Floer homology and the heat flow

Dietmar A. Salamon Joa Weber *

ETH-Zürich

April 2003, revised October 2005

Abstract

We study the heat flow in the loop space of a closed Riemannian manifold M as an adiabatic limit of the Floer equations in the cotangent bundle. Our main application is a proof that the Floer homology of the cotangent bundle, for the Hamiltonian function kinetic plus potential energy, is naturally isomorphic to the homology of the loop space.

1 Introduction

Let M be a closed Riemannian manifold and denote by $\mathcal{L}M$ the free loop space. Consider the classical action functional

$$\mathcal{S}_V(x) = \int_0^1 \left(\frac{1}{2}|\dot{x}(t)|^2 - V(t, x(t)) \right) dt$$

for $x : S^1 \to M$. Here and throughout we identify $S^1 = \mathbb{R}/\mathbb{Z}$ and think of $x \in C^\infty(S^1, M)$ as a smooth map $x : \mathbb{R} \to M$ which satisfies $x(t+1) = x(t)$. The potential is a smooth function $V : S^1 \times M \to \mathbb{R}$ and we write $V_t(x) := V(t, x)$. The critical points of $\mathcal{S}_V$ are the 1-periodic solutions of the ODE

$$\nabla_t \dot{x} = -\nabla V_t(x), \tag{1}$$

where ∇V_t denotes the gradient and $\nabla_t \dot{x}$ denotes the Levi-Civita connection. Let $\mathcal{P} = \mathcal{P}(V)$ denote the set of 1-periodic solutions $x : S^1 \to M$ of (1). In the case $V = 0$ these are the closed geodesics. Via the Legendre transformation the solutions of (1) can also be interpreted as the critical points of the symplectic action $\mathcal{A}_V : \mathcal{L}T^*M \to \mathbb{R}$ given by

$$\mathcal{A}_V(z) = \int_0^1 \Big(\langle y(t), \dot{x}(t) \rangle - H(t, x(t), y(t)) \Big) dt$$

*Partial financial support by TH-Projekt 00321.

where $z = (x, y) : S^1 \to T^*M$ and the Hamiltonian $H = H_V : S^1 \times T^*M \to \mathbb{R}$ is given by

$$H(t, x, y) = \frac{1}{2}|y|^2 + V(t, x) \tag{2}$$

for $y \in T_x^*M$. A loop $z(t) = (x(t), y(t))$ in T^*M is a critical point of $\mathcal{A}_V$ iff x is a solution of (1) and $y(t) \in T_{x(t)}^*M$ is related to $\dot{x}(t) \in T_{x(t)}M$ via the isomorphism $TM \to T^*M$ induced by the Riemannian metric. For such loops z the symplectic action $\mathcal{A}_V(z)$ agrees with the classical action $\mathcal{S}_V(x)$.

The negative L^2 gradient flow of the classical action gives rise to a Morse-Witten complex which computes the homology of the loop space. For a regular value a of $\mathcal{S}_V$ we shall denote by $\mathrm{HM}_*^a(\mathcal{L}M, \mathcal{S}_V)$ the homology of the Morse-Witten complex of the functional $\mathcal{S}_V$ corresponding to the solutions of (1) with $\mathcal{S}_V(x) \le a$. Here we assume that $\mathcal{S}_V$ is a Morse function and its gradient flow satisfies the Morse-Smale condition (i.e. the stable and unstable manifolds intersect transversally, see [3] for the unstable manifold). As in the finite dimensional case one can show that the Morse-Witten homology $\mathrm{HM}_*^a(\mathcal{L}M, \mathcal{S}_V)$ is naturally isomorphic to the singular homology of the sublevel set

$$\mathcal{L}^a M = \{x \in \mathcal{L}M \mid \mathcal{S}_V(x) \le a\}.$$

On the other hand one can use the L^2 gradient flow of $\mathcal{A}_V$ to construct Floer homology groups $\mathrm{HF}_*^a(T^*M, H_V)$. Our main result is the following.

Theorem 1.1. *Assume $\mathcal{S}_V$ is Morse and a is either a regular value of $\mathcal{S}_V$ or is equal to infinity. Then there is a natural isomorphism*

$$\mathrm{HF}_*^a(T^*M, H_V; R) \cong \mathrm{HM}_*^a(\mathcal{L}M, \mathcal{S}_V; R)$$

for every principal ideal domain R. If M is not simply connected then there is a separate isomorphism for each component of the loop space. The isomorphism commutes with the homomorphisms $\mathrm{HF}_*^a(T^*M, H_V) \to \mathrm{HF}_*^b(T^*M, H_V)$ and $\mathrm{HM}_*^a(\mathcal{L}M, \mathcal{S}_V) \to \mathrm{HM}_*^b(\mathcal{L}M, \mathcal{S}_V)$ for $a < b$.

Corollary 1.2. *Let $\mathcal{S}_V$ and a be as in Theorem 1.1. Then there is a natural isomorphism*

$$\mathrm{HF}_*^a(T^*M, H_V; R) \cong \mathrm{H}_*(\mathcal{L}^a M; R)$$

for every principal ideal domain R. If M is not simply connected then there is a separate isomorphism for each component of the loop space. The isomorphism commutes with the homomorphisms $\mathrm{HF}_*^a(T^*M, H_V) \to \mathrm{HF}_*^b(T^*M, H_V)$ and $\mathrm{H}_*(\mathcal{L}^a M) \to \mathrm{H}_*(\mathcal{L}^b M)$ for $a < b$.

Proof. The definition of the Morse homology groups involve a perturbation $\mathcal{V} : \mathcal{L}M \to \mathbb{R}$ (of the function $x \mapsto \int_0^1 V_t(x(t))\, dt$) that satisfies the hypotheses $(V0{-}V4)$ of Section 2 and the transversality requirements of Theorem A.6. Now Theorem A.7 in Appendix A (proved in the forthcoming paper [31]) asserts that the Morse homology group $\mathrm{HM}_*^a(\mathcal{L}M, \mathcal{S}_V; R) \cong \mathrm{HM}_*^a(\mathcal{L}M, \mathcal{S}_{\mathcal{V}}; R)$ is naturally isomorphic to the singular homology group $\mathrm{H}_*(\mathcal{L}^a M; R)$. Hence the assertion follows from Theorem 1.1. □

Both the Morse-Witten homology $HM_*^a(\mathcal{L}M, \mathcal{S}_V)$ and the Floer homology $\mathrm{HF}_*^a(T^*M, H_V)$ are based on the same chain complex C_*^a which is generated by the solutions of (1) and graded by the Morse index (as critical points of $\mathcal{S}_V$). In [28] it is shown that this Morse index agrees, up to a universal additive constant zero or one, with minus the Conley-Zehnder index. Thus it remains to compare the boundary operators and this will be done by considering an adiabatic limit with a family of metrics on T^*M which scales the vertical part down to zero. Another approach to Corollary 1.2 is contained in Viterbo's paper [25]. While the present paper was being completed a new proof of Corollary 1.2 was given by Abbondandolo and Schwarz [1]. Some recent applications of Corollary 1.2 can be found in [30]; these applications require the statement with action windows and fixed homotopy classes of loops.

The Floer chain complex and its adiabatic limit

We assume throughout that $\mathcal{S}_V$ is a Morse function on the loop space, i.e. that the 1-periodic solutions of (1) are all nondegenerate. (For a proof that this holds for a generic potential V see [28].) Under this assumption the set

$$\mathcal{P}^a(V) := \{x \in \mathcal{P}(V) \,|\, \mathcal{S}_V(x) \leq a\}$$

is finite for every real number a. Moreover, each critical point $x \in \mathcal{P}(V)$ has well defined stable and unstable manifolds with respect to the (negative) L^2 gradient flow (see for example Davies [3]). Call $\mathcal{S}_V$ **Morse–Smale** if it is a Morse function and the unstable manifold $W^u(y)$ intersects the stable manifold $W^s(x)$ transversally for any two critical points $x, y \in \mathcal{P}(V)$.

Assume $\mathcal{S}_V$ is a Morse function and consider the $\mathbb{Z}$-module

$$C^a = C^a(V) = \bigoplus_{x \in \mathcal{P}^a(V)} \mathbb{Z}x.$$

If $\mathcal{S}_V$ and $\mathcal{A}_V$ are Morse–Smale then this module carries two boundary operators. The first is defined by counting the (negative) gradient flow lines of $\mathcal{S}_V$. They are solutions $u : \mathbb{R} \times S^1 \to M$ of the **heat equation**

$$\partial_s u - \nabla_t \partial_t u - \nabla V_t(u) = 0 \tag{3}$$

satisfying

$$\lim_{s \to \pm\infty} u(s,t) = x^{\pm}(t), \qquad \lim_{s \to \pm\infty} \partial_s u = 0, \tag{4}$$

where $x^{\pm} \in \mathcal{P}(V)$. The limits are uniform in t. The space of solutions of (3) and (4) will be denoted by $\mathcal{M}^0(x^-, x^+; V)$. The Morse–Smale hypothesis guarantees that, for every pair $x^{\pm} \in \mathcal{P}^a(V)$, the space $\mathcal{M}^0(x^-, x^+; V)$ is a smooth manifold whose dimension is equal to the difference of the Morse indices. In the case of Morse index difference one it follows that the quotient $\mathcal{M}^0(x^-, x^+; V)/\mathbb{R}$ by the (free) time shift action is a finite set. Counting the number of solutions with appropriate signs gives rise to a boundary operator on $C^a(V)$. The homology $\mathrm{HM}^a_*(\mathcal{L}M,\mathcal{S}_V)$ of the resulting chain complex is naturally isomorphic to the singular homology of the loop space for every regular value a of $\mathcal{S}_V$:

$$\mathrm{HM}^a_*(\mathcal{L}M,\mathcal{S}_V)\cong \mathrm{H}_*(\mathcal{L}^aM;\mathbb{Z}),\qquad \mathcal{L}^aM:=\{x\in\mathcal{L}M\,|\,\mathcal{S}_V(x)\le a\}\,.$$

The details of this isomorphism will be established in a separate paper (see Appendix A for a summary of the relevant results).

The second boundary operator is defined by counting the negative gradient flow lines of the symplectic action functional $\mathcal{A}_V$. These are the solutions $(u,v):\mathbb{R}\times S^1\to TM$ of the **Floer equations**

$$\partial_s u-\nabla_t v-\nabla V_t(u)=0,\qquad \nabla_s v+\partial_t u-v=0,\tag{5}$$

$$\lim_{s\to\pm\infty}u(s,t)=x^\pm(t),\qquad \lim_{s\to\pm\infty}v(s,t)=\dot{x}^\pm(t).\tag{6}$$

Here we also assume that $\partial_s u$ and $\nabla_s v$ converge to zero, uniformly in t, as $|s|$ tends to infinity. For notational simplicity we identify the tangent and cotangent bundles of M via the metric. Counting the index-1 solutions of (5) and (6) with appropriate signs we obtain the Floer boundary operator. We wish to prove that the resulting Floer homology groups $\mathrm{HF}^a_*(T^*M,H_V)$ are naturally isomorphic to $\mathrm{HM}^a_*(\mathcal{L}M,\mathcal{S}_V)$. To construct this isomorphism we modify equation (5) by introducing a small parameter ε as follows

$$\partial_s u-\nabla_t v-\nabla V(t,u)=0,\qquad \nabla_s v+\varepsilon^{-2}(\partial_t u-v)=0.\tag{7}$$

The space of solutions of (7) and (6) will be denoted by $\mathcal{M}^\varepsilon(x^-,x^+;V)$. The Floer homology groups for different values of ε are isomorphic (see Remark 1.3 below). Thus the task at hand is to prove that, for $\varepsilon>0$ sufficiently small, there is a one-to-one correspondence between the solutions of (3) and those of (7). A first indication, why one might expect such a correspondence, is the energy identity

$$\begin{aligned}E^\varepsilon(u,v)&=\frac{1}{2}\int_{-\infty}^{\infty}\int_0^1\left(|\partial_s u|^2+|\nabla_t v+\nabla V_t(u)|^2+\varepsilon^2|\nabla_s v|^2+\varepsilon^{-2}|\partial_t u-v|^2\right)\\&=\mathcal{S}_V(x^-)-\mathcal{S}_V(x^+)\end{aligned}\tag{8}$$

for the solutions of (7) and (6). It shows that $\partial_t u-v$ must converge to zero in the L^2 norm as ε tends to zero. If $\partial_t u=v$ then the first equation in (7) is equivalent to (3).

Remark 1.3. Let M be a Riemannian manifold. Then the tangent space of the cotangent bundle T^*M at a point (x,y) with $y\in T^*_xM$ can be identified with the direct sum $T_xM\oplus T^*_xM$. The isomorphism takes the derivative $\dot{z}(t)$ of a curve $\mathbb{R}\to T^*M:t\mapsto z(t)=(x(t),y(t))$ to the pair $(\dot{x}(t),\nabla_t y(t))$. With this identification the almost complex structure J_ε and the metric G_ε on T^*M, given by

$$J_\varepsilon=\begin{pmatrix}0&-\varepsilon g^{-1}\\ \varepsilon^{-1}g&0\end{pmatrix},\qquad G_\varepsilon=\begin{pmatrix}\varepsilon^{-1}g&0\\0&\varepsilon g^{-1}\end{pmatrix},$$

are compatible with the standard symplectic form ω on T^*M. Here we denote by $g : TM \to T^*M$ the isomorphism induced by the metric. The case $\varepsilon = 1$ corresponds to the standard almost complex structure. The Floer equations for the almost complex structure J_ε and the Hamiltonian (2) are

$$\partial_s w - J_\varepsilon(w)(\partial_t w - X_{H_t}(w)) = 0.$$

If we write $w(s,t) = (u(s,t), v(s,t))$ with $v(s,t) \in T^*_{u(s,t)}M$ then this equation has the form

$$\partial_s u - \varepsilon g^{-1}\nabla_t v - \varepsilon \nabla V_t(u) = 0, \qquad \nabla_s v + \varepsilon^{-1} g \partial_t u - \varepsilon^{-1} v = 0. \tag{9}$$

A function $w = (u,v)$ is a solution of (9) if and only if the functions $\tilde{u}(s,t) := u(\varepsilon^{-1}s, t)$ and $\tilde{v}(s,t) := g^{-1}v(\varepsilon^{-1}s, t)$ satisfy (7). In view of this discussion it follows from the Floer homotopy argument that the Floer homology defined with the solutions of (7) is independent of the choice of $\varepsilon > 0$. The only nonstandard aspect of this argument is the apriori estimate of Section 5 with $\varepsilon = 1$ which carries over verbatim to the time dependent Floer equation. For the standard theory see [8, 22, 23].

Assume $\mathcal{S}_V$ is Morse–Smale. Then we shall prove that, for every $a \in \mathbb{R}$, there exists an $\varepsilon_0 > 0$ such that, for $0 < \varepsilon < \varepsilon_0$ and every pair $x^+, x^- \in \mathcal{P}^a(V)$ with Morse index difference one, there is a natural bijective correspondence between the (shift equivalence classes of) solutions of (3), (4) and those of (7), (6). This will follow from Theorems 4.1 and 10.1 below.

It is an open question if the function $\mathcal{S}_V$ is Morse–Smale (with respect to the L^2 metric on the loop space) for a generic potential V. However, it is easy to establish transversality for a general class of abstract perturbations $\mathcal{V} : \mathcal{L}M \to \mathbb{R}$ (see Section 2). We shall use these perturbations to prove Theorem 1.1 in general.

The general outline of the proof is similar to that of the Atiyah–Floer conjecture in [5] which compares two elliptic PDEs via an adiabatic limit argument. By contrast our adiabatic limit theorem compares elliptic with parabolic equations. This leads to new features in the analysis that are related to the fact that the parabolic equation requires different scaling in space and time directions.

The present paper is organized as follows. The next section introduces a relevant class of abstract perturbations $\mathcal{V} : \mathcal{L}M \to \mathbb{R}$. Section 3 explains the relevant linearized operators and states the estimates for the right inverse. These are proved in Appendices C and D. In Section 4 we construct a map $\mathcal{T}^\varepsilon : \mathcal{M}^0(x^-, x^+; \mathcal{V}) \to \mathcal{M}^\varepsilon(x^-, x^+; \mathcal{V})$ which assigns to every parabolic cylinder of index one a nearby Floer connecting orbit for $\varepsilon > 0$ sufficiently small. The existence of this map was established in the thesis of the second author [27], where the results of Section 3, Section 4, and Appendix D were proved. Sections 5, 6, and 7 are of preparatory nature and establish uniform estimates for the solutions of (7). Section 5 shows that the solutions of (7) (with fixed endpoints) are all contained in a fixed compact subset of T^*M that is independent of ε. The proof uses an inequality

$$\left(\varepsilon^2 \partial_s^2 + \partial_t^2 - \partial_s\right)|v|^2 \geq -c(|v|^2 + 1). \tag{10}$$

Integrating this inequality over the t-variable and using a bound on the action one first obtains an estimate for $\sup_s \int_0^1 |v(s,t)|^2\,dt$; using (10) again gives the required estimate for $\sup|v|$. Section 6 then gives estimates for the first, and Section 7 for the second derivatives. In each case the operator $\varepsilon^2\partial_s^2 + \partial_t^2 - \partial_s$ reappears and the axioms on the perturbation $\mathcal{V}$ in Section 2 require that the estimate is first established in an integrated form. Section 8 deals with exponential decay, Section 9 establishes local surjectivity of the map $\mathcal{T}^\varepsilon$ by a time-shift argument, and in Section 10 we prove that $\mathcal{T}^\varepsilon$ is bijective. Things are put together in Section 11 where we compare orientations and prove Theorem 1.1. Appendix A summarizes some results about the heat flow (3) which will be proved in [31]. In Appendix B we prove several mean value inequalities that play a central role in our apriori estimates of Sections 5, 6, and 7.

2 Perturbations

In this section we introduce a class of perturbations of equations (3) and (7) for which transversality is easy to achieve. The perturbations take the form of smooth maps $\mathcal{V} : \mathcal{L}M \to \mathbb{R}$. For $x \in \mathcal{L}M$ let $\operatorname{grad}\mathcal{V}(x) \in \Omega^0(S^1, x^*TM)$ denote the L^2-gradient of $\mathcal{V}$; it is defined by

$$\int_0^1 \langle \operatorname{grad}\mathcal{V}(u), \partial_s u\rangle\,dt := \frac{d}{ds}\mathcal{V}(u)$$

for every smooth path $\mathbb{R} \to \mathcal{L}M : s \mapsto u(s,\cdot)$. The **covariant Hessian** of $\mathcal{V}$ at a loop $x : S^1 \to M$ is the operator

$$\mathcal{H}_\mathcal{V}(x) : \Omega^0(S^1, x^*TM) \to \Omega^0(S^1, x^*TM)$$

defined by

$$\mathcal{H}_\mathcal{V}(u)\partial_s u := \nabla_s \operatorname{grad}\mathcal{V}(u)$$

for every smooth map $\mathbb{R} \to \mathcal{L}M : s \mapsto u(s,\cdot)$. The axiom $(V1)$ below asserts that this Hessian is a zeroth order operator. We impose the following conditions on $\mathcal{V}$; here $|\cdot|$ denotes the pointwise absolute value at $(s,t) \in \mathbb{R}\times S^1$ and $\|\cdot\|_{L^p}$ denotes the L^p-norm over S^1 at time s. Although condition $(V1)$, the first part of $(V2)$, and $(V3)$ are all special cases of $(V4)$ we state the axioms in the form below, because some of our results don't require all the conditions to hold.

(V0) $\mathcal{V}$ is continuous with respect to the C^0 topology on $\mathcal{L}M$. Moreover, there is a constant $C > 0$ such that

$$\sup_{x\in\mathcal{L}M} |\mathcal{V}(x)| + \sup_{x\in\mathcal{L}M} \|\operatorname{grad}\mathcal{V}(x)\|_{L^\infty(S^1)} \le C.$$

(V1) There is a constant $C > 0$ such that

$$|\nabla_s \operatorname{grad}\mathcal{V}(u)| \le C\big(|\partial_s u| + \|\partial_s u\|_{L^1}\big),$$
$$|\nabla_t \operatorname{grad}\mathcal{V}(u)| \le C\Big(1 + |\partial_t u|\Big)$$

for every smooth map $\mathbb{R} \to \mathcal{L}M : s \mapsto u(s,\cdot)$ and every $(s,t) \in \mathbb{R}\times S^1$.

(V2) There is a constant $C > 0$ such that

$$|\nabla_s \nabla_s \mathrm{grad}\, \mathcal{V}(u)| \le C\Big(|\nabla_s \partial_s u| + \|\nabla_s \partial_s u\|_{L^1} + \big(|\partial_s u| + \|\partial_s u\|_{L^2}\big)^2\Big),$$
$$|\nabla_t \nabla_s \mathrm{grad}\, \mathcal{V}(u)| \le C\Big(|\nabla_t \partial_s u| + \big(1 + |\partial_t u|\big)\big(|\partial_s u| + \|\partial_s u\|_{L^1}\big)\Big)$$

and

$$|\nabla_s \nabla_s \mathrm{grad}\, \mathcal{V}(u) - \mathcal{H}_{\mathcal{V}}(u)\nabla_s \partial_s u| \le C\big(|\partial_s u| + \|\partial_s u\|_{L^2}\big)^2$$

for every smooth map $\mathbb{R} \to \mathcal{L}M : s \mapsto u(s,\cdot)$ and every $(s,t) \in \mathbb{R} \times S^1$.

(V3) There is a constant $C > 0$ such that

$$\begin{aligned}
|\nabla_s \nabla_s \nabla_s \mathrm{grad}\, \mathcal{V}(u)| \le C\Big(&|\nabla_s \nabla_s \partial_s u| + \|\nabla_s \nabla_s \partial_s u\|_{L^1} \\
&+ \big(|\nabla_s \partial_s u| + \|\nabla_s \partial_s u\|_{L^2}\big)\big(|\partial_s u| + \|\partial_s u\|_{L^2}\big) \\
&+ \big(|\partial_s u| + \|\partial_s u\|_{L^\infty}\big)\big(|\partial_s u| + \|\partial_s u\|_{L^2}\big)^2\Big), \\
|\nabla_t \nabla_s \nabla_s \mathrm{grad}\, \mathcal{V}(u)| \le C\Big(&|\nabla_t \nabla_s \partial_s u| + |\nabla_t \partial_s u|\,\big(|\partial_s u| + \|\partial_s u\|_{L^1}\big) \\
&+ \big(1 + |\partial_t u|\big)\big(|\nabla_s \partial_s u| + \|\nabla_s \partial_s u\|_{L^1}\big) \\
&+ \big(1 + |\partial_t u|\big)\big(|\partial_s u| + \|\partial_s u\|_{L^2}\big)^2\Big), \\
|\nabla_t \nabla_t \nabla_s \mathrm{grad}\, \mathcal{V}(u)| \le C\Big(&|\nabla_t \nabla_t \partial_s u| + \big(1 + |\partial_t u|\big)\,|\nabla_t \partial_s u| \\
&+ \big(1 + |\partial_t u|^2 + |\nabla_t \partial_t u|\big)\big(|\partial_s u| + \|\partial_s u\|_{L^1}\big)\Big)
\end{aligned}$$

for every smooth map $\mathbb{R} \to \mathcal{L}M : s \mapsto u(s,\cdot)$ and every $(s,t) \in \mathbb{R} \times S^1$.

(V4) For any two integers $k > 0$ and $\ell \ge 0$ there is a constant $C = C(k,\ell)$ such that

$$\big|\nabla_t^\ell \nabla_s^k \mathrm{grad}\, \mathcal{V}(u)\big| \le C \sum_{k_j,\ell_j} \left(\prod_{\substack{j \\ \ell_j > 0}} \Big|\nabla_t^{\ell_j} \nabla_s^{k_j} u\Big| \right) \prod_{\substack{j \\ \ell_j = 0}} \left(\big|\nabla_s^{k_j} u\big| + \big\|\nabla_s^{k_j} u\big\|_{L^{p_j}} \right)$$

for every smooth map $\mathbb{R} \to \mathcal{L}M : s \mapsto u(s,\cdot)$ and every $(s,t) \in \mathbb{R}\times S^1$; here $p_j \ge 1$ and $\sum_{\ell_j=0} 1/p_j = 1$; the sum runs over all partitions $k_1 + \cdots + k_m = k$ and $\ell_1 + \cdots + \ell_m \le \ell$ such that $k_j + \ell_j \ge 1$ for all j. For $k = 0$ the same inequality holds with an additional summand C on the right.

Remark 2.1. The archetypal example of a perturbation is

$$\mathcal{V}(x) := \rho\left(\|x - x_0\|_{L^2}^2\right) \int_0^1 V_t(x(t))\, dt,$$

where $\rho : \mathbb{R} \to [0,1]$ is a smooth cutoff function, $x_0 : S^1 \to M$ is a smooth loop, and $x - x_0$ denotes the difference in some ambient Euclidean space into which M is (isometrically) embedded. Any such perturbation satisfies $(V0 - V4)$.

Remark 2.2. If

$$\mathcal{V}(x) = \int_0^1 V_t(x(t))\,dt$$

then

$$\operatorname{grad}\mathcal{V}(x) = \nabla V_t(x), \qquad \mathcal{H}_{\mathcal{V}}(x)\xi = \nabla_\xi \nabla V_t(x),$$

for $x \in \mathcal{L}M$ and $\xi \in \Omega^0(S^1, x^*TM)$.

With an abstract perturbation $\mathcal{V}$ the classical and symplectic action are given by

$$\mathcal{S}_{\mathcal{V}}(x) = \frac{1}{2}\int_0^1 |\dot{x}(t)|^2\,dt - \mathcal{V}(x)$$

and

$$\mathcal{A}_{\mathcal{V}}(x,y) = \int_0^1 \left(\langle y(t), \dot{x}(t)\rangle - \frac{1}{2}|y(t)|^2\right)dt - \mathcal{V}(x)$$

for $x \in \mathcal{L}M$ and $y \in \Omega^0(S^1, x^*T^*M)$. Equation (7) has the form

$$\partial_s u - \nabla_t v - \operatorname{grad}\mathcal{V}(u) = 0, \qquad \nabla_s v + \varepsilon^{-2}(\partial_t u - v) = 0, \tag{11}$$

and the limit equation is

$$\partial_s u - \nabla_t \partial_t u - \operatorname{grad}\mathcal{V}(u) = 0. \tag{12}$$

Here $\operatorname{grad}\mathcal{V}(u)$ denotes the value of $\operatorname{grad}\mathcal{V}$ on the loop $t \mapsto u(s,t)$. The relevant set of critical points consists of the loops $x : S^1 \to M$ that satisfy the differential equation $\nabla_t \dot{x} = -\operatorname{grad}\mathcal{V}(x)$ and will be denoted by $\mathcal{P}(\mathcal{V})$. The subset $\mathcal{P}^a(\mathcal{V}) \subset \mathcal{P}(\mathcal{V})$ consists of all critical points x with $\mathcal{S}_{\mathcal{V}}(x) \le a$.

Assume that the critical points of $\mathcal{S}_{\mathcal{V}}$ are all nondegenerate. Then every solution (u,v) of (11) with finite energy $E^\varepsilon(u,v) := \int_{\mathbb{R}\times S^1}\left(|\partial_s u|^2 + \varepsilon^2|\nabla_s v|^2\right) < \infty$ converges exponentially and in the C^∞-topology to critical points of $\mathcal{A}_{\mathcal{V}}$ as s tends to $\pm\infty$. Moreover, every sequence of solutions of (11) with a uniform bound on the first derivatives (and $\varepsilon > 0$ fixed) has a subsequence that converges in the C^∞-topology. These two assertions follow by an easy adaptation of the standard arguments (with Hamiltonian perturbations) to the present case (see for example [22]).

3 The linearized operators

Throughout this section we fix a perturbation $\mathcal{V}$ that satisfies $(V0 - V4)$. Linearizing the heat equation (12) gives rise to the operator

$$\mathcal{D}_u^0 : \Omega^0(\mathbb{R}\times S^1, u^*TM) \to \Omega^0(\mathbb{R}\times S^1, u^*TM)$$

given by

$$\mathcal{D}_u^0\xi = \nabla_s\xi - \nabla_t\nabla_t\xi - R(\xi, \partial_t u)\partial_t u - \mathcal{H}_{\mathcal{V}}(u)\xi, \tag{13}$$

for every element ξ of the set $\Omega^0(\mathbb{R}\times S^1, u^*TM)$ of smooth vector fields along u. Here R denotes the Riemann curvature tensor. If $\mathcal{S}_\mathcal{V}$ is Morse then this is a Fredholm operator between appropriate Sobolev completions. More precisely, define

$$\mathcal{L}_u = \mathcal{L}_u^p, \qquad \mathcal{W}_u = \mathcal{W}_u^p$$

as the completions of the space of smooth compactly supported sections of the pullback tangent bundle $u^*TM \to \mathbb{R}\times S^1$ with respect to the norms

$$\|\xi\|_\mathcal{L} = \left(\int_{-\infty}^\infty \int_0^1 |\xi|^p \, dtds\right)^{1/p},$$

$$\|\xi\|_\mathcal{W} = \left(\int_{-\infty}^\infty \int_0^1 |\xi|^p + |\nabla_s\xi|^p + |\nabla_t\nabla_t\xi|^p \, dtds\right)^{1/p}.$$

Then $\mathcal{D}_u^0 : \mathcal{W}_u^p \to \mathcal{L}_u^p$ is a Fredholm operator for $p > 1$ (Theorem A.4) with index

$$\mathrm{index}\mathcal{D}_u^0 = \mathrm{ind}_\mathcal{V}(x^-) - \mathrm{ind}_\mathcal{V}(x^+).$$

Here $\mathrm{ind}_\mathcal{V}(x)$ denotes the Morse index, i.e. the number of negative eigenvalues of the Hessian of $\mathcal{S}_\mathcal{V}$. This Hessian is given by

$$A^0(x)\xi = -\nabla_t\nabla_t\xi - R(\xi,\dot{x})\dot{x} - \mathcal{H}_\mathcal{V}(x)\xi,$$

where $\mathcal{H}_\mathcal{V}$ denotes the covariant Hessian of $\mathcal{V}$ (see Section 2). The Morse–Smale condition asserts that the operator $\mathcal{D}_u^0$ is surjective for every finite energy solution of (12). That this condition can be achieved by a generic perturbation $\mathcal{V}$ is proved in [31] (see Appendix A).

Linearizing equation (11) gives rise to the first order differential operator

$$\mathcal{D}_{u,v}^\varepsilon : W^{1,p}(\mathbb{R}\times S^1, u^*TM \oplus u^*TM) \to L^p(\mathbb{R}\times S^1, u^*TM\oplus u^*TM)$$

given by

$$\mathcal{D}_{u,v}^\varepsilon \begin{pmatrix}\xi\\ \eta\end{pmatrix} = \begin{pmatrix}\nabla_s\xi - \nabla_t\eta - R(\xi,\partial_t u)v - \mathcal{H}_\mathcal{V}(u)\xi \\ \nabla_s\eta + R(\xi,\partial_s u)v + \varepsilon^{-2}(\nabla_t\xi - \eta)\end{pmatrix} \tag{14}$$

for $(\xi,\eta) \in W^{1,p}(\mathbb{R}\times S^1, u^*TM\oplus u^*TM)$.

Remark 3.1. Assume $\mathcal{S}_\mathcal{V}$ is Morse and let $p > 1$. Then $\mathcal{D}_{u,v}^\varepsilon$ is a Fredholm operator for every pair (u,v) that satisfies (6) and its index is given by

$$\mathrm{index}\,\mathcal{D}_{u,v}^\varepsilon = \mathrm{ind}_\mathcal{V}(x^-) - \mathrm{ind}_\mathcal{V}(x^+).$$

To see this rescale u and v as in Remark 1.3. Then the operator on the rescaled vector fields $\tilde{\xi}(s,t) := \xi(\varepsilon^{-1}s,t)$ and $\tilde{\eta}(s,t) := g^{-1}\eta(\varepsilon^{-1}s,t)$ has the same form as in Floer's original papers [8] with the almost complex structure J_ε of Remark 1.3. That this operator is Fredholm was proved in [6, 23, 19] for $p = 2$. An elegant proof of the Fredholm property for general $p > 1$ was given by Donaldson [4] for the instanton case; it adapts easily to the symplectic case [22]. The Fredholm index can be expressed as a difference of the Conley–Zehnder indices [23, 19]. That it agrees with the difference of the Morse indices was proved in [27, 28].

Let us now fix a solution u of (12) and define $v := \partial_t u$. For this pair (u,v) we must prove that the operator $\mathcal{D}_u^\varepsilon := \mathcal{D}_{u,\partial_t u}^\varepsilon$ is onto for $\varepsilon > 0$ sufficiently small and prove an estimate for the right inverse which is independent of ε. We will establish this under the assumption that the operator $\mathcal{D}_u^0$ is onto. To obtain uniform estimates for the inverse with constants independent of ε we must work with suitable ε-dependent norms, which in case $p = 2$ are suggested by the energy identity (8). For compactly supported vector fields $\zeta = (\xi, \eta) \in \Omega^0(\mathbb{R} \times S^1, u^*TM \oplus u^*TM)$ define

$$\|\zeta\|_{0,p,\varepsilon} = \left(\int_{-\infty}^{\infty} \int_0^1 \left(|\xi|^p + \varepsilon^p |\eta|^p \right) dtds \right)^{1/p},$$

$$\begin{aligned}\|\zeta\|_{1,p,\varepsilon} = \Big(\int_{-\infty}^{\infty} \int_0^1 \big(&|\xi|^p + \varepsilon^p |\eta|^p + \varepsilon^p |\nabla_t \xi|^p + \varepsilon^{2p} |\nabla_t \eta|^p \\ &+ \varepsilon^{2p} |\nabla_s \xi|^p + \varepsilon^{3p} |\nabla_s \eta|^p \big) dtds \Big)^{1/p}.\end{aligned}$$

Theorem 3.2. *Let $(u,v) : \mathbb{R} \times S^1 \to TM$ be a smooth map such that v and the derivatives $\partial_s u, \partial_t u, \nabla_t \partial_s u, \nabla_t \partial_t u$ are bounded and $\lim_{s \to \pm\infty} u(s,t)$ exists, uniformly in t. Then, for every $p > 1$, there are positive constants c and ε_0 such that, for every $\varepsilon \in (0, \varepsilon_0)$ and every $\zeta = (\xi, \eta) \in W^{1,p}(\mathbb{R} \times S^1, u^*TM \oplus u^*TM)$, we have*

$$\begin{aligned}&\varepsilon^{-1} \|\nabla_t \xi - \eta\|_{L^p} + \|\nabla_t \eta\|_{L^p} + \|\nabla_s \xi\|_{L^p} + \varepsilon \|\nabla_s \eta\|_{L^p} \\ &\le c \left(\left\| \mathcal{D}_{u,v}^\varepsilon \zeta \right\|_{0,p,\varepsilon} + \|\xi\|_{L^p} + \varepsilon^2 \|\eta\|_{L^p} \right).\end{aligned} \tag{15}$$

The formal adjoint operator $(\mathcal{D}_{u,v}^\varepsilon)^*$ defined below satisfies the same estimate. Moreover, the constants c and ε_0 are invariant under s-shifts of u.

The formal adjoint operator

$$(\mathcal{D}_{u,v}^\varepsilon)^* : W^{2,p}(\mathbb{R} \times S^1, u^*TM \oplus u^*TM) \to W^{1,p}(\mathbb{R} \times S^1, u^*TM \oplus u^*TM)$$

with respect to the $(0,2,\varepsilon)$-inner product associated to the $(0,2,\varepsilon)$-norm has the form

$$(\mathcal{D}_{u,v}^\varepsilon)^* \begin{pmatrix} \xi \\ \eta \end{pmatrix} = \begin{pmatrix} -\nabla_s \xi - \nabla_t \eta - R(\xi, v)\partial_t u - \mathcal{H}_\mathcal{V}(u)\xi + \varepsilon^2 R(\eta, v)\partial_s u \\ -\nabla_s \eta + \varepsilon^{-2}(\nabla_t \xi - \eta) \end{pmatrix}$$

for $\xi, \eta \in W^{1,p}(\mathbb{R} \times S^1, u^*TM)$. We shall also use the operator

$$\pi_\varepsilon : L^p(S^1, x^*TM) \times L^p(S^1, x^*TM) \to W^{1,p}(S^1, x^*TM)$$

given by

$$\pi_\varepsilon(\xi, \eta) = (\mathbb{1} - \varepsilon \nabla_t \nabla_t)^{-1}(\xi - \varepsilon^2 \nabla_t \eta)$$

for $x \in \mathcal{L}M$ and $\xi, \eta \in \Omega^0(S^1, x^*TM)$. This operator, for the loop $x(t) = u(s,t)$, will be applied to the pair $(\xi(s,\cdot), \eta(s,\cdot))$. The rationale for introducing this operator is explained in Appendix D.

Theorem 3.3. *Assume $\mathcal{S}_\mathcal{V}$ is Morse-Smale and let $u \in \mathcal{M}^0(x^-, x^+; \mathcal{V})$. Then, for every $p > 1$, there are positive constants c and ε_0 (invariant under s-shifts of u) such that, for every $\varepsilon \in (0, \varepsilon_0)$ the following are true. The operator $\mathcal{D}_u^\varepsilon := \mathcal{D}_{u,\partial_t u}^\varepsilon$ is onto and for every pair*

$$\zeta^* := (\xi^*, \eta^*) \in \mathrm{im}\,(\mathcal{D}_u^\varepsilon)^* \subset W^{1,p}(\mathbb{R} \times S^1, u^*TM \oplus u^*TM)$$

we have

$$\|\xi^*\|_p + \varepsilon^{1/2} \|\eta^*\|_p + \varepsilon^{1/2} \|\nabla_t \xi^*\|_p \le c \left(\varepsilon \|\mathcal{D}_u^\varepsilon \zeta^*\|_{0,p,\varepsilon} + \|\pi_\varepsilon(\mathcal{D}_u^\varepsilon \zeta^*)\|_p \right), \quad (16)$$

$$\|\zeta^*\|_{1,p,\varepsilon} \le c \left(\varepsilon \|\mathcal{D}_u^\varepsilon \zeta^*\|_{0,p,\varepsilon} + \|\pi_\varepsilon(\mathcal{D}_u^\varepsilon \zeta^*)\|_p \right). \quad (17)$$

The proofs of Theorems 3.2 and 3.3 are given in Appendix D. They are based on a simplified form of Theorem 3.2 for flat manifolds with $\mathcal{V} = 0$ which is proved in Appendix C. In particular, Corollary C.3 shows that the ε-weights on the left hand side of equation (15) appear in a natural manner by a rescaling argument and, for $p = 2$, these terms can be interpreted as a linearized version of the energy. This was in fact the motivation for introducing the above ε-dependent norms. The proof of Theorem 3.3 is based on Theorem 3.2 and a comparison of the operators $\mathcal{D}_u^0$ and $\mathcal{D}_u^\varepsilon$.

