ISSN 1364-0380 (on line) 1465-3060 (printed) **565**

Geometry & Topology
Volume 8 (2004) 565–610
Published: 31 March 2004

The Gromov invariant and the Donaldson–Smith standard surface count

Michael Usher

Department of Mathematics, MIT
Cambridge, MA 02139–4307, USA

Email: `usher@math.mit.edu`

Abstract

Simon Donaldson and Ivan Smith recently studied symplectic surfaces in symplectic 4–manifolds X by introducing an invariant $\mathcal{DS}$ associated to any Lefschetz fibration on blowups of X which counts holomorphic sections of a relative Hilbert scheme that is constructed from the fibration. Smith has shown that $\mathcal{DS}$ satisfies a duality relation identical to that satisfied by the Gromov invariant Gr introduced by Clifford Taubes, which led Smith to conjecture that $\mathcal{DS} = Gr$ provided that the fibration has high enough degree. This paper proves that conjecture. The crucial technical ingredient is an argument which allows us to work with curves C in the blown-up 4–manifold that are made holomorphic by an almost complex structure which is integrable near C and with respect to which the fibration is a pseudoholomorphic map.

AMS Classification numbers Primary: 53D45

Secondary: 57R17

Keywords: Pseudoholomorphic curves, symplectic Lefschetz fibrations, Gromov–Witten invariants

Proposed: Yasha Eliashberg | Received: 18 December 2003
Seconded: Ronald Fintushel, Ronald Stern | Accepted: 26 March 2004

1 Introduction

Let (X, ω) be a symplectic 4–manifold. Since the publication of Simon Donaldson's famous paper [2] it has been realized that a fruitful way of studying X is to construct a symplectic Lefschetz fibration $f\colon X' \to S^2$ on a suitable blow-up X' of X. One application of Lefschetz fibration techniques has been the work of Donaldson and Ivan Smith in [4] and [14] toward re-proving results concerning holomorphic curves in X which were originally obtained by Cliff Taubes in his seminal study of the Seiberg–Witten equations on symplectic manifolds. In [15], Taubes constructs a "Gromov invariant" $Gr(\alpha)$ which counts embedded, not necessarily connected, pseudoholomorphic submanifolds of X which are Poincaré dual to a class $\alpha \in H^2(X;\mathbb{Z})$, and in his other papers (collected in [16]) he identifies Gr with the Seiberg–Witten invariants. From the charge–conjugation symmetry in Seiberg–Witten theory there then follows the surprising Taubes duality relation that, where κ is the canonical class of X (ie, the first Chern class of the cotangent bundle), $Gr(\alpha) = \pm Gr(\kappa - \alpha)$, provided that $b^+(X) > 1$.

One might reasonably expect that a formula such as the Taubes duality relation could be proven in a more hands-on way than that provided by Seiberg–Witten theory, and Donaldson and Smith have indeed provided a somewhat more intuitive framework for understanding it. After perturbing ω to make its cohomology class rational and then scaling it to make it integral, Donaldson's construction gives, for large enough k, symplectic Lefschetz pencils $f_k\colon X \setminus B_k \to S^2$ (B_k being a set of $k^2[\omega]^2$ points obtained as the common vanishing locus of two sections of a line bundle over X) which lift to symplectic Lefschetz fibrations $f'_k\colon X'_k \to S^2$ where $\pi_k\colon X'_k \to X$ is the blowup of X along B_k; the fibers of f'_k are Poincaré dual to $k\pi_k^*[\omega]$. From any symplectic Lefschetz fibration $f\colon X' \to S^2$ and for any natural number r Donaldson and Smith [4] construct the "relative Hilbert scheme" $F\colon X_r(f) \to S^2$ whose fiber over a regular value t of f is the symmetric product $S^r f^{-1}(t)$; this is a smooth manifold that can be given a (continuous family of) symplectic structure(s) by the Thurston trick. A section of F then naturally corresponds to a closed set in X' which intersects each fiber of f r times (possibly counting multiplicities). So if we take an almost complex structure j on X' with respect to which the fibration $f\colon X' \to S^2$ is a pseudoholomorphic map (so that in particular the fibers of f are j–holomorphic and therefore intersect other j–holomorphic curves locally positively), then a holomorphic curve Poincaré dual to some class α and not having any fiber components will, to use Smith's words, "tautologically correspond" to a section of $X_r(f)$. This section will further be holomorphic with respect to the almost complex structure $\mathbb{J}_j$ on $X_r(f)$ obtained from j as follows: a tangent vector V at a point $\{p_1, \ldots, p_r\} \in X_r(f)$ where each $p_i \in f^{-1}(t)$ amounts to a collection of tangent vectors $v_i \in T_{p_i}X'$ such that all of the $\pi_* v_i \in T_t S^2$ are the same, and $\mathbb{J}_j V$ is defined as the collection of vectors $\{jv_1 \ldots, jv_r\}$. (The assumption that f is a pseudoholomorphic map with respect to j ensures that the 'horizontal parts' $\pi_* jv_i$ all agree, so that the collection $\{jv_1 \ldots, jv_r\}$ is in fact a well-defined tangent vector to $X_r(f)$; both Section 5 of [14] and a previous version of this paper assert that $\mathbb{J}_j$ can be constructed if j is merely assumed to make the fibers of f holomorphic, but this is not the case.) Conversely, a section s of $X_r(f)$ naturally corresponds to a closed set C_s in X' meeting each fiber r times with multiplicities, and s is $\mathbb{J}_j$–holomorphic exactly if C_s is a j–holomorphic subset of X'. Moreover, as Smith shows, there is just one homotopy class c_α of sections of $X_r(f)$ which tautologically correspond to closed sets in any given class α, and the expected complex dimension $d(\alpha)$ of the moduli space of such sections is the same as the expected dimension of the moduli space involved in the construction of the Gromov invariant. So it seems appropriate to try to count holomorphic curves in X by counting holomorphic sections of the various $X_r(f)$ in the corresponding homotopy classes. Accordingly, in [14] (and earlier in [4] for the special case $\alpha = \kappa$), the *standard surface count* $\mathcal{DS}_{(X,f)}(\alpha)$ is defined to be the Gromov–Witten invariant counting sections s of $X_r(f)$ in the class c_α with the property that, for a generic choice of $d(\alpha)$ points z_i in X, the value $s(f(z_i))$ is a divisor in $S^r f(z_i)$ containing the point z_i. Note that such sections will then descend to closed sets in X containing each of the points z_i. Actually, in order to count curves in X and not X' α should be a class in X, and the standard surface count will count sections of $X_r(f)$ in the class $c_{\pi_k^*(\alpha)}$; it's straightforward to see that $Gr(\pi_k^*(\alpha)) = Gr(\alpha)$. k here needs to be taken large enough that the relevant moduli space of sections of $X_r(f)$ is compact; we can ensure that this will be true if $k[\omega]^2 > \omega \cdot \alpha$, since in this case the section component of any cusp curve resulting from bubbling would descend to a possibly-singular symplectic submanifold of X' on which $\pi_k^*\omega$ evaluates negatively, which is impossible. With this compactness result understood, the Gromov–Witten invariant in question may be defined using the original definition given by Yongbin Ruan and Gang Tian in [10]; recourse to virtual moduli techniques is not necessary.

The main theorem of [14], proven using Serre duality on the fibers of f and the special structure of the Abel–Jacobi map from $X_r(f)$ to a similarly-defined "relative Picard scheme" $P_r(f)$, is that $\mathcal{DS}_{(X,f)}(\alpha) = \pm\mathcal{DS}_{(X,f)}(\kappa - \alpha)$, provided that $b^+(X) > b_1(X) + 1$ (and Smith in fact gives at least a sketch of a proof whenever $b^+(X) > 2$) and that the degree of the Lefschetz fibration is sufficiently high.

Smith's theorem would thus provide a new proof of Taubes duality under a somewhat weaker constraint on the Betti numbers if it were the case that (as Smith conjectures)

$$\mathcal{DS}_{(X,f)}(\alpha) = Gr(\alpha) \tag{1.1}$$

Even without this, the duality theorem is strong enough to yield several of the topological consequences of Taubes duality: for instance, the main theorem of [4] gives the existence of a symplectic surface Poincaré dual to κ; see also Section 7.1 of [14] for new Seiberg–Witten theory-free proofs of several other symplectic topological results of the mid-1990s. The tautological correspondence discussed above would seem to provide a route to proving the conjecture (1.1), but one encounters some difficulties with this. While the tautological correspondence implies that the moduli space of $\mathbb{J}$–holomorphic sections of $X_r(f)$ agrees set-theoretically with the space of j–holomorphic submanifolds of X, it is not obvious whether the weights assigned to each of the sections and curves in the definitions of the respective invariants will agree. This might seem especially worrisome in light of the fact that the invariant Gr counts some multiply-covered square-zero tori with weights other than ± 1 in order to account for the wall crossing that occurs under a variation of the complex structure when a sequence of embedded curves converges to a double cover of a square-zero torus.

This paper confirms, however, that the weights agree. The main theorem is:

Theorem 1.1 *Let $f\colon (X,\omega) \to S^2$ be a symplectic Lefschetz fibration and $\alpha \in H^2(X,\mathbb{Z})$ any class such that $\omega\cdot\alpha < \omega\cdot(fiber)$. Then $\mathcal{DS}_{(X,f)}(\alpha) = Gr(\alpha)$.*

The hypothesis of the theorem is satisfied, for instance, for Lefschetz fibrations f of sufficiently high degree obtained by Donaldson's construction applied to some symplectic manifold X_0 (X will be a blow-up of X_0) where α is the pullback of some cohomology class of X_0. In particular, the theorem implies that the standard surface count for such classes is independent of the degree of the fibration provided that the degree is high enough. It is not known whether this fact can be proven by comparing the standard surface counts directly rather than equating them with the Gromov invariant, though Smith has suggested that the stabilization procedure discussed in [1] and [13] might provide a route for doing so.

Combining the above Theorem 1.1 with Theorem 1.1 of [14], we thus recover:

Corollary 1.2 (Taubes) *Let (X,ω) be a symplectic 4–manifold with $b^+(X) > b_1(X)+1$ and canonical class κ. Then for any $\alpha \in H^2(X;\mathbb{Z})$, $Gr(\alpha) = \pm Gr(\kappa-\alpha)$.*

While the requirement on the Betti numbers here is stronger than that of Taubes (who only needed $b^+(X) > 1$), the proof of Corollary 1.2 via the path created by Donaldson and Smith and completed by Theorem 1.1 avoids the difficult gauge-theoretic arguments of [16] and also remains more explicitly within the realm of symplectic geometry.

We now briefly describe the proof of Theorem 1.1 and the organization of this paper. Our basic approach is to try to arrange to use, for some j making f pseudoholomorphic, the j–moduli space to compute Gr and the $\mathbb{J}_j$–moduli space to compute $\mathcal{DS}$, and to show that the contribution of each curve in the former moduli space to Gr is the same as the contribution of its associated section to $\mathcal{DS}$. In Section 2, we justify the use of such j in the computation of Gr. In Section 3, we refine our choice of j to allow $\mathbb{J}_j$ to be used to compute $\mathcal{DS}$, at least when there are no multiple covers in the relevant moduli spaces. For a non-multiply-covered curve C, then, we show that its contributions to Gr and $\mathcal{DS}$ agree by, in Section 4, directly comparing the spectral flows for C and for its associated section s_C of $X_r(f)$. This comparison relies on the construction of an almost complex structure which makes both C and f holomorphic and which is integrable near C. Although for an arbitrary curve C such an almost complex structure may not exist, the constructions of Section 3 enable us to reduce to the case where each curve at issue does admit such an almost complex structure nearby by first delicately perturbing the original almost complex structure on X. We use this result in Section 4 to set up corresponding spectral flows in X and $X_r(f)$ and show that the signs of the spectral flows are the same, which proves that curves with no multiply-covered components contribute in the same way to $\mathcal{DS}$ and Gr.

For curves with multiply covered components, such a direct comparison is not possible because the almost complex structure $\mathbb{J}$ is generally non-differentiable at the image of the section of $X_r(f)$ associated to such a curve. Nonetheless, we see in Section 5 that the contribution of such a j–holomorphic curve C to the invariant $\mathcal{DS}$ is still a well-defined quantity which remains unchanged under especially nice variations of j and C and which is the same as the contribution of C to Gr in the case where j is integrable and nondegenerate in an appropriate sense. To obtain this contribution, we take a smooth almost complex structure J which is close in Hölder norm to $\mathbb{J}$; because Gromov compactness remains true in the Hölder context, this results in the section s of $X_r(f)$ tautologically corresponding to C being perturbed into some number (possibly zero) of J–holomorphic sections which are constrained to lie in some small neighborhood of the original section s, and the contribution of C to $\mathcal{DS}$ is then obtained as the signed count of these nearby sections. We then deduce the agreement of $\mathcal{DS}$ and Gr by effectively showing that any rule for assigning contributions of j–holomorphic curves in the 4–manifold X which satisfies the invariance properties of the contributions to $\mathcal{DS}$ and agrees with the contributions to Gr in the integrable case must in fact yield Taubes' Gromov invariant. Essentially, the fact that $\mathcal{DS}$ is independent of the almost complex structure used to define it forces the contributions to $\mathcal{DS}$ to satisfy wall crossing formulas identical to those introduced by Taubes for Gr in [15]. Since the results of Section 3 allow us to assume that our curves admit integrable complex structures nearby which make the fibration holomorphic, and we know that contributions to $\mathcal{DS}$ and Gr are the same in the integrable case, the wall crossing formulas lead to the result that $\mathcal{DS} = Gr$ in all cases. This approach could also be used to show the agreement of $\mathcal{DS}$ and Gr for non-multiply covered curves, but the direct comparison used in Section 4 seems to provide a more concrete way of understanding the correspondence between the two invariants, and most of the lemmas needed for this direct proof are also necessary for the indirect proof given in Section 5, so we present both approaches.

Throughout the paper, just as in this introduction, a lowercase j will denote an almost complex structure on the 4–manifold, and an uppercase J (or $\mathbb{J}$) will denote an almost complex structure on the relative Hilbert scheme. When the complex structure on the domain of a holomorphic curve appears, it will be denoted by i.

This results of this paper are also contained in my thesis [17]. I would like to thank my advisor Gang Tian for suggesting this interesting problem and for many helpful conversations while this work was in progress.

2 Good almost complex structures I

Let $f\colon X \to S^2$ be a symplectic Lefschetz fibration and $\alpha \in H^2(X,\mathbb{Z})$. As mentioned in the introduction, if j is an almost complex structure on X with respect to which f is pseudoholomorphic, we have a tautological correspondence $\mathcal{M}_X^j(\alpha) = \mathcal{MS}_{X_r(f)}^{\mathbb{J}_j}(c_\alpha)$ between the space of j–holomorphic submanifolds of X Poincaré dual to α with no fiber components and the space of $\mathbb{J}_j$–holomorphic sections of $X_r(f)$ in the corresponding homotopy class. In light of this, to show that $Gr(\alpha)$ agrees with $\mathcal{DS}_{(X,f)}(\alpha)$, we would like, if possible, to use such an almost complex structure j to compute the former and the corresponding $\mathbb{J}_j$ to compute the latter. Two obstacles exist to carrying this out: first, the requirement that j make f holomorphic is a rather stringent one, so it is not immediately clear that the moduli spaces of j–holomorphic submanifolds will be generically well-behaved; second, the almost complex structure $\mathbb{J}_j$ is only Hölder continuous, and so does not fit into the general machinery for constructing Gromov–Witten invariants such as $\mathcal{DS}$. The first obstacle will be overcome in this section. The second obstacle is more serious, and will receive its share of attention in due course.

We will, in general, work with Lefschetz fibrations such that $\omega \cdot \alpha < \omega \cdot (fiber)$ for whatever classes α we consider; note that this requirement can always be fulfilled by fibrations obtained by Donaldson's construction, and ensures that j–holomorphic curves in class α never have any fiber components.

By a *branch point* of a j–holomorphic curve C we will mean a point at which C is tangent to one of the fibers of f.

Lemma 2.1 *Let $f\colon (X,\omega) \to (S^2, \omega_{FS})$ be a symplectic Lefschetz fibration and let $\alpha \in H^2(X,\mathbb{Z})$ be such that $d = d(\alpha) \geq 0$ and $\omega \cdot \alpha < \omega \cdot (fiber)$. Let $\mathcal{S}$ denote the set of pairs (j,Ω) where j is an almost complex structure on X making f holomorphic and Ω is a set of d distinct points of f, and let $\mathcal{S}^0 \subset \mathcal{S}$ denote the set for which:*

1. *(j,Ω) is nondegenerate in the sense of Taubes [15]; in particular, where $\mathcal{M}_X^{j,\Omega}(\alpha)$ denotes the set of j–holomorphic curves Poincaré dual to α passing through all the points of Ω, $\mathcal{M}_X^{j,\Omega}(\alpha)$ is a finite set consisting of embedded curves.*
2. *Each member of $\mathcal{M}_X^{j,\Omega}(\alpha)$ misses all critical points of f.*
3. *No curve in $\mathcal{M}_X^{j,\Omega}(\alpha)$ meets any of the branch points of any of the other curves.*

Then $\mathcal{S}^0$ is open and dense in $\mathcal{S}$.

Proof As usual for statements such as the assertion that Condition 1 is dense, the key is the proof that the map $\mathcal{F}$ defined from

$$\mathcal{U} = \{(i,u,j,\Omega) | (j,\Omega) \in \mathcal{S}, u\colon \Sigma \looparrowright X, \Omega \subset Im(u),\ u \in W^{k,p}\}$$

to a bundle with fiber $W^{k-1,p}(\Lambda^{0,1}T^*\Sigma \otimes u^*TX)$ by $(i,u,j,\Omega) \mapsto \bar{\partial}_{i,j}u$ is submersive at all zeroes. (i denotes the complex structure on the domain curve Σ.)

