= -\frac{1}{3} \left(\frac{\pi}{\omega}\right)^2 (\theta_{10}^4 + \theta_{00}^4). \end{aligned}$$

Note that  $\omega$  in [2] is our  $\omega/2$ .

The blow-up formula is further simplified as

$$\lim_{\varepsilon_1, \varepsilon_2 \rightarrow 0} \frac{\widehat{Z}_{c_1=C}(\varepsilon_1, \varepsilon_2, a, m; t; \Lambda)}{Z(\varepsilon_1, \varepsilon_2, a, m; \Lambda)} = -e^{ut^2/6} \sigma(t) \Lambda$$

(cf. [27, §6.3]). This is the form of Fintushel–Stern’s blow-up formula for the Donaldson invariants if we replace the curve appropriately, i.e., the Seiberg–Witten curve for the *pure* theory.

**§6.2. Seiberg–Witten differential**

In this subsection, we revert the roles of  $u$  and  $a$ . We consider  $u$  as a variable and introduce the cubic curve (6.7). Then we write  $a$  as the integral of a certain differential form  $dS$  on the Seiberg–Witten curve. The prepotential  $F_0$  is defined in a similar way. It is the usual framework to relate the Seiberg–Witten curve and the partition function. This is not necessary for our computation of derivatives of  $F_0$ , but we explain it for completeness.

Let  $Q(x)$  be the right hand side of (6.7). We set

$$dS := -\frac{1}{2\pi\sqrt{-1}} \frac{Q'(x)dx}{4xy}.$$

Then we define

$$(6.10) \quad a = \int_A dS.$$

Let us explain a motivation of this definition. We differentiate (6.7) to get

$$2ydy = Q'(x)dx.$$

Therefore

$$dS = -\frac{1}{2\pi\sqrt{-1}} \frac{dy}{2x}.$$

We differentiate (6.7) by  $u$  after setting  $y$  to be constant:

$$0 = Q'(x) \frac{\partial x}{\partial u} \Big|_{y=\text{const}} + 4x^2.$$

Hence

$$-2\pi\sqrt{-1} \frac{\partial}{\partial u} dS \Big|_{y=\text{const}} = -\frac{dy}{2x^2} \frac{\partial x}{\partial u} \Big|_{y=\text{const}} = \frac{2dy}{Q'(x)} = \frac{dx}{y}.$$

This is compatible with (6.2). And we have (6.10) up to a constant independent of  $u$ .

Note that  $dS$  has a pole at  $x = 0$ . We have  $y = \pm\Lambda^3$ , hence the residue is

$$\text{Res}_{x=0, y=\pm\Lambda^3} dS = \frac{\mp m}{2\pi\sqrt{-1}}.$$

Therefore we need to specify the  $A$ -cycle in (6.10), otherwise the residue is well-defined only up to  $\mathbb{Z} \frac{m}{2\pi\sqrt{-1}}$ . This is possible by studying the perturbative part of the integral, but we leave the details to [32].

We assume that the perturbative part has a correct behavior and sketch how to prove that the above  $a$ , denoted by  $a^{\text{SW}}$ , coincides with the original  $a$ . We consider both  $a$  and  $a^{\text{SW}}$  as functions in  $u, m, \Lambda$ . The latter is a convergent power

series in  $\Lambda$  in a bounded region of  $u, m$  avoiding the discriminant of the Seiberg–Witten curves. We consider it as a formal power series and compare with  $a$ . Since  $da/du = da^{\text{SW}}/du$ ,  $a - a^{\text{SW}}$  is a function of  $m$  and  $\Lambda$  only. On the other hand, both  $a$  and  $a^{\text{SW}}$  are homogeneous of degree 1 if we put  $\deg u = 2, \deg m = \deg \Lambda = 1$ . For  $a$ , this is a cohomological degree. And for  $a^{\text{SW}}$ , it can be shown as in [27, §2.2]. Since  $a$  and  $a^{\text{SW}}$  have the same perturbative part (by our assumption),  $a - a^{\text{SW}}$  is a power series in  $\Lambda^3$ , so it must be of the form  $\sum_{n=1}^{\infty} a_n m^{1-3n} \Lambda^{3n}$  with constants  $a_n$ , for degree reasons. But as both  $a$  and  $a^{\text{SW}}$  are regular at  $m = 0$ , we must have  $a_n = 0$  for all  $n$ , i.e.,  $a = a^{\text{SW}}$ .

The remaining part of the derivation of  $F_0$  from the Seiberg–Witten curves can be discussed as usual. Let us sketch it again. We define the dual variable  $a^D$  as the integral of  $dS$  over the  $B$ -cycle. We consider  $a$  as a variable, and  $u$  as a function in  $a$ . Then we define  $F_0$  by

$$a^D = -\frac{1}{2\pi\sqrt{-1}} \frac{\partial F_0}{\partial a}.$$

This is compatible with (6.1) and determine  $F_0$  up to a constant independent of  $a$ . We fix the ambiguity so that  $F_0$  satisfies (5.8). ( $F_0$  satisfies (5.8) up to a constant as in [27, §2.4].)

**§6.3. Genus 1 part**

We next determine the coefficients  $A$  and  $B$ . This was done in [27, §7.1] for the pure theory. We use the same method.

Consider the blow-up formula (5.4) for  $c_1 = C$  and take the coefficient of  $t^0 \cdot (\varepsilon_1 + \varepsilon_2)$ . By (5.7) it is zero. As in [27] we get

$$\frac{\partial}{\partial a} \left( A - \frac{1}{3} B \right) = -\frac{1}{3} \frac{\partial}{\partial a} \log \theta'_{11}(0).$$

Therefore we have

$$\exp \left( A - \frac{1}{3} B \right) = C \theta'_{11}(0)^{-1/3}$$

for some constant  $C$  independent of  $a$ . Together with (6.3) we get

$$\begin{aligned} \exp A &= (C^3 \Lambda^{-1} \omega^{-1})^{1/2}, \\ \exp B &= (C \Lambda^{-1} \omega^{-1})^{3/2} \theta'_{11}(0) = C^{3/2} (2\pi)^{-1/2} \Lambda^{-3/2} \Delta^{1/8}, \end{aligned}$$

where  $\Delta = 16(\pi/\omega)^{12} (\theta'_{11}(0)/\pi)^8$  is the discriminant.

The perturbative part of  $\exp A$  is

$$\left( \frac{-2\sqrt{-1}a}{\Lambda} \right)^{1/2}.$$

On the other hand,  $\omega^{-1/2} = (-2\pi\sqrt{-1})^{-1/2}(\partial u/\partial a)^{1/2}$  has

$$(-2\pi\sqrt{-1})^{-1/2}\sqrt{2a}.$$

Therefore

$$\exp A \left( \frac{-\sqrt{-1} \partial u}{\Lambda \partial a} \right)^{-1/2}$$

has perturbative part 1. On the other hand, from the discussion above, this is a constant independent of  $a$ . From degree considerations as in [27], it is a homogeneous element. However the instanton part is a formal power series in  $\Lambda/a$  and  $m/a$ . Therefore it must be 1. Hence

$$(6.11) \quad \exp A = \left( \frac{-\sqrt{-1} \partial u}{\Lambda \partial a} \right)^{1/2}, \quad \exp B = \sqrt{-1} \Lambda^{-3/2} \Delta^{1/8}.$$

**§6.4. Derivatives of  $F_0$**

We will redo the computation in this subsection at the point  $a = m$  again later, so the reader can safely skip to the next section. But we just want to point out that the derivatives of  $F_0$  can be computed before specializing  $a = m$ .

Let us rewrite the blow-up formula (5.6) for  $c_1 = 0$  in terms of the  $\sigma$ -function:

$$(6.12) \quad \lim_{\varepsilon_1, \varepsilon_2 \rightarrow 0} \frac{\widehat{Z}_{c_1=0}(\varepsilon_1, \varepsilon_2, a, m; t; \Lambda)}{Z(\varepsilon_1, \varepsilon_2, a, m; \Lambda)} \\ = \theta_{01}(0) \exp \left[ -\frac{1}{8} \frac{\partial^2 F_0}{\partial m^2} + A - B - \frac{\eta}{\omega} \left( \frac{\omega}{4\pi\sqrt{-1}} \frac{\partial^2 F_0}{\partial a \partial m} \right)^2 \right. \\ \left. + t \left\{ \frac{1}{\gamma} \left( -m + \frac{1}{2} \frac{\partial^2 F_0}{\partial \log \Lambda \partial m} \right) + \frac{\eta}{2\pi\sqrt{-1}} \frac{\partial^2 F_0}{\partial a \partial m} \right\} + \frac{ut^2}{6} \right] \\ \times \sigma_3 \left( t - \frac{\omega}{4\pi\sqrt{-1}} \frac{\partial^2 F_0}{\partial a \partial m} \right),$$

where  $\eta = \zeta(\omega/2) = \pi^2 E_2(\tau)/(6\omega)$ . Taking the coefficients of  $t^0, t^1, t^2$  and comparing with (5.5), we get

$$(6.13) \quad \theta_{01}(0) \exp \left[ -\frac{1}{8} \frac{\partial^2 F_0}{\partial m^2} + A - B - \eta \omega \left( \frac{1}{4\pi\sqrt{-1}} \frac{\partial^2 F_0}{\partial a \partial m} \right)^2 \right] \sigma_3 \left( -\frac{\omega}{4\pi\sqrt{-1}} \frac{\partial^2 F_0}{\partial a \partial m} \right) = 1,$$

$$(6.14) \quad \frac{1}{\gamma} \left( m + \frac{1}{2} \frac{\partial^2 F_0}{\partial \log \Lambda \partial m} \right) + \frac{\eta}{2\pi\sqrt{-1}} \frac{\partial^2 F_0}{\partial a \partial m} + \frac{d}{dt}(\log \sigma_3) \left( -\frac{\omega}{4\pi\sqrt{-1}} \frac{\partial^2 F_0}{\partial a \partial m} \right) = 0,$$

$$(6.15) \quad \frac{u}{3} + \frac{d^2}{dt^2}(\log \sigma_3) \left( -\frac{\omega}{4\pi\sqrt{-1}} \frac{\partial^2 F_0}{\partial a \partial m} \right) = 0.$$

Since the second derivative of  $\log \sigma$  is  $(-1)$  times the Weierstrass  $\wp$ -function, we have

$$(6.16) \quad \frac{u}{3} = -\frac{d^2}{dt^2}(\log \sigma_3) \left( -\frac{\omega}{4\pi\sqrt{-1}} \frac{\partial^2 F_0}{\partial a \partial m} \right) = \wp \left( \frac{\omega\tau}{2} - \frac{\omega}{4\pi\sqrt{-1}} \frac{\partial^2 F_0}{\partial a \partial m} \right)$$

from the last equation. Therefore

$$-\frac{\omega}{4\pi\sqrt{-1}} \frac{\partial^2 F_0}{\partial a \partial m} = \int_{\infty}^0 \frac{dx}{y} - \frac{\omega_3}{2} = \int_{e_3-u/3}^0 \frac{dx}{y},$$

where  $y$  is as in (6.7) and  $u/3$  is replaced by 0 since the quadratic term of (6.7) is  $4u$ . Note that this 0 is the point where  $dS$  has a pole.

