Fukaya-Seidel category and gauge theory

Andriy Haydys
Imperial College London and University of Bielefeld

May 25, 2011

Abstract

Given a J-holomorphic Morse function on a symplectic manifold, a new construction of the Fukaya-Seidel category is outlined. Applying this construction in an infinite dimensional case, a Fukaya-Seidel-type category is associated to a smooth three-manifold. In this case the construction is based on a five-dimensional gauge theory.

1 Introduction

This paper consists of two major parts. In the first part, based on the idea of Seidel [Sei1] we outline a construction of the Fukaya-Seidel category, which is associated to a symplectic manifold M equipped with a compatible almost complex structure J and a J-holomorphic Morse function f. This construction does not rely on the notion of a vanishing cycle but emphasizes instead the role of the antigradient flow lines of $\mathrm{Re}\,(e^{i\theta}f)$. In the second part, this construction is applied in the infinite dimensional case of the complex Chern-Simons functional. The corresponding construction conjecturally associates a Fukaya-Seidel-type category to a smooth three-manifold.

Our motivation originated from the suggestion to use higher dimensional gauge theory in studies of low dimensional manifolds as outlined in [Hay]. Namely, suppose we are granted a construction that associates a higher dimensional manifold W_X to each lower dimensional manifold X from a suitable subclass and possibly equipped with an additional structure. The manifold W_X is assumed to be of dimension $6, 7$ or 8 and endowed with an $SU(3), G_2$ or $Spin(7)$ structure respectively. Then, by counting higher dimensional instantons on W_X we should obtain an invariant of X. The construction studied in [Hay] in detail associates to each smooth spin four-manifold the total space of its spinor bundle.

Another construction of a similar nature associates to X^4 the total space of the "twisted spinor bundle" $\underline{\mathbb{R}} \oplus \Lambda^2_+T^*X$. Then $Spin(7)$-instantons invariant along each fibre are solutions of the Vafa-Witten equations [VW], while $Spin(7)$-instantons invariant only along the fibres of $\Lambda^2_+T^*X$ can be interpreted as antigradient flow lines of a function, whose critical points are solutions of the Vafa-Witten equations. It turns out that these flow lines can be obtained from certain elliptic equations on a general five-manifold W^5 equipped with a nonvanishing vector field by specializing to the case $W = X^4 \times \mathbb{R}$ just like flow lines of the real Chern-Simons functional are obtained from the anti-self-duality equations on $X^4 = Y^3 \times \mathbb{R}$. Specializing further to $W = Y^3 \times \mathbb{R}^2$ we obtain a construction of a Fukaya-type A_∞-category (this requires some extra choices) just like specialization of the anti-self-duality equations to $\Sigma^2 \times \mathbb{R}^2$ leads to the construction of the Fukaya A_∞-category associated to Σ. At this point an important distinction from the case of Riemann surfaces emerges. Namely, the construction involves a natural holomorphic function, the complex Chern-Simons functional, and this has significant implications for the flavour of the construction.

Having said this though, we do not appeal in this paper to higher dimensional asd equations but rather begin directly with the formulation of the five-dimensional gauge theory. From this perspective the most interesting theory is obtained via reduction to three-manifolds, where the construction of the A_∞-category admits a finite-dimensional interpretation in the framework of symplectic geometry.

This paper is organized as follows. In Section 2 we describe the construction of the Fukaya-Seidel category in the finite-dimensional case. From one point of view, this construction is a generalization of a Floer theory, where *generators* of the homology groups are antigradient flow lines of the real part of a holomorphic Morse function connecting a pair of critical points. Then the Floer differential is obtained from pseudoholomorphic planes with a Hamiltonian perturbation satisfying certain asymptotic conditions (see (4) for more details).

Sections 3 and 4 are devoted to the formulation of the five-dimensional gauge theory and its various dimensional reductions. In Section 5 we describe applications of the equations obtained in the previous sections to low dimensional topology. In particular, one can (conjecturally) associate an integer to a five-manifold, Floer-type homology groups to a four-manifold and a Fukaya-Seidel-type category to a three-manifold. In dimension three, critical points correspond to flat G^c-connections on Y, flow lines correspond to Vafa-Witten-type instantons on $Y \times \mathbb{R}$ and pseudoholomorphic planes correspond to "five-dimensional instantons" on $Y \times \mathbb{R}^2$. This should be a part of a multi-tier (extended) quantum field theory [Fre] but we do not study this aspect in the current paper.

The constructions described in this paper may also be useful in other settings, for instance in the context of Calabi-Yau threefolds. Here the critical points of the holomorphic Chern-Simons functional correspond to holomorphic vector bundles over a Calabi-Yau threefold Z, flow lines correspond to G_2-instantons on $Z \times \mathbb{R}$ and pseudoholomorphic planes correspond to $Spin(7)$-instantons on $Z \times \mathbb{R}^2$.

Many aspects of this paper are related to ideas of various authors. As it has been already mentioned above, our construction of the Fukaya-Seidel category in the finite dimensional case is a modification of Seidel's idea. The equation we utilize for the definition of the structure maps in the Fukaya-Seidel A_∞-category was used in the context of mirror symmetry by Fan-Jarvis-Ruan [FJR] ("Witten equation") in the case of quasi-homogeneous polynomials. The antigradient flow lines of the real part of the holomorphic Chern-Simons functional appeared in [KW] for the first time and were further studied by Witten [Wit1, Wit3]. Donaldson and Segal [DS] used antigradient flow lines of the real part of the holomorphic Chern-Simons functional in the context of Calabi-Yau threefolds.

Acknowledgements: I thank S.Bauer, S.Donaldson, V.Pidstrygach, Y.Ruan and P.Seidel for helpful discussions and also B.Himpel for reading the draft of this paper. I acknowledge the financial support of the German Research Foundation (DFG) and the

hospitality of Imperial College London, where part of this work was carried out.

2 Fukaya-Seidel categories of symplectic Picard-Lefschetz fibrations

In [Sei1] Seidel describes the construction of a Fukaya category associated to an exact Morse fibration of a symplectic manifold in terms of vanishing cycles. In the first part of this section we describe — omitting (important) technical details — an alternative approach, which does not rely on the notion of a vanishing cycle. The rest of the section is devoted to some analytic properties of the objects involved in the construction.

2.1 Outline of the construction

Let (M^{2n}, ω) be an exact symplectic manifold. Choose a compatible almost complex structure J such that $g(\cdot,\cdot) = \omega(\cdot, J\cdot)$ is a Riemannian metric on M. Let $f\colon M \to \mathbb{C}$ be a proper J-holomorphic function. Assume that f has finitely many critical points of Morse-type lying in pairwise different fibres. It is convenient to choose a linear order on the set of critical points: $m_1 < m_2 < \cdots < m_k$. Denote also $z_j = f(m_j)$.

Let us briefly recall the basic ingredients of the Fukaya-Seidel A_∞-category (see [Sei1, Sei2] for details). For the sake of simplicity we consider the ungraded version with coefficients in $\mathbb{Z}/2\mathbb{Z}$ ("preliminary version" in the terminology of [Sei2]). Choose a basepoint $z_0 \in \mathbb{C}$ distinct from critical values and a collection of paths connecting z_0 with each z_j missing the remaining critical values. Let $L_j \subset f^{-1}(z_0)$ be the vanishing cycle of m_j associated to the path connecting z_0 with z_j. Denote by Γ the ordered collection $(L_1, \dots, L_k)$. Seidel associates to Γ a directed Fukaya A_∞-category $Lag^{\rightarrow}(\Gamma)$, whose objects are vanishing cycles L_j and morphisms are Floer chain complexes as follows. First recall that an A_∞-structure is a collection of maps

$$\mu^d\colon hom(L_{j_d}, L_{j_{d+1}}) \otimes \cdots \otimes hom(L_{j_1}, L_{j_2}) \longrightarrow hom(L_{j_1}, L_{j_{d+1}}), \quad d = 1, 2, 3, \dots$$

satisfying certain quadratic relations and by the directedness we have

$$hom(L_j, L_k) = \begin{cases} CF(L_j, L_k) & j < k, \\ \mathbb{Z}/2 \cdot id & j = k, \\ 0 & j > k. \end{cases}$$

The Floer complex $CF(L_j, L_k)$ is generated by the points of $L_j \cap L_k$ and the map μ^1 is the Floer differential, which counts pseudoholomorphic strips such that one boundary component is mapped to L_j and the other component is mapped to L_k. The maps μ^d for $d \geq 2$ are defined similarly by counting pseudoholomorphic discs with $d+1$ punctures on the boundary. The resulting A_∞-category $Lag^{\rightarrow}(\Gamma)$ depends on the choices made but Seidel shows that the derived category $D^b(Lag^{\rightarrow}(\Gamma))$ is an invariant of the Morse fibration.

With this understood we now give another construction of the Fukaya-Seidel A_∞-category. Choose a basepoint z_0 that does not lie on any straight line determined by a pair of critical values.[1] Pick a pair of critical points (m_-, m_+) and denote $\theta_\pm = \arg(z_\pm - z_0) \in (-\pi, \pi]$. Put

$$f = f_0 + i f_1, \qquad v_j = \operatorname{grad} f_j \quad \text{for } j = 0, 1$$

and denote by $\gamma_m^\pm$ the solution of the Cauchy problem

$$\dot\gamma_m^\pm + \cos\theta_\pm\, v_0 + \sin\theta_\pm\, v_1 = 0, \quad \gamma_m^\pm(0) = m \in f^{-1}(z_0).$$

Notice that the image of $f \circ \gamma_m^\pm \colon \mathbb{R} \to \mathbb{C}$ is contained in a straight line passing through z_0 and $z_\pm$. Then the vanishing cycle $L_\pm$ of $m_\pm$ associated with the segment $\overline{z_0 z_\pm}$ can be conveniently described as

$$L_\pm = \bigl\{ m \in f^{-1}(z_0) \bigm| \lim_{t\to+\infty} \gamma_m^\pm(t) = m_\pm \bigr\}.$$

Then, if we denote

$$\theta_0(t) = \begin{cases} \theta_+ & t \le 0, \\ \arg i(z_- - z_0) = \theta_- \pm \pi & t > 0, \end{cases}$$

the set $L_+ \cap L_-$ can be identified with the space of solutions to the problem

$$\dot\gamma + \cos\theta_0(t)\, v_0 + \sin\theta_0(t)\, v_1 = 0, \qquad \lim_{t\to\pm\infty} \gamma(t) = m_\mp. \tag{1}$$

Here solutions are understood to be smooth on $\mathbb{R} \setminus \{0\}$ and continuous at $t = 0$. We call solutions of (1) *broken flow lines* of f connecting m_- and m_+ and denote by $\mathcal{M}_0(m_-; m_+)$ the space of all solutions. Notice that for each broken antigradient flow line γ the image of $f \circ \gamma$ lies on the curve $\overline{z_- z_0 z_+}$ and $f \circ \gamma(0) = z_0$.

It will be convenient in the sequel to regularize θ_0. To do this choose a smooth function $\beta \colon \mathbb{R} \to [0, 1]$ such that $\beta(t) = 0$ for $t \le -1$ and $\beta(t) = 1$ for $t \ge 1$. Put

$$\theta_\lambda(t) = \theta_+ + \beta(\lambda^{-1} t)(\theta_- + \pi - \theta_+), \tag{2}$$

where $\lambda > 0$ is a parameter. The graph of the function $\theta_\lambda(t)$ is sketched in Fig.1.

Denote by $\mathcal{M}_\lambda = \mathcal{M}_\lambda(m_-, m_+)$ the space of solutions to the problem

$$\dot\gamma + \cos\theta_\lambda(t)\, v_0 + \sin\theta_\lambda(t)\, v_1 = 0, \qquad \lim_{t\to\pm\infty} \gamma(t) = m_\mp. \tag{3}$$

We call solutions of equations (3) broken flow lines, as well.

Remark 2.1. It follows from Lemma B.1 *(ii)* that for any solution γ of equations (3) with sufficiently small λ the image of $f \circ \gamma$ is arbitrarily close (in C^0-topology) to the curve $\overline{z_+ z_0 z_-}$. Moreover, it is shown in Lemma 2.5 that there exists a natural bijective correspondence between $\mathcal{M}_\lambda$ and $\mathcal{M}_0 \cong L_- \cap L_+$ provided L_- and L_+ intersect transversally in $f^{-1}(z_0)$ and λ is small enough.

[1] It is not absolutely necessary to require that z_0 does not lie on any straight line passing through a pair of critical values. However the case when z_0 lies on some straight line containing two critical values differs in details and requires a separate treatment.

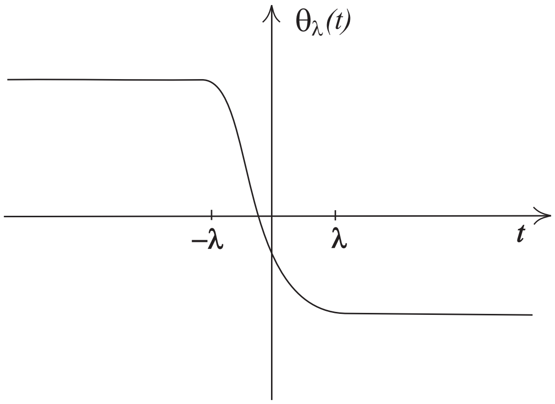

Figure 1: Graph of θ_λ.

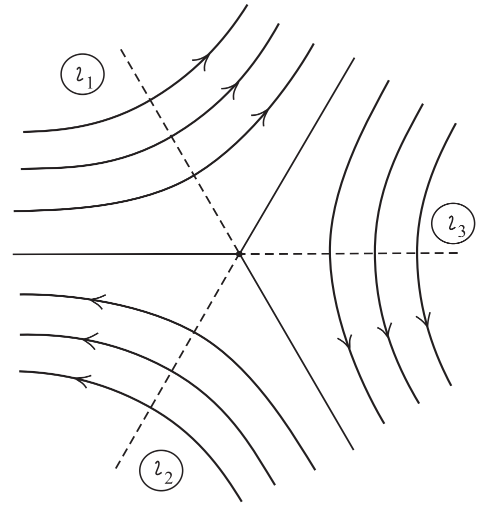

Figure 2: The domain Ω with three long necks.