To construct a solution of (11) near a parabolic cylinder it is useful to combine Theorems 3.2 and 3.3 into the following corollary. This corollary involves an ε-dependent norm which at first glance appears to be somewhat less natural but plays a useful role for technical reasons.

Given a smooth map $u : \mathbb{R} \times S^1 \to M$ and a compactly supported pair of vector fields $\zeta = (\xi, \eta) \in \Omega^0(\mathbb{R} \times S^1, u^*TM \oplus u^*TM)$ we define

$$\begin{aligned} |||\zeta|||_\varepsilon := &\|\xi\|_p + \varepsilon^{1/2} \|\eta\|_p + \varepsilon^{1/2} \|\nabla_t \xi\|_p + \|\eta - \nabla_t \xi\|_p + \varepsilon^2 \|\nabla_s \eta\|_p \\ &+ \varepsilon \|\nabla_t \eta\|_p + \varepsilon \|\nabla_s \xi\|_p + \varepsilon^{3/2p} \|\xi\|_\infty + \varepsilon^{1/2+2/p} \|\eta\|_\infty. \end{aligned} \quad (18)$$

For small ε this norm is much bigger than the $(1, p, \varepsilon)$-norm. If the last two summands on the right hand side of (18) are dropped one obtains an equivalent norm with a factor independent of ε (see (20) below).

Corollary 3.4. *Assume $\mathcal{S}_\mathcal{V}$ is Morse-Smale and let $u \in \mathcal{M}^0(x^-, x^+; \mathcal{V})$. Then, for every $p > 1$, there are positive constants c and ε_0 such that, for every $\varepsilon \in (0, \varepsilon_0)$ the following holds. If*

$$\zeta = (\xi, \eta) \in \mathrm{im}\,(\mathcal{D}_u^\varepsilon)^*, \qquad \zeta' = (\xi', \eta') := \mathcal{D}_u^\varepsilon \zeta,$$

then

$$|||\zeta|||_\varepsilon \le c \left(\|\xi'\|_p + \varepsilon^{3/2} \|\eta'\|_p \right). \quad (19)$$

Proof. Let c_2 be the constant of Theorem 3.2 and c_3 be the constant of Theorem 3.3. Then, by Theorem 3.3,

$$\begin{aligned}
&\|\xi\|_p + \varepsilon^{1/2}\|\eta\|_p + \varepsilon^{1/2}\|\nabla_t\xi\|_p \\
&\le c_3\left(\varepsilon\|\xi'\|_p + \varepsilon^2\|\eta'\|_p + \left\|(\mathbb{1} - \varepsilon\nabla_t\nabla_t)^{-1}(\xi' - \varepsilon^2\nabla_t\eta')\right\|_p\right) \\
&\le c_3\left((1+\varepsilon)\|\xi'\|_p + (\varepsilon^2 + \kappa_p\varepsilon^{3/2})\|\eta'\|_p\right) \\
&\le c_4\left(\|\xi'\|_p + \varepsilon^{3/2}\|\eta'\|_p\right).
\end{aligned}$$

Here the second step follows from Lemma D.3. Combining the last estimate with Theorem 3.2 we obtain

$$\begin{aligned}
&\|\eta - \nabla_t\xi\|_p + \varepsilon\|\nabla_t\eta\|_p + \varepsilon\|\nabla_s\xi\|_p + \varepsilon^2\|\nabla_s\eta\|_p \\
&\le c_2\varepsilon\left(\|\xi'\|_p + \varepsilon\|\eta'\|_p + \|\xi\|_p + \varepsilon^2\|\eta\|_p\right) \\
&\le c_2\left(\varepsilon\|\xi'\|_p + \varepsilon^2\|\eta'\|_p + c_4\varepsilon\left(\|\xi'\|_p + \varepsilon^{3/2}\|\eta'\|_p\right)\right) \\
&\le c_2(1+c_4)\left(\varepsilon\|\xi'\|_p + \varepsilon^2\|\eta'\|_p\right).
\end{aligned}$$

Now let c_5 be the constant of Lemma 3.5 below. Then

$$\begin{aligned}
\varepsilon^{3/2p}\|\xi\|_\infty &\le c_5\left(\|\xi\|_p + \varepsilon^{1/2}\|\nabla_t\xi\|_p + \varepsilon\|\nabla_s\xi\|_p\right), \\
\varepsilon^{1/2+2/p}\|\eta\|_\infty &\le c_5\left(\varepsilon^{1/2}\|\eta\|_p + \varepsilon\|\nabla_t\eta\|_p + \varepsilon^2\|\nabla_s\eta\|_p\right).
\end{aligned} \tag{20}$$

(Here we used the cases $(\beta_1,\beta_2) = (1/2,1)$ and $(\beta_1,\beta_2) = (1/2,3/2)$.) Combining these four estimates we obtain (19). □

The second estimate in the proof of Corollary 3.4 shows that one can obtain a stronger estimate than (19) from Theorems 3.2 and 3.3. Namely, (19) continues to hold if $|||\zeta|||_\varepsilon$ is replaced by the stronger norm where the L^p norms of $\nabla_t\xi - \eta$, $\nabla_t\eta$, $\nabla_s\xi$, and $\nabla_s\eta$ are multiplied by an additional factor $\varepsilon^{-1/2}$. The reason for not using this stronger norm lies in the proof of Theorem 4.1. In the first step of the iteration we solve an equation of the form $\mathcal{D}_u^\varepsilon\zeta_0 = \zeta' = (0,\eta')$ where η' is bounded (in L^p) with all its derivatives. Our goal in this first step is to obtain the sharpest possible estimate for ζ_0 and its first derivatives. We shall see that this estimate has the form $|||\zeta_0|||_\varepsilon \le c\varepsilon^2$ and that such an estimate in terms of ε^2 cannot be obtained with the stronger norm indicated above.

Lemma 3.5. *Let $u \in C^\infty(\mathbb{R}\times S^1, M)$ such that $\|\partial_s u\|_\infty$ and $\|\partial_t u\|_\infty$ are finite and $\lim_{s\to\pm\infty} u(s,t)$ exists, uniformly in t. Then, for every $p > 2$, there is a constant $c > 0$ such that*

$$\|\xi\|_\infty \le c\varepsilon^{-(\beta_1+\beta_2)/p}\left(\|\xi\|_p + \varepsilon^{\beta_1}\|\nabla_t\xi\|_p + \varepsilon^{\beta_2}\|\nabla_s\xi\|_p\right)$$

for every $\varepsilon \in (0,1]$, every pair of nonnegative real numbers β_1 and β_2, and every compactly supported vector field $\xi \in \Omega^0(\mathbb{R}\times S^1, u^*TM)$.

Proof. Define $\tilde{u} : Z_\varepsilon := \mathbb{R} \times \left(\mathbb{R}/\varepsilon^{-\beta_1}\mathbb{Z}\right) \to M$ and $\tilde{\xi} \in \Omega^0(Z_\varepsilon, \tilde{u}^*TM)$ by

$$\tilde{u}(s,t) := u(\varepsilon^{\beta_2}s, \varepsilon^{\beta_1}t), \qquad \tilde{\xi}(s,t) := \xi(\varepsilon^{\beta_2}s, \varepsilon^{\beta_1}t).$$

The estimate is equivalent to the Sobolev inequality

$$\|\tilde{\xi}\|_\infty \le c\left(\|\tilde{\xi}\|_p + \|\nabla_t\tilde{\xi}\|_p + \|\nabla_s\tilde{\xi}\|_p\right)$$

with a uniform constant $c = c(p, \|\partial_s u\|_\infty, \|\partial_t u\|_\infty)$ that is independent of $\varepsilon \in (0,1]$. (To see how the L^∞ bounds on $\partial_s u$ and $\partial_t u$ enter the estimate, embedd M into some Euclidean space and use the Gauss-Weingarten formula.) □

4 Existence and uniqueness

Throughout this section we fix a perturbation $\mathcal{V}$ that satisfies $(V0 - V4)$. In the next theorem we denote by

$$\Phi(x,\xi) : T_xM \to T_{\exp_x(\xi)}M$$

parallel transport along the geodesic $\tau \mapsto \exp_x(\tau\xi)$.

Theorem 4.1 (Existence). *Assume $\mathcal{S}_\mathcal{V}$ is Morse–Smale and fix two constants $a \in \mathbb{R}$ and $p > 2$. Then there are positive constants c and ε_0 such that the following holds. For every $\varepsilon \in (0, \varepsilon_0)$, every pair $x^\pm \in \mathcal{P}^a(\mathcal{V})$ of index difference one, and every $u \in \mathcal{M}^0(x^-, x^+; \mathcal{V})$, there exists a pair $(u^\varepsilon, v^\varepsilon) \in \mathcal{M}^\varepsilon(x^-, x^+; \mathcal{V})$ of the form*

$$u^\varepsilon = \exp_u(\xi), \qquad v^\varepsilon = \Phi(u,\xi)(\partial_t u + \eta), \qquad (\xi,\eta) \in \mathrm{im}\,(\mathcal{D}_u^\varepsilon)^*,$$

where ξ and η satisfy the inequalities

$$\begin{aligned} &\|\nabla_t\xi - \eta\|_{L^p} + \|\xi\|_{L^p} + \varepsilon^{1/2}\|\eta\|_{L^p} + \varepsilon^{1/2}\|\nabla_t\xi\|_{L^p} \\ &+ \varepsilon\|\nabla_t\eta\|_{L^p} + \varepsilon\|\nabla_s\xi\|_{L^p} + \varepsilon^2\|\nabla_s\eta\|_{L^p} \le c\varepsilon^2 \end{aligned} \tag{21}$$

and

$$\|\xi\|_{L^\infty} \le c\varepsilon^{2-3/2p}, \qquad \|\eta\|_{L^\infty} \le c\varepsilon^{3/2-2/p}. \tag{22}$$

Remark 4.2. The estimates (21) and (22) can be summarized in the form

$$|\!|\!|\zeta|\!|\!|_\varepsilon \le c\varepsilon^2$$

for $\zeta := (\xi, \eta)$ (with a larger constant c).

Theorem 4.3 (Uniqueness). *Assume $\mathcal{S}_\mathcal{V}$ is Morse–Smale and fix two constants $a \in \mathbb{R}$ and $C > 0$. Then there are positive constants δ and ε_0 such that, for every $\varepsilon \in (0, \varepsilon_0)$, every pair $x^\pm \in \mathcal{P}^a(\mathcal{V})$ of index difference one, and every $u \in \mathcal{M}^0(x^-, x^+; \mathcal{V})$ the following holds. If*

$$(\xi_i, \eta_i) \in \mathrm{im}\,(\mathcal{D}_u^\varepsilon)^*, \qquad \|\xi_i\|_{L^\infty} \le \delta\varepsilon^{1/2}, \qquad \|\eta_i\|_{L^\infty} \le C, \tag{23}$$

for $i = 1, 2$ and the pairs

$$u_i^\varepsilon := \exp_u(\xi_i), \qquad v_i^\varepsilon := \Phi(u, \xi_i)(\partial_t u + \eta_i),$$

belong to the moduli space $\mathcal{M}^\varepsilon(x^-, x^+; \mathcal{V})$, then $(u_1^\varepsilon, v_1^\varepsilon) = (u_2^\varepsilon, v_2^\varepsilon)$.

In the hypotheses of Theorem 4.3 we did not specify the Sobolev space to which $\zeta_i = (\xi_i, \eta_i)$ is required to belong. The reason is that ζ_i is smooth and, by exponential decay, belongs to the Sobolev space $W^{k,p}(\mathbb{R} \times S^1, u^*TM \oplus u^*TM)$ for every integer $k \geq 0$ and every $p \geq 1$.

Definition 4.4. *Assume $\mathcal{S}_\mathcal{V}$ is Morse–Smale and fix three constants $a \in \mathbb{R}$, $C > 0$, and $p > 2$. Choose positive constants ε_0, δ, and c such that the assertions of Theorem 4.1 and 4.3 hold with these constants. Shrink ε_0 so that $c\varepsilon_0^{1/2} < \delta$ and $c\varepsilon_0^{1/2} \leq C$. Define the map*

$$\mathcal{T}^\varepsilon : \mathcal{M}^0(x^-, x^+; \mathcal{V}) \to \mathcal{M}^\varepsilon(x^-, x^+; \mathcal{V})$$

by

$$\mathcal{T}^\varepsilon(u) := (u^\varepsilon, v^\varepsilon), \qquad u^\varepsilon := \exp_u(\xi), \qquad v^\varepsilon := \Phi(u, \xi)(\partial_t u + \eta),$$

where the pair $(\xi, \eta) \in \mathrm{im}\,(\mathcal{D}_u^\varepsilon)^*$ is chosen such that (21) and (22) are satisfied and $(\exp_u(\xi), \Phi(u, \xi)(\partial_t u + \eta)) \in \mathcal{M}^\varepsilon(x^-, x^+; \mathcal{V})$. Such a pair (ξ, η) exists, by Theorem 4.1, and is unique, by Theorem 4.3. The map $\mathcal{T}^\varepsilon$ is shift equivariant.

The proof of Theorem 4.1 is based on the Newton–Picard iteration method to detect a zero of a map near an approximate zero. The first step is to define a suitable map between Banach spaces. In order to do so let $(u, v) : \mathbb{R} \times S^1 \to TM$ be a smooth map and consider the map $\mathcal{F}_{u,v}^\varepsilon : W^{1,p}(\mathbb{R} \times S^1, u^*TM \oplus u^*TM) \to L^p(\mathbb{R} \times S^1, u^*TM \oplus u^*TM)$ given by

$$\mathcal{F}_{u,v}^\varepsilon \begin{pmatrix} \xi \\ \eta \end{pmatrix} := \begin{pmatrix} \Phi(u,\xi)^{-1} & 0 \\ 0 & \Phi(u,\xi)^{-1} \end{pmatrix} \mathcal{F}_\varepsilon \begin{pmatrix} \exp_u \xi \\ \Phi(u,\xi)(v + \eta) \end{pmatrix}, \tag{24}$$

where

$$\mathcal{F}_\varepsilon \begin{pmatrix} u^\varepsilon \\ v^\varepsilon \end{pmatrix} := \begin{pmatrix} \partial_s u^\varepsilon - \nabla_t v^\varepsilon - \operatorname{grad} \mathcal{V}(u^\varepsilon) \\ \nabla_s v^\varepsilon + \varepsilon^{-2}(\partial_t u^\varepsilon - v^\varepsilon) \end{pmatrix}. \tag{25}$$

Thus, abbreviating $\Phi := \Phi(u, \xi)$, we have

$$\mathcal{F}_{u,v}^\varepsilon \begin{pmatrix} \xi \\ \eta \end{pmatrix} := \begin{pmatrix} \Phi^{-1}\left(\partial_s \exp_u(\xi) - \nabla_t(\Phi(v + \eta)) - \operatorname{grad} \mathcal{V}(\exp_u(\xi))\right) \\ \Phi^{-1}\left(\nabla_s(\Phi(v + \eta)) + \varepsilon^{-2}\partial_t \exp_u(\xi)\right) - \varepsilon^{-2}(v + \eta) \end{pmatrix}.$$

Moreover, the differential of $\mathcal{F}_{u,v}^\varepsilon$ at the origin is given by $d\mathcal{F}_{u,v}^\varepsilon(0,0) = \mathcal{D}_{u,v}^\varepsilon$ (see [27, Appendix A.3]).

One of the key ingredients in the iteration is to have control over the variation of derivatives. This is provided by the following quadratic estimates.

Proposition 4.5. *There exists a constant $\delta > 0$ with the following significance. For every $p > 1$ and every $c_0 > 0$ there is a constant $c > 0$ such that the following is true. Let $(u, v) : \mathbb{R} \times S^1 \to TM$ be a smooth map and $Z = (X, Y), \zeta = (\xi, \eta) \in \Omega^0(\mathbb{R} \times S^1, u^*TM \oplus u^*TM)$ be two pairs of vector fields along u such that*

$$\|\partial_s u\|_\infty + \|\partial_t u\|_\infty + \|v\|_\infty \le c_0, \quad \|\xi\|_\infty + \|X\|_\infty \le \delta, \quad \|\eta\|_\infty + \|Y\|_\infty \le c_0.$$

Then the vector fields F_1, F_2 along u, defined by

$$\mathcal{F}^\varepsilon_{u,v}(Z + \zeta) - \mathcal{F}^\varepsilon_{u,v}(Z) - d\mathcal{F}^\varepsilon_{u,v}(Z)\zeta =: \begin{pmatrix} F_1 \\ F_2 \end{pmatrix},$$

satisfy the inequalities

$$\begin{aligned}
\|F_1\|_p \le c\,\|\xi\|_\infty &\Big(\|\xi\|_p + \|\eta\|_p + \|\nabla_t \xi\|_p + \|\nabla_s \xi\|_p \|\xi\|_\infty\Big) \\
&+ c\Big(\|\nabla_t X\|_p + \|\nabla_s X\|_p\Big) \|\xi\|_\infty^2 + c\,\|\nabla_t X\|_p \|\xi\|_\infty \|\eta\|_\infty \\
&+ c\,\|X\|_\infty \Big(\|\nabla_s \xi\|_p \|\xi\|_\infty + \|\nabla_t \xi\|_p \|\eta\|_\infty\Big), \\
\|F_2\|_p \le c\,\|\xi\|_\infty &\Big(\varepsilon^{-2} \|\xi\|_p + \|\eta\|_p + \|\nabla_s \xi\|_p + \varepsilon^{-2} \|\nabla_t \xi\|_p \|\xi\|_\infty\Big) \\
&+ c\Big(\|\nabla_s X\|_p + \varepsilon^{-2} \|\nabla_t X\|_p\Big) \|\xi\|_\infty^2 + c\,\|\nabla_s X\|_p \|\xi\|_\infty \|\eta\|_\infty \\
&+ c\,\|X\|_\infty \Big(\varepsilon^{-2} \|\nabla_t \xi\|_p \|\xi\|_\infty + \|\nabla_s \xi\|_p \|\eta\|_\infty\Big).
\end{aligned}$$

Proposition 4.6. *There exists a constant $\delta > 0$ with the following significance. For every $p > 1$ and every $c_0 > 0$ there is a constant $c > 0$ such that the following is true. Let $(u, v) : \mathbb{R} \times S^1 \to TM$ be a smooth map and $Z = (X, Y), \zeta = (\xi, \eta) \in \Omega^0(\mathbb{R} \times S^1, u^*TM \oplus u^*TM)$ be two pairs of vector fields along u such that*

$$\|\partial_s u\|_\infty + \|\partial_t u\|_\infty + \|v\|_\infty \le c_0, \qquad \|X\|_\infty \le \delta, \qquad \|Y\|_\infty \le c_0.$$

Then the vector fields F_1, F_2 along u, defined by

$$d\mathcal{F}^\varepsilon_{u,v}(Z)\zeta - d\mathcal{F}^\varepsilon_{u,v}(0)\zeta =: \begin{pmatrix} F_1 \\ F_2 \end{pmatrix},$$

satisfy the inequalities

$$\begin{aligned}
\|F_1\|_p \le c\,\|\xi\|_\infty &\Big(\|X\|_p + \|Y\|_p + \|\nabla_t X\|_p + \|\nabla_s X\|_p \|X\|_\infty\Big) \\
&+ c\,\|X\|_\infty \Big(\|\eta\|_p + \|\nabla_t \xi\|_p + \|\nabla_s \xi\|_p \|X\|_\infty + \|\nabla_t X\|_p \|\eta\|_\infty\Big), \\
\|F_2\|_p \le c\,\|\xi\|_\infty &\Big(\varepsilon^{-2} \|X\|_p + \varepsilon^{-2} \|\nabla_t X\|_p \|X\|_\infty + \|Y\|_p + \|\nabla_s X\|_p\Big) \\
&+ c\,\|X\|_\infty \Big(\varepsilon^{-2} \|\nabla_t \xi\|_p \|X\|_\infty + \|\eta\|_p + \|\nabla_s \xi\|_p + \|\nabla_s X\|_p \|\eta\|_\infty\Big).
\end{aligned}$$

For the proof of Propositions 4.5 and 4.6 we refer to [27, Chapter 5]. To understand the estimate of Proposition 4.6 note that η and Y appear only as zeroth order terms, that $\nabla_s\xi$ and $\nabla_s X$ appear only in cubic terms in F_1, and that $\nabla_t\xi$ and $\nabla_t X$ appear only in cubic terms in F_2. This follows from the fact that the first component of $\mathcal{F}_\varepsilon$ is linear in $\partial_s u$ and the second component is linear in $\partial_t u$. In Proposition 4.5 we have included cubic terms that arise when the derivative hits X. In this case we must use the L^∞ norms on the factors ξ and η and can profit from the fact that $\nabla_s X$ and $\nabla_t X$ will be small in L^p. The constant δ appears as a condition for the pointwise quadratic estimates in suitable coordinate charts on M.

We now reformulate the quadratic estimates in terms of the norm (18).

Corollary 4.7. *There exists a constant $\delta > 0$ with the following significance. For every $p > 1$ and every $c_0 > 0$ there is a constant $c > 0$ such that the following holds. If (u,v), $Z = (X,Y)$ and $\zeta = (\xi,\eta)$ satisfy the hypotheses of Proposition 4.5 then*

$$\begin{aligned}&\left\|\mathcal{F}^\varepsilon_{u,v}(Z+\zeta) - \mathcal{F}^\varepsilon_{u,v}(Z) - d\mathcal{F}^\varepsilon_{u,v}(Z)\zeta\right\|_{0,p,\varepsilon^{3/2}}\\ &\le c|||\zeta|||_\varepsilon \left(\varepsilon^{-1/2}\left\|\xi\right\|_\infty + \varepsilon^{-1}\left\|\xi\right\|_\infty^2\right) + c\varepsilon^{-1-3/2p}|||Z|||_\varepsilon|||\zeta|||_\varepsilon\left(\left\|\xi\right\|_\infty + \varepsilon^{1/2}\left\|\eta\right\|_\infty\right).\end{aligned}$$

If (u,v), $Z = (X,Y)$ and $\zeta = (\xi,\eta)$ satisfy the hypotheses of Proposition 4.6 then

$$\left\|d\mathcal{F}^\varepsilon_{u,v}(Z)\zeta - d\mathcal{F}^\varepsilon_{u,v}(0)\zeta\right\|_{0,p,\varepsilon^{3/2}} \le c\left(\varepsilon^{-1/2-3/2p}|||Z|||_\varepsilon + \varepsilon^{-1-7/2p}|||Z|||_\varepsilon^2\right)|||\zeta|||_\varepsilon.$$

Proof. The result follows from Propositions 4.5 and 4.6 via term by term inspection. In particular, we must use the inequalities

$$\left\|\xi\right\|_\infty \le \varepsilon^{-3/2p}|||\zeta|||_\varepsilon, \qquad \left\|\eta\right\|_\infty \le \varepsilon^{-1/2-2/p}|||\zeta|||_\varepsilon, \qquad \left\|X\right\|_\infty \le \varepsilon^{-3/2p}|||Z|||_\varepsilon$$

at various places. These follow from the definition of the norm (18). □

Proof of Theorem 4.1. Given $u \in \mathcal{M}^0(x^-,x^+;\mathcal{V})$ with $x^\pm \in \mathcal{P}^a(\mathcal{V})$ we aim to detect an element of $\mathcal{M}^\varepsilon(x^-,x^+;\mathcal{V})$ near u. We set $v := \partial_t u$ and carry out the Newton–Picard iteration method for the map $\mathcal{F}^\varepsilon_u := \mathcal{F}^\varepsilon_{u,\partial_t u}$. Key ingredients are a small initial value, a uniformly bounded right inverse and control over the variation of derivatives (which is provided by the quadratic estimates above). Because $\mathcal{S}_\mathcal{V}$ is Morse-Smale, the sets $\mathcal{P}^a(\mathcal{V})$ and $\mathcal{M}^0(x^-,x^+;\mathcal{V})/\mathbb{R}$ are finite (the latter in addition relies on the assumption of index difference one). All constants appearing below turn out to be invariant under s-shifts of u. Hence they can be chosen to depend on a only.

Since $\mathcal{M}^0(x^-,x^+;\mathcal{V})/\mathbb{R}$ is a finite set it follows from Theorems A.1 and A.2 that there is a constant $c_0 > 0$ such that

$$\left\|\partial_s u\right\|_\infty + \left\|\partial_t u\right\|_\infty + \left\|\nabla_t\partial_t u\right\|_\infty \le c_0 \tag{26}$$

and

$$\left\|\nabla_t\partial_s u\right\|_\infty + \left\|\nabla_t\partial_s u\right\|_p + \left\|\nabla_t\nabla_t\partial_s u\right\|_p \le c_0 \tag{27}$$

for every $u \in \mathcal{M}^0(x^-, x^+; \mathcal{V})$. Thus the assumptions in Theorem 3.2, Theorem 3.3, Proposition 4.5, Proposition 4.6 and Lemma 3.5 are satisfied. Moreover, by (27) the value of the initial point $Z_0 := 0$ is indeed small with respect to the $(0, p, \varepsilon)$-norm:

$$\|\mathcal{F}_u^\varepsilon(0)\|_{0,p,\varepsilon} = \|\mathcal{F}^\varepsilon(u, \partial_t u)\|_{0,p,\varepsilon} = \left\| \begin{pmatrix} 0 \\ \nabla_s \partial_t u \end{pmatrix} \right\|_{0,p,\varepsilon} \leq c_0 \varepsilon. \tag{28}$$

Here we used in addition (24), (25) and the parabolic equations. Define the initial correction term $\zeta_0 = (\xi_0, \eta_0)$ by

$$\zeta_0 := -\mathcal{D}_u^{\varepsilon\,*} (\mathcal{D}_u^\varepsilon \mathcal{D}_u^{\varepsilon\,*})^{-1} \mathcal{F}_u^\varepsilon(0).$$

Recursively, for $\nu \in \mathbb{N}$, define the sequence of correction terms $\zeta_\nu = (\xi_\nu, \eta_\nu)$ by

$$\zeta_\nu := -\mathcal{D}_u^{\varepsilon\,*} (\mathcal{D}_u^\varepsilon \mathcal{D}_u^{\varepsilon\,*})^{-1} \mathcal{F}_u^\varepsilon(Z_\nu), \qquad Z_\nu = (X_\nu, Y_\nu) := \sum_{\ell=0}^{\nu-1} \zeta_\ell. \tag{29}$$

We prove by induction that there is a constant $c > 0$ such that

$$\|\!|\zeta_\nu|\!\|_\varepsilon \leq \frac{c}{2^\nu} \varepsilon^2, \qquad \|\mathcal{F}_u^\varepsilon(Z_{\nu+1})\|_{0,p,\varepsilon^{3/2}} \leq \frac{c}{2^\nu} \varepsilon^{7/2 - 3/2p}. \tag{H_ν}$$

Initial Step: $\nu = 0$. By definition of ζ_0 we have

$$\mathcal{D}_u^\varepsilon \zeta_0 = -\mathcal{F}_u^\varepsilon(0) = \begin{pmatrix} 0 \\ -\nabla_s \partial_t u \end{pmatrix}.$$

Thus, by Theorem 3.3 (with constant $c_1 > 0$),

$$\begin{aligned} \|\xi_0\|_p + \varepsilon^{1/2} \|\eta_0\|_p + \varepsilon^{1/2} \|\nabla_t \xi_0\|_p &\leq c_1 \left(\varepsilon \|(0, \nabla_s \partial_t u)\|_{0,p,\varepsilon} + \|\pi_\varepsilon(0, \nabla_s \partial_t u)\|_p \right) \\ &\leq c_1 \left(\varepsilon^2 \|\nabla_s \partial_t u\|_p + \varepsilon^2 \|\nabla_t \nabla_s \partial_t u\|_p \right) \\ &\leq c_0 c_1 \varepsilon^2. \end{aligned}$$

Here the second inequality follows from Lemma D.3 and the last from (27). By Theorem 3.2 (with constant $c_2 > 0$),

$$\begin{aligned} &\|\nabla_t \xi_0 - \eta_0\|_p + \varepsilon \|\nabla_t \eta_0\|_p + \varepsilon \|\nabla_s \xi_0\|_p + \varepsilon^2 \|\nabla_s \eta_0\|_p \\ &\leq c_2 \varepsilon \left(\|(0, \nabla_s \partial_t u)\|_{0,p,\varepsilon} + \|\xi_0\|_p + \varepsilon^2 \|\eta_0\|_p \right) \\ &\leq c_2 \varepsilon \left(\varepsilon \|\nabla_s \partial_t u\|_p + c_0 c_1 \varepsilon^2 \right) \\ &\leq c_0 c_2 (1 + c_1 \varepsilon) \varepsilon^2. \end{aligned}$$

The last inequality follows again from (27). Combining these two estimates with (20) we obtain

$$\varepsilon^{3/2p} \|\xi_0\|_\infty + \varepsilon^{1/2 + 2/p} \|\eta_0\|_\infty \leq \|\!|\zeta_0|\!\|_\varepsilon \leq c \varepsilon^2. \tag{30}$$

with a suitable constant $c > 0$ (depending only on c_0, c_1, c_2 and the constant of Lemma 3.5). This proves the first estimate in (H_ν) for $\nu = 0$. To prove the second estimate we observe that $Z_1 = \zeta_0$ and hence, by Proposition 4.5 (with constant $c_3 > 0$),

$$\begin{aligned}
&\|\mathcal{F}_u^\varepsilon(Z_1)\|_{0,p,\varepsilon^{3/2}} \\
&= \|\mathcal{F}_u^\varepsilon(\zeta_0) - \mathcal{F}_u^\varepsilon(0) - \mathcal{D}_u^\varepsilon \zeta_0\|_{0,p,\varepsilon^{3/2}} \\
&\le c_3 \|\xi_0\|_\infty \Big(\|\xi_0\|_p + \|\eta_0\|_p + \|\nabla_t \xi_0\|_p + \|\nabla_s \xi_0\|_p \|\xi_0\|_\infty \Big) \\
&\quad + c_3 \varepsilon^{3/2} \|\xi_0\|_\infty \Big(\varepsilon^{-2} \|\xi_0\|_p + \|\eta_0\|_p + \|\nabla_s \xi_0\|_p + \varepsilon^{-2} \|\nabla_t \xi_0\|_p \|\xi_0\|_\infty \Big) \\
&\le c\varepsilon^{7/2-3/2p}.
\end{aligned}$$

with a suitable constant $c > 0$ (depending only on c_0, c_1, c_2 and the constant of Lemma 3.5). Thus we have proved (H_ν) for $\nu = 0$. From now on we fix the constant c for which the estimate (H_0) has been established.

Induction step: $\boldsymbol{\nu - 1 \Rightarrow \nu}$. Let $\nu \ge 1$ and assume that $(H_0), \dots, (H_{\nu-1})$ are true. Then

$$\begin{aligned}
|\!|\!| Z_\nu |\!|\!|_\varepsilon &\le \sum_{\ell=0}^{\nu-1} |\!|\!| \zeta_\ell |\!|\!|_\varepsilon \le c\varepsilon^2 \sum_{\ell=0}^{\nu-1} 2^{-\ell} \le 2c\varepsilon^2, \\
\|\mathcal{F}_u^\varepsilon(Z_\nu)\|_{0,p,\varepsilon^{3/2}} &\le \frac{c}{2^{\nu-1}} \varepsilon^{7/2-3/2p}.
\end{aligned}$$

By (29) we have

$$\mathcal{D}_u^\varepsilon \zeta_\nu = -\mathcal{F}_u^\varepsilon(Z_\nu), \qquad \zeta_\nu \in \mathrm{im}(\mathcal{D}_u^\varepsilon)^*.$$

Hence, by Corollary 3.4, (with constant $c_4 > 0$),

$$|\!|\!| \zeta_\nu |\!|\!|_\varepsilon \le c_4 \, \|\mathcal{F}_u^\varepsilon(Z_\nu)\|_{0,p,\varepsilon^{3/2}} \le \frac{cc_4}{2^{\nu-1}} \varepsilon^{7/2-3/2p} \le \frac{c}{2^\nu} \varepsilon^2. \tag{31}$$

The last inequality holds whenever $c_4 \varepsilon^{3/2-3/2p} \le 1/2$.

By what we have just proved the vector fields Z_ν and ζ_ν satisfy the requirements of Corollary 4.7 (with the constant $c_5 > 0$). Hence

$$\begin{aligned}
\|\mathcal{F}_u^\varepsilon(Z_{\nu+1})\|_{0,p,\varepsilon^{3/2}} &\le \|\mathcal{F}_u^\varepsilon(Z_\nu + \zeta_\nu) - \mathcal{F}_u^\varepsilon(Z_\nu) - d\mathcal{F}_u^\varepsilon(Z_\nu)\zeta_\nu\|_{0,p,\varepsilon^{3/2}} \\
&\quad + \|d\mathcal{F}_u^\varepsilon(Z_\nu)\zeta_\nu - \mathcal{D}_u^\varepsilon \zeta_\nu\|_{0,p,\varepsilon^{3/2}} \\
&\le c_5 \Big(\varepsilon^{-1/2} \|\xi_\nu\|_\infty + \varepsilon^{-1} \|\xi_\nu\|_\infty^2 \Big) |\!|\!| \zeta_\nu |\!|\!|_\varepsilon \\
&\quad + c_5 \varepsilon^{-1-3/2p} |\!|\!| Z_\nu |\!|\!|_\varepsilon \Big(\|\xi_\nu\|_\infty + \varepsilon^{1/2} \|\eta_\nu\|_\infty \Big) |\!|\!| \zeta_\nu |\!|\!|_\varepsilon \\
&\quad + c_5 \varepsilon^{-1/2-3/2p} |\!|\!| Z_\nu |\!|\!|_\varepsilon |\!|\!| \zeta_\nu |\!|\!|_\varepsilon + c_5 \varepsilon^{-1-7/2p} |\!|\!| Z_\nu |\!|\!|_\varepsilon^2 |\!|\!| \zeta_\nu |\!|\!|_\varepsilon \\
&\le c_5 \Big(c\varepsilon^{3/2-3/2p} + c^2 \varepsilon^{3-3/p} \Big) |\!|\!| \zeta_\nu |\!|\!|_\varepsilon + 2c^2 c_5 \varepsilon^{3-7/2p} |\!|\!| \zeta_\nu |\!|\!|_\varepsilon \\
&\quad + 2cc_5 \varepsilon^{3/2-3/2p} |\!|\!| \zeta_\nu |\!|\!|_\varepsilon + 4c^2 c_5 \varepsilon^{3-7/2p} |\!|\!| \zeta_\nu |\!|\!|_\varepsilon \\
&\le \frac{1}{2c_4} |\!|\!| \zeta_\nu |\!|\!|_\varepsilon \\
&\le \frac{c}{2^\nu} \varepsilon^{7/2-3/2p}.
\end{aligned}$$

In the third step we have used the inequalities

$$\|\xi_\nu\|_\infty \le \varepsilon^{-3/2p} |||\zeta_\nu|||_\varepsilon \le c\varepsilon^{2-3/2p}$$

and

$$\|\xi_\nu\|_\infty + \varepsilon^{1/2} \|\eta_\nu\|_\infty \le \varepsilon^{-2/p} |||\zeta_\nu|||_\varepsilon \le c\varepsilon^{2-2/p}$$

as well as $|||Z|||_\nu \le 2c\varepsilon^2$. The fourth step holds for ε sufficiently small, and the last step follows from (31). This completes the induction and proves (H_ν) for every ν.

It follows from (H_ν) that Z_ν is a Cauchy sequence with respect to $|||\cdot|||_\varepsilon$. Denote its limit by

$$\zeta := \lim_{\nu\to\infty} Z_\nu = \sum_{\nu=0}^{\infty} \zeta_\nu.$$

By construction and by (H_ν), the limit satisfies

$$|||\zeta|||_\varepsilon \le 2c\varepsilon^2, \qquad \mathcal{F}_\nu^\varepsilon(\zeta) = 0, \qquad \zeta \in \mathrm{im}\,(\mathcal{D}_u^\varepsilon)^*.$$

Hence, by (24), the pair

$$(u^\varepsilon, v^\varepsilon) := (exp_u(\xi), \Phi(u,\xi)(\partial_t u + \eta))$$

is a solution of (11). Since $|||\zeta|||_\varepsilon$ is finite so is the L^p-norm of $(\partial_s u^\varepsilon, \nabla_s v^\varepsilon)$. Hence, by the standard elliptic bootstrapping arguments for pseudoholomorphic curves, the shifted functions $u^\varepsilon(s+\cdot,\cdot), v^\varepsilon(s+\cdot,\cdot)$ converge in the C^∞ topology on every compact set as s tends to $\pm\infty$. Since $\zeta \in W^{1,p}$, the limits must be the periodic orbits $x^\pm$ and, moreover, the pair $(\partial_s u^\varepsilon(s,t), \nabla_s v^\varepsilon(s,t))$ converges to zero, uniformly in t, as s tends to $\pm\infty$. Hence $(u^\varepsilon, v^\varepsilon) \in \mathcal{M}^\varepsilon(x^-,x^+;\mathcal{V})$. Evidently, each step in the iteration including the constants in the estimates is invariant under time shift. This proves Theorem 4.1. □

Proof of Theorem 4.3. Fix a constant $p > 2$ and an index one parabolic cylinder $u \in \mathcal{M}^0(x^-,x^+;\mathcal{V})$. Denote $v := \partial_t u$ and $\mathcal{F}_u^\varepsilon := \mathcal{F}_{u,\partial_t u}^\varepsilon$. As in the proof of Theorem 4.1, the map u satisfies the estimates (26) and (27). Denote by

$$\mathcal{T}^\varepsilon(u) = (u^\varepsilon, v^\varepsilon) = (\exp_u(X), \Phi(u,X)(\partial_t u + Y))$$

the solution of (11) constructed in Theorem 4.1 and let $Z := (X,Y)$. Then

$$Z \in \mathrm{im}\,(\mathcal{D}_u^\varepsilon)^*, \qquad \mathcal{F}_u^\varepsilon(Z) = 0, \qquad |||Z|||_\varepsilon \le c\varepsilon^2$$

for a suitable constant $c > 0$. Now suppose $(u^\varepsilon, v^\varepsilon) \in \mathcal{M}^\varepsilon(x^-,x^+;\mathcal{V})$ satisfies the hypotheses of the theorem. This means that there is a pair

$$\zeta = (\xi,\eta) \in W^{1,p}(\mathbb{R}\times S^1, u^*TM \oplus u^*TM)$$

such that

$$\zeta \in \mathrm{im}\,(\mathcal{D}_u^\varepsilon)^*, \qquad \mathcal{F}_u^\varepsilon(\zeta) = 0, \qquad \|\xi\|_\infty \le \delta\varepsilon^{1/2}, \qquad \|\eta\|_\infty \le C.$$

The difference

$$\zeta' := (\xi', \eta') := \zeta - Z$$

satisfies the inequalities

$$\|\xi'\|_\infty \le \delta\varepsilon^{1/2} + c\varepsilon^{2-3/2p} \le 2\delta\varepsilon^{1/2}, \qquad \|\eta'\|_\infty \le C + c\varepsilon^{3/2-2/p} \le 2C,$$

provided that ε is sufficiently small. Hence, by Corollary 3.4 (with a constant $c_1 > 0$) and Corollary 4.7 (with a constant $c_2 > 0$), we have

$$\begin{aligned}
\|\!|\zeta'|\!\|_\varepsilon &\le c_1 \|\mathcal{D}_u^\varepsilon \zeta'\|_{0,p,\varepsilon^{3/2}} \\
&\le c_1 \|\mathcal{F}_u^\varepsilon(Z+\zeta') - \mathcal{F}_u^\varepsilon(Z) - d\mathcal{F}_u^\varepsilon(Z)\zeta'\|_{0,p,\varepsilon^{3/2}} \\
&\quad + c_1 \|d\mathcal{F}_u^\varepsilon(Z)\zeta' - d\mathcal{F}_u^\varepsilon(0)\zeta'\|_{0,p,\varepsilon^{3/2}} \\
&\le c_1c_2 \left(\varepsilon^{-1/2}\|\xi'\|_\infty + \varepsilon^{-1}\|\xi'\|_\infty^2\right) \|\!|\zeta'|\!\|_\varepsilon \\
&\quad + c_1c_2\varepsilon^{-1-3/2p}\|\!|Z|\!\|_\varepsilon \Big(\|\xi'\|_\infty + \varepsilon^{1/2}\|\eta'\|_\infty\Big)\|\!|\zeta'|\!\|_\varepsilon \\
&\quad + c_1c_2\varepsilon^{-1/2-3/2p}\|\!|Z|\!\|_\varepsilon\|\!|\zeta'|\!\|_\varepsilon + c_1c_2\varepsilon^{-1-7/2p}\|\!|Z|\!\|_\varepsilon^2\|\!|\zeta'|\!\|_\varepsilon \\
&\le c_1c_2\left(2\delta + 4\delta^2\right)\|\!|\zeta'|\!\|_\varepsilon + cc_1c_2\varepsilon^{3/2-3/2p}\Big(2\delta + 2C\Big)\|\!|\zeta'|\!\|_\varepsilon \\
&\quad + cc_1c_2\varepsilon^{3/2-3/2p}\|\!|\zeta'|\!\|_\varepsilon + c^2c_1c_2\varepsilon^{3-7/2p}\|\!|\zeta'|\!\|_\varepsilon \\
&\le \frac{1}{2}\|\!|\zeta'|\!\|_\varepsilon.
\end{aligned}$$

The last inequality holds when δ and ε are sufficiently small. It follows that $\zeta' = 0$ and this proves the theorem. □

5 An apriori estimate

In this section we prove that the solutions of (11) and (6) are all contained in a fixed compact subset of T^*M that is independent of ε. Recall that the energy of such a solution is given by

$$E^\varepsilon(u,v) = \int_{-\infty}^{\infty}\int_0^1 \left(\|\partial_s u\|^2 + \varepsilon^2\|\nabla_s v\|^2\right) dtds = \mathcal{S}_\mathcal{V}(x^-) - \mathcal{S}_\mathcal{V}(x^+).$$

Theorem 5.1. *Fix a constant $c_0 > 0$ and a perturbation $\mathcal{V} : \mathcal{L}M \to \mathbb{R}$ that satisfies $(V0)$ and $(V1)$. Then there is a constant $C = C(c_0, \mathcal{V}) > 0$ such that the following holds. If $0 < \varepsilon \le 1$ and $(u,v) : \mathbb{R}\times S^1 \to TM$ is a solution of (11) such that*

$$E^\varepsilon(u,v) \le c_0, \qquad \sup_{s\in\mathbb{R}} \mathcal{A}_\mathcal{V}(u(s,\cdot), v(s,\cdot)) \le c_0 \tag{32}$$

then $\|v\|_\infty \le C$.