**§7. Partition functions at the singular point**

Recall that we need to specialize  $a = m$  in Theorem 4.5. At this point, the Seiberg–Witten curve is singular, and many formulas are simplified.

**§7.1. The special point  $a = m$**

Recall that the period  $\tau$  of the Seiberg–Witten curve was given by the second derivative of  $F_0$  with respect to  $a$  (see (6.1)). Its perturbative part is given by (5.3). In particular,  $q = \exp(\pi\sqrt{-1}\tau)$  vanishes at  $a = m$  since it contains a factor  $-a + m$ . Therefore  $\theta_{00} \rightarrow 1$ ,  $\theta_{01} \rightarrow 1$ ,  $\theta_{10} \rightarrow 0$  at  $a = m$ , and hence we have  $e_2 = e_3$  from (6.9). The cycle encircling  $e_2, e_3$  vanishes and the curve develops singularities.

The blow-up formula (6.12) is *not* suitable for the specialization  $e_2 = e_3$ , as it contains an expression  $\partial^2 F_0 / \partial a \partial m$ , which has  $\log(-a + m) / \Lambda$  in the perturbative part. We observe that

$$\frac{\omega_3}{2} - \frac{\omega}{4\pi\sqrt{-1}} \frac{\partial^2 F_0}{\partial a \partial m} = -\frac{\omega}{4\pi\sqrt{-1}} \left( \frac{\partial^2 F_0}{\partial a^2} + \frac{\partial^2 F_0}{\partial a \partial m} \right)$$

does not contain the term  $\log(-a + m) / \Lambda$  in the perturbative part. Hence we can evaluate this term at  $a = m$ . Therefore we use  $\sigma$  instead of  $\sigma_3$  in (6.12):

$$\begin{aligned} & \lim_{\varepsilon_1, \varepsilon_2 \rightarrow 0} \frac{\widehat{Z}_{e_1=0}(\varepsilon_1, \varepsilon_2, \vec{a}, \vec{m}; t; \Lambda)}{Z(\varepsilon_1, \varepsilon_2, \vec{a}, \vec{m}; \Lambda)} \\ &= \sqrt{-1} \Lambda \exp \left[ -\frac{1}{8} \left( \frac{\partial^2 F_0}{\partial m^2} + 2 \frac{\partial^2 F_0}{\partial a \partial m} + \frac{\partial^2 F_0}{\partial a^2} \right) - \eta\omega \left\{ \frac{1}{4\pi\sqrt{-1}} \left( \frac{\partial^2 F_0}{\partial a \partial m} + \frac{\partial^2 F_0}{\partial a^2} \right) \right\}^2 \right. \\ & \quad \left. + \frac{t}{\gamma} \left\{ \frac{1}{2} \left( \frac{\partial^2 F_0}{\partial \log \Lambda \partial m} - 2m \right) \right\} + \pi\sqrt{-1} \frac{t}{\omega} + \frac{t\eta}{2\pi\sqrt{-1}} \left( \frac{\partial^2 F_0}{\partial a \partial m} + \frac{\partial^2 F_0}{\partial a^2} \right) + \frac{t^2 u}{6} \right] \\ & \quad \times \sigma \left( t - \frac{\omega}{4\pi\sqrt{-1}} \left( \frac{\partial^2 F_0}{\partial a \partial m} + \frac{\partial^2 F_0}{\partial a^2} \right) \right). \end{aligned}$$

Here we have used

$$\begin{aligned} \theta_{01}(z) &= \exp\left(\frac{\pi\sqrt{-1}}{4}\tau + \pi\sqrt{-1}\left(z + \frac{1}{2}\right)\right)\theta_{11}\left(z + \frac{1}{2}\tau\right), \\ \sigma(t) &= \omega \exp\left(\frac{\eta}{\omega}t^2\right)\frac{\theta_{11}(t/\omega)}{\theta'_{11}(0)} = \omega \exp\left(\frac{\pi^2}{6\omega^2}E_2(\tau)t^2\right)\frac{\theta_{11}(t/\omega)}{\theta'_{11}(0)}. \end{aligned}$$

Now we can specialize  $e_2 = e_3$  (see, e.g., [2, §13.15] for relevant formulas): the  $\sigma$ -function becomes

$$\sigma(t) = \frac{\omega}{\pi} \sin\left(\frac{\pi}{\omega}t\right) \exp\left[\frac{1}{6}\left(\frac{\pi}{\omega}\right)^2 t^2\right].$$

We also note that

$$\eta\omega = \pi^2/6$$

at  $e_2 = e_3$ . Therefore

$$\begin{aligned} \lim_{\varepsilon_1, \varepsilon_2 \rightarrow 0} \frac{\widehat{Z}_{c_1=0}(\varepsilon_1, \varepsilon_2, \vec{a}, \vec{m}; t; \Lambda)}{Z(\varepsilon_1, \varepsilon_2, \vec{a}, \vec{m}; \Lambda)} &= \frac{\sqrt{-1}\omega\Lambda}{\pi} \exp\left[-\frac{1}{8}\left(\frac{\partial^2 F_0}{\partial m^2} + 2\frac{\partial^2 F_0}{\partial a\partial m} + \frac{\partial^2 F_0}{\partial a^2}\right)\right. \\ &\quad \left. + \frac{t}{\gamma}\left\{\frac{1}{2}\left(\frac{\partial^2 F_0}{\partial \log \Lambda \partial m} - 2m\right)\right\} + \pi\sqrt{-1}\frac{t}{\omega} + \frac{t^2}{6}\left(u + \left(\frac{\pi}{\omega}\right)^2\right)\right] \\ &\quad \times \sin\left(\frac{\pi}{\omega}t - \frac{1}{4\sqrt{-1}}\left(\frac{\partial^2 F_0}{\partial a\partial m} + \frac{\partial^2 F_0}{\partial a^2}\right)\right). \end{aligned}$$

As before, we take the coefficients of  $t^0, t^1, t^2$ , compare with (5.5) and get

$$(7.1) \quad 1 = \frac{\sqrt{-1}\omega\Lambda}{\pi} \exp\left[-\frac{1}{8}\left(\frac{\partial^2 F_0}{\partial m^2} + 2\frac{\partial^2 F_0}{\partial a\partial m} + \frac{\partial^2 F_0}{\partial a^2}\right)\right] \times \sin\left(-\frac{1}{4\sqrt{-1}}\left(\frac{\partial^2 F_0}{\partial a\partial m} + \frac{\partial^2 F_0}{\partial a^2}\right)\right),$$

$$(7.2) \quad 0 = \frac{1}{\gamma}\left\{\frac{1}{2}\left(\frac{\partial^2 F_0}{\partial \log \Lambda \partial m} - 2m\right)\right\} + \frac{\pi\sqrt{-1}}{\omega} + \frac{\pi}{\omega} \cot\left(-\frac{1}{4\sqrt{-1}}\left(\frac{\partial^2 F_0}{\partial a\partial m} + \frac{\partial^2 F_0}{\partial a^2}\right)\right),$$

$$(7.3) \quad 0 = \frac{1}{3}\left(u + \left(\frac{\pi}{\omega}\right)^2\right) - \left(\frac{\pi}{\omega}\right)^2 \sin^{-2}\left(-\frac{1}{4\sqrt{-1}}\left(\frac{\partial^2 F_0}{\partial a\partial m} + \frac{\partial^2 F_0}{\partial a^2}\right)\right).$$

**§7.2. Miscellaneous identities**

We assume  $a = m$  hereafter, and solve equations (7.1, 7.2, 7.3) to write down various derivatives of  $F_0$  explicitly.

Since  $e_1 - u/3 = (2/3)(\pi/\omega)^2 - u/3$ ,  $e_2 - u/3 = e_3 - u/3 = -(1/3)(\pi/\omega)^2 - u/3$  is a solution of  $0 = 4x^2(x + u) + 4a\Lambda^3x + \Lambda^6$ , we have

$$4\left(x + \frac{u}{3} + \frac{1}{3}\left(\frac{\pi}{\omega}\right)^2\right)^2\left(x + \frac{u}{3} - \frac{2}{3}\left(\frac{\pi}{\omega}\right)^2\right) = 4x^2(x + u) + 4a\Lambda^3x + \Lambda^6.$$

Thus

$$(7.4) \quad \left(u + \left(\frac{\pi}{\omega}\right)^2\right)^2\left(u - 2\left(\frac{\pi}{\omega}\right)^2\right) = \frac{27}{4}\Lambda^6,$$

$$(7.5) \quad \left(u + \left(\frac{\pi}{\omega}\right)^2\right)\left(u - \left(\frac{\pi}{\omega}\right)^2\right) = 3a\Lambda^3.$$

This suggests the possibility to replace  $a$  by  $u - (\pi/\omega)^2$  or  $u + (\pi/\omega)^2$ . Therefore we write various functions in terms of  $u$  and  $\pi/\omega$  instead of  $a$ . In fact, we will find that it is even more natural to introduce a function  $T$  given by

$$T := \frac{1}{3}\left(u + \left(\frac{\pi}{\omega}\right)^2\right) = \frac{1}{3}\left(u - \frac{1}{4}\left(\frac{\partial u}{\partial a}\right)^2\right).$$

Up to a constant multiple, this is the *contact term* for surfaces in the physics literature, say in [22, 19]. It will give the contribution of the intersection number  $(\alpha^2)$  in Donaldson invariants in view of our formula in Theorem 4.5, thanks to (7.7) proved just below.