Now as in the proof of Proposition 3.2 of [10] (but using a $\bar\partial$–operator equal to one-half of theirs) , the linearization at a zero (i,u,j,Ω) is given by

$$\mathcal{F}_*(\beta,\xi,y,\vec{v}) = D_u\xi + \frac{1}{2}(y\circ du\circ i + j\circ du\circ\beta)$$

Here D_u is elliptic, β is a variation in the complex structure on Σ (and so can be viewed as a member of $H^{0,1}_i(T_{\mathbb{C}}\Sigma)$) and y is a j–antilinear endomorphism of TX that (in order that $\exp_j y$ have the compatibility property) preserves $T^{vt}X$ and pushes forward trivially to S^2, so with respect to the splitting $TX = T^{vt}X\oplus T^{hor}X$ (T^{hor} being the symplectic complement of T^{vt}; of course this splitting only exists away from $Crit(f)$) y is given in block form as

$$y = \begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix}$$

where all entries are j–antilinear.

Now suppose $\eta\in W^{k-1,p}(\Lambda^{0,1}T^*\Sigma\otimes u^*TX)$, so that η is a complex-antilinear map $T\Sigma\to u^*TX$, and take a point $x_0\in\Sigma$ for which $d(f\circ u)(x_0)$ is injective. Let v be a generator for $T^{1,0}_{x_0}\Sigma$; then $du(i(v))\in(T^{1,0}X)_{u(x_0)}$ and $du(i(\bar v))\in(T^{0,1}X)_{u(x_0)}$ are tangent to $u(\Sigma)$ and so have nonzero horizontal components. We take $y(u(x_0)) = \begin{pmatrix} 0 & b \\ 0 & 0 \end{pmatrix}$ where

$$b\colon\ T^{hor}_{u(x_0)}\to T^{vt}_{u(x_0)}$$

is a j–antilinear map with $b(du(v)^{hor}) = (\eta(v))^{vt}$ and $b(du(\bar v)^{hor}) = (\eta(\bar v))^{vt}$. Since complex antilinear maps are precisely those maps interchanging $T^{1,0}$ with $T^{0,1}$ this is certainly possible.

Suppose now that $\eta\in\operatorname{coker}(\mathcal{F}_*)_{(i,u,j,\Omega)}$. The above considerations show that for any point $x_0\notin Crit(f\circ u)$ there is y such that

$$\mathcal{F}_*(0,0,y,0)(x_0) = \eta^{vt}(x_0). \tag{2.1}$$

Cutting off y by some function χ supported near x_0, if $\eta^{vt}(x_0)\neq 0$ we can arrange that

$$\int_\Sigma\langle\mathcal{F}_*(0,0,\chi y,0),\eta\rangle = \int_\Sigma\langle\mathcal{F}_*(0,0,\chi y,0),\eta^{vt}\rangle > 0,$$

contradicting the supposition that $\eta\in\operatorname{coker}(\mathcal{F}_*)_{(i,u,j,\Omega)}$. η^{vt} must therefore be zero at every point not in $Crit(f\circ u)$.

Meanwhile, letting η^C denote the projection of η (which is an antilinear map $T\Sigma\to u^*TX$) to TC where $C = Im(u)$, η^C then is an element of the cokernel of the linearization at (i, id) of the map $(i', v) \mapsto \bar{\partial}_{i',i} v$, i' being a complex structure on Σ and v being a map $\Sigma \to \Sigma$. But the statement that this cokernel vanishes is just the statement that the set of complex structures on Σ is unobstructed at i (for the cokernel of the map $v \to \bar{\partial}_{i,i} v$ is $H^1(T_{\mathbb{C}}\Sigma)$, which is the same as the space through which the almost complex structures i' vary infinitesimally, and the relevant linearization just sends a variation β in the complex structure on Σ to $i\beta/2$). So in fact $\eta^C = 0$.

Now at any point x on Σ at which $(f \circ u)_*(x) \neq 0$, TC and $T^{vt}X$ together span TX, so since $\eta^C(x) = \eta^{vt}(x) = 0$ we have $\eta(x) = 0$. But the assumption on the size of the fibers ensures that $(f \circ u)_*(x) \neq 0$ for all but finitely many x, so η vanishes at all but finitely many x, and hence at all x since elliptic regularity implies that η is smooth. This proves that $(\mathcal{F}_*)_{(i,u,j,\Omega)}$ is submersive whenever $\mathcal{F}(i, u, j, \Omega) = 0$. The Sard–Smale theorem applied to the projection $(i, u, j, \Omega) \mapsto (j, \Omega)$ then gives that Condition 1 in the lemma is a dense (indeed, generic) condition; that it is an open condition just follows from the fact that having excess kernel is a closed condition on the linearizations of the $\bar{\partial}$, so that degeneracy is a closed condition on (j, Ω).

As for Conditions 2 and 3, from the implicit function theorem for the $\bar{\partial}$–equation it immediately follows that both are open conditions on $(j, \Omega) \in \mathcal{S}$ satisfying Condition 1, so it suffices to show denseness. To begin, we need to adjust the incidence condition set Ω so that it is disjoint from the critical locus of f and from all of the branch points of all of the curves of $\mathcal{M}_X^{j,\Omega}(\alpha)$. So given a nondegenerate pair (j, Ω) we first perturb Ω to be disjoint from $crit(f)$ while (j, Ω) remains nondegenerate; then, supposing a point $p \in \Omega$ is a branch point of some $C_0 \in \mathcal{M}_X^{j,\Omega}(\alpha)$, we change Ω by replacing p by some p' on C_0 which is not a branch point of C_0 and is close enough to p that for each other curve $C \in \mathcal{M}_X^{j,\Omega}(\alpha)$ which does not have a branch point at p, moving p to p' has the effect of replacing C in the moduli space by some C' which also does not have a branch point at p' (this is possible by the implicit function theorem). Denoting the new incidence set by Ω', the number of curves of $\mathcal{M}_X^{j,\Omega'}(\alpha)$ having a branch point at p' is one fewer than the number of curves of $\mathcal{M}_X^{j,\Omega}(\alpha)$ having a branch point at p, and so repeating the process we eventually arrange that no curve in $\mathcal{M}_X^{j,\Omega}(\alpha)$ has a branch point at any point of Ω.

So now assume $(j, \Omega) \in \mathcal{S}$ with Ω missing both $Crit(f)$ and all branch points of all curves in $\mathcal{M}_X^{j,\Omega}(\alpha)$. Let

$$\mathcal{M}_X^{j,\Omega}(\alpha) = \{[u_1], \ldots, [u_r]\}$$

where $[u_m]$ denotes the equivalence class of a map u_m under the action of $Aut(\Sigma_m)$, Σ_m being the (not necessarily connected) domain of u_m. For each

m, enumerate the points of Σ_m which are mapped by u_m either to $Crit(f)$ or to an intersection point with one of the other curves as $p_{m,1}, \ldots, p_{m,l}$, so in particular none of the $u_m(p_{m,k})$ lie in Ω. Take small, disjoint neighborhoods $U_{m,k}$ of the $p_{m,k}$ such that $u_m(U_{m,k})$ misses Ω and $u_m(U_{m,k} \setminus \frac{1}{2}U_{m,k})$ misses each of the other curves and also misses $Crit(f)$, and take *local* sections $\xi_{m,k}$ of $u_m^* T^{vt} X$ over $U_{m,k}$ such that $D_{u_m}\xi_{m,k} = 0$ and $\xi_{m,k}(p_{m,k}) \neq 0$ (this is certainly possible, as the $\xi_{m,k}$ only need to be defined on small discs, on which the equation $D_{u_m}\xi_{m,k} = 0$ has many solutions). Now for each m glue the $\xi_{m,k}$ together to form $\xi_m \in \Gamma(u_m^* T^{vt} X)$ by using cutoff functions which are 1 on $\frac{1}{2}U_{m,k}$ and 0 outside $U_{m,k}$. Then since $D_{u_m}\xi_{m,k} = 0$ the sections $D_{u_m}\xi_m$ will be supported in

$$A_m = \bigcup_k (U_{m,k} \setminus \frac{1}{2}U_{m,k}).$$

Now according to page 28 of [8], the linearization D_{u_m} may be expressed with respect to a j–Hermitian connection ∇ by the formula

$$(D_{u_m}\xi)(v) = \frac{1}{2}(\nabla_v \xi + j(u_m)\nabla_{iv}\xi) + \frac{1}{8}N_j((u_m)_* v, \xi) \tag{2.2}$$

where N_j is the Nijenhuis tensor. Our sections ξ_m are vertically-valued, so the first two terms above will be vertical tangent vectors; in fact, the last term will be as well, because where z is the pullback of the local coordinate on S^2 and w a holomorphic coordinate on the fibers, the anti-holomorphic tangent space for j can be written

$$T_j^{0,1}X = \langle \partial_{\bar z} + b(z,w)\partial_w, \partial_{\bar w} \rangle,$$

in terms of which one finds

$$N_j(\partial_{\bar z}, \partial_{\bar w}) = 4(\partial_{\bar w} b)\partial_w. \tag{2.3}$$

So if ξ is a vertically-valued vector field, the right-hand side of Equation 2.2 is also vertically-valued for any v, ie, D_{u_m} maps $W^{k,p}(u_m^* T^{vt} X)$ to $W^{k-1,p}(\Lambda^{0,1}T^*\Sigma_m \otimes u_m^* T^{vt} X)$ (and not just to $W^{k-1,p}(\Lambda^{0,1}T^*\Sigma_m \otimes u_m^* TX)$). Now

$$D_{u_m}\xi_m \in W^{k-1,p}(\Lambda^{0,1}T^*\Sigma_m \otimes u_m^* T^{vt} X)$$

is supported in A_m, so (using that $u_m(A_m)$ misses $Crit(f)$) as in (2.1) we can find a perturbation y_m of the almost complex structure j supported near $u_m(A_m)$ such that

$$\mathcal{F}_*(0, \xi_m, y_m, 0) = D_{u_m}\xi_m + \frac{1}{2} y_m \circ du_m \circ m = 0.$$

Since the $u_m(\bar{A_m})$ are disjoint, we can paste these y_m together to obtain a global perturbation y with $\mathcal{F}_*(0, \xi_m, y, 0) = 0$ for each m. For $t > 0$ small enough that $(\exp_j(ty), \Omega)$ remains nondegenerate, the holomorphic curves for the complex structure $\exp_j(ty)$ will be approximated in any $W^{k,p}$ norm ($p > 2$) to order $C\|\exp_j(ty) - j\|_{C^1}\|t\xi_m\|_{W^{k,p}} \leq Ct^2$ by the curves $\exp_{u_m}(t\xi_m)$ (using, for example, the implicit function theorem as formulated in Theorem 3.3.4 and Proposition 3.3.5 of [8]). Now since $\xi_m(p_{m,k}) \neq 0$, the $\exp_{u_m}(t\xi_m)$ will have their branch points moved vertically with respect to where they were before; in particular, these curves will no longer pass through $Crit(f)$, and their branch points will no longer meet other curves. Similarly (for t suitably small, and k appropriately large chosen at the beginning of the procedure) any set of curves within Ct^2 of these in $W^{k,p}$–norm will satisfy these conditions as well. So for t small enough, $(\exp_j(ty), \Omega)$ will obey conditions 1 through 3 of the lemma. (j, Ω) was an arbitrary nondegenerate pair, so it follows that $\mathcal{S}^0$ is dense. □

As has been mentioned above, the almost complex structure $\mathbb{J}_j$ that we would in principle like to use to evaluate $\mathcal{DS}$ is generally only Hölder continuous; however, under certain favorable circumstances we shall see presently that it is somewhat better-behaved. To wit, assume that our almost complex structure j is given locally by

$$T_j^{0,1} = \langle \partial_{\bar{z}} + b(z,w)\partial_w, \partial_{\bar{w}} \rangle,$$

where z is the pullback of the coordinate on the base and w a coordinate on the fibers. Then, following [11], where σ_k denotes the kth elementary symmetric polynomial, the function

$$\hat{b}_d(z, w_1, \dots, w_r) = \sum_{k=1}^{r} \sigma_{d-1}(w_1, \dots, \widehat{w_k}, \dots, w_r) b(z, w_k)$$

on $\mathbb{C} \times \mathbb{C}^r$ is symmetric in the w_k and so descends to a function $b_d(z, \sigma_1, \dots, \sigma_r)$ on $\mathbb{C} \times S^r\mathbb{C}$, and our almost complex structure $\mathbb{J}_j$ on $X_r(f)$ is given locally by

$$T_{\mathbb{J}_j}^{0,1} = \langle \partial_{\bar{z}} + \sum_{d=1}^{r} b_d(z, \sigma_1, \dots, \sigma_r)\partial_{\sigma_d}, \partial_{\bar{\sigma}_1}, \dots, \partial_{\bar{\sigma}_r} \rangle.$$

The nondifferentiability of $\mathbb{J}_j$ can then be understood in terms of the fact that smooth symmetric functions on $\mathbb{C}^r$ such as $\hat{b}_d(z, \cdot)$ generally only descend to Hölder continuous functions in the standard coordinates $\sigma_1, \dots, \sigma_r$ on $S^r\mathbb{C}$ (when $r = 2$, for example, consider the function $\bar{w}_1 w_2 + w_1 \bar{w}_2$). On the other hand, *holomorphic* symmetric functions on $\mathbb{C}^r$ descend to holomorphic (and in particular smooth) functions on the symmetric product, so when $\partial_{\bar{w}} b = 0$, the functions b_d are holomorphic in the vertical coordinates, and so $\mathbb{J}_j$ is smooth. Furthermore, note that by Equation 2.3, b is holomorphic in w exactly when j is integrable on the neighborhood under consideration; moreover, computing the Nijenhuis tensor of $\mathbb{J}_j$ shows that $\mathbb{J}_j$ is integrable exactly when $\partial_{\bar{\sigma}_k} b_l = 0$ for all k and l. This sets the stage for the following proposition, which foreshadows some of the constructions in the next two sections:

Proposition 2.2 *Let $C \in \mathcal{M}_X^{j,\Omega}(\alpha)$ where (j,Ω) is as in Lemma 2.1, and let s_C be the corresponding section of $X_r(f)$. If j is integrable on a neighborhood of C, then $\mathbb{J}_j$ is integrable on a neighborhood of s_C. More generally, if j is only integrable on neighborhoods of each of the branch points of C, then $\mathbb{J}_j$ is still smooth on a neighborhood of s_C.*

Proof The first statement follows directly from the above argument. As for the second statement, note that the only place where our functions b_d above ever fail to be smooth is in the diagonal stratum Δ of $\mathbb{C} \times S^r\mathbb{C}$ where two or more points in the divisor in $S^r\mathbb{C}$ come together. A suitably small neighborhood of s_C only approaches this stratum in a region whose differentiable structure for the vertical coordinates is just that of the Cartesian product of symmetric products of neighborhoods of all the branch points in some fiber (where smoothness is taken care of by the integrability assumption) with copies of $\mathbb{C}$ corresponding to neighborhoods of each of the other points of C which lie in the same fiber. □

We close this section with a proposition which shows that if $\mathbb{J}_j$ can be assumed smooth, then its moduli spaces will generically be well-behaved. We make here a statement about generic almost complex structures from a set $\mathcal{S}^1$ which at this point in the paper has not yet been proved to be nonempty; rest assured that it will be seen to be nonempty in the following section.

Proposition 2.3 *For generic (j,Ω) in the set $\mathcal{S}^1$ consisting of members of the set $\mathcal{S}^0$ from Lemma 2.1 which satisfy the additional property that j is integrable near every branch point of every curve C in $\mathcal{M}_X^{j,\Omega}(\alpha)$, the linearization of the operator $\bar{\partial}_{\mathbb{J}_j}$ is surjective at each of the sections s_C.*

Proof We would like to adapt the usual method of constructing a universal moduli space $\mathcal{U} = \{(s,j,\Omega) | \bar{\partial}_{\mathbb{J}_j} u = 0, (j,\Omega) \in \mathcal{S}^1, \Omega \subset C_s\}$, appealing to the implicit function theorem to show that $\mathcal{U}$ is a smooth Banach manifold, and then applying the Sard–Smale theorem to the projection from $\mathcal{U}$ onto the second factor (ie, $\mathcal{S}^1$) to obtain the statement of the proposition. Just as in the proof of Lemma 2.1, this line of argument will work as long as we can show that the map $(s,j,\Omega) \mapsto \bar{\partial}_{\mathbb{J}_j} s$ is transverse to zero.

