THE BLOW UP FORMULA FOR SU(3) DONALDSON INVARIANTS

LUCAS CULLER

1. Introduction

Donaldson's invariants of smooth 4-manifolds were originally defined using moduli spaces of anti-self dual connections on principal bundles with structure group $SU(2)$ or $SO(3) = PU(2)$. These invariants and their corresponding generating functions have many interesting structural features which have been described in a mathematically rigorous way. For instance, a paper of Fintushel and Stern [3] shows that the invariants of a 4-manifold X and its blowup $\hat{X} = X\#\overline{\mathbb{CP}}^2$ are related in a predictably way, and that the relationship can be expressed in terms of a theta function on the elliptic curve

$$y^2 = x^3 + ax^2 + x$$

over the ring $\mathbb{Q}[a]$. More generally, the structure theorem of Kronheimer-Mrowka expresses the Donaldson invariants of a 4-manifold of simple type in terms of exponential functions (i.e. theta functions on a degenerate genus 1 curve) and there is a conjectural structure theorem for 4-manifolds not of simple type, in terms of Jacobi elliptic functions.

It is natural to ask if Donaldson invariants can be defined for higher rank bundles, and if so, whether or not these structural features have some natural generalization. Indeed, the physics literature ([6],[2]) has achieved results in this direction, which indicate that the blow-up formula and structure theorem for $SU(N)$ invariants can be expressed in terms of function theory on the hyperelliptic curve

$$\Sigma_N : y^2 = (x^N + a_2x^{N-2} + \cdots + a_{N-1}x + a_N)^2 + 4$$

over the ring

$$\Lambda = H^*(BPU(N); \mathbb{Q}) = \mathbb{Q}[a_2, \ldots, a_N].$$

In particular it is shown in [2] that the blow-up function is a theta function on the Jacobian of this curve, generalizing the result of Fintushel and Stern. Unfortunately their derivation relies on the use of quantum field theory techniques, which do not appear to be mathematically rigorous.

One route to overcoming these difficulties, pursued by Nakajima and Yoshioka in [7], is through Nekrasov's partition function. Using the localization formula in equivariant cohomology, they rigorously define certain integrals which are analogous to the Donaldson invariants of the plane blown up at a single point. They compute these integrals using techniques of algebraic geometry, and thereby arrive at a blow-up formula which agrees with the one from the physics literature.

It is not clear, however, what either of the methods above has to do with moduli spaces of instantons on general 4-manifolds. In this paper we take a more direct route. In [4], Peter Kronheimer has given a rigorous definition of $PU(N)$ invariants for general N. Using his definition we can directly attempt to compute the blowup function, using the method of

Fintushel and Stern. For small values of N we can compute the blowup function recursively, and for $N = 3$ we can explicitly identify it with a theta function on the Jacobian of the curve Σ_3.

2. SU(3) Donaldson Invariants

Given a Riemannian 4-manifold X with $b^+(X) > 2$h, and a $U(N)$-bundle $E \to X$ one can consider the moduli space $\mathcal{M}_{X,c}$ of connections A with fixed trace θ, whose associated adjoint connection ad$A$ is anti-self dual. An index computation shows that this moduli space has virtual dimension

$$\dim \mathcal{M}_{X,E} = 4Nc_2(E) - (2N-2)c_1(E)^2 + (1-N^2)\left(\frac{\chi(X)+\sigma(X)}{2}\right)$$

In general the moduli space is quite singular, but with suitable perturbations of the ASD equations and a generic metric assumption it can be made smooth, provided $c_1(E)$ is coprime to N in the sense that there is a homology class $\sigma \in H_2(X)$ with

$$\langle c_1(E), \sigma\rangle \equiv 1 \mod N.$$

Once a smooth moduli space has been obtained, one can define polynomial invariants. Given a homology class $\Sigma \in H_2(M)$, Kronheimer introduces explicit divisors V_Σ in the moduli space of connections, by considering sections of certain determinant line bundles. If $\dim \mathcal{M}_E = 2k$ then the intersection

$$\mathcal{M}_{X,E} \cap V_{\Sigma_1} \cap \cdots \cap V_{\Sigma_k}$$

is a finite set of points, whose signed count remains constant under variations of the metric and perturbation data. Thus one obtains well defined polynomial invariants much as in the case of $SO(3)$ Donaldson theory. Following the notation of [3] and the conventions of [4] we will write

$$D_{X,c}(\mu_2(\Sigma_1)\cdots\mu_2(\Sigma_k)) = (2N)^{-k}\cdot \#\left(\mathcal{M}_E \cap V_{\Sigma_1} \cap \cdots \cap V_{\Sigma_k}\right)$$

where $c = c_1(E)$ is coprime to N. Following Kronheimer we extend this formula to the case where c is not coprime to N, using the blowup formula

$$D_{X,c}(z) = D_{X,c+e}(\mu_2(e)^{N-1}z),$$

which can be proved straightforwardly when c is already coprime to N and taken as a definition otherwise.