The perturbative parts of  $u$  and  $(1/4)(\partial u/\partial a)^2$  cancel out, so the perturbative part of  $T$  is 0. We calculate its leading term from the combinatorial expression of (3.3) for the ‘1-instanton part’  $M(2, n = 1)$  to find

$$(7.6) \quad T = \frac{1}{2a}\Lambda^3 + O(\Lambda^6).$$

By (6.4) together with  $E_2(\tau) = 1$  when  $a = m$ , we have

$$(7.7) \quad \frac{\partial^2 F_0}{\partial(\log \Lambda)^2} = -3\left(u + \left(\frac{\pi}{\omega}\right)^2\right) = -9T.$$

Since  $\Delta$  vanishes at  $a = m$ , we have  $\frac{\partial \Delta}{\partial a} + \frac{\partial \Delta}{\partial m} = 0$ . Therefore we get

$$0 = \left(3u^2 - 2a^2u - \frac{9}{2}a\Lambda^3\right)\left(\frac{\partial u}{\partial a} + \frac{\partial u}{\partial m}\right) - 2u^2a - \frac{9}{2}u\Lambda^3 + 12a^2\Lambda^3$$

from (6.8). Using (7.4, 7.5), we find

$$(7.8) \quad \begin{aligned} -2u^2a - \frac{9}{2}u\Lambda^3 + 12a^2\Lambda^3 &= \frac{4}{\Lambda^3}\left(\frac{\pi}{\omega}\right)^6 T, \\ 3u^2 - 2a^2u - \frac{9}{2}a\Lambda^3 &= -\frac{4}{\Lambda^6}\left(\frac{\pi}{\omega}\right)^6 T^2. \end{aligned}$$

For example, the upper equality is checked as

$$\begin{aligned} -2u^2a - \frac{9}{2}u\Lambda^3 + 12a^2\Lambda^3 &= \frac{1}{3\Lambda^3} \left[ -2u^2 \left( u^2 - \left( \frac{\pi}{\omega} \right)^4 \right) + 4 \left( u^2 - \left( \frac{\pi}{\omega} \right)^4 \right)^2 - \frac{27}{2}u\Lambda^6 \right] \\ &= \frac{4}{3\Lambda^3} \left( \frac{\pi}{\omega} \right)^6 \left( u + \left( \frac{\pi}{\omega} \right)^2 \right), \end{aligned}$$

where the first equality is from (7.5) and the second from (7.4). The lower equality can be proved in a similar way.

Therefore

$$(7.9) \quad -\frac{1}{\gamma} \left( \frac{\partial^2 F_0}{\partial \log \Lambda \partial m} - 2m \right) - \frac{2\pi\sqrt{-1}}{\omega} = \frac{\partial u}{\partial a} + \frac{\partial u}{\partial m} = \Lambda^3 T^{-1}.$$

Plugging (7.9) to (7.2), we obtain

$$\frac{\pi}{\omega} \cot \left( -\frac{1}{4\sqrt{-1}} \left( \frac{\partial^2 F_0}{\partial a \partial m} + \frac{\partial^2 F_0}{\partial a^2} \right) \right) = \frac{1}{2} \Lambda^3 T^{-1}.$$

The left hand side is

$$\frac{\pi\sqrt{-1}}{\omega} \frac{\exp \left[ -\frac{1}{2} \left( \frac{\partial^2 F_0}{\partial a \partial m} + \frac{\partial^2 F_0}{\partial a^2} \right) \right] + 1}{\exp \left[ -\frac{1}{2} \left( \frac{\partial^2 F_0}{\partial a \partial m} + \frac{\partial^2 F_0}{\partial a^2} \right) \right] - 1}.$$

Hence

$$(7.10) \quad \begin{aligned} &\exp \left[ -\frac{1}{2} \left( \frac{\partial^2 F_0}{\partial a \partial m} + \frac{\partial^2 F_0}{\partial a^2} \right) \right] \\ &= - \left( \frac{2\pi\sqrt{-1}}{\omega} + \Lambda^3 T^{-1} \right) / \left( \frac{2\pi\sqrt{-1}}{\omega} - \Lambda^3 T^{-1} \right) = \frac{1}{4} T^{-1} \left( \frac{2\pi\sqrt{-1}}{\omega} + \Lambda^3 T^{-1} \right)^2, \end{aligned}$$

where we have used (7.4) in the last equality.

By (7.1) and (7.3) we have

$$(7.11) \quad \exp \left[ -\frac{1}{4} \left( \frac{\partial^2 F_0}{\partial m^2} + 2 \frac{\partial^2 F_0}{\partial a \partial m} + \frac{\partial^2 F_0}{\partial a^2} \right) \right] = -\frac{1}{\Lambda^2} T.$$

By (7.10) and (7.11) we obtain

$$(7.12) \quad \exp \left[ -\frac{1}{2} \left( \frac{\partial^2 F_0}{\partial m^2} + \frac{\partial^2 F_0}{\partial a \partial m} \right) \right] = \frac{4T^3}{\Lambda^4} \left( \frac{2\pi\sqrt{-1}}{\omega} + \Lambda^3 T^{-1} \right)^{-2}.$$

### §7.3. Computation of instanton parts

Since we will express Mochizuki's formula in terms of *instanton parts* of derivatives of  $F_0$  and  $A, B$ , we need to compute them. Since their perturbative parts are explicit functions, we just subtract them from the full partition functions. We denote instanton parts by putting 'inst' as sub/superscripts.

We have

$$(7.13) \quad \frac{1}{\gamma} \frac{\partial F_0^{\text{inst}}}{\partial \log \Lambda} = \frac{1}{\gamma} \left( \frac{\partial F_0}{\partial \log \Lambda} + 2a^2 \right) = -u + a^2$$

from the perturbative part of  $\partial F_0 / \log \Lambda$  and the definition of  $u$  in (5.8).

Since  $\exp[-(1/4)(\partial^2 F_0 / \partial m^2 + 2\partial^2 F_0 / \partial a \partial m + \partial^2 F_0 / \partial a^2)]$  has  $-(2a/\Lambda)^{-1}$  as the perturbative part, we get

$$(7.14) \quad \exp \left[ -\frac{1}{4} \left( \frac{\partial^2 F_0^{\text{inst}}}{\partial m^2} + 2 \frac{\partial^2 F_0^{\text{inst}}}{\partial a \partial m} + \frac{\partial^2 F_0^{\text{inst}}}{\partial a^2} \right) \right] = \frac{2a}{\Lambda^3} T$$

from (7.11). Note that the left hand side starts with 1 as a formal power series in  $\Lambda$ . This is compatible with the expansion of the right hand side in (7.6).

In the same way, we get

$$(7.15) \quad \exp \left[ -\frac{1}{2} \left( \frac{\partial^2 F_0^{\text{inst}}}{\partial a \partial m} + \frac{\partial^2 F_0^{\text{inst}}}{\partial a^2} \right) \right] = \frac{1}{4} \left( \frac{2a}{\Lambda} \right)^3 T^{-1} \left( \frac{2\pi\sqrt{-1}}{\omega} + \Lambda^3 T^{-1} \right)^2$$

from (7.10), and

$$(7.16) \quad \exp \left[ -\frac{1}{2} \left( \frac{\partial^2 F_0^{\text{inst}}}{\partial m^2} + \frac{\partial^2 F_0^{\text{inst}}}{\partial a \partial m} \right) \right] = \frac{2T^3}{\Lambda^3 a} \left( \frac{2\pi\sqrt{-1}}{\omega} + \Lambda^3 T^{-1} \right)^{-2}$$

from (7.12).

We need a trick to compute the instanton part of  $q = \exp(-(1/2)\partial^2 F_0 / \partial a^2)$ , since it vanishes at  $a = m$ . For the moment we no longer set  $a = m$  and consider

$$\begin{aligned} q_{\text{inst}}^2 &= \exp \left( -\frac{\partial^2 F_0^{\text{inst}}}{\partial a^2} \right) = q^2 \left( \frac{-2\sqrt{-1}a}{\Lambda} \right)^8 \left( \frac{(m+a)(m-a)}{\Lambda^2} \right)^{-1} \\ &= \frac{q^2}{m-a} \left( \frac{-2\sqrt{-1}a}{\Lambda} \right)^8 \left( \frac{m+a}{\Lambda^2} \right)^{-1}. \end{aligned}$$

Since  $q$  vanishes at  $a = m$ , we get

$$q_{\text{inst}}^2 \Big|_{m=a} = -\Lambda \frac{\partial(q^2)}{\partial a} \Big|_{m=a} \left( \frac{2a}{\Lambda} \right)^7.$$

The discriminant  $\Delta$  has an expansion  $\omega^{12}\Delta = (2\pi)^{12}(q^2 - 24q^4 + \dots)$ , so

$$\omega^{12} \frac{\partial \Delta}{\partial a} \Big|_{m=a} = \frac{\partial}{\partial a}(\omega^{12}\Delta) \Big|_{m=a} = (2\pi)^{12} \frac{\partial(q^2)}{\partial a} \Big|_{m=a}.$$

We differentiate (6.8) with respect to  $a$  to get

$$(7.17) \quad \frac{\partial \Delta}{\partial a} \Big|_{m=a} = -16\Lambda^6 \left( 3u^2 - 2a^2u - \frac{9}{2}a\Lambda^3 \right) \frac{\partial u}{\partial a} = -\sqrt{-1} \left( \frac{2\pi}{\omega} \right)^7 T^2,$$

where we have used (7.8). Therefore

$$(7.18) \quad g_{\text{inst}}^2 = \exp\left(-\frac{\partial^2 F_0^{\text{inst}}}{\partial a^2}\right) = \Lambda\sqrt{-1} \left(\frac{2\pi}{\omega}\right)^{-5} \left(\frac{2a}{\Lambda}\right)^7 T^2.$$

It is to be understood that all functions are evaluated at  $a = m$  unless an equation contains an expression ‘ $m - a$ ’.