For $\varepsilon = 1$ and $\mathcal{V}(x) = \int_0^1 V_t(x(t))\,dt$ this result was proved by Cieliebak [2, Theorem 5.4]. His proof combines the 2-dimensional maximum principle and the Krein-Rutman theorem. Our proof is based on the following L^2-estimate.

Proposition 5.2. *Fix a constant $c_0 > 0$ and a perturbation $\mathcal{V} : \mathcal{L}M \to \mathbb{R}$ that satisfies $(V0)$ and $(V1)$. Then there is a constant $c = c(c_0, \mathcal{V}) > 0$ such that the following holds. If $0 < \varepsilon \le 1$ and $(u, v) : \mathbb{R} \times S^1 \to TM$ is a solution of (11) that satisfies (32) then*

$$\sup_{s \in \mathbb{R}} \int_0^1 |v(s,t)|^2 \, dt \le c.$$

Proof. Define $F : \mathbb{R} \to \mathbb{R}$ by

$$F(s) := \int_0^1 |v(s,t)|^2 \, dt.$$

We prove that there is a constant $\mu = \mu(\mathcal{V}) > 0$ such that

$$\varepsilon^2 F'' - F' + \mu F + 1 \ge 0. \tag{33}$$

To see this we abbreviate

$$L_\varepsilon := \varepsilon^2 \partial_s^2 + \partial_t^2 - \partial_s, \qquad \mathcal{L}_\varepsilon := \varepsilon^2 \nabla_s \nabla_s + \nabla_t \nabla_t - \nabla_s.$$

By (11), we have

$$\mathcal{L}_\varepsilon v = -\nabla_t \operatorname{grad} \mathcal{V}(u) \tag{34}$$

and hence

$$\begin{aligned} L_\varepsilon \frac{|v|^2}{2} &= \varepsilon^2 |\nabla_s v|^2 + |\nabla_t v|^2 + \langle \mathcal{L}_\varepsilon v, v \rangle \\ &= \varepsilon^2 |\nabla_s v|^2 + |\nabla_t v|^2 - \langle \nabla_t \operatorname{grad} \mathcal{V}(u), v \rangle \\ &\ge \varepsilon^2 |\nabla_s v|^2 + |\nabla_t v|^2 - C\bigl(1 + |\partial_t u|\bigr) |v| \\ &\ge \varepsilon^2 |\nabla_s v|^2 + |\nabla_t v|^2 - C\bigl(1 + |v| + \varepsilon^2 |\nabla_s v|\bigr) |v| \\ &\ge \frac{\varepsilon^2}{2} |\nabla_s v|^2 + |\nabla_t v|^2 - \left(\frac{C^2}{2} + C + \frac{\varepsilon^2 C^2}{2} \right) |v|^2 - \frac{1}{2} \\ &\ge -\bigl(C + C^2\bigr) |v|^2 - \frac{1}{2}. \end{aligned} \tag{35}$$

Here C is the constant in $(V1)$. Integrating this inequality over the interval $0 \le t \le 1$ gives (33) with $\mu := 2C + 2C^2$. It follows from (33) and Lemma B.3 with f replaced by $f + 1/\mu$ and $r := 1/2$ that

$$F(s) \le F(s) + \frac{1}{\mu} \le 16 c_2 e^{\mu/4} \int_{s-1}^{s+1} \left(F(\sigma) + \frac{1}{\mu} \right) d\sigma \tag{36}$$

for every $s \in \mathbb{R}$.

Next we observe that, by (32), we have

$$\begin{aligned}
c_0 &\geq \mathcal{A}_{\mathcal{V}}(u(s,\cdot), v(s,\cdot)) \\
&= \int_0^1 \left(\langle v(s,t), \partial_t u(s,t)\rangle - \frac{|v(s,t)|^2}{2} \right) dt - \mathcal{V}(u(s,\cdot)) \\
&= \int_0^1 \left(\frac{|v(s,t)|^2}{2} - \varepsilon^2 \langle v(s,t), \nabla_s v(s,t)\rangle \right) dt - \mathcal{V}(u(s,\cdot)) \\
&\geq \int_0^1 \left(\frac{|v(s,t)|^2}{4} - \varepsilon^4 |\nabla_s v(s,t)|^2 \right) dt - C.
\end{aligned}$$

Here C is the constant in $(V0)$ and we have used the fact that $\partial_t u = v - \varepsilon^2 \nabla_s v$. This implies

$$F(s) \leq 4 \left(c_0 + C + \int_0^1 \varepsilon^2 |\nabla_s v(s,t)|^2 \, dt \right)$$

for every $s \in \mathbb{R}$. Integrating this inequality we obtain

$$\int_{s-1}^{s+1} F(\sigma)\, d\sigma \leq 8c_0 + 8C + 8E^\varepsilon(u,v) \leq 16c_0 + 8C.$$

Now the assertion follows from (36). □

Proof of Theorem 5.1. In the proof of Proposition 5.2 we have seen that there is a constant $\mu = \mu(\mathcal{V}) > 0$ such that every solution (u,v) of (11) with $0 < \varepsilon \leq 1$ satisfies the inequality

$$L_\varepsilon |v|^2 \geq -\mu |v|^2 - 1. \tag{37}$$

Now let $(s_0, t_0) \in \mathbb{R} \times S^1$ and apply Lemma B.2 with $r = 1$ to the function $w : \mathbb{R} \times \mathbb{R} \supset P_1^\varepsilon \to \mathbb{R}$, given by $w(s,t) := |v(s+s_0, t+t_0)|^2 + 1/\mu$:

$$\begin{aligned}
|v(s_0,t_0)|^2 &\leq 2c_2 e^\mu \int_{-1-\varepsilon}^{\varepsilon} \int_{-1}^{1} \left(|v(s+s_0, t+t_0)|^2 + \frac{1}{\mu} \right) dt ds \\
&\leq 12 c_2 e^\mu \left(\frac{1}{\mu} + \sup_{s\in\mathbb{R}} \int_0^1 |v(s,t)|^2 \, dt \right).
\end{aligned}$$

Hence the result follows from Proposition 5.2. □

6 Gradient bounds

Theorem 6.1. *Fix a constant $c_0 > 0$ and a perturbation $\mathcal{V} : \mathcal{L}M \to \mathbb{R}$ that satisfies $(V0 - V3)$. Then there is a constant $C = C(c_0, \mathcal{V}) > 0$ such that the following holds. If $0 < \varepsilon \leq 1$ and $(u,v) : \mathbb{R} \times S^1 \to TM$ is a solution of (11)* *that satisfies (32), i.e.* $E^\varepsilon(u,v) \le c_0$ *and* $\sup_{s\in\mathbb{R}} \mathcal{A}_\mathcal{V}(u(s,\cdot),v(s,\cdot)) \le c_0$*, then*

$$\begin{aligned}&|\partial_s u(s,t)|^2 + |\nabla_s v(s,t)|^2\\&+\int_{s-1/2}^{s+1/2}\int_0^1 \Big(|\nabla_t\partial_s u|^2 + |\nabla_s\partial_s u|^2 + |\nabla_t\nabla_s v|^2 + \varepsilon^2|\nabla_s\nabla_s v|^2\Big)\\&\le CE^\varepsilon_{[s-1,s+1]}(u,v)\end{aligned} \tag{38}$$

for all s *and* t*. Here* $E^\varepsilon_I(u,v)$ *denotes the energy of* (u,v) *over the domain* $I\times S^1$*.*

Remark 6.2. Note that (38) implies the estimate

$$\|\partial_t u - v\|_{L^\infty} \le \varepsilon^2\sqrt{CE^\varepsilon(u,v)}$$

for every solution $(u,v):\mathbb{R}\times S^1\to TM$ of (11) that satisfies (32).

The proof of Theorem 6.1 has five steps. The first step is a bubbling argument and establishes a weak form of the required L^∞ estimate (with $\partial_s u$ replaced by $\varepsilon^2\partial_s u$ and $\nabla_s v$ replaced by $\varepsilon^3\nabla_s v$). The second step establishes an L^2-version of the estimate, namely an estimate for $\|\partial_s u(s,\cdot)\|_{L^2(S^1)} + \varepsilon\|\nabla_s v(s,\cdot)\|_{L^2(S^1)}$. The third step is an auxiliary result of the same type for the second derivatives. The fourth step establishes the L^∞ bound with $\nabla_s v$ replaced by $\varepsilon\nabla_s v$. The final step then proves the theorem in full.

Lemma 6.3. *Fix a constant* $c_0 > 0$ *and a perturbation* $\mathcal{V}:\mathcal{L}M\to\mathbb{R}$ *that satisfies* $(V0-V1)$*. Then the following holds.*

(i) *For every* $\delta>0$ *there is an* $\varepsilon_0>0$ *such that every solution* $(u,v):\mathbb{R}\times S^1\to M$ *of (11) and (32) with* $0<\varepsilon\le\varepsilon_0$ *satisfies the inequality*

$$\varepsilon^2\|\partial_s u\|_\infty + \varepsilon^3\|\nabla_s v\|_\infty \le \delta. \tag{39}$$

(ii) *For every* $\varepsilon_0>0$ *there is a constant* $c>0$ *such that every solution* $(u,v):\mathbb{R}\times S^1\to M$ *of (11) and (32) with* $\varepsilon_0\le\varepsilon\le 1$ *satisfies*

$$\|\partial_s u\|_\infty + \|\nabla_s v\|_\infty \le c.$$

Proof. We prove (i). Suppose, by contradiction, that the result is false. Then there is a sequence of solutions $(u_\nu,v_\nu):\mathbb{R}\times S^1\to M$ of (11) with $\varepsilon_\nu>0$ satisfying

$$E^{\varepsilon_\nu}(u_\nu,v_\nu)\le c_0,\qquad \sup_{s\in\mathbb{R}}\mathcal{A}_\mathcal{V}(u_\nu(s,\cdot),v_\nu(s,\cdot))\le c_0,\qquad \lim_{\nu\to\infty}\varepsilon_\nu = 0,$$

and

$$\varepsilon_\nu^2\|\partial_s u_\nu\|_\infty + \varepsilon_\nu^3\|\nabla_s v_\nu\|_\infty \ge 2\delta$$

for suitable constants $c_0>0$ and $\delta>0$. Since (u_ν,v_ν) has finite energy the functions $|\partial_s u_\nu(s,t)|$ and $|\nabla_s v_\nu(s,t)|$ converge to zero as $|s|$ tends to infinity.

Hence the function $|\partial_s u_\nu| + \varepsilon_\nu |\nabla_s v_\nu|$ takes on its maximum at some point $z_\nu = s_\nu + it_\nu$, i.e.

$$c_\nu := \sup_{\mathbb{R}\times S^1} \left(|\partial_s u_\nu| + \varepsilon_\nu \left|\nabla_s v_\nu\right|\right) = |\partial_s u_\nu(s_\nu, t_\nu)| + \varepsilon_\nu \left|\nabla_s v_\nu(s_\nu, t_\nu)\right|$$

and

$$\varepsilon_\nu^2 c_\nu \ge \delta. \tag{40}$$

Applying a time shift and using the periodicity in t we may assume without loss of generality that $s_\nu = 0$ and $0 \le t_\nu \le 1$.

Now consider the sequence

$$\tilde{w}_\nu = (\tilde{u}_\nu, \tilde{v}_\nu) : \mathbb{R}^2 \to TM$$

defined by

$$\tilde{u}_\nu(s,t) := u_\nu\left(\frac{s}{c_\nu}, t_\nu + \frac{t}{\varepsilon_\nu c_\nu}\right), \qquad \tilde{v}_\nu(s,t) := \varepsilon_\nu v_\nu\left(\frac{s}{c_\nu}, t_\nu + \frac{t}{\varepsilon_\nu c_\nu}\right).$$

This sequence satisfies the partial differential equation

$$\partial_s \tilde{u}_\nu - \nabla_t \tilde{v}_\nu = \frac{1}{c_\nu}\xi_\nu, \qquad \nabla_s \tilde{v}_\nu + \partial_t \tilde{u}_\nu = \frac{1}{{\varepsilon_\nu}^2 c_\nu}\tilde{v}_\nu, \tag{41}$$

where

$$\xi_\nu(s,t) := \operatorname{grad}\mathcal{V}(u_\nu(s/c_\nu, \cdot))(t_\nu + t/\varepsilon_\nu c_\nu) \in T_{\tilde{u}_\nu(s,t)}M.$$

By definition of c_ν we have

$$|\partial_s \tilde{u}_\nu(0, t_\nu)| + |\nabla_s \tilde{v}_\nu(0, t_\nu)| = 1 \tag{42}$$

and

$$|\partial_s \tilde{u}_\nu(s,t)| + |\nabla_s \tilde{v}_\nu(s,t)| \le 1$$

for all s and t. Since $|\tilde{v}_\nu|$ is uniformly bounded, by Theorem 5.1, and $|\xi_\nu|$ is uniformly bounded, by axiom $(V0)$, it then follows from (41) that $\tilde{u}_\nu$ and $\tilde{v}_\nu$ are uniformly bounded in C^1. Moreover, it follows from $(V1)$ that

$$|\nabla_t \xi_\nu(s,t)| \le \frac{C}{\varepsilon_\nu c_\nu}\Big(1 + |\partial_t u_\nu(s/c_\nu, t_\nu + t/\varepsilon_\nu c_\nu)|\Big) = C\left(\frac{1}{\varepsilon_\nu c_\nu} + |\partial_t \tilde{u}_\nu(s,t)|\right)$$

and

$$|\nabla_s \xi_\nu(s,t)| \le \frac{C}{c_\nu}\Big(|\partial_s u_\nu(s/c_\nu, t_\nu + t/\varepsilon_\nu c_\nu)| + \|\partial_s u_\nu(s/c_\nu, \cdot)\|_{L^1(S^1)}\Big) \le 2C.$$

Since the sequence $1/\varepsilon_\nu^2 c_\nu$ is bounded, by (40), it now follows from (41) that $\partial_s \tilde{u}_\nu - \nabla_t \tilde{v}_\nu$ and $\nabla_s \tilde{v}_\nu + \partial_t \tilde{u}_\nu$ are uniformly bounded in C^1, and hence in $W^{1,p}$ for any $p > 2$ and on any compact subset of $\mathbb{R}^2$. Since $\tilde{u}_\nu$ and $\tilde{v}_\nu$ are uniformly bounded in C^1, this implies that they are also uniformly bounded in $W^{2,p}$ over every compact subset of $\mathbb{R}^2$, by the standard elliptic bootstrapping techniques for J-holomorphic curves (see [16, Appendix B]). Hence, by the Arzéla–Ascoli theorem, there is a subsequence that converges in the C^1 topology to a solution $(\tilde{u}, \tilde{v})$ of the partial differential equation

$$\partial_s \tilde{u} - \nabla_t \tilde{v} = 0, \qquad \nabla_s \tilde{v} + \partial_t \tilde{u} = \lambda \tilde{v},$$

where $\lambda = \lim_{\nu\to\infty} 1/\varepsilon_\nu^2 c_\nu$. Since v_ν is uniformly bounded and $\varepsilon_\nu \to 0$ we have $\tilde{v} \equiv 0$ and so $\tilde{u}$ is constant. On the other hand it follows from (42) that $(\tilde{u}, \tilde{v})$ is nonconstant; contradiction. This proves (i).

The proof of (ii) is almost word by word the same, except that ε_ν no longer converges to zero while c_ν still diverges to infinity. So the limit $\tilde{w} = (\tilde{u}, \tilde{v}) : \mathbb{C} \to TM \cong T^*M$ is a J-holomorphic curve with finite energy and, by removal of singularities, extends to a nonconstant J-holomorphic sphere $\tilde{w} : S^2 \to T^*M$, which cannot exist since the symplectic form on T^*M is exact. □

The second step in the proof of Theorem 6.1 is to prove an integrated version of the estimate with $\nabla_s v$ replaced by $\varepsilon \nabla_s v$.

Lemma 6.4. *Fix a constant $c_0 > 0$ and a perturbation $\mathcal{V} : \mathcal{L}M \to \mathbb{R}$ that satisfies $(V0 - V2)$. Then there is a constant $C = C(c_0, \mathcal{V}) > 0$ such that the following holds. If $0 < \varepsilon \leq 1$ and $(u, v) : \mathbb{R} \times S^1 \to TM$ is a solution of (11) that satisfies (32) then, for every $s \in \mathbb{R}$,*

$$\begin{aligned} &\int_0^1 \left(|\partial_s u(s,t)|^2 + \varepsilon^2 |\nabla_s v(s,t)|^2 \right) dt \\ &+ \int_{s-1/4}^{s+1/4} \int_0^1 \left(|\nabla_t \partial_s u|^2 + \varepsilon^2 |\nabla_s \partial_s u|^2 + \varepsilon^2 |\nabla_t \nabla_s v|^2 + \varepsilon^4 |\nabla_s \nabla_s v|^2 \right) \\ &\leq C E^\varepsilon_{[s-1/2, s+1/2]}(u, v). \end{aligned} \tag{43}$$

Corollary 6.5. *Fix a constant $c_0 > 0$ and a perturbation $\mathcal{V} : \mathcal{L}M \to \mathbb{R}$ that satisfies $(V0 - V2)$. Then there is a constant $C = C(c_0, \mathcal{V}) > 0$ such that the following holds. If $0 < \varepsilon \leq 1$ and $(u, v) : \mathbb{R} \times S^1 \to TM$ is a solution of (11) that satisfies (32), then*

$$\int_{s-1/4}^{s+1/4} \int_0^1 |\nabla_s v|^2 \leq C E^\varepsilon_{[s-1/2, s+1/2]}(u, v)$$

for every $s \in \mathbb{R}$.

Proof. Since $\nabla_s v = \nabla_t \partial_s u + \varepsilon^2 \nabla_s \nabla_s v$ this estimate follows immediately from Lemma 6.4. □

Proof of Lemma 6.4. Define the functions $f, g : \mathbb{R} \times S^1 \to \mathbb{R}$ by

$$f := \frac{1}{2} \left(|\partial_s u|^2 + \varepsilon^2 |\nabla_s v|^2 \right)$$

and

$$g := \frac{1}{2} \left(|\nabla_t \partial_s u|^2 + \varepsilon^2 |\nabla_s \partial_s u|^2 + \varepsilon^2 |\nabla_t \nabla_s v|^2 + \varepsilon^4 |\nabla_s \nabla_s v|^2 \right),$$

and abbreviate

$$F(s) := \int_0^1 f(s,t)\,dt, \qquad G(s) := \int_0^1 g(s,t)\,dt.$$

Recall the definition of $L_\varepsilon := \varepsilon^2\partial_s^2 + \partial_t^2 - \partial_s$ and $\mathcal{L}_\varepsilon := \varepsilon^2\nabla_s\nabla_s + \nabla_t\nabla_t - \nabla_s$ in the proof of Proposition 5.2. Then

$$L_\varepsilon f = 2g + U + \varepsilon^2 V, \qquad U := \langle \partial_s u, \mathcal{L}_\varepsilon \partial_s u\rangle, \qquad V := \langle \nabla_s v, \mathcal{L}_\varepsilon \nabla_s v\rangle. \tag{44}$$

We shall prove that U and V satisfy the pointwise inequality

$$|U| + \varepsilon^2|V| \le \mu f + \frac{1}{2}\left(g + \|\partial_s u\|^2_{L^2(S^1)} + \varepsilon^4 \|\nabla_s\partial_s u\|^2_{L^2(S^1)}\right) \tag{45}$$

for a suitable constant $\mu > 0$. Inserting this inequality in (44) gives

$$L_\varepsilon f + \mu f + F \ge g + \frac{1}{2}(g - G).$$

Now integrate over the interval $0 \le t \le 1$ to obtain

$$\varepsilon^2 F'' - F' + (\mu+1)F \ge G.$$

With this understood the result follows from Lemma B.3 and Lemma B.6 via a covering argument.

To prove (45) we observe that, by (11),

$$\begin{aligned}
\mathcal{L}_\varepsilon \partial_s u &= \varepsilon^2\nabla_s\nabla_s\left(\nabla_t v + \operatorname{grad}\mathcal{V}(u)\right) + \nabla_t\nabla_s\left(v - \varepsilon^2\nabla_s v\right)\\
&\quad - \nabla_s\left(\nabla_t v + \operatorname{grad}\mathcal{V}(u)\right)\\
&= \varepsilon^2\left[\nabla_s\nabla_s, \nabla_t\right]v + \left[\nabla_t, \nabla_s\right]v - \nabla_s\operatorname{grad}\mathcal{V}(u) + \varepsilon^2\nabla_s\nabla_s\operatorname{grad}\mathcal{V}(u)\\
&= 2\varepsilon^2 R(\partial_s u, \partial_t u)\nabla_s v + \varepsilon^2\left(\nabla_{\partial_s u}R\right)(\partial_s u, \partial_t u)v - R(\partial_s u, \partial_t u)v\\
&\quad + \varepsilon^2 R(\nabla_s\partial_s u, \partial_t u)v + \varepsilon^2 R(\partial_s u, \nabla_s\partial_t u)v\\
&\quad - \nabla_s\operatorname{grad}\mathcal{V}(u) + \varepsilon^2\nabla_s\nabla_s\operatorname{grad}\mathcal{V}(u).
\end{aligned} \tag{46}$$

Now fix a sufficiently small constant $\delta > 0$ and choose $\varepsilon_0 > 0$ such that the assertion of Lemma 6.3 (i) holds. Choose $C > 0$ such that the assertion of Theorem 5.1 holds and assume $0 < \varepsilon \le \varepsilon_0 \le \delta/C$. Then, by Theorem 5.1 and Lemma 6.3, we have

$$\varepsilon^2\|\partial_s u\|_\infty \le \delta, \qquad \varepsilon^3\|\nabla_s v\|_\infty \le \delta, \qquad \|v\|_\infty \le C, \qquad \varepsilon\|\partial_t u\|_\infty \le 2\delta. \tag{47}$$

The last estimate uses the identity $\partial_t u = v - \varepsilon^2\nabla_s v$. Now take the pointwise inner product of (46) with $\partial_s u$ and estimate the resulting seven expressions separately. By (47) and $(V1)$, the terms four, five, and six are bounded by the right hand side of (45). For the last term we find, by $(V2)$,

$$\begin{aligned}
\varepsilon^2 \left|\langle \partial_s u, \nabla_s \nabla_s \mathrm{grad}\,\mathcal{V}(u)\rangle\right| &\le \varepsilon^2 C\,|\partial_s u| \left(|\nabla_s \partial_s u| + \|\nabla_s \partial_s u\|_{L^2(S^1)}\right) \\
&\quad + \varepsilon^2 C\,|\partial_s u| \left(|\partial_s u| + \|\partial_s u\|_{L^2(S^1)}\right)^2 \\
&\le \varepsilon^2 C\,|\partial_s u| \left(|\nabla_s \partial_s u| + \|\nabla_s \partial_s u\|_{L^2(S^1)}\right) \\
&\quad + 2C\delta\,|\partial_s u| \left(|\partial_s u| + \|\partial_s u\|_{L^2(S^1)}\right) \\
&\le \mu f + \frac{1}{8}\left(g + \|\partial_s u\|_{L^2(S^1)}^2 + \varepsilon^4 \|\nabla_s \partial_s u\|_{L^2(S^1)}^2\right).
\end{aligned}$$

For the first three terms on the right in (46) we argue as follows. Differentiate the equation $v = \partial_t u + \varepsilon^2 \nabla_s v$ covariantly with respect to s to obtain

$$\nabla_s v = \nabla_s \partial_t u + \varepsilon^2 \nabla_s \nabla_s v, \qquad \partial_t u = v - \varepsilon^2 \nabla_t \partial_s u - \varepsilon^4 \nabla_s \nabla_s v. \tag{48}$$

Now express half the first term on the right in (46) in the form

$$\begin{aligned}
&\varepsilon^2 \langle \partial_s u, R(\partial_s u, \partial_t u)\nabla_s v\rangle \\
&= \varepsilon^2 \langle \partial_s u, R(\partial_s u, v)\nabla_t \partial_s u\rangle + \varepsilon^4 \langle \partial_s u, R(\partial_s u, v)\nabla_s \nabla_s v\rangle \\
&\quad - \varepsilon^4 \langle \partial_s u, R(\partial_s u, \nabla_t \partial_s u)\nabla_t \partial_s u\rangle - \varepsilon^6 \langle \partial_s u, R(\partial_s u, \nabla_t \partial_s u)\nabla_s \nabla_s v\rangle \\
&\quad - \varepsilon^6 \langle \partial_s u, R(\partial_s u, \nabla_s \nabla_s v)\nabla_t \partial_s u\rangle - \varepsilon^8 \langle \partial_s u, R(\partial_s u, \nabla_s \nabla_s v)\nabla_s \nabla_s v\rangle.
\end{aligned}$$

Here we have replaced $\partial_t u$ and $\nabla_s v$ by the expressions in (48). In the first two terms we eliminate one of the factors $\partial_s u$ by using the inequality $\varepsilon^2 |\partial_s u| \le \delta$ and in the last four terms we eliminate both factors $\partial_s u$ by the same inequality. The next two terms in our expression for U have the form

$$\varepsilon^2 \langle \partial_s u, (\nabla_{\partial_s u} R)\,(\partial_s u, \partial_t u)v\rangle - \langle \partial_s u, R(\partial_s u, \partial_t u)v\rangle.$$

Replace $\partial_t u$ by the expression in (48) and elimate in each of the resulting summands one or two of the factors $\varepsilon^2 \partial_s u$ as above. This proves the required estimate for U and $0 < \varepsilon \le \varepsilon_0$.

To estimate V we observe that, by (11),

$$\begin{aligned}
\mathcal{L}_\varepsilon \nabla_s v &= \nabla_s \nabla_s (\varepsilon^2 \nabla_s v - v) + \nabla_t \nabla_s \nabla_t v + \nabla_t([\nabla_t, \nabla_s] v) \\
&= -\nabla_s \nabla_t \partial_s u + \nabla_t \nabla_s (\partial_s u - \mathrm{grad}\,\mathcal{V}(u)) - \nabla_t (R(\partial_s u, \partial_t u)v) \\
&= -R(\partial_s u, \partial_t u)\partial_s u - \nabla_t \nabla_s \mathrm{grad}\,\mathcal{V}(u) - \nabla_t (R(\partial_s u, \partial_t u)v) \\
&= -2R(\partial_s u, \partial_t u)\partial_s u + R(\partial_s u, \partial_t u)\mathrm{grad}\,\mathcal{V}(u) \\
&\quad - (\nabla_{\partial_t u} R)\,(\partial_s u, \partial_t u)v - R(\nabla_t \partial_s u, \partial_t u)v - R(\partial_s u, \nabla_t \partial_t u)v \\
&\quad - \nabla_t \nabla_s \mathrm{grad}\,\mathcal{V}(u).
\end{aligned} \tag{49}$$

The last step uses the identity $\nabla_t v = \partial_s u - \mathrm{grad}\,\mathcal{V}(u)$. Now take the pointwise inner product with $\varepsilon^2 \nabla_s v$. Then the first term has the same form as the one dicussed above. In the second and fourth term we estimate $\varepsilon|\partial_t u|$ by 2δ and we use $(V0)$. For the last term we find, by $(V2)$,

$$\begin{aligned}
\varepsilon^2 \left|\langle \nabla_s v, \nabla_t \nabla_s \mathrm{grad}\, \mathcal{V}(u)\rangle\right| &\le \varepsilon^2 C \left|\nabla_s v\right| \Big(\left|\nabla_t \partial_s u\right| + \left|\partial_s u\right| + \left\|\partial_s u\right\|_{L^2(S^1)} \\
&\qquad + \left|\partial_t u\right| \left(\left|\partial_s u\right| + \left\|\partial_s u\right\|_{L^2(S^1)} \right) \Big) \\
&\le \varepsilon^2 C \left|\nabla_s v\right| \Big(\left|\nabla_t \partial_s u\right| + \left|\partial_s u\right| + \left\|\partial_s u\right\|_{L^2(S^1)} \Big) \\
&\quad + 2\varepsilon C\delta \left|\nabla_s v\right| \Big(\left|\partial_s u\right| + \left\|\partial_s u\right\|_{L^2(S^1)} \Big) \\
&\le \mu f + \frac{1}{8} \Big(g + \left\|\partial_s u\right\|_{L^2(S^1)}^2 \Big) .
\end{aligned}$$

This leaves the terms three and five. In the third term we estimate $\varepsilon^2|\partial_t u|^2$ by $4\delta^2$ and use the identity

$$\nabla_s v = \nabla_s \partial_t u + \varepsilon^2 \nabla_s \nabla_s v$$

of (48). For term five we use the identity

$$\nabla_t \partial_t u = \nabla_t (v - \varepsilon^2 \nabla_s v) = \partial_s u - \mathrm{grad}\, \mathcal{V}(u) - \varepsilon^2 \nabla_t \nabla_s v$$

to obtain the expression

$$\begin{aligned}
&\varepsilon^2 \langle \nabla_s v, R(\partial_s u, \partial_s u - \mathrm{grad}\, \mathcal{V}(u) - \varepsilon^2 \nabla_t \nabla_s v) v\rangle \\
&= -\varepsilon^2 \langle \nabla_s v, R(\partial_s u, \mathrm{grad}\, \mathcal{V}(u)) v\rangle - \varepsilon^4 \langle \nabla_s v, R(\partial_s u, \nabla_t \nabla_s v) v\rangle.
\end{aligned}$$

In the last summand we use the estimate $\varepsilon^2|\partial_s u| \le \delta$. This proves (45) for $0 < \varepsilon \le \varepsilon_0$. For $\varepsilon_0 \le \varepsilon \le 1$ the estimate (45) follows immediately from (46), (49), and Lemma 6.3 (ii). □

The third step in the proof of Theorem 6.1 is to estimate the summand $\varepsilon^4 \left\|\nabla_s \partial_s u\right\|_{L^2(S^1)}^2$ in (45) in terms of the energy. This is the content of the following lemma.

Lemma 6.6. *Fix a constant $c_0 > 0$ and a perturbation $\mathcal{V} : \mathcal{L}M \to \mathbb{R}$ that satisfies $(V0 - V3)$. Then there is a constant $C = C(c_0, \mathcal{V}) > 0$ such that the following holds. If $0 < \varepsilon \le 1$ and $(u, v) : \mathbb{R} \times S^1 \to TM$ is a solution of (11) that satisfies (32) then, for every $s \in \mathbb{R}$,*

$$\begin{aligned}
&\int_0^1 \Big(\varepsilon^2 \left|\nabla_t \partial_s u\right|^2 + \varepsilon^4 \left|\nabla_s \partial_s u\right|^2 + \varepsilon^4 \left|\nabla_t \nabla_s v\right|^2 + \varepsilon^6 \left|\nabla_s \nabla_s v\right|^2 \Big) \\
&\le C E^{\varepsilon}_{[s-1/2, s+1/2]}(u, v).
\end{aligned} \tag{50}$$

Proof. Define f_1 and g_1 by

$$2f_1 := \left|\partial_s u\right|^2 + \varepsilon^2 \left|\nabla_s v\right|^2 + \varepsilon^2 \left|\nabla_t \partial_s u\right|^2 + \varepsilon^4 \left|\nabla_s \partial_s u\right|^2 + \varepsilon^4 \left|\nabla_t \nabla_s v\right|^2$$

and

$$
\begin{aligned}
2g_1 := &\left|\nabla_t\partial_s u\right|^2 + \varepsilon^2\left|\nabla_s\partial_s u\right|^2 + \varepsilon^2\left|\nabla_t\nabla_s v\right|^2 + \varepsilon^4\left|\nabla_s\nabla_s v\right|^2 \\
&+ \varepsilon^2\left|\nabla_t\nabla_t\partial_s u\right|^2 + \varepsilon^4\left|\nabla_s\nabla_t\partial_s u\right|^2 + \varepsilon^4\left|\nabla_t\nabla_s\partial_s u\right|^2 \\
&+ \varepsilon^6\left|\nabla_s\nabla_s\partial_s u\right|^2 + \varepsilon^4\left|\nabla_t\nabla_t\nabla_s v\right|^2 + \varepsilon^6\left|\nabla_s\nabla_t\nabla_s v\right|^2
\end{aligned}
$$

and abbreviate $F_1(s) := \int_0^1 f_1(s,t)\,dt$ and $G_1(s) := \int_0^1 g_1(s,t)\,dt$. Then

$$
L_\varepsilon f_1 = 2g_1 + U + \varepsilon^2 V + \varepsilon^2 U_t + \varepsilon^4 U_s + \varepsilon^4 V_t \tag{51}
$$

where $U := \langle\partial_s u, \mathcal{L}_\varepsilon\partial_s u\rangle$ and $V := \langle\nabla_s v, \mathcal{L}_\varepsilon\nabla_s v\rangle$ as in Lemma 6.4 and

$$
U_t := \langle\nabla_t\partial_s u, \mathcal{L}_\varepsilon\nabla_t\partial_s u\rangle, \qquad U_s := \langle\nabla_s\partial_s u, \mathcal{L}_\varepsilon\nabla_s\partial_s u\rangle,
$$

$$
V_t := \langle\nabla_t\nabla_s v, \mathcal{L}_\varepsilon\nabla_t\nabla_s v\rangle.
$$

We shall prove the estimate

$$
\begin{aligned}
&|U| + \varepsilon^2|V| + \varepsilon^2|U_t| + \varepsilon^4|U_s + V_t| \\
&\le \mu f_1 + \frac{1}{2}\left(g_1 + \|\partial_s u\|^2_{L^2(S^1)} + \varepsilon^4\|\nabla_s\partial_s u\|^2_{L^2(S^1)} + \varepsilon^8\|\nabla_s\nabla_s\partial_s u\|^2_{L^2(S^1)}\right) \\
&\le \mu f_1 + F_1 + g_1 + G_1
\end{aligned} \tag{52}
$$

for a suitable constant $\mu > 0$. By (51) and (52), $L_\varepsilon f_1 + \mu f_1 + F_1 \ge g_1 - G_1$. Integrating this inequality over the interval $0 \le t \le 1$ gives

$$
\varepsilon^2 F_1'' - F_1' + (\mu+1)F_1 \ge 0.
$$

Hence it follows from Lemma B.3 with $r := 1/5$ that

$$
F_1(s) \le c\int_{s-1/4}^{s+1/4} F_1(\sigma)\,d\sigma \le c\left(1 + \frac{C\varepsilon^2}{2}\right)E^\varepsilon_{[s-1/2,s+1/2]}(u,v).
$$

Here $c := 100c_2e^{(\mu+1)/25}$, where c_2 is the constant of Lemma B.3, and the second inequality follows from Lemma 6.4. Now use Lemma 6.4 again and the identity $\varepsilon^2\nabla_s\nabla_s v = \nabla_s v - \nabla_s\partial_t u$ to estimate the term $\varepsilon^6\left|\nabla_s\nabla_s v\right|^2$.