To properly state our results we will need to be able to evaluate other cohomology classes not described by Kronheimer. We begin by describing the rational cohomology ring of $BPU(N)$. Note that there is a map $\pi : BSU(N) \to BPU(N)$ with homotopy fiber $B(\mathbb{Z}/N\mathbb{Z})$. Since the rational cohomology of $B(\mathbb{Z}/N\mathbb{Z})$ is trivial, the Serre spectral sequence shows that the map $\pi^* : H^*(BPU(N);\mathbb{Q}) \to H^*(BSU(N);\mathbb{Q})$ is an isomorphism. We denote by a_k the unique class on $BPU(N)$ that maps to c_k under π^*. Thus if $P = P(E)$ is the projectivization of a vector bundle E whose determinant is divisible by N, we have

$$a_k(P) = c_k(E \otimes (\det E)^{-1/N})$$

In general, given any $PU(N)$ bundle P we can construct an associated bundle W_P, which reduces to

$$W_P = E^{\otimes N} \otimes (\det E)^{-1}$$

when $P = P(E)$. The classes a_k can be written explicitly as rational polynomials in the chern classes of W_P, due to the formula

$$ch(E) = (\exp(c_1(E))ch(W))^{1/N}.$$

The a_k defined in this way generate the rational cohomology ring of $BPU(N)$, which we write as:

$$\Lambda = H^*(BPU(N)) = \mathbb{Q}[a_2, \ldots, a_N].$$

From this we can obtain a description of the rational cohomology ring of $\mathcal{B}^*_E$, the configuration space of irreducible connections on a fixed $U(N)$ bundle $E \to X$ with fixed determinant θ mod the determinant 1 gauge group. Namely, there is a $PU(N)$ bundle

$$\mathbb{P} = \mathbb{P}_E = \mathcal{A}_E \times P/\mathcal{G} \to \mathcal{B}^*_E \times X$$

and hence for any homology class $z \in H_*(X)$ and any $a \in H^*(BPU(N))$ we have a μ class defined via the slant product:

$$\mu_a(z) = a(\mathbb{P}_E)/[z].$$

For simplicity, we will often (abusively) shorten our notation for μ classes as follows:

$$\mu_k(z) = \mu_{a_k}(z)$$
$$a_k = \mu_{a_k}([\mathrm{pt}])$$

with the latter notation justified by the fact that $\mu_a([\mathrm{pt}]) = a(\mathbb{P}_E)$. We denote by

$$\begin{aligned}\mathbb{A}(X) &= \mathbb{Q}[\{\mu_a(z) | a \in H^*(BPU(N)), z \in H_0(X;\mathbb{Z}) \oplus H_2(X;\mathbb{Z})\}] \\ &= \Lambda[\{\mu_k(\sigma) | \sigma \in H_2(X,\mathbb{Z}), k \in \mathbb{N}]\end{aligned}$$

the part of the rational cohomology of $\mathcal{B}^*_E$ generated by μ classes arising from points and surfaces. Given a moduli space $\mathcal{M}_{X,E} \subset \mathcal{B}^*_E$ of dimension d, and cohomology classes $\mu_{k_i}(z_i)$ whose codimensions sum to d, we would like to compute the pairing

$$\langle \mu_{k_1}(z_1) \cdots \mu_{k_r}(z_r), [\mathcal{M}_{\mathcal{X},\mathcal{E}}] \rangle$$

It is simplest to describe how to evaluate classes the a_k. To represent these classes we make use of the fact that they are rational characteristic classes of the universal $PU(N)$ bundle $\mathbb{P}_E \to \mathcal{B}^*_E \times \{\mathrm{pt}\}$. We can therefore express them as rational linear combinations of chern classes of the associated bundle $\mathbb{W}_E$ described above. But it is well known that the chern class c_k can be represented by a stratified subvariety, by considering the locus where $N^N - k + 1$ sections of $\mathbb{W}_E$ become linearly dependent. By intersecting these stratified subvarieties with $\mathcal{M}_{X,E} \times \mathrm{pt}$ we obtained a finite set of signed points whose count does not depend on the choice of generic metric or perturbation.

More generally, we proceed as follows. Given a homology class $z \subset X$ we represent it by a closed submanifold $Z \subset X$. Given a class $a \in \Lambda$ we represent the restriction of $a(\mathbb{P})$ to $\mathcal{B}^* \times Z$ by a stratified subvariety $\mathcal{V}_a$ as above (or more precisely, by a formal rational combination of such subvarieties). To evaluate the class $\mu_a(z)$ on a moduli space of ASD connections $\mathcal{M} \subset \mathcal{B}^*$ we intersect $\mathcal{M} \times Z$ with $\mathcal{V}_a$. To evaluate more classes we can intersect $\mathcal{M} \times Z_1 \times \cdots Z_r$ with the preimages of the $\mathcal{V}_i$ in the ambient space $\mathcal{B}^* \times X \times \cdots \times X$. When this intersection is zero dimensional, it is compact and hence its signed count is well-defined. We denote by $D_{X,c}(\mu_{k_1}(z_1) \cdots \mu_{k_r}(z_r))$ the signed count of points obtained in this way.

3. Moduli Spaces Near a Negative Sphere

Proposition 1. *Let $E_k = kL \oplus (N-k)\mathbb{C}$ be the $U(N)$ bundle over N_1 with $c_1(E_k) = k\sigma_1$. Then for a generic cylindrical end metric on N_1, the completely reducible ASD connection*

$$A_k = k\lambda \oplus (N-k)$$

has energy strictly less than that of any other ASD connection on E_k.

Proof. Note that if A is an irreducible ASD connection on a vector bundle F over N_1, it must live in a moduli space of nonnegative virtual dimension. By the index formula we get

$$4nc_2(F) - (2n-2)c_1(F)^2 \geq n^2 - 1,$$

and rewriting this in terms of the chern character of F yields

$$ch_2(F) \geq \frac{n^2 + 2k^2 - 1}{n},$$

where $n = \dim F$ and $c_1(F) = k\sigma_1$.