Further, notice that the Floer differential μ^1 should take broken flow lines as input and should return a formal linear combinations of broken flow lines as output. With $m_\pm$ as above, pick additionally two solutions $\gamma_\pm$ of equations (3). Then the role of holomorphic strips with boundary on $L_\pm$ in our framework is played by solutions of the problem

$$\begin{aligned} &\partial_s u + J\big(\partial_t u + \cos\theta_\lambda(t)\, v_0 + \sin\theta_\lambda(t)\, v_1\big) = 0, \qquad u\colon \mathbb{R}^2_{s,t} \to M, \\ &\lim_{t\to\pm\infty} u(s,t) = m_\mp, \qquad \lim_{s\to\pm\infty} u(s,t) = \gamma_\mp(t) \end{aligned} \tag{4}$$

with finite energy. Here the limits are understood in the uniform topology and the energy of u is defined by

$$E(u) = \int_{\mathbb{R}^2} \big(|\partial_s u|^2 + |\partial_t u + \cos\theta_\lambda(t)\, v_0 + \sin\theta_\lambda(t)\, v_1|^2\big) ds \wedge dt.$$

Notice that the first equation in (4) is the pseudoholomorphic map equation with a Hamiltonian perturbation. Namely, the time-dependent Hamiltonian function here is $\operatorname{Im}(e^{-i\theta_\lambda(t)} f)$.

It is very instructive to see a relation between solutions of (4) and pseudoholomorphic strips as in Seidel's approach. This is outlined in Appendix A. However, instead of proving that such a connection indeed holds, we deal with equations (4) directly, since in view of the intended applications it is important to have direct proofs of the basic properties (Fredholmness, transversality, compactness etc.). In this paper we prove that problem (4) admits a Fredholm formulation.

Next we show how to define the map μ^2 in our framework. Let Ω be a (non-compact) Riemann surface containing three "long necks". By this we mean a triple of holomorphic embeddings

$$\imath_j\colon \{z \mid \operatorname{Re} z < 0\} \to \Omega, \qquad j = 1, 2, 3$$

with disjoint images. To be more explicit, we choose the complex plane $\mathbb{C}$ as a model for Ω (see Fig. 2), where the embedding $\imath_1$ is given in polar coordinates by $(r, \varphi) \mapsto$

$(r, \frac{2}{3}(\varphi+\pi))$, $\frac{\pi}{2} < \varphi < \frac{3\pi}{2}$ and the other two embeddings are defined similarly. The curves shown on the figure are of the form $t \mapsto \imath_j(s,t)$. This is our analogue of the "pair of pants" surface.

Further, pick any three critical points, say $m_1 < m_2 < m_3$, and a pair (γ_1, γ_2) of broken flow lines connecting m_1 with m_2 and m_2 with m_3 respectively. More precisely, γ_1 and γ_2 are solutions to the equations

$$\begin{gathered}\dot\gamma_j + \cos\theta_{j,\lambda}(t)\, v_0 + \sin\theta_{j,\lambda}(t)\, v_1 = 0,\\ \lim_{t\to+\infty} \gamma_j(t) = m_j, \quad \lim_{t\to-\infty}\gamma_j(t) = m_{j+1}.\end{gathered}$$

Here $\theta_{j,\lambda}(t)$ is defined by replacing (θ_-, θ_+) in formula (2) with (θ_j, θ_{j+1}). Then $\mu^2(\gamma_1,\gamma_2)$ should be a formal linear combination of broken flow lines connecting m_1 with m_3. Pick any such flow line, i.e. a solution to the problem[2]

$$\begin{gathered}\dot\gamma_3 + \cos\theta_{3,\lambda}(t)\, v_0 + \sin\theta_{3,\lambda}(t)\, v_1 = 0,\\ \lim_{t\to+\infty} \gamma_3(t) = m_1, \quad \lim_{t\to-\infty}\gamma_3(t) = m_3,\end{gathered}$$

and also choose a nowhere vanishing form $\omega \in \Omega^{0,1}(\Omega)$ such that $\imath_j^*\omega = \frac{1}{2}e^{i\theta_{j,\lambda}(t)}d\bar z$ for $j = 1,2$ and $\imath_3^*\omega = \frac{1}{2}e^{-i\theta_{3,\lambda}(t)}d\bar z$. Then the multiplicity of γ_3 can conjecturally be defined by counting solutions to the equations

$$\begin{aligned}&\bar\partial u + \omega\otimes v_0(u) = 0, \qquad u\colon \Omega\to M,\\ &\lim_{s\to-\infty} u\circ\imath_j(s,t) = \gamma_j(t) \text{ for } j=1,2 \quad\text{and}\quad \lim_{s\to-\infty} u\circ\imath_3(s,t) = \gamma_3(-t),\\ &\lim_{t\to+\infty} u\circ\imath_j(s,t) = m_j, \qquad \lim_{t\to-\infty} u\circ\imath_j(s,t) = m_{j+1}, \qquad j=1,2,3,\end{aligned}$$

where $\omega\otimes v_0(u) \in \Omega^{0,1}(\Omega; u^*TM)$. Notice that over the long necks the above equations are of a similar form as equations (4).

The analogue of holomorphic discs with $d+1$ punctures on the boundary involved in the definition of μ^d are defined in a similar manner.

Let us briefly summarize. Assuming the standard compactness and transversality results, we can (conjecturally) construct a directed A_∞-category $\mathcal{A}(f,J)$ as follows. The objects of $\mathcal{A}(f,J)$ are critical points of f. For any pair (m_-, m_+) of critical points, denote by $CF(m_-,m_+)$ the vector space generated by $\mathcal{M}_\lambda(m_-;m_+)$ and put

$$hom_{\mathcal{A}(f,J)}(m_-,m_+) = \begin{cases} CF(m_-,m_+) & m_- < m_+,\\ \mathbb{Z}/2\cdot id & m_- = m_+,\\ 0 & m_- > m_+.\end{cases}$$

For $\gamma_\pm \in \mathcal{M}_\lambda(m_-;m_+)$ denote by $\mathcal{M}^0_\lambda(\gamma_-,\gamma_+)$ the zero-dimensional component of the space $\{u \mid u \text{ solves } (4)\}/\mathbb{R}$ and put[3]

$$\mu^1(\gamma_-) = \sum_{\gamma_+}\big(\#\mathcal{M}^0_\lambda(\gamma_-,\gamma_+) \bmod 2\big)\,\gamma_+.$$

[2]Our convention is that for $m_- < m_+$ a broken flow line goes from m_+ to m_- as t varies between $-\infty$ and $+\infty$ and therefore the asymmetry between γ_3 and γ_1,γ_2.

[3]It is clear that some perturbations are necessary to make sense of $\#\mathcal{M}^0_\lambda(\gamma_-,\gamma_+)$.

The maps μ^d for $d \geq 2$ are defined in a similar manner and together with μ^1 (conjecturally) combine to an A_∞-structure. Clearly, $\mathcal{A}(f, J)$ depends on the various choices involved in the construction. However, as explained in [Sei1] the derived category $D^b(\mathcal{A}(f, J))$ should not depend on these choices. Moreover, assume $(f_\tau, J_\tau),\ \tau \in [0, 1]$ is a continuous family such that f_τ is a J_τ-holomorphic function, whose critical points lie in pairwise different fibres for all τ. Then $D^b(\mathcal{A}(f_0, J_0))$ is equivalent to $D^b(\mathcal{A}(f_1, J_1))$.

Remark 2.2. Our main example is the complex Chern-Simons functional, which takes values in $\mathbb{C}/\mathbb{Z}$ rather than in $\mathbb{C}$. In this case, the construction outlined above does not immediately apply. However, we may proceed as follows. Assume that each line $\ell_r = \{z \mid \operatorname{Re} z = r \mod \mathbb{Z}\}$ contains at most one critical value of f (possibly after a perturbation). Pick r such that the line ℓ_r does not contain any critical value of f and "cut" the cylinder $\mathbb{C}/\mathbb{Z}$ along ℓ_r to obtain a holomorphic function f_r with values in $(0, 1) \times \mathbb{R}$. In other words, consider only those flow lines γ of f for which the image of $f \circ \gamma$ does not intersect the line ℓ_r. Then $D^b(\mathcal{A}(f_r))$ does not depend on r as long as r varies in a connected interval I such that $I \times \mathbb{R}$ does not contain any critical value of f. In this way we obtain a collection of k triangulated categories $\big(D^b(\mathcal{A}(f_{r_j}))\big)_{j=1}^{k}$, which is well-defined up to a cyclic permutation. Here k is the number of critical values of f.

Clearly, this procedure is not optimal but allows to avoid complications related to the Novikov theory.

2.2 The action functional

With $m_\pm$ and f as in the preceding subsection denote

$$\Gamma(m_-; m_+) = \{\gamma \in C^\infty(\mathbb{R}; M) \mid \gamma(t) \to m_\pm \text{ as } t \to \mp\infty \text{ and } |\dot\gamma| \text{ is rapidly decaying}\}$$

and choose a reference curve $\gamma_0 \in \Gamma(m_-; m_+)$. For another curve $\gamma \in \Gamma(m_-; m_+)$ assume there exists a smooth map $u \colon \mathbb{R}^2 \to M$ with the following properties:

- ◇ $u(s, \cdot) \in \Gamma(m_-; m_+)$ for all s and $\lim_{s\to+\infty} u(s, t) = \gamma(t),\ \lim_{s\to-\infty} u(s, t) = \gamma_0(t)$;
- ◇ $|\partial_s u|$ is rapidly decaying.

Clearly, such a map exists if and only if γ and γ_0 are homotopic relative to their endpoints or, equivalently, if and only if the loop $\gamma_0^{-1} * \gamma$ is contractible.

The action functional is defined by

$$\mathcal{F}(\gamma, u) = \int_{\mathbb{R}^2} u^*\omega + \int_{\mathbb{R}} \operatorname{Im}\big(e^{-i\theta_\lambda(t)} f\,(\gamma\,(t))\big)\, dt.$$

Let v^t denote the time-dependent vector field $\operatorname{grad}\operatorname{Re}\big(e^{-i\theta_\lambda(t)} f\big) = \cos\theta_\lambda(t)\, v_0 + \sin\theta_\lambda(t)\, v_1$. A standard computation shows that $d\mathcal{F}(\xi) = -\int_{\mathbb{R}} \omega(\xi, \dot\gamma + v^t)\, dt$, where ξ is a vector field along γ. Therefore $\operatorname{grad}\mathcal{F} = J(\dot\gamma + v^t)$ and it follows that the critical points of the functional $\mathcal{F}$ (at least formally) are broken flow lines of f connecting m_+ and m_-. Similarly, the antigradient flow lines of $\mathcal{F}$ are solutions of equations (4).

Remark 2.3. If ω is exact, say $\omega = d\lambda$, then up to a constant we have

$$\mathcal{F}(\gamma) = -\int_{\mathbb{R}} \gamma^*\lambda + \int_{\mathbb{R}} \operatorname{Im}\big(e^{-i\theta_\lambda(t)} f\,(\gamma\,(t))\big)\, dt.$$

Thus, if M is exact, $\mathcal{F}$ is a well-defined function on $\Gamma(m_-; m_+)$.

2.3 Fredholm property

In this section we show that problem (4) admits a Fredholm formulation. This is done by adopting the well-known technique for elliptic operators on manifolds with tubular ends (as in [Don2], for instance). Some extra care is necessary, since in our case tubular ends are modeled on $Y \times \mathbb{R}$ with non-compact Y (in our case $Y = \mathbb{R}$, in fact).

For the proof of Theorem 2.6 below we need the following general result, which we state in somewhat greater generality than we need, since this may also be useful in other settings. The statement of the lemma below may well be known to analysts. The case $p = 2$ can be found in [Kor], but for the reader's convenience we reproduce the argument here. For relevant definitions and some basic facts consult [Shu, App.1].

Lemma 2.4. *Let E, F be vector bundles of bounded geometry over a Riemannian manifold M of bounded geometry. Suppose $A\colon \Gamma(E) \to \Gamma(F)$ is a C^∞-bounded, uniformly elliptic differential operator of order m. If in addition $A\colon W^{m,2}(E) \to L^2(F)$ is an isomorphism, then $A\colon W^{m+k,p}(E) \to W^{k,p}(F)$ is an isomorphism for all $k \in \mathbb{R}$ and $p \in (1, \infty)$.*

Proof. We first prove the statement in the case $M = \mathbb{R}^n$ and trivial bundles E, F. The proof is based on a standard regularity result. Namely, if $A\xi \in W^{k,p}$ for some k, p and $\xi \in \bigcup_l W^{l,p}$, then $\xi \in W^{k+m,p}$ and we have the a priori estimate

$$\|\xi\|_{k+m,p} \leq C\big(\|A\xi\|_{k,p} + \|\xi\|_{l,p}\big) \qquad \text{for any } l < k. \tag{5}$$

First consider $p = 2$. It easily follows from the above estimate that $A\colon W^{m+k,2} \to W^{k,2}$ is an isomorphism for all $k > 0$. To prove the statement for $k < 0$, recall that $W^{-k,2}$ is the dual space of $W^{k,2}$. Therefore the dual operator $A^*\colon L^2 \to W^{-m,2}$ is an isomorphism and hence also an isomorphism as a map $A^*\colon W^{k,2} \to W^{k-m,2}$ for all $k > 0$. Applying the same trick again, we obtain the statement for $k < 0$.

Let A^{-1} denote the inverse map, which is well-defined on any $W^{k,2}$ and hence on any $W^{k,p}$. From the above mentioned regularity follows that $A^{-1}(W^{k,p}) \subset W^{k+m,p}$. Further, with the help of the Sobolev embedding theorem, inequality (5), and the estimate $\|\xi\|_{k+m,2} \leq C(k)\|A\xi\|_{k,2}$ it is easy to obtain the inequality

$$\|\xi\|_{k+m,p} \leq C(k,p)\|A\xi\|_{k,p}. \tag{6}$$

This finishes the proof in the case $M = \mathbb{R}^n$. Then by a standard patching argument as in [Shu, App.1], one obtains that an estimate similar to (6) also holds for E, F, M as in the statement of the lemma. This finishes the proof. $\square$

Let $W^{1,2}(\mathbb{R}; M)$ denote the Sobolev space of maps $\gamma\colon \mathbb{R} \to M$ with square integrable first derivative. Introduce the Banach manifold

$$W^{1,2}_{m_-;\,m_+} = \{\, \gamma \in W^{1,2}(\mathbb{R}; M) \mid \lim_{t\to\pm\infty} \gamma(t) = m_\mp \}$$

and the vector bundle $\mathcal{E} \to W^{1,2}_{m_-;\,m_+}$, whose fibre at γ is the Hilbert space $L^2(\gamma^*TM)$. The section Φ_λ of $\mathcal{E}$ given by $\Phi_\lambda(\gamma) = \dot\gamma + \cos\theta_\lambda(t)\, v_0 + \sin\theta_\lambda(t)\, v_1$ is Fredholm [Sal1] with vanishing index, since the Morse indices of m_+ and m_- are equal. Here $m_\pm$ is regarded as a critical point of $\mathrm{Re}\,(e^{-i\theta_\pm} f)$. Clearly, $\Phi_\lambda^{-1}(0) = \mathcal{M}_\lambda(m_-; m_+)$.