It remains to prove (52). For the terms $|U| + \varepsilon^2|V|$ the estimate was established in (45). To estimate the term $\varepsilon^2|U_t|$ write

$$
\begin{aligned}
\mathcal{L}_\varepsilon\nabla_t\partial_s u &= \nabla_t\mathcal{L}_\varepsilon\partial_s u + \varepsilon^2[\nabla_s\nabla_s, \nabla_t]\partial_s u - [\nabla_s, \nabla_t]\partial_s u \\
&= \nabla_t\mathcal{L}_\varepsilon\partial_s u + \varepsilon^2\nabla_s\left(R(\partial_s u, \partial_t u)\partial_s u\right)) \\
&\quad + \varepsilon^2 R(\partial_s u, \partial_t u)\nabla_s\partial_s u - R(\partial_s u, \partial_t u)\partial_s u \\
&= 2\varepsilon^2\nabla_t\left(R(\partial_s u, \partial_t u)\nabla_s v\right) + \varepsilon^2\nabla_t\left(\left(\nabla_{\partial_s u}R\right)(\partial_s u, \partial_t u)v\right) \\
&\quad - \nabla_t\left(R(\partial_s u, \partial_t u)v\right) \\
&\quad + \varepsilon^2\nabla_t\left(R(\nabla_s\partial_s u, \partial_t u)v\right) + \varepsilon^2\nabla_t\left(R(\partial_s u, \nabla_s\partial_t u)v\right) \\
&\quad - \nabla_t\nabla_s \operatorname{grad}\mathcal{V}(u) + \varepsilon^2\nabla_t\nabla_s\nabla_s\operatorname{grad}\mathcal{V}(u) \\
&\quad + \varepsilon^2\nabla_s\left(R(\partial_s u, \partial_t u)\partial_s u\right)) \\
&\quad + \varepsilon^2 R(\partial_s u, \partial_t u)\nabla_s\partial_s u - R(\partial_s u, \partial_t u)\partial_s u.
\end{aligned} \tag{53}
$$

The last equation follows from (46). Now take the pointwise inner product with $\varepsilon^2\nabla_t\partial_s u$. We begin by explaining how to estimate the first term. We encounter an expression of the form $\varepsilon^4\langle\nabla_t\partial_s u, (\nabla_{\partial_t u}R)(\partial_s u, \partial_t u)\nabla_s v\rangle$. Here we can use the identity

$$\nabla_s v = \nabla_s\partial_t u + \varepsilon^2\nabla_s\nabla_s v$$

to obtain an inequality

$$|\nabla_s v|\,|\nabla_t\partial_s u| \le 3g_1$$

By (47) we can estimate the product $\varepsilon^4\,|\partial_s u|\,|\partial_t u|^2$ by a small constant. Another expression we encounter is $\varepsilon^4\langle\nabla_t\partial_s u, R(\nabla_t\partial_s u, \partial_t u)\nabla_s v\rangle$; by (47), we have $\varepsilon^4\,|\partial_t u|\,|\nabla_s v| \le 2\delta^2$ and so the expression can be estimated by a small constant times g_1. Then we encounter the expression $\varepsilon^4\langle\nabla_t\partial_s u, R(\partial_s u, \nabla_t\partial_t u)\nabla_s v\rangle$; here we use the identity

$$\nabla_t\partial_t u = \nabla_t(v - \varepsilon^2\nabla_s v) = \partial_s u - \operatorname{grad}\mathcal{V}(u) - \varepsilon^2\nabla_t\nabla_s v;$$

the crucial observation is that the summand $\partial_s u$ can be dropped when inserting this formula in $R(\partial_s u, \nabla_t\partial_t u)$; in the summand $\varepsilon^4\langle\nabla_t\partial_s u, R(\partial_s u, \operatorname{grad}\mathcal{V}(u))\nabla_s v\rangle$ we use $(V0)$ and $\varepsilon^2\,|\partial_s u| \le \delta$; for the summand $\varepsilon^6\langle\nabla_t\partial_s u, R(\partial_s u, \nabla_t\nabla_s v)\nabla_s v\rangle$ we use $\varepsilon^5\,|\partial_s u|\,|\nabla_s v| \le \delta^2$ and $\varepsilon\,|\nabla_t\partial_s u|\,|\nabla_t\nabla_s v| \le Cg_1$. The last expression we encounter is $\varepsilon^4\langle\nabla_t\partial_s u, R(\partial_s u, \partial_t u)\nabla_t\nabla_s v\rangle$; here we use $\varepsilon^3\,|\partial_s u|\,|\partial_t u| \le 2\delta^2$, by (47), and again $\varepsilon\,|\nabla_t\partial_s u|\,|\nabla_t\nabla_s v| \le Cg_1$. This deals with the first term; the next two terms can be estimated by the same method.

In the fourth term we encounter the expression $\varepsilon^4\langle\nabla_t\partial_s u, R(\nabla_t\nabla_s\partial_s u, \partial_t u)v\rangle$; here we use $\varepsilon\,|\partial_t u| \le 2\delta$ and $\varepsilon^2\,|\nabla_t\partial_s u|\,|\nabla_t\nabla_s\partial_s u| \le Cg_1$. Another expression is $\varepsilon^4\langle\nabla_t\partial_s u, R(\nabla_s\partial_s u, \partial_t u)\nabla_t v\rangle$; here we use $\nabla_t v = \partial_s u - \operatorname{grad}\mathcal{V}(u)$ and the inequalities $\varepsilon^3\,|\partial_s u|\,|\partial_t u| \le 2\delta^2$ and $\varepsilon\,|\nabla_t\partial_s u|\,|\nabla_s\partial_s u| \le Cg_1$. A third expression is $\varepsilon^4\langle\nabla_t\partial_s u, R(\nabla_s\partial_s u, \nabla_t\partial_t u)v\rangle$; here we use the formula

$$\begin{aligned}\varepsilon^2\nabla_s\partial_s u + \nabla_t\partial_t u &= \varepsilon^2\nabla_s(\nabla_t v + \operatorname{grad}\mathcal{V}(u)) + \nabla_t(v - \varepsilon^2\nabla_s v)\\ &= \partial_s u + \varepsilon^2 R(\partial_s u, \partial_t u)v - \operatorname{grad}\mathcal{V}(u) + \varepsilon^2\nabla_s\operatorname{grad}\mathcal{V}(u);\end{aligned}\tag{54}$$

so the curvature term can be estimated by

$$|R(\nabla_s\partial_s u, \nabla_t\partial_t u)| \le C\,|\nabla_s\partial_s u|\left(1 + |\partial_s u| + \varepsilon^2\,|\partial_s u|\,|\partial_t u|\right). \tag{55}$$

Here we have used $(V0)$ and $(V1)$ and the fact that the term $\|\partial_s u\|_{L^1(S^1)}$ is uniformly bounded, by Lemma 6.4. This completes the discussion of the fourth term. The fifth term is similar, except that the cubic expression in the second derivatives vanishes. The last three terms can be disposed of similarly; the only new expression that appears is $\varepsilon^4\langle\nabla_t\partial_s u, (\nabla_{\partial_s u}R)(\partial_s u, \partial_t u)\partial_s u\rangle$; here we use $\partial_t u = v - \varepsilon^2\nabla_s v$ and the inequalities $\varepsilon^2\,|\partial_s u| \le \delta$ as well as $|\nabla_t\partial_s u|\,|\partial_s u| \le g_1 + f_1$ and $|\nabla_t\partial_s u|\,|\nabla_s v| \le 3g_1$.

This leaves the terms involving $\operatorname{grad}\mathcal{V}$. For $\varepsilon^2\langle\nabla_t\partial_s u, \nabla_t\nabla_s\operatorname{grad}\mathcal{V}(u)\rangle$ we use $(V2)$ and for $\varepsilon^4\langle\nabla_t\partial_s u, \nabla_t\nabla_s\nabla_s\operatorname{grad}\mathcal{V}(u)\rangle$ we use $(V3)$. Both terms can be estimated by $C\varepsilon(f_1 + g_1 + \|\partial_s u\|_{L^2(S^1)}^2 + \varepsilon^4\,\|\nabla_s\partial_s u\|_{L^2(S^1)}^2)$. This completes the estimate of $\varepsilon^2\,|U_t|$.

To estimate the term $\varepsilon^4 |U_s + V_t|$ write

$$
\begin{aligned}
\mathcal{L}_\varepsilon \nabla_s \partial_s u &= \nabla_s \mathcal{L}_\varepsilon \partial_s u + [\nabla_t \nabla_t, \nabla_s] \partial_s u \\
&= \nabla_s \mathcal{L}_\varepsilon \partial_s u - \nabla_t \left(R(\partial_s u, \partial_t u) \partial_s u \right) - R(\partial_s u, \partial_t u) \nabla_t \partial_s u \\
&= 2\varepsilon^2 \nabla_s \left(R(\partial_s u, \partial_t u) \nabla_s v \right) + \varepsilon^2 \nabla_s \left(\left(\nabla_{\partial_s u} R \right) (\partial_s u, \partial_t u) v \right) \\
&\quad - \nabla_s \left(R(\partial_s u, \partial_t u) v \right) \\
&\quad + \varepsilon^2 \nabla_s \left(R(\nabla_s \partial_s u, \partial_t u) v \right) + \varepsilon^2 \nabla_s \left(R(\partial_s u, \nabla_s \partial_t u) v \right) \\
&\quad - \nabla_s \nabla_s \mathrm{grad}\, \mathcal{V}(u) + \varepsilon^2 \nabla_s \nabla_s \nabla_s \mathrm{grad}\, \mathcal{V}(u) \\
&\quad - \nabla_t \left(R(\partial_s u, \partial_t u) \partial_s u \right) - R(\partial_s u, \partial_t u) \nabla_t \partial_s u
\end{aligned} \tag{56}
$$

(where the last equation follows from (46)) and

$$
\begin{aligned}
\mathcal{L}_\varepsilon \nabla_t \nabla_s v &= \nabla_t \mathcal{L}_\varepsilon \nabla_s v + \varepsilon^2 [\nabla_s \nabla_s, \nabla_t] \nabla_s v - [\nabla_s, \nabla_t] \nabla_s v \\
&= \nabla_t \mathcal{L}_\varepsilon \nabla_s v + \varepsilon^2 \nabla_s \left(R(\partial_s u, \partial_t u) \nabla_s v) \right) \\
&\quad + \varepsilon^2 R(\partial_s u, \partial_t u) \nabla_s \nabla_s v - R(\partial_s u, \partial_t u) \nabla_s v \\
&= -2 \nabla_t \left(R(\partial_s u, \partial_t u) \partial_s u \right) + \nabla_t \left(R(\partial_s u, \partial_t u) \mathrm{grad}\, \mathcal{V}(u) \right) \\
&\quad - \nabla_t \left(\left(\nabla_{\partial_t u} R \right) (\partial_s u, \partial_t u) v \right) \\
&\quad - \nabla_t \left(R(\nabla_t \partial_s u, \partial_t u) v \right) - \nabla_t \left(R(\partial_s u, \nabla_t \partial_t u) v \right) \\
&\quad - \nabla_t \nabla_t \nabla_s \mathrm{grad}\, \mathcal{V}(u) + \varepsilon^2 \nabla_s \left(R(\partial_s u, \partial_t u) \nabla_s v) \right) \\
&\quad + \varepsilon^2 R(\partial_s u, \partial_t u) \nabla_s \nabla_s v - R(\partial_s u, \partial_t u) \nabla_s v
\end{aligned} \tag{57}
$$

(where the last equation follows from (49)). The terms that require special attention are those involving $\mathrm{grad}\, \mathcal{V}$ and the cubic terms in the second derivatives. The cubic terms in the second derivatives are

$$
\begin{aligned}
U_{s0} &:= 2\varepsilon^6 \langle \nabla_s \partial_s u, R(\nabla_s \partial_s u, \nabla_s \partial_t u) v \rangle, \\
V_{t0} &:= 2\varepsilon^4 \langle \nabla_t \nabla_s v, R(\nabla_t \partial_t u, \nabla_t \partial_s u) v \rangle.
\end{aligned}
$$

Now insert

$$
\nabla_s \partial_s u = \nabla_s \left(\nabla_t v + \mathrm{grad}\, \mathcal{V}(u) \right), \qquad \nabla_t \partial_t u = \nabla_t \left(v - \varepsilon^2 \nabla_s v \right)
$$

into U_{s0} and V_{t0}, respectively. Then the only difficult remaining terms are the ones involving again three second derivatives. After replacing $\nabla_s \nabla_t v$ by $\nabla_t \nabla_s v + R(\partial_s u, \partial_t u) v$ we obtain

$$
\begin{aligned}
U_{s1} &:= 2\varepsilon^6 \langle \nabla_s \partial_s u, R(\nabla_t \nabla_s v, \nabla_s \partial_t u) v \rangle, \\
V_{t1} &:= -2\varepsilon^6 \langle \nabla_t \nabla_s v, R(\nabla_t \nabla_s v, \nabla_t \partial_s u) v \rangle.
\end{aligned}
$$

The sum is

$$
\begin{aligned}
U_{s1} + V_{t1} &= 2\varepsilon^6 \langle \nabla_s \partial_s u - \nabla_t \nabla_s v, R(\nabla_t \nabla_s v, \nabla_s \partial_t u) v \rangle \\
&= 2\varepsilon^6 \langle \nabla_s (\partial_s u - \nabla_t v) + R(\partial_s u, \partial_t u) v, R(\nabla_t \nabla_s v, \nabla_s \partial_t u) v \rangle \\
&= 2\varepsilon^6 \langle \nabla_s \mathrm{grad}\, \mathcal{V}(u) + R(\partial_s u, \partial_t u) v, R(\nabla_t \nabla_s v, \nabla_s \partial_t u) v \rangle
\end{aligned}
$$

and can be estimated in the required fashion. In particular we can use $(V1)$ and Lemma 6.3 to get a pointwise bound on $\varepsilon^2 \nabla_s \mathrm{grad}\, \mathcal{V}(u)$.

The terms involving $\mathrm{grad}\,\mathcal{V}$ can be estimated by

$$
\begin{aligned}
&\varepsilon^6 \left|\langle \nabla_s\partial_s u, \nabla_s\nabla_s\nabla_s \mathrm{grad}\,\mathcal{V}(u)\rangle\right| + \varepsilon^4 \left|\langle \nabla_s\partial_s u, \nabla_s\nabla_s \mathrm{grad}\,\mathcal{V}(u)\rangle\right| \\
&+ \varepsilon^4 \left|\langle \nabla_t\nabla_s v, R(\partial_s u, \partial_t u)\nabla_t \mathrm{grad}\,\mathcal{V}(u)\rangle\right| + \varepsilon^4 \left|\langle \nabla_t\nabla_s v, \nabla_t\nabla_t\nabla_s \mathrm{grad}\,\mathcal{V}(u)\rangle\right| \\
&\le C\varepsilon^2 \left|\nabla_s\partial_s u\right| \Big(\varepsilon^4 \left|\nabla_s\nabla_s\partial_s u\right| + \varepsilon^2 \left|\nabla_s\partial_s u\right| + \left|\partial_s u\right|\Big) \\
&\quad + C\varepsilon^2 \left|\nabla_s\partial_s u\right| \Big(\varepsilon^4 \left\|\nabla_s\nabla_s\partial_s u\right\|_{L^2(S^1)} + \varepsilon^2 \left\|\nabla_s\partial_s u\right\|_{L^2(S^1)} + \left\|\partial_s u\right\|_{L^2(S^1)}\Big) \\
&\quad + C\varepsilon^2 \left|\nabla_t\nabla_s v\right| \Big(\varepsilon^2 \left|\nabla_t\nabla_t\partial_s u\right| + \varepsilon \left|\nabla_t\partial_s u\right| + \left|\partial_s u\right| + \left\|\partial_s u\right\|_{L^2(S^1)}\Big) \\
&\quad + C\varepsilon^4 \left|\nabla_t\nabla_s v\right|^2 \\
&\le \mu f_1 + \frac{1}{8}\Big(g_1 + \left\|\partial_s u\right\|^2_{L^2(S^1)} + \varepsilon^4 \left\|\nabla_s\partial_s u\right\|^2_{L^2(S^1)} + \varepsilon^8 \left\|\nabla_s\nabla_s\partial_s u\right\|^2_{L^2(S^1)}\Big).
\end{aligned}
$$

Here the first inequality follows from $(V1-3)$; it also uses the identity

$$\nabla_t\partial_t u = \partial_s u - \nabla_s \mathrm{grad}\,\mathcal{V}(u) - \varepsilon^2\nabla_t\nabla_s v$$

and the fact that $\varepsilon^2 \left|\partial_s u\right|$ and $\varepsilon\left|\partial_t u\right|$ are uniformly bounded (Lemma 6.3). All the other summands appearing in our expression for $\varepsilon^4 \left|U_s + V_t\right|$ can be estimated by the same arguments as for $\varepsilon^2 \left|U_t\right|$. This implies (52) for small ε. For $\varepsilon_0 \le \varepsilon \le 1$ and μ sufficiently large the estimate (52) follows from (53), (56), (57) and Lemma 6.3 (ii). The same tricks as above are needed to deal with the cubic terms in the second derivatives but no care needs to be taken concerning the value of ε. This completes the proof of Lemma 6.6. □

The fourth step in the proof of Theorem 6.1 is to establish the L^∞ estimate with $\nabla_s v$ replaced by $\varepsilon\nabla_s v$.

Lemma 6.7. *Fix a constant $c_0 > 0$ and a perturbation $\mathcal{V} : \mathcal{L}M \to \mathbb{R}$ that satisfies $(V0 - V3)$. Then there is a constant $C = C(c_0, \mathcal{V}) > 0$ such that the following holds. If $0 < \varepsilon \le 1$ and $(u, v) : \mathbb{R}\times S^1 \to TM$ is a solution of (11) that satisfies (32), i.e. $E^\varepsilon(u, v) \le c_0$ and $\sup_{s\in\mathbb{R}} \mathcal{A}_\mathcal{V}(u(s,\cdot), v(s,\cdot)) \le c_0$, then*

$$\left|\partial_s u(s,t)\right|^2 + \varepsilon^2 \left|\nabla_s v(s,t)\right|^2 \le C E^\varepsilon_{[s-1,s+1]}(u, v) \le C c_0 \tag{58}$$

for all s and t.

Proof. Let f, g, F, G, U, V, and μ be as in the proof of Lemma 6.4. Choose a constant $C > 0$ such that the assertions of Lemmas 6.4 and 6.6 hold with this constant. Then, by (44) and (45), we have

$$
\begin{aligned}
L_\varepsilon f &= 2g + U + \varepsilon^2 V \\
&\ge -\mu f - \frac{1}{2}\Big(\left\|\partial_s u\right\|^2_{L^2(S^1)} + \varepsilon^4 \left\|\nabla_s\partial_s u\right\|^2_{L^2(S^1)}\Big) \\
&\ge -\mu f - C E^\varepsilon_{[s-1/2,s+1/2]}(u, v)
\end{aligned}
$$

for all $(s,t) \in \mathbb{R} \times S^1$. Let $s_0 \in \mathbb{R}$ and denote

$$a := \frac{C}{\mu} E^{\varepsilon}_{[s_0-1,s_0+1]}(u,v).$$

Then

$$L_\varepsilon(f+a) + \mu(f+a) \geq 0$$

for $s_0 - 1/2 \leq s \leq s_0 + 1/2$. Hence we may apply Lemma B.2 with $r = 1/3$ to the function $w(s,t) := f(s_0+s, t_0+t) + a$:

$$\begin{aligned}
f(s_0,t_0) &\leq 54c_2 e^{\mu/9} \int_{s_0-1/9-\varepsilon/3}^{s_0+\varepsilon/3} \int_0^1 (f(s,t)+a)\, dtds \\
&\leq 54c_2 e^{\mu/9} \int_{s_0-1/2}^{s_0+1/2} \int_0^1 \left(\frac{1}{2} |\partial_s u(s,t)|^2 + \frac{\varepsilon^2}{2} |\nabla_s v(s,t)|^2 + a \right) dtds \\
&\leq 54c_2 e^{\mu/9} \Big(E^{\varepsilon}_{[s_0-1,s_0+1]}(u,v) + a \Big) \\
&= 54c_2 e^{\mu/9} \left(1 + \frac{C}{\mu} \right) E^{\varepsilon}_{[s_0-1,s_0+1]}(u,v).
\end{aligned}$$

This proves the lemma. □

Proof of Theorem 6.1. Define f_2 and g_2 by

$$\begin{aligned}
2f_2 &:= |\partial_s u|^2 + |\nabla_s v|^2, \\
2g_2 &:= |\nabla_t \partial_s u|^2 + \varepsilon^2 |\nabla_s \partial_s u|^2 + |\nabla_t \nabla_s v|^2 + \varepsilon^2 |\nabla_s \nabla_s v|^2
\end{aligned}$$

and abbreviate $F_2(s) := \int_0^1 f_2(s,t)\, dt$ and $G_2(s) := \int_0^1 g_2(s,t)\, dt$. Then

$$L_\varepsilon f_2 = 2g_2 + U + V \tag{59}$$

where U and V are as in the proof of Lemma 6.4. These functions satisfy the estimate

$$|U| + |V| \leq \mu f_2 + \frac{1}{2} \Big(g_2 + \|\partial_s u\|^2_{L^2} + \varepsilon^4 \|\nabla_s \partial_s u\|^2_{L^2} \Big) \tag{60}$$

for a suitable constant $\mu > 0$; here $\|\cdot\|_{L^2}$ denotes the L^2-norm over the circle at time s. This follows from (46) and (49) via term by term inspection. (We use the fact that $|\partial_s u|$, $\varepsilon |\nabla_s v|$, and $|\partial_t u|$ are uniformly bounded, by Lemma 6.7.)

By Lemmas 6.4 and 6.6, we have

$$\int_0^1 \Big(|\partial_s u(s,t)|^2 + \varepsilon^4 |\nabla_s \partial_s u(s,t)|^2 \Big)\, dt \leq C E^{\varepsilon}_{[s-1/2,s+1/2]}(u,v)$$

for a suitable constant C and every $s \in \mathbb{R}$. Hence it follows from (59) and (60) that

$$L_\varepsilon f_2(s,t) \geq -\mu f_2(s,t) - C E^{\varepsilon}_{[s-1/2,s+1/2]}(u,v)$$

for all $(s,t)\in\mathbb{R}\times S^1$. Fix a number s_0 and abbreviate

$$a := \frac{C}{\mu}E^{\varepsilon}_{[s_0-1,s_0+1]}(u,v).$$

Then

$$L_\varepsilon(f_2+a)+\mu(f_2+a)\ge 0$$

for $s_0-1/2\le s\le s_0+1/2$. Hence we may apply Lemma B.2 with $r=1/3$ to the function $w(s,t):=f_2(s_0+s,t_0+t)+a$:

$$\begin{aligned}
f_2(s_0,t_0) &\le 54c_2e^{\mu/9}\int_{s_0-1/9-\varepsilon/3}^{s_0+\varepsilon/3}\int_0^1 (f_2(s,t)+a)\,dtds \\
&\le 54c_2e^{\mu/9}\int_{s_0-1/2}^{s_0+1/2}\int_0^1\left(\frac{1}{2}\left|\partial_s u(s,t)\right|^2+\frac{1}{2}\left|\nabla_s v(s,t)\right|^2+a\right)\,dtds \\
&\le c_3\Bigl(E^{\varepsilon}_{[s_0-1,s_0+1]}(u,v)+a\Bigr) \\
&= c_3\left(1+\frac{C}{\mu}\right)E^{\varepsilon}_{[s_0-1,s_0+1]}(u,v).
\end{aligned}$$

Here the third inequality, with a suitable constant c_3, follows from Corollary 6.5. This proves the pointwise estimate.

To prove the L^2-estimate integrate (59) and (60) over $0\le t\le 1$ to obtain

$$\varepsilon^2F_2''-F_2'+(\mu+1)F_2\ge G_2$$

for every $s\in\mathbb{R}$. Hence, by Lemma B.6 and a covering argument, we have

$$\int_{-1/2}^{1/2}G_2(s)\,ds\le c_4\int_{-3/4}^{3/4}F_2(s)\,ds$$

for every $s\in\mathbb{R}$ and a constant $c_4>0$ that depends only on R, r, and μ. Now it follows from Corollary 6.5 that

$$\int_{-3/4}^{3/4}F_2(s)\,ds\le c_5E^{\varepsilon}_{[s-1,s+1]}(u,v)$$

for every $s>0$ and some constant $c_5=c_5(c_0,\mathcal{V})>0$. Hence

$$\int_{-1/2}^{1/2}\int_0^1\Bigl(\left|\nabla_t\partial_s u\right|^2+\left|\nabla_t\nabla_s v\right|^2+\varepsilon^2\left|\nabla_s\nabla_s v\right|^2\Bigr)\,dtds\le 2c_4c_5E^{\varepsilon}_{[s-1,s+1]}(u,v).$$

The estimate for $\nabla_s\partial_s u$ now follows from the identity

$$\nabla_s\partial_s u=\nabla_s\nabla_t v+\nabla_s\mathrm{grad}\,\mathcal{V}(u)=\nabla_t\nabla_s v+R(\partial_s u,\partial_t u)v+\nabla_s\mathrm{grad}\,\mathcal{V}(u).$$

This proves Theorem 6.1. □

7 Estimates of the second derivatives

Theorem 7.1. *Fix a constant $c_0 > 0$ and a perturbation $\mathcal{V} : \mathcal{L}M \to \mathbb{R}$ that satisfies $(V0 - V4)$. Then there is a constant $C = C(c_0, \mathcal{V}) > 0$ such that the following holds. If $0 < \varepsilon \leq 1$ and $(u, v) : \mathbb{R} \times S^1 \to TM$ is a solution of (11) that satisfies (32) then*

$$\begin{aligned} &\|\nabla_t \partial_s u\|_{L^p([-T,T]\times S^1)} + \|\nabla_s \partial_s u\|_{L^p([-T,T]\times S^1)} \\ &+ \|\nabla_t \nabla_s v\|_{L^p([-T,T]\times S^1)} + \|\nabla_s \nabla_s v\|_{L^p([-T,T]\times S^1)} \\ &\leq c\sqrt{E^\varepsilon_{[-T-1,T+1]}(u,v)} \end{aligned} \tag{61}$$

for $T > 1$ and $2 \leq p \leq \infty$.

For $p = 2$ the estimate, with $\nabla_s \nabla_s v$ replaced by $\varepsilon \nabla_s \nabla_s v$, was established in Theorem 6.1. The strategy is to prove the estimate for $p = \infty$ and, as a byproduct, to get rid of the factor ε for $p = 2$ (see Corollary 7.3 below). The result for general p then follows by interpolation.

Lemma 7.2. *Fix a constant $c_0 > 0$ and a perturbation $\mathcal{V} : \mathcal{L}M \to \mathbb{R}$ that satisfies $(V0 - V3)$. Then there is a constant $C = C(c_0, \mathcal{V}) > 0$ such that the following holds. If $0 < \varepsilon \leq 1$ and $(u, v) : \mathbb{R} \times S^1 \to TM$ is a solution of (11) that satisfies (32), then*

$$\begin{aligned} &\int_0^1 \left(|\nabla_t \partial_s u(s,t)|^2 + |\nabla_s \partial_s u(s,t)|^2 + |\nabla_t \nabla_s v(s,t)|^2 + \varepsilon^2 |\nabla_s \nabla_s v(s,t)|^2 \right) dt \\ &+ \int_{s-1/4}^{s+1/4} \int_0^1 \left(|\nabla_t \nabla_t \partial_s u|^2 + |\nabla_t \nabla_s \partial_s u|^2 + \varepsilon^2 |\nabla_s \nabla_s \partial_s u|^2 \right) \\ &+ \int_{s-1/4}^{s+1/4} \int_0^1 \left(|\nabla_t \nabla_t \nabla_s v|^2 + \varepsilon^2 |\nabla_t \nabla_s \nabla_s v|^2 + \varepsilon^4 |\nabla_s \nabla_s \nabla_s v|^2 \right) \\ &\leq C E^\varepsilon_{[s-1/2,s+1/2]}(u,v) \end{aligned}$$

for every $s \in \mathbb{R}$.

Corollary 7.3. *Fix a constant $c_0 > 0$ and a perturbation $\mathcal{V} : \mathcal{L}M \to \mathbb{R}$ that satisfies $(V0 - V3)$. Then there is a constant $C = C(c_0, \mathcal{V}) > 0$ such that the following holds. If $0 < \varepsilon \leq 1$ and $(u, v) : \mathbb{R} \times S^1 \to TM$ is a solution of (11) that satisfies (32), then*

$$\int_{s-1/4}^{s+1/4} \int_0^1 |\nabla_s \nabla_s v|^2 \leq C E^\varepsilon_{[s-1/2,s+1/2]}(u,v)$$

for every $s \in \mathbb{R}$.

Proof. Since

$$\nabla_s \nabla_s v = \nabla_s \nabla_s \partial_t u + \varepsilon^2 \nabla_s \nabla_s \nabla_s v = R(\partial_s u, \partial_t u)\partial_s u + \nabla_t \nabla_s \partial_s u + \varepsilon^2 \nabla_s \nabla_s \nabla_s v$$

this estimate follows immediately from Lemma 7.2. □

Proof of Lemma 7.2. Define f_3 and g_3 by

$$2f_3 := |\partial_s u|^2 + |\nabla_s v|^2 + |\nabla_t \partial_s u|^2 + |\nabla_s \partial_s u|^2 + |\nabla_t \nabla_s v|^2 + \varepsilon^2 |\nabla_s \nabla_s v|^2$$

and

$$\begin{aligned}
2g_3 :=\ & |\nabla_t \partial_s u|^2 + \varepsilon^2 |\nabla_s \partial_s u|^2 + |\nabla_t \nabla_s v|^2 + \varepsilon^2 |\nabla_s \nabla_s v|^2 \\
& + |\nabla_t \nabla_t \partial_s u|^2 + \varepsilon^2 |\nabla_s \nabla_t \partial_s u|^2 + |\nabla_t \nabla_s \partial_s u|^2 + \varepsilon^2 |\nabla_s \nabla_s \partial_s u|^2 \\
& + |\nabla_t \nabla_t \nabla_s v|^2 + \varepsilon^2 |\nabla_s \nabla_t \nabla_s v|^2 + \varepsilon^2 |\nabla_t \nabla_s \nabla_s v|^2 + \varepsilon^4 |\nabla_s \nabla_s \nabla_s v|^2
\end{aligned}$$

and abbreviate

$$F_3(s) := \int_0^1 f_3(s,t)\, dt, \qquad G_3(s) := \int_0^1 g_3(s,t)\, dt.$$

Then

$$L_\varepsilon f_3 = 2g_3 + U + V + U_t + U_s + V_t + \varepsilon^2 V_s \tag{62}$$

where U, V, U_t, U_s, V_t are as in the proof of Lemma 6.6 and V_s is defined analogously. These functions satisfy the estimate

$$\begin{aligned}
& |U| + |V| + |U_t| + |U_s + V_t| + \varepsilon^2 |V_s| \\
& \le \mu f_3 + \frac{1}{2}\Big(g_3 + \|\partial_s u\|_{L^2}^2 + \|\nabla_s \partial_s u\|_{L^2}^2 + \varepsilon^4 \|\nabla_s \nabla_s \partial_s u\|_{L^2}^2 \Big) \\
& \le \mu f_3 + F_3 + \frac{1}{2}(g_3 + G_3)
\end{aligned} \tag{63}$$

for a suitable constant $\mu > 0$; here $\|\cdot\|_{L^2}$ denotes the L^2-norm over the circle at time s. For U and V this follows from (45) and (49) in the proof of Lemma 6.4. For U_t this follows from (53) and for $U_s + V_t$ from (56) and (57) by the same arguments as in the proof of Lemma 6.6. The improved estimate (63) follows by combining these arguments with Theorem 6.1. For V_s we use the formula

$$\begin{aligned}
\mathcal{L}_\varepsilon \nabla_s \nabla_s v &= \nabla_s \mathcal{L}_\varepsilon \nabla_s v + [\nabla_t \nabla_t, \nabla_s] \nabla_s v \\
&= \nabla_s \mathcal{L}_\varepsilon \nabla_s v \\
&\quad - \nabla_t \left(R(\partial_s u, \partial_t u) \nabla_s v \right) - R(\partial_s u, \partial_t u) \nabla_t \nabla_s v \\
&= -2\nabla_s \left(R(\partial_s u, \partial_t u) \partial_s u \right) + \nabla_s \left(R(\partial_s u, \partial_t u) \operatorname{grad} \mathcal{V}(u) \right) \\
&\quad - \nabla_s \left((\nabla_{\partial_t u} R)(\partial_s u, \partial_t u) v \right) \\
&\quad - \nabla_s \left(R(\nabla_t \partial_s u, \partial_t u) v \right) - \nabla_s \left(R(\partial_s u, \nabla_t \partial_t u) v \right) \\
&\quad - \nabla_s \nabla_t \nabla_s \operatorname{grad} \mathcal{V}(u) \\
&\quad - \nabla_t \left(R(\partial_s u, \partial_t u) \nabla_s v \right) - R(\partial_s u, \partial_t u) \nabla_t \nabla_s v.
\end{aligned} \tag{64}$$

(The last equation uses (49).) The desired estimate now follows from a term by term inspection; since all the first derivatives are uniformly bounded, by Theorem 6.1, we only need to examine the second and third derivatives; in particular, the cubic term $\varepsilon^2 \langle \nabla_s \nabla_s v, R(\nabla_s \partial_s u, \nabla_t \partial_t u) v \rangle$ can be estimated by $C\varepsilon^2 |\nabla_s \nabla_s v| \, |\nabla_s \partial_s u|$ (see (55) in the proof of Lemma 6.6).

It follows from (62) and (63) that

$$L_\varepsilon f_3 + \mu f_3 + F_3 \geq g_3 + \frac{1}{2}(g_3 - G_3).$$

Integrating this inequality over the interval $0 \leq t \leq 1$ gives

$$\varepsilon^2 F_3'' - F_3' + (\mu + 1) F_3 \geq G_3.$$

By Theorem 6.1 and Corollary 6.5 we have

$$\int_{s-1/4}^{s+1/4} F_3(s)\, ds \leq C E^\varepsilon_{[s-1/2, s+1/2]}(u, v)$$

for a suitable constant $C = C(c_0, \mathcal{V}) > 0$. Hence the estimate for the second derivatives follows from Lemma B.3 with $r := 1/5$. The estimate for the third derivatives follows from Lemma B.6 and a covering argument. □

Lemma 7.4. *Fix a constant $c_0 > 0$ and a perturbation $\mathcal{V} : \mathcal{L}M \to \mathbb{R}$ that satisfies $(V0 - V3)$. Then there is a constant $c = c(c_0, \mathcal{V}) > 0$ such that the following holds. If $0 < \varepsilon \leq 1$ and $(u, v) : \mathbb{R} \times S^1 \to TM$ is a solution of (11) that satisfies (32), then*

$$\begin{aligned} &\|\nabla_t \partial_t u\|_{L^\infty} + \varepsilon \|\nabla_s \partial_t u\|_{L^\infty} + \varepsilon^2 \|\nabla_s \partial_s u\|_{L^\infty} \\ &+ \varepsilon \|\nabla_t \nabla_t v\|_{L^\infty} + \varepsilon^2 \|\nabla_t \nabla_s v\|_{L^\infty} + \varepsilon^3 \|\nabla_s \nabla_s v\|_{L^\infty} \leq c. \end{aligned}$$

Proof. For every solution (u, v) of (11) define

$$\tilde{u}(s, t) := u(\varepsilon s, t), \qquad \tilde{v}(s, t) := \varepsilon v(\varepsilon s, t).$$

Then

$$\partial_s \tilde{u} - \nabla_t \tilde{v} = \varepsilon \mathrm{grad}\, \mathcal{V}(\tilde{u}), \qquad \nabla_s \tilde{v} + \partial_t \tilde{u} = \frac{\tilde{v}}{\varepsilon}. \tag{65}$$

By Theorem 6.1, Lemma 7.2, and $(V0 - V3)$, the function $\tilde{w} := (\tilde{u}, \tilde{v})$ and the vector field

$$\zeta(s, t) := (\varepsilon\, \mathrm{grad}\, \mathcal{V}(u(\varepsilon s, \cdot))(t), v(\varepsilon s, t))$$

along $\tilde{w}$ are both uniformly bounded in $W^{3,2}$ (under the assumption (32)); here we use the identities

$$\begin{aligned} \nabla_t \partial_t u &= \partial_s u - \mathrm{grad}\, \mathcal{V}(u) - \varepsilon^2 \nabla_t \nabla_s v, \\ \nabla_t \nabla_t \partial_t u &= \nabla_t \partial_s u - \nabla_t \mathrm{grad}\, \mathcal{V}(u) - \varepsilon^2 \nabla_t \nabla_t \nabla_s v, \\ \nabla_t \nabla_t v &= \nabla_s v - \varepsilon^2 \nabla_s \nabla_s v - \nabla_t \mathrm{grad}\, \mathcal{V}(u), \\ \nabla_t \nabla_t \nabla_t v &= \nabla_t \nabla_s v - \varepsilon^2 \nabla_t \nabla_s \nabla_s v - \nabla_t \nabla_t \mathrm{grad}\, \mathcal{V}(u). \end{aligned}$$

It follows that $\tilde{w}$ and ζ are both uniformly bounded in $W^{2,p}$ for any $p > 2$. Since

$$\partial_s \tilde{w} + J(\tilde{w}) \partial_t \tilde{w} = \zeta$$

it follows from [16, Proposition B.4.9] that $\tilde{u}$ and $\tilde{v}$ are uniformly bounded in $W^{3,p}$ and hence in C^2. This proves the lemma. □

Lemma 7.5. *Fix a constant $c_0 > 0$ and a perturbation $\mathcal{V} : \mathcal{L}M \to \mathbb{R}$ that satisfies $(V0 - V4)$. Then there is a constant $C = C(c_0, \mathcal{V}) > 0$ such that the following holds. If $0 < \varepsilon \leq 1$ and $(u, v) : \mathbb{R} \times S^1 \to TM$ is a solution of (11) that satisfies (32), then*

$$\int_0^1 \varepsilon^4 \left|\nabla_s \nabla_s \partial_s u(s,t)\right|^2 \, dt \leq C E^{\varepsilon}_{[s-1/2,s+1/2]}(u,v)$$

for every $s \in \mathbb{R}$.