Now suppose A is an arbitrary ASD connection on the bundle E_k. Then $A = \bigoplus_i A_i$ where each A_i is an irreducible connection on a subbundle $F_i \subset E_k$ of dimension n_i. If we write $c_1(F_i) = k_i\sigma_1$ then by additivity of the chern character

$$ch_2(E_k) = \sum_i ch_2(F_i) \geq \sum_i \frac{n_i^2 + 2k_i^2 - 1}{n_i^2},$$

and therefore

$$\begin{aligned} \dim \mathcal{M}(A) &\geq 1 - N^2 + 2c_1(F)^2 + Nch_2(F) \\ &= 1 - N^2 + 2k(N-k) - 2kN + \sum_i \frac{N}{n_i}\left(n_i^2 + 2k_i^2 - 1\right) \\ &= \dim \mathcal{M}(A_k) + \sum_i \frac{N}{n_i}\left((n_i - k_i)^2 + k_i^2 - 1\right). \end{aligned}$$

Note that equality holds precisely when each n_i is 1. Thus A_k is the unique connection of minimal energy, as claimed.

∎

Proposition 2. *Let $E_k = kL \oplus (N-k)\mathbb{C}$ denote the $U(N)$ bundle over N_2 with $c_1(E_k) = k\sigma_2$, and let $\eta_r = r\chi \oplus (N-r)\mathbf{1}$ be a flat connection on the restriction of E_k to $L_2 = \partial N_2$. Then for a generic cylindrical end metric on N_2, the completely reducible ASD connection*

$$A_{k,r} = \begin{cases} \frac{k+r}{2}\lambda \oplus (N-r)\mathbf{1} \oplus \frac{r-k}{2}\lambda^{-1} & r \geq k \\ \frac{k-r}{2}\lambda^2 \oplus r\lambda \oplus (N - \frac{k+r}{2})\mathbf{1} & k \geq r \end{cases}$$

has minimal energy among all ASD connections on E_k limiting on η_r.

Proof. Let A be another ASD connection with $c_1(A) = k\sigma_2$ and limiting connection η_r. Again write $A = \bigoplus A_i$, where each A_i is an irreducible connection on a subbundle $F_i \subset E_k$. Let $\dim F_i = n_i$, $c_1(F_i) = k_i\sigma_2$, and suppose that A_i limits on the flat connection η_{r_i}. Each A_i lives in a moduli space of positive virtual dimension, hence by the index formula

$$4n_i c_2(A_i) - (2n_i - 2)c_1(A_i)^2 + \left(\frac{1-n_i^2}{2}\right) + \left(\frac{1 - r_i^2 - (n_i - r_i)^2}{2}\right) \geq 0.$$

Arguing along the lines of Proposition 1, we obtain an energy inequality

$$4Nc_2(A) - (2N-2)c_1(A)^2 \geq k(N-k) + \sum_i \frac{N}{n_i}(n_i^2 - r_i n_i + r_i^2 - k_i n_i + k_i^2 - 1)$$

and therefore

$$4Nc_2(A) - (2N-2)c_1(A)^2 \geq \begin{cases} k(N-k) + (r-k)N + \sum_i \frac{N}{n_i}((n_i - r_i)^2 + k_i^2 - 1) & r \geq k \\ k(N-k) + (k-r)N + \sum_i \frac{N}{n_i}((n_i - k_i)^2 + r_i^2 - 1) & k \geq r. \end{cases}$$

On the other hand, a straightforward computation shows

$$4Nc_2(A_{k,r}) - (2N-2)c_1(A_{k,r})^2 = \begin{cases} k(N-k) + (r-k)N & r \geq k \\ k(N-k) + (k-r)N & k \geq r, \end{cases}$$

so $A_{k,r}$ has minimal energy as claimed. □

Proposition 3. *Let $N = 3$ and let $\mathcal{M}_{0,2}$ be the ASD moduli space on N_2 containing the connection $A_{0,2}$. Then for generic cylindrical end metrics on N_2, $\mathcal{M}_{0,2}$ is the disjoint union of a finite set of irreducible connections and a set of reducible connections of the form $B \oplus \mathbf{1}$ for some ASD $SU(2)$ connection B. The space of all such B is a disjoint union of finitely many circles and a noncompact component diffeomorphic to the interval $[0, \infty)$.*

Proof. Note that the virtual dimension of the moduli space $\mathcal{M}_{0,2}$ is zero, hence all irreducible connections must be isolated, for generic metrics (and perturbations). Any reducible connection must be of the form $B \oplus \lambda^k$ but if k is nonzero then the inequality in Proposition 2 is strict, hence the connection does not have minimal energy. Thus every reducible is of the form $B \oplus \mathbf{1}$ as claimed.

Now consider the moduli space of ASD $SU(2)$ connections on N_2 with $c_1(B) = 1$, limiting on $\eta_2 = \chi \oplus \chi$. The virtual dimension of this moduli space is 1, hence the irreducible locus is a smooth 1-manifold, hence it consists of a union of circles and intervals. There are two types of noncompactness possible - a family of irreducibles can either limit on a reducible or lose energy at infinity. The latter ends are in bijection with flow lines from $\chi \oplus \chi$ to $\mathbf{1} \oplus \mathbf{1}$. There is one reducible, $\lambda \oplus \lambda^{-1}$, so there is at least one of each type of end, and there is a component of the $SU(2)$ moduli space joining these, which is diffeomorphic to $[0, \infty)$.