Lemma 2.5. *Assume the vanishing cycles L_- and L_+ intersect transversally in $f^{-1}(0)$. Then there exists $\lambda_0 > 0$ such that Φ_λ intersects the zero section transversally for all $\lambda \in (0, \lambda_0)$. Moreover there exists a natural bijective correspondence between $\mathcal{M}_\lambda(m_-; m_+)$ and $\mathcal{M}_0(m_-; m_+)$ provided $\lambda \in (0, \lambda_0)$.*

The proof of this lemma is somewhat technical and is deferred to Appendix B.

Fix any sufficiently small λ and two distinct solutions $\gamma_\pm = \gamma_{\lambda,\pm}$ of equations (3). For $k \geq 1$ and $p \in [2, +\infty)$ introduce the Banach manifold

$$W^{k,p}_{\gamma_-,\gamma_+} = \{u \in W^{k,p}(\mathbb{R}^2; M) \mid \lim_{t\to\pm\infty} u(s,t) = m_\mp,\ \lim_{s\to\pm\infty} u(s,t) = \gamma_\mp(t)\},$$

where the limits are understood in the uniform topology. Let $\mathcal{F} \to W^{k,p}_{\gamma_-,\gamma_+}$ be the vector bundle with the fibre $\mathcal{F}_u = W^{k-1,p}(\mathbb{C}; u^*TM)$. Define the section $\Upsilon = \Upsilon_\lambda \in \Gamma(\mathcal{F})$ by

$$\Upsilon(u) = \partial_s u + J\big(\partial_t u + \cos\theta_\lambda(t) v_0(u) + \sin\theta_\lambda(t) v_1(u)\big).$$

Then the linearization $D_u\Upsilon\colon T_u W^{k,p}_{\gamma_-,\gamma_+} = W^{k,p}(\mathbb{R}^2; u^*TM) \to W^{k-1,p}(\mathbb{R}^2; u^*TM)$ is given by

$$\begin{aligned} D_u\Upsilon(\xi) &= \nabla_s \xi + J\big(\nabla_t \xi + \cos\theta_\lambda \nabla_\xi v_0 + \sin\theta_\lambda(t) \nabla_\xi v_1\big) + \nabla_\xi J\big(\partial_t u + v^t\big) \\ &= \nabla_s \xi + J\nabla_t \xi + \cos\theta_\lambda \nabla_\xi v_1 - \sin\theta_\lambda \nabla_\xi v_0 + \nabla_\xi J(\partial_t u), \end{aligned} \tag{7}$$

where $v^t = \cos\theta_\lambda v_0 + \sin\theta_\lambda v_1$.

Theorem 2.6. *Assume the hypothesis of Lemma 2.5 is satisfied. Then for each zero u of Υ with $E(u) < \infty$, i.e. a solution of the equations*

$$\begin{gathered} \partial_s u + J\big(\partial_t u + \cos\theta_\lambda(t) v_0(u) + \sin\theta_\lambda(t) v_1(u)\big) = 0, \\ \lim_{t\to\pm\infty} u(s,t) = m_\mp, \qquad \lim_{s\to\pm\infty} u(s,t) = \gamma_\mp(t) \end{gathered}$$

with a finite energy, the linearization $D_u\Upsilon\colon W^{k,p}(\mathbb{R}^2; u^*TM) \to W^{k-1,p}(\mathbb{R}^2; u^*TM)$ is Fredholm.

Proof. Since most of the arguments involved in the proof of this theorem are well-known and documented, we outline here the main steps only. More details can be found in [Sal2, SZ, Don2] and references therein.

Observe that from the finiteness of $E(u)$ we obtain that $\partial_t u - v^t(u) \to 0$ uniformly as $s \to \pm\infty$, i.e. the family of curves $\gamma_s(t) = u(s,t)$ converges to $\gamma_\mp(t)$ in the C^1-topology as $s \to \pm\infty$. This follows from an estimate for the energy density in the same vein as in the proof of Proposition 1.21 in [Sal2] (see also references therein). The rest of the proof consists of three steps.

Step 1. *For any solution γ of (3) denote $E = \pi^*\gamma^*TM$, where $\pi\colon \mathbb{R}^2_{s,t} \to \mathbb{R}_t$ is the projection. Consider the linear operator*

$$L(\xi) = \partial_s\xi + J(\nabla_t \xi + B_t(\gamma)\xi), \qquad \xi \in \Gamma(E),$$

where B_t is a section of $End(TM)$, namely $B_t\xi = \nabla_\xi v^t$. Then $L\colon W^{k,p}(E) \to W^{k-1,p}(E)$ is an isomorphism for all $k \in \mathbb{R}$ and $p \in (1, +\infty)$.

In view of Lemma 2.4 it is enough to prove the statement for $k = 1, p = 2$. Choosing a unitary trivialization of γ^*TM, the operator L can be represented in the form

$$L(\xi_0) = \partial_s \xi_0 + J_0\, \partial_t \xi_0 + S(t)\xi_0, \quad \xi_0 \colon \mathbb{R}^2 \to \mathbb{R}^{2n}, \qquad J_0 = \begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}.$$

Here $S(t) = S(t)^T$ is a real-valued $2n \times 2n$-matrix. Denote $A = J_0\partial_t + S(t)$ and observe that in the chosen trivialization the operator $-J_0 A$ represents the linearization of Φ_λ at the point γ. Hence by Lemma 2.5 the operator $A \colon W^{1,2}(\mathbb{R}; \mathbb{R}^{2n}) \to L^2(\mathbb{R}; \mathbb{R}^{2n})$ is an isomorphism so that we have an estimate of the shape

$$\|\xi_1\|_{W^{1,2}} \le C_1 \|A\xi_1\|_{L^2}, \qquad \xi_1 \in W^{1,2}(\mathbb{R}; \mathbb{R}^{2n}). \tag{8}$$

Pick any $\eta_0 \in L^2(\mathbb{R}^2; \mathbb{C}^n) \cap C^\infty(\mathbb{R}^2; \mathbb{C}^n)$ and apply the Fourier transform in the variable s to the equation $L(\xi_0) = \eta_0$ to obtain

$$i\sigma \hat{\xi}_0(\sigma, t) + A\hat{\xi}_0(\sigma, t) = \hat{\eta}_0(\sigma, t).$$

Since A is symmetric and 0 does not belong to the spectrum of A, the above equation is solvable for any real σ. Applying the inverse Fourier transform we obtain a solution ξ_0 of the initial equation $L(\xi_0) = \eta_0$. Moreover, with the help of (8) an easy computation yields the estimate $\|\xi_0\|_{W^{1,2}} \le C_2 \|\eta_0\|_{L^2}$. This implies that L is an isomorphism.

Step 2. *Denote $Q_R = \{(s,t) \in \mathbb{R}^2 \mid \max\{|s|, |t|\} \le R\}$. Then there exists $R > 0$ such that for any $\xi \in \Gamma(u^*TM)$ with $\operatorname{supp} \xi \subset \mathbb{R}^2 \setminus Q_R$ we have an estimate of the shape*

$$\|\xi\|_{W^{k,p}} \le C_3 \|D_u \Upsilon(\xi)\|_{W^{k-1,p}}. \tag{9}$$

Here the constant C_3 depends on k, p and R, but not on ξ.

Obviously, Step 2 will follow if we can prove estimate (9) in the following four cases: The support of ξ is contained in the left, right, upper and lower half-planes.

We first assume that $\operatorname{supp} \xi \subset H_R = \{s \ge R\}$, where $R > 0$ will be chosen below. With the help of the parallel transport along the family of curves $s \mapsto u(s,t),\ s \ge 0$, we obtain an isomorphism $u^*TM|_{H_0} \cong \pi^*\gamma_-^*TM|_{H_0}$. Since $u(s,t) \to \gamma_-(t)$ as $s \to +\infty$, it follows from Step 1 that there exist $R > 0$ and $C_3 > 0$ such that $\|\xi\|_{W^{k,p}} \le C_3 \|D_u \Upsilon(\xi)\|_{W^{k-1,p}}$ provided $\operatorname{supp} \xi \subset H_R$. The case $\operatorname{supp} \xi \subset \{s \le -R\}$ is analyzed similarly.

In fact the same arguments apply for the cases $\operatorname{supp} \xi \subset \{t \ge R\}$ and $\operatorname{supp} \xi \subset \{t \le -R\}$. Indeed, $B_t(m_\pm)$ are the Hessians of $\operatorname{Re}(e^{-i\theta_\pm} f)$ at the corresponding points for $|t| \ge 1$. Hence $B_t(m_\pm)$ are nondegenerate and independent of t provided $|t| \ge 1$. Therefore the same arguments as in the proof of Step 1 yield that the operators $\partial_s + J_0(\partial_t + B_t(m_\mp)) \colon W^{k,p}(\mathbb{R}^2; \mathbb{R}^{2n}) \to W^{k-1,p}(\mathbb{R}^2; \mathbb{R}^{2n})$ are isomorphisms. Hence if $\operatorname{supp} \xi \subset \{|t| \ge R\}$ and R is sufficiently large, estimate (9) holds. This finishes the proof of Step 2.

Step 3. *We prove the theorem.*

With the help of Step 2 and the a priori estimate $\|\xi\|_{W^{k,p}} \le C_4\big(\|\xi\|_{L^p} + \|D_u \Upsilon(\xi)\|_{W^{k-1,p}}\big)$ we obtain that the inequality

$$\|\xi\|_{W^{k,p}(\mathbb{R}^2)} \le C_5 \Big(\|\xi\|_{L^p(Q_R)} + \|D_u \Upsilon(\xi)\|_{W^{k-1,p}(\mathbb{R}^2)} \Big)$$

holds for all $\xi \in W^{k,p}(u^*TM)$ provided R is large enough. It follows that $D_u\Upsilon$ has a finite-dimensional kernel and closed range. The cokernel of $D_u\Upsilon$ is the kernel of its formal adjoint operator, which is of the similar structure as $D_u\Upsilon$, and therefore is also finite-dimensional. □

To compute the index of $D_u\Upsilon$ we need some preparation first. Notice that $\mathrm{Ind}(D_u\Upsilon)$ depends neither on k nor on p, so we can put $k = 1,\ p = 2$. Further, to an arbitrary C^1–curve $\gamma\colon \mathbb{R} \to M$ satisfying

$$\lim_{t\to\pm\infty} \gamma(t) = m_\mp \quad \text{and} \quad \lim_{t\to\pm\infty} \dot\gamma(t) = 0 \tag{10}$$

we associate a pair of Lagrangian subspaces in $T_{\gamma(0)}M$ as follows. Consider the operator

$$A\colon C^\infty(\gamma^*TM) \to C^\infty(\gamma^*TM), \qquad A\xi = J\nabla_t\,\xi + S\xi,$$

where S is a zero-order operator, namely $S\xi = \nabla_\xi\,\tilde v^t + (\nabla_\xi J)\dot\gamma,\ \tilde v^t = \cos\theta_\lambda(t)v_1 - \sin\theta_\lambda(t)v_0$ and λ is sufficiently small (compare with (7)). Notice that $\lim\limits_{t\to\mp\infty} S = S^\pm \in End(T_{m_\pm}M)$ is the Hessian of $\mathrm{Im}\,(e^{-i\theta_\pm}f)$ at $m_\pm$. Since at the point $m = m_\pm$ we have

$$JS^\pm\xi = J\nabla_\xi\,\tilde v^t = -\nabla_\xi\,(\cos\theta_\pm v_0 + \sin\theta_\pm v_1) = -\,\mathrm{Hess}\,\mathrm{Re}\,\big(e^{-i\theta_\pm}f\big)(\xi),$$

the endomorphism $JS^\pm$ is self-adjoint with the vanishing signature (i.e. $JS^\pm$ has n positive and n negative eigenvalues).

Denote by $\xi_{\mathrm{v}},\ \mathrm{v} \in \mathrm{T}_{\gamma(0)}\mathrm{M}$, a solution of the Cauchy problem $A\xi_{\mathrm{v}} = 0,\ \xi_{\mathrm{v}}(0) = \mathrm{v}$ and put

$$\Lambda^\pm = \{\, \mathrm{v} \in \mathrm{T}_{\gamma(0)}\mathrm{M} \ \mid\ \lim_{\mathrm{t}\to\mp\infty} \xi_{\mathrm{v}}(\mathrm{t}) = 0\}.$$

Then $\Lambda^\pm$ are Lagrangian subspaces. Indeed, a straightforward computation shows that $\omega(\xi_{\mathrm{v}}(t), \xi_{\mathrm{w}}(t))$ does not depend on t for any $\mathrm{v}, \mathrm{w} \in \mathrm{T}_{\gamma(0)}\mathrm{M}$. Therefore, if $\mathrm{v}, \mathrm{w} \in \Lambda^+$, then $\omega(\mathrm{v}, \mathrm{w}) = 0$ since $\omega(\xi_{\mathrm{v}}(t), \xi_{\mathrm{w}}(t))$ vanishes at $-\infty$. Besides, $\dim \Lambda^+ = n$ since the signature of JS^+ vanishes.

Remark 2.7. If $\mathrm{v} \in \Lambda^\pm$, then ξ_{v} decays exponentially fast at $\mp\infty$ since $JS^\mp$ is nondegenerate and self-adjoint. Hence, the kernel of the operator $A\colon W^{1,2}(\gamma^*TM) \to L^2(\gamma^*TM)$ can be identified with $\Lambda^+ \cap \Lambda^-$. In particular, $\ker A$ is nontrivial if and only if $\Lambda^+ \cap \Lambda^- \neq \{0\}$.