Arguing as before, it's enough to show that, for a section s with $\bar{\partial}_{\mathbb{J}_j} s = 0$, where D_s^* denotes the formal adjoint of D_s, and where i denotes the complex structure on S^2, the following holds: if $D_s^*\eta = 0$, and if, for every variation y in the complex structure j on X among almost complex structures j' with $(j', \Omega) \in \mathcal{S}^1$, letting Y denote the variation in $\mathbb{J}_j$ induced by y, we have that

$$\int_{S^2} \langle \eta, Y(s) \circ ds \circ i \rangle = 0, \tag{2.4}$$

then $\eta \equiv 0$. If η were nonzero, then it would be nonzero at some $t_0 \in S^2$ which is not the image under f of any of the branch points of C_s, so assume this to be the case. Now η is a $s^*T^{vt}X_r(f)$–valued (0,1)–form, so giving its value at t_0 is equivalent to giving r maps $\eta_k\colon T_{t_0}S^2 \to T^{vt}_{s_k(t_0)}X$ $(r = 1, \ldots, k)$, where the $s_k(t_0)$ are the points in the fiber Σ_{t_0} over t_0 of the Lefschetz fibration which correspond to the point $s(t_0) \in S^r\Sigma_{t_0}$ (our assumption on t_0 ensures that these are all distinct). $\eta(t_0)$ being nonzero implies that one of these cotangent vectors (say η_m) is nonzero. Then s_m is a local holomorphic section of $X \to S^2$ around t_0, and exactly as in the proof of Lemma 2.1 we may find a perturbation y_0 of the almost complex structure near $s_m(t_0)$ such that

$$y_0(s_m(t_0)) \circ ds_m(t_0) \circ i = \eta_m$$

and y_0 preserves the pseudoholomorphicity of the fibration f. Multiplying y_0 by a smooth cutoff supported in a suitably small neighborhood of $s_m(t_0) \in X$, we obtain a variation y of the complex structure on X whose associated variation Y in $\mathbb{J}_j$ violates (2.4); note that since y is supported away from the nodes of the curves of $\mathcal{M}_X^{j,\Omega}(\alpha)$, the variation will also not disrupt the integrability condition in the definition of $\mathcal{S}^1$. This contradiction shows that η must vanish everywhere, and hence that $(s, j, \Omega) \mapsto \bar{\partial}_{\mathbb{J}_j} s$ is indeed transverse to zero, so that the universal space $\mathcal{U}$ will be a manifold and the usual Sard–Smale theorem argument implies the proposition. □

3 Good almost complex structures II

We fix a symplectic Lefschetz fibration $f\colon X \to S^2$ and a class $\alpha \in H^2(X, \mathbb{Z})$. Assume unless otherwise stated that $(j, \Omega) \in \mathcal{S}^0$, so that each curve $C \in \mathcal{M}_X^{j,\Omega}(\alpha)$ is identified by the tautological correspondence with a section s_C of $X_r(f)$ which misses the critical locus. Assume also that α cannot be decomposed as a sum of classes each of which pairs positively with ω and one of which, say β, satisfies $\kappa \cdot \beta = \beta \cdot \beta = 0$. Then the contribution of $C \in \mathcal{M}_X^{j,\Omega}(\alpha)$ to the invariant $Gr(\alpha)$ is found by considering a path of operators D_t acting on sections of the disc normal bundle U_C of C such that D_0 is the $\bar{\partial}$ operator obtained from the complex structure j_0 on U_C given by pulling back $j|_C$ to U_C

via the Levi–Civita connection, while D_1 is the $\bar\partial$ operator obtained by viewing U_C as a tubular neighborhood of C in X and restricting j to U_C (see section 2 of [15]). If the path (D_t) misses the stratum of operators with 2-dimensional kernel and meets the stratum with one-dimensional kernel transversely, then the contribution of C to $Gr(\alpha)$ is given by -1 raised to a power equal to the number of times it meets this latter stratum; more generally the contribution is found by orienting the zero-dimensional space $\ker D_1$ so that the corresponding orientation of $\det(D_1) = \Lambda^{max}\ker D_1 = \Lambda^{max}\ker D_1 \otimes (\Lambda^{max}\operatorname{coker} D_1)^*$ agrees with the natural orientation of the bundle $\bigcup_t \det(D_t)\times\{t\}$ which restricts to $t = 0$ as the complex orientation of $\det(D_0)$ (since j_0 is integrable, one has

$$D_0\xi = \frac{1}{2}(\nabla\xi + j(u)\nabla\xi\circ i) + \frac{1}{8}N_{j_0}(\partial_j u, \xi) = \frac{1}{2}(\nabla\xi + j(u)\nabla\xi\circ i) \qquad (3.1)$$

where $u\colon (\Sigma, i)\to X$ is an embedding of C, ∇ is a j–hermitian connection, and N is the Nijenhuis tensor, using remark 3.3.1 of [8]. D_0 therefore commutes with j_0, giving $\det(D_0)$ a natural (complex) orientation).

As for $\mathcal{DS}$, if J is a smooth regular almost complex structure on $X_r(f)$ and $s\in\mathcal{MS}^{J,\Omega}_{X_r(f)}(c_\alpha)$, the contribution of s to $\mathcal{DS}_{(X,f)}(\alpha)$ is similarly obtained by the spectral flow. Owing to the tautological correspondence, we would prefer to replace this smooth J with the almost complex structure $\mathbb{J}_j$. In general this is problematic because of the nondifferentiability of $\mathbb{J}_j$, but let us suppose for a moment that we have found some way to get around this, by choosing j as in Proposition 2.3. $\mathbb{J}_j$ is then smooth and nondegenerate (ie, the linearization of $\bar\partial_{\mathbb{J}_j}$ is surjective) at each of the sections in the set $\mathcal{MS}^{\mathbb{J}_j,\Omega}_{X_r(f)}(c_\alpha)$ of $\mathbb{J}_j$–holomorphic sections descending to curves which pass through Ω, which makes the following simple observation relevant.

Proposition 3.1 *Assume J is an almost complex structure on $X_r(f)$ which is Hölder continuous globally and smooth and nondegenerate at each member s of $\mathcal{MS}^{J,\Omega}_{X_r(f)}(c_\alpha)$. Then $\mathcal{DS}_{(X,f)}(\alpha)$ may be computed as the sum of the spectral flows of the linearizations of $\bar\partial_J$ at the sections s.*

Proof If J were globally smooth this would just be the definition of $\mathcal{DS}$. As it stands, we can find a sequence of smooth almost complex structures J_n agreeing with J on an open subset U of its smooth locus which contains the images of all members of $\mathcal{MS}^{J,\Omega}_{X_r(f)}(c_\alpha)$ such that J_n converges to J in Hölder norm. According to [12], Gromov compactness holds assuming only Hölder convergence of the almost complex structures, so since there are no sections in $\mathcal{MS}^{J,\Omega}_{X_r(f)}(c_\alpha)$ meeting $X_r(f)\setminus\bar U$, for large enough n there must not be any sections in $\mathcal{MS}^{J_n,\Omega}_{X_r(f)}(c_\alpha)$ meeting that region either. But then since J_n agrees with J on U, we must have $\mathcal{MS}^{J_n,\Omega}_{X_r(f)}(c_\alpha) = \mathcal{MS}^{J,\Omega}_{X_r(f)}(c_\alpha)$. Moreover, the spectral flow for a J'–holomorphic section s depends only on the restriction of J' to a neighborhood of s, so since J and J_n agree near all members of $\mathcal{MS}^{J,\Omega}_{X_r(f)}(c_\alpha)$, they will both give the same spectral flows. Using J_n to compute $\mathcal{DS}$ then proves the proposition. $\square$

Assuming then that we can contrive to use the almost complex structure $\mathbb{J}_j$ to compute $\mathcal{DS}$, we would like to arrange that the spectral flows for j on the disc normal bundle and for $\mathbb{J}_j$ on the disc bundle in $s_C^* T^{vt} X_r(f)$ correspond in some natural way. Now since D_0 on $U_C \subset X$ comes from a complex structure which does not preserve the fibers of f (rather, it preserves the fibers of the normal bundle) and so does not naturally correspond to any complex structure on a neighborhood of $Im(s_C)$ in $X_r(f)$, this at first seems a tall order. However, the key observation is that rather than starting the spectral flow at D_0 we can instead start it at the $\bar{\partial}$ operator $\tilde{D}$ corresponding to any integrable complex structure $\tilde{j}$ on U_C. Indeed, if j_t is a path of (not-necessarily integrable) almost complex structures from j_0 to $\tilde{j}$ then the operators $D_t\xi = \frac{1}{2}(\nabla^t\xi + j(u)\nabla^t\xi \circ i)$ (∇^t being a j_t–Hermitian connection) form a family of complex linear operators which by (3.1) agree at the endpoints with D_0 and $\tilde{D}$, so the complex orientation of $\bigcup \det(D_t) \times \{t\}$ agrees at the endpoints of D_0 and $\tilde{D}$. So by taking the path used to find the contribution of C to Gr to have $D_{1/2} = \tilde{D}$, the orientation induced on $\det(D_1)$ by $\bigcup_{t\in[0,1]} \det(D_t) \times \{t\}$ and the complex orientation on $\det(D_0)$ is the same as that induced by $\bigcup_{t\in[1/2,1]} \det(D_t) \times \{t\}$ and the complex orientation of $\det(D_{1/2}) = \det(\tilde{D})$.

The upshot is that for both Gr and $\mathcal{DS}$ we can obtain the contribution of a given curve (or section) by starting the spectral flow at any complex structure which is integrable on a neighborhood of the curve (or section) and makes the curve (or section) holomorphic. By Proposition 2.2, if $\tilde{j}$ makes f pseudo-holomorphic and is integrable on an open set $U \subset X$ then the corresponding almost complex structure $\mathbb{J}_{\tilde{j}}$ is integrable on the corresponding neighborhood in $X_r(f)$. So if we can take (j,Ω) to belong to the set $\mathcal{S}^1$ of Proposition 2.3 (a set we have not yet shown to be nonempty), we can hope to have the spectral flows correspond if we can find an almost complex structure $\tilde{j}$ integrable on a neighborhood of any given member C of $\mathcal{M}^{j,\Omega}_X(\alpha)$ which makes both C and f holomorphic. We will see later on that given such a $(j,\Omega) \in \mathcal{S}^1$, constructing $\tilde{j}$ is fairly easy, so we turn now to the task of replacing our original pair (j,Ω), assumed to be as in Lemma 2.1, by a pair belonging to $\mathcal{S}^1$.

Accordingly, let $C \in \mathcal{M}_X^{j,\Omega}(\alpha)$ where $(j,\Omega) \in \mathcal{S}^0$, and let $u\colon \Sigma \to X$ be an embedding of C. Restrict attention to a small neighborhood U of a branch point p of C; note that by Condition 3 of Lemma 2.1, U may be taken small enough to miss all of the other curves in $\mathcal{M}_X^{j,\Omega}(\alpha)$; also, as is shown in the proof of that Lemma, U can be taken small enough to miss Ω. Let w be a j–holomorphic coordinate on the fibers, and let z be the pullback of the holomorphic coordinate on the base S^2, translated so that p has coordinates $(0,0)$. Then j is determined by giving a function b such that the anti-holomorphic tangent space for j is

$$T_j^{0,1} = \langle \partial_{\bar{z}} + b(z,w)\partial_w, \partial_{\bar{w}} \rangle \tag{3.2}$$

From Equation 2.3, a complex structure defined by such an expression is integrable exactly when $b_{\bar{w}} \equiv 0$.

In general, we cannot hope to realize our initial goal of finding an almost complex structure making both f and C holomorphic which is integrable on a neighborhood of C. The problem may be explained as follows. If our almost complex structure is to have the form (3.2), the condition that C be holomorphic determines $b|_C$ uniquely. In regions not containing any points of $crit(f|_C)$ this doesn't create a problem, since at least after shrinking the region so that each connected component of its intersection with any fiber contains only one point of C, $b|_C$ can be extended to the region arbitrarily, say by prescribing b to be locally constant on each fiber. When C is tangent to the fiber $\{w = 0\}$ at $(0,0)$, though, we have that $\partial_{\bar{w}} \in T_{(0,0)}C \otimes \mathbb{C}$, and so $b_{\bar{w}}(0,0)$ is determined by $b|_C$ (which is in turn determined by C).

More concretely, assuming the tangency between C and the fiber at $(0,0)$ to be of second order, we can write $C = \{z = g(w)\}$ where, after scaling w, g is a function of form $g(w) = w^2 + O(3)$. A routine computation shows that for C to be holomorphic with respect to an almost complex structure defined by (3.2), we must have

$$b(g(w), w) = \frac{-g_{\bar{w}}}{|g_w|^2 - |g_{\bar{w}}|^2} \tag{3.3}$$

from which one finds by using the Taylor expansion of g to Taylor-expand the right-hand side that $b_{\bar{w}}(0,0) = -\frac{1}{8} g_{w\bar{w}\bar{w}\bar{w}}(0)$, which has no a priori reason to be zero.

Evidently, then, in order to construct an almost complex structure $\tilde{j}$ as above, or even to find a pair $(j_1, \Omega) \in \mathcal{S}^1$, so that j_1 is integrable in neighborhoods of all of the branch points of all of the curves in $\mathcal{M}_X^{j_1,\Omega}(\alpha)$, we will have to move the j–holomorphic curves C. We show now how to arrange to do so.

Let j, Ω, C, u, p, and U be as above. We will construct almost complex structures j_ϵ which are integrable on increasingly small neighborhoods of p and the linearization of whose $\bar{\partial}$ operators (considered as acting on sections of the normal bundle $N = N_C = N_C X$) are increasingly close to the linearization of $\bar{\partial}_j$. For the latter condition one might initially expect that the j_ϵ would need to be C^1–close to j, which the above considerations indicate would be impossible in the all-too-likely event that $b_{\bar{w}}(0,0) \neq 0$. However, the only directional derivatives of the complex structure which enter into the formula for the linearization are those in the direction of the section being acted on, so since normal vectors of C near p have small vertical components the disagreement between the vertical derivatives of j_ϵ and j will turn out not to pose a problem.

To begin, we fix r and ϵ_0 such that the set

$$D^z_{3r} \times D^w_{3\epsilon_0} := \{(z,w) \mid |z| < 3r, |w| < 3\epsilon_0\}$$

is disjoint from all curves of $\mathcal{M}^{j,\Omega}_X(\alpha)$ except for C. Let $\beta(z)$ (resp. $\chi(w)$) be a cutoff function which is 1 on D^z_r (resp. $D^w_{\epsilon_0}$) and 0 outside D^z_{2r} (resp. $D^w_{2\epsilon_0}$). Let

$$C_0 = \sup\{|\nabla\beta|, |\nabla\chi|/\epsilon_0\}$$

(so we can certainly take $C_0 \leq \max\{2/r, 2\}$). Where

$$T^{0,1}_j = \langle \partial_{\bar{z}} + b(z,w)\partial_w, \partial_{\bar{w}} \rangle$$

for each $\epsilon < \epsilon_0$ we define almost complex structures j_ϵ by

$$T^{0,1}_{j_\epsilon} = \langle \partial_{\bar{z}} + b_\epsilon(z,w)\partial_w, \partial_{\bar{w}} \rangle \tag{3.4}$$

where

$$b_\epsilon(z,w) = \beta(z)\chi\left(\frac{\epsilon_0 w}{\epsilon}\right) b(z,0) + \left(1 - \beta(z)\chi\left(\frac{\epsilon_0 w}{\epsilon}\right)\right) b(z,w)$$

So within the region $D^z_r \times D^w_\epsilon$ we have $(b_\epsilon)_{\bar{w}} \equiv 0$, meaning that j_ϵ is integrable, while outside the region $D^z_{2r} \times D^w_{2\epsilon}$ j_ϵ agrees with j. Further,

$$|b(z,w) - b_\epsilon(z,w)| = |\beta(z)\chi(\epsilon_0 w/\epsilon)(b(z,w) - b(z,0))| \leq 2\epsilon \|b\|_{C^1} \tag{3.5}$$

(since the expression is zero for $|w| > 2\epsilon$),

$$\begin{aligned} |\nabla_z(b - b_\epsilon)| &\leq |\nabla_z\beta||\chi(\epsilon_0 w/\epsilon)(b(z,w) - b(z,0))| \\ &\qquad + \beta(z)\chi(\epsilon_0 w/\epsilon)|\nabla_z(b(z,w) - b(z,0)| \\ &\leq 2C_0\epsilon\|b\|_{C^1} + 2\epsilon\|b\|_{C^2} \end{aligned} \tag{3.6}$$

and

$$\begin{aligned} |\nabla_w(b - b_\epsilon)| &\leq |\nabla_w\chi(\epsilon_0 w/\epsilon)||\beta(z)(b(z,w) - b(z,0))| + \beta(z)\chi(\epsilon_0 w/\epsilon)|\nabla_w b(z,w)| \\ &\leq \frac{C_0}{\epsilon} 2\epsilon\|b\|_{C^1} + \|b\|_{C^1} = (2C_0 + 1)\|b\|_{C^1} \end{aligned} \tag{3.7}$$

C is tangent to $\{w = 0\}$ at $(0,0)$, so after scaling z we can write C as $\{z = w^n + O(n+1)\}$ for some $n > 1$. It follows that there is a constant C_1 such that if ξ is a normal vector to C based at $(z,w) \in C$ then $|\xi^{vt}| \leq C_1|w|^{n-1}|\xi|$. Hence since $|\partial_\xi(b_\epsilon - b)| = 0$ if $|w| > 2\epsilon$, equations (3.6) and (3.7) give that

$$\begin{aligned} |\partial_\xi(b_\epsilon - b)| &\leq |\xi^{hor}||\nabla_z(b_\epsilon - b)| + |\xi^{vt}||\nabla_w(b_\epsilon - b)| \\ &\leq 2(C_0\|b\|_{C^1} + \|b\|_{C^2})\epsilon|\xi| + (2C_0+1)\|b\|_{C^1}C_1(2\,\epsilon)^{n-1}|\xi| \end{aligned} \quad (3.8)$$

We summarize what we have found in:

Lemma 3.2 *The almost complex structures given by (3.4) are integrable in $D_r^z \times D_\epsilon^w$ and agree with j outside $D_{2r}^z \times D_{2\epsilon}^w$. Further there is a constant C_2 depending only on j and the curve C such that $\|j_\epsilon - j\|_{C^0} \leq C_2\epsilon$ and, for any vector ξ normal to C, $|\partial_\xi j_\epsilon - \partial_\xi j| \leq C_2\epsilon|\xi|$.*

Now for any almost complex structure J on X, the linearization of $\bar{\partial}_J$ at a map $u\colon (\Sigma, i) \to (X,J)$ is given by

$$D_u^J\xi = \frac{1}{2}(\nabla^J\xi + J(u)\circ\nabla^J\xi\circ i) + \frac{1}{2}(\nabla_\xi^J J)(u)\partial_J(u)\circ i$$

where ∇^J is the Levi–Civita connection of the metric associated to J (this is equation 3.2 of [8]; they view D as acting on sections of u^*TX, but we may equally well view it as a map $\Gamma(u^*N_C) \to \Gamma(u^*N_C \otimes T^{0,1}C)$, as in [15]). Now the difference between ∇^{j_ϵ} and ∇^j is controlled by the C^0 norm of $j_\epsilon - j$, as is $\partial_{j_\epsilon}(u) - \partial_j(u)$, so in the only terms in which the derivatives of j_ϵ and j come into play in $(D_u^{j_\epsilon} - D_u^j)\xi$, the complex structure is being differentiated in the direction ξ. Lemma 3.2 thus implies:

Corollary 3.3 *There is a constant C_3 such that the linearizations*

$$D_u^{j_\epsilon}, D_u^j\colon W^{1,p}(u^*N_C) \to L^p(u^*N_C \otimes T^{0,1}C)$$

obey $\|D_u^{j_\epsilon}\xi - D_u^j\xi\|_{L^p} \leq C_3\epsilon\|\xi\|_{W^{1,p}}$.