Substituting (7.18) into (7.15) we get

$$\exp\left(-\frac{\partial^2 F_0^{\text{inst}}}{\partial a \partial m}\right) = \frac{1}{\sqrt{-1}\Lambda} T^{-4} \left(\frac{2a}{\Lambda}\right)^{-1} \left(\frac{2\pi}{\omega}\right)^5 \left(\frac{\pi\sqrt{-1}}{\omega} + \frac{\Lambda^3}{2T}\right)^4.$$

Then we substitute this into (7.16) to get

$$(7.19) \quad \exp\left(-\frac{\partial^2 F_0^{\text{inst}}}{\partial m^2}\right) = \frac{\sqrt{-1}}{\Lambda^7} T^{10} \left(\frac{2a}{\Lambda}\right)^{-1} \left(\frac{2\pi}{\omega}\right)^{-5} \left(\frac{\pi\sqrt{-1}}{\omega} + \frac{\Lambda^3}{2T}\right)^{-8}.$$

Let us consider instanton parts of other derivatives: we have

$$(7.20) \quad \frac{\partial^2 F_0^{\text{inst}}}{\partial a \partial \log \Lambda} + \frac{\partial^2 F_0^{\text{inst}}}{\partial m \partial \log \Lambda} = 6a - 3\Lambda^3 T^{-1}$$

from the definition (5.8) of  $u$  and (7.9). We also have

$$(7.21) \quad \frac{\partial^2 F_0^{\text{inst}}}{\partial m \partial \log \Lambda} = -3 \left( \Lambda^3 T^{-1} + \frac{2\pi\sqrt{-1}}{\omega} \right),$$

from the definition (6.2) of  $\omega$  and its perturbative part.

From (6.11) we have

$$(7.22) \quad \exp A^{\text{inst}} = \left( \frac{1}{2a} \frac{\partial u}{\partial a} \right)^{1/2} = \left( \frac{1}{\sqrt{-1} a \omega} \pi \right)^{1/2}.$$

In order to compute the instanton part of  $B$ , we use the same technique as for  $q$ , since  $B$  also vanishes at  $a = m$ . The perturbative part of  $B$  is

$(1/2) \log(-2\sqrt{-1}a/\Lambda) + (1/8) \log((m-a)(m+a)/\Lambda^2)$ . Thus we have

$$\exp B^{\text{inst}} = \sqrt{-1} \Lambda^{-3/2} \left(\frac{\Delta}{m-a}\right)^{1/8} \left(\frac{-2\sqrt{-1}a}{\Lambda}\right)^{-1/2} \left(\frac{m+a}{\Lambda^2}\right)^{-1/8}$$

from (6.11). Therefore at  $m = a$ , we have

$$\exp 8B^{\text{inst}} = \Lambda^{-11} \left(-\frac{\partial\Delta}{\partial a}\right) \left(\frac{2a}{\Lambda}\right)^{-5}.$$

Using (7.17), we get

$$(7.23) \quad \exp 8B^{\text{inst}} = \sqrt{-1} \Lambda^{-11} \left(\frac{2\pi}{\omega}\right)^7 \left(\frac{2a}{\Lambda}\right)^{-5} T^2.$$

**§7.4. The variable  $\phi$**

In the partition function, we need to substitute  $\Lambda^{4/3}a^{-1/3}$  into  $\Lambda$ . We denote the substitution by  $\bullet|_{\Lambda=\Lambda^{4/3}a^{-1/3}}$ .

Let  $\mathfrak{X} := T|_{\Lambda=\Lambda^{4/3}a^{-1/3}}$ . By (7.6) it has the expansion  $\Lambda^4/(2a^2) + \dots$ . So we can choose the branch of its square root so that it starts as  $\sqrt{\mathfrak{X}} = \Lambda^2/(\sqrt{2}a) + \dots$ . We set

$$(7.24) \quad \phi := \sqrt{\mathfrak{X}}/\Lambda.$$

From (7.4,7.5) we have

$$\begin{aligned} & \left(u|_{\Lambda=\Lambda^{4/3}a^{-1/3}} + \left(\frac{\pi}{\omega}\right)^2 \Big|_{\Lambda=\Lambda^{4/3}a^{-1/3}}\right)^2 \left(u|_{\Lambda=\Lambda^{4/3}a^{-1/3}} - 2\left(\frac{\pi}{\omega}\right)^2 \Big|_{\Lambda=\Lambda^{4/3}a^{-1/3}}\right) \\ & \hspace{20em} = \frac{27}{4} \Lambda^8 a^{-2}, \\ & \left(u|_{\Lambda=\Lambda^{4/3}a^{-1/3}} + \left(\frac{\pi}{\omega}\right)^2 \Big|_{\Lambda=\Lambda^{4/3}a^{-1/3}}\right) \left(u|_{\Lambda=\Lambda^{4/3}a^{-1/3}} - \left(\frac{\pi}{\omega}\right)^2 \Big|_{\Lambda=\Lambda^{4/3}a^{-1/3}}\right) \\ & \hspace{20em} = 3\Lambda^4. \end{aligned}$$

From the second equation and the definition of  $\phi$ , we get

$$(7.25) \quad \begin{aligned} \frac{1}{\Lambda^2}(u)|_{\Lambda=\Lambda^{4/3}a^{-1/3}} &= \frac{1}{2}(3\phi^2 + \phi^{-2}), \\ \frac{1}{\Lambda^2} \left(\left(\frac{\pi}{\omega}\right)^2\right) \Big|_{\Lambda=\Lambda^{4/3}a^{-1/3}} &= \frac{1}{2}(3\phi^2 - \phi^{-2}). \end{aligned}$$

Substituting this to the first equation, we obtain

$$(7.26) \quad \frac{1}{4} \Lambda^2 a^{-2} = \phi^4 \left(-\frac{1}{2}\phi^2 + \frac{1}{2\phi^2}\right) = \frac{1}{2} \phi^2 (-\phi^4 + 1).$$

Therefore

$$(7.27) \quad \frac{da}{a} = -\frac{d\phi}{\phi(1-\phi^4)}(1-3\phi^4).$$

Taking square roots in (7.25, 7.26), we get

$$(7.28) \quad \frac{2\pi\sqrt{-1}}{\omega\Lambda} = -\sqrt{2}\phi^{-1}\sqrt{1-3\phi^4}, \quad \frac{\Lambda}{a} = \sqrt{2}\phi\sqrt{1-\phi^4}.$$

By the above formulas, all the terms computed in §7.3 can be expressed in terms of  $\phi$  only. Hence we will treat  $\phi$  as a variable instead of  $a$ . We will write the differential  $\mathcal{B}(\xi_1, \xi; a)da$ , of which we take the residue in Mochizuki's formula, in terms of  $\phi$ . The explicit formula will be given in the next section, but it is already clear that it will involve several square roots and rational expressions in  $\phi$ , when we expand it as a series in  $x$  and  $z$ . We will see that square roots, in fact, do not appear, so we get a rational differential in  $\phi$  defined over  $\mathbb{P}^1$ .

We will use the residue theorem to rewrite Mochizuki's formula as the sum of the residues at other poles in the next section. But it is instructive to see the meaning of poles at this stage.

Since  $\phi = \Lambda/(\sqrt{2}a) + \dots$ , we have  $\phi = 0$  at  $a = \infty$ . By (7.26) there are other points  $\phi^4 = 1$  giving  $a = \infty$ . By (7.27) they are indeed poles of the differential. From (7.25) we have

$$u^2|_{\Lambda=\Lambda^{4/3}a^{-1/3}} = 4\Lambda^4.$$

As  $u$  is coupled with the variable  $x$  for the  $\mu$ -class of the point in the formula in Theorem 4.5, these correspond to the KM-simple type condition in Definition 2.2. In [40] it was noted that the Seiberg–Witten curve (for the pure theory) has singularities at those points, and they give the Seiberg–Witten invariant contribution to Donaldson invariants. Therefore even before the actual calculation, it is natural to expect that the residues at  $\phi^4 = 1$  give what is expected in Witten's conjecture (1.1).

There are other poles, which is already seen in (7.28), at  $\phi^4 = 1/3$ . At those points  $\pi/\omega = (\sqrt{-1}/2)\partial u/\partial a$  vanishes. This means that the Seiberg–Witten curve completely degenerates as we have  $e_1 = e_2 = e_3$ . It is called a *superconformal point* in the physics literature, and is the origin of the superconformal simple type condition [20]. Therefore it is natural to expect that the residue at  $\phi^4 = 1/3$  is related to the superconformal simple type condition. We will see that this is indeed so.

§8. Computation

§8.1. Explicit expression of the differential

Substituting all terms computed in the previous section into the formula in Theorem 4.5, we obtain

$$\begin{aligned}
 (8.1) \quad \mathcal{B}(\tilde{\xi}_1, \xi; a) da &= -(-1)^{((\xi, \xi + K_X) - (K_X^2) - (K_X, \tilde{\xi}_1))/2 + \chi_h(X)} \frac{1 - 3\phi^4}{1 - \phi^4} \frac{d\phi}{\phi} \\
 &\times \exp \left[ -\frac{\Lambda^2}{2} (3\phi^2 + \phi^{-2})x - \frac{1}{2} \phi^2 \Lambda^2 (\alpha^2) z^2 \right] \phi^{-((\xi - K_X)^2) - (K_X)^2 - 3\chi_h(X)} \\
 &\times \left( \frac{1}{\sqrt{2}} \phi^{-2} (\sqrt{1 - \phi^4} - \sqrt{1 - 3\phi^4}) \right)^{(\xi - K_X, \tilde{\xi}_1)} \\
 &\times \exp \left( \frac{\Lambda}{\sqrt{2}} \phi^{-1} (\sqrt{1 - 3\phi^4}(\tilde{\xi}_1, \alpha)z - \sqrt{1 - \phi^4}(\xi - K_X, \alpha)z) \right) \\
 &\times (\sqrt{2} \sqrt{1 - 3\phi^4})^{(K_X^2) - \chi_h(X)}.
 \end{aligned}$$

This is a simple substitution except that we need to take square roots or 8<sup>th</sup> roots for some expressions. For example, the term with  $((\xi - K_X)^2)$  is

$$2^{-1/2} \left( \frac{2a}{\Lambda} \right) \exp \left[ \frac{1}{8} \left( \frac{\partial^2 F_0^{\text{inst}}}{\partial m^2} + 2 \frac{\partial^2 F_0^{\text{inst}}}{\partial a \partial m} + \frac{\partial^2 F_0^{\text{inst}}}{\partial a^2} \right) \right] \Big|_{\Lambda = \Lambda^{4/3} a^{-1/3}}.$$

From (7.14) the square of this is equal to  $\Lambda^2/\mathfrak{I}$ . Since the leading term of the above is  $\sqrt{2}a/\Lambda$ , we find that it is equal to  $\Lambda/\sqrt{\mathfrak{I}} = \phi^{-1}$  from our choice of  $\sqrt{\mathfrak{I}}$ . We use the same argument for other expressions involving square roots.