Proof. Define f_4 and g_4 by

$$2f_4 := \left|\partial_s u\right|^2 + \left|\nabla_s v\right|^2 + \left|\nabla_t \partial_s u\right|^2 + \left|\nabla_t \nabla_t \partial_s u\right|^2,$$

$$\begin{aligned} 2g_4 := & \left|\nabla_t \partial_s u\right|^2 + \varepsilon^2 \left|\nabla_s \partial_s u\right|^2 + \left|\nabla_t \nabla_s v\right|^2 + \varepsilon^2 \left|\nabla_s \nabla_s v\right|^2 \\ & + \left|\nabla_t \nabla_t \partial_s u\right|^2 + \varepsilon^2 \left|\nabla_s \nabla_t \partial_s u\right|^2 + \left|\nabla_t \nabla_t \nabla_t \partial_s u\right|^2 + \varepsilon^2 \left|\nabla_s \nabla_t \nabla_t \partial_s u\right|^2, \end{aligned}$$

and abbreviate $F_4(s) := \int_0^1 f_4(s,t)\, dt$ and $G_4(s) := \int_0^1 g_4(s,t)\, dt$. Then

$$L_\varepsilon f_4 = 2g_4 + U + V + U_t + U_{tt}, \tag{66}$$

where U, V, U_t are as in Lemma 6.6 and $U_{tt} := \langle \nabla_t \nabla_t \partial_s u, \mathcal{L}_\varepsilon \nabla_t \nabla_t \partial_s u \rangle$. We shall prove that there is a constant $\mu > 0$ such that

$$|U| + |V| + |U_t| + |U_{tt}| \leq \mu f_4 + \frac{1}{2}\left(g_4 + \left\|\partial_s u\right\|^2_{L^2(S^1)} + \varepsilon^2 \left\|\nabla_s \partial_s u\right\|^2_{L^2(S^1)}\right) \tag{67}$$

It follows from (66) and (67) that

$$L_\varepsilon f_4 + \mu f_4 + F_4 \geq g_4 + \frac{1}{2}(g_4 - G_4).$$

Integrating this inequality over the interval $0 \leq t \leq 1$ gives

$$\varepsilon^2 F_4'' - F_4' + (\mu + 1) F_4 \geq 0.$$

By Theorem 6.1 and Lemma 7.2, we have

$$\int_{s-1/4}^{s+1/4} F_4(\sigma)\, d\sigma \leq c E^{\varepsilon}_{[s-1/2,s+1/2]}(u,v)$$

for a suitable constant $c = c(c_0, \mathcal{V})$. Hence, by Lemma B.3 with $r = 1/5$, there is a constant $C = C(c_0, \mathcal{V})$ such that $F_4(s) \leq C E^{\varepsilon}_{[s-1/2,s+1/2]}(u,v)$ for every $s \in \mathbb{R}$; this gives

$$\int_0^1 \left|\nabla_t \nabla_t \partial_s u(s,t)\right|^2 \, dt \leq C E^{\varepsilon}_{[s-1/2,s+1/2]}(u,v).$$

Now use (46) and

$$\varepsilon^2 \nabla_s \nabla_s \partial_s u = \mathcal{L}_\varepsilon \partial_s u - \nabla_t \nabla_t \partial_s u + \nabla_s \partial_s u$$

to get the required estimate for $\varepsilon^4 \left|\nabla_s \nabla_s \partial_s u\right|$.

For U and V the estimate (67) was established in the proof of Theorem 6.1; for U_t it follows from (53) via the arguments used in the proof of Lemma 6.6. For U_{tt} we use the identity

$$
\begin{aligned}
\mathcal{L}_\varepsilon \nabla_t \nabla_t \partial_s u &= \nabla_t \mathcal{L}_\varepsilon \nabla_t \partial_s u + \varepsilon^2 [\nabla_s \nabla_s, \nabla_t] \nabla_t \partial_s u - [\nabla_s, \nabla_t] \nabla_t \partial_s u \\
&= \nabla_t \mathcal{L}_\varepsilon \partial_s u + \varepsilon^2 \nabla_s \left(R(\partial_s u, \partial_t u) \nabla_t \partial_s u \right) \\
&\quad + \varepsilon^2 R(\partial_s u, \partial_t u) \nabla_s \nabla_t \partial_s u - R(\partial_s u, \partial_t u) \nabla_t \partial_s u \\
&= 2\varepsilon^2 \nabla_t \nabla_t \left(R(\partial_s u, \partial_t u) \nabla_s v \right) + \varepsilon^2 \nabla_t \nabla_t \left(\left(\nabla_{\partial_s u} R \right) (\partial_s u, \partial_t u) v \right) \\
&\quad - \nabla_t \nabla_t \left(R(\partial_s u, \partial_t u) v \right) \\
&\quad + \varepsilon^2 \nabla_t \nabla_t \left(R(\nabla_s \partial_s u, \partial_t u) v \right) + \varepsilon^2 \nabla_t \nabla_t \left(R(\partial_s u, \nabla_t \partial_s u) v \right) \\
&\quad - \nabla_t \nabla_t \nabla_s \mathrm{grad}\, \mathcal{V}(u) + \varepsilon^2 \nabla_t \nabla_t \nabla_s \nabla_s \mathrm{grad}\, \mathcal{V}(u) \\
&\quad + \varepsilon^2 \nabla_t \nabla_s \left(R(\partial_s u, \partial_t u) \partial_s u \right) \\
&\quad + \varepsilon^2 \nabla_t \left(R(\partial_s u, \partial_t u) \nabla_s \partial_s u \right) - \nabla_t \left(R(\partial_s u, \partial_t u) \partial_s u \right) \\
&\quad + \varepsilon^2 \nabla_s \left(R(\partial_s u, \partial_t u) \nabla_t \partial_s u \right) \\
&\quad + \varepsilon^2 R(\partial_s u, \partial_t u) \nabla_s \nabla_t \partial_s u - R(\partial_s u, \partial_t u) \nabla_t \partial_s u.
\end{aligned} \tag{68}
$$

Here the last equation follows from (53). To establish (67) we now use the pointwise estimates on the first derivatives in Theorem 6.1 and the pointwise estimates on the second derivatives in Lemma 7.4. The term by term analysis shows that all the second, third, and fourth order factors appear with the appropriate powers of ε. This proves (67) and the lemma. □

Proof of Theorem 7.1. For $p = 2$ the estimate (61) follows from Theorem 6.1 and Corollary 7.3. To prove it for $p = \infty$ define f_5 and g_5 by

$$
2f_5 := \left|\partial_s u\right|^2 + \left|\nabla_s v\right|^2 + \left|\nabla_t \partial_s u\right|^2 + \left|\nabla_s \partial_s u\right|^2 + \left|\nabla_t \nabla_s v\right|^2 + \left|\nabla_s \nabla_s v\right|^2
$$

and

$$
\begin{aligned}
2g_3 := &\left|\nabla_t \partial_s u\right|^2 + \varepsilon^2 \left|\nabla_s \partial_s u\right|^2 + \left|\nabla_t \nabla_s v\right|^2 + \varepsilon^2 \left|\nabla_s \nabla_s v\right|^2 \\
&+ \left|\nabla_t \nabla_t \partial_s u\right|^2 + \varepsilon^2 \left|\nabla_s \nabla_t \partial_s u\right|^2 + \left|\nabla_t \nabla_s \partial_s u\right|^2 + \varepsilon^2 \left|\nabla_s \nabla_s \partial_s u\right|^2 \\
&+ \left|\nabla_t \nabla_t \nabla_s v\right|^2 + \varepsilon^2 \left|\nabla_s \nabla_t \nabla_s v\right|^2 + \left|\nabla_t \nabla_s \nabla_s v\right|^2 + \varepsilon^2 \left|\nabla_s \nabla_s \nabla_s v\right|^2 .
\end{aligned}
$$

Then

$$
L_\varepsilon f_5 = 2g_5 + U + V + U_t + U_s + V_t + V_s \tag{69}
$$

where U, V, U_t, U_s, V_t, and V_s are as in Lemma 6.6. These functions satisfy the estimate

$$
\begin{aligned}
&|U| + |V| + |U_t| + |U_s + V_t| + |V_s| \\
&\le \mu f_5 + g_5 + \left\|\partial_s u\right\|_{L^2}^2 + \left\|\nabla_s \partial_s u\right\|_{L^2}^2 + \varepsilon^4 \left\|\nabla_s \nabla_s \partial_s u\right\|_{L^2}^2
\end{aligned} \tag{70}
$$

for all $(s,t) \in \mathbb{R} \times S^1$ and a suitable constant $\mu > 0$. To see this one argues as in the proof of Lemma 7.2 and notices that the factor ε^2 in front of $|V_s|$ is no longer needed. (It can now be dropped since, by Corollary 7.3, the L^2-norm of f_5 is controlled by the energy.)

By (69) and (70), we have

$$\begin{aligned} L_\varepsilon f_5 + \mu f_5 &\geq - \|\partial_s u\|^2_{L^2(S^1)} - \|\nabla_s \partial_s u\|^2_{L^2(S^1)} - \varepsilon^4 \|\nabla_s \nabla_s \partial_s u\|^2_{L^2(S^1)} \\ &\geq -C E^\varepsilon_{[s-1/2,s+1/2]}(u,v) \end{aligned}$$

for every $s \in \mathbb{R}$ and suitable positive constants μ and C. Here the last inequality follows from Lemmas 7.2 and 7.5. Let $s_0 \in \mathbb{R}$ and denote

$$a := \frac{C}{\mu} E^\varepsilon_{[s_0-1,s_0+1]}(u,v).$$

Then

$$L_\varepsilon(f_5 + a) + \mu(f_5 + a) \geq 0$$

for $s_0 - 1/2 \leq s \leq s_0 + 1/2$. Hence we may apply Lemma B.2 with $r = 1/3$ to the function $w(s,t) := f_6(s_0 + s, t_0 + t) + a$:

$$\begin{aligned} f_5(s_0,t_0) &\leq 54 c_2 e^{\mu/9} \int_{s_0-1/9-\varepsilon/3}^{s_0+\varepsilon/3} \int_0^1 \big(f_5(s,t) + a\big)\, dtds \\ &\leq 54 c_2 e^{\mu/9} \int_{s_0-1/2}^{s_0+1/2} \int_0^1 \big(f_5(s) + a\big)\, dtds \\ &\leq c_3 \Big(E^\varepsilon_{[s_0-1,s_0+1]}(u,v) + a\Big) \\ &= c_3 \left(1 + \frac{C}{\mu}\right) E^\varepsilon_{[s_0-1,s_0+1]}(u,v). \end{aligned}$$

Here the third inequality, with a suitable constant $c_3 = c_3(c_0, \mathcal{V}) > 0$, follows from Theorem 6.1 and Corollaries 6.5 and 7.3. This proves (61) for $p = \infty$. To prove the result for general p we apply the interpolation inequality

$$\|\xi\|_{L^p} \leq \|\xi\|_{L^\infty}^{1-2/p} \|\xi\|_{L^2}^{2/p}$$

to the terms on the left hand side of the estimate and use the results for $p = 2$ and $p = \infty$. This proves the theorem. □

8 Uniform exponential decay

Theorem 8.1. *Fix a perturbation $\mathcal{V} : \mathcal{L}M \to \mathbb{R}$ that satisfies $(V0 - V3)$. Suppose $\mathcal{S}_\mathcal{V}$ is Morse and let $a \in \mathbb{R}$ be a regular value of $\mathcal{S}_\mathcal{V}$. Then there exist positive constants δ, c, ρ such that the following holds. If $x^\pm \in \mathcal{P}^a(\mathcal{V})$, $0 < \varepsilon \leq 1$, $T_0 > 0$, and $(u,v) \in \mathcal{M}^\varepsilon(x^-, x^+; \mathcal{V})$ satisfies*

$$E^\varepsilon_{\mathbb{R}\setminus[-T_0,T_0]}(u,v) < \delta, \tag{71}$$

then

$$E^\varepsilon_{\mathbb{R}\setminus[-T,T]}(u,v) \leq c e^{-\rho(T-T_0)} E^\varepsilon_{\mathbb{R}\setminus[-T_0,T_0]}(u,v)$$

for every $T \geq T_0 + 1$.

Corollary 8.2. *Fix a perturbation $\mathcal{V} : \mathcal{L}M \to \mathbb{R}$ that satisfies $(V0 - V3)$. Suppose $\mathcal{S}_{\mathcal{V}}$ is Morse and let $x^{\pm} \in \mathcal{P}(\mathcal{V})$. Then there exist positive constants δ, c, ρ such that the following holds. If $0 < \varepsilon \le 1$, $T_0 > 0$, and $(u, v) \in \mathcal{M}^{\varepsilon}(x^-, x^+; \mathcal{V})$ satisfies (71) then*

$$|\partial_s u(s,t)|^2 + |\nabla_s v(s,t)|^2 \le c e^{-\rho|s|} E^{\varepsilon}_{\mathbb{R}\setminus[-T_0,T_0]}(u,v) \tag{72}$$

for every $|s| \ge T_0 + 2$.

Proof. Theorem 6.1 and Theorem 8.1. $\square$

The proof of Theorem 8.1 is based on the following two lemmas.

Lemma 8.3 (The Hessian). *Fix a perturbation $\mathcal{V} : \mathcal{L}M \to \mathbb{R}$ that satisfies $(V0 - V2)$. Suppose $\mathcal{S}_{\mathcal{V}}$ is Morse and fix $a \in \mathbb{R}$. Then there are positive constants δ_0 and c such that the following is true. If $x_0 \in \mathcal{P}^a(\mathcal{V})$ and $(x, y) \in C^\infty(S^1, TM)$ satisfy*

$$x = \exp_{x_0}(\xi_0), \qquad y = \Phi(x_0, \xi_0)(\partial_t x_0 + \eta_0), \qquad \|\xi_0\|_{W^{1,2}} + \|\eta_0\|_\infty \le \delta_0,$$

then

$$\begin{aligned} &\|\xi\|_2 + \|\nabla_t \xi\|_2 + \|\eta\|_2 + \|\nabla_t \eta\|_2 \\ &\le c\left(\|\nabla_t \eta + R(\xi, \partial_t x)y + \mathcal{H}_{\mathcal{V}}(x)\xi\|_2 + \|\nabla_t \xi - \eta\|_2\right) \end{aligned}$$

for all $\xi, \eta \in \Omega^0(S^1, x^*TM)$.

Proof. The operator

$$A^{\varepsilon}(x,y)(\xi,\eta) := (-\nabla_t \eta - R(\xi, \partial_t x)y - \mathcal{H}_{\mathcal{V}}(x)\xi, \nabla_t \xi - \eta)$$

on $L^2(S^1, x^*TM \oplus x^*TM)$ with dense domain $W^{1,2}(S^1, x^*TM \oplus x^*TM)$ is self-adjoint if $y = \partial_t x$. In the case $(x, y) = (x_0, \partial_t x_0)$ it is bijective, because $\mathcal{S}_{\mathcal{V}}$ is Morse. Hence the result is a consequence of the open mapping theorem. Since bijectivity is preserved under small perturbations (with respect to the operator norm), the result for general pairs (x, y) follows from continuous dependence of the operator family on the pair (x, y) with respect to the $W^{1,2}$-topology on x and the L^∞-topology on y. The set $\mathcal{P}^a(\mathcal{V})$ is finite, because $\mathcal{S}_{\mathcal{V}}$ is Morse (see [28]). Hence we may choose the same constants δ_0 and c for all $x_0 \in \mathcal{P}^a(\mathcal{V})$. $\square$

Lemma 8.4. *Fix a perturbation $\mathcal{V} : \mathcal{L}M \to \mathbb{R}$ that satisfies $(V0)$. Suppose $\mathcal{S}_{\mathcal{V}}$ is Morse and let $a \in \mathbb{R}$ be a regular value of $\mathcal{S}_{\mathcal{V}}$. Then, for every $\delta_0 > 0$, there is a constant $\delta_1 > 0$ such that the following is true. Let $(x, y) : S^1 \to TM$ be a smooth loop such that*

$$\mathcal{A}_{\mathcal{V}}(x, gy) \le a, \qquad \|\nabla_t y + \operatorname{grad} \mathcal{V}(x)\|_\infty + \|\partial_t x - y\|_\infty < \delta_1$$

(where the isomorphism $g : TM \to T^*M$ is induced by the metric). Then there is a periodic orbit $x_0 \in \mathcal{P}^a(\mathcal{V})$ and a pair of vector fields $\xi_0, \eta_0 \in \Omega^0(S^1, {x_0}^*TM)$ such that

$$x = \exp_{x_0}(\xi_0), \qquad y = \Phi(x_0, \xi_0)(\partial_t x_0 + \eta_0),$$

and

$$\|\xi_0\|_\infty + \|\nabla_t \xi_0\|_\infty + \|\eta_0\|_\infty + \|\nabla_t \eta_0\|_\infty \le \delta_0.$$

Proof. First note that

$$\begin{aligned}\frac{1}{2}\int_0^1 |y(t)|^2 &= \mathcal{A}_{\mathcal{V}}(x,gy)+\mathcal{V}(x)-\int_0^t \langle y(t),\dot{x}(t)-y(t)\rangle\,dt \\ &\le a+C+\int_0^1\left(\frac{1}{4}|y(t)|^2+|\dot{x}(t)-y(t)|^2\right)dt,\end{aligned}$$

where C is the constant in $(V0)$. Hence, assuming $\delta_1 \le 1$, we have

$$\|y\|_2^2 \le 4\,(a+C+1).$$

Now

$$\begin{aligned}\left|\frac{d}{dt}|y|^2\right| &= \left|2\langle y,\nabla_t y+\operatorname{grad}\mathcal{V}(x)\rangle-2\langle y,\operatorname{grad}\mathcal{V}(x)\rangle\right| \\ &\le 2\,(\delta_1+C)\,|y| \le (C+1)^2+|y|^2.\end{aligned}$$

Integrate this inequality to obtain

$$|y(t_1)|^2-|y(t_0)|^2 \le (C+1)^2+\|y\|_2^2$$

for $t_0,t_1\in[0,1]$. Integrating again over the interval $0\le t_0\le 1$ gives

$$\|y\|_\infty \le \sqrt{(C+1)^2+2\,\|y\|_2^2} \le c \tag{73}$$

where $c^2 := (C+1)^2+8\,(a+C+1)$.

Now suppose that the assertion is wrong. Then there is a $\delta_0>0$ and a sequence of smooth loops $(x_\nu,y_\nu):S^1\to TM$ satisfying

$$\mathcal{A}_{\mathcal{V}}(x_\nu,gy_\nu)\le a, \qquad \lim_{\nu\to\infty}\left(\|\nabla_t y_\nu+\operatorname{grad}\mathcal{V}(x_\nu)\|_\infty+\|\partial_t x_\nu-y_\nu\|_\infty\right)=0,$$

but not the conclusion of the lemma for the given constant δ_0. By (73), we have $\sup_\nu\|y_\nu\|_\infty<\infty$. Hence $\sup_\nu\|\partial_t x_\nu\|_\infty<\infty$ and also $\sup_\nu\|\nabla_t y_\nu\|_\infty<\infty$. Hence, by the Arzela–Ascoli theorem, there exists a subsequence, still denoted by (x_ν,y_ν), that converges in the C^0-topology. Our assumptions guarantee that this subsequence actually converges in the C^1-topology. Let $(x_0,y_0):S^1\to TM$ be the limit. Then $\partial_t x_0=y_0$ and $\nabla_t y_0+\operatorname{grad}\mathcal{V}(x_0)=0$. Hence $x_0\in\mathcal{P}^a(\mathcal{V})$ and (x_ν,y_ν) converges to $(x_0,\partial_t x_0)$ in the C^1-topology. This contradicts our assumption on the sequence (x_ν,y_ν) and hence proves the lemma. □

Proof of Theorem 8.1. To begin with note that $\mathcal{S}_{\mathcal{V}}(x)\ge -C_0$ for every $x\in\mathcal{P}(\mathcal{V})$, where C_0 is the constant in $(V0)$. Hence, with $c_0:=a+C_0$, we have

$$x^\pm\in\mathcal{P}^a(\mathcal{V}) \qquad\Longrightarrow\qquad \mathcal{S}_{\mathcal{V}}(x^-)\le c_0,\quad \mathcal{S}_{\mathcal{V}}(x^-)-\mathcal{S}_{\mathcal{V}}(x^+)\le c_0.$$

Let $C>0$ be the constant of Theorem 6.1 with this choice of c_0. Let δ_0 and c be the constants of Lemma 8.3 and $\delta_1>0$ the constant of Lemma 8.4 associated to a and δ_0. Then choose $\delta>0$ such that $\sqrt{C\delta}\le\delta_1$. Below we will shrink the constants δ_1 and δ further if necessary.

In the remainder of the proof we will sometimes use the notation $u_s(t) := u(s,t)$ and $v_s(t) := v(s,t)$. Moreover, $\|\cdot\|$ will always denote the L^2 norm on S^1 and $\|\cdot\|_\infty$ the L^∞ norm on S^1.

Now let $x^\pm \in \mathcal{P}^a(\mathcal{V})$, $0 < \varepsilon \le 1$, and $T_0 > 0$, and suppose $(u,v) \in \mathcal{M}^\varepsilon(x^-, x^+; \mathcal{V})$ satisfies (71). Then, by Theorem 6.1, we have

$$\|\partial_s u_s\|_\infty + \|\nabla_s v_s\|_\infty \le \sqrt{C E^\varepsilon_{[s-1,s+1]}(u,v)} \le \sqrt{C\delta} \le \delta_1 \tag{74}$$

for $|s| \ge T_0 + 1$. Hence, by Lemma 8.4, we know that, for every $s \in \mathbb{R}$ with $|s| \ge T_0 + 1$, there is a periodic orbit $x_s \in \mathcal{P}^a(\mathcal{V})$ such that the C^1-distance between (u_s, v_s) and $(x_s, \partial_t x_s)$ is bounded by δ_0. Hence we can apply Lemma 8.3 to the pair (u_s, v_s) and the vector fields $(\partial_s u_s, \nabla_s v_s)$ for $|s| \ge T_0 + 1$. Since

$$\nabla_t \nabla_s v + R(\partial_s u, \partial_t u)v + \mathcal{H}_\mathcal{V}(u)\partial_s u = \nabla_s \partial_s u, \qquad \nabla_t \partial_s u - \nabla_s v = -\varepsilon^2 \nabla_s \nabla_s v,$$

we obtain from Lemma 8.3 that

$$\begin{aligned} &\|\partial_s u_s\|^2 + \|\nabla_t \partial_s u_s\|^2 + \|\nabla_s v_s\|^2 + \|\nabla_t \nabla_s v_s\|^2 \\ &\le c\left(\|\nabla_s \partial_s u_s\|^2 + \varepsilon^4 \|\nabla_s \nabla_s v_s\|^2\right). \end{aligned} \tag{75}$$

for $|s| \ge T_0 + 1$.

Define the function $F : \mathbb{R} \to [0,\infty)$ by

$$F(s) := \frac{1}{2}\int_0^1 \left(|\partial_s u(s,t)|^2 + \varepsilon^2 |\nabla_s v(s,t)|^2\right) dt.$$

We shall prove that

$$F''(s) \ge \frac{1}{c} F(s) \tag{76}$$

for $|s| \ge T_0 + 1$. The proof of (76) is based on the identity

$$\begin{aligned} F''(s) &= 2\|\nabla_s \partial_s u\|^2 + 2\varepsilon^2 \|\nabla_s \nabla_s v\|^2 \\ &\quad + \langle \partial_s u, \nabla_s \nabla_s \operatorname{grad} \mathcal{V}(u) - \mathcal{H}_\mathcal{V}(u) \nabla_s \partial_s u \rangle \\ &\quad + \langle \partial_s u, 3R(\partial_s u, \partial_t u)\nabla_s v \rangle + \langle \partial_s u, (\nabla_{\partial_s u} R)(\partial_s u, \partial_t u) v \rangle \\ &\quad + \langle \partial_s u, R(\partial_s u, \nabla_s v) v \rangle - \varepsilon^2 \langle \partial_s u, R(\partial_s u, \nabla_s \nabla_s v) v \rangle \\ &\quad + \varepsilon^2 \langle \partial_s u, R(\nabla_s v, v) \nabla_s \partial_s u \rangle \end{aligned} \tag{77}$$

which will be proved below. Here all norms and inner products are understood in $L^2(S^1, u_s^* TM)$ and we have dropped the subscript s for u_s and v_s. The L^∞ norms of v and $\partial_t u = v - \varepsilon^2 \nabla_s v$ are uniformly bounded, by Theorems 5.1 and 6.1. Hence there is a constant $c' > 0$ such that

$$\begin{aligned} F''(s) &\ge 2\|\nabla_s \partial_s u_s\|^2 + 2\varepsilon^2 \|\nabla_s \nabla_s v_s\|^2 \\ &\quad - c' \|\partial_s u_s\|_\infty \left(\|\partial_s u_s\|^2 + \|\partial_s u_s\| \|\nabla_s v_s\|\right) \\ &\quad - c'\varepsilon^2 \|\partial_s u_s\|_\infty \left(\|\partial_s u_s\| \|\nabla_s \nabla_s v\| + \|\nabla_s v_s\| \|\nabla_s \partial_s u_s\|\right) \\ &\ge \|\nabla_s \partial_s u_s\|^2 + \varepsilon^2 \|\nabla_s \nabla_s v_s\|^2. \end{aligned}$$

Here the first inequality uses $(V2)$. To understand the last step note that, by (74), we have $\|\partial_s u_s\|_\infty \leq \sqrt{C\delta}$ and so the inequality follows from (75), provided that $\delta > 0$ is sufficiently small. Now use (75) again to obtain (76).

Thus we have proved that $F''(s) \geq \rho^2 F(s)$ for $|s| \geq T_0 + 1$, where $\rho := c^{-1/2}$. Since $F(s)$ does not diverge to infinity as $|s| \to \infty$ it follows by standard arguments (see for example [5, 22]) that $F(s) \leq e^{-\rho(s-T_0-1)} F(T_0+1)$ for $s \geq T_0 + 1$ and similarly for $s \leq -T_0 - 1$.

It remains to prove (77). By direct computation,

$$F''(s) = \|\nabla_s \partial_s u\|^2 + \varepsilon^2 \|\nabla_s \nabla_s v\|^2 + G(s) + H(s),$$

where

$$\begin{aligned}
G(s) &:= \langle \partial_s u, \nabla_s \nabla_s \partial_s u \rangle \\
&= \langle \partial_s u, \nabla_s \nabla_s (\nabla_t v + \operatorname{grad} \mathcal{V}(u)) \rangle \\
&= \langle \partial_s u, [\nabla_s \nabla_s, \nabla_t] v + \nabla_s \nabla_s \operatorname{grad} \mathcal{V}(u) + \nabla_t \nabla_s \nabla_s v \rangle \\
&= \langle \partial_s u, [\nabla_s \nabla_s, \nabla_t] v + \nabla_s \nabla_s \operatorname{grad} \mathcal{V}(u) \rangle - \langle \nabla_s \partial_t u, \nabla_s \nabla_s v \rangle \\
&= \langle \partial_s u, \nabla_s [\nabla_s, \nabla_t] v + [\nabla_s, \nabla_t] \nabla_s v + \nabla_s \nabla_s \operatorname{grad} \mathcal{V}(u) \rangle \\
&\quad - \left\langle \nabla_s (v - \varepsilon^2 \nabla_s v), \nabla_s \nabla_s v \right\rangle, \\
H(s) &:= \varepsilon^2 \langle \nabla_s v, \nabla_s \nabla_s \nabla_s v \rangle \\
&= \langle \nabla_s v, \nabla_s \nabla_s (v - \partial_t u) \rangle \\
&= \langle \nabla_s v, \nabla_s \nabla_s v - [\nabla_s, \nabla_t] \partial_s u - \nabla_t \nabla_s \partial_s u \rangle \\
&= \langle \nabla_s v, \nabla_s \nabla_s v - [\nabla_s, \nabla_t] \partial_s u \rangle + \langle \nabla_t \nabla_s v, \nabla_s \partial_s u \rangle \\
&= \langle \nabla_s v, \nabla_s \nabla_s v - [\nabla_s, \nabla_t] \partial_s u \rangle + \langle [\nabla_t, \nabla_s] v + \nabla_s (\partial_s u - \operatorname{grad} \mathcal{V}(u)), \nabla_s \partial_s u \rangle .
\end{aligned}$$

Here all inner products are in $L^2(S^1, u_s^* TM)$; in each formula the fourth step uses integration by parts. The sum is

$$\begin{aligned}
G(s) + H(s) = \|\nabla_s \partial_s u\|^2 &+ \varepsilon^2 \|\nabla_s \nabla_s v\|^2 \\
&+ \langle \partial_s u, \nabla_s \nabla_s \operatorname{grad} \mathcal{V}(u) \rangle - \langle \nabla_s \operatorname{grad} \mathcal{V}(u), \nabla_s \partial_s u \rangle \\
&+ \langle \partial_s u, 3R(\partial_s u, \partial_t u) \nabla_s v \rangle + \langle \partial_s u, (\nabla_{\partial_s u} R)(\partial_s u, \partial_t u) v \rangle \\
&+ \langle \partial_s u, R(\partial_s u, \nabla_s \partial_t u) v \rangle \\
&+ \langle \partial_s u, R(\nabla_s \partial_s u, \partial_t u) v \rangle - \langle R(\partial_s u, \partial_t u) v, \nabla_s \partial_s u \rangle.
\end{aligned}$$

To obtain (77) replace $\nabla_s \partial_t u$ by $\nabla_s v - \varepsilon^2 \nabla_s \nabla_s v$. Moreover, by the first Bianchi identity, the last two terms can be expressed in the form

$$\begin{aligned}
&\langle \partial_s u, R(\nabla_s \partial_s u, \partial_t u) v \rangle - \langle R(\partial_s u, \partial_t u) v, \nabla_s \partial_s u \rangle \\
&= \langle \partial_s u, R(\nabla_s \partial_s u, \partial_t u) v \rangle + \langle \partial_s u, R(v, \nabla_s \partial_s u) \partial_t u \rangle \\
&= -\langle \partial_s u, R(\partial_t u, v) \nabla_s \partial_s u \rangle \\
&= \langle \partial_s u, R(v - \partial_t u, v) \nabla_s \partial_s u \rangle \\
&= \varepsilon^2 \langle \partial_s u, R(\nabla_s v, v) \nabla_s \partial_s u \rangle
\end{aligned}$$

This proves (77) and the theorem. □

9 Time shift

The next theorem establishes local surjectivity for the map $\mathcal{T}^\varepsilon$ constructed in Definition 4.4. The idea is to prove that, after a suitable time shift, the pair $\zeta = (\xi, \eta)$ with $u^\varepsilon = \exp_u(\xi)$ and $v^\varepsilon = \Phi(u, \xi)(\partial_t u + \eta)$ satisfies the hypothesis $\zeta \in \mathrm{im}\,(\mathcal{D}_u^\varepsilon)^*$ of Theorem 4.3. The neighbourhood, in which the next theorem establishes surjectivity, depends on ε.

Theorem 9.1. *Fix a perturbation $\mathcal{V} : \mathcal{L}M \to \mathbb{R}$ that satisfies $(V0 - V4)$. Assume $\mathcal{S}_\mathcal{V}$ is Morse–Smale and fix a regular value $a \in \mathbb{R}$ of $\mathcal{S}_\mathcal{V}$. Fix two constants $C > 0$ and $p > 1$. Then there are positive constants δ, ε_0, and c such that $\varepsilon_0 \le 1$ and the following holds. If $x^\pm \in \mathcal{P}^a(\mathcal{V})$ is a pair of index difference one,*

$$u \in \mathcal{M}^0(x^-, x^+; \mathcal{V}), \qquad (u^\varepsilon, v^\varepsilon) \in \mathcal{M}^\varepsilon(x^-, x^+; \mathcal{V})$$

with $0 < \varepsilon \le \varepsilon_0$, and

$$u^\varepsilon = \exp_u(\xi^\varepsilon),$$

where $\xi^\varepsilon \in \Omega^0(\mathbb{R} \times S^1, u^*TM)$ satisfies

$$\|\xi^\varepsilon\|_\infty \le \delta \varepsilon^{1/2}, \qquad \|\xi^\varepsilon\|_p \le \delta \varepsilon^{1/2}, \qquad \|\nabla_t \xi^\varepsilon\|_p \le C, \tag{78}$$

then there is a real number σ such that

$$(u^\varepsilon, v^\varepsilon) = \mathcal{T}^\varepsilon(u(\sigma + \cdot, \cdot)), \qquad |\sigma| < c(\|\xi^\varepsilon\|_p + \varepsilon^2).$$

Proof. It suffices to prove the result for a fixed pair $x^\pm \in \mathcal{P}^a(\mathcal{V})$ of index difference one and a fixed parabolic cylinder $u \in \mathcal{M}^0(x^-, x^+; \mathcal{V})$. (The assumptions and conclusions of the theorem are invariant under simultaneous time shift of u and $(u^\varepsilon, v^\varepsilon)$; up to time shift there are only finitely many index one parabolic cylinders with $\mathcal{S}_\mathcal{V} \le a$.) Define

$$c^* := \mathcal{S}_\mathcal{V}(x^-) - \mathcal{S}_\mathcal{V}(x^+) > 0. \tag{79}$$

Then, by the energy identity (107), we have $\|\partial_s u\|^2_{L^2(\mathbb{R} \times S^1)} = c^*$. Let $(u^\varepsilon, v^\varepsilon) \in \mathcal{M}^\varepsilon(x^-, x^+; \mathcal{V})$ with $\varepsilon \in (0, 1]$. Denote the time shift of u by

$$u_\sigma(s, t) := u(s + \sigma, t)$$

for $\sigma \in \mathbb{R}$ and define $\zeta = \zeta(\sigma) = (\xi, \eta)$ by

$$u^\varepsilon = exp_{u_\sigma}(\xi), \qquad v^\varepsilon = \Phi(u_\sigma, \xi)\,(\partial_t u_\sigma + \eta)\,. \tag{80}$$

The pair (ξ, η) is well defined whenever $\sigma \|\partial_s u\|_{L^\infty} + \|\xi^\varepsilon\|_{L^\infty}$ is smaller than the injectivity radius ρ_M of M (i.e. when σ and $\delta\varepsilon^{1/2}$ are sufficiently small). We assume throughout that

$$\delta \varepsilon^{1/2} \le \frac{\rho_M}{2}$$

and choose $\sigma_0 > 0$ so that $\sigma_0 \|\partial_s u\|_{L^\infty} < \rho_M/2$.

By Theorem A.1 and Theorem 5.1, there is a constant $c_0 > 0$ such that, for every $\varepsilon \in (0,1]$ and every $(u^\varepsilon, v^\varepsilon) \in \mathcal{M}^\varepsilon(x^-, x^+; \mathcal{V})$, we have

$$\|\partial_s u\|_\infty + \|\partial_t u\|_\infty + \|v^\varepsilon\|_\infty \le c_0. \tag{81}$$

It follows from (80) and (81) that $\|\eta(\sigma)\|_\infty \le c_0$ for every $\sigma \in [-\sigma_0, \sigma_0]$. Choose $\delta_0 > 0$ so small that the assertion of the Uniqueness Theorem 4.3 holds with $C = c_0$ and $\delta = \delta_0$. We shall prove that for every sufficiently small $\varepsilon > 0$ there is a $\sigma \in [-\sigma_0, \sigma_0]$ such that

$$\zeta(\sigma) \in \mathrm{im}(\mathcal{D}^\varepsilon_{u_\sigma})^*, \qquad \|\xi(\sigma)\|_\infty \le \delta_0 \varepsilon^{1/2}, \qquad \|\eta(\sigma)\|_\infty \le c_0. \tag{82}$$

It then follows from Theorem 4.3 that $(u^\varepsilon, v^\varepsilon) = \mathcal{T}^\varepsilon(u_\sigma)$. The proof of (82) will take five steps and uses the following estimate. Choose $q > 1$ such that $1/p + 1/q = 1$. Then, by parabolic exponential decay (see Theorem A.2), there is a constant $c_1 > 0$ such that, for $r = p, q, \infty$,

$$\|\partial_s u\|_r + \|\nabla_t \partial_s u\|_r + \|\nabla_s \partial_s u\|_r + \|\nabla_s \nabla_t \partial_s u\|_r \le c_1. \tag{83}$$

Step 1. *For $\sigma \in [-\sigma_0, \sigma_0]$ and $\varepsilon > 0$ sufficiently small define*

$$\theta^\varepsilon(\sigma) := -\langle Z^\varepsilon_\sigma, \zeta \rangle_\varepsilon,$$

where $\zeta = \zeta(\sigma)$ is given by (80),

$$Z^\varepsilon := \begin{pmatrix} X^\varepsilon \\ Y^\varepsilon \end{pmatrix} := \begin{pmatrix} \partial_s u \\ \nabla_t \partial_s u \end{pmatrix} - \begin{pmatrix} \xi^* \\ \eta^* \end{pmatrix},$$

$$\zeta^* := \begin{pmatrix} \xi^* \\ \eta^* \end{pmatrix} := (\mathcal{D}^\varepsilon_u)^* \left(\mathcal{D}^\varepsilon_u (\mathcal{D}^\varepsilon_u)^*\right)^{-1} \mathcal{D}^\varepsilon_u \begin{pmatrix} \partial_s u \\ \nabla_t \partial_s u \end{pmatrix},$$

and Z^ε_σ denotes the time shift of Z^ε. Then $\theta^\varepsilon(\sigma) = 0$ if and only if $\zeta \in \mathrm{im}(\mathcal{D}^\varepsilon_{u_\sigma})^*$.

For $\varepsilon > 0$ sufficiently small, the operator $\mathcal{D}^\varepsilon_u$ is onto, by Theorem 3.3, and, by assumption, it has index one (see Remark 3.1). Hence Z^ε is well defined and spans the kernel of $\mathcal{D}^\varepsilon_u$; so

$$(Z^\varepsilon)^\perp = (\ker \mathcal{D}^\varepsilon_u)^\perp = \mathrm{im}(\mathcal{D}^\varepsilon_u)^*$$

where $\perp$ denotes the orthogonal complement with respect to the ε-dependent L^2 inner product. It remains to prove that $Z^\varepsilon \ne 0$ for $\varepsilon > 0$ sufficiently small. To see this note that $\partial_s u \ne 0$ and so the $(0,2,\varepsilon)$-norm of the pair $(\partial_s u, \nabla_t \partial_s u)$ is bounded below by a positive constant (the parabolic energy identity gives $\sqrt{c^*}$ as a lower bound). On the other hand,

$$\zeta^* = (\mathcal{D}^\varepsilon_u)^* \left(\mathcal{D}^\varepsilon_u (\mathcal{D}^\varepsilon_u)^*\right)^{-1} \begin{pmatrix} 0 \\ \nabla_s \nabla_t \partial_s u \end{pmatrix} \tag{84}$$

and $\|\pi_\varepsilon(0, \nabla_s \nabla_t \partial_s u)\|_{L^p}$ tends to zero as $\varepsilon \to 0$. Hence, by Theorem 3.3, the $(0,2,\varepsilon)$-norm of ζ^* converges to zero as ε tends to zero. It follows that $Z^\varepsilon \ne 0$ for $\varepsilon > 0$ sufficiently small and this proves Step 1.

STEP 2. *There are positive constants ε_0 and c_2 such that*

$$|\theta^\varepsilon(0)| \le c_2 \left(\|\xi^\varepsilon\|_p + \varepsilon^2 \right)$$

for $0 < \varepsilon \le \varepsilon_0$ and every $(u^\varepsilon = \exp_u(\xi^\varepsilon), v^\varepsilon) \in \mathcal{M}^\varepsilon(x^-, x^+; \mathcal{V})$ satisfying (78).