Now let D^ϵ denote the operator

$$D_u^{j_\epsilon} \oplus (ev_\Omega)_*\colon W^{1,p}(u^*N_C) \to L^p(u^*N_C \otimes T^{0,1}C) \oplus \bigoplus_{q\in\Omega} T_qX$$

and likewise $D = D_u^j \oplus (ev_\Omega)_*$. D and all of the D^ϵ are then Fredholm of index zero, and j being nondegenerate in the sense of Taubes [15] amounts to the statement that D is surjective and hence has a two-sided (since $\mathrm{ind}(D) = 0$) bounded inverse, which we denote Q.

Lemma 3.4 *Let $\epsilon_n \to 0$ and let ξ_n be a bounded sequence in $W^{1,p}(u^*N_C)$ with $D^{\epsilon_n}\xi_n \to 0$. Then $\xi_n \to 0$.*

Proof The proof is based on the elliptic estimate

$$\|\xi\|_{W^{1,p}} \leq c(\|D^j_u\xi\|_{L^p} + \|\xi\|_{L^p}) \tag{3.9}$$

(for this estimate, see Lemma B.4.6 in [8], for example). Where ϵ_n, ξ_n are as in the hypothesis, we have

$$\begin{aligned}\|\xi_n - \xi_m\|_{W^{1,p}} &\leq c\big(\|D^j_u\xi_n - D^j_u\xi_m\|_{L^p} + \|\xi_n - \xi_m\|_{L^p}\big)\\ &= c\Big(\|(D^j_u - D^{j_{\epsilon_n}}_u)\xi_n - (D^j_u - D^{j_{\epsilon_m}}_u)\xi_m + D^{j_{\epsilon_n}}_u\xi_n - D^{j_{\epsilon_m}}_u\xi_m\|_{L^p}\\ &\qquad + \|\xi_n - \xi_m\|_{L^p}\Big)\\ &\leq c\big(C_3(\epsilon_n\|\xi_n\|_{W^{1,p}} + \epsilon_m\|\xi_m\|_{W^{1,p}}) + \|D^{j_{\epsilon_n}}_u\xi_n\|_{L^p} + \|D^{j_{\epsilon_m}}_u\xi_m\|_{L^p}\\ &\qquad + \|\xi_n - \xi_m\|_{L^p}\big)\end{aligned} \tag{3.10}$$

Now since $\{\xi_n\}$ is a bounded sequence in $W^{1,p}$, by Rellich compactness it has a subsequence which is Cauchy in L^p, and this fact along with the hypothesis of the lemma imply that, after passing to a subsequence, the right hand side tends to zero as $m, n \to \infty$. $\{\xi_n\}$ is therefore in fact Cauchy in $W^{1,p}$; say $\xi_n \to \xi$. Then

$$D\xi = (D - D^{\epsilon_n})\xi + D^{\epsilon_n}(\xi - \xi_n) + D^{\epsilon_n}\xi_n \to 0$$

by Corollary 3.3 and the facts that $\xi_n \to \xi$ and $D^{\epsilon_n}\xi_n \to 0$. But D is injective, so $\xi = 0$. So the ξ_n have a subsequence converging to zero. If the entire sequence did not converge to zero, we could take a subsequence bounded away from zero and apply the argument to that subsequence, obtaining a contradiction which proves the lemma. □

Corollary 3.5 (i) *There is $\epsilon_1 > 0$ such that D^ϵ is bijective for all $\epsilon < \epsilon_1$.*

(ii) *Denoting $Q^\epsilon = (D^\epsilon)^{-1}$, for any sequence $\epsilon_n \to 0$ we have $\|Q^{\epsilon_n} - Q\| \to 0$.*

Proof If (i) were false we could find $\epsilon_n \to 0$ and ξ_n with $\|\xi_n\|_{W^{1,p}} = 1$ and $D^{\epsilon_n}\xi_n = 0$. This is prohibited by Lemma 3.4.

For (ii), were this not the case for some sequence $\{\epsilon_n\}$, we could find η_n with L^p norm 1 such that $Q^{\epsilon_n}\eta_n - Q\eta_n \nrightarrow 0$. But then

$$\begin{aligned}\|D^{\epsilon_n}(Q^{\epsilon_n}\eta_n - Q\eta_n)\|_{L^p} &= \|D^{\epsilon_n}Q^{\epsilon_n}\eta_n + (D - D^{\epsilon_n})Q\eta_n - DQ\eta_n\|_{W^{1,p}}\\ &= \|\eta_n + (D - D^{\epsilon_n})Q^{\eta_n} - \eta_n\|_{W^{1,p}} \leq C_3\|Q\|\epsilon_n \to 0\end{aligned}$$

violating Lemma 3.4 (with $\xi_n = Q^{\epsilon_n}\eta_n - Q\eta_n$) once again. □

Corollary 3.5 (ii) in particular implies that there is $\epsilon_2 < \epsilon_1$ such that if $\epsilon < \epsilon_2$ then $\|Q^\epsilon\| \leq \|Q\|+1$ (for otherwise we could find $\epsilon_n \to 0$ with $\|Q^{\epsilon_n} - Q\| \geq 1$). Note that in general, where $u\colon (\Sigma, i) \to X$ denotes the (fixed) embedding of C, we have $\bar{\partial}_{j_\epsilon} u = \bar{\partial}_{j_\epsilon} u - \bar{\partial}_j u = \frac{1}{2}(j_\epsilon - j) \circ du \circ i$, so since $\|j_\epsilon - j\|_{C^0} \leq C_2 \epsilon$ and $j_\epsilon = j$ outside $D^z_{2r} \times D^w_{2\epsilon}$ (a region whose intersection with C has area proportional to ϵ^2), we have, for some constant C_4 related to C_2 and $\|du\|_{L^\infty}$, a bound

$$\|\bar{\partial}_{j_\epsilon} u\|_{L^p} \leq C_4 \epsilon^{1+2/p} \tag{3.11}$$

for $p > 2$. Fix such a p. This puts us into position to prove:

Lemma 3.6 *There are constants C_5 and $\epsilon_3 > 0$ such that for $\epsilon < \epsilon_3$ there exists $\eta_\epsilon \in L^p(u^* N_C \otimes T^{0,1} C) \oplus \bigoplus_{q \in \Omega} T_q X$ such that $\bar{\partial}_{j_\epsilon}(\exp_u(Q^\epsilon \eta_\epsilon)) = 0$ and $\|Q^\epsilon \eta_\epsilon\|_{W^{1,p}} \leq C_5 \epsilon^{1+2/p}$.*

Proof This is a direct application of Theorem 3.3.4 of [8] (whose proof adapts without change to the case where the domain and range consist of sections of $u^* N_C$ rather than $u^* TX$). In McDuff and Salamon's notation we take $c_0 = \max\{\|Q\| + 1, \|du\|_{L^p}, vol(\Sigma)\}$ and $\xi = 0$. The theorem gives δ and c independent of ϵ such that if $\|Q^\epsilon\| \leq c_0$ (as we have arranged to be the case for $\epsilon < \epsilon_2$) and $\|\bar{\partial}_{j_\epsilon} u\|_{L^p} \leq \delta$ then there is η_ϵ with $\bar{\partial}_{j_\epsilon}(\exp_u(Q^\epsilon \eta_\epsilon)) = 0$ and $\|Q^\epsilon \eta_\epsilon\| \leq \|\bar{\partial}_{j_\epsilon} u\|_{L^p}$, so we simply take $\epsilon_3 < \epsilon_2$ so small that $C_4 \epsilon_3^{1+2/p} \leq \delta$ and then $C_5 = cC_4$ □

For $\epsilon < \epsilon_3$, let $\xi_\epsilon = Q^\epsilon \eta_\epsilon$ and $u_\epsilon = \exp_u \xi_\epsilon$. We need to consider how the branch points of the curve $C_\epsilon = u_\epsilon(\Sigma)$ relate to those of C. Our intent is to carry out this construction sequentially for every branch point of C: at each step in the procedure, then, we replace j by an almost complex structure which is integrable in some neighborhood of the branch point under consideration, which has the effect of moving the curve somewhat; we may assume inductively that at each of the previous steps our procedure has resulted in the branch points being considered getting replaced by branch points p' contained in some neighborhood U' on which the new almost complex structure is integrable. For the present step, we need to ensure that two things hold when ϵ is sufficiently small:

(i) That the branch points q of C_ϵ that are not close to p are close enough to other branch points p' of C that if the neighborhood U' as above (on which j and so also j_ϵ is integrable) has already been constructed around p', then $q \in U'$; and

(ii) That the branch points of C_ϵ which are close to p fall into the neighborhood $D^z_r \times D^w_\epsilon$ on which j_ϵ is integrable.

The first statement is somewhat easier, since every j_ϵ agrees with j outside $D^z_{2r} \times D^w_{2\epsilon_0}$, and so where V is a small neighborhood of $D^z_{2r} \times D^w_{2\epsilon_0}$ it follows from elliptic bootstrapping that on $\Sigma \setminus u^{-1}(V)$ the $W^{1,p}$ bound on ξ_ϵ implies C^k bounds for all k. Now all branch points p' of C other than p lie in V, so for any such p', since $f \circ u_\epsilon$ is holomorphic and tends to $f \circ u$ in any C^k norm near p', for any neighborhood U' of $u(p')$, if ϵ is small enough U' will contain some number k of branch points $q_1, \ldots, q_k$ of C_ϵ such that, where n_q denotes the ramification index of a point q on the curve (equivalently, the order of tangency at q between the curve and the fiber), we have

$$\sum_m (n_{q_m} - 1) = n_{p'} - 1.$$

Conversely, at any $x \in \Sigma \setminus u^{-1}(V)$, the derivative of $f \circ u_\epsilon$ at x will be approximated to order $\epsilon^{1+2/p}$ by that of $f \circ u$ at x. In particular, if $u_\epsilon(x)$ is a branch point, ie if $(f \circ u_\epsilon)_*$ is zero at x, then $(f \circ u)_*(x) = O(\epsilon^{1+2/p})$, which if ϵ is small enough will force $u(x)$ (and so also the new branch point $u_\epsilon(x)$, which is a distance $O(\epsilon^{1+2/p})$ from $u(x)$) to be contained in any previously-specified neighborhood of the branch locus of C. This proves assertion (i) above.

Since the sum of the numbers $n_q - 1$ where q is a branch point of C_ϵ is the same as the corresponding number for C by the Hurwitz formula applied to the holomorphic maps $f \circ u_\epsilon$ and $f \circ u$, the sum of these numbers for just the branch points of C_ϵ contained in $D^z_{2r} \times D^w_{2\epsilon_0}$ must then $n_p - 1$, n_p being the ramification index of p as a branch point of C (for by what we've shown above, the sum of the $n_q - 1$ for q lying outside this set also has not been changed by replacing C with C_ϵ).

As such, p is replaced either by a single branch point of C_ϵ with ramification index n_p or by some collection of branch points (all in $D^z_{2r} \times D^w_{2\epsilon_0}$) each with ramification index strictly less than n_p. In the former case, in the usual coordinates (z, w) around p, since both j and j_ϵ preserve all of the fibers $\{z = const\}$, as in Section 2 of [7] we may write C as $\{z = w^{n_p} + O(n_p + 1)\}$ and C_ϵ as $\{z = z_0 + k(w - w_0)^{n_p} + O(n_p + 1)\}$ for some k, where (z_0, w_0) is the position of the new branch point. But from Lemma 3.6 and the Sobolev Embedding theorem we have an estimate $\|\xi_\epsilon\|_{C^{1-2/p}} \le K \epsilon^{1+2/p}$, which leads z_0, $k-1$, and w_0 to all be bounded by a constant times $\epsilon^{1+2/p}$. So if ϵ is small enough, the new node (z_0, w_0) will fall into the region $D^z_r \times D^w_\epsilon$ on which j_ϵ is integrable, thanks to the fact that $\epsilon^{1+2/p} \ll \epsilon$.

If instead p is replaced by distinct branch points with lower ramification indices, they in principle may not be so close, but then we can apply our construction near each of these new branch points. Because at each step we either succeed or lower the index, the process will eventually terminate (at the latest, when the index has been lowered to two).

We should note that at each stage of the process the moduli space only changes in the way that we have been anticipating. Namely, with the notation as above, we have:

Lemma 3.7 *Write* $\mathcal{M}_X^{j,\Omega}(\alpha) = \{[u],[v_1],\ldots,[v_r]\}$. *Then for* ϵ *sufficiently small,*

$$\mathcal{M}_X^{j_\epsilon,\Omega}(\alpha) = \{[u_\epsilon],[v_1],\ldots,[v_r]\}.$$

Proof That $\{[u_\epsilon],[v_1],\ldots,[v_r]\} \subset \mathcal{M}_X^{j_\epsilon,\Omega}(\alpha)$ is clear, since u_ϵ is j_ϵ–holomorphic and passes through Ω by construction (for it agrees with u on the u–preimages of all the points of Ω), and since the $Im(v_k)$ are all contained in the set on which j_ϵ agrees with j.

To show the reverse inclusion, assume to the contrary that there exists a sequence $\epsilon_n \to 0$ and $v_n\colon \Sigma_n \to X$ with $[v_n] \in \mathcal{M}_X^{j_{\epsilon_n},\Omega}(\alpha) \setminus \{[u_{\epsilon_n}],[v_1],\ldots,[v_r]\}$. Now the almost complex structures j_{ϵ_n} converge in the C^0 norm to j, so by Gromov compactness (generalized to the case of C^0 convergence of the almost complex structures by Theorem 1 of [6]), after passing to a subsequence there would be $[v] \in \mathcal{M}_X^{j,\Omega}(\alpha)$ with $[v_{\epsilon_n}] \to [v]$ in any $W^{1,p}$ norm. Now if $[v]$ were one of the $[v_k]$ this would of course be impossible, since the $[v_{\epsilon_n}]$ would then all eventually miss $D^z_{3r} \times D^w_{3\epsilon_0}$, so the $Im(v_{\epsilon_n})$ would be contained in the region where $j_{\epsilon_n} = j$, implying that the v_{ϵ_n} are j–holomorphic curves passing through Ω, which we assumed they were not.

So suppose $[v_{\epsilon_n}] \to [u]$ in C^0. Now $u_{\epsilon_n} = \exp_u \xi_n$ with $\|\xi_n\|_{W^{1,p}} \leq C_5\epsilon_n^{1+2/p}$, so $\|u_{\epsilon_n} - v_{\epsilon_n}\|_{W^{1,p}} \to 0$ as $n \to \infty$ for an appropriate parametrization of the v_{ϵ_n}. But, using the uniform boundedness of the right inverses Q^ϵ of the linearizations $D_u^{j_\epsilon}$ at u, Proposition 3.3.5 of [8] gives some δ such that $\|u_{\epsilon_n} - v_{\epsilon_n}\|_{C^0} \geq \delta$ for all n, a contradiction which proves the lemma. □

Lemma 3.7 and the facts noted before it now let us prove the following:

Theorem 3.8 *There is a constant* C_8 *such that for* ϵ *sufficiently small there exists an almost complex structure* $\tilde{j}_\epsilon$ *with* $\|\tilde{j}_\epsilon - j\|_{C^0} \leq C_8\epsilon$ *having the property that, where* $\mathcal{M}_X^{\tilde{j}_\epsilon,\Omega}(\alpha) = \{[u_1^\epsilon],\ldots,[u_r^\epsilon]\}$, $\tilde{j}_\epsilon$ *is integrable on a neighborhood of* *each point of* $crit(f|_{Im(u_i^\epsilon)})$*. Moreover* $\tilde{j}_\epsilon \in \mathcal{S}^0$*, and* $\mathbb{J}_{\tilde{j}_\epsilon}$ *is a regular almost complex structure on* $X_r(f)$*.*

Proof Our construction shows how to modify j into j_ϵ having the desired property in a small neighborhood of one branch point of one of the curves, say C, of $\mathcal{M}_X^{j,\Omega}(\alpha)$ without perturbing the other curves in $\mathcal{M}_X^{j,\Omega}(\alpha)$, and, as noted above, the construction can then be repeated at the other (slightly perturbed) branch points of C, moving C to a curve C' near *all* of the branch points of which our new almost complex structure has the desired property. Because the almost complex structure remains unchanged near the other curves, we can apply the same procedure sequentially to all of the curves of $\mathcal{M}_X^{j,\Omega}(\alpha)$; this entails only finitely many steps, at the end of which we obtain $\tilde{j}_\epsilon$, which is regular by construction.