For a general metric on L_2, there might be more noncompact components of the $SU(2)$ moduli space, such that both ends correspond to losing energy at infinity. But we can rule these out by choosing a round metric on L_2. In this case, a round S^3 double covers L_2, so the flow lines must arise from the ADHM construction and be equivariant under the $\mathbb{Z}/2\mathbb{Z}$ deck transformation group. But the ADHM construction shows that the moduli space of flow lines on S^3 (of appropriate energy, mod translation) is isomorphic to hyperbolic 4-space and $\mathbb{Z}/2\mathbb{Z}$ acts on it isometrically, from which it is easy to see that the moduli space (being a 0-manifold) must be a single point. So there is a unique flow line from $\chi \oplus \chi$ to $1 \oplus 1$ of the appropriate energy, hence no such additional components. □

4. Some Initial Conditions

Proposition 4. *Let Y be a smooth 4-manifold and let $X = Y \# \overline{\mathbb{CP}^2}$. Then for any class $z \in \mathbb{A}(Y)$ we have:*

$$D_X(z) = D_Y(z)$$

Proof. When we stretch the neck joining Y and $\overline{\mathbb{CP}^2}$ we see that for dimension reasons the limiting connection on $\overline{\mathbb{CP}^2}$ must be trivial. Hence we are gluing a moduli space $\mathcal{M}_Y$ to the trivial connection, and the restriction map $\mathcal{M}_X \to \mathcal{M}_Y$ is an isomorphism. Since the class z is pulled back from $\mathcal{M}_Y$ we get the identity $D_X(z) = D_Y(z)$ as desired. □

Proposition 5. *Let X be a smooth 4-manifold containing an embedded sphere e of self-intersection -1, and let $I = (i_2, \ldots, i_N)$ be a multi-index of integers satisfying the inequality*

$$0 < 2i_2 + 4i_3 + \cdots + (2N-2)i_N < 4N.$$

Then for any class $z \in \mathbb{A}(e^\perp)$ we have:

$$D_X(\mu(e)^I z) = 0$$

Proof. As above the restriction map $\mathcal{M}_X \to \mathcal{M}_Y$ is an isomorphism. But in this case we are trying to evaluate an extra class $\mu(e)^I$ which has positive dimension and is pulled back from a point. Hence $D_X(\mu(e)^I z)$ is trivial. □

Proposition 6. *Let X be a smooth 4-manifold containing an embedded sphere e of self-intersection -1. Then for any class $z \in \mathbb{A}(e^\perp)$ we have:*

$$D_{X,e}(\mu_2(e)^{N-1} z) = D_X(z)$$

Proof. See Kronheimer [4]. □

Proposition 7. *Let $N = 3$ and let X be a smooth 4-manifold containing an embedded sphere σ of self-intersection -2. Then for any class $z \in \mathbb{A}(\sigma^\perp)$ we have:*

$$D_{X,c}(\mu_2(\sigma)^2 z) = -2D_{c+\sigma}(z)$$

Proof. Note that a neighbrohood of σ is diffeomorphic to N_2, so $X = Y \cup_{L_2} N_2$ for some Y. The moduli spaces involved in the above identity are obtained by taking a single moduli space of connections on Y and gluing two different moduli spaces on N_2. It follows (using an argument similar to [3]) that $D_{X,c}(\mu_2(\sigma)^2 z)$ and $D_{X,c+\sigma}(z)$ are related by a universal constant of proportionality:

$$D_{X,c}(\mu_2(\sigma)^2 z) = M D_{X,c+\sigma}(z).$$

To evaluate the constant M we can use a single example. Consider a K3 surface blown up at two points, and write $\sigma = e_1 - e_2$, where e_i are the exceptional spheres. Then σ is represented by a sphere of self-intersection -2. We can choose c and z such that $D_{X,c}(z) \neq 0$. Then

$$\begin{aligned} D_{c+e_1+e_2}(\mu_2(\sigma)^2\mu_2(e_1+e_2)^2 z) &= D_{c+e_1+e_2}((\mu_2(e_1)^4 - 2\mu_2(e_1)\mu_2(e_2) + \mu_2(e_2)^4)z) \\ &= -2D_{c+e_1+e_2}(\mu_2(e_1)^2\mu_2(e_2)^2 z) \\ &= -2D_c(z) \\ &= -2D_{c-e_1}(\mu_2(-e_1-e_2))^2 z) \\ &= -2D_{c+2e_1}(\mu_2(e_1+e_2)^2 z) \\ &= -2D_{c+e_1+e_2+\sigma}(\mu_2(e_1+e_2)^2 z) \end{aligned}$$

Since both sides are nonzero we conclude that $M = -2$, which completes the proof. $\square$

Proposition 8. *Let $N = 3$ and let $e \subset X$ be an exceptional sphere. Then for any class $z \in \mathbb{A}(e^\perp)$ we have:*

$$D_X(\mu_2(e)^6 z) = -6D_X(z)$$

Proof. Let Y denote X blown up at a point. From Proposition 7 we get:

$$D_{Y,c}(\mu_2(e_1+e_2)^2\mu_2(e_1-e_2)^4 z) = -2D_{Y,c+e_1+e_2}(\mu_2(e_1-e_2)^4 z)$$

Then from Proposition 5 we get:

$$D_{Y,c}(\mu_2(e_1)^6 + \mu_2(e_2)^6) = -2D_{Y,c+e_1+e_2}(6\mu_2(e_1)^2\mu_2(e_2)^2 z).$$

And finally Proposition 6 gives the desired conclusion:

$$2D_{X,c}(\mu_2(e)^6) = -2D_{X,c}(6z)$$

$\square$

5. A Recursive Formula

Theorem 1. *Let X be a smooth 4-manifold containing an embedded sphere σ of self-intersection -2. Then for any class $z \in \mathbb{A}(\sigma^\perp)$ and any $c \in \sigma^\perp$ we have:*

$$D_c(\mu_2(\sigma)^4 z) = -4D_c(a_2\mu_2(\sigma)^2 z) - 3D_c(\mu_3(\sigma)^2 z) \tag{1}$$

$$D_c(\mu_2(\sigma)^3\mu_3(\sigma)z) = -3D_c(a_3\mu_2(\sigma)^2 z) - D_c(a_2\mu_2(\sigma)\mu_3(\sigma)z) \tag{2}$$

Proof. Write $X = X_0 \coprod N_2$, where N_2 is a neighborhood of the sphere σ. By stretching the neck, we see that $\mathcal{M}$ can be written as a union of two open sets U and V, where connections in U are approximately trivial on N_2 and connections in V are approximately in the moduli space $\mathcal{M}_{0,2}$, which was described in the previous section.