When we expand  $\tilde{\mathcal{A}}(\xi_1, y; a)$  into a formal power series in  $z, x$  as  $\sum_{k,l} \mathcal{A}_{k,l} z^k x^l$ , we will be interested in the case  $k + 2l = 4n - (\xi^2) - 3\chi_h(X) = \dim M_H(y)$ , otherwise the residue at  $\phi = 0$  vanishes for cohomology degree reasons. Note also that  $\dim M_H(y) \equiv -(\xi^2) - 3\chi_h(X) \pmod{4}$  is independent of  $n$ . Therefore we decompose  $\mathcal{B}(\tilde{\xi}_1, \xi; a)$  as

$$\mathcal{B}(\tilde{\xi}_1, \xi; a) = \mathcal{B}^{(0)}(\tilde{\xi}_1, \xi; a) + \mathcal{B}^{(1)}(\tilde{\xi}_1, \xi; a) + \mathcal{B}^{(2)}(\tilde{\xi}_1, \xi; a) + \mathcal{B}^{(3)}(\tilde{\xi}_1, \xi; a)$$

according to  $(k + 2l) \pmod{4}$ . If we write variables  $(x, z)$ , those are given explicitly as

$$\mathcal{B}^{(p)}(\tilde{\xi}_1, \xi; a)(x, z) = \frac{1}{4} \sum_{q=0}^3 (\sqrt{-1})^{-qp} \mathcal{B}(\tilde{\xi}_1, \xi; a) ((-1)^q x, (\sqrt{-1})^q z).$$

We will be concerned with  $\mathcal{B}^{(\dim M_H(y))}(\tilde{\xi}_1, \xi; a)$ , where we understand  $\dim M_H(y)$  modulo 4 as explained above.

We will be interested in the sum over all Seiberg–Witten classes  $\tilde{\xi}_1$ . Therefore we can combine the contribution for  $\tilde{\xi}_1$  and  $-\tilde{\xi}_1$ , using  $\text{SW}(-\tilde{\xi}_1) =$

$(-1)^{\chi_h(X)} \text{SW}(\tilde{\xi}_1)$ . Hence we will be interested in

$$(8.2) \quad \mathcal{B}^{(\dim M_H(y))}(\tilde{\xi}_1, \xi; a) + (-1)^{\chi_h(X)} \mathcal{B}^{(\dim M_H(y))}(-\tilde{\xi}_1, \xi; a).$$

**Proposition 8.3.** (1) *The sum  $\mathcal{B}^{(p)}(\tilde{\xi}_1, \xi; a)da + (-1)^{\chi_h(X)} \mathcal{B}^{(p)}(-\tilde{\xi}_1, \xi; a)da$  is unchanged under the sign change of  $\sqrt{1 - 3\phi^4}$ .*

(2) *Suppose that  $p \equiv \dim M_H(y) \pmod 2$ . Then  $\mathcal{B}^{(p)}(\tilde{\xi}_1, \xi; a)da$  is unchanged under the simultaneous sign change of  $\sqrt{1 - \phi^4}$  and  $\sqrt{1 - 3\phi^4}$ .*

*In particular, if  $p \equiv \dim M_H(y) \pmod 2$ , the combination  $\mathcal{B}^{(p)}(\tilde{\xi}_1, \xi; a)da + (-1)^{\chi_h(X)} \mathcal{B}^{(p)}(-\tilde{\xi}_1, \xi; a)da$  contains even powers of  $\sqrt{1 - 3\phi^4}$  and  $\sqrt{1 - \phi^4}$ , and hence is a rational 1-form in  $\phi$ .*

(3) *The sum  $\mathcal{B}^{(\dim M_H(y))}(\tilde{\xi}_1, \xi; a)da + (-1)^{\chi_h(X)} \mathcal{B}^{(\dim M_H(y))}(-\tilde{\xi}_1, \xi; a)da$  is a rational 1-form in  $\phi^4$ .*

*Proof.* (1) Looking at (8.1), we see that replacing  $\sqrt{1 - 3\phi^4}$  by  $-\sqrt{1 - 3\phi^4}$  has the same effect as replacing  $\tilde{\xi}_1$  by  $-\tilde{\xi}_1$  together with multiplication by  $(-1)^{\chi_h(X)}$ , as

$$\frac{1}{\sqrt{2}\phi^2}(\sqrt{1 - \phi^4} + \sqrt{1 - 3\phi^4}) = \left\{ \frac{1}{\sqrt{2}\phi^2}(\sqrt{1 - \phi^4} - \sqrt{1 - 3\phi^4}) \right\}^{-1}$$

and

$$(-1)^{(K_X, K_X - \tilde{\xi}_1)/2} = (-1)^{(K_X, K_X + \tilde{\xi}_1)/2} (-1)^{(K_X^2)}.$$

Therefore the sum  $\mathcal{B}^{(p)}(\tilde{\xi}_1, \xi; a)da + (-1)^{\chi_h(X)} \mathcal{B}^{(p)}(-\tilde{\xi}_1, \xi; a)da$  is unchanged.

(2) Looking at (8.1), we find that the replacement of  $\sqrt{1 - \phi^4}$ ,  $\sqrt{1 - 3\phi^4}$  by  $-\sqrt{1 - \phi^4}$ ,  $-\sqrt{1 - 3\phi^4}$  has the same effect as the replacement of  $(x, z)$  by  $(x, -z)$  together with multiplication by  $(-1)^{(\xi - K_X, \tilde{\xi}_1) + (K_X^2) - \chi_h(X)}$ . From the definition, the first replacement gives multiplication by  $(-1)^p$ . Now the assertion follows from the following:

$$(8.4) \quad \begin{aligned} (\xi - K_X, \tilde{\xi}_1) + \dim M_H(y) &\equiv (\xi - K_X, \tilde{\xi}_1) + (\xi^2) + \chi_h(X) \\ &\equiv (\xi - K_X, K_X) + (\xi, K_X) + \chi_h(X) \equiv (K_X^2) - \chi_h(X) \pmod 2. \end{aligned}$$

For a later purpose we need a refinement:

$$(8.5) \quad \begin{aligned} (\xi - K_X, \tilde{\xi}_1) + (K_X^2) + 3\chi_h(X) \\ &\equiv (\xi - K_X, \tilde{\xi}_1) + (K_X^2) - (\xi^2) - \dim M_H(y) \pmod 4 \\ &= (\xi - K_X, \tilde{\xi}_1 - K_X) + (\xi, K_X - \xi) - \dim M_H(y). \end{aligned}$$

(3) Looking at (8.1) again, we find that the replacement of  $\phi$  by  $\sqrt{-1}\phi$  has the same effect as the replacement  $(x, z)$  by  $(-x, -\sqrt{-1}z)$  together with multiplication by

$$(-1)^{(\xi - K_X, \tilde{\xi}_1)} (\sqrt{-1})^{-((\xi - K_X)^2) - (K_X)^2 - 3\chi_h(X)}.$$

The first replacement gives multiplication by  $(\sqrt{-1})^{-\dim M_H(y)}$ . Therefore the assertion follows from

$$\begin{aligned}
 & -\dim M_H(y) - 2(\xi - K_X, \tilde{\xi}_1) - \{((\xi - K_X)^2) + (K_X^2) + 3\chi_h(X)\} \\
 & \equiv (\xi^2) + 3\chi_h(X) - 2(\xi - K_X, K_X) - \{((\xi - K_X)^2) + (K_X^2) + 3\chi_h(X)\} \\
 & \equiv 0 \pmod{4}. \quad \square
 \end{aligned}$$

From the form of  $\mathcal{B}(\tilde{\xi}_1, \xi; a)da$  in (8.1), we find that the differential (8.2) has poles possibly only at  $\phi^4 = 0, \infty, 1$  and  $1/3$ . Mochizuki’s formula is given by the residue at  $\phi^4 = 0$ . The power of  $\phi$  containing  $-(\xi - K_X)^2$  is *very* negative since  $\xi$  is sufficiently ample when we apply Mochizuki’s formula to compute Donaldson invariants. Hence it is not so easy to compute the residue at  $\phi^4 = 0$  directly. Therefore we use the residue theorem

$$\left( \operatorname{Res}_{\phi^4=0} + \operatorname{Res}_{\phi^4=\infty} + \operatorname{Res}_{\phi^4=1} + \operatorname{Res}_{\phi^4=1/3} \right) [\text{the differential (8.2)}] = 0,$$

to compute the residues at  $\infty, 1, 1/3$  instead.

**Remark 8.6.** The differential in (8.1) was originally only defined in the formal neighborhood of  $\phi = 0$ , corresponding to  $a = \infty$ . Its extension to  $\mathbb{P}^1$  is obvious from its explicit form, but is quite mysterious if we remember that it is originally given by the partition function. We do not have a geometric explanation, and a similar mystery appears also in the study of mirror symmetry, where the analytic continuation is far from obvious in the *A*-model, but is clear in the *B*-model.

**§8.2. Residue at  $\phi = \infty$**

We first treat the simplest (possible) pole  $\phi = \infty$ . Recall that we expand the 1-form  $\mathcal{B}(\tilde{\xi}_1, \xi; a)da$  as a formal power series in  $x, z$  and take the coefficients of  $x^k z^l$  with  $k + 2l = 4n - (\xi^2) - 3\chi_h(X) = \dim M_H(y)$ . Let us denote this part as  $\mathcal{B}^{[\dim M_H(y)]}(\tilde{\xi}_1, \xi; a)da$ . The residue at  $\phi^4 = 0$  is the same as that of  $\tilde{\mathcal{A}}(\xi_1, y; a)$  for cohomological degree reasons, but it is not equal to  $\tilde{\mathcal{A}}(\xi_1, y; a)$  itself as we still take the sum over all  $n$ . Recall that when we use Mochizuki’s formula in Theorem 4.1, we expand  $\mathcal{B}^{(\dim M_H(y))}(\tilde{\xi}_1, \xi; a)da$  in  $x, z$ , compute the residue at  $\phi^4 = 0$ , and then take the sum over  $y$ . Thus we actually need to compute the residue of  $\mathcal{B}^{[\dim M_H(y)]}(\tilde{\xi}_1, \xi; a)da$ .

**Proposition 8.7.**  $\mathcal{B}^{[\dim M_H(y)]}(\tilde{\xi}_1, \xi; a)da + (-1)^{\chi_h(X)} \mathcal{B}^{[\dim M_H(y)]}(-\tilde{\xi}_1, \xi; a)da$  is regular at  $\phi^4 = \infty$  if  $\chi(y) > 0$ .