If $\mathbb{J}_{\tilde{j}_\epsilon}$ is not already regular, Proposition 2.3 shows that it will become so after generic perturbations of $\tilde{j}_\epsilon$ supported away from the critical loci of the $f|_{Im(u_i^\epsilon)}$ and the points of Ω. Provided they are small enough, such perturbations will not change the other properties asserted in the theorem. □

Corollary 3.9 *In computing the invariant* $Gr(\alpha)$*, we can use an almost complex structure* j_1 *from the set* $\mathcal{S}^1$ *of Proposition 2.3, and in computing the invariant* $\mathcal{DS}_{(X,f)}(\alpha)$*, we can use the complex structure* $\mathbb{J}_{j_1}$*.*

4 Comparing the spectral flows

We now fix an almost complex structure j_1 as in Corollary 3.9, which we assume to have been constructed by the procedure in the preceding section. C will now denote a fixed member of $\mathcal{M}_X^{j_1,\Omega}(\alpha)$ with $u\colon (\Sigma, i) \to (X, j_1)$ a fixed embedding of C. The assumption on α at the start of the preceding section ensures that C will not have any components which are multiply covered square-zero tori; for more general α we now instead simply assume that this is true for C. We will show in this section that the contribution of C to $Gr(\alpha)$ is the same as that of the associated section s_C to $\mathcal{DS}_{(X,f)}(\alpha)$.

Lemma 4.1 *There is a neighborhood* U *of* C *and an integrable almost complex structure* $\tilde{j}$ *on* U *which makes both* f *and* C *holomorphic.*

Proof Let $Crit(f|_C) = \{p_1, \ldots, p_n\}$. By our construction of j_1, there are neighborhoods $V_1, \ldots, V_n$ of the p_k on which j_1 is given by

$$T^{0,1}_{j_1} = \langle \partial_{\bar{z}} + b(z,0)\partial_w, \partial_{\bar{w}} \rangle.$$

Since all of the branch points of C are contained within $\cup_k V_k$, we may cover $C \setminus \cup_k V_k$ by open sets U_α such that for each fiber $f^{-1}(t)$, $U_\alpha \cap f^{-1}(t)$ only contains at most one point of C. In each U_α, then, $C \cap U_\alpha$ is given as a graph

$$\{w_\alpha = \lambda_\alpha(z)\},$$

where w_α is a j_1–holomorphic coordinate on the fibers; in such coordinates $C \cap U_\alpha$ will be holomorphic with respect to an almost complex structure given by $T^{0,1} = \langle \partial_{\bar{z}} + b(z, w_\alpha)\partial_{w_\alpha}, \partial_{\bar{w}_\alpha}\rangle$ exactly if $b(z, \lambda_\alpha(z)) = \frac{\partial\lambda}{\partial\bar{z}}$. We therefore simply define $\tilde{j}_\alpha$ on U_α by

$$T^{0,1}_{\tilde{j}_\alpha} = \langle \partial_{\bar{z}} + \frac{\partial\lambda}{\partial\bar{z}}\partial_{w_\alpha}, \partial_{\bar{w}_\alpha}\rangle.$$

Geometrically, the $j_1|_{V_k}$ and the $\tilde{j}_\alpha$ are all uniquely determined by the fact that they restrict to the fibers as j_1, make C and f holomorphic, and have defining functions b which do not vary vertically, so in particular they agree on the overlaps of their domains and so piece together to form a complex structure $\tilde{j}$ on the set $U = \bigcup_k V_k \cup \bigcup_\alpha U_\alpha$, which is integrable by Equation 2.3 and so enjoys the properties stated in the lemma. □

Lemma 4.2 *Let $\mathcal{J}(U, f, C)$ denote the set of almost complex structures on U making both C and f holomorphic which are integrable near each branch point of C. Let $\mathcal{J}^{int}(U, f, C)$ be the subset of $\mathcal{J}(U, f, C)$ consisting of almost complex structures integrable near all of C. Then the maps*

$$\mathcal{F}\colon\ H^{0,1}_i(T_{\mathbb{C}}\Sigma) \times W^{1,p}(u^*TX) \times \mathcal{J}(U, f, C) \to L^p(u^*TX \otimes T^{0,1}\Sigma)$$
$$(\beta, \xi, j) \mapsto D^j_u\xi + \frac{1}{2}j \circ du \circ \beta,$$

$$\widehat{\mathcal{F}}\colon\ W^{1,p}(s_C^*T^{vt}X_r(f)) \times \mathcal{J}(U, f, C) \to L^p(s_C^*T^{vt}X_r(f) \otimes T^{0,1}S^2)$$
$$(\zeta, j) \mapsto D^{\mathbb{J}_j}_{s_C}\zeta,$$

$$\mathcal{F}'\colon\ H^{0,1}_i(T_{\mathbb{C}}\Sigma) \times W^{1,p}(u^*TX) \times \mathcal{J}^{int}(U, f, C) \to L^p(u^*TX \otimes T^{0,1}\Sigma)$$
$$(\beta, \xi, j) \mapsto D^j_u\xi + \frac{1}{2}j \circ du \circ \beta,$$

and

$$\widehat{\mathcal{F}'}\colon\ W^{1,p}(s_C^*T^{vt}X_r(f)) \times \mathcal{J}^{int}(U, f, C) \to L^p(s_C^*T^{vt}X_r(f) \otimes T^{0,1}S^2)$$
$$(\zeta, j) \mapsto D^{\mathbb{J}_j}_{s_C}\zeta$$

are each submersive at all zeros whose section component is not identically zero.

Proof Suppose $\mathcal{F}(\beta,\xi,j)=0$. The linearization of $\mathcal{F}$ at (β,ξ,j) is given by

$$\begin{aligned}\mathcal{F}_*(\gamma,\mu,y) &= D_u^j\mu + \left(\frac{d}{dt}\bigg|_{t=0} D_u^{exp_j(ty)}\right)\xi + \frac{1}{2}j\circ du\circ\gamma \\ &= D_u^j\mu + \frac{1}{2}(\nabla_\xi y)\circ du\circ i + \frac{1}{2}j\circ du\circ\gamma, \qquad (4.1)\end{aligned}$$

where ∇ is the Levi–Civita connection of the metric associated to j. We assume ξ is not identically zero, so that by Aronzajn's theorem it does not vanish identically on any open subset. If η were a nonzero element of coker $\mathcal{F}_*$, as in the usual argument find $x_0\in\Sigma$ with $u(x_0)\notin Crit(f|_C)$ and $\eta(x_0)$ and $\xi(x_0)$ both nonzero. Near $u(x_0)$, using the Levi–Civita connection of the metric associated to j, TX splits as $T^{vt}X\oplus TC$, and with respect to this splitting y (in order to be tangent to $\mathcal{J}(U,f,C)$) is permitted to have any block decomposition of form

$$y=\begin{pmatrix} a & b\\ 0 & 0\end{pmatrix} \qquad (4.2)$$

where all entries are j–antilinear and, in order that C remain holomorphic, $b|_C=0$, so $\nabla_\xi y$ can have any block decomposition of form $\begin{pmatrix} a' & b'\\ 0 & 0\end{pmatrix}$ where all entries are j–antilinear. We have $0\neq\eta(x_0)\in(u^*TX\otimes T^{0,1}\Sigma)_{x_0}$, and $u(x_0)\notin crit(f|_C)$, so $(\eta(x_0))^{vt}\neq 0$. Hence similarly to the proof of Lemma 2.1 we can take $b'(x_0)$ and $c'(x_0)$ such that

$$\begin{pmatrix} 0 & b'(x_0)\\ 0 & 0\end{pmatrix} du\circ i(v)=(\eta(x_0)(v))^{vt} \qquad \begin{pmatrix} 0 & b'(x_0)\\ 0 & 0\end{pmatrix} du\circ i(\bar v)=(\eta(x_0)(\bar v))^{vt}$$

where v generates $T^{1,0}_{x_0}\Sigma$. We then take y supported in a small neighborhood of $u(x_0)$ so that $a=0$ in the decomposition (4.2) and so that

$$(\nabla_\xi y)(x_0)=\begin{pmatrix} 0 & b'(x_0)\\ 0 & 0\end{pmatrix}$$

By taking the small neighborhood appropriately, unless the vertical projection $\eta^{vt}(x_0)$ of $\eta(x_0)$ is zero we can thus arrange that

$$\int\langle\eta,\mathcal{F}_*(0,y)\rangle\neq 0,$$

in contradiction with the assumption that η belonged to the cokernel of $\mathcal{F}_*$. This shows that any $\eta\in\operatorname{coker}\mathcal{F}_*$ must have η^{vt} identically zero. Then arguing just as in the proof of Lemma 2.1, we consider the projection η^C of η onto TC; once again η^C would give an element of the cokernel of the linearization at (i,id) of the map $(i',v)\mapsto\bar\partial_{i',i}v$ acting on pairs consisting of complex structures i' on Σ and maps $v\colon\Sigma\to\Sigma$, and the vanishing of this cokernel is just the statement that the space of complex structures on Σ is unobstructed at i. η^C is therefore also zero, so since TC and $T^{vt}X$ span TX at all but finitely many points of C, we conclude that η vanishes identically, proving the Lemma for $\mathcal{F}$.

The proof of the transversality of $\widehat{\mathcal{F}}$ proceeds in essentially the same way; if $\eta \in \mathrm{coker}(\widehat{\mathcal{F}}_*)_{(\zeta,j)}$ with $\widehat{\mathcal{F}}(\zeta,j)=0$ is nonzero at some t (which we can assume to be a regular value for $f|_C$), then as in the proof of Lemma 2.3, for at least one point p_0 among the r points of X appearing in the divisor $s_C(t)$, η descends to a nonzero $T^{vt}_{p_0}X$–valued cotangent vector at p_0, and we can use a perturbation y supported near p_0 similar to that above to obtain the desired contradiction.

As for $\mathcal{F}'$ and $\widehat{\mathcal{F}'}$, for which the almost complex structure is required to be integrable near C, the allowed perturbations y include anything in the block form

$$y = \begin{pmatrix} 0 & b \\ 0 & 0 \end{pmatrix}$$

where b varies holomorphically in the vertical variable w (as can be seen from Equation 2.3). So (aside from j–antilinearity) we only require that for any vertical vector ζ we have $\nabla_{j\zeta} b = j\nabla_\zeta b$. For a particular tangent vector ξ at $u(x_0)$, then, we still have the freedom to make $\nabla_\xi b$ any antilinear map that we choose, so we can just duplicate the proof of the submersivity of $\mathcal{F}$ and $\widehat{\mathcal{F}}$ to see that $\mathcal{F}'$ and $\widehat{\mathcal{F}'}$ are also submersive at all zeros where ξ is not identically zero. ◻

Corollary 4.3 *There is a neighborhood U of C and an integrable almost complex structure $\tilde{j}$ on U such that $\tilde{j}$ makes both f and C holomorphic, and such that the linearization $\mathcal{D}_u^{\tilde{j}}$ of the operator $(i,u)\mapsto \bar{\partial}_{i,\tilde{j}}u$ at the embedding of C is surjective, as is the linearization of $\bar{\partial}_{\mathbb{J}_{\tilde{j}}}$ at s_C*

Proof We have just shown that the map $\mathcal{F}'\colon H^{0,1}_i(T_{\mathbb{C}}\Sigma)\times (W^{1,p}(u^*TX)\setminus \{0\})\times \mathcal{J}^{int}(U,f,C)\to L^p(u^*TX\otimes T^{0,1}\Sigma)$ which sends (β,ξ,j) to $\mathcal{D}^j_u(\beta,\xi) = D^j_u\xi + \frac{1}{2}j\circ du\circ\beta$ is submersive at all zeros, so that the subset $\{(\beta,\xi,j) : \mathcal{D}^j_u(\beta,\xi)=0, \xi\not\equiv 0\}$ is a smooth manifold. As usual, applying the Sard–Smale theorem to the projection onto the second factor we obtain that for generic $j\in\mathcal{J}^{int}(U,f,C)$,

$$\ker\left((\beta,\xi)\mapsto D^j_u\xi + \frac{1}{2}j\circ du\circ\beta\right)\setminus\{0\} = \ker\mathcal{D}^j_u\setminus\{0\}$$

is a smooth manifold of the expected dimension. The correctness of the expected dimension for generic $j\in\mathcal{J}^{int}(U,f,C)$ of course translates directly to the surjectivity of the linearization $\mathcal{D}_u^j$ for such j. Likewise, the submersivity of $\widehat{\mathcal{F}'}$ shows that the linearization of $\bar{\partial}_{\mathbb{J}_{\tilde{j}}}$ at s_C is surjective for generic $j \in \mathcal{J}^{int}(U,f,C)$. So since Lemma 4.1 shows that $\mathcal{J}^{int}(U,f,C)$ is nonempty, the corollary follows. □

$\tilde{j}$ shall now denote an almost complex structure of the type obtained by Corollary 4.3.

Lemma 4.4 *There are paths j_t of almost complex structures on U connecting $j_0 := \tilde{j}$ to j_1 for which every j_t makes both f and C holomorphic. Moreover, for a dense set of such paths:*

(i) *The path j_t is transverse to the set of almost complex structures j for which the linearization D^j of the $\bar{\partial}_j$ operator at u (acting on normal sections) has excess kernel.*

(ii) *The path $\mathbb{J}_{j_t}$ of complex structures on the subset $\mathbb{U}$ of $X_r(f)$ corresponding to U is transverse to the set of almost complex structures J for which the linearization D^J of the $\bar{\partial}_J$ operator at s_C (acting on sections of $s_C^* T^{vt} X_r(f)$) has excess kernel.*

Proof In local coordinates near C, the almost complex structures j_1 and $\tilde{j}$ are given as

$$T^{0,1}_{j_1} = \langle \partial_{\bar{z}} + b_1(z,w)\partial_w, \partial_{\bar{w}} \rangle$$

and

$$T^{0,1}_{\tilde{j}} = \langle \partial_{\bar{z}} + \tilde{b}(z,w)\partial_w, \partial_{\bar{w}} \rangle.$$

Here we necessarily have $b_1|_C = \tilde{b}|_C$ since both j_1 and $\tilde{j}$ make C holomorphic, so to define a path j_t we can simply set

$$T^{0,1}_{j_t} = \langle \partial_{\bar{z}} + ((1-t)\tilde{b}(z,w) + tb_1(z,w))\partial_w, \partial_{\bar{w}} \rangle;$$

on each chart (this obviously pieces together to give an almost complex structure on all of C); since $(1-t)\tilde{b} + tb_1|_C = b_1|_C = \tilde{b}|_C$, C will be j_t–holomorphic for each t.

As for statements (i) and (ii), Lemma 4.2 implies that the map with domain

$$H^{0,1}_i(T_{\mathbb{C}}\Sigma) \times (W^{1,p}(u^* N_C) \setminus \{0\}) \times (W^{1,p}(s_C^* T^{vt} X_r(f)) \setminus \{0\}) \times \mathcal{J}(U,f,C)$$

defined by

$$(\beta, \xi, \zeta, j) \mapsto (\mathcal{D}_u^j \xi, D_{s_C}^{\mathbb{J}_j} \zeta)$$

is transverse to zero, so that its zero set is a smooth manifold and we obtain using the Sard–Smale theorem that, letting $\mathbb{U}$ refer to the connected component containing s_C in the open subset of $X_r(f)$ consisting of unordered r–tuples of points in $U \subset X$ that lie in the same fiber,

$$\mathcal{S}^1 = \{j \in \mathcal{J}(U,f,C) \mid (j,\Omega),(\mathbb{J}_j,\Omega) \text{ are nondegenerate on } U \text{ and } \mathbb{U} \text{ respectively}\}$$

is open and dense; here nondegeneracy of $(\mathbb{J}_j,\Omega)$ means that the direct sum $\mathbb{D}^j$ of $D^{\mathbb{J}_j}_{s_C}$ with the evaluation map that tautologically corresponds to $(ev_\Omega)_*$ is bijective, while as in [15] nondegeneracy of (j,Ω) means that $D^j_u \oplus (ev_\Omega)_*$ is bijective, which is implied for generic Ω by the surjectivity of $\mathcal{D}^j_u$.

Theorem 4.3.10 of [3] shows then that a dense set of paths from j_0 to j_1 consists of paths which only cross the locus for which either D^j or $D^{\mathbb{J}_j}$ has excess kernel transversely. (Alternately we could of course prove a parametrized version of Lemma 4.2 and apply the Sard–Smale theorem to the projection to the space of paths in $\mathcal{J}(U,f,C)$). □

Lemma 4.5 *For every $j \in \mathcal{J}(U,f,C)$ we have*

$$\ker(D^j_u \oplus (ev_\Omega)_*) = 0 \iff \ker \mathbb{D}^j = 0.$$

Proof Suppose that $\ker(D^j_u \oplus (ev_\Omega)_*) \neq 0$ and let $0 \neq \xi \in \ker(D^j_u \oplus (ev_\Omega)_*)$. $\xi \in W^{1,p} \subset C^0$, so for n sufficiently large $Im(\exp_u(\xi/n)) \subset U$. Let η_n be the sections of $s_C^* T^{vt} X_r(f)$ such that $\exp_{s_C} \eta_n$ tautologically corresponds to $\exp_u(\xi/n)$.

By the construction of $\mathbb{J}_j$, for any point t in the domain of s_C, $|\bar{\partial}_{\mathbb{J}_j}(\exp_{s_C}\eta_n)(t)|$ would be comparable to the maximum of the $|\bar{\partial}_j(\exp_u(\xi/n))|$ at the r points corresponding to $s_C(t)$, and similarly for $|\eta_n(t)|$ and the $|\xi/n|$ at the corresponding points, but for the fact that the end q of a normal vector based at a point $p_1 \in C$ will lie vertically over some other point $p_2 \in C$, which tends to increase distances as we pass to $X_r(f)$ since the (vertical) distance from p_2 to q will be larger than the length of the normal vector. However, for any compact subset K of $C \setminus crit(f|_C)$ normal vectors of small enough norm based at some $p_1 \in K$ will correspond to vertical vectors based at some p_2 lying not too far outside of K (and still outside of $crit(f|_C)$), and the norms of the normal vector and the associated vertical vector will be comparable by some constant (depending on the set K).