We can describe the open set V as follows. After first cutting down by the class z, there is a certain moduli space $\mathcal{Z}$ of irreducible connections on X_0, and this moduli space has a universal bundle $\mathcal{Q} \to \mathcal{Z}$. The open set V then fibers over $\mathcal{Z}$ with fiber $\tilde{\mathcal{M}}_{0,2}$. More precisely, it can be described by an associated bundle construction:

$$V = \mathcal{Q} \times_\Gamma \tilde{\mathcal{M}}_{0,2}$$

where $\Gamma = P(U(2) \times U(1))$ is the stabilizer of the limiting flat connection $\eta_{0,2}$. As long as the dimension of $\mathcal{Z}$ is less than 8, it is gauranteed to be compact (by considering moduli spaces of flow lines from $\eta_{0,2}$ to $\mathbf{1}$). In the situations we must consider to prove our identities, $\mathcal{Z}$ will have dimensions 4 and 6.

To compute the classes $\mu_i(\sigma)$ we need to calculate the universal bundle P over $\mathcal{M} \times \sigma$. Note that this universal bundle can be trivialized on the set of all connections that are sufficiently close to the trivial connection on a small neighborhood of σ. Hence both classes $\mu_i(\sigma)$ are compactly supported within the open set V. So we need only compute the universal bundle over $V \times \sigma$. In fact, from our description of $\mathcal{M}_{0,2}$ we know that V retracts onto a codimension two subset D, which is obtained by gluing connections in Z to the completely reducible connection $A_{0,2} = \lambda \oplus \mathbf{1} \oplus \lambda^{-1}$. Hence we only need to compute P as a $PU(3)$ bundle over D.

To describe the universal bundle on $V \times \sigma$ observe that there is a Γ-equivariant universal bundle $\tilde{P}$ on $\tilde{\mathcal{M}}_{0,2} \times \sigma$, and that P is obtained from $\tilde{P}$ by the universal bundle construction. Hence we would like to compute $\tilde{P}$ as an equivariant bundle over the orbit of $A_{0,2}$. This orbit can be identified with the flag manifold $F_{1,1,1}$, and under this identification the bundle $\tilde{P}$ is given by:

$$\tilde{P} = \mathcal{P}(\lambda r_1 \oplus \lambda^{-1} r_2 \oplus s)$$

where r_1, r_2, and s are chern roots of the universal $P(U(1) \times U(1) \times U(1))$ bundle over $E\Gamma/P(U(1) \times U(1) \times U(1))$. Thus we have:

$$a_2(\tilde{P}) = a_2 + (r_1 - r_2)\lambda - \lambda^2$$

$$a_2(\tilde{P}) = a_3 + (r_1 - r_2)s\lambda - \lambda^2 s$$

and slanting by $[\sigma]$ we get:

$$\mu_2(\sigma) = r_2 - r_1$$

$$\mu_3(\sigma) = (r_2 - r_1)s = \mu_2(\sigma)s$$

Since $r_1 + r_2 + s = 0$ we have:

$$s^2 = (r_1 + r_2)^2 = (r_1 - r_2)^2 + 4r_1 r_2$$

and on the other hand,

$$a_2 = r_1 r_2 + (r_1 + r_2)s = r_1 r_2 - s^2 = -3r_1 r_2 (r_1 - r_2)^2$$

from which we arrive our first identity:

$$\mu_3(\sigma)^2 = \mu_2(\sigma)^4 + 4a_2\mu_2(\sigma)^2 + 4\mu_3(\sigma)^2$$

Multiplying by s we get:

$$\mu_2(\sigma)^3\mu_3(\sigma) = -4a_2\mu_2(\sigma)\mu_3(\sigma) - 3s^3\mu_2(\sigma)^2$$

Now, since s is a chern root we have:

$$s^3 = -a_2 s + a_3$$

and substituting gives the second identity:

$$\mu_2(\sigma)^3\mu_3(\sigma) = -a_2\mu_2(\sigma)\mu_3(\sigma) - 3a_3\mu_2(\sigma)^2$$

□

Corollary 1. *For any $I = (i_2, i_3)$ there is a class $b_I \in \Lambda$ such that*

$$D_X(\mu(e)^I z) = D_X(b_I z))$$

for any 4-manifold X with $b^+(X) > 1$, any exceptional sphere $e \subset X$, and any class $z \in \mathbb{A}(e^\perp)$.

Proof. By considering an arbitrary 4-manifold blown up twice, setting $\sigma = e_1 + e_2$, applying Theorem 1 to the classes $z = \mu(e_1 - e_2)^I$, and using the initial conditions proved in the previous section, we get an effective recursion that allows us to inductively compute $D_c(\mu(e)^I z)$ in terms of $D_c(a^I z)$. Explicitly we use a combination of identities (1) and (2) to compute $D_c(\mu_2(e)^i \mu_3(e)^j)$ by induction on i, for $j = 0, 1, 2$. Then we extend to arbitrary j using identity (1). □

Definition 1. *We define the $SU(3)$ blowup function to be the formal power series*

$$B(t_2, t_3) = \sum_{i,j=0}^{\infty} b_{ij} \frac{t_2^i t_3^j}{i!j!}$$

where the coefficients b_{ij} are computed according to the recursion of Corollary 1.