*Proof.* Terms appearing in (8.1) have the following order of vanishing at  $\phi = \infty$ :

$$\begin{aligned} \text{Order}_{\phi=\infty}(\phi) &= -1, & \text{Order}_{\phi=\infty}\left(\frac{d\phi}{\phi}\right) &= -1, & \text{Order}_{\phi=\infty}(3\phi^2 + \phi^{-2})^k &= -2k, \\ \text{Order}_{\phi=\infty}(\phi^{-1}\sqrt{1-\phi^4})^l &= -l, & \text{Order}_{\phi=\infty}(\phi^{-1}\sqrt{1-3\phi^4})^l &= -l, \\ \text{Order}_{\phi=\infty}(\sqrt{1-3\phi^4}) &= -2 \end{aligned}$$

Therefore  $\mathcal{B}^{[\dim M_H(y)]}(\tilde{\xi}_1, \xi) + (-1)^{\chi_h(X)} \mathcal{B}^{[\dim M_H(y)]}(-\tilde{\xi}_1, \xi)$  has a zero of order at least

$$\begin{aligned} -1 + [((\xi - K_X)^2) + (K_X^2) + 3\chi(\mathcal{O}_X)] - \dim M_H(y) - 2(K_X^2) + 2\chi(\mathcal{O}_X) \\ = (\xi, \xi - 2K_X) + 5\chi(\mathcal{O}_X) - \dim M_H(y) - 1. \end{aligned}$$

This is equal to  $4\chi(y) - 1$ . The assertion follows. □

**§8.3. Residue at  $\phi^4 = 1$**

Next we study the residue at  $\phi^4 = 1$ . We will show that it is identified with Witten’s formula.

By (8.1), the residue of  $\mathcal{B}(\tilde{\xi}_1, \xi; a)da$  at  $\phi = 1$  is given by

$$\begin{aligned} -\frac{1}{2}(-1)^{((\xi, \xi + K_X) - (K_X^2) - (K_X, \tilde{\xi}_1))/2 + \chi_h(X)} e^{-2\Lambda^2 x - \frac{1}{2}\Lambda^2(\alpha^2)z^2} (\sqrt{2}\sqrt{-2})^{(K_X^2) - \chi_h(X)} \\ \times \left(\frac{1}{\sqrt{2}}(-\sqrt{-2})\right)^{(\xi - K_X, \tilde{\xi}_1)} \exp\left(\frac{\Lambda}{\sqrt{2}}\sqrt{-2}(\tilde{\xi}_1, \alpha)z\right) \\ = -(-1)^{((\xi, \xi + K_X) - (K_X^2) - (K_X, \tilde{\xi}_1))/2 + (K_X^2)} 2^{(K_X^2) - \chi_h(X) - 1} \\ \times \exp\left[-2\Lambda^2 x - \frac{1}{2}\Lambda^2(\alpha^2)z^2\right] (\sqrt{-1})^{-\{(\xi - K_X, \tilde{\xi}_1) + (K_X^2) - \chi_h(X)\}} \exp(\Lambda\sqrt{-1}(\tilde{\xi}_1, \alpha)z). \end{aligned}$$

By (8.5),

$$(\xi - K_X, \tilde{\xi}_1) + (K_X^2) - \chi_h(X) \equiv (\xi - K_X, \tilde{\xi}_1 - K_X) + (\xi, K_X - \xi) - \dim M_H(y) \pmod{4}.$$

We combine the first two terms, which are even, with the factor coming from  $((K_X^2) + (K_X, \tilde{\xi}_1))/2$ :

$$\begin{aligned} -\frac{(K_X, K_X + \tilde{\xi}_1)}{2} + (\xi - K_X, \frac{\tilde{\xi}_1 - K_X}{2}) + \frac{(\xi, K_X - \xi)}{2} \\ = -\frac{(\xi, \xi - \tilde{\xi}_1)}{2} - (K_X, \tilde{\xi}_1) \equiv \frac{(\xi, \xi - \tilde{\xi}_1)}{2} - (K_X^2) \pmod{2}. \end{aligned}$$

Hence we get

$$\begin{aligned} & \operatorname{Res}_{\phi=1} \mathcal{B}(\tilde{\xi}_1, \xi; a) da \\ &= -(-1)^{((\xi, \xi + K_X) + (\xi, \xi - \tilde{\xi}_1))/2} 2^{(K_X^2) - \chi_h(X) - 1} \exp[-2\Lambda^2 x - \frac{1}{2}\Lambda^2(\alpha^2)z^2] \\ & \quad \times (\sqrt{-1})^{\dim M_H(y)} \exp(\Lambda\sqrt{-1}(\tilde{\xi}_1, \alpha)z) \end{aligned}$$

and

$$\begin{aligned} & \frac{1}{2} \operatorname{Res}_{\phi=1} [\mathcal{B}^{(\dim M_H(y))}(\tilde{\xi}_1, \xi; a) da + (-1)^{\chi_h(X)} \mathcal{B}^{(\dim M_H(y))}(-\tilde{\xi}_1, \xi; a) da] \\ &= -(-1)^{((\xi, \xi + K_X) + (\xi, \xi - \tilde{\xi}_1))/2} 2^{(K_X^2) - \chi_h(X) - 3} \\ & \quad \times [e^{-2\Lambda^2 x - \frac{1}{2}\Lambda^2(\alpha^2)z^2} \{(\sqrt{-1})^{\dim M_H(y)} e^{\Lambda\sqrt{-1}(\tilde{\xi}_1, \alpha)z} \\ & \quad \quad \quad + (\sqrt{-1})^{-\dim M_H(y)} e^{-\Lambda\sqrt{-1}(\tilde{\xi}_1, \alpha)z}\} \\ & \quad + e^{2\Lambda^2 x + \frac{1}{2}\Lambda^2(\alpha^2)z^2} \{e^{-\Lambda(\tilde{\xi}_1, \alpha)z} + (-1)^{-\dim M_H(y)} e^{\Lambda(\tilde{\xi}_1, \alpha)z}\}], \end{aligned}$$

where we have used  $(\xi, \tilde{\xi}_1) + \chi_h(X) \equiv \dim M_H(y) \pmod{2}$  (cf. (8.4)). The residues at  $\phi = \sqrt{-1}, -1, -\sqrt{-1}$  are the same as above by Proposition 8.3(3). Thus we multiply the above by 4 for the contribution from  $\phi^4 = 1$ .

This contribution satisfies the KM-simple type condition, i.e., it is killed by  $(\partial/\partial x)^2 - 4\Lambda^4$ . If we consider the contribution to the Donaldson series  $\mathcal{D}^\xi$ , we get

$$-(-1)^{\frac{(\xi, \xi + K_X)}{2}} 2^{(K_X^2) - \chi_h(X) + 1} e^{2\Lambda^2 x + \frac{1}{2}\Lambda^2(\alpha^2)z^2} \sum_{\tilde{\xi}_1} \operatorname{SW}(\tilde{\xi}_1) (-1)^{\frac{(\xi, \xi - \tilde{\xi}_1)}{2}} e^{-\Lambda(\tilde{\xi}_1, \alpha)z}.$$

Replacing  $\tilde{\xi}_1$  by  $-\tilde{\xi}_1$ , removing the sign factor  $(-1)^{(\xi, \xi + K_X)/2}$  as in §2.2 and multiplying with the 2 from Mochizuki’s convention, we get the right hand side of (1.1) with the opposite sign. Therefore, if the residue at  $\phi^4 = 1/3$  vanishes, we obtain (1.1).

**§8.4. Residue at  $\phi^4 = 1/3$**

**Proposition 8.8.** *Suppose that  $X$  is of superconformal simple type. Then*

$$\sum_{\tilde{\xi}_1} \operatorname{SW}(\tilde{\xi}_1) \mathcal{B}^{(\dim M_H(y))}(\tilde{\xi}_1, \xi; a) da$$

*is regular at  $\phi^4 = 1/3$ .*

*Proof.* Let

$$\begin{aligned} f(\lambda) := & \sum_{\tilde{\xi}_1} (-1)^{((K_X, K_X + \tilde{\xi}_1))/2} \operatorname{SW}(\tilde{\xi}_1) \lambda^{(\xi - K_X, \tilde{\xi}_1)} \\ & \times \{(-\lambda + \lambda^{-1})(\tilde{\xi}_1, \alpha) - (\lambda + \lambda^{-1})(\xi - K_X, \alpha)\}^k, \end{aligned}$$

where we assume  $k$  has the same parity as  $\dim M_H(y)$ . By (8.4), we have

$$f(\lambda) = (-1)^{\chi_h(X) - (K_X^2)} f(-\lambda).$$

By the superconformal simple type condition, we have

$$f^{(n)}(1) = 0 \quad \text{for } n = 0, \dots, \chi_h(X) - (K_X^2) - 3.$$

Therefore  $f(\lambda) \in (\lambda - 1)^{\chi_h(X) - (K_X^2) - 2} \mathbb{C}[\lambda^\pm]$ . Since  $f(-\lambda)$  is equal to  $f(\lambda)$  up to sign, we also have  $f(\lambda) \in (\lambda + 1)^{\chi_h(X) - (K_X^2) - 2} \mathbb{C}[\lambda^\pm]$ . Therefore

$$f(\lambda) \in (\lambda - \lambda^{-1})^{\chi_h(X) - (K_X^2) - 2} \mathbb{C}[\lambda^\pm].$$

From this we have the assertion by substituting  $(1/\sqrt{2})\phi^{-2}(\sqrt{1 - \phi^4} - \sqrt{1 - 3\phi^4})$  to  $\lambda$ . □

Next we study the converse direction. We fix  $\xi^\circ$  and consider

$$\xi = K_X + t(\xi^\circ - K_X)$$

with  $t \in 2\mathbb{Z}_{\geq 0} + 1$  as a function in  $t$ . Suppose  $\xi^\circ - K_X$  is ample.