Since as $n \to \infty$, $\exp_u(\xi/n)$ approaches the embedding u of C, we can then conclude the following: given ϵ, let $V_\epsilon \subset C$ be the ϵ–neighborhood of $crit(f|_C)$

in C. Then there are N and $C_{1,\epsilon}, C_{2,\epsilon}, C_{3,\epsilon}, C_{4,\epsilon}$ such that for $n \geq N$ we have:

$$C_{1,\epsilon}\|\xi/n\|_{W^{1,p}(C\backslash V_{2\epsilon})} \leq \|\eta_n\|_{W^{1,p}(s_{C\backslash V_\epsilon})} \leq C_{2,\epsilon}\|\xi/n\|_{W^{1,p}(C\backslash V_{\epsilon/2})} \tag{4.3}$$

and

$$\begin{aligned} C_{3,\epsilon}\|\bar\partial_j \exp_u(\xi/n)\|_{L^p(C\backslash V_{2\epsilon})} &\leq \|\bar\partial_{\mathbb{J}_j}(\exp_{s_C}\eta_n)\|_{L^p(s_{C\backslash V_\epsilon})} \\ &\leq C_{4,\epsilon}\|\exp_u(\xi/n)\|_{L^p(C\backslash V_{\epsilon/2})} \end{aligned} \tag{4.4}$$

Now since $D^j\xi = 0$, there is a constant C_5 such that, for any ϵ, n we have

$$\|\bar\partial_j \exp_u(\xi/n)\|_{L^p(C\backslash V_\epsilon)} \leq C_5\|\xi/n\|^2_{W^{1,p}(C\backslash V_\epsilon)}$$

Also, by Aronzajn's theorem, ξ does not vanish on any open set, so writing $C_{6,\epsilon} = \frac{\|\xi\|_{W^{1,p}(C\backslash V_{\epsilon/2})}}{\|\xi\|_{W^{1,p}(C\backslash V_{2\epsilon})}}$, we have, independently of n,

$$\|\xi/n\|_{W^{1,p}(C\backslash V_{\epsilon/2})} \leq C_{6,\epsilon}\|\xi/n\|_{W^{1,p}(C\backslash V_{2\epsilon})}$$

We hence obtain, for all n

$$\begin{aligned} \|\bar\partial_{\mathbb{J}_j}(\exp_{s_C}\eta_n)\|_{L^p s_(C\backslash V_\epsilon)} &\leq C_{4,\epsilon}\|\exp_u(\xi/n)\|_{L^p(C\backslash V_{\epsilon/2})} \\ &\leq C_{4,\epsilon}C_5\|\xi/n\|_{W^{1,p}(C\backslash V_{\epsilon/2})} \\ &\leq C_{4,\epsilon}C_5C_{6,\epsilon}^2\|\xi/n\|^2_{W^{1,p}(C\backslash V_{2\epsilon})} \\ &\leq \frac{C_{4,\epsilon}C_5C_{6,\epsilon}^2}{C_{1,\epsilon}^2}\|\eta_n\|^2_{W^{1,p}(s_{C\backslash V_\epsilon})} \end{aligned}$$

So we have $W^{1,p}$ sections $\eta_n \to 0$ of $s_C^* T^{vt}X_r(f)$ such that, for each ϵ,

$$\frac{\|\bar\partial_{\mathbb{J}_j}(\exp_{s_C}\eta_n)\|_{L^p(s_{C\backslash V_\epsilon})}}{\|\eta_n\|_{W^{1,p}(s_{C\backslash V_\epsilon})}} \to 0 \tag{4.5}$$

We now show how to obtain from (4.5) an element of the kernel of the linearization $D^{\mathbb{J}_j}_{s_C}$.

Fix ϵ and consider the linearization D_ϵ of $\bar\partial_{\mathbb{J}_j}$ at $s_{C\backslash V_\epsilon}$, acting on $W^{1,p}$ sections of the bundle $E_\epsilon = s^*_{C\backslash V_\epsilon}T^{vt}X_r(f)$. Let $r_n\colon E_\epsilon \to E_\epsilon$ be the bundle endomorphism given by fiberwise multiplication by $\frac{1}{\|\eta_n\|_{W^{1,p}(s_{C\backslash V_\epsilon})}}$. Identifying a neighborhood of the zero section in E_ϵ with a neighborhood of $s_{C\backslash V_\epsilon}$, we have that, fixing k small enough that each $Im\left(\exp_{s_{C\backslash V_\epsilon}}\left(\frac{k\eta_n}{\|\eta_n\|_{W^{1,p}(s_{C\backslash V_\epsilon})}}\right)\right)$ is in this neighborhood (which is possible since the $\eta_n/\|\eta_n\|$ are C^0–bounded),

$$\bar\partial_{r_n^*\mathbb{J}_j}\left(\exp_{s_{C\backslash V_\epsilon}}\left(\frac{k\eta_n}{\|\eta_n\|_{W^{1,p}(s_{C\backslash V_\epsilon})}}\right)\right) = \frac{k}{\|\eta_n\|_{W^{1,p}(s_{C\backslash V_\epsilon})}}\bar\partial_{\mathbb{J}_j}\exp_{s_{C\backslash V_\epsilon}}\eta_n \to 0,$$

and each $\frac{k\eta_n}{\|\eta_n\|_{W^{1,p}(s_{C\setminus V_\epsilon})}}$ has norm k. Write $\zeta_n = \frac{k\eta_n}{\|\eta_n\|_{W^{1,p}(s_{C\setminus V_\epsilon})}}$.

Now since r_n is multiplication by $\frac{1}{\|\eta_n\|_{W^{1,p}(s_{C\setminus V_\epsilon})}}$, which tends to ∞ with n, we have that

$$\lim_{n\to\infty} D_\epsilon \zeta_n = \lim_{n\to\infty} \bar{\partial}_{r_n^*\mathbb{J}_j}(\exp_{s_{C\setminus V_\epsilon}} \zeta_n) = 0$$

By Rellich compactness, after passing to a subsequence the ζ_n L^p–converge to some $\zeta^\epsilon \in L^p$; since the ζ_n have norm bounded away from zero, $\zeta^\epsilon \neq 0$. Where D_ϵ^* is the formal adjoint of D_ϵ, we then have that, for each $\beta \in W^{1,q}(\Lambda^{0,1}M_P \otimes s^*_{C\setminus V_\epsilon} T^{vt}X_r(f))$ $(1/p+1/q=1)$,

$$\langle \zeta^\epsilon, D_\epsilon^*\beta\rangle = \lim_{n\to\infty}\langle \zeta_n, D_\epsilon^*\beta\rangle = \lim_{n\to\infty}\langle D_\epsilon\zeta_n, \beta\rangle = 0$$

So ζ^ϵ is a weak solution to $D_\epsilon\zeta_\epsilon = 0$; by elliptic regularity this implies that ζ^ϵ is in fact in $W^{1,p}$ with $D_\epsilon\zeta^\epsilon = 0$.

All of the ζ^ϵ so constructed agree up to scale on the overlaps of their domains (since they are limits of rescaled versions of the η_n, and the η_n do not vary with ϵ); also if we require that the tubular neighborhoods of $s_{C\setminus V_\epsilon}$ used in the construction are all contained in a common tubular neighborhood of s_C, the $\exp_{s_C}\zeta^\epsilon$ will all be contained in this neighborhood, so that the norms of the ζ^ϵ will be bounded, say by M, as $\epsilon \to 0$. So we can rescale the ζ^ϵ to all agree on their domains with a common section $\zeta \in W^{1,p}(s_C^* T^{vt}X_r(f))$ defined on the complement from the finite set of critical values of $f|_C$ which is nonzero (since all of the ζ^ϵ are) and has $D_\epsilon\zeta = 0$ for every $\epsilon > 0$. Moreover the norm of ζ on any compact subset of its domain is at most M, so by removal of singularities ζ extends to all of S^2, and $\zeta \in \ker D_{s_C}^{\mathbb{J}_j}$. Further, since $\xi \in \ker(ev_\Omega)_*$, it readily follows from the construction that ζ is in the kernel of the corresponding linearization of the corresponding evaluation map on $X_r(f)$, so that $0 \neq \zeta \in \ker\mathbb{D}^j$, proving the forward implication in the statement of Lemma 4.5.

The reverse implication can be proven in just the same way, by taking an element $0 \neq \eta \in \ker\mathbb{D}^j$ and extracting a normal section ξ from the curves tautologically corresponding to the $\exp(\eta/n)$ which lies in the kernel of the restriction of $(D_u^j \oplus (ev_\Omega)_*)$ to any set missing $crit(f|_C)$ Once again, removal of singularities then implies that ξ extends to give a global nonzero element of $\ker(D_u^j \oplus (ev_\Omega)_*)$. □

This directly yields the theorem promised at the beginning of the section.

Theorem 4.6 *The contribution of C to $Gr(\alpha)$ is the same as that of s_C to $\mathcal{DS}_{(X,f)}(\alpha)$.*

Proof Take a path j_t as in Lemma 4.4, so that j_t is transverse to the set of j for which either $D_u^j \oplus (ev_\Omega)_*$ or $\mathbb{D}^j$ has nonzero kernel. Since $N_{\tilde{j}} = 0$, we have $N_{\mathbb{J}_{\tilde{j}}} = 0$, so by the remarks at the start of Section 3 the contribution of C to Gr may be computed from the spectral flow of the path of operators $D_u^{j_t} \oplus (ev_\Omega)_*$, while that of s_C to $\mathcal{DS}$ may be computed from the spectral flow of the path $\mathbb{D}^{j_t}$. By Lemma 4.5, for every t the operator $D_u^{j_t} \oplus (ev_\Omega)_*$ has a kernel if and only if $\mathbb{D}^{j_t}$ does, so the number of eigenvalue crossings for positive t, each of which is known to be transverse, will be the same. The two contributions are then both equal to negative one to this common number of crossings. □

5 Multiple covers of square-zero tori

For curves with square-zero toroidal components, the difficulties involved in comparing the contributions to Gr and $\mathcal{DS}$ are more serious. On the Gr side, as Taubes showed in [15], if C is a j–holomorphic square-zero torus, not only C but also each of its multiple covers contributes to Gr, according to a prescription which depends on the spectral flows not only of the linearization D of the $\bar{\partial}$ operator on the normal bundle N_C but also of the three operators D_ι corresponding to D which act on sections of the bundle obtained by twisting N_C by the real line bundles with Stiefel–Whitney class ι. From the standpoint of the tautological correspondence, it is encouraging that multiple covers of square-zero tori contribute to Gr, since such covers do tautologically correspond to $\mathbb{J}_j$–holomorphic sections of $X_r(f)$ for appropriate r. These sections are more difficult to analyze, though, because they are contained in the diagonal stratum Δ of $X_r(f)$, so the problems stemming from the nondifferentiability of $\mathbb{J}_j$ cannot be evaded by modifying j to be integrable near the branch points.

Throughout this section, all almost complex structures j defined on some region of X that we consider will be assumed to make the restriction of f to that region pseudoholomorphic.

As in Definition 4.1 of [15], a j–holomorphic square-zero torus C will be called m–nondegenerate if, for each holomorphic cover $\tilde{C} \to C$ of degree at most m, the operator $\tilde{D}$ obtained by pulling back the linearization D (which acts on $\Gamma(u^*N_C)$ if u is the map of C into X) by the cover $\tilde{C} \to C$ has trivial kernel. j will be called m–nondegenerate for some fixed cohomology class $\alpha \in H^2(X,\mathbb{Z})$

with $\alpha^2 = \kappa \cdot \alpha = 0$ if every j–holomorphic curve C with $[C] = PD(\alpha)$ is m–nondegenerate. Lemma 5.4 of [15] shows that m–nondegeneracy is an open and dense condition on j.

For any integer m, if C is a j–holomorphic square-zero torus Poincaré dual to the class α, where j is m–nondegenerate and is as in Lemma 2.1, we can define the contribution $r'_j(C,m)$ of m–fold covers of C to $\mathcal{DS}_{(X,f)}(m\alpha)$ as follows. Take a small tubular neighborhood U of C which does not meet any of the other j–holomorphic curves Poincaré dual to any $k\alpha$ where $k \le m$ (this is possible since the nondegeneracy of j ensures that there are only finitely many such curves and since $\alpha^2 = 0$) and which misses the critical points of the fibration. Where r is the intersection number with the fibers of f, let $\mathbb{U}$ be the neighborhood of the section s_{mC} of $X_{mr}(f)$ tautologically corresponding to U, so $\mathbb{J}_j$ is Hölder continuous (say C^γ) on $\mathbb{U}$ and s_{mC} is the only $\mathbb{J}_j$–holomorphic section in its homotopy class which meets $\mathbb{U}$. Let V be an open set with closure contained in $\mathbb{U}$ and containing the image of s_{mC}; then it follows readily from Gromov compactness that there is $\epsilon > 0$ such that if J is any almost complex structure with $\|J - \mathbb{J}_j\|_{C^\gamma} < \epsilon$ then any J–holomorphic curve meeting $\mathbb{U}$ must in fact be contained in V. $r'_j(C,m)$ is then defined as the usual signed count of all J–holomorphic sections homotopic to s_{mC} and contained in V where J is a generic almost complex structure which is smooth on V and has $\|J - \mathbb{J}_j\|_{C^\gamma} < \epsilon$. The usual cobordism argument (using cobordisms which stay Hölder-close to $\mathbb{J}_j$ so that sections in the parametrized moduli spaces don't wander outside of V) shows that this count is independent of the choice of J. Similarly, for any $\beta \in H^2(X,\mathbb{Z})$, defining the contribution to $\mathcal{DS}_{(X,f)}(\beta)$ of any disjoint union of j–holomorphic curves with multiplicities with homology classes adding to $PD(\beta)$ by smoothing $\mathbb{J}_j$ near the associated section of $X_r(f)$, one notes that $\mathcal{DS}_{(X,f)}(\beta)$ is indeed the sum of all the contributions of all such unions, so the terminology is not misleading.

Note that this definition of the contribution of m–fold covers of C to $\mathcal{DS}$ makes sense even if C is itself a multiple cover. If C is a k–fold cover of C', then the section s_{lC} associated to an l–fold cover of C is just the same as the section $s_{klC'}$, and $r'_j(C,l)$ is defined by perturbing the almost complex structure on the relative Hilbert scheme near this section. In particular, we have $r'_j(C,l) = r'_j(C',kl)$.

Lemma 5.1 *Let j_t $(0 \le t \le 1)$ be a path of almost complex structures which make f holomorphic such that every j_t is m–non-degenerate, and let C_t be a path of embedded square-zero tori in X such that $\{(C_t,t) | 0 \le t \le 1\}$ is one of* *the connected components of the parametrized moduli space of j_t–holomorphic curves homologous to C_0. Then $r'_{j_0}(C_0,m)=r'_{j_1}(C_1,m)$.*

Proof Because all of the j_t are m–non-degenerate, there is an open neighborhood U of $\cup_t C_t\times\{t\}\subset X\times[0,1]$ such that no curve in homology class $k[C_t]$ for any $k\leq m$ meets U (for otherwise Gromov compactness would give either a j_t–holomorphic curve in class $k[C_t]$ meeting C_t in an isolated point, which is impossible since $[C_t]^2=0$, or a sequence of curves distinct from C_t which converge to a k–fold cover of C_t, which is prohibited by m–non-degeneracy). Where r is the intersection number of C_t with the fibers of f, let $\mathbb{U}$ be the neighborhood of $\cup_t Im(s_{mC})\times\{t\}$ tautologically corresponding to U and V some neighborhood of $\cup_t Im(s_{mC})\times\{t\}$ compactly contained in $\mathbb{U}$. Let J_t be a family of smooth almost complex structures on $X_{mr}(f)$ which are sufficiently Hölder-close to $\mathbb{J}_{j_t}$ that each J_t–holomorphic section meeting $\mathbb{U}$ is contained in V, taken so that J_0 and J_1 are both regular and the path J_t is suitably generic. Now $\{(s,t)|\,\bar{\partial}_{J_t}s=0\}$ of course gives an oriented cobordism between the moduli spaces of J_0 and J_1–holomorphic sections in the relevant homotopy class, and moreover, since none of the members of $\{(s,t)|\,\bar{\partial}_{J_t}s=0\}$ even meet the open set $\mathbb{U}\setminus\bar{V}$, this cobordism restricts to a cobordism between the set of J_0–sections contained in V and the set of J_1–sections contained in V. Since the $r'_{j_k}(C_k,m)$ $(k=0,1)$ are precisely the signed count of these sections, it follows that $r'_{j_0}(C_0,m)=r'_{j_1}(C_1,m)$. □

A major reason that the analysis of multiply-covered pseudoholomorphic curves is generally more difficult is that when multiply-covered curves are allowed the argument that is generally used to show the submersivity of the "universal map" $(u,j)\mapsto\bar{\partial}_j u$ breaks down. As a consequence, for instance, as far as the author can tell it is not possible to ensure that a square-zero torus C will admit any almost complex structures near it which both make it m–nondegenerate and are integrable if $m>1$. In the semi-positive context in which we presently find ourselves, the standard way to navigate around this difficulty, following [9] and [10], is to construct our invariants from solutions to the inhomogeneous Cauchy–Riemann equation

$$(\bar{\partial}_j u)(p)=\nu(p,u(p)),\tag{5.1}$$

where the domain of the map $u\colon\Sigma\to X$ is viewed as contained in a "good cover" of the universal curve $\bar{\mathcal{U}}_{g,n}$ which is itself embedded in some $\mathbb{P}^N$, and ν is a section of the bundle $Hom(\pi_1^*T\mathbb{P}^N,\pi_2^*TX)\to\mathbb{P}^N\times X$ which is antilinear with respect to the standard complex structure on $\mathbb{P}^N$ and the almost complex structure j on X (see Definitions 2.1 and 2.2 of [10] for details; note however in our case since we are counting curves which may not be connected, we need to replace $\bar{\mathcal{U}}_{g,n}$ with the universal space $\bar{\mathcal{U}}^{(m)}_{\chi,n}$ of curves with at most m components, n marked points, and total Euler characteristic χ). Solutions to this equation are called (j,ν)–holomorphic curves. ν is called an inhomogeneous term.