Further, pick any two curves $\gamma_\pm$ satisfying (10) such that the associated pairs of Lagrangian subspaces are transversal. Let $u\colon \mathbb{R}^2 \to M$ be any C^1–map such that each curve $\gamma_s(t) = u(s,t)$ also satisfies (10) and $\gamma_s \to \gamma_\pm$ as $s \to \mp\infty$ in the C^1-topology. With the help of the relative Maslov index for Lagrangian pairs [RS] we associate to the triple $(\gamma^+, \gamma^-; u)$ an integer $\mu(\gamma^+, \gamma^-; u)$, which is referred to as the *relative Maslov index*. To define $\mu(\gamma^+, \gamma^-; u)$ denote by $\mathcal{L}(TM)$ the Lagrangian Grassmannian bundle and put $\bar\gamma(s) = \gamma_s(0) = u(s,0)$. Then we obtain a pair of sections (Λ^+, Λ^-) of the bundle $\bar\gamma^*\mathcal{L}(TM)$ such that the subspaces $\Lambda^+(s)$ and $\Lambda^-(s)$ are transversal for $s = \pm\infty$. Choose a unitary trivialization of $\bar\gamma^*TM$ and represent $\Lambda^\pm$ by a pair of curves $\Lambda_0^\pm\colon \mathbb{R} \to \mathcal{L}(\mathbb{R}^{2n})$. It is said that a crossing, i.e. a point s_0 such that $\Lambda^+(s_0) \cap \Lambda^-(s_0) \neq 0$, is *regular* if the associated crossing form [RS] $\Gamma(\Lambda^+,\Lambda^-,s_0)\colon \Lambda^+(s_0)\cap\Lambda^-(s_0)\to\mathbb{R}$ is nondegenerate. If all crossings are regular, then the number

$$\mu(\gamma^+,\gamma^-;u)=\mu(\Lambda_0^+,\Lambda_0^-)=\sum_{s_0 \text{ is crossing}} \operatorname{sign}\Gamma(\Lambda^+,\Lambda^-,s_0)\ \in\mathbb{Z}$$

does not depend on the choice of the unitary trivialization, i.e. the relative Maslov index is well-defined.

Proposition 2.8. *With the same notations as in Theorem 2.6, the index of $D_u\Upsilon$ is given by*

$$\operatorname{Ind}(D_u\Upsilon)=\mu(\gamma^+,\gamma^-;u).$$

Proof. We follow the line of argument in [Sal2].

Choose a C^1–small perturbation $\hat u$ of the map u with the following properties: $D_{\hat u}\Upsilon$ is Fredholm, $\operatorname{Ind}(D_{\hat u}\Upsilon)=\operatorname{Ind}(D_u\Upsilon)$, the Lagrangian pairs associated to the curves $\hat u(\pm\infty,t)$ are transversal, and there exists $T>0$ such that $\hat u(s,\pm t)=m_\mp$ for all $t\geq T$. Choose also a unitary trivialization of $\hat u^*TM$ and write $D_{\hat u}\Upsilon$ in the form

$$L(\xi)=\partial_s\xi+A(s)\xi,\qquad \xi\colon\mathbb{R}^2\to\mathbb{R}^{2n},$$

where $A(s)\xi=J_0\partial_t\xi+S(s,t)\xi$ and the matrix $S(s,t)$ is symmetric for all (s,t). By the choice of $\hat u$ we also have $S(s,\pm t)=S^\mp$ for $t\geq T$, where $S^\mp$ represents the Hessian of $\operatorname{Im}(e^{-i\theta_\mp}f)$ at $m_\mp$.

Further, denote $\mu_0=\min\{|\mu|\ :\ \ker(J_0S^\pm-\mu)\neq 0\}>0$ and consider $A(s)$ as an unbounded operator in $L^2(\mathbb{R};\mathbb{R}^{2n})$ with the domain $W^{1,2}(\mathbb{R};\mathbb{R}^{2n})$. Then for all $s\in\mathbb{R}$ any point of the spectrum of $A(s)$ from the interval $(-\mu_0,\mu_0)$ is an eigenvalue. Indeed, for any $(s,\mu)\in\mathbb{R}\times(-\mu_0,\mu_0)$ the operator $A(s)-\mu\colon W^{1,2}(\mathbb{R};\mathbb{R}^{2n})\to L^2(\mathbb{R};\mathbb{R}^{2n})$ is Fredholm, since $J_0S^\pm-\mu$ is nondegenerate. Moreover, $\operatorname{Ind}(A(s)-\mu)=\operatorname{Ind}A(s)=\operatorname{Ind}(-J_0A(s))=0$ since the indices of $-J_0A(s)$ and the linearization of Φ_λ coincide. Hence, if μ is not an eigenvalue of $A(s)$, then $A(s)-\mu\colon W^{1,2}(\mathbb{R};\mathbb{R}^{2n})\to L^2(\mathbb{R};\mathbb{R}^{2n})$ is bijective and therefore μ belongs to the resolvent set of $A(s)$.

From the above observation follows [APS] that the index of L can be computed with the help of the spectral flow of $A(s)$. Namely, a point s_0 is said to be a regular crossing of the family $A(s)$ if $\ker A(s_0)\neq 0$ and the crossing form

$$\Gamma(A,s_0)\xi=\langle\xi,(\partial_sA)\xi\rangle_{L^2}=\langle\xi,\partial_sS(s_0,\cdot)\xi\rangle_{L^2},\qquad \xi\in\ker A(s_0)$$

is nondegenerate. Then, if $A(s)$ has only regular intersection points, we have: $\operatorname{Ind}L=\sum_{s_0}\operatorname{sign}\Gamma(A,s_0)$, where the summation runs over all crossings s_0.

It follows from Remark 2.7 that crossings of $(\Lambda_0^+,\Lambda_0^-)$ and $A(\cdot)$ coincide. Therefore to complete the proof it suffices to show that under the natural identification $\Lambda_0^+(s_0)\cap\Lambda_0^-(s_0)\cong\ker A(s_0)$ the associated crossing forms coincide at each crossing s_0 (we can assume that only regular crossings occur).

Let $\Xi(s,t)$ be the solution operator of $A(s)$, i.e. $\Xi(s,t)$ is a square matrix of dimension $2n$ satisfying

$$J_0\partial_t\Xi+S(s,t)\Xi=0,\quad \Xi(s,0)=\mathbb{1}.$$

Since $S(s,t)$ is symmetric, $\Xi(s,t)\in Sp(2n;\mathbb{R})$ for all (s,t). From the equality

$$\partial_t(\Xi^T J_0\partial_s\Xi) = -(\Xi^T S J_0)J_0\partial_s\Xi + \Xi^T J_0\partial_s(J_0 S\Xi) = -\Xi^T\partial_s S\,\Xi$$

we obtain $\big\langle \Xi\xi_0,\, \partial_s S\Xi\xi_0\big\rangle = -\partial_t\,\big\langle \Xi\xi_0,\, J_0\partial_s\Xi\xi_0\big\rangle = \partial_t\,\omega_0\big(\Xi\xi_0,\, \partial_s\Xi\xi_0\big)$, where $\xi_0\in\mathbb{R}^{2n}$. Hence, for any crossing s_0 and any $\xi_0\in\Lambda_0^+(s_0)\cap\Lambda_0^-(s_0)$ we have:

$$\begin{aligned}\Gamma(A,s_0)\xi_0 &= \int_{-\infty}^{+\infty}\big\langle \Xi(s_0,t)\xi_0,\, \partial_s S(s_0,t)\Xi(s_0,t)\xi_0\big\rangle\,dt\\ &= \lim_{t\to+\infty}\omega_0\big(\Xi(s_0,t)\xi_0,\, \partial_s\Xi(s_0,t)\xi_0\big) - \lim_{t\to-\infty}\omega_0\big(\Xi(s_0,t)\xi_0,\, \partial_s\Xi(s_0,t)\xi_0\big).\end{aligned}\tag{11}$$

On the other hand, for ξ_0 as above and for all s from a sufficiently small neighbourhood of s_0 there exists $\xi^-(s)\in\Lambda^-(s_0)$ such that $\xi_0+\xi^-(s)\in\Lambda^+(s)$, i.e.

$$\lim_{t\to+\infty}\Xi(s_0,t)\xi^-(s) = 0 \quad\text{and}\quad \lim_{t\to-\infty}\Xi(s,t)(\xi_0+\xi^-(s)) = 0.$$

For $t\le -T$ we must have $\Xi(s,t)(\xi_0+\xi^-(s)) = \sum_{j=1}^n c_j(s)e^{\lambda_j t}$, where $\lambda_1,\dots,\lambda_n$ are positive eigenvalues of the matrix J_0S^+. Hence, $\partial_s\Xi(s_0,t)\xi_0 + \Xi(s_0,t)\partial_s\xi^-(s_0)\to 0$ as $t\to-\infty$ and this in turn implies

$$\begin{aligned}\omega_0\big(\xi_0, \partial_s\xi^-(s_0)\big) &= \omega_0\big(\Xi(s_0,t)\xi_0, \Xi(s_0,t)\partial_s\xi^-(s_0)\big)\\ &= -\lim_{t\to-\infty}\omega_0\big(\Xi(s_0,t)\xi_0, \partial_s\Xi(s_0,t)\xi_0\big).\end{aligned}\tag{12}$$

Similarly, there also exists $\xi^+(s)\in\Lambda^+(s_0)$ for all s sufficiently close to s_0 such that $\xi_0+\xi^+(s)\in\Lambda^-(s)$. Arguing as above, wee see that

$$\omega_0\big(\xi_0, \partial_s\xi^+(s_0)\big) = -\lim_{t\to+\infty}\omega_0\big(\Xi(s_0,t)\xi_0, \partial_s\Xi(s_0,t)\xi_0\big).\tag{13}$$

Since by definition $\Gamma(\Lambda^+,\Lambda^-,s_0)\xi_0 = \omega_0\big(\xi_0,\partial_s\xi^-(s_0)\big) - \omega_0\big(\xi_0,\partial_s\xi^+(s_0)\big)$, combining (11)-(13) we finally obtain $\Gamma(\Lambda^+,\Lambda^-,s_0)\xi_0 = \Gamma(A,s_0)\xi_0$. This finishes the proof. □

3 A gauge theory on 5-manifolds

Let E be a five-dimensional oriented Euclidean vector space with a preferred vector $\mathrm{v}\in E$ of unit norm. Let $\eta(\cdot) = \langle \mathrm{v},\cdot\rangle$ denote the corresponding 1-form. Then the linear map

$$T_\eta\colon \Lambda^2E^*\longrightarrow\Lambda^2E^*,\quad \omega\mapsto *(\omega\wedge\eta)$$

has three eigenvalues $\{-1,0,+1\}$ and the space Λ^2E^* decomposes as the direct sum of the corresponding eigenspaces:

$$\Lambda^2E^*\cong\Lambda^2_-E^*\oplus\Lambda^2_0E^*\oplus\Lambda^2_+E^*.$$

Indeed, denote by H the orthogonal complement of v. Then $\Lambda^2E^*\cong\Lambda^2H^*\oplus H^*$ and one easily checks that the following subspaces $\Lambda^2_\pm H^*$ and H^* are eigenspaces of T_η, where $\Lambda^2_\pm H^*$ denote the eigenspaces of the four-dimensional Hodge star operator. In other words, $\Lambda^2_\pm E^*\cong\Lambda^2_\pm H^*$ and $\Lambda^2_0E^*\cong H^*$.

Identify the Clifford algebra of E with ΛE and recall the following description of the Clifford multiplication

$$\begin{aligned} &Cl\colon E^*\otimes\Lambda E^*\longrightarrow\Lambda E^*, && Cl=Cl'+Cl'',\\ &Cl'\colon E^*\otimes\Lambda^pE^*\cong E\otimes\Lambda^pE^*\xrightarrow{\;c\;}\Lambda^{p-1}E^*, && c(e\otimes\omega)=-\imath_e\omega,\\ &Cl''\colon E^*\otimes\Lambda^pE^*\xrightarrow{\;\cdot\wedge\cdot\;}\Lambda^{p+1}E^*. \end{aligned}$$

In particular, by restriction we get a map $Cl'\colon E^*\otimes\Lambda^2_+E^*\longrightarrow E^*$, which is essentially the four-dimensional homomorphism $H^*\otimes\Lambda^2_+H^*\longrightarrow H^*$.

Observe that $\Lambda^2_+H^*$ has a natural structure of a Lie algebra as a three-dimensional oriented Euclidean vector space. For an arbitrary Lie algebra $\mathfrak{g}$ denote $V=\Lambda^2_+H^*\otimes\mathfrak{g}$ and consider the linear map $\sigma\colon V\otimes V\to V,\ \sigma=\frac{1}{2}[\cdot\,,\cdot]_{\Lambda^2_+H^*}\otimes[\cdot\,,\cdot]_{\mathfrak{g}}$. Choosing a Lie algebra isomorphism $\Lambda^2_+H^*\cong\mathbb{R}^3$, for $\xi=e_1\otimes\xi_1+e_2\otimes\xi_2+e_3\otimes\xi_3$ we obtain

$$\sigma(\xi,\xi)=e_1\otimes[\xi_2,\xi_3]+e_2\otimes[\xi_3,\xi_1]+e_3\otimes[\xi_1,\xi_2].$$

Let (W^5,g) be an arbitrary oriented Riemannian five-manifold with a preferred vector field v of pointwise unit norm. Denote $\eta(\cdot)=g(v,\cdot)\in\Omega^1(W)$ and $\mathcal{H}=\ker\eta\subset TW$. As described above, we have the following splittings:

$$\begin{aligned} &\Omega^1(W)=\Omega^1_h(W)\oplus\Omega^0(W)\eta, \qquad \Omega^1_h(W)=\Gamma(\mathcal{H}^*),\\ &\Omega^2(W)=\Omega^2_-(W)\oplus\Omega^2_0(W)\oplus\Omega^2_+(W). \end{aligned}$$

Let $P\to W$ be a principal G-bundle, where G is a compact Lie group. Denote by $\mathcal{A}(P)$ the space of connections on P and by $ad\,P$ the adjoint bundle of Lie algebras. Consider the following equations for a pair $(A,B)\in\mathcal{A}(P)\times\Omega^2_+(ad\,P)=\mathcal{B}$:

$$\begin{aligned} &\imath_vF_A-\delta^+_A\,B=0,\\ &F^+_A-\nabla^A_v\,B-\sigma(B,B)=0, \end{aligned}\qquad(14)$$

where the operator $\delta^+_A\colon\Omega^2_+(ad\,P)\to\Omega^1_h(ad\,P)$ is defined by the composition

$$\delta^+_A\colon\Gamma(\Lambda^2_+\mathcal{H}^*\otimes ad\,P)\xrightarrow{\;\nabla^{LC,\,A}\;}\Gamma(T^*W\otimes\Lambda^2_+\mathcal{H}^*\otimes ad\,P)\xrightarrow{\;Cl'\otimes id\;}\Gamma(\mathcal{H}^*\otimes ad\,P).$$

Here $\nabla^{LC,\,A}$ denotes the tensor product of A and the connection on $\Lambda^2_+\mathcal{H}^*$ induced by the Levi-Civita connection (we do not assume that $\Lambda^2_+\mathcal{H}^*$ is preserved by the Levi-Civita connection). It is convenient to define a map $\Phi\colon\mathcal{B}\to\Omega^1_h(ad\,P)\times\Omega^2_+(ad\,P)$ by the left hand side of equations (14).