6. Theta Functions on a Genus 2 Jacobian

Definition 2. *Let $\Lambda \subset \mathbb{C}^g$ be a full lattice. We say that an entire function $\theta : \mathbb{C}^g \to \mathbb{C}$ is quasiperiodic with respect to Λ if for every $\lambda \in \Lambda$ there exist constants $a_\lambda \in \mathrm{Hom}(\mathbb{C}^g, C)$ and $b_\lambda \in \mathbb{C}$ such that*

$$\theta(x + \lambda) = e^{a_\lambda x + b_\lambda} \theta(x)$$

for all $x \in \mathbb{C}^g$.

For example, any Gaussian $f(x) = e^{x^T Ax + b^T x + c}$ is periodic. So is the Riemann theta function

$$\theta(x) = \sum_{n \in \mathbb{Z}^g} e^{\pi i n^T \Omega n} e^{2\pi i n^T z}$$

if the lattice takes the form $\Lambda = \mathbb{Z}^g \oplus \Omega\mathbb{Z}^g$ for Ω a symmetric matrix with positive definite imaginary part. If two quasiperiodic function have the same automorphy factors a_λ,b_λ, then any linear combination of them is quasiperiodic (with the same automorphy factors). The product of any two quasiperiodic functions is quasiperiodic (with different automorphy factors). A linear map followed by a quasiperiodic function is quasiperiodic (with respect to a different lattice).

Note that we can regard any quasiperiodic function as a section of a certain line bundle over $\mathbb{C}^g/\Lambda$, whose transition maps can be derived from the automorphy factors. We will use capital letters to denote the line bundle corresponding to a set of automorphy factors, so if θ is quasiperiodic, we will say that it is a section of the line bundle Θ.

Definition 3. *Let $\theta(x)$ be an analytic function on $\mathbb{C}^g$. We define the Hirota differentials $\theta[I](x)$ to be the unique functions such that*

$$\theta(x + y)\theta(x - y) = \sum_I \theta[I](x) \frac{y^I}{I!}$$

where the sum ranges over all multi-indices $I = (i_1, \ldots, i_g)$. In other words,

$$\theta[I](x) = \frac{\partial^{|I|}}{\partial y^I}\theta(x+y)\theta(x-y)|_{y=0}$$

Proposition 9. *The Hirota differentials of an arbitrary analytic (or $\mathcal{C}^\infty$) function θ satisfy the following properties:*

(1) $\theta[I] = 0$ *if I has odd degree.*
(2) $\theta[ij] = 2\partial_{ij}\log\theta$
(3) $\theta[ijkl] = 2\partial_{ijkl}\log\theta + 4\left(\partial_{ij}\log\theta\cdot\partial_{kl}\log\theta + \partial_{ik}\log\theta\cdot\partial_{jl}\log\theta + \partial_{il}\log\theta\cdot\partial_{jk}\log\theta\right)$

Proof. Straightforward computation. □

Proposition 10. *Let θ be a quasiperiodic function with respect to a full lattice $\Lambda \subset \mathbb{C}^g$, so that it defines a section of a line bundle Θ. Then for any $x \in \mathbb{C}^g$ the function*

$$\phi_t(x) = \theta(x+t)\theta(x-t)$$

is a section of 2Θ.

Proof. By definition, sections of Θ correspond to quasiperiodic functions ψ satisfying identities

$$\psi(x+\lambda) = e^{a_\lambda x + b_\lambda}\psi(x)$$

for every $\lambda \in \Lambda$. Thus sections of 2Θ correspond to those satisfying

$$\phi(x+\lambda) = e^{2a_\lambda x + 2b_\lambda}\phi(x)$$

It is then straightforward to check that this identity is satisfied by ϕ_t. □

Corollary 2. *If θ is a quasiperiodic function, then its Hirota differentials $\theta[I]$ are all sections of 2Θ.*

Proof. Since $\theta(x+y)\theta(x-y)$ is quasiperiodic in x, so are all its derivatives with respect to y. □

Proposition 11. *Let Σ be a curve of genus 2, and let θ be the Riemann theta function on its Jacobian $J\Sigma$. Then the Hirota differentials $\theta[I]$ with $\deg I = 0{,}2$ form a basis of $H^0(J\Sigma, 2\Theta)$.*

Proof. Translating and multiplying by a Gaussian, we can assume θ is an even function whose Hessian vanishes at the origin. In this case it suffices to prove that the 3 components of the logarithmic Hessian of θ are linearly independent. If there were a linear relation, then we would have $\partial_v\partial_w\log\theta = 0$ for some vectors v and w. But this implies that an irreducible component of the zero locus of θ is invariant under a 1-parameter group of translations, which is absurd since the zero locus is a genus 2 curve. □

Corollary 3. *The Riemann theta function on $J\Sigma$ satisfies a system of differential equations of the form:*

$$\theta[ijkl] = 2A_{ijkl}\theta^2 + \sum_{pq} C_{ijkl}^{pq}\theta[pq]$$

or equivalently,

$$P_{ijkl} + 2\left(P_{ij}P_{kl} + P_{ik}P_{jl} + P_{il}P_{jk}\right) = A_{ijkl} + \sum_{pq} C_{ijkl}^{pq}P_{pq}$$

where $P_I = \partial_I\log\theta$.