**Proposition 8.9.** *Assume that for any given  $M_1$  there exists  $M_2 > 0$  such that the coefficients of  $x^l z^k$  ( $2l + k \leq M_1$ ) in*

$$\text{Res}_{\phi^4=1/3} \left( \sum_{\xi_1} \text{SW}(\tilde{\xi}_1) \mathcal{B}^{(\dim M_H(y))}(\tilde{\xi}_1, K_X + t(\xi^\circ - K_X); a) da \right)$$

*are assumed to be independent of any integer  $t > M_2$  up to sign. Then  $X$  is of superconformal simple type.*

The assumption above is satisfied thanks to Theorem 4.1: For a given  $M_1$ , we only need to consider finitely many  $n$  for  $\tilde{\mathcal{A}}(\xi_1, (2, \xi, n); a)$ . Then the assumption  $\chi(y) = \chi((2, \xi, n)) > 0$  is satisfied for  $t > M_2$ . So Theorem 4.1 is applicable. But the left hand side of (4.2) is independent of  $t$ . Therefore  $X$  is of superconformal simple type. Then the residue at  $\phi^4 = 1/3$  vanishes by the previous proposition, and the sum of the residues at  $\phi^4 = 0$  and  $\phi^4 = 1$  is zero. This proves Witten’s conjecture (1.1).

Before starting the proof of Proposition 8.9, we give some preparation.

Replacing  $\tilde{\xi}_1$  by  $-\tilde{\xi}_1$  if necessary, we may assume  $(\xi^\circ - K_X, \tilde{\xi}_1) \geq 0$ . We expand (8.1) by using the binomial theorem:

$$\begin{aligned}
 & \frac{1}{2}(\mathcal{B}^{\dim M_H(y)}(\tilde{\xi}_1, \xi; a)da + (-1)^{\chi_h(X)}\mathcal{B}^{\dim M_H(y)}(-\tilde{\xi}_1, \xi; a)da) \\
 &= -(-1)^{((\xi, \xi + K_X) - (K_X^2) - (K_X, \tilde{\xi}_1))/2} 2^{(K_X^2) - \chi_h(X) + 1} \frac{\phi^4}{1 - \phi^4} \frac{d\phi}{\phi} \\
 & \quad \times \sum_{i, j, k, l} \phi^{-((\xi - K_X)^2 + (K_X^2) - 5\chi_h(X) + k + 2l)} \\
 & \quad \times (-1)^{i+k-j} \Lambda^k \binom{(\xi - K_X, \tilde{\xi}_1)}{i} \binom{k}{j} (\tilde{\xi}_1, \alpha)^j (\xi - K_X, \alpha)^{k-j} \\
 & \quad \times \left(\frac{1 - \phi^4}{2\phi^4}\right)^{((\xi - K_X, \tilde{\xi}_1) - i + k - j)/2} \left(\frac{1 - 3\phi^4}{2\phi^4}\right)^{(i+j+(K_X^2) - \chi_h(X) + 2)/2} \\
 & \quad \times \left(-\frac{\Lambda^2}{2}(3 + \phi^{-4})x - \frac{1}{2}\Lambda^2(\alpha^2)z^2\right)^l \frac{1}{l!} \frac{z^k}{k!},
 \end{aligned}$$

where the summation runs over

$$2l + k \equiv \dim M_H(y) \pmod{4}, \quad i + j + (K_X^2) - \chi_h(X) \equiv 0 \pmod{2}.$$

Moreover, since we are interested in the residue at  $\phi^4 = 1/3$ , we only need to consider terms with

$$(8.10) \quad i + j + (K_X^2) - \chi_h(X) + 2 \leq -2.$$

We put

$$\zeta = \frac{1 - 3\phi^4}{2\phi^4}.$$

Then the above is equal to

$$\begin{aligned}
 (8.11) \quad & -(-1)^{((\xi, \xi + K_X) - (K_X^2) - (K_X, \tilde{\xi}_1))/2} 2^{(K_X^2) - \chi_h(X) - 1} d\zeta \\
 & \times \sum_{i, j, k, l} (2\zeta + 3)^{\{((\xi - K_X)^2) - (K_X^2) + 5\chi_h(X) - k - 2l\}/4 - 1} \\
 & \quad \times (-1)^{i+k-j} \Lambda^k \binom{(\xi - K_X, \tilde{\xi}_1)}{i} \binom{k}{j} (\tilde{\xi}_1, \alpha)^j (\xi - K_X, \alpha)^{k-j} \\
 & \quad \times (\zeta + 1)^{((\xi - K_X, \tilde{\xi}_1) - i + k - j)/2 - 1} \zeta^{(i+j+(K_X^2) - \chi_h(X) + 2)/2} \\
 & \quad \times \left(-\Lambda^2(\zeta + 3)x - \frac{1}{2}\Lambda^2(\alpha^2)z^2\right)^l \frac{1}{l!} \frac{z^k}{k!}.
 \end{aligned}$$

In order to illustrate the idea of the proof, let us first consider the simplest nontrivial case  $(K_X^2) - \chi_h(X) = -5$ . (The case  $(K_X^2) - \chi_h(X) = -4$  is too simple.) We only need to consider terms with  $i + j = 1$ , i.e.,  $i = 1, j = 0$  and  $i = 0, j = 1$

by (8.10). Then, up to a constant, the residue of (8.11) is

$$\begin{aligned}
 & (-1)^{(K_X, K_X + \tilde{\xi}_1)/2} \sum_{k,l} 3^{\{((\xi - K_X)^2) - (K_X^2) + 5\chi_h(X) - k - 2l\}/4} (-\Lambda)^k \\
 & \quad \times \left( -3\Lambda^2 x - \frac{1}{2}\Lambda^2(\alpha^2)z^2 \right)^l \frac{1}{l!} \frac{z^k}{k!} \\
 & \quad \times ((\xi - K_X, \tilde{\xi}_1)(\xi - K_X, \alpha)^k + k(\tilde{\xi}_1, \alpha)(\xi - K_X, \alpha)^{k-1}).
 \end{aligned}$$

Since  $\Lambda, x, z$  are formal variables, each term for individual  $k, l$  must be independent of  $t$ . Since  $(\xi - K_X, \tilde{\xi}_1)(\xi - K_X, \alpha)^k$  and  $k(\tilde{\xi}_1, \alpha)(\xi - K_X, \alpha)^{k-1}$  have different degrees in  $t$  (the former has degree  $k + 1$ , the latter has  $k - 1$ ), they cannot cancel out. Therefore we must have

$$\sum_{\tilde{\xi}_1} (-1)^{(K_X, K_X + \tilde{\xi}_1)/2} (\tilde{\xi}_1, \alpha) \text{SW}(\tilde{\xi}_1) = 0.$$

This is the superconformal simple type condition when  $(K_X^2) - \chi_h(X) = -5$ .

*Proof of Proposition 8.9.* For the same reason as in the special case  $(K_X^2) - \chi_h(X) = -5$ , each term for individual  $k, l$  must be independent of  $t$ .

We expand terms in (8.11) as

$$\begin{aligned}
 & (2\zeta + 3)^{\{((\xi - K_X)^2) - (K_X^2) + 5\chi_h(X) - k - 2l\}/4 - 1} (\zeta + 1)^{\{(\xi - K_X, \tilde{\xi}_1) - i + k - j\}/2 - 1} \\
 & = 3^{\{((\xi - K_X)^2) - (K_X^2) + 5\chi_h(X) - k - 2l\}/4 - 1} \\
 & \quad \times \left[ 1 + \zeta \left\{ \frac{2}{3} \left( \frac{1}{4}(-K_X^2) + 5\chi_h(X) - 2l \right) - 1 \right\} + \frac{-i - j}{2} - 1 + \frac{k}{3} \right. \\
 & \quad \left. + \frac{(\xi - K_X, \tilde{\xi}_1)}{2} + \frac{1}{6}((\xi - K_X)^2) \right\} + \dots \Big].
 \end{aligned}$$

The coefficient of  $\zeta^L$  in [ ] has the leading term (as a polynomial in  $k$ )

$$(8.12) \quad \sum_{L_1 + L_2 = L} \frac{(-k)^{L_1}}{4^{L_1}} \frac{2^{L_1}}{3^{L_1}} \frac{1}{L_1!} \times \frac{k^{L_2}}{2^{L_2}} \frac{1}{L_2!} = \frac{k^L}{3^L L!} \neq 0.$$

Moreover the coefficient of  $\zeta^L$  is a polynomial in  $(\xi - K_X, \tilde{\xi}_1)$  whose degree is at most  $L$ .

When we multiply the above expression with  $\zeta^{(i+j+(K_X^2)-\chi_h(X)+2)/2}$  in (8.11), it contributes to the residue at  $\zeta = 0$  only if

$$i + j = -2L + (\chi_h(X) - (K_X^2) - 4).$$

And the residue is a linear combination of

$$(\xi - K_X, \tilde{\xi}_1)^{p+q} ((K_X - \xi)^2)^r (\alpha, \tilde{\xi}_1)^j (\xi - K_X, \alpha)^{k-j}$$

for various  $p, q, r, j$ . ( $k$  is fixed, as we explained at the beginning.) Here  $q$  and  $r$  come from the above expansion, and  $p$  appears when we expand  $(\xi^{-K_X}, \tilde{\xi}_1)$ . Therefore we have

$$(8.13) \quad \begin{aligned} 0 \leq p \leq i \text{ and } i \neq 0 \text{ implies } p \neq 0, \\ q + r \leq L, \quad q, r \geq 0. \end{aligned}$$

Up to the factor  $3^{-((\xi - K_X)^2)}$ , each term is a polynomial in  $t$  with degree  $m := k - j + p + q + 2r$ . We will consider each coefficient of  $t^m$  whose sum over  $p, q, r, j$  and over the Seiberg–Witten classes  $\tilde{\xi}_1$  must be 0 by our assumption.

We set  $j_{\max} := \chi_h(X) - (K_X^2) - 4$ . Then  $j \leq j_{\max}$  and the equality holds if and only if  $i = L = 0$ . We will check the superconformal simple type condition (1.4) by descending induction on  $n$ . Starting from  $n = j_{\max}$ , we check it for  $n = j_{\max} - 2, j_{\max} - 4$  and so on. (As  $\text{SW}(-\tilde{\xi}_1) = (-1)^{\chi_h(X)} \text{SW}(\tilde{\xi}_1)$ , the assertions for  $n = j_{\max} - 1, j_{\max} - 3, \dots$  are automatically true. See also §2.5.)