Remark 3.1. Equations (14) were independently discovered by Witten [Wit2] from a different perspective. A partial case with $B\equiv 0$ has been studied by Fan [Fan].

Remark 3.2. The total space of $\Lambda^2_+\mathcal{H}\to W$ is an eight-manifold equipped with a natural $Spin(7)$-structure, which is induced by the Riemannian metric and orientation on W. This $Spin(7)$- structure can be constructed in a similar manner as in [BS]. Then, following the line of argument in [Hay], one can show that solutions of equations (14) correspond to $Spin(7)$-instantons on $\Lambda^2_+\mathcal{H}$ invariant along each fibre.

The gauge group $\mathcal{G}(P)$ acts on the configuration space $\mathcal{B}$ on the right

$$(A, B) \cdot g = (A \cdot g,\ ad_{g^{-1}} B), \qquad g \in \mathcal{G}(P),$$

where g acts on the first component by the usual gauge transformation. The infinitesimal action at a point (A, B) is given by the map

$$K\colon \Omega^0(ad\, P) \longrightarrow \Omega^1(ad\, P) \oplus \Omega^2_+(ad\, P), \qquad \xi \mapsto \big(d_A\xi,\ [B, \xi]\big).$$

Notice also that the map Φ is $\mathcal{G}(P)$-equivariant.

The standard computation yields

$$\delta\Phi_{(A,\,B)}\begin{pmatrix}\alpha\\ \beta\end{pmatrix} = \begin{pmatrix} \imath_v(d_A\alpha) - \delta_A^+\beta + \alpha\cdot B\\ d_A^+\alpha - \nabla_v^A\beta - [\alpha(v), B] - 2\sigma(B, \beta)\end{pmatrix}, \qquad (\alpha, \beta) \in T_{(A,B)}\mathcal{B},$$

where the term $\alpha \cdot B \in \Omega^1_h(ad\, P)$ is constructed algebraically from α and B, namely $\alpha\cdot B = Cl' \otimes [\cdot\,,\cdot]_{\mathfrak{g}}(\alpha\otimes B)$. Thus we get the deformation complex at the point (A, B):

$$0 \to \Omega^0(ad\, P) \xrightarrow{K} \Omega^1(ad\, P) \oplus \Omega^2_+(ad\, P) \xrightarrow{\delta\Phi} \Omega^1_h(adP) \oplus \Omega^2_+(ad\, P) \to 0. \tag{15}$$

Lemma 3.3. *Deformation complex (15) is elliptic.*

The statement of this Lemma follows immediately from Remark 3.2. Alternatively, one can consider equations (14) on $\mathbb{R}^5$ and show that the symbol of $K^* + \delta\Phi$ is modelled on the octonionic multiplication. We omit the details.

4 Dimensional reductions

Before turning our attention to the dimensional reductions of equations (14) a little digression is in place. Suppose a Lie group $\mathcal{G}$ acts *freely* and isometrically on a Riemannian manifold M. Identify a $\mathcal{G}$-invariant function $f\colon M \to \mathbb{R}$ with a function $\hat f\colon M/\mathcal{G} \to \mathbb{R}$. Then critical points of $\hat f$ correspond to orbits of solutions to the equation $\operatorname{grad} f = K_\xi$, where $\xi \in Lie(\mathcal{G})$ and K_ξ is the Killing vector field corresponding to ξ. But the invariance of f implies $\langle \operatorname{grad} f, K_{\xi'}\rangle = 0$ for any $\xi' \in Lie(\mathcal{G})$ so that we necessarily have $\operatorname{grad} f = 0$ for any point on M projecting to a critical point of $\hat f$.

Similarly, a curve $m\colon \mathbb{R} \to M$ projects to an antigradient flow of $\hat f$ if and only if there exists $\xi\colon \mathbb{R} \to Lie(\mathcal{G})$ such that

$$\dot m = -\operatorname{grad} f + K_\xi. \tag{16}$$

The Lie group $\{g\colon \mathbb{R} \to \mathcal{G}\}$ acts on solutions of equation (16) and the orbits are in bijective correspondence with the antigradient flow lines of $\hat f$. Further, we may consider only those solutions of (16), which are horizontal with respect to the natural connection. This gives a bijection between ordinary flow lines of f modulo $\mathcal{G}$ and flow lines of $\hat f$.

The upshot is that $\mathcal{G}$-invariance of f implies that equation (16) is equivalent to the ordinary antigradient flow equation of f. It will be important to switch freely between these two approaches in an infinite-dimensional setup. The reasons will be clear below.

4.1 Dimension four

Let X be a closed oriented Riemannian four-manifold. Below we consider equations (14) on $(W,v)=(X\times\mathbb{R}_t,\frac{\partial}{\partial t})$ endowed with the product metric.

Denote by $pr\colon X\times\mathbb{R}\to X$ the canonical projection and set $P=pr^*P_X$, where $P_X\to X$ is a principal G-bundle. Think of $B\in\Omega^2_+(X\times\mathbb{R};pr^*ad\,P_X)$ as a map $b\colon\mathbb{R}\to\Omega^2_+(X;ad\,P_X)$. Similarly $A\in\mathcal{A}(pr^*P_X)$ can be seen as a map $(a,c)\colon\mathbb{R}\to\mathcal{A}(P_X)\times\Omega^0(ad\,P_X)$, where c is a Higgs field. Then equations (14) are easily seen to become

$$\begin{aligned}\dot a&=\delta_a^+b+d_ac,\\ \dot b&=F_a^+-\sigma(b,b)-[c,b],\end{aligned}\tag{17}$$

where $\delta_a^+=(d_a^+)^*$. These equations turn out to be the antigradient flow equations of a certain function. Indeed, consider the function

$$h\colon\Lambda^2_+H^*\otimes\mathfrak{g}\to\mathbb{R},\qquad h(\mathrm{w})=\frac{1}{3}\langle\mathrm{w},\sigma(\mathrm{w})\rangle.$$

Choose an isomorphism $\Lambda^2_+H^*\cong\mathbb{R}^3$ and write $\mathrm{w}=\sum_{i=1}^3e_i\otimes\xi_i$. Then we have $h(\mathrm{w})=\langle\xi_1,[\xi_2,\xi_3]\rangle$ and therefore $\operatorname{grad}h(\mathrm{w})=\sigma(\mathrm{w})$. Since h is equivariant with respect to both $SO(3)$ and G, we obtain a well-defined map $\Omega^2_+(ad\,P_X)\to C^\infty(X)$ denoted by the same letter.

Denote $\mathcal{B}=\mathcal{A}(P)\times\Omega^2_+(ad\,P)/\mathcal{G}(P)$. As usual, $\mathcal{B}^*\subset\mathcal{B}$ denotes the quotient space of irreducible points. The negative L^2–gradient of the function

$$U\colon\mathcal{B}\to\mathbb{R},\qquad U(a,b)=-\langle F_a^+,b\rangle_{L^2}+\int_X h(b)\,vol_X$$

is $(\delta_a^+b,\,F_a^+-\sigma(b,b))$. Hence, assuming there are no reducible solutions, equations (17) represent the antigradient flow equations of the function $\hat U\colon\mathcal{B}^*\to\mathbb{R}$ as in (16) with $\xi=c$.

We summarize our computations in the following proposition.

Proposition 4.1. *If there are no reducible solutions, equations (17) represent antigradient flow equations of the function $\hat U\colon\mathcal{B}^*\to\mathbb{R}$.* □

The critical points of the function U are solutions of the Vafa-Witten equations [VW]:

$$\begin{aligned}&\delta_a^+b+d_ac=0,\\ &F_a^+-\sigma(b,b)+[b,c]=0.\end{aligned}$$

These equations are elliptic and the expected dimension of the moduli space is zero.

As we have seen, the $\mathcal{G}(P)$-invariance of U implies that for each irreducible solution of the Vafa-Witten equations we have $(d_ac,[b,c])=0$, i.e. in the absence of reducible solutions the above equations are equivalent to

$$\begin{aligned}&\delta_a^+b=0,\\ &F_a^+-\sigma(b,b)=0.\end{aligned}\tag{18}$$

Notice that the Weitzenböck formula

$$2d_a^+\delta_a^+=(\nabla^a)^*\nabla^a-2W^++\frac{s}{3}+\sigma(F_a^+,\cdot),$$

for the operator $d_a^+\delta_a^+ \colon \Omega^2_+(ad\,P) \to \Omega^2_+(ad\,P)$ yields

$$\begin{aligned} 4\|\delta_a^+ b\|^2 + \|F_a^+ - \sigma(b,b)\|^2 &= 2\|\nabla^a b\|^2 - 4\langle W^+(b), b\rangle + \frac{2}{3}\langle sb, b\rangle + 2\langle F_a^+, \sigma(b,b)\rangle \\ &\quad + \|F_a^+\|^2 + \|\sigma(b,b)\|^2 - 2\langle F_a^+, \sigma(b,b)\rangle \\ &= 2\|\nabla^a b\|^2 - 4\langle W^+(b), b\rangle + \frac{2}{3}\langle sb, b\rangle + \|F_a^+\|^2 + \|\sigma(b,b)\|^2. \end{aligned}$$

Proposition 4.2 ([VW])**.** *If the operator $-W^+ + \frac{1}{6}s$ is pointwise non-negative definite on $\Lambda^2_+ T^*X$, then for any irreducible solution (a,b) of the Vafa-Witten equations the following holds: $F_a^+ = 0$, $\nabla^a b = 0$.* □

4.2 Dimension three

In this section various forms of equations (14) are studied on $Y^3 \times \mathbb{R}^2$, where Y is a closed oriented Riemannian three-manifold.

Similarly as in the instanton Floer theory, consider solutions of Vafa-Witten equations (18) on $X = Y \times \mathbb{R}$. Assuming a is in a temporal gauge, we obtain the following system of equations

$$\begin{aligned} \dot a &= -*(F_a - \tfrac{1}{2}[b\wedge b]), \\ \dot b &= *d_a b, \\ 0 &= \delta_a b, \end{aligned} \tag{19}$$

where (a,b) is interpreted as a curve in $\mathcal{A}(P)\times\Omega^1(ad\,P) \cong T^*\mathcal{A}(P)$. Here we have also used the isomorphism $\Gamma(\pi^*T^*Y) \cong \Omega^2_+(Y\times\mathbb{R})$, $\omega \mapsto \frac{1}{2}(*_3\omega + ds\wedge\omega)$, where $\pi\colon Y\times\mathbb{R}\to Y$ is the projection.

Observe that $T^*\mathcal{A}(P)$ is a (flat) Kähler manifold and the action of the gauge group is Hamiltonian. The momentum map is given by

$$\mu\colon T^*\mathcal{A}(P) \to \Omega^0(ad\,P), \qquad \mu(a,b) = \delta_a b. \tag{20}$$

Denote $N = \mu^{-1}(0) = \{(a,b) \mid \delta_a b = 0\} \subset T^*\mathcal{A}(P)$. It follows from the very definition of the momentum map that $d\mu$ is surjective at (a,b) if and only if the gauge group acts locally freely at (a,b). Therefore, the subset N^* consisting of all irreducible points of N is a submanifold. Hence, $N^*/\mathcal{G}(P)$ is a Kähler manifold.

Consider the map

$$f_0\colon \mathcal{A}(P)\times\Omega^1(ad\,P) \to \mathbb{R}/\mathbb{Z}, \qquad f_0(a,b) = 8\pi^2\vartheta(a) - \tfrac{1}{2}\langle b, *d_a b\rangle_{L^2},$$

where ϑ is the Chern-Simons function. It is easy to check that the vector field grad $f_0 = \big(*(F_a - \frac{1}{2}[b\wedge b]),\ -*d_a b\big)$ is tangent to N^* at each point. Therefore critical points of the restriction of f_0 to N^* are solutions to Hitchin's equations[4] [Hit]:

$$\begin{aligned} F_a - \tfrac{1}{2}[b\wedge b] &= 0, \\ d_a b &= 0, \\ \delta_a b &= 0. \end{aligned} \tag{21}$$

[4] Hitchin studied these equations in the case of two-dimensional base manifolds.

More accurately, in the same manner as described at the beginning of this section, orbits of irreducible solutions to (21) correspond to critical points of $\hat{f}_0\colon N^*/\mathcal{G}(P) \to \mathbb{R}$. Similarly, orbits of (19) correspond to the flow lines of $\hat{f}_0$.

Remark 4.3. Denote by G^c the complexified Lie group and by $\mathcal{P} = P \times_G G^c$ the principal G^c-bundle associated with P. Any connection on $\mathcal{P}$ can be written in the form $\mathcal{A} = a + ib$, where $(a, b) \in \mathcal{A}(P) \times \Omega^1(ad\, P)$ and vice versa any pair (a, b) combines to a G^c connection $\mathcal{A}$. Then $\mathcal{A}$ is flat if and only if the first two equations of (21) are satisfied. The last equation, i.e. the vanishing of the moment map, has been analyzed in [Don1, Cor].

Remark 4.4. Hitchin's equations can be obtained from $SU(3)$-asd equations along similar lines as in Remark 3.2. Namely, the total space of T^*Y is equipped with an $SU(3)$-structure. Then $SU(3)$-instantons invariant along each fibre are solutions of Hitchin's equations.

We can also consider equations (14) on $W = \mathbb{R}_t \times Y \times \mathbb{R}_s$ with $v = -\frac{\partial}{\partial s}$. Write $A = a + e\, ds + c\, dt$, where a is a family of connections on $P \to Y$. Consider first only t-invariant solutions with $c = e = 0$. A computation yields the following system:

$$\begin{aligned} \dot{a} &= - *d_a b, \\ \dot{b} &= - *(F_a - \tfrac{1}{2}[b \wedge b]), \\ 0 &= \delta_a b, \end{aligned} \tag{22}$$

where the dots denote the derivative with respect to the variable s. Equations (22) and (19) have appeared in [KW] and were further studied in [Wit1, Wit3].

Consider the function

$$f_1\colon T^*\mathcal{A}(P) \to \mathbb{R}, \qquad f_1(a, b) = \langle F_a, *b\rangle_{L^2} - \int_Y h(b)\, vol_Y.$$

Since $\operatorname{grad} f_1 = (*d_a b, *(F_a - \frac{1}{2}[b \wedge b]))$ is tangent to N^* at each point we conclude that the moduli of solutions to equations (22) correspond to antigradient flow lines of $\hat{f}_1\colon N^*/\mathcal{G}(P) \to \mathbb{R}$.