Imitating very closely the proof of Lemma 4.2, one can see that for any given $m \geq 1$ and for any fixed $(j,0)$–holomorphic curve C and for generic inhomogeneous terms ν which

(a) vanish along the graphs of the embedding u of C and of all of its covers up to degree m,

(b) take values in $T^{vt}X$ (rather than just TX),

(c) are "holomorphic in the X variable" in the sense that $\nabla_{(0,j\zeta)}\nu = j\nabla_{(0,\zeta)}\nu$ for $\zeta \in TX$ (and $(0,\zeta) \in T(\mathbb{P}^N \times X)$), and

(d) have the following "coherence" property: where $u\colon \Sigma \to X$ is embedding of C and $\phi'\colon \Sigma' \to \Sigma$ and $\phi''\colon \Sigma'' \to \Sigma$ are any two holomorphic, possibly disconnected, m–fold covers of Σ, for each $p \in \Sigma$ and each $x \in X$ close to $u(p)$ the unordered m–tuples $\{\nu(p',x) : \phi'(p') = p\}$ and $\{\nu(p'',x) : \phi''(p'') = p\}$ are the same,

all of the covers of C of degree m will be nondegenerate as (j,ν)–holomorphic curves (ie, the linearization of the equation (5.1) will be surjective at each of these covers). The point of condition (c) above is that it ensures that these linearizations are all complex linear if j is integrable near C. The point of condition (d) is that it ensures that there is an inhomogeneous term μ on $X_{mr}(f)$ such that the equation for a (j,ν)–holomorphic curve in class $m[C]$ near C is the same as the equation for a $(\mathbb{J}_j,\mu)$–holomorphic section of $X_{mr}(f)$ near s_{mC} which descends to a cycle in class $m[C]$. ν satisfying this condition may easily be constructed: any choice of m perturbation terms $\nu_1,\ldots,\nu_m \in \Gamma(\overline{Hom}(T\Sigma, u^*T^{vt}X))$ which vanish near the branch points of C can be assembled into perturbation terms near each of the holomorphic m–fold covers, and we can use cutoff functions to put these together in order to form a coherent inhomogeneous term $\nu \in \Gamma(\overline{Hom}(\pi_1^*T\mathbb{P}^N, \pi_2^*TX))$. Since the curves giving m–fold covers of Σ in $\bar{\mathcal{U}}^{(m)}_{\chi=0,n}$ are separated from each other, the coherence condition does not make the proof of generic nondegeneracy any more difficult. The reason that we can imitate the proof of Lemma 4.2 using inhomogeneous terms but not using almost complex structures is of course that we need the freedom to vary the linearization of the equation on individual small neighborhoods in the domain while leaving it unchanged elsewhere, and for, say, a k–fold cover, varying the almost complex structure on a small neighborhood in X has the effect of varying the linearization on k different neighborhoods of the domain all in the same way.

A pair (j, ν) such that ν satisfies conditions (b) through (d) with respect to all (j, ν)–holomorphic curves C will be called *admissible*. We will slightly enlarge the class of data we study as follows: instead of only considering pairs (C, j) where C is j–holomorphic, we consider triples (C, j, ν) where C is j–holomorphic, ν vanishes along the graphs of the embedding of C and all of its covers up to degree m, and (j, ν) is admissible; such a triple will be called m–nondegenerate if all of the covers of C of degree m or lower are nondegenerate as (j, ν)–holomorphic curves. The admissible pair (j, ν) will itself be called m–nondegenerate if (C, j, ν) is m–nondegenerate for each (j, ν)–holomorphic curve C. We can then define the contribution $r'_{j,\nu}(C, m)$ to $\mathcal{DS}$ if (C, j, ν) is m–nondegenerate: the nondegeneracy implies that there is a neighborhood U of C which does not meet any other (j, ν)–holomorphic curves in class $k[C]$ for $k \leq m$. We have a tautologically-corresponding inhomogeneous term μ on $X_{mr}(f)$, and we may perturb the almost complex structure $\mathbb{J}_j$ to a smooth almost complex structure J such that (J, μ) is nondegenerate on a neighborhood V of s_{mC} contained in the set tautologically corresponding to U; we then count (J, μ) holomorphic sections according to the prescription in [9]. (Gromov compactness in the context of solutions to the inhomogeneous Cauchy–Riemann equation is needed here; this result appears as Proposition 3.1 of [9].) The proof of Lemma 5.1 then goes through to show:

Corollary 5.2 *Let (j_t, ν_t) $(0 \leq t \leq 1)$ be a path of m–nondegenerate admissible pairs, and let C_t be a path of embedded square-zero tori in X such that $\{(C_t, t) | 0 \leq t \leq 1\}$ is one of the connected components of the parametrized moduli space of (j_t, ν_t)–holomorphic curves homologous to C_0. Then $r'_{j_0,\nu_0}(C_0, m) = r'_{j_1,\nu_1}(C_1, m)$.*

Now assume that (C, j, ν) is m–nondegenerate and that j is *integrable* near C. $\mathbb{J}_j$ is then smooth (and even integrable) near s_{mC}; the argument in the proof of Lemma 4.5 shows that $(\mathbb{J}_j, \mu)$ will then also be nondegenerate (and even if it weren't, it would become so after a suitable perturbation of ν among inhomogeneous terms satisfying conditions (a) through (d)), so in computing $r'_{j,\nu}(C, m)$ we don't need to perturb $\mathbb{J}_j$ at all. So since the linearization of the equation $\bar{\partial}_{\mathbb{J}_j} s = \mu$ at s_{mC} is complex-linear and since s_{mC} is the only solution to that equation in V, we obtain (using Corollary 5.2):

Lemma 5.3 *If (j,ν) is an admissible pair and C a j–holomorphic square-zero torus such that j is integrable near C, and if the m–non-degenerate pair (j',ν') with C j'–holomorphic is sufficiently close to j, then $r'_{j',\nu'}(C,m)=1$ for every m.*

Our basic strategy in proving that multiple covers of square-zero tori contribute identically to $\mathcal{DS}$ and Gr will be, using an almost complex structure j as in Corollary 3.9, to investigate how the contributions $r'_{j_t,\nu_t}(C,m)$ vary as we move among admissible pairs such that C is j_t–holomorphic along a path from an m–nondegenerate pair (j_0,ν_0) with j_0 integrable near C to the pair $(j,0)$ where j is the given nondegenerate almost complex structure. This requires a digression into the chamber structure of almost complex structures on X, which was investigated extensively by Taubes in [15]. For simplicity of exposition, we will generally work in the homogeneous context $\nu=0$ below; since the wall crossing results that follow only depend on the basic shape of the differential equations involved and their linearizations, the results below will remain valid when "j_t" is replaced by "(j_t,ν_t)."

Where $\mathcal{M}_{1,1}$ is the moduli space of smooth pointed complex tori, consider the bundle $\mathcal{G}\to\mathcal{M}_{1,1}$ whose fiber over the curve C is the set of 1–jets at C of almost complex structures on the trivial complex line bundle over C. Any such 1–jet gives rise to four linearizations D_ι of the $\bar{\partial}$ operator on the bundles $\underline{\mathbb{C}}\otimes L_\iota$ over C, where L_ι is the real line bundle over C with Stiefel–Whitney class $\iota\in H^1(C,\mathbb{Z}/2)$. Taubes shows that the set $\mathcal{D}_\iota$ of points of $\mathcal{G}$ whose corresponding linearization has a nontrivial kernel is a subvariety of real codimension at least 1, and that the set of elements of $\mathcal{D}_\iota$ either corresponding to a linearization with two-or-greater-dimensional kernel or belonging to some other $\mathcal{D}_{\iota'}$ has real codimension at least 2 in $\mathcal{G}$. Identical results apply when we instead take the fiber of $\mathcal{G}$ to consist of 1–jets of admissible pairs (j,ν).

A path $\gamma=(u_t,C_t,j_t)_{t\in[0,1]}$ of j_t–holomorphic immersions $u_t\colon C_t\to X$ (each C_t belonging to $\mathcal{M}_{1,1}$; more commonly we will just denote such paths by (C_t,j_t), suppressing the map and identifying C_t with its image in X) then gives rise to a path $\tilde{\gamma}$ in $\mathcal{G}$; we say γ crosses a wall at $t=t_0$ if $\tilde{\gamma}$ meets one of the codimension-one sets $\mathcal{D}_\iota$ transversely at t_0. (Note that it's not essential that the u_t be embeddings, and in fact the case where u_t is a double cover will be of some relevance later on). The path components of $\mathcal{G}\setminus\cup_\iota\mathcal{D}_\iota$ are called *chambers*. For any m, Part 5 of Lemma 5.8, Lemma 5.9, and Lemma 5.10 of [15] show (among other things) that for a generic path (C_t,j_t), the only t_0 for which j_{t_0} fails to be m–nondegenerate near C_{t_0} are those t_0 for which (C_{t_0},j_{t_0}) is on a wall. The proofs of the results concerning connectivity and regularity of almost complex structures which make f holomorphic from sections 2 through 4 may easily be modified to show that the corresponding statement is true for paths j_t generic among paths of almost complex structures which make f holomorphic. On a similar note, if a path (C_t, j_t), where each j_t is an almost complex structure which makes f holomorphic, remains in the same chamber except for one point at which it touches a wall, the arguments in the proofs of Lemmas 2.1 and 4.4 show that the path may be perturbed to a path which remains entirely within the chamber and for which the almost complex structure continues to make f holomorphic.

In general, with the convention that $r'_j(C,0) = 1$, we will organize the contributions $r'_j(C,m)$ into a generating function (to be viewed as a formal power series; we are not making any convergence assertions here)

$$P'_j(C,z) = \sum_{m \geq 0} r'_j(C,m) z^m.$$

Strictly speaking, this power series should be truncated after the term corresponding to the largest m for which j is m–non-degenerate and the fibration satisfies $\omega \cdot (fiber) > m\omega \cdot \alpha$. However, by working with suitably generic j and suitably high-degree Lefschetz fibrations given by Donaldson's construction, we can fix this m to be as large as we want at the start of the argument.

Proposition 5.4 *If $\alpha^2 = \kappa \cdot \alpha = 0$ and j is m–nondegenerate for each m under consideration, the total contribution of all disjoint unions of possibly-multiply-covered tori in classes proportional to $PD(\alpha)$ to the standard surface count $\mathcal{DS}_{(X,f)}(n\alpha)$ is the coefficient of z^n in the product*

$$\prod_k \prod_{C \in \mathcal{M}_X^{j,\varnothing}(k\alpha)} P'_j(C, z^k).$$

Proof Let C_i be j–holomorphic tori in class $k_i\alpha$, and write $r = \alpha \cdot (fiber)$. The contribution of a disjoint union of m_i–fold covers of the C_i to $\mathcal{DS}_{(X,f)}(\sum m_i k_i \alpha)$ may be found by using an almost complex structure J on $X_{\sum m_i k_i r}(f)$ obtained by pushing forward generic smooth almost complex structures J_i on the $X_{m_i k_i r}(f)$ via the "divisor addition" map $\prod S^{m_i k_i r}\Sigma_t \to S^{\sum m_i k_i r}\Sigma_t$. This is because the J–holomorphic sections will just be fiberwise sums of the J_i–holomorphic sections, which are in turn close to the sections s_{C_i}, and the C_i are assumed disjoint, so that J (which is smooth away from the diagonal) will be smooth near each J–holomorphic section, putting us in the situation of Proposition 3.1. We may then conclude that the total contribution of such a disjoint union of covers is $\prod_i r'_j(C, m_i)$, since J–holomorphic sections are obtained precisely by adding together J_i–holomorphic sections under the divisor addition map, and there are $\prod_i r'_j(C, m_i)$ ways to do this. Organizing these contributions into a generating function then yields the proposition. □

We now fix an embedding u of a square-zero torus C and consider paths j_t ($t \in (-\epsilon, \epsilon)$) of almost complex structures making u and f holomorphic. If (C, j_t) crosses a wall at $t = 0$ we would like to compare the $r'_{j_t}(C, m)$ for small negative values of t to those for small positive values. We note again that we are taking $\nu_t = 0$ for ease of exposition, but the following lemma and its proof go through unchanged to the case when we instead have a family (j_t, ν_t) of admissible pairs with (C, j_t, ν_t) crossing a wall just at $t = 0$.

Lemma 5.5 *Assume that (C, j_t) crosses the wall $\mathcal{D}_0$ at $t = 0$ and that the path j_t is generic among paths of almost complex structures making both C and f holomorphic. Then there is a path of j_t–holomorphic tori C_t such that:*

(1) *For each t the set of j_t–holomorphic tori homologous to C in a suitably small tubular neighborhood U of C is $\{C, C_t\}$.*

(2) $C_0 = C$

(3) *For $0 < |t| < \epsilon$, (C_t, j_{-t}) and (C, j_t) are connected by a path (C'_s, j'_s) with every j'_s making f holomorphic and every C'_s m–nondegenerate.*

Moreover, there are small regular perturbations j'_t of the path j_t supported near $t = 0$ with the property that there are no j'_0–holomorphic curves in any homology class $k[C]$ contained in U

Proof We mimic the argument on pp. 863–864 of [15]. Let D be the linearization of $\bar{\partial}_{j_0}$ at the embedding u of C. For small $|t| > 0$, the equation for a section v_t of N_C to have the property that $\exp_u v_t$ is j_t–holomorphic has the form

$$Dv_t + R(t, v_t, \nabla v_t) = 0 \tag{5.2}$$

where the Taylor expansion of R begins at order 2 (in the case considered in [15] there is an additional term proportional to t times the derivative with respect to t of the projection to N_C of the restriction of j_t to $T_{0,1}C$, but in the present context this term vanishes since all the j_t make C holomorphic.) Generically D will have a one-dimensional kernel and cokernel, so let s span $\ker D$ and write $v_t = as + w$ where a is small and w is L^2–orthogonal to s; the implicit function theorem lets us solve the equation obtained by projecting (5.2) orthogonal to $\operatorname{coker} D$ for w in terms of t and a, so to determine the structure of the j_t moduli space it remains to solve for a in terms of t. Now when we project (5.2) onto $\operatorname{coker} D$ we obtain an identification of the moduli space in question with the zero set of a function whose Taylor series begins

$$c_1 t^2 + c_2 ta + c_3 a^2. \tag{5.3}$$

Now since $a = 0$ is a solution for all t (corresponding to the curve C, which is j_t–holomorphic for all t), we have $c_1 = 0$. Since (C, j_t) is nondegenerate except at $t = 0$, the solution $a = 0$ is nondegenerate for $t \neq 0$, which forces $c_2 \neq 0$. Moreover, as in [15], $c_3 \neq 0$ because of the transversality of the path j_t to the wall. It follows that provided the tubular neighborhood U and the interval $(-\epsilon, \epsilon)$ are taken small enough, the j_t–moduli space is as described in the statement of the Lemma.

Moreover, since the two zeros a of $c_2 ta + c_3 a^2$ are oppositely-oriented, for each t the spectral flows of the linearizations at C and C_t of $\bar{\partial}_{j_t}$ will be opposite. Since the sign of the spectral flow for C changes as t crosses zero, the spectral flows of (C, j_{-t}) and (C_t, j_t) therefore have the same sign (ie, the number of eigenvalue crossings that occur in the flow is the same modulo 2). Now consider the path

$$t \mapsto \begin{cases} (C, j_t) & t \leq 0 \\ (C_t, j_t) & t \geq 0 \end{cases} \tag{5.4}$$

The only t at which this path touches a wall is $t = 0$, and we know that the signs of the spectral flows at the endpoints are the same. Although curves whose spectral flows have the same sign may in general lie in different chambers, when this happens they are separated by at least two walls, not one, so it follows that (C, j_{-t}) and (C_t, j_t) must lie in the same chamber when $0 < t < \epsilon$ (and, by an identical argument, when $-\epsilon < t < 0$ as well). An appropriate perturbation of (5.4) to a path remaining in this chamber will then have the property stated in part 3 of the lemma.