We first prove that that there are positive constants ε_0 and c_3 such that

$$\|X^\varepsilon\|_q + \|Y^\varepsilon\|_q \le c_3 \tag{85}$$

for $0 < \varepsilon \le \varepsilon_0$. For the summands $\partial_s u$ of X^ε and $\nabla_s \partial_t u$ of Y^ε this follows from (83) with $r = q$. Moreover, by (84) and Theorem 3.3, we have

$$\begin{aligned} \|\xi^*\|_q + \varepsilon^{1/2} \|\eta^*\|_q &\le c_4 \left(\varepsilon \|(0, \nabla_s \nabla_t \partial_s u)\|_{0,q,\varepsilon} + \|\pi_\varepsilon(0, \nabla_s \nabla_t \partial_s u)\|_q \right) \\ &\le c_4 \varepsilon^{3/2} (\varepsilon^{1/2} + \kappa_q) \|\nabla_s \nabla_t \partial_s u\|_q . \end{aligned} \tag{86}$$

The last step uses Lemma D.3 with constant $\kappa_q > 1$. This proves (85). It follows from (85) that

$$|\theta^\varepsilon(0)| \le c_3 \left(\|\xi^\varepsilon\|_p + \varepsilon^2 \|\eta^\varepsilon\|_p \right), \tag{87}$$

where $\eta^\varepsilon \in \Omega^0(\mathbb{R} \times S^1, u^*TM)$ is defined by

$$v^\varepsilon =: \Phi(u, \xi^\varepsilon)(\partial_t u^\varepsilon + \eta^\varepsilon).$$

Define the linear maps $E_i(x, \xi) : T_x M \to T_{exp_x(\xi)} M$ by the formula

$$\frac{d}{d\tau} exp_x(\xi) =: E_1(x, \xi) \partial_\tau x + E_2(x, \xi) \nabla_\tau \xi \tag{88}$$

for every smooth path $x : \mathbb{R} \to M$ and every vector field $\xi \in \Omega^0(\mathbb{R}, x^*TM)$ along x. Abbreviate $\Phi := \Phi(u, \xi^\varepsilon)$ and $E_i := E_i(u, \xi^\varepsilon)$ for $i = 1, 2$. Then

$$\begin{aligned} \eta^\varepsilon &= \Phi^{-1} v^\varepsilon - \partial_t u \\ &= \Phi^{-1}(v^\varepsilon - \partial_t u^\varepsilon) + \Phi^{-1}(E_1 \partial_t u + E_2 \nabla_t \xi_\varepsilon) - \partial_t u \\ &= \varepsilon^2 \Phi^{-1} \nabla_s v^\varepsilon + \Phi^{-1} E_2 \nabla_t \xi^\varepsilon + (\Phi^{-1} E_1 - 1\!\!1) \partial_t u. \end{aligned}$$

By Corollary 8.2, there is a constant c_5 such that $\|\nabla_s v^\varepsilon\|_p \le c_5$. Moreover, there is a constant $c_6 > 0$ such that $\left\|\Phi^{-1} E_1 - 1\!\!1\right\|_p \le c_6 \|\xi^\varepsilon\|_p$. Hence there is another constant $c_7 > 0$ such that

$$\|\eta^\varepsilon\|_p \le c_7 \left(\varepsilon + \|\nabla_t \xi^\varepsilon\|_p + \|\xi^\varepsilon\|_p \right) \le c_7 \left(\|\xi^\varepsilon\|_p + C + 1 \right).$$

Combining this with (87) proves Step 2.

STEP 3. *There is a constant $c_8 > 0$ such that*

$$\|\xi(\sigma)\|_\infty \le \delta \varepsilon^{1/2} + c_8 |\sigma|, \qquad \|\eta(\sigma)\|_\infty \le c_0,$$

$$\|\nabla_s \xi\|_p \le c_8, \quad \|\nabla_\sigma \xi + \partial_s u_\sigma\|_p \le c_8 \left(|\sigma| + \delta \varepsilon^{1/2} \right), \quad \|\xi(\sigma)\|_p \le \delta \varepsilon^{1/2} + c_8 |\sigma|$$

for $0 < \varepsilon \le \varepsilon_0$ and $|\sigma| \le \sigma_0$.

For every $\sigma \in \mathbb{R}$, we have

$$d\left(u(s+\sigma,t),u(s,t)\right) \le L(\gamma) \le |\sigma| \left\|\partial_s u\right\|_\infty,$$

where $\gamma(r) := u(s+r\sigma,t)$, $0 \le r \le 1$. Moreover, by (78), $d(u(s,t),u^\varepsilon(s,t)) \le \delta\varepsilon^{1/2}$. Hence the first estimate of Step 3 follows from the triangle inequality. The second estimate follows from the identity

$$\eta(\sigma) = \Phi(u_\sigma,\xi(\sigma))^{-1}v^\varepsilon - \partial_t u_\sigma$$

and (81). To prove the next two estimates we differentiate the identity

$$\exp_{u_\sigma}(\xi(\sigma)) = u^\varepsilon$$

with respect to σ and s to obtain

$$E_1(u_\sigma,\xi)\partial_s u_\sigma + E_2(u_\sigma,\xi)\nabla_\sigma\xi = 0, \qquad E_1(u_\sigma,\xi)\partial_s u_\sigma + E_2(u_\sigma,\xi)\nabla_s\xi = \partial_s u^\varepsilon.$$

By the energy identities the L^2 norms of $\partial_s u$ and $\partial_s u^\varepsilon$ are uniformly bounded and hence, so is the L^2 norm of $\nabla_s\xi$. Moreover,

$$\left\|\nabla_\sigma\xi + \partial_s u_\sigma\right\|_p = \left\|\left(E_2^{-1}E_1 - \mathbb{1}\right)\partial_s u\right\|_p \le c_9\left\|\xi(\sigma)\right\|_\infty \le c_{10}\left(|\sigma| + \delta\varepsilon^{1/2}\right).$$

Hence the L^p norm of $\nabla_\sigma\xi$ is uniformly bounded. Now differentiate the function $\sigma \mapsto \left\|\xi(\sigma)\right\|_p$ to obtain the inequality $\left\|\xi(\sigma)\right\|_p \le \left\|\xi(0)\right\|_p + c_{11}|\sigma|$. Then the last inequality in Step 3 follows from (78).

STEP 4. *Shrinking σ_0 and ε_0, if necessary, we have*

$$\frac{d}{d\sigma}\theta^\varepsilon(\sigma) \ge \frac{c^*}{2}$$

for $0 < \varepsilon \le \varepsilon_0$ and $|\sigma| \le \sigma_0$, where c^* is defined by (79).

We will investigate the two terms in the sum

$$\frac{d}{d\sigma}\theta^\varepsilon(\sigma) = -\frac{d}{d\sigma}\left\langle X_\sigma^\varepsilon,\xi(\sigma)\right\rangle - \varepsilon^2\frac{d}{d\sigma}\left\langle Y_\sigma^\varepsilon,\eta(\sigma)\right\rangle \tag{89}$$

separately. The key term is $\langle X_\sigma^\varepsilon,\nabla_\sigma\xi\rangle$. We have seen that X_σ^ε is L^q-close to $\partial_s u_\sigma$ and $\nabla_\sigma\xi$ is L^p-close to $-\partial_s u_\sigma$. We shall prove that all the other terms are small and hence $\partial_\sigma\theta^\varepsilon$ is approximately equal to $\left\|\partial_s u\right\|_2^2$. More precisely, for the first term in (89) we obtain

$$\begin{aligned}
-\frac{d}{d\sigma}\left\langle X_\sigma^\varepsilon,\xi\right\rangle &= -\langle X_\sigma^\varepsilon,\nabla_\sigma\xi\rangle - \langle\nabla_s X_\sigma^\varepsilon,\xi\rangle \\
&= \left\|\partial_s u\right\|_2^2 - \langle X_\sigma^\varepsilon,\partial_s u_\sigma + \nabla_\sigma\xi\rangle - \langle\xi^*,\partial_s u_\sigma\rangle \\
&\quad - \langle\nabla_s\partial_s u_\sigma,\xi\rangle - \langle\xi_\sigma^*,\nabla_s\xi\rangle \\
&\ge \left\|\partial_s u\right\|_2^2 - c_{12}\left(\left\|\partial_s u_\sigma + \nabla_\sigma\xi\right\|_p + \left\|\xi^*\right\|_q + \left\|\xi\right\|_p\right) \\
&\ge \left\|\partial_s u\right\|_2^2 - c_{13}\left(|\sigma| + \delta\varepsilon^{1/2} + \varepsilon^{3/2}\right).
\end{aligned}$$

Here the second step follows from integration by parts. The third step uses the inequalities $\|X^\varepsilon\|_q \le c$ (see (85)), $\|\partial_s u\|_p + \|\nabla_s \partial_s u\|_q \le c$ (see (83)), and $\|\nabla_s \xi\|_p \le c$ (see Step 3). The last step uses Step 3 and (86).

To estimate the second term in (89) we differentiate the identity

$$\Phi(u_\sigma, \xi(\sigma))(\partial_t u_\sigma + \eta(\sigma)) = v^\varepsilon$$

with respect to σ to obtain

$$\|\nabla_s \partial_t u_\sigma + \nabla_\sigma \eta\|_p \le c_{14} \left(\|\partial_s u\|_p + \|\nabla_\sigma \xi\|_p \right) \le c_{15}.$$

In the first inequality we have used the fact that the L^∞ norms of $\eta(\sigma)$ and $\partial_t u_\sigma$ are uniformly bounded. In the second inequality we have used Step 3. Combining this estimate with (83) we find that the L^p norm of $\nabla_\sigma \eta$ is uniformly bounded. Differentiating the same identity with respect to s we obtain

$$\|\nabla_s \partial_t u_\sigma + \nabla_s \eta\|_p \le c_{16} \left(\|\nabla_s v^\varepsilon\|_p + \|\partial_s u\|_p + \|\nabla_s \xi\|_p \right) \le c_{17}.$$

Here the last inequality follows from Step 3 and Corollary 8.2. Using (83) again, we obtain that the L^p norm of $\nabla_s \eta$ is uniformly bounded. Now

$$\begin{aligned} \varepsilon^2 \frac{d}{d\sigma} \langle Y_\sigma^\varepsilon, \eta \rangle &= \varepsilon^2 \langle \nabla_s Y_\sigma^\varepsilon, \eta \rangle + \varepsilon^2 \langle Y_\sigma^\varepsilon, \nabla_\sigma \eta \rangle \\ &= -\varepsilon^2 \langle Y_\sigma^\varepsilon, \nabla_s \eta \rangle + \varepsilon^2 \langle Y_\sigma^\varepsilon, \nabla_\sigma \eta \rangle \\ &\le c_{18} \varepsilon^2. \end{aligned}$$

In the last estimate we have used (85) and the uniform estimates on the L^p norms of $\nabla_\sigma \eta$ and $\nabla_s \eta$. Putting things together we obtain

$$\frac{d}{d\sigma} \theta^\varepsilon(\sigma) \ge \|\partial_s u\|_2^2 - c_{19} \left(|\sigma| + \varepsilon^{1/2} \right).$$

Since $\|\partial_s u\|_2^2 = c^*$, the assertion of Step 4 holds whenever $0 < \varepsilon \le \varepsilon_0$, $|\sigma| \le \sigma_0$, and $c_{19}(\sigma_0 + \varepsilon_0^{1/2}) \le c^*/2$.

STEP 5. *We prove Theorem 9.1.*

Suppose the pair $(u^\varepsilon, v^\varepsilon)$ satisfies the requirements of the theorem with ε and δ sufficiently small. Then, by Steps 2 and 4, there is a $\sigma \in [-\sigma_0, \sigma_0]$ such that

$$\theta^\varepsilon(\sigma) = 0, \qquad |\sigma| \le c_{20}(\|\xi^\varepsilon\|_p + \varepsilon^2), \qquad c_{20} := \frac{2c_2}{c^*}.$$

Let $\xi := \xi(\sigma)$ and $\eta := \eta(\sigma)$. Then, by Step 1, we have $\zeta := (\xi, \eta) \in \mathrm{im}\,(\mathcal{D}_{u_\sigma}^\varepsilon)^*$ and, by Step 3,

$$\|\xi\|_\infty \le (\delta + c_8 c_{20}(\delta + \varepsilon^{3/2}))\varepsilon^{1/2}, \qquad \|\eta\|_\infty \le c_0.$$

If $\delta + c_8 c_{20}(\delta + \varepsilon^{3/2}) \le \delta_0$ then, by Theorem 4.3, $(u^\varepsilon, v^\varepsilon) = \mathcal{T}^\varepsilon(u_\sigma)$. $\square$

10 Surjectivity

Theorem 10.1. *Fix a perturbation $\mathcal{V} : \mathcal{L}M \to \mathbb{R}$ that satisfies $(V0 - V4)$. Assume $\mathcal{S}_\mathcal{V}$ is Morse–Smale and fix a constant $a \in \mathbb{R}$. Then there is a constant $\varepsilon_0 > 0$ such that, for every $\varepsilon \in (0, \varepsilon_0)$ and every pair $x^\pm \in \mathcal{P}^a(\mathcal{V})$ of index difference one, the map $\mathcal{T}^\varepsilon : \mathcal{M}^0(x^-, x^+; \mathcal{V}) \to \mathcal{M}^\varepsilon(x^-, x^+; \mathcal{V})$, constructed in Definition 4.4, is bijective.*

The proof relies on the following two lemmas. We use the notation

$$E_{[-T,T]}(u) := \int_{-T}^{T} \int_0^1 |\partial_s u|^2 \, dtds$$

for $u : \mathbb{R} \times S^1 \to M$ and $T > 0$.

Lemma 10.2. *Assume $\mathcal{S}_\mathcal{V}$ is a Morse function. Let $x^\pm \in \mathcal{P}(\mathcal{V})$ and $(u_i, v_i) \in \mathcal{M}^{\varepsilon_i}(x^-, x^+; \mathcal{V})$ where ε_i is a sequence of positive real numbers converging to zero. Then there is a pair $x_0, x_1 \in \mathcal{P}(\mathcal{V})$, a parabolic cylinder $u \in \mathcal{M}^0(x_0, x_1; \mathcal{V})$, and a subsequence, still denoted by (u_i, v_i), such that the following holds.*

- **(i)** *(u_i, v_i) converges to $(u, \partial_t u)$ strongly in C^1 and weakly in $W^{2,p}$ on every compact subset of $\mathbb{R} \times S^1$ and for every $p > 1$. Moreover, $v_i - \partial_t u_i$ converges to zero in the C^1 norm on every compact subset of $\mathbb{R} \times S^1$.*

- **(ii)** *For all $s \in \mathbb{R}$ and $T > 0$,*

$$\mathcal{S}_\mathcal{V}(u(s, \cdot)) = \lim_{i \to \infty} \mathcal{A}_\mathcal{V}(u_i(s, \cdot), v_i(s, \cdot)),$$
$$E_{[-T,T]}(u) = \lim_{i \to \infty} E^{\varepsilon_i}_{[-T,T]}(u_i, v_i).$$

Proof. By the energy identity (8) and Theorems 5.1, 6.1, and 7.1, there is a constant $c > 0$ such that

$$\|v_i\|_\infty + \|\partial_t u_i\|_\infty + \|\partial_s u_i\|_\infty + \|\nabla_t v_i\|_\infty + \|\nabla_s v_i\|_\infty \le c, \tag{90}$$

$$\|\nabla_s \partial_t u_i\|_p + \|\nabla_s \partial_s u_i\|_p + \|\nabla_t \nabla_s v_i\|_p + \|\nabla_s \nabla_s v_i\|_p \le c, \tag{91}$$

$$\|\nabla_t \partial_t u_i\|_\infty + \|\nabla_t \nabla_t v_i\|_\infty \le c \tag{92}$$

for every $i \in \mathbb{N}$ and every $p \in [2, \infty]$. In (90) the estimate for $\nabla_t v_i$ follows from the one for $\partial_s u_i$ and the identity $\nabla_t v_i = \partial_s u_i - \operatorname{grad} \mathcal{V}(u_i)$. The estimate for $\partial_t u_i$ follows from the ones for v_i and $\nabla_s v_i$ and the identity $\partial_t u_i = v_i - \varepsilon_i^2 \nabla_s v_i$. In (92) the estimate for $\nabla_t \partial_t u_i$ follows from the ones for $\nabla_t v_i$ and $\nabla_t \nabla_s v_i$ and the identity $\nabla_t \partial_t u_i = \nabla_t v_i - \varepsilon_i^2 \nabla_t \nabla_s v_i$. The estimate for $\nabla_t \nabla_t v_i$ follows from the ones for $\nabla_t \partial_s u_i$ and $\partial_t u_i$ and the identity $\nabla_t \nabla_t v_i = \nabla_t \partial_s u_i - \nabla_t \operatorname{grad} \mathcal{V}(u_i)$.

By (90), (91), and (92) the sequence (u_i, v_i) is bounded in C^2 and hence in $W^{2,p}([-T, T] \times S^1)$ for every $T > 0$ and every $p > 1$. Hence, by the Arzela-Ascoli theorem and the Banach–Alaoglu theorem, a suitable subsequence, still denoted by (u_i, v_i), converges strongly in C^1 and weakly in $W^{2,p}$ on every compact subset of $\mathbb{R} \times S^1$ to some C^2-funtion $(u, v) : \mathbb{R} \times S^1 \to TM$. By (90) and (91), the sequence

$$v_i - \partial_t u_i = \varepsilon_i^2 \nabla_s v_i$$

converges to zero in the C^1 norm. Hence $v = \partial_t u$. Moreover, the sequence

$$\partial_s u_i - \nabla_t \partial_t u_i - \operatorname{grad} \mathcal{V}(u_i) = \varepsilon_i^2 \nabla_t \nabla_s v_i$$

converges to zero in the sup-norm, by (91), so the limit $u : \mathbb{R} \times S^1 \to M$ satisfies the parabolic equation (12). By the parabolic regularity theorem A.3, u is smooth and so is $v = \partial_t u$. This proves (i).

To prove (ii) note that

$$\begin{aligned} E_{[-T,T]}(u) &= \lim_{i\to\infty} \int_{-T}^{T} \int_0^1 |\partial_s u_i|^2 \, dsdt \\ &= \lim_{i\to\infty} \int_{-T}^{T} \int_0^1 \left(|\partial_s u_i|^2 + \varepsilon_i^2 |\nabla_s v_i|^2 \right) dsdt \\ &= \lim_{i\to\infty} E^{\varepsilon_i}_{[-T,T]}(u_i, v_i) \end{aligned}$$

for every T; here the third identity follows from (90). Hence the limit u has finite energy and so belongs to the moduli space $\mathcal{M}^0(x_0, x_1; \mathcal{V})$ for some pair $x_0, x_1 \in \mathcal{P}(\mathcal{V})$. To prove convergence of the symplectic action at time s note that

$$\mathcal{V}(u(s,\cdot)) = \lim_{i\to\infty} \mathcal{V}(u_i(s,\cdot)),$$

because $\mathcal{V}$ is continuous with respect to the C^0 topology on $\mathcal{L}M$. Moreover

$$\begin{aligned} \mathcal{S}_0(u(s,\cdot)) &= \frac{1}{2} \int_0^1 |\partial_t u(s,t)|^2 \, dt \\ &= \lim_{i\to\infty} \int_0^1 \left(\langle \partial_t u_i(s,t), v_i(s,t) \rangle - \frac{1}{2} |v_i(s,t)|^2 \right) dt \\ &= \lim_{i\to\infty} \mathcal{A}_0(u_i(s,\cdot), v_i(s,\cdot)). \end{aligned}$$

Here the second equality follows from the fact that $\partial_t u_i(s,\cdot)$ and $v_i(s,\cdot)$ both converge to $\partial_t u(s,\cdot)$ in the sup-norm. This proves the lemma. □

Lemma 10.3. *Assume $\mathcal{S}_\mathcal{V}$ is a Morse function. Let $x^\pm \in \mathcal{P}(\mathcal{V})$ and $(u_i, v_i) \in \mathcal{M}^{\varepsilon_i}(x^-, x^+; \mathcal{V})$ where ε_i is a sequence of positive real numbers converging to zero. Then there exist periodic orbits $x^- = x^0, x^1, \ldots, x^\ell = x^+ \in \mathcal{P}(\mathcal{V})$, parabolic cylinders $u^k \in \mathcal{M}^0(x^{k-1}, x^k; \mathcal{V})$ for $k \in \{1, \ldots, \ell\}$, a subsequence, still denoted by (u_i, v_i), and sequences $s_i^k \in \mathbb{R}$, $k \in \{1, \ldots, \ell\}$, such that the following holds.*

(i) *For every $k \in \{1, \ldots, \ell\}$ the sequence $(s,t) \mapsto (u_i(s_i^k + s, t), v_i(s_i^k + s, t))$ converges to $(u^k, \partial_t u^k)$ as in Lemma 10.2.*

(ii) *$s_i^k - s_i^{k-1}$ diverges to infinity for $k = 2, \ldots, \ell$ and $\partial_s u^k \not\equiv 0$ for $k = 1, \ldots, \ell$.*

(iii) *For every $k \in \{0, \dots, \ell\}$ and every $\rho > 0$ there is a constant $T > 0$ such that, for every i and every $(s,t) \in \mathbb{R} \times S^1$,*

$$s_i^k + T \le s \le s_i^{k+1} - T \qquad \Longrightarrow \qquad d(u_i(s,t), x^k(t)) < \rho.$$

(Here we abbreviate $s_i^0 := -\infty$ and $s_i^{\ell+1} := \infty$.)

Proof. Denote $a := \mathcal{S}_\mathcal{V}(x^-)$ and choose $\rho > 0$ so small that $d(x(t), x'(t)) > 2\rho$ for every $t \in \mathbb{R}$ and any two distinct periodic orbits $x, x' \in \mathcal{P}^a(\mathcal{V})$. Choose s_i^1 such that

$$\sup_{s \le s_i^1} \sup_t d(x^-(t), u_i(s,t)) \le \rho, \qquad \sup_t d(x^-(t), u_i(s_i^1, t)) = \rho. \tag{93}$$

Passing to a subsequence we may assume, by Lemma 10.2, that the sequence $(u_i(s_i^1 + \cdot, \cdot), v_i(s_i^1 + \cdot, \cdot))$ converges in the required sense to a parabolic cylinder $u^1 \in \mathcal{M}^0(x^0, x^1; \mathcal{V})$, where $x^0, x^1 \in \mathcal{P}^a(\mathcal{V})$. By (93), we have $x^0 = x^-$ and $x^1 \neq x^0$. Hence $\partial_s u^1 \not\equiv 0$ and so $\mathcal{S}_\mathcal{V}(x^1) < \mathcal{S}_\mathcal{V}(x^0)$. If $x^1 = x^+$ the lemma is proved. If $x^1 \neq x^+$ choose $T > 0$ such that $d(u^1(s,t), x^1(t)) < \rho$ for every t and every $s \ge T$. Passing to a subsequence, we may assume that $d(u_i(s_i^1 + T, t), x^1(t)) < \rho$ for every t. Since $x^1 \neq x^+$ there exists a sequence $s_i^2 > s_i^1 + T$ such that

$$\sup_{s_i^1 + T \le s \le s_i^2} \sup_t d(x^1(t), u_i(s,t)) \le \rho, \qquad \sup_t d(x^1(t), u_i(s_i^2, t)) = \rho.$$

The difference $s_i^2 - s_i^1$ diverges to infinity and, by Lemma 10.2, there is a further subsequence such that $(u_i(s_i^2 + \cdot, \cdot), v_i(s_i^2 + \cdot, \cdot))$ converges to a parabolic cylinder $u^2 \in \mathcal{M}^0(x^1, x^2; \mathcal{V})$, where $\mathcal{S}_\mathcal{V}(x^2) < \mathcal{S}_\mathcal{V}(x^1)$. Continue by induction. The induction can only terminate if $x^\ell = x^+$. It must terminate because $\mathcal{P}^a(\mathcal{V})$ is a finite set. This proves the lemma. □

Proof of Theorem 10.1. By Theorem 4.3 the map $\mathcal{T}^\varepsilon$ is injective for $\varepsilon > 0$ sufficiently small. We will prove surjectivity by contradiction.

Assume the result is false. Then there exist periodic orbits $x^\pm \in \mathcal{P}^a(\mathcal{V})$ of Morse index difference one and sequences $\varepsilon_i > 0$ and $(u_i, v_i) \in \mathcal{M}^{\varepsilon_i}(x^-, x^+; \mathcal{V})$ such that

$$\lim_{i \to \infty} \varepsilon_i = 0, \qquad (u_i, v_i) \notin \mathcal{T}^{\varepsilon_i}(\mathcal{M}^0(x^-, x^+; \mathcal{V})). \tag{94}$$

Applying a time shift, if necessary, we assume without loss of generality that

$$\mathcal{A}_\mathcal{V}\big(u_i(0,\cdot), v_i(0,\cdot)\big) = \frac{1}{2}\big(\mathcal{S}_\mathcal{V}(x^-) + \mathcal{S}_\mathcal{V}(x^+)\big). \tag{95}$$

Fix a constant $p > 2$. We shall prove in two steps that, after passing to a subsequence if necessary, there is a sequence $u_i^0 \in \mathcal{M}^0(x^-, x^+; \mathcal{V})$ and a constant $C > 0$ such that

$$u_i = exp_{u_i^0}(\xi_i),$$

where the sequence $\xi_i \in \Omega^0(\mathbb{R} \times S^1, (u_i^0)^* TM)$ satisfies

$$\lim_{i \to \infty} \varepsilon_i^{-1/2}\big(\|\xi_i\|_\infty + \|\xi_i\|_p\big) = 0, \qquad \|\nabla_t \xi_i\|_p \le C. \tag{96}$$

Hence it follows from Theorem 9.1 that, for i sufficiently large, there is a real number σ_i such that $(u_i, v_i) = \mathcal{T}^{\varepsilon_i}(u_i^0(\sigma_i + \cdot, \cdot))$. This contradicts (94) and hence proves Theorem 10.1.

STEP 1. *For every $\delta > 0$ there is a constant $T_0 > 0$ such that*

$$E^{\varepsilon_i}_{\mathbb{R}\setminus[-T_0,T_0]}(u_i, v_i) < \delta \tag{97}$$

for every $i \in \mathbb{N}$.

Assume, by contradiction, that the statement is false. Then there is a constant $\delta > 0$, a sequence of positive real numbers $T_i \to \infty$, and a subsequence, still denoted by $(\varepsilon_i, u_i, v_i)$, such that, for every $i \in \mathbb{N}$,

$$E^{\varepsilon_i}_{[-T_i,T_i]}(u_i, v_i) \leq \mathcal{S}_{\mathcal{V}}(x^-) - \mathcal{S}_{\mathcal{V}}(x^+) - \delta. \tag{98}$$

Choose a further subsequence, still denoted by (u_i, v_i), that converges as in Lemma 10.3 to a finite collection of parabolic cylinders $u^k \in \mathcal{M}^0(x^{k-1}, x^k; \mathcal{V})$, $k = 1, \ldots, \ell$, with $x^- = x^0, x^1, \ldots, x^{\ell-1}, x^\ell = x^+ \in \mathcal{P}(\mathcal{V})$. We claim that $\ell \geq 2$. Otherwise, $u_i(s_i + \cdot, \cdot)$ converges to $u := u^1 \in \mathcal{M}^0(x^-, x^+; \mathcal{V})$ as in Lemma 10.2 (for some sequence $s_i \in \mathbb{R}$). By (95) and Lemma 10.2 (ii), the sequence s_i must be bounded. By (98), this implies that

$$E_{[-T,T]}(u) = \int_{-T}^{T} \int_0^1 |\partial_s u|^2 \; dtds \leq \mathcal{S}_{\mathcal{V}}(x^-) - \mathcal{S}_{\mathcal{V}}(x^+) - \delta$$

for every $T > 0$. This contradicts the fact that u connects x^- with x^+. Thus we have proved that $\ell \geq 2$ as claimed. Since $\mathcal{S}_{\mathcal{V}}$ is Morse–Smale it follows that the Morse index difference of x^- and x^+ is at least two. This contradicts our assumption and proves Step 1.

STEP 2. *For i sufficiently large there is a parabolic cylinder $u_i^0 \in \mathcal{M}^0(x^-, x^+; \mathcal{V})$ and a vector field $\xi_i \in \Omega^0(\mathbb{R} \times S^1, (u_i^0)^*TM)$ such that $u_i = exp_{u_i^0}(\xi_i)$ and ξ_i satisfies (96).*

Let δ, c and ρ denote the constants in Theorem 8.1 and choose $T_0 > 0$, according to Step 1, such that (97) holds with this constant δ. Then, by Corollary 8.2,

$$|\partial_s u_i(s,t)|^2 + |\nabla_s v_i(s,t)|^2 \leq c_3 e^{-\rho|s|} E^{\varepsilon_i}_{\mathbb{R}\setminus[-T_0,T_0]}(u_i, v_i) \tag{99}$$

for $|s| \geq T_0 + 2$ and a suitable constant $c_3 > 0$. By Theorem 5.1 and Theorem 6.1, there is a constant $c_4 > 0$ such that

$$\|v_i\|_\infty + \|\partial_s u_i\|_\infty + \|\partial_t u_i\|_\infty + \|\nabla_s v_i\|_\infty \leq c_4 \tag{100}$$

for every i. Here we have also used the identity $\partial_t u_i = v_i - \varepsilon_i^2 \nabla_s v_i$. It follows from (99) and (100) that there is a constant $c_5 \geq c_4$ such that

$$|\partial_s u_i(s,t)| + |\nabla_s v_i(s,t)| \leq \frac{c_5}{1 + s^2}$$

for every $(s,t) \in \mathbb{R}\times S^1$ and every $i \in \mathbb{N}$. Moreover, it follows from Theorem 7.1 that

$$\left\|\partial_s u_i - \nabla_t\partial_t u_i - \operatorname{grad}\mathcal{V}(u_i)\right\|_p = \varepsilon_i^2 \left\|\nabla_t\nabla_s v_i\right\|_p \leq c_6\varepsilon_i^2.$$

for a suitable constant $c_6 > 0$. Now let $\delta_0 = \delta_0(p,c_5)$ and $c = c(p,c_5)$ be the constants in the parabolic implicit function theorem A.5. Then the function u_i satisfies the hypotheses of Theorem A.5, whenever $c_6\varepsilon_i^2 < \delta_0$. Hence, for i sufficiently large, there is a parabolic cylinder $u_i^0 \in \mathcal{M}^0(x^-,x^+;\mathcal{V})$ and a vector field $\xi_i \in \Omega^0(\mathbb{R}\times S^1,(u_i^0)^*TM)$ such that

$$u_i = exp_{u_i^0}(\xi_i),$$

$$\left\|\xi_i\right\|_{\mathcal{W}_{u_i^0}} \leq c_7\left\|\partial_s u_i - \nabla_t\partial_t u_i - \operatorname{grad}\mathcal{V}(u_i)\right\|_p \leq c_6 c_7\varepsilon_i^2.$$

By the Sobolev embedding theorem, we have

$$\left\|\xi_i\right\|_\infty \leq c_8\left\|\xi_i\right\|_{\mathcal{W}_{u_i^0}} \leq c_6c_7c_8\varepsilon_i^2$$

for large i. Moreover, by definition of the $\mathcal{W}_{u_i^0}$-norm and Lemma D.4 we have

$$\left\|\xi_i\right\|_p + \left\|\nabla_t\xi_i\right\|_p \leq 2\left\|\xi_i\right\|_{\mathcal{W}_{u_i^0}} \leq 2c_6c_7\varepsilon_i^2.$$

Hence ξ_i satisfies (96). This proves Step 2 and the theorem. □

Corollary 10.4. *Fix a perturbation $\mathcal{V} : \mathcal{L}M \to \mathbb{R}$ that satisfies $(V0 - V4)$. Assume $\mathcal{S}_\mathcal{V}$ is Morse–Smale and fix a regular value a of $\mathcal{S}_\mathcal{V}$. Then there is a constant $\varepsilon_0 > 0$ such that, for every $\varepsilon \in (0,\varepsilon_0]$, the following holds.*

(i) *If $x^\pm \in \mathcal{P}^a(\mathcal{V})$ have index difference less than or equal to zero and $x^+ \neq x^-$ then $\mathcal{M}^\varepsilon(x^-,x^+;\mathcal{V}) = \emptyset$.*

(ii) *If $x^\pm \in \mathcal{P}^a(\mathcal{V})$ have index difference one then*

$$\#\mathcal{M}^0(x^-,x^+;\mathcal{V})/\mathbb{R} = \#\mathcal{M}^\varepsilon(x^-,x^+;\mathcal{V})/\mathbb{R}.$$

(iii) *If $x^\pm \in \mathcal{P}^a(\mathcal{V})$ have index difference one and $(u,v) \in \mathcal{M}^\varepsilon(x^-,x^+;\mathcal{V})$ then $\mathcal{D}^\varepsilon_{u,v}$ is surjective.*

Proof. Assertion (i) follows from Lemma 10.3. Assertion (ii) follows from Theorems 4.1 and 10.1. Assertion (iii) follows from Theorems 4.1, 3.3, and 10.1. □

11 Proof of the main result

Theorem 11.1. *The assertion of Theorem 1.1 holds with $\mathbb{Z}_2$-coefficients.*

Proof. Let V_t be a potential such that $\mathcal{S}_V$ is a Morse function on the loop space and denote
$$\mathcal{V}(x) := \int_0^1 V_t(x(t))\,dt.$$
Fix a regular value a of $\mathcal{S}_V$. Choose a sequence of perturbations $\mathcal{V}_i : \mathcal{L}M \to \mathbb{R}$, converging to $\mathcal{V}$ in the C^∞ topology, such that $\mathcal{S}_{\mathcal{V}_i} : \mathcal{L}M \to \mathbb{R}$ is Morse–Smale for every i. We may assume without loss of generality that the perturbations agree with $\mathcal{V}$ near the critical points and that $\mathcal{P}(\mathcal{V}_i) = \mathcal{P}(V)$ for all i. Let $\varepsilon_i > 0$ be the constant of Corollary 10.4 for $\mathcal{V} = \mathcal{V}_i$. Then, by Corollary 10.4,
$$\#\mathcal{M}^0(x^-, x^+; \mathcal{V}_i)/\mathbb{R} = \#\mathcal{M}^{\varepsilon_i}(x^-, x^+; \mathcal{V}_i)/\mathbb{R}$$
for every pair $x^\pm \in \mathcal{P}^a(V)$ with index difference one. Hence the Floer boundary operator on the chain complex
$$C^a(V; \mathbb{Z}_2) := \bigoplus_{x \in \mathcal{P}^a(V)} \mathbb{Z}_2 x,$$
defined by counting modulo 2 the solutions of (11) with $\mathcal{V} = \mathcal{V}_i$ and $\varepsilon = \varepsilon_i$ agrees with the Morse boundary operator defined by counting the solutions of (12) with $\mathcal{V} = \mathcal{V}_i$. Let us denote the resulting Floer homology groups by $\mathrm{HF}_*^a(T^*M, \mathcal{V}_i, \varepsilon_i; \mathbb{Z}_2)$. Then, by what we have just observed, there is a natural isomorphism
$$\mathrm{HF}_*^a(T^*M, \mathcal{V}_i, \varepsilon_i; \mathbb{Z}_2) \cong \mathrm{HM}_*^a(\mathcal{L}M, \mathcal{S}_{\mathcal{V}_i}; \mathbb{Z}_2) \cong \mathrm{H}_*(\{\mathcal{S}_{\mathcal{V}_i} \le a\}; \mathbb{Z}_2).$$
Here the last isomorphism follows from Theorem A.7. The assertion of Theorem 1.1 with $\mathbb{Z}_2$ coefficients now follows from the isomorphisms
$$\mathrm{HF}_*^a(T^*M, H_V; \mathbb{Z}_2) \cong \mathrm{HF}_*^a(T^*M; \mathcal{V}_i, \varepsilon_i; \mathbb{Z}_2)$$
and
$$\mathrm{H}_*(\{\mathcal{S}_{\mathcal{V}_i} \le a\}; \mathbb{Z}_2) \cong \mathrm{H}_*(\{\mathcal{S}_V \le a\}; \mathbb{Z}_2)$$
for i sufficiently large. Here the second isomorphism follows by varying the level a and noting that the inclusions $\{\mathcal{S}_V \le a\} \hookrightarrow \{\mathcal{S}_{\mathcal{V}_i} \le b\} \hookrightarrow \{\mathcal{S}_V \le c\}$ are homotopy equivalences for $a < b < c$, c sufficiently close to a, and i sufficiently large. To understand the isomorphism on Floer homology, we first recall that the Floer homology groups $\mathrm{HF}_*^a(T^*M, H_V; \mathbb{Z}_2)$ (for a nonregular Hamiltonian H_V and a regular value a of the symplectic action $\mathcal{A}_V$) are defined in terms of almost complex structures J and nearby Hamiltonian functions H, such that (J, H) is a regular pair in the sense of Floer; one then defines $\mathrm{HF}_*^a(T^*M, H_V; \mathbb{Z}_2) := \mathrm{HF}_*^a(T^*M, H, J; \mathbb{Z}_2)$ and observes that the resulting Floer homology groups are independent of J and of the nearby Hamiltonian H. Now let $J = J_{\varepsilon_i}$ be the almost complex structure of Remark 1.3 and choose a J_{ε_i}-regular Hamiltonian $H = H_V + W$ with W sufficiently close to zero and of compact support. Then the Floer equation for the pair (J_{ε_i}, H) can be written in the form
$$\partial_s u + \nabla_t v = \nabla V_t(u) + \nabla_1 W_t(u, v), \qquad \varepsilon_i^2 \nabla_s v + \partial_t u = v + \nabla_2 W_t(u, v). \quad (101)$$

Now the standard Floer homotopy argument can be used to relate the Floer complex associated to (101) to that of

$$\partial_s u + \nabla_t v = \operatorname{grad} \mathcal{V}_i(u), \qquad \varepsilon_i^2 \nabla_s v + \partial_t u = v. \tag{102}$$

To carry the standard theory over to the present case one must establish the apriori bound of Section 5 for the solutions of the time dependent Floer equation. To establish these one can use the arguments of Section 5 (with eps=1) in an almost word by word fashion. In particular, the key estimate (35) remains valid if the metric and perturbation depend on s. Once this is understood, the standard Floer homotopy arguments apply (see [8, 22, 23] for example). This shows that $\mathrm{HF}_*^a(T^*M, H_V; \mathbb{Z}_2)$ is isomorphic to $\mathrm{HF}_*^a(T^*M, \mathcal{V}_i, \varepsilon_i; \mathbb{Z}_2)$ for i sufficiently large. This proves Theorem 1.1 with $\mathbb{Z}_2$ coefficients when $a < \infty$. For $a = \infty$ the result follows from naturality and a direct limit argument. □

To prove the result with integer coefficients it remains to examine the orientations of the moduli spaces. The first step is a result about abstract Fredholm operators on Hilbert spaces.

Let $W \subset H$ be an inclusion of Hilbert spaces that is compact and has a dense image. Let $\mathbb{R} \mapsto \mathcal{L}(W, H) : s \mapsto A(s)$ be a family of bounded linear operators satisfying the following conditions.

(A1) The map $s \mapsto A(s)$ is continuously differentiable in the norm topology. Moreover, there is a constant $c > 0$ such that

$$\|A(s)\xi\|_H + \|\dot{A}(s)\xi\|_H \le c\,\|\xi\|_W$$

for every $s \in \mathbb{R}$ and every $\xi \in W$.

(A2) The operators $A(s)$ are uniformly self-adjoint. This means that, for each s, the operator $A(s)$, when considered as an unbounded operator on H, is self adjoint, and there is a constant c such that

$$\|\xi\|_W \le c\,(\|A(s)\xi\|_H + \|\xi\|_H)$$

for every $s \in \mathbb{R}$ and every $\xi \in W$.