Let us examine the functions f_0 and f_1 more closely. Since $\operatorname{grad} f_1 = J \operatorname{grad} f_0$, where J is the constant complex structure on $T^*\mathcal{A}(P) \cong \Omega^1(ad\, P) \otimes \mathbb{C}$, we obtain that the function $f = f_0 + i f_1$ is J-holomorphic. Writing (a, b) as a G^c-connection $\mathcal{A}$ as in Remark 4.3 it is easy to check that f is the complex Chern-Simons functional

$$\mathrm{CS}(\mathcal{A}) = \frac{1}{2} \int_Y \left(\langle \mathcal{A} \wedge d\mathcal{A}\rangle + \tfrac{1}{3}\langle \mathcal{A} \wedge [\mathcal{A} \wedge \mathcal{A}]\rangle \right).$$

Here we interpret $\mathcal{A}$ as a $\mathfrak{g}_{\mathbb{C}}$-valued 1-form on Y, and $\langle\cdot,\cdot\rangle\colon \mathfrak{g}_{\mathbb{C}} \otimes \mathfrak{g}_{\mathbb{C}} \to \mathbb{C}$ denotes the $\mathbb{C}$-linear extension of the scalar product on $\mathfrak{g}$.

Further, let us consider equations (14) on $(W, v) = (Y \times \mathbb{R}^2_{s,t}, \frac{\partial}{\partial t})$. The standard reduction procedure yields the following system

$$\begin{aligned} &\partial_s a - \partial_t b + [b, c] - d_a e + *\big(F_a - \tfrac{1}{2}[b \wedge b]\big) = 0, \\ &\partial_t a + \partial_s b - [b, e] - d_a c - *d_a b = 0, \\ &\partial_t e - \partial_s c + [c, e] + \delta_a b = 0. \end{aligned} \tag{23}$$

Here a is a connection on the pull-back of $P = P_Y$ to $Y \times \mathbb{R}^2$, b is a 1-form and $c,\ e$ are 0-forms with values in the adjoint bundle of Lie algebras. It is easy to check that these equations are symplectic vortex equations [CGS] with a Hamiltonian perturbation for the following data: The target space is $T^*\mathcal{A}(P)$ equipped with the Hamiltonian action of the gauge group $\mathcal{G}(P)$, $u = (a,\, b)\colon \mathbb{R}^2 \to T^*\mathcal{A}(P),\ A = e\, ds + c\, dt$, and the perturbation is $\sigma = \frac{1}{2}\mathrm{Im}\,(f dz) = \frac{1}{2}(f_0\, dt + f_1\, ds)$.

Notice also that we are free to rotate the coordinates s and t or, equivalently, to rotate the initial vector field $v = \partial_t$. This is in turn equivalent to the choice of the Hamiltonian perturbation $\sigma = \frac{1}{2}\mathrm{Im}\,(e^{i\theta} f dz)$ and the resulting equations are

$$\begin{aligned}
&\partial_s a - \partial_t b + [b, c] - d_a e - \sin\theta\, *d_a b + \cos\theta\, *\big(F_a - \tfrac{1}{2}[b \wedge b]\big) = 0,\\
&\partial_t a + \partial_s b - [b, e] - d_a c - \cos\theta\, *d_a b - \sin\theta\, *\big(F_a - \tfrac{1}{2}[b \wedge b]\big) = 0,\\
&\partial_t e - \partial_s c + [c, e] + \delta_a b = 0.
\end{aligned} \tag{24}$$

Remark 4.5. The above description of equations (24) is analogous to the interpretation of the anti-self-duality equations on $\mathbb{C} \times \Sigma$ as symplectic vortex equations [CGS]. A new phenomenon here is the appearance of the Hamiltonian perturbation. Notice also that the adiabatic limit procedure as in [CGS, GS] for equations (24) yields (at least formally) holomorphic planes to $T^*\mathcal{A}(P)/\!\!/\mathcal{G}(P)$ with a Hamiltonian perturbation.

For solutions of equations (24) invariant with respect to s we obtain the following system

$$\begin{aligned}
\dot a &= \quad \cos\theta\, *d_a b + \sin\theta\, *\big(F_a - \tfrac{1}{2}[b \wedge b]\big) + d_a c + [b, e],\\
\dot b &= -\sin\theta\, *d_a b + \cos\theta\, *\big(F_a - \tfrac{1}{2}[b \wedge b]\big) - d_a e + [b, c],\\
\dot e &= -\delta_a b + [e, c].
\end{aligned} \tag{25}$$

Notice that if $c = e = 0$ we obtain equations (19) and (22) for $\theta = -\pi/2$ and $\theta = \pi$ respectively.

It will be helpful in the sequel to give an abstract interpretation for equations (25). To do so, let M, ω, J, and f be as in Section 2 except that M is acted upon by a Lie group $\mathcal{G}$ and f is $\mathcal{G}$-invariant. The action of $\mathcal{G}$ is assumed to be Hamiltonian with moment map $\mu\colon M \to \mathfrak{G} = Lie(\mathcal{G})$. Consider the following equations for a curve (γ, ξ, η) in $M \times \mathfrak{G} \times \mathfrak{G}$:

$$\begin{aligned}
\dot\gamma &= \sin\theta\, \mathrm{grad}\, f_0(\gamma) + \cos\theta\, \mathrm{grad}\, f_1(\gamma) + K_\xi(\gamma) - JK_\eta(\gamma),\\
\dot\eta &= -\mu(\gamma) - [\xi, \eta],
\end{aligned} \tag{26}$$

where K is the Killing vector field. Clearly, we obtain equations (25) from (26) putting $M = T^*\mathcal{A}(P)$.

Further, observe that for any $\zeta, \rho \in \mathfrak{G}$ the following equalities hold:

$$d\big(\omega(K_\zeta, K_\rho)\big) = d\,\imath_{K_\rho}(\imath_{K_\zeta}\omega) = \mathcal{L}_{K_\rho}(\imath_{K_\zeta}\omega) - \imath_{K_\rho} d(\imath_{K_\zeta}\omega) = \imath_{K_{[\rho,\zeta]}}\omega = -d\,\langle \mu, [\zeta, \rho]\rangle.$$

Here the second equality follows from Cartan's equation. Hence, $\langle \mu, [\zeta, \rho]\rangle = -\omega(K_\zeta, K_\rho) = g(K_\zeta, JK_\rho)$. Therefore for any solution of equations (26) we have

$$\begin{aligned}
\frac{d}{dt}\langle \mu(\gamma), \eta\rangle &= \omega(K_\eta, \dot\gamma) + \langle \mu, \dot\eta\rangle = g(JK_\eta, K_\xi) - g(JK_\eta, JK_\eta) - \langle \mu, \mu\rangle - \langle \mu, [\xi, \eta]\rangle\\
&= -g(K_\eta, K_\eta) - \langle \mu, \mu\rangle \le 0.
\end{aligned} \tag{27}$$

Here the first equality follows from the definition of the momentum map, the second one from equations (26) and the $\mathcal{G}$-invariance of f, and the last one from the equation $\langle \mu, [\xi, \eta] \rangle = g(K_\xi, JK_\eta)$. Hence, for any solution of equations (26) the function $\langle \mu(\gamma), \eta \rangle$ is non-increasing.

We will be interested below in solutions (γ, ξ, η) of (26) satisfying the following condition

$$(\gamma, \xi, \eta) \longrightarrow (m_\pm, 0, 0) \qquad \text{as} \quad t \to \mp\infty, \tag{28}$$

where $m_\pm$ are critical points of f. For any such solution $\langle \mu(\gamma), \eta \rangle$ vanishes at $\pm\infty$ and hence vanishes everywhere. Then from (27) we conclude that η and $\mu \circ \gamma$ vanish everywhere, i.e. under condition (28) equations (26) reduce to

$$\dot\gamma = \sin\theta \operatorname{grad} f_0 + \cos\theta \operatorname{grad} f_1 + K_\xi, \qquad \mu(\gamma) = 0.$$

From the discussion at the beginning of Section 4 we obtain that these equations are equivalent to

$$\dot\gamma = \sin\theta \operatorname{grad} f_0 + \cos\theta \operatorname{grad} f_1, \qquad \mu(\gamma) = 0. \tag{29}$$

Summing up, we have that under condition (28) systems (26) and (29) are equivalent. Applying this conclusion in the case $M = T^*\mathcal{A}(P)$ we obtain that for $\theta = -\pi/2$ equations (25) together with the condition

$$(a, b, c, e) \longrightarrow (a_\pm, b_\pm, 0, 0) \qquad \text{as} \quad t \to \pm\infty,$$

where $(a_\pm, b_\pm)$ are solutions of Hitchin's equations, are equivalent to equations (19) together with $(a, b) \to (a_\pm, b_\pm)$ as $t \to \mp\infty$. The upshot is that while equations (19) and (25) with $\theta = -\pi/2$ are essentially equivalent, only the latter are elliptic.

Remark 4.6. In a similar spirit as above, one obtains an elliptic form of Hitchin's equations on a three manifold by considering solutions of equations (14) on $(Y \times \mathbb{R}^2, \partial_t)$ invariant along $\mathbb{R}^2$. The corresponding equations are easily obtained from (23).

5 Invariants

In this section we outline constructions of invariants assigned to five-, four-, and three-manifolds arising from gauge theories described in the preceding sections. It is clear that the constructions described below need an appropriate analytic justification. We postpone it to subsequent papers and restrict ourselves to some examples. In particular the problems of compactness and transversality are not studied here. Throughout this section the coefficient ring is $\mathbb{Z}/2\mathbb{Z}$ in all constructions for the sake of simplicity.

The expected dimension of the moduli space of solutions to equations (14) for closed five-manifolds is zero. Therefore, assuming compactness and transversality, an algebraic count associates a number to closed five-manifolds. More accurately, this number depends on the isomorphism class of P and on the class of the vector field v in $\pi_0(\mathfrak{X}_0(W))$, where $\mathfrak{X}_0(W)$ denotes the space of all vector fields on W without zeros.

Let us now consider the dimension four. The corresponding construction is very similar to the instanton Floer theory, so we are very brief here. Assume the moduli space of solutions to the Vafa-Witten equations $\mathcal{M}_{VW}$ is compact and zero-dimensional (for the case $\dim \mathcal{M}_{VW} > 0$ see example below). The index of the Hessian on $X^4 \times S^1$ vanishes and therefore the relative Morse index of a pair of critical points is an integer.[5] The Floer differential counts the moduli space of finite-energy solutions to equations (14) on $X \times \mathbb{R}$ converging to solutions of the Vafa-Witten equations at $\pm\infty$. As a result, for a smooth four-manifold equipped with a principal G-bundle Floer-type homology groups can conjecturally be constructed.

Example 5.1. Let X be a Kähler surface with a non-negative scalar curvature. Then Proposition 4.2 applies and, therefore, $\mathcal{M}_{VW} = \mathcal{M}_{asd}$ assuming all asd connections are irreducible and non-degenerate. If $\dim \mathcal{M}_{asd} > 0$ the function U is not Morse but rather Morse-Bott. Then choosing a suitable perturbation, which is essentially a Morse function h on $\mathcal{M}_{asd}$, one obtains the Morse-Witten complex of h. The details can be found for instance in [BH]. In other words, the corresponding Floer homology groups are homology groups of $\mathcal{M}_{asd}$. Notice that this agrees perfectly with the Vafa-Witten theory: The Vafa-Witten invariant, which counts solutions of the Vafa-Witten equations, is the Euler characteristic of $\mathcal{M}_{asd}$ provided the only solutions of the Vafa-Witten equations are asd instantons.

It is worth pointing out that the above reasoning is valid if $\mathcal{M}_{asd}$ admits a compactification, which is itself a manifold. Notice that the Euler characteristic of $\mathcal{M}_{asd}$ in [VW] is taken as the Euler characteristic of the Gieseker compactification.

Further, let us consider dimension three. Let (Y, g) be a closed oriented Riemannian three-manifold. Pick a nontrivial principal G-bundle $P \to Y$ and assume that all solutions to Hitchin's equations are irreducible (thus we exclude the case $G = SU(2)$) and the moduli space is finite, say $\{\mathcal{A}_1, \dots, \mathcal{A}_k\}$. Recall that this is the critical set of the complex Chern-Simons functional and therefore we can conjecturally construct a corresponding collection of k Fukaya-Seidel A_∞-categories[6] $\mathcal{A}_j(Y)$ along similar lines as in Section 2.1.

Thus, the objects of $\mathcal{A}_j(Y)$ are classes of solutions $\mathcal{A}_l$ to Hitchin's equations. These are assumed to be ordered by indexes: $\mathcal{A}_1 < \dots < \mathcal{A}_k$. Moreover we can assume that the ordering is chosen so that $\operatorname{Re}\mathrm{CS}(\mathcal{A}_1) \le \dots \le \operatorname{Re}\mathrm{CS}(\mathcal{A}_k)$, where $\operatorname{Re}\mathrm{CS}(\mathcal{A}_l)$ is understood to take values in $[0, 1)$. For the sake of simplicity suppose also that all inequalities are strict: $\operatorname{Re}\mathrm{CS}(\mathcal{A}_1) < \dots < \operatorname{Re}\mathrm{CS}(\mathcal{A}_k)$. Recall that for any pair $\mathcal{A}_\pm \in \{\mathcal{A}_1, \dots, \mathcal{A}_k\}$, $\mathcal{A}_- < \mathcal{A}_+$, the space $hom(\mathcal{A}_-, \mathcal{A}_+)$ is generated by the broken flow lines of the complex Chern-Simons functional connecting $\mathcal{A}_-$ with $\mathcal{A}_+$. More precisely, as described in Remark 2.2, we consider only those broken flow lines γ for which the image of $\mathrm{CS}\circ\gamma$ does not intersect the set $\big(\operatorname{Re}\mathrm{CS}(\mathcal{A}_j), \operatorname{Re}\mathrm{CS}(\mathcal{A}_{j+1})\big) \times \mathbb{R}$. Recall also that the flow lines of the complex Chern-Simons functional can conveniently be described as moduli of solutions to equations (25) satisfying the asymptotic conditions

$$(a, b, c, e) \longrightarrow (a^0_\pm, b^0_\pm, 0, 0) \qquad \text{as} \quad t \to \mp\infty, \tag{30}$$

where $(a^0_\pm, b^0_\pm)$ are solutions of Hitchin's equations representing $\mathcal{A}_\pm$.

[5] In general, there is no a distinguished critical point as in the $SU(2)$-instanton Floer theory, so that we are left with the relative grading only.