Proof. By Proposition 9, we know that the left hand side is a ratio $\frac{\phi(x)}{\theta(x)^2}$, where ϕ is a section of 2Θ. By Proposition 11, ϕ is a linear combination of Hirota differentials of order ≤ 2. Dividing by θ^2 and applying Proposition 9 again, the system of equations follows. $\square$

Observe that if $\theta(x)$ satisfies a system of the above form, then so does

$$e^{x^T Ax+b^t x+c}\theta(Mx+t)$$

for any matrices A, C, covector b and constants c, t. In other words, multiplying by a gaussian and precomposing with an affine transformation preserves the form of the equations. Ignoring this ambiguity, we can show that the Riemann theta function is characterized by the above system of equations. More precisely:

Theorem 2. *Suppose that θ is a function whose logarithmic derivatives satisfy a system of equations of the form:*

$$\begin{aligned}
P_{1111}+6P_{11}^2 &= A_{1111}+G_2P_{11}-2G_1P_{12}+G_0P_{22}\\
P_{1112}+6P_{11}P_{12} &= A_{1112}+G_3P_{11}-2G_2P_{12}+G_1P_{22}\\
P_{1122}+2P_{11}P_{22}+4P_{12}^2 &= A_{1122}+G_4P_{11}-2G_3P_{12}+G_2P_{22}\\
P_{1222}+6P_{12}P_{22} &= A_{1222}+G_5P_{11}-2G_4P_{12}+G_3P_{22}\\
P_{2222}+6P_{22}^2 &= A_{2222}+G_6P_{11}-2G_5P_{12}+G_4P_{22}
\end{aligned}$$

that the coefficients A_{ijkl} are given by:

$$\begin{aligned}
A_{1111} &= \frac{1}{3}\left(G_0G_4-4G_1G_3+3G_2^2\right)\\
A_{1112} &= \frac{1}{6}\left(G_0G_5-3G_1G_4+2G_2G_3\right)\\
A_{1122} &= \frac{1}{18}\left(G_0G_6-9G_2G_4+8G_3^2\right)\\
A_{1222} &= \frac{1}{6}\left(G_1G_6-3G_2G_5+2G_3G_4\right)\\
A_{2222} &= \frac{1}{3}\left(G_2G_6-2G_3G_5+3G_4^2\right)
\end{aligned}$$

and that θ is not itself a Gaussian. Suppose, moreover, that the homogeneous polynomial

$$G(x,y)=\sum_{i+j=6}G_i\frac{x^iy^j}{i!j!}$$

has exactly 6 nondegenerate zeroes in $\mathbb{P}^1$. Then, up to a multiplying by a Gaussian and precomposing with an affine transformation, θ is the Riemann theta function on the Jacobian of the hyperelliptic curve

$$y^2=g(x)=G(x,1)$$

Proof. This is proved in Baker [1]. To give some sense of what is involved in the argument, we will explain how to recover the hyperelliptic curve. The idea is as follows. First differentiate each of the 5 equations above with respect to each of the two variables. This yields a set of equations like:

$$P_{11112}+12P_{11}P_{112}=G_2P_{112}-2G_1P_{122}+G_0P_{222}$$

If we eliminate all 5th derivatives, we are left with system of 4 equations of the form:

$$M_{i1}P_{111} + M_{i2}P_{112} + M_{i3}P_{122} + M_{i4}P_{222} = 0$$

where the coefficients M_{ij} lie in $H^0(J\Sigma, 2\Theta)$. Since θ is not a Gaussian, at least one of the P_{ijk} is nonzero. Hence the 4 sections of $H^0(J\Sigma, 2\Theta)$ satisfy a nondegenerate degree 4 equation,

$$\det M = 0$$

This is the equation of the Kummer surface in $\mathbb{P}^3$. When we intersect it with the plane at infinity (i.e. the image of the theta divisor) we get a doubled conic. The Kummer surface has six singular points along this conic, which an explicit computation shows are precisely the roots of $G(x, y)$. □

We also need to make a remark about initial conditions. Clearly the function P is determined by $P_I(0)$ for $|I| \leq 3$. But as a consequence of Theorem 2, these initial conditions are subject to the constraint that the second derivatives must lie on the Kummer surface. There is, of course, no constraint on the zeroth or first derivatives, since they can be modified by multiplying by an exponential, which does not change the equations. Finally, it is proved in Baker that the third derivatives can be expressed in terms of the lower derivatives, which gives some actually constraints on the set of valid initial conditions. In the case where the initial conditions are finite and specify a singular point of the Kummer surface, these constraints simply say that the function θ is (up to third order) an even function times an exponential.

7. The Blowup Function

Theorem 3. *The $SU(3)$ blowup function is a quasiperiodic function on the Jacobian of the hyperelliptic curve*

$$y^2 = (x^3 + a_2x + a_3)^2 - 4.$$

Proof. We can express the recursion of Theorem 1 using the following identities:

$$\begin{aligned} &D_c(\mu_2(e_1 - e_2)^4 \exp(t(e_1 + e_2))) \\ &\qquad = -4D_c(\mu_2(e_1 - e_2)^2 \exp(t(e_1 + e_2))) - 3D_c(\mu_3(e_1 - e_2)^2 \exp(t(e_1 + e_2))) \\ &D_c(\mu_2(e_1 - e_2)^3 \mu_2(e_1 - e_2) \exp(t(e_1 + e_2))) \\ &\qquad = -3D_c(a_3\mu_2(e_1 - e_2)^2 \exp(t(e_1 + e_2))) - a_2D_c(\mu_3(e_1 - e_2)^2 \exp(t(e_1 + e_2))) \end{aligned}$$