(A3) There are invertible operators $A^\pm : W \to H$ such that

$$\lim_{s \to \pm\infty} \left\| A(s) - A^\pm \right\|_{\mathcal{L}(W,H)} = 0.$$

(A4) The operator $A(s)$ has finitely many negative eigenvalues for every $s \in \mathbb{R}$.

Denote by $\mathcal{S}(W, H)$ the set of invertible self-adjoint operators $A : W \to H$ with finitely many negative eigenvalues. For $A \in \mathcal{S}(W, H)$ denote by $E(A)$ the direct sum of the eigenspaces of A with negative eigenvalues. Given $A^\pm \in \mathcal{S}(W, H)$ denote by $\mathcal{P}(A^-, A^+)$ the set of functions $A : \mathbb{R} \to \mathcal{L}(W, H)$ that satisfy (A1-4) and by $\mathcal{P}$ the union of the spaces $\mathcal{P}(A^-, A^+)$ over all pairs $A^\pm \in \mathcal{S}(W, H)$. This is an open subset of a Banach space.

Denote

$$\mathcal{W} := L^2(\mathbb{R}, W) \cap W^{1,2}(\mathbb{R}, H), \qquad \mathcal{H} := L^2(\mathbb{R}, H)$$

and, for every pair $A^{\pm} \in \mathcal{S}(W, H)$ and every $A \in \mathcal{P}(A^-, A^+)$, consider the operator $\mathcal{D}_A : \mathcal{W} \to \mathcal{H}$ defined by

$$(\mathcal{D}_A \xi)(s) := \dot{\xi}(s) + A(s)\xi(s)$$

for $\xi \in \mathcal{W}$. This operator is Fredholm and its index is the spectral flow, i.e.

$$\mathrm{index}(\mathcal{D}_A) = \dim E(A^-) - \dim E(A^+)$$

(see Robbin–Salamon [19]). The formal adjoint operator $\mathcal{D}_A^* : \mathcal{W} \to \mathcal{H}$ is given by $\mathcal{D}_A^* \eta = -\dot{\eta} + A\eta$. Denote by

$$\det(\mathcal{D}_A) := \Lambda^{\max}(\ker \mathcal{D}_A) \otimes \Lambda^{\max}(\ker (\mathcal{D}_A)^*)$$

the determinant line of $\mathcal{D}_A$ and by $\mathrm{Or}(\mathcal{D}_A)$ the set of orientations of $\det(\mathcal{D}_A)$. For $A \in \mathcal{S}(W, H)$ denote by $\mathrm{Or}(A)$ the set of orientations of $E(A)$.

Remark 11.2 (The finite dimensional case). Assume $W = H = \mathbb{R}^n$. Let $A^{\pm}$ be nonsingular symmetric $(n \times n)$-matrices and $A \in \mathcal{P}(A^-, A^+)$. Suppose that $A(s) = A^{\pm}$ for $\pm s \geq T$. Define $\Phi(s, s_0) \in \mathbb{R}^{n \times n}$ by

$$\partial_s \Phi(s, s_0) + A(s)\Phi(s, s_0) = 0, \qquad \Phi(s_0, s_0) = \mathbb{1}.$$

Define

$$E^{\pm}(s) := \left\{ \xi \in \mathbb{R}^n \mid \lim_{r \to \pm\infty} \Phi(r, s)\xi = 0 \right\}.$$

Then $E^-(s) = E(A^-)$ for $s \leq -T$ and $E^+(s) = E(A^+)^{\perp}$ for $s \geq T$. Moreover,

$$\ker \mathcal{D}_A \cong E^-(s) \cap E^+(s), \qquad (\mathrm{im}\mathcal{D}_A)^{\perp} \cong (E^-(s) + E^+(s))^{\perp}.$$

Hence there is a natural map

$$\tau_A : \mathrm{Or}(A^-) \times \mathrm{Or}(A^+) \to \mathrm{Or}(\mathcal{D}_A)$$

defined as follows. Given orientations of $E(A^-) \cong E^-(s)$ and $E(A^+) \cong E^+(s)^{\perp}$, pick any basis $u_1, \ldots, u_\ell$ of $E^-(s) \cap E^+(s) \cong \ker \mathcal{D}_A$. Extend it to a positive basis of $E^-(s)$ by picking a suitable basis $v_1, \ldots, v_m$ of $E^-(s) \cap E^+(s)^{\perp}$. Now extend the vectors v_j to a positive basis of $E^+(s)^{\perp}$ by picking a suitable basis $w_1, \ldots, w_n$ of $(E^-(s) + E^+(s))^{\perp} \cong (\mathrm{im}\, \mathcal{D}_A)^{\perp}$. Then the bases $u_1, \ldots, u_\ell$ of $\mathit{ker}\, \mathcal{D}_A$ and $w_1, \ldots, w_n$ of $(\mathrm{im}\, \mathcal{D}_A)^{\perp}$ determine the induced orientation of $\det(\mathcal{D}_A)$. Note that this is well defined (a sign change in the u_i leads to a sign change in the w_k).

Remark 11.3 (Catenation). Let $A_0, A_1, A_2 \in \mathcal{S}(W,H)$ and suppose that $A_{01} \in \mathcal{P}(A_0, A_1)$ and $A_{12} \in \mathcal{P}(A_1, A_2)$ satisfy

$$A_{01}(s) = \begin{cases} A_0 & \text{if } s \le -T, \\ A_1 & \text{if } s \ge T, \end{cases} \qquad A_{12}(s) = \begin{cases} A_1 & \text{if } s \le -T, \\ A_2 & \text{if } s \ge T. \end{cases} \tag{103}$$

For $R > T$ define $A_{02}^R \in \mathcal{P}(A_0, A_2)$ by

$$A_{02}^R(s) = \begin{cases} A_{01}(s+R) & \text{if } s \le 0, \\ A_{12}(s-R) & \text{if } s \ge 0. \end{cases} \tag{104}$$

If $\mathcal{D}_{A_{01}}$ and $\mathcal{D}_{A_{12}}$ are onto then, for R sufficiently large, the operator $\mathcal{D}_{A_{02}^R}$ is onto and there is a natural isomorphism

$$S^R : \ker \mathcal{D}_{A_{01}} \oplus \ker \mathcal{D}_{A_{12}} \to \ker \mathcal{D}_{A_{02}^R}$$

The isomorphism S^R is defined by composing a pre-gluing operator with the orthogonal projection onto the kernel. That this gives an isomorphism follows from exponential decay estimates for the elements in the kernel and a uniform estimate for suitable right inverses of the operators $\mathcal{D}_{A_{02}^R}$ (see for example [22]).

Theorem 11.4. *There is a family of maps*

$$\tau_A : \mathrm{Or}(A^-) \times \mathrm{Or}(A^+) \to \mathrm{Or}(\mathcal{D}_A),$$

one for each pair of Hilbert spaces $W \subset H$ with a compact dense inclusion, each pair $A^\pm \in \mathcal{S}(W,H)$, and each $A \in \mathcal{P}(A^-, A^+)$, satisfying the following axioms.

(Equivariant) *τ_A is equivariant with respect to the $\mathbb{Z}_2$-action on each factor.*

(Homotopy) *The map $(A, o^-, o^+) \mapsto (A, \tau_A(o^-, o^+))$ from the topological space $\{(A, o^-, o^+) \,|\, A \in \mathcal{P},\, o^\pm \in \mathrm{Or}(A^\pm)\}$ to $\{(A, o) \,|\, A \in \mathcal{P},\, o \in \mathrm{Or}(\mathcal{D}_A)\}$ is continuous.*

(Naturality) *Let $\Phi(s) : (W,H) \to (W', H')$ be a family of (pairs of) Hilbert space isomorphisms that is continuously differentiable in the operator norm on H and continuous in the operator norm on W. Suppose that there exist Hilbert space isomorphisms $\Phi^\pm : (W,H) \to (W',H')$ such that $\Phi(s)$ converges to $\Phi^\pm$ in the operator norm on both spaces and $\dot\Phi(s)$ converges to zero in $\mathcal{L}(H)$ as $s \to \pm\infty$. Then*

$$\tau_{\Phi_* A}(\Phi_*^- o^-, \Phi_*^+ o^+) = \Phi_* \tau_A(o^-, o^+)$$

for all $A^\pm \in \mathcal{S}(W,H)$, $A \in \mathcal{P}(A^-, A^+)$, and $o^\pm \in \mathrm{Or}(A^\pm)$.

(Direct Sum) *If $A_j^\pm \in \mathcal{S}(W_j, H_j)$ and $A_j \in \mathcal{P}(A_j^-, A_j^+)$ for $j = 0, 1$ then*

$$\tau_{A_0 \oplus A_1}(o_0^- \otimes o_1^-, o_0^+ \otimes o_1^+) = \tau_{A_0}(o_0^-, o_0^+) \otimes \tau_{A_1}(o_1^-, o_1^+).$$

for all $o_j^\pm \in \mathrm{Or}(A_j^\pm)$.

(Catenation) *Let $A_0, A_1, A_2 \in \mathcal{S}(W,H)$, suppose that $A_{01} \in \mathcal{P}(A_0, A_1)$ and $A_{12} \in \mathcal{P}(A_1, A_2)$ satisfy (103) and, for $R > T$, define A_{02}^R by (104). Assume $\mathcal{D}_{A_{01}}$ and $\mathcal{D}_{A_{12}}$ are onto. Then $\mathcal{D}_{A_{02}^R}$ is onto for large R and*

$$\tau_{A_{02}}(o_0, o_2) = \sigma^R\left(\tau_{A_{01}}(o_0, o_1), \tau_{A_{12}}(o_1, o_2)\right).$$

for $o_0 \in \mathrm{Or}(A_0)$, $o_1 \in \mathrm{Or}(A_1)$, and $o_2 \in \mathrm{Or}(A_2)$. Here the map

$$\sigma^R : \det(\mathcal{D}_{A_{01}}) \times \det(\mathcal{D}_{A_{12}}) \to \det(\mathcal{D}_{A_{02}^R})$$

is induced by the isomorphism S^R of Remark 11.3.

(Constant) *If $A(s) \equiv A^+ = A^-$ and $o^+ = o^- \in \mathrm{Or}(A^\pm)$ then $\tau_A(o^-, o^+)$ is the standard orientation of $\det(\mathcal{D}_A) \cong \mathbb{R}$.*

(Normalization) *If $W = H = \mathbb{R}^n$ then τ_A is the map defined in Remark 11.2.*

The maps τ_A are uniquely determined by the (Homotopy), (Direct Sum), (Constant), and (Normalization) axioms.

Theorem 11.4 is standard with the techniques of [9] (although the assumptions are not quite the same as in the work of Floer and Hofer).

Proof of Theorem 1.1. Assume $\mathcal{S}_\mathcal{V}$ is Morse–Smale. For $x \in \mathcal{P}(\mathcal{V})$ denote by $W^u(x)$ the unstable manifold of x with respect to the negative gradient flow of $\mathcal{S}_\mathcal{V}$. Thus $W^u(x)$ is the space of all smooth loops $y : S^1 \to M$ such that there exists a solution $u : (-\infty, 0] \times S^1 \to M$ of the nonlinear heat equation (12) that converges to x as $s \to -\infty$ and satisfies $u(0,t) = y(t)$. Then $W^u(x)$ is a finite dimensional manifold (see for example [3]). It is diffeomorphic to $\mathbb{R}^k$ where $k = \mathrm{ind}_\mathcal{V}(x)$ is the Morse index of x as a critical point of $\mathcal{S}_\mathcal{V}$. Fix an orientation of $W^u(x)$ for every periodic orbit $x \in \mathcal{P}(\mathcal{V})$. These orientations determine a system of coherent orientations for the heat flow as follows.

Fix a pair $x^\pm \in \mathcal{P}(\mathcal{V})$ of periodic orbits that represent the same component of $\mathcal{L}M$. Denote by $\mathcal{P}^0(x^-, x^+)$ the set of smooth maps $u : \mathbb{R} \times S^1 \to M$ such that $u(s,\cdot)$ converges to $x^\pm$ in the C^2 norm and $\partial_s u(s,\cdot)$ converges to zero in the C^1 norm as s tends to $\pm\infty$. Then, in a suitable trivialization of the tangent bundle u^*TM, the linearized operator $\mathcal{D}_u^0$ has the form of an operator $\mathcal{D}_A$ as in Theorem 11.4 where the spaces $E(A^\pm)$ correspond to the tangent spaces $T_{x^\pm}W^u(x^\pm)$ of the unstable manifolds. Hence, by Theorem 11.4, the given orientations of the unstable manifolds determine orientations

$$\nu^0(u) \in \mathrm{Or}(\det(\mathcal{D}_u^0))$$

of the determinant lines for all $u \in \mathcal{P}^0(x^-, x^+)$ and all $x^\pm \in \mathcal{P}(\mathcal{V})$. By the *(Naturality)* axiom, these orientations are independent of the choice of the trivializations used to define them. By the *(Catenation)* axiom, they form a system of *coherent orientations* in the sense of Floer–Hofer [9].

Next we show how the coherent orientations for the heat flow induce a system of coherent orientations

$$\nu^\varepsilon(u,v) \in \mathrm{Or}(\det(\mathcal{D}^\varepsilon_{u,v}))$$

for the Floer equations (11). Let us denote by $\mathcal{P}(x^-,x^+)$ the set of smooth maps $(u,v):\mathbb{R}\times S^1 \to TM$ such that $(u(s,\cdot),v(s,\cdot))$ converges to $(x^\pm,\dot{x}^\pm)$ in the C^1 norm and $(\partial_s u, \nabla_s v)$ converges to zero, uniformly in t, as s tends to $\pm\infty$. By the obvious homotopy arguments it suffices to assume $u \in \mathcal{P}^0(x^-,x^+)$ and $v = \partial_t u$. We abbreviate

$$\mathcal{D}^\varepsilon_u := \mathcal{D}^\varepsilon_{u,\partial_t u}.$$

It follows from the definition of the operators in (14) that

$$\mathcal{D}^0_u \xi = 0 \qquad \Longrightarrow \qquad \mathcal{D}^\varepsilon_u \begin{pmatrix} \xi \\ \nabla_t \xi \end{pmatrix} = \begin{pmatrix} 0 \\ \nabla_s \nabla_t \xi + R(\xi,\partial_s u)\partial_t u \end{pmatrix}.$$

Hence $\mathcal{D}^\varepsilon_u(\xi,\nabla_t\xi)$ is small in the $(0,2,\varepsilon)$-norm. If the operator $\mathcal{D}^0_u$ is onto then the estimate of Theorem 3.3 shows that the operator $\mathcal{D}^\varepsilon_u$ is onto as well. As both operators have the same index, their kernels have the same dimension. Hence, again by Theorem 3.3, the map

$$\ker\, \mathcal{D}^0_u \to \ker\, \mathcal{D}^\varepsilon_u : \xi \mapsto \begin{pmatrix} \xi \\ \nabla_t \xi \end{pmatrix} - \mathcal{D}^{\varepsilon\,*}_u \left(\mathcal{D}^\varepsilon_u \mathcal{D}^{\varepsilon\,*}_u\right)^{-1} \mathcal{D}^\varepsilon_u \begin{pmatrix} \xi \\ \nabla_t \xi \end{pmatrix} \tag{105}$$

is an isomorphism between the kernels and we define $\nu^\varepsilon(u,\partial_t u)$ to be the image of $\nu^0(u)$ under the induced isomorphism of the top exterior powers. If $\mathcal{D}^0_u$ is not onto we obtain a similar isomorphism between the determinant lines of $\mathcal{D}^0_u$ and $\mathcal{D}^\varepsilon_u$ by augmenting the operators first to make them surjective. That the resulting orientations $\nu^\varepsilon(u,\partial_t u)$ of the operators $\mathcal{D}^\varepsilon_u$ satisfy the *(Catenation)* axiom follows from a linear version of the standard pregluing construction; namely an *approximate* basis of the kernel of the glued operator can be obtained from pairs of basis elements of the two operators involved in the gluing construction. This gives an isomorphism

$$\ker \mathcal{D}^0_{u_{01}} \times \ker \mathcal{D}^0_{u_{12}} \to \ker \mathcal{D}^0_{u_{02}}$$

and similarly for ε. These isomorphisms commute with the isomorphisms (105) up to small perturbations of order ε. Hence the orientations $\nu^\varepsilon(u,v)$ satisfy the *(Catenation)* axiom and so form a system of coherent orientations for the Floer equations.

Now assume that $x^\pm \in \mathcal{P}(\mathcal{V})$ have Morse index difference one. Consider the map

$$\mathcal{T}^\varepsilon : \mathcal{M}^0(x^-,x^+;\mathcal{V}) \to \mathcal{M}^\varepsilon(x^-,x^+;\mathcal{V})$$

of Definition 4.4 and recall that, by Theorem 10.1, it is bijective. It follows from the proof of Theorem 9.1 that the map $\mathcal{T}^\varepsilon$ satisfies the following. Let $u \in \mathcal{M}^0(x^-,x^+;\mathcal{V})$ and

$$(u^\varepsilon,v^\varepsilon) := \mathcal{T}^\varepsilon(u) \in \mathcal{M}^\varepsilon(x^-,x^+;\mathcal{V}).$$

Then the vector $\partial_s u \in \ker \mathcal{D}_u^0$ is positively oriented with respect to $\nu^0(u)$ if and only if the vector $(\partial_s u^\varepsilon, \nabla_s v^\varepsilon) \in \ker \mathcal{D}^\varepsilon_{u^\varepsilon, v^\varepsilon}$ is positively oriented with respect to $\nu^\varepsilon(u^\varepsilon, v^\varepsilon)$. In other words, fix a positive generator $Z \in \ker \mathcal{D}_u^0$ and let $Z^\varepsilon \in \ker \mathcal{D}^\varepsilon_{u^\varepsilon, v^\varepsilon}$ be its image under the composition

$$\ker \mathcal{D}_u^0 \to \ker \mathcal{D}^\varepsilon_{u, \partial_t u} \to \ker \mathcal{D}^\varepsilon_{u^\varepsilon, v^\varepsilon}$$

of the isomorphism (105) with a parallel transport map. Then Z^ε determines the orientation $\nu^\varepsilon(u^\varepsilon, v^\varepsilon)$ of the kernel of $\mathcal{D}^\varepsilon_{u^\varepsilon, v^\varepsilon}$. Moreover, the first component dominates the ε-inner product of $(\partial_s u^\varepsilon, \nabla_s v^\varepsilon)$ with Z^ε which therefore has the same sign as the L^2-inner product of $\partial_s u$ with Z. This brief sketch shows that the bijection $\mathcal{T}^\varepsilon$ preserves the signs for the definitions of the two boundary operators. Hence the Morse complex of the heat flow has the same boundary operator as the Floer complex for ε sufficiently small. Hence the resulting homologies are naturally isomorphic, i.e. for every regular value a of $\mathcal{S}_\mathcal{V}$ there is a constant $\varepsilon_0 > 0$ such that

$$\mathrm{HF}^a_*(T^*M, \mathcal{V}, \varepsilon; \mathbb{Z}) \cong \mathrm{HM}^a_*(\mathcal{L}M, \mathcal{S}_\mathcal{V}; \mathbb{Z})$$

for $0 < \varepsilon \le \varepsilon_0$. In fact, we have established this isomorphism on the chain level and with integer coefficients. To complete the proof of Theorem 1.1 one can now argue as in the proof of Theorem 11.1 to show that, given a potential V such that $\mathcal{S}_V$ is Morse and a regular value a of $\mathcal{S}_V$, we have two isomorphisms

$$\mathrm{HF}^a_*(T^*M, H_V; \mathbb{Z}) \cong \mathrm{HF}^a_*(T^*M, \mathcal{V}, \varepsilon; \mathbb{Z})$$

and

$$\mathrm{HM}^a_*(\mathcal{L}M, \mathcal{S}_\mathcal{V}; \mathbb{Z}) \cong \mathrm{H}_*(\{\mathcal{S}_V \le a\}; \mathbb{Z})$$

for a suitable perturbation $\mathcal{V}$ and $\varepsilon > 0$ sufficiently small. This proves the result for integer coefficients and $a < \infty$. The argument for general coefficient rings is exactly the same.

To prove the result for $a = \infty$ we observe that, by construction, our isomorphisms $\mathrm{HF}^a_*(T^*M, H_V; \mathbb{Z}) \cong \mathrm{H}_*(\{\mathcal{S}_V \le a\}; \mathbb{Z})$ intertwine the homomorphisms

$$\mathrm{HF}^a_*(T^*M, H_V; \mathbb{Z}) \to \mathrm{HF}^b_*(T^*M, H_V; \mathbb{Z}), \qquad \mathrm{H}_*(\{\mathcal{S}_V \le a\}) \to \mathrm{H}_*(\{\mathcal{S}_V \le b\})$$

for $a < b$. Hence the result for $a = \infty$ follows by taking the direct limit $a \to \infty$ and noting that there are natural isomorphisms

$$\mathrm{HF}_*(T^*M, H_V) \cong \varinjlim_{a \in \mathbb{R}} \mathrm{HF}^a_*(T^*M, H_V)$$

and

$$\mathrm{H}_*(\mathcal{L}M) \cong \varinjlim_{a \in \mathbb{R}} \mathrm{H}_*(\{\mathcal{S}_V \le a\}).$$

This proves Theorem 1.1. $\square$

A The heat flow

In this appendix we summarize results from [31] that are used in this paper. We assume throughout this appendix that M is a closed Riemannian manifold. Let $\mathcal{V} : \mathcal{L}M \to \mathbb{R}$ be a smooth function that satisfies the axioms $(V0 - V4)$. Consider the action functional

$$\mathcal{S}_{\mathcal{V}}(x) = \frac{1}{2}\int_0^1 |\dot{x}(t)|^2 \, dt - \mathcal{V}(x)$$

and the corresponding heat equation

$$\partial_s u - \nabla_t \partial_t u - \operatorname{grad} \mathcal{V}(u) = 0 \tag{106}$$

for smooth functions $\mathbb{R} \times S^1 \to M : (s,t) \mapsto u(s,t)$. In the following we denote by $\mathcal{P}(\mathcal{V}) \subset C^\infty(S^1, M)$ the set of critical points x of $\mathcal{S}_{\mathcal{V}}$ (i.e. of solutions of the equation $\nabla_t \dot{x} + \operatorname{grad} \mathcal{V}(x) = 0$), and by $\mathcal{P}^a(\mathcal{V})$ the set of all $x \in \mathcal{P}(\mathcal{V})$ with action $\mathcal{S}_{\mathcal{V}}(x) \leq a$. For two nondegenerate critical points $x^\pm \in \mathcal{P}(\mathcal{V})$ we denote by $\mathcal{M}^0(x^-, x^+; \mathcal{V})$ the set of all solutions u of (106) that converge to $x^\pm(t)$ as $s \to \pm\infty$. The energy of such a solution is given by

$$E(u) := \int_{-\infty}^{\infty} \int_0^1 |\partial_s u|^2 \, dtds = \mathcal{S}_{\mathcal{V}}(x^-) - \mathcal{S}_{\mathcal{V}}(x^+). \tag{107}$$

Theorem A.1 (Apriori estimates). *Fix a perturbation $\mathcal{V} : \mathcal{L}M \to \mathbb{R}$ that satisfies $(V0 - V1)$ and a constant $c_0 > 0$. Then there is a constant $C = C(c_0, \mathcal{V}) > 0$ such that the following holds. If $u : \mathbb{R} \times S^1 \to M$ is a solution of (106) such that $\mathcal{S}_{\mathcal{V}}(u(s,\cdot)) \leq c_0$ for every $s \in \mathbb{R}$ then*

$$\|\partial_s u\|_\infty + \|\partial_t u\|_\infty + \|\nabla_t \partial_t u\|_\infty \leq C.$$

Theorem A.2 (Exponential decay). *Fix a perturbation $\mathcal{V} : \mathcal{L}M \to \mathbb{R}$ that satisfies $(V0 - V4)$ and assume $\mathcal{S}_{\mathcal{V}}$ is Morse.*

(F) *Let $u : [0,\infty) \times S^1 \to M$ be a solution of (106). Then there are positive constants ρ and $c_1, c_2, c_3, \ldots$ such that*

$$\|\partial_s u\|_{C^k([T,\infty)\times S^1)} \leq c_k e^{-\rho T}$$

for every $T \geq 1$. Moreover, there is a periodic orbit $x \in \mathcal{P}(\mathcal{V})$ such that $u(s,t)$ converges to $x(t)$ as $s \to \infty$.

(B) *Let $u : (-\infty, 0] \times S^1 \to M$ be a solution of (106) with finite energy. Then there are positive constants ρ and $c_1, c_2, c_3, \ldots$ such that*

$$\|\partial_s u\|_{C^k((-\infty,-T]\times S^1)} \leq c_k e^{-\rho T}$$

for every $T \geq 1$. Moreover, there is a periodic orbit $x \in \mathcal{P}(\mathcal{V})$ such that $u(s,t)$ converges to $x(t)$ as $s \to -\infty$.

Theorem A.3 (Regularity). *Fix a constant $p > 2$ and a perturbation $\mathcal{V} : \mathcal{L}M \to \mathbb{R}$ that satisfies $(V0-V4)$. Let $u : \mathbb{R}\times S^1 \to M$ be a continuous function which is locally of class $W^{1,p}$. Assume further that u is a weak solution of (106). Then u is smooth.*

The covariant Hessian of $\mathcal{S}_\mathcal{V}$ at a loop $x : S^1 \to M$ is the operator $A(x) : W^{2,2}(S^1, x^*TM) \to L^2(S^1, x^*TM)$, given by

$$A(x)\xi := -\nabla_t\nabla_t\xi - R(\xi, \dot{x})\dot{x} - \mathcal{H}_\mathcal{V}(x)\xi.$$

This operator is self-adjoint with respect to the standard L^2 inner product on $\Omega^0(S^1, x^*TM)$. In this notation the linearized operator $\mathcal{D}_u^0 : \mathcal{W}_u^p \to \mathcal{L}_u^p$ is given by

$$\mathcal{D}_u^0\xi := \nabla_s\xi + A(u_s)\xi$$

where $u_s(t) := u(s,t)$. (See Section 3 for the definition of the spaces $\mathcal{W}_u = \mathcal{W}_u^p$ and $\mathcal{L}_u = \mathcal{L}_u^p$.)

Theorem A.4 (Fredholm). *Fix a perturbation $\mathcal{V} : \mathcal{L}M \to \mathbb{R}$ that satisfies $(V0-V4)$ and assume $\mathcal{S}_\mathcal{V}$ is Morse. Let $x^\pm \in \mathcal{P}(\mathcal{V})$ and $u : \mathbb{R}\times S^1 \to M$ be a smooth map such that $u(s,\cdot)$ converges to $x^\pm$ in the C^2 norm and $\partial_s u$ converges uniformly to zero as $s \to \pm\infty$. Then, for every $p > 1$, the operator $\mathcal{D}_u^0 : \mathcal{W}_u^p \to \mathcal{L}_u^p$ is Fredholm and its Fredholm index is given by*

$$\operatorname{index}\mathcal{D}_u^0 = \operatorname{ind}_\mathcal{V}(x^-) - \operatorname{ind}_\mathcal{V}(x^+).$$

Here $\operatorname{ind}_\mathcal{V}(x^\pm)$ denotes the Morse index of $x^\pm$, i.e. the number of negative eigenvalues of $A(x^\pm)$.

Theorem A.5 (Implicit function theorem). *Fix a perturbation $\mathcal{V} : \mathcal{L}M \to \mathbb{R}$ that satisfies $(V0-V4)$. Assume $\mathcal{S}_\mathcal{V}$ is Morse and that $\mathcal{D}_u^0$ is onto for every $u \in \mathcal{M}^0(x^-, x^+; \mathcal{V})$ and every pair $x^\pm \in \mathcal{P}^a(\mathcal{V})$. Fix two critical points $x^\pm \in \mathcal{P}^a(\mathcal{V})$ with Morse index difference one. Then, for all $c_0 > 0$ and $p > 2$, there exist positive constants δ_0 and c such that the following holds. If $u : \mathbb{R}\times S^1 \to M$ is a smooth map such that $\lim_{s\to\pm\infty} u(s,\cdot) = x^\pm(\cdot)$ exists, uniformly in t, and such that*

$$|\partial_s u(s,t)| \le \frac{c_0}{1+s^2}, \qquad |\partial_t u(s,t)| \le c_0$$

for all $(s,t) \in \mathbb{R}\times S^1$ and

$$\|\partial_s u - \nabla_t\partial_t u - \operatorname{grad}\mathcal{V}(u)\|_p \le \delta_0.$$

Then there exist elements $u_0 \in \mathcal{M}^0(x^-, x^+; \mathcal{V})$ and $\xi \in \operatorname{im}(\mathcal{D}_{u_0}^0)^* \cap \mathcal{W}_{u_0}$ satisfying

$$u = \exp_{u_0}(\xi), \qquad \|\xi\|_{\mathcal{W}_{u_0}} \le c\,\|\partial_s u - \nabla_t\partial_t u - \operatorname{grad}\mathcal{V}(u)\|_p.$$

Theorem A.6 (Transversality). *For a generic perturbation $\mathcal{V} : \mathcal{L}M \to \mathbb{R}$ satisfying $(V0-V4)$ the function $\mathcal{S}_\mathcal{V} : \mathcal{L}M \to \mathbb{R}$ is* **Morse–Smale** *in the sense that every critical point x of $\mathcal{S}_\mathcal{V}$ is nondegenerate (i.e. the Hessian $A(x)$ is bijective) and every finite energy solution $u : \mathbb{R}\times S^1 \to M$ of (106) is regular (i.e. the Fredholm operator $\mathcal{D}_u^0$ is surjective).*

Theorem A.7. *Let $\mathcal{V} : \mathcal{L}M \to \mathbb{R}$ be a perturbation that satisfies $(V0 - V4)$ and assume that $\mathcal{S}_\mathcal{V}$ is Morse–Smale. Then, for every regular value a of $\mathcal{S}_\mathcal{V}$ and every principal ideal domain R, there is a natural isomorphism*

$$\mathrm{HM}_*^a(\mathcal{L}M, \mathcal{S}_\mathcal{V}; R) \cong \mathrm{H}_*(\mathcal{L}^a M; R), \qquad \mathcal{L}^a M := \{x \in \mathcal{L}M \mid \mathcal{S}_\mathcal{V}(x) \leq a\}.$$

If M is not simply connected then there is a separate isomorphism for each component of the loop space. The isomorphism commutes with the homomorphisms $\mathrm{HM}_*^a(\mathcal{L}M, \mathcal{S}_\mathcal{V}) \to \mathrm{HM}_*^b(\mathcal{L}M, \mathcal{S}_\mathcal{V})$ and $\mathrm{H}_*(\mathcal{L}^a M) \to \mathrm{H}_*(\mathcal{L}^b M)$ for $a < b$.

Comments on the proofs

The apriori estimate. A uniform bound on $\int_0^1 |\partial_t u|^2\, dt$ is obvious, by axiom $(V0)$ and the fact that $\mathcal{S}_\mathcal{V}(u)$ decreases along the solutions of equation (106). Using the estimate (35) with $\varepsilon = 0$ one then gets

$$(\partial_t^2 - \partial_s)|\partial_t u|^2 \geq -\mu|\partial_t u|^2 - 1.$$

One can now use Lemma B.1 to obtain the pointwise estimate.

Compactness, regularity, and exponential decay. For compactness and Fredholm theory parabolic regularity theorems play a central role - a good reference is [15]. The exponential decay estimates in the standard Floer theory (eg [20, 22]) carry over to the parabolic case. The proof involves parabolic (instead of elliptic) regularity theorems.

Fredholm theory. In a suitable Hilbert space setting the linearized operator is of the Atiyah-Patodi-Singer type, and hence is Fredholm (in the Morse case) - see for example [19]. As far as the Fredholm index is concerned, the story is slightly simpler than in some other comparable situations, as all the critical points have finite Morse indices (and their differences are the Fredholm indices).

The implicit function theorem A.5 now says that an approximate solution with surjective Fredholm operator is close to a true solution. The proof here is analogous to similar theorems in the literature (e.g. [16, Theorem 3.5.2]). There are some additional subtleties in the parabolic case that were resolved by the second author in an unpublished manuscript on which [31] will be based. The subtleties are related to the appearence of terms of the form $|\partial_t u|^2$ in the quadratic estimate.

Transversality. This is the first of the two main points. The perturbations in Section 2 were introduced precisely to make the transversality theory work. The key issue is unique continuation: two solutions of equation (106) which agree at time $s = s_0$ agree for all time. The proof is based on the beautiful Agmon-Nirenberg technique (for the linearized equation differentiate $\log\|x(s)\|$ when $\dot{x}(s) + A(s)x(s) = y(s)$). This argument also works in backward time.

Now one can apply the usual Thom–Smale transversality theory. There's a universal moduli space of all pairs $(\mathcal{V}, u)$ such that u satisfies equation (106). The key is to construct a suitable Hilbert space of perturbations of the type discussed in Section 2. This can be done by using a countable family of perturbations in the specific form of Remark 2.1. Then one has to prove that this universal moduli space is indeed a Hilbert manifold, and that is where the cutoff functions in Remark 2.1 will play the crucial role, which lead to global perturbations, in the sense that they depend on the whole loop, but they can be localized near any given specific loop.

With this understood the rest of the argument is standard: the regular values of the projection $(\mathcal{V}, u) \mapsto \mathcal{V}$ on the universal moduli space are the desired regular perturbations for which the gradient flow of $\mathcal{S}_{\mathcal{V}}$ is Morse–Smale.

Once transversality is established one can define the chain complex by counting the gradient flow lines of $\mathcal{S}_{\mathcal{V}}$.

Computing the Morse homology of $\mathcal{S}_{\mathcal{V}}$. This is the second main point. The homology of the loop space can be computed via Morse theory - from a countable cell complex with one cell for each perturbed closed geodesic (see [18] for the path case). The difference between the present approach and the standard theory is that we are led by the adiabatic limit analysis to consider the L^2-gradient flow of $\mathcal{S}_{\mathcal{V}}$ (and hence a parabolic pde) while the standard theory works with the $W^{1,2}$-gradient flow (and hence with a potentially simpler ode - though still in an infinite dimensional setting).

Let us assume that the critical points of $\mathcal{S}_{\mathcal{V}}$ all have different critical values, and let c be such a critical value with a critical point x of Morse index k. Then the relative homology of the subsets $\mathcal{L}^{c+\varepsilon} := \{\mathcal{S}_{\mathcal{V}} \le c+\varepsilon\}$ and $\mathcal{L}^{c-\varepsilon}$ is one copy of $\mathbb{Z}$ in dimension k (see [18]). Let us assume for simplicity that the critical point on the next higher critical level c' has a critical point x' of Morse index $k+1$. Then there's a boundary map

$$\mathbb{Z} = \mathrm{H}_{k+1}(\mathcal{L}^{c'+\varepsilon}, \mathcal{L}^{c+\varepsilon}) \longrightarrow \mathrm{H}_k(\mathcal{L}^{c+\varepsilon}, \mathcal{L}^{c-\varepsilon}) = \mathbb{Z} \tag{108}$$

determined by a single integer. The key point is now to prove that this integer agrees with the algebraic number of solutions of equation (106) connecting x' to x. Finite dimensional versions of this argument can be found in [17, 7, 21].

A relevant issue in the adaptation of this argument to the infinite dimensional case is the observation that Morse critical points of parabolic gradient flows have (finite dimensional) unstable manifolds. A proof which follows the finite dimensional case can be found in [3], for the Yang–Mills functional over Riemann surfaces, but the argument remains valid in greater generality.

If two critical points of index difference 1 do not lie on adjacent critical levels, one has to replace the sublevel sets by appropriate cells, either following Milnor's book [18], or adapting Conley's construction of index pairs to parabolic pdes.

Once it has been established that the boundary map (108) is given by counting connecting orbits, it follows by standard arguments in homology theory that the operator obtained by counting the solutions of (106) defines indeed a chain complex, and that its homology is isomorphic to the singular homology of the loop space. (See [17] and also [21, 26, 29] for a finite dimensional version of this argument.)

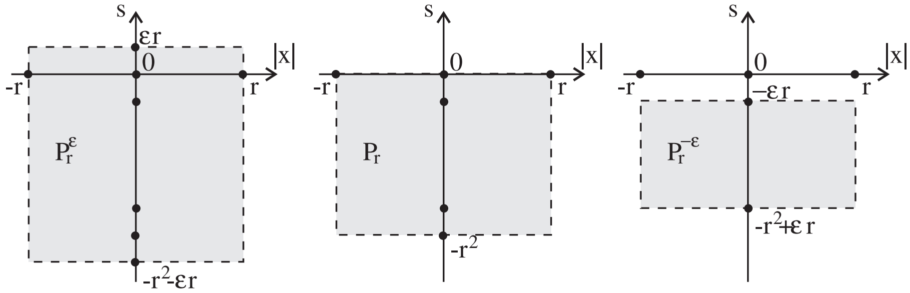

Figure 1: Parabolic cylinders

B Mean value inequalities

Let n be a positive integer and denote by

$$\Delta := {\partial_1}^2 + \cdots + {\partial_n}^2$$

the standard Laplacian on $\mathbb{R}^n$. Given positive real numbers r and ε let $B_r = B_r(0)$ be the open ball of radius r in $\mathbb{R}^n$ and define the *parabolic cylinders* $P_r, P_r^\varepsilon, P_r^{-\varepsilon} \subset \mathbb{R}^{n+1}$ by

$$\begin{aligned} P_r^\varepsilon &:= (-r^2 - \varepsilon r, \varepsilon r) \times B_r, \\ P_r &:= (-r^2, 0) \times B_r, \\ P_r^{-\varepsilon} &:= (-r^2 + \varepsilon r, -\varepsilon r) \times B_r. \end{aligned}$$

(See Figure 1.) For $r \le 2\varepsilon$ we simply define $P_r^{-\varepsilon} := \emptyset$. The elements of P_r are denoted by $(s, x) = (s, x_1, \ldots, x_n)$.

Lemma B.1. *For every $n \in \mathbb{N}$ there is a constant $c_n > 0$ such that the following holds for every $r \in (0, 1]$. If $a \ge 0$ and $w : \mathbb{R} \times \mathbb{R}^n \supset P_r \to \mathbb{R}$ is C^1 in the s-variable and C^2 in the x-variable such that*

$$(\Delta - \partial_s) w \ge -aw, \qquad w \ge 0,$$

then

$$w(0) \le \frac{c_n e^{ar^2}}{r^{n+2}} \int_{P_r} w.$$

Proof. For $a = 0$ this is a special case of a theorem by Gruber for parabolic differential operators with variable coefficients. (See Gruber [12, Theorem 2.1] with $p = 1$, $\theta = 1$, $\lambda = 1$, $\sigma = 1/2$, $R = r$ and $f = 0$; for an another proof see Lieberman [14, Theorem 7.21] with $R = r$, $p = 1$, $\rho = 1/2$, $f = 0$.)