[6] $\mathcal{A}_j(Y)$ will also depend on the metric as well as on the choice of P.

Further, the Floer differential $\mu^1\colon hom(\mathcal{A}_-, \mathcal{A}_+) \to hom(\mathcal{A}_-, \mathcal{A}_+)$ is obtained by counting moduli of finite-energy pseudoholomorphic planes with a Hamiltonian perturbation satisfying suitable conditions at infinity. In our case, by Remark 4.5 these pseudoholomorphic planes can (formally) be interpreted as solutions to equations (24), which are in turn interpreted as solutions of equations (14) on $W = Y \times \mathbb{R}^2$.

Summing up, choose any admissible pair $\mathcal{B}_\pm$ of gauge equivalence classes of finite-energy solutions to equations (25) and (30). Then *define* the map μ^1 by counting moduli of solutions to equations (14) on $(W, v) = (Y \times \mathbb{R}^2, \cos\theta\,\partial_t + \sin\theta\,\partial_s)$ with the following boundary conditions

$$\begin{aligned} &(a,b,c,e) \to (a_\pm(t), b_\pm(t), 0, 0) && \text{as} \quad s \to \mp\infty,\\ &(a,b,c,e) \to (a^0_\pm, b^0_\pm, 0, 0) && \text{as} \quad t \to \mp\infty, \end{aligned}$$

where $(a_\pm(t), b_\pm(t))$ represents the class $\mathcal{B}_\pm$.

To define the map μ^2, one considers finite-energy solutions of equations (14) on $W = Y \times \Omega$ satisfying appropriate boundary conditions, where Ω is as shown in Fig.2. The maps μ^d for $d \geq 3$ are defined similarly and conjecturally the whole collection $\{\mu^d\}$ combines to an A_∞-structure.

Notice that the change of orientation on Y is equivalent to multiplication of f by -1 and hence does not affect $\mathcal{A}_j(Y)$. On the other hand, $\mathcal{A}_j(Y)$ depends on the Riemannian metric g. However, as explained in [Sei1] the derived category $D^b(\mathcal{A}_j(Y))$ should be independent of g.

A Pseudoholomorphic strips and pseudoholomorphic planes

In this appendix we outline (without proof) a connection between pseudoholomorphic planes with a Hamiltonian perturbation and pseudoholomorphic strips with Lagrangian boundary conditions. To do so, pick a pair (m_-, m_+) of critical points of f and assume that the interval $\overline{z_- z_+}$ does not contain any other critical point, where $z_\pm = f(m_\pm)$. It is convenient to choose the midpoint of $\overline{z_- z_+}$ as the basepoint. We deviate from our convention on the choice of the basepoint for the convenience of exposition only, namely to avoid differential equations with non-smooth coefficients.

Replacing f with $e^{-i\theta_\pm}(f - z_0)$ if necessary we may assume that $z_\pm = \pm T, T > 0$ and hence $z_0 = 0, \theta_0(t) \equiv 0$. We establish a relation between solutions of the equations

$$\begin{aligned} &\partial_s u + J(\partial_t u + v_0) = 0, \qquad u\colon \mathbb{R}^2_{s,t} \to M,\\ &\lim_{t\to\pm\infty} u(s,t) = m_\mp, \qquad \lim_{s\to\pm\infty} u(s,t) = \gamma_\mp(t) \end{aligned} \tag{31}$$

and pseudoholomorphic strips in two steps. In the first step we relate solutions of equations (31) to solutions of the problem

$$\begin{aligned} &\partial_s u_0 + J\Big(\partial_\tau u_0 + \frac{1}{\|v_0\|^2} v_0\Big) = 0, \qquad (s,\tau) \in \mathbb{R} \times (-T, T),\\ &u_0(s, \pm T) = m_\mp, \qquad \lim_{s\to\pm\infty} u_0(s,\tau) = \gamma_{0,\pm}(\tau), \end{aligned} \tag{32}$$

where $\gamma_{0,\pm}$ satisfies the equations

$$\begin{aligned} &\frac{d}{d\tau}\gamma_0 + \frac{1}{\|v_0\|^2} v_0 = 0, \qquad \tau \in (-T, T), \\ &\gamma_0(\pm T) = m_{\mp}. \end{aligned} \tag{33}$$

In the second step we show how to relate solutions of (32) to pseudoholomorphic strips.

Step 1. It is an elementary fact that equations (33) are equivalent to the antigradient flow equations for f_0. Nevertheless it is instructive to examine this equivalence more closely. Consider the family of equations

$$\begin{aligned} &\dot\gamma_\lambda + \frac{1}{\lambda + (1-\lambda)\|v_0\|^2}\, v_0 = 0, \qquad \gamma_\lambda\colon \mathbb{R} \to M, \\ &\lim_{t\to\pm\infty} \gamma_\lambda(t) = m_{\mp}, \end{aligned} \tag{34}$$

where $\lambda \in (0, 1]$, and fix a parametrization by the condition $f_0 \circ \gamma_\lambda(0) = 0$. Pick any solution γ_1 of equations (34) for $\lambda = 1$, i.e. an antigradient flow line of f_0, and consider the following family of diffeomorphisms

$$\tau_\lambda\colon \mathbb{R} \to \mathbb{R}, \quad \tau_\lambda(t) = \lambda t + (1-\lambda) f_0 \circ \gamma_1(t), \qquad \lambda \in (0, 1].$$

It is straightforward to check that $\gamma_\lambda = \gamma_1 \circ \tau_\lambda^{-1}$ is a solution of (34), and this establishes a bijective correspondence between antigradient flow lines of f_0 and solutions of equations (34). This correspondence is also valid for $\lambda = 0$, but in this case τ_0 maps $\mathbb{R}$ bijectively onto the interval $(-T, T)$. If we extend γ_0 by the constant values outside $(-T, T)$, then γ_λ converges to γ_0 in $C^0(\mathbb{R}; M)$ as $\lambda \to 0$ (in fact, in any reasonable topology).

With this understood, consider the family of equations

$$\begin{aligned} &\partial_s u_\lambda + J\Big(\partial_t u_\lambda + \frac{1}{\lambda + (1-\lambda)\|v_0\|^2}\, v_0\Big) = 0, \qquad (s, t) \in \mathbb{R}^2 \\ &\lim_{t\to\pm\infty} u_\lambda(s, t) = m_{\mp}, \qquad \lim_{s\to\pm\infty} u_\lambda(s, t) = \gamma_{\lambda,\pm}(t). \end{aligned} \tag{35}$$

For these equations explicit correspondence between solutions for different values of λ is not available anymore, but it is reasonable to expect that u_λ converges to a solution of (32) as $\lambda \to 0$.

Step 2. Let $L_\pm(\tau) \subset f^{-1}(\tau),\ \tau \in (-T, T)$, denote the vanishing cycle of $m_\pm$ associated with the segment $[\tau, \pm T]$. Consider the family of equations

$$\begin{aligned} &\partial_s u_\mu + J\Big(\partial_t u_\mu + \frac{1-\mu}{\|v_0\|^2}\, v_0\Big) = 0, \qquad (s, \tau) \in \mathbb{R} \times (-T, T) \\ &u_\mu(s, \pm T) \in L_\pm(\pm(1-\mu)T), \qquad \lim_{s\to\pm\infty} u_\mu(s, \tau) = \gamma_{0,\pm}((1-\mu)\tau) \end{aligned}$$

with $\mu \in [0, 1]$. Clearly, for $\mu = 0$ we obtain equations (32), whereas for $\mu = 1$ we have holomorphic strips as in the classical definition of the Floer differential. Notice that the images of such holomorphic strips lie in the fibre of f.

Remark A.1. Pick a solution u_0 of equations (32) and denote $f \circ u_0 = \varphi + i\psi$. It follows from the holomorphicity of f that φ and ψ satisfy the inhomogeneous Cauchy-Riemann equations

$$\partial_s \varphi - \partial_\tau \psi = 0, \quad \partial_s \psi + \partial_\tau \varphi + 1 = 0$$

and therefore both functions are harmonic. Moreover, the holomorphicity of f also implies that $f_1 \circ \gamma_\pm(\tau)$ is constant it τ and therefore vanishes everywhere, since $f_1 \circ \gamma_\pm(\pm T) = f_1(m_\pm) = 0$. We conclude that ψ vanishes as $\tau = \pm T$ and as $s \to \pm\infty$ and thus vanishes everywhere. Therefore $\varphi(s, \tau) = -\tau$. We see that unlike for pseudoholomorphic strips, images of solutions to equations (32) do not lie in a fixed fibre of f, but rather the fibre of $u_0(s, \tau)$ varies in a controlled (and very explicit) manner for any u_0.

Notice also that at the first glance equation (32) has singularities. Namely, if a solution u_0 hits a critical point of f at a single point (s_0, τ_0), then φ and ψ are harmonic in $\mathbb{R} \times (-T, T) \setminus \{(s_0, \tau_0)\}$ and continuous at (s_0, τ_0). Hence the singularity is removable and the above argument shows that the image of $f \circ u_0$ is the segment $(-T, T)$. Since by assumption the segment $(-T, T)$ does not contain any critical values, we conclude that a priori a solution of (32) cannot hit a critical point of f in an interior point.

B Transversality for broken flow lines

This appendix is devoted to the proof of Lemma 2.5.

Denote by $\rho \colon M \to \mathbb{R}$ the non-negative function $m \mapsto |v_0(m)|^2 = |v_1(m)|^2$. Replacing f by $e^{-i\theta_+}(f - z_0)$ if necessary we may assume that $z_0 = 0$, $z_+ = r \in \mathbb{R}_{>0}$ and $\theta_- \neq 0, \pi$. For the sake of definiteness we assume that $\theta_- \in (0, \pi)$. Observe that under these assumptions $\theta_\lambda(t) = -(\pi - \theta_-)\beta(\lambda^{-1}t) \in [-\pi + \theta_-, 0]$ for all t.

Recall the equations for broken flow lines

$$\dot\gamma_\lambda + \cos\theta_\lambda(t)\, v_0 + \sin\theta_\lambda(t)\, v_1 = 0, \qquad \lim_{t \to \pm\infty} \gamma_\lambda(t) = m_\mp. \tag{36}$$

Lemma B.1. *Denote $R = \{z \in \mathbb{C} \mid 0 \le \operatorname{Im} z \le \operatorname{Im} z_-,\ -r\sin\theta_- \le \operatorname{Im} e^{-i\theta_-} z \le 0\}$, i.e. R is the parallelogram spanned by the vectors $\overrightarrow{z_0 z_+}$ and $\overrightarrow{z_0 z_-}$. Then the following holds:*

(i) The image of any solution γ_λ to equations (36) is contained in $f^{-1}(R)$;

(ii) There exist positive constants C_1, C_2, and C_3 such that

$$\left|\operatorname{Im} f \circ \gamma_\lambda(t)\right| \le C_1 \lambda \qquad \text{for all } t \le 0, \tag{37}$$

$$\left|\operatorname{Im} e^{-i\theta_-} f \circ \gamma_\lambda(t)\right| \le C_2 \lambda \qquad \text{for all } t \ge 0, \tag{38}$$

$$\left|\operatorname{Re} f \circ \gamma_\lambda(\pm\lambda)\right| \le C_3 \lambda \tag{39}$$

holds for all λ and all solutions γ_λ of equations (36).

Proof. Recall that with our conventions $\theta_\lambda(t)$ takes values in $[-\pi + \theta_-, 0]$. Then it follows from equations (36) that

$$\frac{d}{dt} \operatorname{Im} f \circ \gamma_\lambda = -|v_1|^2 \sin\theta_\lambda(t) \ge 0.$$

From the boundary conditions at $\pm\infty$ we get that the inequality $0 \le \operatorname{Im} f \circ \gamma_\lambda(t) \le \operatorname{Im} z_-$ holds for all t. The inequality $-r \sin\theta_- \le \operatorname{Im}\big(e^{-i\theta_-} f \circ \gamma_\lambda(t)\big) \le 0$ can be established by applying similar arguments. This proves assertion *(i)*.

Further, notice that $f^{-1}(R)$ is compact since f is proper by assumption. Hence, there exists $C_2 > 0$ such that $\rho \le C_2$ on $f^{-1}(R)$. Observe also that equations (36) imply that $\operatorname{Im} f \circ \gamma_\lambda(t) = 0$ for all $t \le -\lambda$. Then for $t \in (-\lambda, 0]$ we obtain

$$\big|\operatorname{Im} f \circ \gamma_\lambda(t)\big| = \Big|\int_{-\lambda}^{t} \frac{d}{dt} \operatorname{Im} f \circ \gamma_\lambda(\tau)\, d\tau\Big| = \Big|\int_{-\lambda}^{t} |v_1|^2 \sin\theta_\lambda(\tau)\, d\tau\Big| \le C_2\lambda,$$

which establishes (37). Inequality (38) is proved similarly.

As we have already observed above the point $f \circ \gamma_\lambda(-\lambda)$ lies on the real axis. Similarly, the point $f \circ \gamma_\lambda(+\lambda)$ lies on the line ℓ, which contains z_- and the origin. Then

$$\begin{aligned}\big|\operatorname{Re} f \circ \gamma_\lambda(-\lambda)\big| &= \sin\theta_-\ \operatorname{dist}\big(f \circ \gamma_\lambda(-\lambda),\, \ell\big)\\ &\le \sin\theta_- \int_{-\lambda}^{\lambda} \Big|\frac{d}{dt} f \circ \gamma_\lambda(t)\Big|\, dt \le \sin\theta_-\, C_2\lambda = C_3\lambda.\end{aligned}$$

This finishes the proof. □

Lemma B.2. *Let $B_\varepsilon(m_\pm)$ be the connected components of $\{\rho < \varepsilon\}$ containing $m_\pm$. Then there exist $\varepsilon_0 \in (0,1)$ and $\lambda_0' \in (0,1)$ such that the following holds: For each $\varepsilon \in (0, \varepsilon_0)$ there exists $T_\varepsilon > 0$ such that for any $\lambda \in (0, \lambda_0')$ and any solution γ_λ of equations* (36) *we have*

$$\gamma_\lambda(-\infty, -T_\varepsilon) \subset B_\varepsilon(m_+) \qquad \textit{and} \qquad \gamma_\lambda(T_\varepsilon, +\infty) \subset B_\varepsilon(m_-).$$

Proof. We will restrict ourselves to establishing the inclusion $\gamma_\lambda(-\infty, -T_\varepsilon) \subset B_\varepsilon(m_+)$. The other inclusion can be obtained along similar lines.