We assume

$$k - j_{\max} \leq m \leq k, \quad k - j_{\max} \equiv m \pmod{2}.$$

We will be interested in  $j + p + q$ , which will appear as  $n$  in (1.4). We first note that

$$(8.14) \quad j + p + q \leq i + j + L = -L + (\chi_h(X) - (K_X^2) - 4) \leq j_{\max}.$$

Equality holds if and only if  $L = q = r = 0$  and  $p = i$ . We next note that

$$(8.15) \quad j + p + q = k - m + 2(p + q + r) \geq k - m.$$

Equality holds if and only if  $p = q = r = 0$  and  $(i, j) = (0, k - m)$ . Thus each coefficient of  $t^m$  is

$$(8.16) \quad A(\alpha, \tilde{\xi}_1)^{k-m} + (\text{higher order terms}),$$

where the “higher order terms” means a sum of monomials with  $j + p + q > k - m$ . By (8.12) we have  $A \neq 0$ . (More precisely, when we vary  $k$  fixing  $k - m$ , we have  $A \neq 0$  for sufficiently large  $k$ .)

We now start the descending induction on  $k - m$ . Start with  $k - m = j_{\max}$ . Then (8.14, 8.15) imply  $j = j_{\max}$  and  $p = q = r = 0$  and  $i = n = 0$ . Thus there are no higher order terms in the above expression, and we get the superconformal simple type condition (1.4) for  $n = j_{\max}$ .

If (1.4) is true for  $n > k - m$ , then the sum of the higher order terms in (8.16) over  $\tilde{\xi}_1$  vanishes. Hence we also get (1.4) with  $n = k - m$ . This completes the proof.  $\square$

### Acknowledgements

The authors thank Takuro Mochizuki for explanations of his results and discussions over the years. The second-named author is grateful to Yuji Tachikawa for discussions on Seiberg–Witten curves, and to Thomas Leness and Gregory Moore for useful comments. The authors thank the anonymous referee for useful comments.

The project began when the second and third-named authors were visiting the International Centre for Theoretical Physics in Aug. 2006. Part of the calculations was done while the second-named author was visiting Mathematical Institute of the University of Bonn, and the third-named author was at Max Planck Institute for Mathematics. The authors thank all the institutes for their hospitality.

The second named author is partly supported by the Grant-in-Aid for Scientific Research (B) (No. 19340006), Japan Society for the Promotion of Science.

The third named author is supported by the Grant-in-Aid for Scientific Research (B) (No. 18340010), (S) (No. 19104002), Japan Society for the Promotion of Science, and Max Planck Institute for Mathematics.

### References

- [1] L. Alvarez-Gaume, M. Marino and F. Zamora, Softly broken  $N = 2$  QCD with massive quark hypermultiplets, I, *Int. J. Modern Phys. A* **13** (1998), 403–430. Zbl 0936.81051 MR 1621192
- [2] A. Erdélyi, W. Magnus, F. Oberhettinger and F. G. Tricomi, *Higher transcendental functions. Vols. I, II*, Based, in part, on notes left by Harry Bateman. McGraw-Hill, New York, 1953. Zbl 0051.30303(Vol. I), Zbl 0052.29502(Vol. II) MR 0058756
- [3] A. Braverman and P. Etingof, Instanton counting via affine Lie algebras II: from Whittaker vectors to the Seiberg–Witten prepotential, in *Studies in Lie theory*, Progr. Math. 243, Birkhäuser, Boston, 2006, 61–78. Zbl 1177.14036 MR 2214246
- [4] S. K. Donaldson, The Seiberg–Witten equations and 4-manifold topology, *Bull. Amer. Math. Soc. (N.S.)* **33** (1996), 45–70. Zbl 0872.57023 MR 1339810
- [5] G. Ellingsrud, L. Göttsche and M. Lehn, On the cobordism class of the Hilbert scheme of a surface, *J. Algebraic Geom.* **10** (2001), 81–100. Zbl 0976.14002 MR 1795551
- [6] P. M. N. Feehan and T. G. Leness,  $PU(2)$  monopoles, I: Regularity, Uhlenbeck compactness, and transversality, *J. Differential Geom.* **49** (1998), 265–410. Zbl 0998.57057 MR 1664908
- [7] ———,  $PU(2)$  monopoles. II. Top-level Seiberg–Witten moduli spaces and Witten’s conjecture in low degrees, *J. Reine Angew. Math.* **538**, (2001), 135–212. Zbl 0983.57025 MR 1855755
- [8] ———, An  $SO(3)$ -monopole cobordism formula relating Donaldson and Seiberg–Witten invariants, <http://www.fiu.edu/~lenesst>.
- [9] ———, Witten’s conjecture for many four-manifolds of simple type, arXiv:math/0609530v2.
- [10] P. M. N. Feehan, P. B. Kronheimer, T. G. Leness and T. S. Mrowka,  $PU(2)$  monopoles and a conjecture of Mariño, Moore, and Peradze, *Math. Res. Lett.* **6** (1999), 169–182. Zbl 0967.57027 MR 1689207

- [11] R. Fintushel and R. J. Stern, The blowup formula for Donaldson invariants, *Ann. of Math.* **143** (1996), 529–546. Zbl 0869.57019 MR 1394968
- [12] ———, Donaldson invariants of 4-manifolds with simple type, *J. Differential Geom.* **42** (1995), 577–633. Zbl 0863.57021 MR 1367403
- [13] R. Friedman and J. W. Morgan, Obstruction bundles, semiregularity, and Seiberg–Witten invariants, *Comm. Anal. Geom.* **7** (1999), 451–495. Zbl 0946.14034 MR 1698386
- [14] O. García-Prada, Seiberg–Witten invariants and vortex equations, in *Symétries quantiques* (Les Houches, 1995), North-Holland, Amsterdam, 1998, 885–934. Zbl 0941.53002 MR 1616352
- [15] L. Göttzsche, H. Nakajima and K. Yoshioka, Instanton counting and Donaldson invariants, *J. Differential Geom.* **80** (2008), 343–390. Zbl 1172.57015 MR 2472477
- [16] L. Göttzsche and D. Zagier, Jacobi forms and the structure of Donaldson invariants for 4-manifolds with  $b_+ = 1$ , *Selecta Math. (N.S.)* **4** (1998), 69–115. Zbl 0924.57025 MR 1623706
- [17] P. B. Kronheimer and T. S. Mrowka, Embedded surfaces and the structure of Donaldson’s polynomial invariants, *J. Differential Geom.* **41** (1995), 573–734. Zbl 0842.57022 MR 1338483
- [18] J. Li, Algebraic geometric interpretation of Donaldson’s polynomial invariants, *J. Differential Geom.* **37** (1993), 417–466. Zbl 0809.14006 MR 1205451
- [19] A. Losev, N. Nekrasov and S. Shatashvili, Issues in topological gauge theory, *Nuclear Phys. B* **534** (1998), 549–611. Zbl 0954.57013 MR 1663467
- [20] M. Mariño, G. Moore and G. Peradze, Superconformal invariance and the geography of four-manifolds, *Comm. Math. Phys.* **205** (1999), 691–735. Zbl 0988.57019 MR 1711332
- [21] T. Mochizuki, *Donaldson type invariants for algebraic surfaces. Transition of moduli stacks*, Lecture Notes in Math. 1972, Springer, Berlin, 2009. Zbl 1177.14003 MR 2508583
- [22] G. Moore and E. Witten, Integration over the  $u$ -plane in Donaldson theory, *Adv. Theor. Math. Phys.* **1** (1997), 298–387. Zbl 0899.57021 MR 1605636
- [23] J. W. Morgan, Comparison of the Donaldson polynomial invariants with their algebro-geometric analogues, *Topology* **32** (1993), 449–488. Zbl 0801.57014 MR 1231956
- [24] ———, *The Seiberg–Witten equations and applications to the topology of smooth four-manifolds*, Math. Notes 44, Princeton Univ. Press, Princeton, NJ, 1996. Zbl 0846.57001 MR 1367507
- [25] H. Nakajima, *Lectures on Hilbert schemes of points on surfaces*, Univ. Lect. Ser. 18, Amer. Math. Soc., 1999. Zbl 0949.14001 MR 1711344
- [26] H. Nakajima and K. Yoshioka, Instanton counting on blowup. I. 4-dimensional pure gauge theory, *Invent. Math.* **162** (2005), 313–355. Zbl 1100.14009 MR 2199008
- [27] ———, Lectures on instanton counting, in *Algebraic structures and moduli spaces*, CRM Proc. Lecture Notes 38, Amer. Math. Soc., Providence, RI, 2004, 31–101. Zbl 1080.14016 MR 2095899
- [28] ———, Instanton counting on blowup. II.  $K$ -theoretic partition function, *Transform. Groups* **10** (2005), 489–519. Zbl 1110.14015
- [29] ———, Perverse coherent sheaves on blow-up. I. A quiver description, arXiv:0802.3120.
- [30] ———, Perverse coherent sheaves on blow-up. II. Wall-crossing and Betti numbers formula, *J. Algebraic Geom.* **20** (2011), 47–100. MR 2729275
- [31] ———, Perverse coherent sheaves on blow-up. III. Blow-up formula from wall-crossing, arXiv:0911.1773v2.
- [32] ———, Instanton counting on blowup. III. Theories with matters, in preparation.

- [33] N. Nekrasov, Seiberg–Witten prepotential from instanton counting, *Adv. Theor. Math. Phys.* **7** (2003), 831–864. Zbl 1056.81068 MR 2045303
- [34] N. Nekrasov and A. Okounkov, Seiberg–Witten prepotential and random partitions, in *The unity of mathematics*, Progr. Math. 244, Birkhäuser Boston, Boston, MA, 2006, 525–596. Zbl pre05050087 MR 2181816
- [35] K. G. O’Grady, Donaldson’s polynomials for  $K3$  surfaces, *J. Differential Geom.* **35** (1992), 415–427. Zbl 0735.57011 MR 1158343
- [36] V. Pidstrigach and A. Tyurin, Localisation of the Donaldson’s invariants along Seiberg–Witten classes, arXiv:dg-ga/9507004.
- [37] N. Seiberg and E. Witten, Electric-magnetic duality, monopole condensation, and confinement in  $N = 2$  supersymmetric Yang–Mills theory, *Nuclear Phys. B* **426** (1994), 19–52; Erratum, *Nuclear Phys. B* **430** (1994), 485–486. Zbl 0996.81510 MR 1293681
- [38] ———, Monopoles, duality and chiral symmetry breaking in  $N = 2$  supersymmetric QCD, *Nuclear Phys. B* **431** (1994), 484–550. Zbl 1020.81911 MR 1306869
- [39] E. Witten, Topological quantum field theory, *Comm. Math. Phys.* **117** (1988), 353–386. Zbl 0656.53078 MR 0953828
- [40] ———, Monopoles and four-manifolds, *Math. Res. Lett.* **1** (1994), 769–796. Zbl 0867.57029 MR 1306021