Writing this in terms of the blowup function $B(t_2, t_3)$ we get

$$\begin{aligned} B_{2222}B - 4B_{222}B_2 + 3B_{22}B_{22} &= -4a_2(B_{22}B - B_2B_2) - 3(B_{33} - B_3B_3) \\ B_{2223}B - 3B_{223}B_2 - B_{222}B_3 + 3B_{22}B_{23} &= -3a_3P_{22} - a_2P_{23} \end{aligned}$$

and setting $Q = \log B$ we get

$$\begin{aligned} Q_{2222} + 6Q_{22}^2 &= -\,4a_2Q_{22} \qquad\qquad - 3Q_{33} \\ Q_{2223} + 6Q_{22}Q_{23} &= -\,3a_3Q_{22} - a_2Q_{23} \end{aligned}$$

which look similar to the first 2 of the 5 equations satisfied by the Riemann theta function. In fact, if we suppose that Q satisfies such a system of 5 equations, there is a unique choice of coefficients for the remaining three equations that is consistent with the initial conditions derived in Section 4. In total the five equations read:

$$\begin{aligned}
Q_{2222} + 6Q_{22}^2 &= \quad - \quad 4a_2Q_{22} \quad - 3Q_{33} \\
Q_{2223} + 6Q_{22}Q_{23} &= \quad - \quad 3a_3Q_{22} - a_2Q_{23} \\
Q_{2233} + 2Q_{22}Q_{33} + 4Q_{23}Q_{23} &= \quad 2 \quad - 3a_3Q_{23} - a_2Q_{33} \\
Q_{2333} + 6Q_{23}Q_{33} &= \quad - \quad a_2a_3Q_{22} + a_2^2Q_{23} - 3a_3Q_{33} \\
Q_{3333} + 6Q_{33}^2 &= -2a_2 - (12+3a_3^2)Q_{22} + 2a_2a_3Q_{23} + a_2^2Q_{33}
\end{aligned}$$

These equations are not quite in the form required by Theorem 2. However, we can bring them into that form if we multiply $B(t_2, t_3)$ by a suitable Gaussian. Specifically, we define

$$\theta(t_1, t_2) = \exp(\frac{3}{10}a_2t_2^2 + \frac{9}{20}a_3t_2t_3 - \frac{1}{10}a_2^2t_3^2)B(t_2, t_3).$$

Writing $P = \log\theta$ and applying we get the system of equations

$$\begin{aligned}
P_{2222} + 6P_{22}^2 &= \frac{9}{25}a_2^2 - \frac{2}{5}a_2P_{22} - 3P_{33} \\
P_{2223} + 6P_{22}P_{23} &= \frac{27}{50}a_2a_3 - \frac{3}{10}a_3P_{22} + \frac{4}{5}a_2P_{23} \\
P_{2233} + 2P_{22}P_{33} + 4P_{23}^2 &= 2 + \frac{27}{50}a_3^2 - \frac{1}{25}a_2^3 - \frac{1}{5}a_2^2P_{22} + \frac{3}{5}a_3P_{23} - \frac{2}{5}a_2P_{33} \\
P_{2333} + 6P_{23}P_{33} &= \frac{-9}{50}a_2^2a_3 - a_2a_3P_{22} + \frac{2}{5}a_2^2P_{23} - \frac{3}{10}a_3P_{33} \\
P_{3333} + 6P_{33}^2 &= \frac{8}{5}a_2 + \frac{1}{25}a_2^4 - (12+3a_3^2)P_{22} + 2a_2a_3P_{23} - \frac{1}{5}a_2^2P_{33}
\end{aligned}$$

The coefficients G_i now clearly take the form of Theorem 2, and a straightforward but tedious computation shows that the coefficients A_{ijkl} satisfy the necessary conditions as well.

It is also straightforward to check that the initial conditions for θ correspond to a singular point of the Kummer surface, and since the first derivatives vanish the integrability constraints simply say that the third derivatives vanish as well (which they do). Thus the $SU(3)$ blowup function can be equivalently defined as the unique solution of the complete set of 5 equations above, with the initial conditions derived in Section 4.

From Theorem 2 we now conclude that the $SU(3)$ blowup function is (the Taylor expansion of) a quasiperiodic function on the Jacobian of a certain hyperelliptic curve Σ. We can compute Σ explicitly using Theorem 2. It is given in hyperelliptic form by the equation:

$$\begin{aligned}
y^2 = g(x) &= -3\frac{x^6}{6!} - \frac{4}{10}a_2\frac{x^4}{4!2!} - \frac{3}{10}a_3\frac{x^3}{3!3!} - \frac{2}{10}a_2^2\frac{x^2}{2!4!} - a_2a_3\frac{x}{1!5!} - \frac{3a_3^2+12}{6!} \\
&= \frac{-1}{240}\left(x^6 + 2a_2x^4 + 2a_3x^3 + a_2^2x^2 + 2a_2a_3x + a_3^2 + 4\right) \\
&= \frac{-1}{240}\left(f(x)^2 + 4\right),
\end{aligned}$$

which is the desired result up to isomorphism. □

References

[1] H.F. Baker. Introduction to the Theory of Multiply Periodic Functions. Cambridge University Press, Cambridge, 1907.

[2] J.D. Edelstein , M. Gomez-Reino and M. Marino. Blowup Formulae in Donaldson-Witten theory and Integrable Hierarchies. Advances in Theoretical and Mathematical Physics, 4 (2000) 503.

[3] R. Fintushel and R. Stern. The Blowup Formula for Donaldson Invariants. Annals of Mathematics, 0 (1993), 1-17.

[4] P.B. Kronheimer. Four Manifold Invariants from Higher Rank Bundles. Journal of Differential Geometry, No. 70 (2005), 59-112.

[5] P.B. Kronheimer and T. Mrowka.

[6] M. Mario and G. Moore, The Donaldson-Witten function for gauge groups of rank larger than one, Comm. Math. Phys. 199 (1998), no. 1, 2569

[7] H. Nakajima and K. Yoshioka. Instanton Counting on Blowup. I. 4-Dimensional Pure Gauge Theory. Inventiones Mathematicae, 162(2):313-355, 2005.