Denote by $\mathcal{U}_+$ the unstable manifold of m_+ regarded as a critical point of $f_0 = \operatorname{Re} f$. Choose ε_0 so small that the restriction of the Hessian $\operatorname{Hess}(f_0)$ to the tangent bundle of $\mathcal{U}_+ \cap B_{\varepsilon_0}(m_+)$ is negative-definite. Redenoting ε_0 if necessary, we can also assume that $\inf\{f_0(m) : m \in B_{\varepsilon_0}(m_\pm)\} \ge 2r/3$.

Further, choose λ_0' so small that the following conditions hold:

- the set $R_0 = \{z \in R \mid 0 \le \operatorname{Im} z \le \lambda_0' \operatorname{Im} z_- \text{ or } -r\lambda_0' \sin\theta_- \le \operatorname{Im} e^{-i\theta_-} z \le 0\}$ contains no critical values of f other than $z_+ = r$ and z_-;
- $C_3\lambda_0' < r/3$, where C_3 is as in inequality (39).

Pick any $(\varepsilon, \lambda) \in (0, \varepsilon_0) \times (0, \lambda_0')$ and any solution γ_λ of equations (36). Notice first that the above choices guarantee that $\gamma_\lambda(-\lambda) \notin B_\varepsilon(m_+)$. Further, with these choices we also have the following property: If $\gamma_\lambda(t_0) \in B_\varepsilon(m_+)$ for some $t_0 < -\lambda_0'$, then $\gamma_\lambda(t) \in B_\varepsilon(m_+)$ for all $t \le t_0$. Indeed, first observe that $\gamma_\lambda(t) \in \mathcal{U}_+$ for all $t \le -\lambda_0'$. Since $d\rho(\cdot) = 2\langle \operatorname{Hess}(f_0)(\cdot), v_0\rangle$ we obtain

$$\frac{d}{dt}\rho \circ \gamma_\lambda = -2\big\langle \operatorname{Hess}(f_0)(\dot\gamma_\lambda), \dot\gamma_\lambda\big\rangle \qquad \text{for all } t \le -\lambda_0'.$$

Hence, if there exists $t_0 \le -\lambda_0'$ such that $\gamma_\lambda(t_0) \in B_\varepsilon(m_+)$, then the function $\rho \circ \gamma_\lambda$ must be increasing for all $t \le t_0$ and therefore $\rho \circ \gamma_\lambda(t) \le \rho \circ \gamma_\lambda(t_0) < \varepsilon$ for all $t \le t_0$.

Put $\varkappa_\varepsilon = \inf\{\rho(m) \mid m \in f^{-1}(R_0) \setminus B_\varepsilon(m_\pm)\} > 0,\ T_\varepsilon = \max\{2r/\varkappa_\varepsilon, 1\}$ and assume $\gamma_\lambda(-T_\varepsilon) \notin B_\varepsilon(m_+)$. Then $\gamma_\lambda(t) \notin B_\varepsilon(m_+)$ for all $t \in [-T_\varepsilon, -\lambda]$ and therefore we obtain

$$\frac{4r}{3} \geq f_0 \circ \gamma_\lambda(-T_\varepsilon) - f_0 \circ \gamma_\lambda(-\lambda) = -\int_{-T_\varepsilon}^{-\lambda} \frac{d}{dt} f_0 \circ \gamma_\lambda(t)\, dt = \int_{-T_\varepsilon}^{-\lambda} |v_0|^2\, dt \geq \varkappa_\varepsilon T_\varepsilon > 2r,$$

which is a contradiction. Thus $\gamma_\lambda(-T_\varepsilon) \in B_\varepsilon(m_+)$ which implies $\gamma_\lambda(-\infty, -T_\varepsilon) \subset B_\varepsilon(m_+)$. $\square$

Proposition B.3. *Let $\lambda_j \to 0,\ \lambda_j \in (0,1)$ be an arbitrary sequence and γ_j be an arbitrary sequence of solutions to equations (36) for $\lambda = \lambda_j$. Then there exists a subsequence γ_{j_k}, which converges in $C^0(\mathbb{R}; M)$ and $\gamma_0 = \lim\limits_{k\to\infty} \gamma_{j_k}$ is a solution of* (1).

Proof. From Lemma B.1 *(i)* we obtain that there exists a constant $C > 0$ such that $|\dot\gamma_j| \leq C$. Then with the help of the Ascoli-Arzela theorem we can find a subsequence γ_{j_k}, which converges to some $\gamma_0 \in C^0(\mathbb{R}; M)$ on each compact interval. Then $\gamma_0 \in C^1(\mathbb{R} \setminus \{0\}; M)$ and satisfies $\dot\gamma_0 + \sin\theta_0\, v_0 + \cos\theta_0\, v_1 = 0$. From Lemma B.1 *(ii)* we obtain that the image of $f \circ \gamma_0$ is contained in $\mathbb{R} \cup \ell$, where ℓ is the straight line containing z_- and the origin. For all $t < 0$ we also have $f_0 \circ \gamma_0(t) = \lim_{k\to\infty} f_0 \circ \gamma_{j_k}(t) \leq r$. Recalling that f is proper we see that the limit $\lim_{t\to-\infty} \gamma_0(t)$ must be a critical point of f_0. Hence, $\lim_{t\to-\infty} \gamma_0(t) = m_+$. Arguing along similar lines, we also obtain $\lim_{t\to+\infty} \gamma_0(t) = m_-$, i.e. $\gamma_0 \in \mathcal{M}_0(m_-; m_+)$.

Fix any $\varepsilon \in (0, \varepsilon_0)$. From Lemma B.2 we obtain that there exist $T'_\varepsilon > 0$ and $N_\varepsilon > 0$ such that $d(\gamma_{j_k}(t), \gamma_0(t)) < \varepsilon$ provided $|t| \geq T'_\varepsilon$ and $k \geq N_\varepsilon$. Since γ_{j_k} is convergent in $C^0(-T'_\varepsilon, T'_\varepsilon)$, we can find $N'_\varepsilon \geq N_\varepsilon$ such that $d(\gamma_{j_k}(t), \gamma_0(t)) < \varepsilon$ for all $t \in [-T'_\varepsilon, T'_\varepsilon]$ provided $k \geq N'_\varepsilon$, i.e. γ_{j_k} converges to γ_0 in $C^0(\mathbb{R}; M)$. $\square$

Proof of Lemma 2.5. Let $\mathcal{U}_+$ denote the unstable manifold of m_+ regarded as a critical point of f_0. Similarly, let $\mathcal{S}_-$ denote the stable manifold of m_- regarded as a critical point of $\mathrm{Re}\,(e^{-i\theta_-} f)$.

Pick a point $m \in L_- \cap L_+ \cong \mathcal{S}_- \cap \mathcal{U}_+$ and observe that $\mathcal{S}_-$ and $\mathcal{U}_+$ are the Lagrangian thimbles of m_- and m_+ associated with the segments $\overline{z_0 z_-}$ and $\overline{z_0 z_+}$ respectively, where z_0 is the origin. Then the hypothesis of the lemma implies that $\mathcal{S}_-$ and $\mathcal{U}_+$ intersect transversally at m.

Let γ_0 be the solution of (1) corresponding to m. Denote by $D_{\gamma_0}\Phi_0$ the linearization of Φ_0 at γ_0. As we have already remarked above, $D_{\gamma_0}\Phi_0 \colon W^{1,2}(\gamma_0^* TM) \to L^2(\gamma_0^* TM)$ is a Fredholm operator of index 0. Moreover, it can be shown in the similar manner as in the proof of Theorem 3.3 in [Sal1] that $\dim\operatorname{coker} D_{\gamma_0}\Phi_0 = \operatorname{codim}(T_m\mathcal{S}_- + T_m\mathcal{U}_+)$. Therefore $\dim\operatorname{coker} D_{\gamma_0}\Phi_0 = 0$ and hence $\dim\ker D_{\gamma_0}\Phi_0 = 0$. Thus we conclude that Φ_0 intersects the zero-section transversally.

It follows from Proposition B.3 that there exists $\lambda_0 > 0$ such that each solution of the equation $\Phi_\lambda(\gamma_\lambda) = 0,\ \lambda \in (0, \lambda_0)$ is contained in a C^0–neighbourhood $U(\gamma_0)$ of some $\gamma_0 \in \Phi_0^{-1}(0)$. Notice that the linearization of Φ_λ at γ_λ can be written in the form

$$\begin{aligned} D_{\gamma_\lambda}\Phi_\lambda(\xi) &= \nabla_{\dot\gamma_\lambda}\xi + \cos\theta_\lambda \nabla_\xi v_0 + \sin\theta_\lambda \nabla_\xi v_1 \\ &= \cos\theta_\lambda\big(\nabla_\xi v_0 - \nabla_{v_0}\xi\big) + \sin\theta_\lambda\big(\nabla_\xi v_1 - \nabla_{v_1}\xi\big). \end{aligned}$$

where $\xi \in W^{1,2}(\gamma^* TM)$. Hence, changing λ_0 if necessary, we can assume that the linearization of Φ_λ is non-degenerate at each $\gamma_\lambda \in \mathcal{M}_\lambda(m_-; m_+)$ contained in $\bigcup_{\gamma_0} U(\gamma_0)$

for $\lambda \in (0, \lambda_0)$, since $\#\Phi_0^{-1}(0) = \#L_- \cap L_+ < \infty$. Thus Φ_λ intersects the zero-section transversally provided $\lambda \leq \lambda_0$.

Consider Φ_λ as a section of $\pi^*\mathcal{E}$, where $\pi\colon W^{1,2}_{m_-;m_+} \times \mathbb{R}_\lambda \to \mathbb{R}_\lambda$ is the canonical projection. Then Φ_λ is continuous and satisfies the hypothesis of the implicit function theorem. Therefore, $\{(\gamma, \lambda) \mid \Phi_\lambda(\gamma) = 0,\ \lambda \in [0, \lambda_0)\}$ is homeomorphic to $\Phi_0^{-1}(0) \times [0, \lambda_0)$. This establishes the bijective correspondence between $\mathcal{M}_0(m_-; m_+)$ and $\mathcal{M}_\lambda(m_-; m_+)$. □

References

[APS] M. F. Atiyah, V. K. Patodi, and I. M. Singer. Spectral asymmetry and Riemannian geometry. III. *Math. Proc. Cambridge Philos. Soc.*, 79(1):71–99, 1976.

[BH] A. Banyaga and D. E. Hurtubise. The Morse-Bott inequalities via a dynamical systems approach. *Ergodic Theory Dynam. Systems*, 29(6):1693–1703, 2009.

[BS] R. L. Bryant and S. M. Salamon. On the construction of some complete metrics with exceptional holonomy. *Duke Math. J.*, 58(3):829–850, 1989.

[CGS] K. Cieliebak, A. R. Gaio, and D. A. Salamon. *J*-holomorphic curves, moment maps, and invariants of Hamiltonian group actions. *Internat. Math. Res. Notices*, (16):831–882, 2000.

[Cor] K. Corlette. Flat *G*-bundles with canonical metrics. *J. Differential Geom.*, 28(3):361–382, 1988.

[Don1] S. K. Donaldson. Twisted harmonic maps and the self-duality equations. *Proc. London Math. Soc. (3)*, 55(1):127–131, 1987.

[Don2] S. K. Donaldson. *Floer homology groups in Yang-Mills theory*, volume 147 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 2002. With the assistance of M. Furuta and D. Kotschick.

[DS] S. Donaldson and E. Segal. Gauge theory in higher dimensions, II. *arXiv:0902.3239*.

[Fan] H. Fan. Half de Rham complexes and line fields on odd-dimensional manifolds. *Trans. Amer. Math. Soc.*, 348(8):2947–2982, 1996.

[FJR] H. Fan, T. J. Jarvis, and Y. Ruan. The Witten equation, mirror symmetry and quantum singularity theory. *arXiv:0712.4021v3*.

[Fre] D. S. Freed. Remarks on Chern-Simons theory. *Bull. Amer. Math. Soc. (N.S.)*, 46(2):221–254, 2009.

[GS] A. R. Gaio and D. A. Salamon. Gromov-Witten invariants of symplectic quotients and adiabatic limits. *J. Symplectic Geom.*, 3(1):55–159, 2005.

[Hay] A. Haydys. Gauge theory, calibrated geometry and harmonic spinors. *arXiv:0902.3738*.

[Hit] N. J. Hitchin. The self-duality equations on a Riemann surface. *Proc. Lond. Math. Soc., III. Ser.*, 55:59–126, 1987.

[Kor] Yu. A. Kordyukov. L^p-theory of elliptic differential operators on manifolds of bounded geometry. *Acta Appl. Math.*, 23(3):223–260, 1991.

[KW] A. Kapustin and E. Witten. Electric-magnetic duality and the geometric Langlands program. *Commun. Number Theory Phys.*, 1(1):1–236, 2007.

[RS] J. Robbin and D. Salamon. The Maslov index for paths. *Topology*, 32(4):827–844, 1993.

[Sal1] D. Salamon. Morse theory, the Conley index and Floer homology. *Bull. London Math. Soc.*, 22(2):113–140, 1990.

[Sal2] D. Salamon. Lectures on Floer homology. In *Symplectic geometry and topology (Park City, UT, 1997)*, volume 7 of *IAS/Park City Math. Ser.*, pages 143–229. Amer. Math. Soc., Providence, RI, 1999.

[Sei1] P. Seidel. Vanishing cycles and mutation. In *European Congress of Mathematics, Vol. II (Barcelona, 2000)*, volume 202 of *Progr. Math.*, pages 65–85. Birkhäuser, Basel, 2001.

[Sei2] P. Seidel. *Fukaya categories and Picard-Lefschetz theory.* Zurich Lectures in Advanced Mathematics. European Mathematical Society (EMS), Zürich, 2008.

[Shu] M. A. Shubin. Spectral theory of elliptic operators on noncompact manifolds. *Astérisque*, (207):5, 35–108, 1992. Méthodes semi-classiques, Vol. 1 (Nantes, 1991).

[SZ] D. Salamon and E. Zehnder. Morse theory for periodic solutions of Hamiltonian systems and the Maslov index. *Comm. Pure Appl. Math.*, 45(10):1303–1360, 1992.

[VW] C. Vafa and E. Witten. A strong coupling test of S-duality. *Nuclear Phys. B*, 431(1-2):3–77, 1994.

[Wit1] E. Witten. Analytic Continuation Of Chern-Simons Theory. *arXiv:1001.2933*.

[Wit2] E. Witten. Fivebranes and knots. *arXiv:1101.3216*.

[Wit3] E. Witten. A new look at the path integral of quantum mechanics. *arXiv:1009.6032v1*.