For the final part of the lemma, consider generic paths $\tilde{j}_s$ of almost complex structures with $\tilde{j}_0 = j_0$ but with the other $\tilde{j}_s$ no longer constrained to make C holomorphic. Then exactly as in [15] the moduli space of $\tilde{j}_s$–holomorphic curves near C will be, for small s, diffeomorphic to the zero set of a function of a whose Taylor series begins $r_1 s + r_2 a^2$ where r_1 and r_2 are nonzero numbers. Taking the sign of s appropriately, we obtain arbitrarily small regular perturbations $\tilde{j}$ of j_0 making no curve near C and homologous to C holomorphic. By taking U small, we can ensure that there were no embedded j_0–holomorphic curves in any class $k[C]$ where $k > 1$ meeting U (this uses the fact that generically (C, j_0) will not be located on any of the walls $\mathcal{D}_\iota$ with $\iota \neq 0$); if the perturbation $\tilde{j}$ of j_0 is taken small enough there will also not be any $\tilde{j}$–holomorphic curves meeting U in any of these classes. Taking a generic perturbation of the path j_t

supported close to zero which passes though $\tilde{j}$ at $t=0$ then gives the desired result. □

Corollary 5.6 *In the context of Lemma 5.5, for* $0<|t|<\epsilon$*,*

$$P'_{j_{-t}}(C,z)=\frac{1}{P'_{j_t}(C,z)}$$

Proof By the third statement in Lemma 5.5 and by Corollary 5.2, we have $P'_{j_{-t}}(C,z)=P'_{j_t}(C_t,z)$. Use a perturbation j'_s on U of the path j_s as in Lemma 5.5 which differs from j_s only for $|s|<t/2$. Assuming the perturbation to be small enough, we may extend j_s and j'_s from the tubular neighborhood U to all of X in such a way that both are regular outside the neighborhood U (for all s) and they agree with each other outside a slightly smaller region V such that no j_s– or j'_s–holomorphic curves are contained in $U\setminus V$. The contributions of all the j'_s holomorphic curves outside U will then be constant in s. Since we can use either $j_{-t}=j'_{-t}$ or j'_0 to evaluate the invariant $\mathcal{DS}$, it follows that the contributions of curves *inside* U will be the same for j_{-t} as for j'_0. Since the former is obtained from the generating function $P'_{j_{-t}}(C,z)P'_{j_{-t}}(C_t,z)=P'_{j_{-t}}(C,z)P'_{j_t}(C,z)$ while the latter is given by the generating function 1 (for there are no j_0 curves in any class $k[C]$ in the region U), the corollary follows. □

Let us now recall some more details in the definition of Gr from [15]. The multiple covers of a j–holomorphic square-zero torus C are given weights $r_j(C,m)$ which are determined by the signs of the spectral flows of each of the four operators D_ι to a complex linear operator. Note that although Taubes did not define a contribution $r_{j,\nu}(C,m)$ when $\nu\neq 0$, these can be defined using the formulas of [5], in which Eleny Ionel and Thomas Parker interpret the Gromov invariant as a combination of the invariants of [10] (which count solutions to the inhomogeneous Cauchy–Riemann equations). As with r', we organize the $r_{j,\nu}(C,m)$ into a generating function $P_{j,\nu}(C,z)=\sum_{m\geq 0}r_{j,\nu}(C,m)z^m$. Assume as we may thanks to Corollary 3.9 that there exists an integrable complex structure j_0 on a neighborhood of C that makes both f and C holomorphic, and let (j_t,ν_t) be a path of admissible pairs with C j_t– holomorphic that connects j_0 to the nondegenerate almost complex structure $j=j_1$, such that (C,j_t,ν_t) is transverse to all walls and meets at most one wall $\mathcal{D}_\iota$ at any given t. Assume the walls are met at $0<t_1<\cdots<t_n<1$. From Taubes' definition of Gr and from Lemma 5.3 and Corollary 5.2, we have

$$P'_{j_t}(C,z)=P_{j_t}(C,z)=\frac{1}{1-z}\text{ for }t<t_1$$

(in the inhomogeneous case this uses the formulas of [5]; see the proof of Corollary 5.9 for more on this). We also know that if (C, j_t) crosses $\mathcal{D}_0$ at t_0, then P and P' both satisfy the transformation rule

$$P_{j_{t_0+\epsilon}}(C,z) = \frac{1}{P_{j_{t_0-\epsilon}}(C,z)} \quad P'_{j_{t_0+\epsilon}}(C,z) = \frac{1}{P'_{j_{t_0-\epsilon}}(C,z)}.$$

So since P and P' are both unchanged when (C, j) varies within a chamber, to show that they agree we need only show that they transform in the same way when (C, j_t) crosses one of the walls $\mathcal{D}_\iota$ where $\iota \neq 0$. To again make contact with the inhomogeneous situation, note that just as the independence of $\mathcal{DS}$ from the almost complex structure and the perturbation on $X_r(f)$ used to define it lead to the wall crossing formulas for the $P'_{j,\nu}$, if we view Gr as a combination of Ruan–Tian invariants, the independence of these invariants from the almost complex structure and the perturbation on X can be considered to lead to wall crossing formulas for the $P_{j,\nu}$ which are identical to the wall crossing formulas written down by Taubes in the case $\nu = 0$.

We now record the following results, which summarize relevant parts of Lemmas 5.10 and 5.11 of [15] and their proofs.

Lemma 5.7 *Assume that (C, j_t) crosses the wall $\mathcal{D}_\iota$ where $\iota \neq 0$ at $t = t_0$. For ϵ sufficiently small, $|t - t_0| < \epsilon$, and for a suitably small neighborhood U of C:*

(1) *The only connected embedded j_t–holomorphic curve homologous to C and meeting U is C itself.*

(2) *The only connected, embedded j_t–holomorphic curves meeting U in any homology class $m[C]$ where $m > 1$ come in a family $\tilde{C}_t$ in class $2[C]$ defined either only for $t > t_0$ or only for $t < t_0$. As $t \to t_0$, suitably chosen embeddings $\tilde{u}_t\colon \tilde{C}_t \to X$ converge to $u \circ \pi\colon \tilde{C}_0 \to X$, where u is the embedding of C and $\pi\colon \tilde{C}_0 \to C$ is a double cover classified by $\iota \in H^1(C, \mathbb{Z}/2)$.*

(3) *The signs of the spectral flows for $(\tilde{C}_{t_0+\delta}, j_{t_0+\delta})$ are the same as those for $(\tilde{C}_0, j_{t_0-\delta})$, where $\tilde{C}_0$ is mapped to X by $u_{t_0-\delta} \circ \pi$ (here δ is any small number having whatever sign is needed for $\tilde{C}_{t_0+\delta}$ to exist).*

Using the information from part 3 of the above lemma, the $r_j(C, m)$ are defined in such a way as to ensure that

$$P_{j_{t_0-\delta}}(C,z) = P_{j_{t_0+\delta}}(C,z) P_{j_{t_0+\delta}}(\tilde{C}_{t_0+\delta}, z), \tag{5.5}$$

which is necessary for Gr to be independent of the almost complex structure used to define it; Taubes finds necessary and sufficient conditions in which the $r_j(C,m)$ should depend on the signs of the spectral flows in order for (5.5) to hold. Meanwhile, the fact that $\mathcal{DS}$ is known a priori to be independent of the almost complex structure J used to define it ensures that

$$P'_{j_{t_0-\delta}}(C,z) = P'_{j_{t_0+\delta}}(C,z)P'_{j_{t_0+\delta}}(\tilde{C}_{t_0+\delta},z), \tag{5.6}$$

as can be seen by the usual method of taking smooth almost complex structures J_t which are Hölder-close enough to the $\mathbb{J}_{j_t}$ that a J_t–holomorphic section in the relevant homotopy classes meets the neighborhood $\mathbb{U}$ if and only if it contributes to one of the terms in (5.6), in which case it is contained in $\mathbb{U}$. If we somehow knew a priori that the $r'_j(C,m)$ depended only on the signs of the spectral flows, then because Taubes' conditions are *necessary* in order to get an invariant it would follow that $P'_{j_t}(C,z)$ has to change as t crosses t_0 in the same way that $P_{j_t}(C,z)$ changes. However, we only know that the $r'_j(C,m)$ are unchanged if we move (C,j) within a chamber; nonetheless it's not difficult to push what we know far enough to get the right transformation rule.

Lemma 5.8 *In the context of Lemma 5.7,*

$$P'_{j_{t_0+\delta}}(C,z) = \frac{P'_{j_{t_0-\delta}}(C,z)}{P'_{j_{t_0-\delta}}(C,z^2)}.$$

Proof Assume that (C,j_t) crosses some $\mathcal{D}_\iota$ with $\iota \neq 0$ precisely at the point t_0, and work in the notation of Lemma 5.7. Observe that, analogously to the situation for crossings of $\mathcal{D}_0$, since (where δ is small and of whichever sign is necessary for the following statements to make sense) $(\tilde{C}_{t_0+\delta}, j_{t_0+\delta})$ and $(\tilde{C}_0, j_{t_0-\delta})$ have identical signs for their spectral flows, and since the path

$$t \mapsto \begin{cases} (\tilde{C}_0, j_t) & t \text{ between } t_0-\delta \text{ and } t_0 \\ (\tilde{C}_t, j_t) & t \text{ between } t_0 \text{ and } t_0+\delta \end{cases} \tag{5.7}$$

only meets a wall at $t=t_0$, $(\tilde{C}_{t_0+\delta}, j_{t_0+\delta})$ and $(\tilde{C}_0, j_{t_0-\delta})$ must lie in the same chamber (their having identical signs for their spectral flows but lying in different chambers would require any path between them to meet two walls). We can therefore perturb the path (5.7) near t_0 to one (say $t \mapsto (C'_t, j'_t)$) which stays entirely within that chamber, with each j'_t making the restriction of f to the neighborhood of C'_t on which it is defined pseudoholomorphic. Hence by Corollary 5.2 we have $r'_{j_{t_0+\delta}}(\tilde{C}_{t_0+\delta},m) = r'_{j_{t_0-\delta}}(\tilde{C}_0,m)$. But $\tilde{C}_0$ is a double cover of C, so in fact $r'_{j_{t_0+\delta}}(\tilde{C}_{t_0+\delta},m) = r'_{j_{t_0-\delta}}(C,2m)$, ie,

$$P'_{j_{t_0+\delta}}(\tilde{C}_{t_0+\delta},z) = P'_{j_{t_0-\delta}}(C,z^2).$$

The lemma then follows immediately from equation 5.6. □

Again, the same wall crossing formula for the $P_{j,\nu}$ for general ν follows in exactly the same way, using the independence of Gr from the data used to define it via the "Ruan–Tian series" that appears in [5].

Corollary 5.9 *Let j be an almost complex structure as in Corollary 3.9 and C a j–holomorphic square-zero torus. Then $r'_{j,\nu}(C,m) = r_{j,\nu}(C,m)$ for all m and ν for which (j,ν) is admissible and (C,j,ν) is m–nondegenerate.*

Proof Let j_t be a path of almost complex structures making f and C holomorphic beginning at an almost complex structure j_0 which is integrable near C and ending at $j = j_1$, and let ν_t be inhomogeneous terms such that each (j_t, ν_t) is admissible and (C, j_t, ν_t) is transverse to all walls; Lemmas 4.1 and 4.4 ensure the existence of such paths. Assume the walls are crossed at the points $t_1 < \cdots < t_n$ (so that in particular (C, j_0, ν_0) is m–nondegenerate). Now it follows from the description of the Gromov invariant in terms of the Ruan–Tian invariants in [5] that $r_{j_0,\nu_0}(C,m) = 1$ for all m: Definition 3.3 and Theorem 4.5 of that paper show that the contribution in question may be computed by assigning to the various m–fold covers of C (including the disconnected ones) weights which add up to 1 when all the linearizations of the inhomogeneous equations are surjective and complex linear. So by Lemma 5.3, for all m $P_{j_t,\nu_t}(C,z) = P'_{j_t,\nu_t}(C,z) = \frac{1}{1-z}$ for all m and all suitably small t, and by Corollary 5.2 $P_{j_t,\nu_t}(C,z)$ and $P'_{j_t,\nu_t}(C,z)$ change only when t passes one of the t_i. By Corollary 5.6 and the construction of Gr (specifically Equation 5.26 of [15]), if the wall $\mathcal{D}_0$ is crossed at t_i the changes in both P and P' are found by taking the reciprocal, while Lemma 5.8 above and Equation 5.28 of [15] tell us that if the wall $\mathcal{D}_\iota$ with $\iota \neq 0$ is crossed at t_i then both P and P' change according to the rule

$$P_{j_{t_i+\delta},\nu_{t_i+\delta}}(C,z) = \frac{P_{j_{t_i-\delta},\nu_{t_i-\delta}}(C,z)}{P_{j_{t_i-\delta},\nu_{t_i-\delta}}(C,z^2)},$$

δ being small and of the same sign as in Lemma 5.7. Hence $P'_{j_1,\nu_1}(C,z) = P_{j_1,\nu_1}(C,z)$, proving the corollary. □

The objects which contribute to $Gr(\alpha)$ are, for generic almost complex structures j, formal sums of form $h = \sum m_i C_i$ where the C_i are disjoint m_i–nondegenerate j–holomorphic curves, the m_i are positive integers which are required to equal 1 unless C_i is a square zero torus, and $\sum m_i[C_i] = PD(\alpha)$.

For curves C_i which are not square zero tori, let $r_j(C,1)$ be the contribution of C to Gr (ie, the sign of the spectral flow of the linearization of $\bar{\partial}_j$), and (assuming j makes f holomorphic and $\mathbb{J}_j$ is regular for C) $r'_j(C,1)$ the contribution of C to $\mathcal{DS}$, so that, by Theorem 4.6, $r'_j(C,1) = r_j(C,1)$. By definition, the contribution of the formal sum h to $Gr(\alpha)$ is $\prod_i r_j(C_i,m_i)$, while the proof of Proposition 5.4 shows that the contribution of h to $\mathcal{DS}_{(X,f)}(\alpha)$ is $\prod_i r'_j(C_i,m_i)$. Thus the previous proposition shows that *every object h which contributes to Gr contributes to $\mathcal{DS}$ in the same way.* To prove that $\mathcal{DS} = Gr$, we need to see that, if we compute $\mathcal{DS}$ using an almost complex structure J Hölder close to a generic $\mathbb{J}_j$, then the only sections contributing to $\mathcal{DS}$ may be viewed as contributions from some disjoint union of j–holomorphic curves in X with only square-zero tori allowed to be multiply covered.

To see this, note that for any $\alpha \in H^2(X,\mathbb{Z})$, by Gromov compactness, if J is close enough to $\mathbb{J}_j$ then any J–holomorphic sections in the class c_α must be contained in some small neighborhood of a section which tautologically corresponds to some (generally disconnected, not embedded) curve in X with total homology class $PD(\alpha)$. Now for generic j, the space of (possibly disconnected) j–holomorphic curves in X which have any singularities (including intersection points of different connected components) or have any components other than square-zero tori or exceptional spheres which are multiply covered has dimension strictly less than the dimension $d(\alpha)$ (This follows by easy algebra using the formula for $d(\alpha)$, and is of course the reason that Gr is not obliged to count singular curves or multiply-covered curves other than square-zero tori). Curves in X with multiply-covered exceptional sphere components may similarly be eliminated by a dimension count: If α is any class represented by a j–holomorphic curve and β is the class of an exceptional sphere, we have $d(\alpha - m\beta) = d(\alpha) - \beta \cdot m\alpha - \frac{1}{2}(m^2 + m) < d(\alpha) - 1$, so for generic choices of $d(\alpha)$ points in X, no union C of a j–holomorphic curve in class $\alpha - m\beta$ with an m–fold cover of the j–holomorphic sphere in class β passes through all $d(\alpha)$ of the points.

Hence in any case, the space of $\mathbb{J}_j$–holomorphic sections tautologically corresponding to curves not counted by Gr has dimension less than the dimension of the space of sections counted by $\mathcal{DS}_{(X,f)}(\alpha)$, which is equal to $d(\alpha)$ by Proposition 4.3 of [14]. In principle, it perhaps could happen that when we perturb $\mathbb{J}_j$ to a smooth almost complex structure J near such a section s_C to find the contribution of C we might obtain a positive-dimensional set of nearby J–holomorphic sections, but because these sections are constrained by Gromov compactness to stay near s_C, for a large open set of choices of the incidence conditions used to cut down the moduli spaces for Gr and $\mathcal{DS}$ to be zero-dimensional, the perturbed sections will still not appear in this moduli space and so will not contribute to $\mathcal{DS}$.

$\mathcal{DS}$ and Gr therefore receive contributions from just the same objects, so since these contributions are equal, Theorem 1.1 follows.

References

[1] **Denis Auroux**, **Ludmil Katzarkov**, *The degree doubling formula for braid monodromies and Lefschetz pencils*, preprint (2000) available at: `http://www.math.polytechnique.fr/cmat/auroux/math.us.html`

[2] **Simon Donaldson**, *Lefschetz pencils on symplectic manifolds*. J. Diff. Geom. 53 (1999) 205–236

[3] **Simon Donaldson**, **Peter Kronheimer**, *The Geometry of Four–Manifolds*, Oxford University Press (1990)

[4] **Simon Donaldson**, **Ivan Smith**, *Lefschetz pencils and the canonical class for symplectic 4–manifolds*, `arXiv:math.SG/0012067`

[5] **Eleny-Nicoleta Ionel**, **Thomas Parker**, *The Gromov invariants of Ruan–Tian and Taubes*, Math. Res. Lett. 4 (1997) 521–532

[6] **Sergei Ivashkovich**, **Vsevolod Shevchishin**, *Gromov Compactness Theorem for Stable Curves*, `arXiv:math.DG/9903047`

[7] **Dusa McDuff**, *The local behaviour of holomorphic curves in almost complex 4–manifolds*, J. Diff. Geom. 34 (1991) 143–164

[8] **Dusa McDuff**, **Dietmar Salamon**, *J–holomorphic Curves and Quantum Cohomology*. University Lecture Series, Volume 6, AMS (1994)

[9] **Yongbin Ruan**, **Gang Tian**, *A mathematical theory of quantum cohomology*, J. Diff. Geom. 42 (1995) 259–367

[10] **Yongbin Ruan**, **Gang Tian**, *Higher genus symplectic invariants and sigma models coupled with gravity*, Invent. Math. 130 (1997) 455–516

[11] **Bernd Siebert**, **Gang Tian**, *Weierstraßpolynomials and plane pseudo-holomorphic curves*, Chinese Ann. Math. B 23 No. 1 (2002) 1–10

[12] **Jean-Claude Sikorav**, *Some Properties of Holomorphic Curves in Almost-Complex Manifolds*, from: "Holomorphic Curves in Symplectic Geometry", (Michele Audin and Jacques Lafontaine, editors), Birkhauser (1993)

[13] **Ivan Smith**, *Lefschetz Pencils and Divisors in moduli space*, Geometry and Topology 5 (2001) 579–608

[14] **Ivan Smith**, *Serre–Taubes Duality for pseudoholomorphic curves*, e-print, `arXiv:math.SG/0106220`

[15] **Clifford Henry Taubes**, *Counting Pseudo-Holomorphic Submanifolds in Dimension 4*, J. Diff. Geom. 44 (1996) 818–893

[16] **Clifford Henry Taubes**, *Seiberg–Witten and Gromov invariants for Symplectic 4–manifolds*, International Press (2000)

[17] **Michael Usher**, *Relative Hilbert scheme methods in pseudoholomorphic geometry*, PhD thesis, MIT (2004)