To prove the result in general assume that w satisfies the hypotheses of the lemma and define $f(s, x) := e^{-as} w(s, x)$. Then

$$(\Delta - \partial_s) f = e^{-as} (\Delta - \partial_s + a) w \ge 0.$$

Hence, by Gruber's theorem,

$$w(0)=f(0)\le\frac{c_n}{r^{n+2}}\int_{P_r}f\le\frac{c_ne^{ar^2}}{r^{n+2}}\int_{P_r}w.$$

This proves the lemma. $\square$

Lemma B.2. *Let c_2 be the constant in Lemma B.1 with $n=2$. Let $\varepsilon\ge 0$, $r\in(0,1]$, and $a\ge 0$. If $w:\mathbb{R}\times\mathbb{R}\supset P_r^\varepsilon\to\mathbb{R}$ is C^1 in the s-variable and C^2 in the t-variable and satisfies*

$$L_\varepsilon w:=\left(\varepsilon^2{\partial_s}^2+{\partial_t}^2-\partial_s\right)w\ge -aw,\qquad w\ge 0,\tag{109}$$

then

$$w(0)\le\frac{2c_2e^{ar^2}}{r^3}\int_{P_r^\varepsilon}w.$$

Proof. The idea of proof was suggested to us by Tom Ilmanen. Define a function W on the domain $P_r\subset\mathbb{R}\times\mathbb{R}^2$ by

$$W(s,t,q):=w(s+\varepsilon q,t).$$

(Note that $(s+\varepsilon q,t)\in P_r^\varepsilon\subset\mathbb{R}\times\mathbb{R}$ for every $(s,t,q)\in P_r\subset\mathbb{R}\times\mathbb{R}^2$.) Then, by assumption, we have

$$\left(\Delta-\partial_s\right)W(s,t,q)=\left(L_\varepsilon w\right)(s+\varepsilon q,t)\ge -aw(s+\varepsilon q,t)=-aW(s,t,q),$$

where $\Delta:=\partial_t^2+\partial_q^2$. Hence it follows from Lemma B.1 with $n=2$ that

$$w(0)=W(0)\le\frac{c_2e^{ar^2}}{r^4}\int_{P_r}W.$$

It remains to estimate the integral on the right hand side:

$$\begin{aligned}\int_{P_r}W&\le\int_{-r}^{r}\int_{-r}^{r}\int_{-r^2}^{0}W(s,t,q)\,dsdqdt\\&=\int_{-r}^{r}\int_{-r}^{r}\int_{-r^2+\varepsilon q}^{\varepsilon q}w(z,t)\,dzdqdt\\&\le\int_{-r}^{r}\int_{-r}^{r}\int_{-r^2-\varepsilon r}^{\varepsilon r}w(z,t)\,dzdqdt\\&=2r\int_{P_r^\varepsilon}w.\end{aligned}\tag{110}$$

The first step uses the fact that $W\ge 0$ and $B_r\subset[-r,r]\times[-r,r]$. The third step uses the fact that $w\ge 0$ and $(-r^2+\varepsilon q,\varepsilon q)\subset(-r^2-\varepsilon r,\varepsilon r)$, since $0\le q\le r$. $\square$

Lemma B.3. *Fix three constants $r \in (0,1]$, $\varepsilon \geq 0$, and $\mu \geq 0$. Let c_2 be the constant of Lemma B.1. If $f : [-r^2 - \varepsilon r, \varepsilon r] \to \mathbb{R}$ is a C^2 function satisfying*

$$\varepsilon^2 f'' - f' + \mu f \geq 0, \qquad f \geq 0,$$

then

$$f(0) \leq \frac{4c_2 e^{\mu r^2}}{r^2} \int_{-r^2-\varepsilon r}^{\varepsilon r} f(s)\, ds.$$

Proof. This follows immediately from Lemma B.2 with $w(s,t) := f(s)$. □

Lemma B.4. *Let $R, r > 0$ and $u : \mathbb{R} \times \mathbb{R}^n \supset P_{R+r} \to \mathbb{R}$ be C^1 in the s-variable and C^2 in the x-variable and $f, g : P_{R+r} \to \mathbb{R}$ be continous functions such that*

$$(\Delta - \partial_s)\, u \geq g - f, \qquad u \geq 0, \qquad f \geq 0, \qquad g \geq 0.$$

Then

$$\int_{P_R} g \leq \int_{P_{R+r}} f + \left(\frac{4}{r^2} + \frac{1}{Rr}\right) \int_{P_{R+r} \setminus P_R} u.$$

Proof. The proof rests on the following two inequalities. Let $B_r \subset \mathbb{R}^n$ be the open ball of radius r centered at zero. Then, for every smooth function $u : \mathbb{R}^n \to [0, \infty)$, we have

$$\int_{\partial B_r} \frac{\partial u}{\partial \nu} = -\frac{n-1}{r} \int_{\partial B_r} u + \frac{d}{dr} \int_{\partial B_r} u \leq \frac{d}{dr} \int_{\partial B_r} u \tag{111}$$

(see [13, Theorem 2.1]). Secondly, every smooth function $u : \mathbb{R} \times \mathbb{R}^n \to [0, \infty)$ satisfies

$$\begin{aligned}
\frac{d}{d\sigma} \int_{-(R+\sigma)^2}^{0} \left(\int_{\partial B_{R+\sigma}} u(s, \cdot) \right) ds &= \int_{-(R+\sigma)^2}^{0} \left(\frac{d}{d\sigma} \int_{\partial B_{R+\sigma}} u(s, \cdot) \right) ds \\
&\quad + 2(R+\sigma) \int_{\partial B_{R+\sigma}} u(-(R+\sigma)^2, \cdot) \qquad (112) \\
&\geq \int_{-(R+\sigma)^2}^{0} \left(\frac{d}{d\sigma} \int_{\partial B_{R+\sigma}} u(s, \cdot) \right) ds.
\end{aligned}$$

Now suppose u, f, g satisfy the assumptions of the lemma. Then, for $0 \leq \sigma \leq r$,

$$\begin{aligned}
&\int_{P_R} g - \int_{P_{R+r}} f \\
&\leq \int_{P_{R+\sigma}} (\Delta u - \partial_s u) \\
&= \int_{-(R+\sigma)^2}^{0} \left(\int_{\partial B_{R+\sigma}} \frac{\partial u}{\partial \nu}(s, \cdot) \right) ds - \int_{B_{R+\sigma}} \Big(u(0, \cdot) - u(-(R+\sigma)^2, \cdot) \Big) dx \\
&\leq \int_{-(R+\sigma)^2}^{0} \left(\frac{d}{d\sigma} \int_{\partial B_{R+\sigma}} u(s, \cdot) \right) ds + \int_{B_{R+\sigma}} u(-(R+\sigma)^2, x)\, dx \\
&\leq \frac{d}{d\sigma} \int_{-(R+\sigma)^2}^{0} \left(\int_{\partial B_{R+\sigma}} u(s, \cdot) \right) ds + \int_{B_{R+\sigma}} u(-(R+\sigma)^2, x)\, dx.
\end{aligned}$$

Here the first step uses the inclusions $P_R \subset P_{R+\sigma} \subset P_{R+r}$. The third step follows from (111) and the last from (112). Now integrate this inequality over the interval $0 \le \sigma \le t$, with $r/2 \le t \le r$, to obtain

$$\begin{aligned}
&\frac{r}{2}\left(\int_{P_R} g - \int_{P_{R+r}} f\right) \\
&\le \int_{-(R+t)^2}^{0} \left(\int_{\partial B_{R+t}} u(s,\cdot)\right) ds + \int_0^r \left(\int_{B_{R+\sigma}} u(-(R+\sigma)^2,\cdot)\right) d\sigma \\
&\le \int_{-(R+r)^2}^{0} \left(\int_{\partial B_{R+t}} u(s,\cdot)\right) ds + \frac{1}{2R}\int_{-(R+r)^2}^{-R^2}\int_{B_{R+r}} u(s,x)\,dxds.
\end{aligned}$$

Here the last step follows by substituting $s = -(R+\sigma)^2$. Integrate this inequality again over the interval $r/2 \le t \le r$ to obtain

$$\frac{r}{2}\left(\int_{P_R} g - \int_{P_{R+r}} f\right) \le \frac{2}{r}\int_{P_{R+r}\setminus P_R} u + \frac{1}{2R}\int_{P_{R+r}\setminus P_R} u.$$

This proves Lemma B.4. $\square$

Lemma B.5. *Let ε, R, r be positive real numbers. Let $u : \mathbb{R}^2 \supset P^{\varepsilon}_{R+r} \to \mathbb{R}$ be a C^2 function and $f, g : P^{\varepsilon}_{R+r} \to \mathbb{R}$ be continous functions such that*

$$\left(\varepsilon^2 {\partial_s}^2 + {\partial_t}^2 - \partial_s\right) u \ge g - f, \qquad u \ge 0, \qquad f \ge 0, \qquad g \ge 0.$$

Then

$$\int_{P^{-\varepsilon}_{R/2}} g \le 2\left(1 + \frac{r}{R}\right)\left(\int_{P^{\varepsilon}_{R+r}} f + \left(\frac{4}{r^2} + \frac{1}{Rr}\right)\int_{P^{\varepsilon}_{R+r}} u\right).$$

Proof. The idea of proof is as in Lemma B.2. Increase the dimension of the domain from two to three and apply Lemma B.4 with $n = 2$. Define functions U, F, G on $P_{R+r} \subset \mathbb{R} \times \mathbb{R}^2$ by

$$U(s,t,q) := u(s+\varepsilon q, t), \quad F(s,t,q) := f(s+\varepsilon q, t), \quad G(s,t,q) := g(s+\varepsilon q, t).$$

The new variable $\sigma := s + \varepsilon q$ satisfies $(\sigma, t) \in P^{\varepsilon}_{R+r} \subset \mathbb{R} \times \mathbb{R}$ whenever $(s,t,q) \in P_{R+r} \subset \mathbb{R} \times \mathbb{R}^2$. Use the differential inequality in the assumption of the lemma to conclude $(\Delta - \partial_s)U \ge G - F$, where $\Delta := \partial_t^2 + \partial_q^2$. Thus Lemma B.4 with $n = 2$ yields

$$\begin{aligned}
\int_{P_R} G &\le \int_{P_{R+r}} F + \left(\frac{4}{r^2} + \frac{1}{Rr}\right)\int_{P_{R+r}} U \\
&\le 2(R+r)\int_{P^{\varepsilon}_{R+r}} f + 2(R+r)\left(\frac{4}{r^2} + \frac{1}{Rr}\right)\int_{P^{\varepsilon}_{R+r}} u.
\end{aligned}$$

The last step uses (110). By definition of G

$$\begin{aligned}
\int_{P_R} G &= \int_{B_R\subset\mathbb{R}^2}\left(\int_{-R^2}^0 g(s+\varepsilon q,t)\,ds\right)dqdt \\
&\ge \int_{-R/2}^{R/2}\int_{-R/2}^{R/2}\int_{-R^2+\varepsilon q}^{\varepsilon q} g(\sigma,t)\,d\sigma dqdt \\
&\ge \int_{-R/2}^{R/2}\int_{-R/2}^{R/2}\int_{-R^2+\varepsilon R/2}^{-\varepsilon R/2} g(\sigma,t)\,d\sigma dqdt \\
&= R\int_{-R/2}^{R/2}\int_{-R^2+\varepsilon R/2}^{-\varepsilon R/2} g(\sigma,t)\,d\sigma dt \\
&\ge R\int_{P_{R/2}^{-\varepsilon}} g.
\end{aligned}$$

This proves the lemma. □

Lemma B.6. *Fix three positive constants r, R, ε and three functions $u, f, g : [-(R+r)^2-\varepsilon(R+r), \varepsilon(R+r)] \to \mathbb{R}$ such that u is C^2 and f, g are continuous. If*

$$\varepsilon^2 u'' - u' \ge g - f, \qquad u \ge 0, \qquad f \ge 0, \qquad g \ge 0,$$

then

$$\begin{aligned}
\int_{-R^2/4+R\varepsilon/2}^{-R\varepsilon/2} g(s)\,ds &\le 4\left(1+\frac{r}{R}\right)^2\int_{-(R+r)^2-\varepsilon(R+r)}^{\varepsilon(R+r)} f(s)\,ds \\
&\quad + 4\left(1+\frac{r}{R}\right)^2\left(\frac{4}{r^2}+\frac{1}{Rr}\right)\int_{-(R+r)^2-\varepsilon(R+r)}^{\varepsilon(R+r)} u(s)\,ds.
\end{aligned}$$

Proof. This follows immediately from Lemma B.5 with u, f, and g independent of the t-variable. □

C Two fundamental L^p estimates

Theorem C.1. *For every $p > 1$ there is a constant $c = c(p) > 0$ such that*

$$\|\partial_s u\|_{L^p} + \|\partial_s v\|_{L^p} \le c\left(\|\partial_s u - \partial_t v\|_{L^p} + \|\partial_s v + \partial_t u - v\|_{L^p}\right) \tag{113}$$

for all $u, v \in C_0^\infty(\mathbb{R}^2)$.

Theorem C.2. *For every $p > 1$ there is a constant $c = c(p) > 0$ such that*

$$\|\partial_s u\|_{L^p} + \|\partial_t\partial_t u\|_{L^p} \le c\,\|\partial_s u - \partial_t\partial_t u\|_{L^p} \tag{114}$$

for every $u \in C_0^\infty(\mathbb{R}^2)$.

If we assume $v = \partial_t u$ then (113) follows from (114) (but not conversely). On the other hand if the term $\partial_s v + \partial_t u - v$ on the right is replaced by $\partial_s v + \partial_t u$, then (113) becomes the Calderon-Zygmund inequality. However, it seems that the estimate (113) in its full strength cannot be deduced directly from the Calderon–Zygmund inequality and the parabolic estimate (114). Theorems C.1 and C.2 will be proved below.

Corollary C.3. *Let $p > 1$ and denote by $c = c(p)$ the constant of Theorem C.1. Then*

$$\|\partial_s u\|_{L^p} + \varepsilon \|\partial_s v\|_{L^p} \le c \left(\|\partial_s u - \partial_t v\|_{L^p} + \varepsilon \left\| \partial_s v + \varepsilon^{-2}(\partial_t u - v) \right\|_{L^p} \right) \quad (115)$$

for every $\varepsilon > 0$ and every pair $u, v \in C_0^\infty(\mathbb{R}^2)$.

Proof. Denote

$$f := \partial_s u - \partial_t v, \qquad g := \partial_s v + \varepsilon^{-2}(\partial_t u - v).$$

Now consider the rescaled functions

$$\tilde{u}(s,t) := u(\varepsilon^2 s, \varepsilon t), \qquad \tilde{v}(s,t) := \varepsilon v(\varepsilon^2 s, \varepsilon t)$$

and

$$\tilde{f}(s,t) := \varepsilon^2 f(\varepsilon^2 s, \varepsilon t), \qquad \tilde{g}(s,t) := \varepsilon^3 g(\varepsilon^2 s, \varepsilon t).$$

Then

$$\partial_s \tilde{u} - \partial_t \tilde{v} = \tilde{f}, \qquad \partial_s \tilde{v} + \partial_t \tilde{u} - \tilde{v} = \tilde{g}.$$

Hence, by Theorem C.1,

$$\left\|\partial_s \tilde{u}\right\|_{L^p} + \left\|\partial_s \tilde{v}\right\|_{L^p} \le c \left(\left\|\tilde{f}\right\|_{L^p} + \left\|\tilde{g}\right\|_{L^p} \right).$$

Now the result follows from the fact that

$$\left\|\partial_s \tilde{u}\right\|_{L^p} = \varepsilon^{2-3/p} \left\|\partial_s u\right\|_{L^p}, \qquad \left\|\tilde{f}\right\|_{L^p} = \varepsilon^{2-3/p} \left\|f\right\|_{L^p},$$

and similarly for the other terms. □

We give a proof of (113) and (114) that is based on the Marcinkiewicz–Mihlin multiplier method. To formulate the result, we consider the Fourier transform

$$\mathcal{F} : L^2(\mathbb{R}^2, \mathbb{C}) \to L^2(\mathbb{R}^2, \mathbb{C}),$$

given by

$$(\mathcal{F} f)(\sigma, \tau) := \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-i(\sigma s + \tau t)} f(s,t)\, ds dt$$

for $f \in L^2(\mathbb{R}^2, \mathbb{C}) \cap L^1(\mathbb{R}^2, \mathbb{C})$. Given a bounded measurable complex valued function $m : \mathbb{R}^2 \to \mathbb{C}$ define the bounded linear operator

$$\mathcal{T}_m : L^2(\mathbb{R}^2, \mathbb{C}) \to L^2(\mathbb{R}^2, \mathbb{C})$$

by

$$\mathcal{T}_m f := \mathcal{F}^{-1}(m \mathcal{F} f).$$

The following theorem is proved in [15].

Theorem C.4 (Marcinkiewicz–Mihlin). *For every $c > 0$ and every $p > 1$ there is a constant $c_p = c_p(c) > 0$ such that the following holds. If $m : \mathbb{R}^2 \to \mathbb{C}$ is a measurable function such that the restriction of m to each of the four open quadrants in $\mathbb{R}^2$ is twice continuously differentiable and*

$$|m(\sigma,\tau)| + |\sigma\partial_\sigma m(\sigma,\tau)| + |\tau\partial_\tau m(\sigma,\tau)| + |\sigma\tau\partial_\sigma\partial_\tau m(\sigma,\tau)| \le c \tag{116}$$

for $\sigma, \tau \in \mathbb{R}\setminus\{0\}$ then

$$f \in L^p(\mathbb{R}^2,\mathbb{C}) \cap L^2(\mathbb{R}^2,\mathbb{C}) \qquad \Longrightarrow \qquad \mathcal{T}_m f \in L^p(\mathbb{R}^2,\mathbb{C})$$

and

$$\|\mathcal{T}_m f\|_{L^p} \le c_p \|f\|_{L^p}$$

for every $f \in L^p(\mathbb{R}^2,\mathbb{C}) \cap L^2(\mathbb{R}^2,\mathbb{C})$.

Remark C.5. The theorem of Marcinkiewicz–Mihlin in its original form is slightly stronger than Theorem C.4, namely condition (116) is replaced by the weaker conditions

$$\sup_{\sigma,\tau} |m(\sigma,\tau)| \le c, \tag{117}$$

$$\sup_{\sigma\ne 0}\int_{2^\ell}^{2^{\ell+1}} |\partial_\tau m(\sigma,\pm\tau)|\, d\tau \le c, \qquad \sup_{\tau\ne 0}\int_{2^k}^{2^{k+1}} |\partial_\sigma m(\pm\sigma,\tau)|\, d\sigma \le c \tag{118}$$

and

$$\int_{2^k}^{2^{k+1}}\int_{2^\ell}^{2^{\ell+1}} |\partial_\sigma\partial_\tau m(\pm\sigma,\pm\tau)|\, d\tau \le c \tag{119}$$

for all integers k and ℓ (and all choices of signs). In this form the result is proved in Stein [24, Theorem 6']. It is easy to see that (116) implies (118) with c replaced by $c\log 2$ and (119) with c replaced by $c(\log 2)^2$.

Proof of Theorem C.2. Let $u \in C_0^\infty(\mathbb{R}^2)$ and define $f \in C_0^\infty(\mathbb{R}^2)$ by

$$f := \partial_s u - \partial_t\partial_t u.$$

Denote the Fourier transforms of f and u by

$$\widehat{f} := \mathcal{F}f, \qquad \widehat{u} := \mathcal{F}u.$$

Then

$$\widehat{f} = i\sigma\widehat{u} + \tau^2\widehat{u}$$

and hence

$$\widehat{\partial_s u} = i\sigma\widehat{u} = \frac{i\sigma}{i\sigma+\tau^2}\widehat{f}.$$

Denote the multiplier in this equation by

$$m(\sigma,\tau) := \frac{i\sigma}{i\sigma+\tau^2}.$$

The formulae

$$\partial_\sigma m = \frac{i\tau^2}{(\tau^2+i\sigma)^2}, \qquad \partial_\tau m = \frac{-2i\sigma\tau}{(\tau^2+i\sigma)^2}, \qquad \partial_\sigma\partial_\tau m = \frac{-2i\tau(\tau^2-i\sigma)}{(\tau^2+i\sigma)^3}$$

show that the functions m, $\sigma\partial_\sigma m$, $\tau\partial_\tau m$, and $\sigma\tau\partial_\sigma\partial_\tau m$ are bounded. (Estimate each factor in the denominator from below by either the real or the imaginary part.) Hence the result follows from Theorem C.4. □

Proof of Theorem C.1. Let $u, v \in C_0^\infty(\mathbb{R}^2)$ and define $f, g \in C_0^\infty(\mathbb{R}^2)$ by

$$f := \partial_s u - \partial_t v, \qquad g := \partial_s v + \partial_t u - v.$$

Then

$$\widehat{f} = i\sigma\widehat{u} - i\tau\widehat{v}, \qquad \widehat{g} = i\sigma\widehat{v} + i\tau\widehat{u} - \widehat{v}.$$

Solving this equation for $\widehat{u}$ and $\widehat{v}$ we find

$$\widehat{u} = \frac{1-i\sigma}{\sigma^2+\tau^2+i\sigma}\widehat{f} - \frac{i\tau}{\sigma^2+\tau^2+i\sigma}\widehat{g},$$

$$\widehat{v} = \frac{i\tau}{\sigma^2+\tau^2+i\sigma}\widehat{f} - \frac{i\sigma}{\sigma^2+\tau^2+i\sigma}\widehat{g},$$

and hence

$$\widehat{\partial_s u} = i\sigma\widehat{u} = \frac{\sigma^2+i\sigma}{\sigma^2+\tau^2+i\sigma}\widehat{f} + \frac{\sigma\tau}{\sigma^2+\tau^2+i\sigma}\widehat{g},$$

$$\widehat{\partial_s v} = i\sigma\widehat{v} = \frac{-\sigma\tau}{\sigma^2+\tau^2+i\sigma}\widehat{f} + \frac{\sigma^2}{\sigma^2+\tau^2+i\sigma}\widehat{g}.$$

The four multipliers in the last two equations satisfy (116). (To see this use separation of the factors in the denominator into real and imaginary parts, as in the proof of Theorem C.2, and the inequality $2\sigma\tau \le \sigma^2+\tau^2$.) Hence the result follows from Theorem C.4. □

D The estimate for the inverse

We begin by proving a weaker version of the estimate in Theorem 3.2. Throughout M is a compact Riemannian manifold.

Proposition D.1. *Let $u \in C^\infty(\mathbb{R}\times S^1, M)$ and $v \in \Omega^0(\mathbb{R}\times S^1, u^*TM)$ such that $\|\partial_s u\|_\infty$, $\|\partial_t u\|_\infty$ and $\|v\|_\infty$ are finite and $\lim_{s\to\pm\infty} u(s,t)$ exists, uniformly in t. Then, for every $p > 1$, there is a constant $c > 0$ such that*

$$\begin{aligned} &\varepsilon^{-1}\|\nabla_t\xi - \eta\|_p + \|\nabla_t\eta\|_p + \|\nabla_s\xi\|_p + \varepsilon\|\nabla_s\eta\|_p \\ &\le c\Big(\|\nabla_s\xi - \nabla_t\eta\|_p + \varepsilon\left\|\nabla_s\eta + \varepsilon^{-2}(\nabla_t\xi - \eta)\right\|_p + \varepsilon^{-1}\|\xi\|_p + \|\eta\|_p\Big) \end{aligned} \tag{120}$$

for every $\varepsilon \in (0,1]$ and every pair of compactly supported vector fields $\zeta = (\xi,\eta) \in \Omega^0(\mathbb{R}\times S^1, u^*TM \oplus u^*TM)$.

Proof. Choose a finite open cover $\{U_\alpha\}_\alpha$ of $\mathbb{R}\times S^1$ with the following properties.

(i) For each α the set $U_\alpha \subset \mathbb{R}\times S^1$ is contractible.

(ii) For each α the closure of the image of U_α under u is contained in a coordinate chart on M.

(iii) There is a constant $T > 0$ and an open cover $\{I_\alpha\}_\alpha$ of S^1 such that $U_\alpha \cap [T,\infty)\times S^1 = [T,\infty)\times I_\alpha$ for every α. Similarly for the interval $(-\infty,-T]$.

Assume first that ξ and η are compactly supported in U_α for some α and denote by $\xi_\alpha, \eta_\alpha : U_\alpha \to \mathbb{R}^n$ the vector fields in local coordinates. By Corollary C.3, there is a constant c_α, depending only on p and the metric, such that

$$\|\partial_s\xi_\alpha\|_p + \varepsilon\|\partial_s\eta_\alpha\|_p \le c_\alpha\Big(\|\partial_s\xi_\alpha - \partial_t\eta_\alpha\|_p + \varepsilon\left\|\partial_s\eta_\alpha + \varepsilon^{-2}(\partial_t\xi_\alpha - \eta_\alpha)\right\|_p\Big).$$

Here we denote by $\|\cdot\|_p$ the L^p norm with respect to the Riemannian metric in the coordinate charts on M. Replacing the partial derivatives ∂_s and ∂_t by the covariant derivatives ∇_s and ∇_t we obtain

$$\begin{aligned}\|\nabla_s\xi\|_p + \varepsilon\|\nabla_s\eta\|_p \le c\Big(&\|\nabla_s\xi - \nabla_t\eta\|_p + \varepsilon\left\|\nabla_s\eta + \varepsilon^{-2}(\nabla_t\xi - \eta)\right\|_p \\ &+ \varepsilon^{-1}\|\xi\|_p + \|\eta\|_p\Big).\end{aligned} \tag{121}$$

for every ξ with support in one of the sets U_α. Here we have used the L^∞ bounds on $\partial_s u$ and $\partial_t u$. Observe that the constant c depends on the Christoffel symbols determined by our coordinate chart on M. Now let $\{\beta_\alpha\}_\alpha$ be a partition of unity subordinate to the cover $\{U_\alpha\}_\alpha$ such that $\|\partial_s\beta_\alpha\|_\infty + \|\partial_t\beta_\alpha\|_\infty < \infty$ for every α. (Note that β_α need not have compact support when U_α is unbounded.) Given any two compactly supported vector fields $\xi,\eta \in \Omega^0(\mathbb{R}\times S^1, u^*TM)$ apply (121) to the (compactly supported) pair $(\beta_\alpha\xi, \beta_\alpha\eta)$ and take the sum to deduce that (121) continues to hold for each pair (ξ,η) with an appropriate larger constant c. This proves the proposition because the first two terms on the left can be estimated by the last two terms and the right hand side. $\square$

Under the assumptions of Proposition D.1 it follows immediately that

$$\|\zeta\|_{1,p,\varepsilon} \le c\left(\varepsilon^2\|\mathcal{D}^\varepsilon_{u,v}\zeta\|_{0,p,\varepsilon} + \|\zeta\|_{0,p,\varepsilon}\right) \tag{122}$$

and similarly for $(\mathcal{D}^\varepsilon_{u,v})^*$. Moreover, note that the difference between Proposition D.1 and Theorem 3.2 lies in the ε-factors in front of $\|\xi\|_p$ and $\|\eta\|_p$ on the right hand sides of the estimates. To prove Theorem 3.2 we must improve these these factors by ε for ξ and by ε^2 for η. This requires the following parabolic estimate. Let $1/p + 1/q = 1$. The *formal adjoint operator*

$$(\mathcal{D}^0_u)^* : \mathcal{W}^q_u \to \mathcal{L}^q_u$$

of $\mathcal{D}^0_u : \mathcal{W}^p_u \to \mathcal{L}^p_u$ is given by

$$(\mathcal{D}^0_u)^*\xi = -\nabla_s\xi - \nabla_t\nabla_t\xi - R(\xi,\partial_t u)\partial_t u - \mathcal{H}_\mathcal{V}(u)\xi. \tag{123}$$

Proposition D.2. *Let $u \in C^\infty(\mathbb{R} \times S^1, M)$ such that $\|\partial_s u\|_\infty$, $\|\partial_t u\|_\infty$ and $\|\nabla_t \partial_t u\|_\infty$ are finite and $\lim_{s\to\pm\infty} u(s,t)$ exists, uniformly in t. Then, for every $p > 1$, there is a constant $c > 0$ such that*

$$\|\nabla_s \xi\|_p + \|\nabla_t \nabla_t \xi\|_p \le c\left(\|\nabla_s \xi - \nabla_t \nabla_t \xi\|_p + \|\xi\|_p\right) \tag{124}$$

for every compactly supported vector field $\xi \in \Omega^0(\mathbb{R} \times S^1, u^*TM)$.

Lemma D.3. *Let $x : S^1 \to M$ be a smooth map, $p > 1$ and*

$$\kappa_p := \begin{cases} p & \text{if } p \ge 2, \\ p/(p-1) & \text{if } p \le 2. \end{cases} \tag{125}$$

Then, for every $\varepsilon > 0$ and every $\xi \in \Omega^0(S^1, x^*TM)$, we have

$$\begin{aligned} \|(\mathbb{1} - \varepsilon \nabla_t \nabla_t)^{-1} \xi\|_p &\le \|\xi\|_p, \\ \sqrt{\varepsilon}\|(\mathbb{1} - \varepsilon \nabla_t \nabla_t)^{-1} \nabla_t \xi\|_p &\le \kappa_p \|\xi\|_p, \\ \varepsilon\|(\mathbb{1} - \varepsilon \nabla_t \nabla_t)^{-1} \nabla_t \nabla_t \xi\|_p &\le 2\|\xi\|_p. \end{aligned}$$

These estimates continue to hold for $u \in C^\infty(\mathbb{R} \times S^1, M)$ and compactly supported vector fields $\xi \in \Omega^0(\mathbb{R} \times S^1, u^*TM)$.

Proof. First consider the case $p \ge 2$: Let $\varepsilon > 0$ and $\xi \in \Omega^0(S^1, x^*TM)$. Define

$$\eta := (\mathbb{1} - \varepsilon \nabla_t \nabla_t)^{-1} \xi.$$

(The operator $(\mathbb{1} - \varepsilon \nabla_t \nabla_t) : W^{2,p}(S^1, x^*TM) \to L^p(S^1, x^*TM)$ is bijective.) Then

$$\begin{aligned} \frac{d^2}{dt^2}|\eta|^p &= \frac{d}{dt}\left(p\,|\eta|^{p-2}\,\langle \nabla_t \eta, \eta\rangle\right) \\ &= p(p-2)|\eta|^{p-4}\langle \nabla_t \eta, \eta\rangle^2 + p|\eta|^{p-2}\left(\langle \nabla_t \nabla_t \eta, \eta\rangle + |\nabla_t \eta|^2\right) \\ &\ge p\varepsilon^{-1}|\eta|^p - p\varepsilon^{-1}|\eta|^{p-2}\langle \xi, \eta\rangle \\ &\ge p\varepsilon^{-1}|\eta|^p - p\varepsilon^{-1}|\eta|^{p-1}|\xi| \\ &\ge \varepsilon^{-1}|\eta|^p - \varepsilon^{-1}|\xi|^p. \end{aligned}$$

The third step uses the identity $\nabla_t \nabla_t \eta = \varepsilon^{-1}\eta - \varepsilon^{-1}\xi$. The last step uses Young's inequality

$$ab \le \frac{a^r}{r} + \frac{b^s}{s}, \qquad \frac{1}{r} + \frac{1}{s} = 1, \tag{126}$$

with $r = p$, $a = |\xi|$ and $s = p/(p-1)$, $b = |\eta|^{p-1}$. Moreover,

$$
\begin{aligned}
&\frac{d}{dt}\Big(|\nabla_t\eta|^{p-2}\langle\nabla_t\eta,\eta\rangle\Big) \\
&= |\nabla_t\eta|^p + |\nabla_t\eta|^{p-2}\langle\nabla_t\nabla_t\eta,\eta\rangle + (p-2)|\nabla_t\eta|^{p-4}\langle\nabla_t\eta,\eta\rangle\langle\nabla_t\nabla_t\eta,\nabla_t\eta\rangle \\
&= |\nabla_t\eta|^p + \varepsilon^{-1}|\nabla_t\eta|^{p-2}|\eta|^2 - \varepsilon^{-1}|\nabla_t\eta|^{p-2}\langle\xi,\eta\rangle \\
&\quad - \varepsilon^{-1}(p-2)|\nabla_t\eta|^{p-4}\langle\nabla_t\eta,\eta\rangle\langle\xi,\nabla_t\eta\rangle + \varepsilon^{-1}(p-2)|\nabla_t\eta|^{p-4}\langle\nabla_t\eta,\eta\rangle^2 \\
&\geq |\nabla_t\eta|^p + \tfrac{1}{2}\varepsilon^{-1}|\nabla_t\eta|^{p-2}|\eta|^2 - \tfrac{p-1}{2}\varepsilon^{-1}|\nabla_t\eta|^{p-2}|\xi|^2 + \tfrac{p-2}{2}\varepsilon^{-1}|\nabla_t\eta|^{p-4}\langle\nabla_t\eta,\eta\rangle^2 \\
&\geq |\nabla_t\eta|^p - \tfrac{p-1}{2}\varepsilon^{-1}|\nabla_t\eta|^{p-2}|\xi|^2 \\
&\geq \tfrac{2}{p}|\nabla_t\eta|^p - \tfrac{2}{p}\big(\tfrac{p-1}{2}\big)^{p/2}\varepsilon^{-p/2}|\xi|^p.
\end{aligned}
$$

The third step uses (126) with $r = s = 2$. The last step uses (126) with $r = p/2$, $a = \frac{p-1}{2}\varepsilon^{-1}|\xi|^2$ and $s = p/(p-2)$, $b = |\nabla_t\eta|^{p-2}$. Now the first two estimates of the lemma follow by integration over S^1, respectively $\mathbb{R}\times S^1$. (The integrals of the left hand sides vanish, by periodicity.) The last estimate is an easy consequence of the first:

$$
\varepsilon\|\nabla_t\nabla_t\eta\|_p = \|\eta-\xi\|_p \leq \|\eta\|_p + \|\xi\|_p \leq 2\|\xi\|_p.
$$

This proves the lemma for $p \geq 2$. Now assume $1 < p < 2$ and let $q := p/(p-1)$. Then $q > 2$ and hence

$$
\begin{aligned}
\sqrt{\varepsilon}\left\|(\mathbb{1}-\varepsilon\nabla_t\nabla_t)^{-1}\nabla_t\xi\right\|_p &= \sqrt{\varepsilon}\sup_{0\neq\eta\in L^q}\frac{\big\langle(\mathbb{1}-\varepsilon\nabla_t\nabla_t)^{-1}\nabla_t\xi,\eta\big\rangle}{\|\eta\|_q} \\
&\leq \sqrt{\varepsilon}\sup_{0\neq\eta\in L^q}\frac{\|\xi\|_p\left\|(\mathbb{1}-\varepsilon\nabla_t\nabla_t)^{-1}\nabla_t\eta\right\|_q}{\|\eta\|_q} \\
&\leq q\,\|\xi\|_p\,.
\end{aligned}
$$

This prove the second estimate for $p < 2$. The other estimates follow similarly. This proves the lemma. □

Lemma D.4. *Let $x \in C^\infty(S^1, M)$ and $p > 1$. Then*

$$
\|\nabla_t\xi\|_p \leq \kappa_p\left(\delta^{-1}\|\xi\|_p + \delta\|\nabla_t\nabla_t\xi\|_p\right)
$$

for $\delta > 0$ and $\xi \in \Omega^0(S^1, x^*TM)$, where κ_p is defined by (125). This estimate continues to hold for $u \in C^\infty(\mathbb{R}\times S^1, M)$ and compactly supported vector fields $\xi \in \Omega^0(\mathbb{R}\times S^1, u^*TM)$.

Proof. Let $1/p + 1/q = 1$. Since the operator

$$
W^{2,q}(S^1, x^*TM) \to L^q(S^1, x^*TM) : \eta \mapsto \delta^{-1}\eta - \delta\nabla_t\nabla_t\eta
$$

is bijective, we have

$$\begin{aligned}
\|\nabla_t\xi\|_p &= \sup_{\eta\in W^{2,q}} \frac{\langle \nabla_t\xi, \delta^{-1}\eta - \delta\nabla_t\nabla_t\eta\rangle}{\|\delta^{-1}\eta - \delta\nabla_t\nabla_t\eta\|_q} \\
&= \sup_{\eta\in W^{2,q}} \frac{-\langle \xi, \delta^{-1}\nabla_t\eta\rangle + \langle \nabla_t\nabla_t\xi, \delta\nabla_t\eta\rangle}{\|\delta^{-1}\eta - \delta\nabla_t\nabla_t\eta\|_q} \\
&\le \left(\delta^{-1}\|\xi\|_p + \delta\|\nabla_t\nabla_t\xi\|_p\right) \sup_{\eta\in W^{2,q}} \frac{\|\nabla_t\eta\|_q}{\|\delta^{-1}\eta - \delta\nabla_t\nabla_t\eta\|_q} \\
&\le \kappa_p \left(\delta^{-1}\|\xi\|_p + \delta\|\nabla_t\nabla_t\xi\|_p\right).
\end{aligned}$$

To prove the last step, denote

$$\zeta := \eta - \delta^2\nabla_t\nabla_t\eta.$$

Then

$$\nabla_t\eta = \left(\mathbb{1} - \delta^2\nabla_t\nabla_t\right)^{-1}\nabla_t\zeta$$

and hence, by Lemma D.3 with $\varepsilon = \delta^2$, we have

$$\|\nabla_t\eta\|_q \le \kappa_q\delta^{-1}\|\zeta\|_q = \kappa_p\left\|\delta^{-1}\eta - \delta\nabla_t\nabla_t\eta\right\|_q.$$

We have used the fact that $\kappa_p = \kappa_q$. This proves the lemma. □

Proof of Proposition D.2. The proof follows the same pattern as that of Proposition D.1. Let $\{U_\alpha\}_\alpha$ be as above. If ξ is (compactly) supported in U_α then, by Theorem C.2,

$$\|\partial_s\xi_\alpha\|_p + \|\partial_t\partial_t\xi_\alpha\|_p \le c_\alpha\|\partial_s\xi_\alpha - \partial_t\partial_t\xi_\alpha\|_p$$

Replacing ∂_s and ∂_t by ∇_s and ∇_t, and using the L^∞ bounds on $\partial_s u$, $\partial_t u$, and $\nabla_t\partial_t u$, we find

$$\|\nabla_s\xi\|_p + \|\nabla_t\nabla_t\xi\|_p \le c\Big(\|\nabla_s\xi - \nabla_t\nabla_t\xi\|_p + \|\xi\|_p + \|\nabla_t\xi\|_p\Big).$$

Using a partition of unity $\{\beta_\alpha\}_\alpha$, subordinate to the cover $\{U_\alpha\}_\alpha$, such that

$$\|\partial_s\beta_\alpha\|_\infty + \|\partial_t\beta_\alpha\|_\infty + \|\partial_t\partial_t\beta_\alpha\|_\infty < \infty,$$

we deduce that the last estimate continues to hold for every compactly supported vector field $\xi \in \Omega^0(\mathbb{R}\times S^1, u^*TM)$. Now apply Lemma D.4 with $\delta cp < 1/2$ to obtain the estimate (124). □

Proof of Theorem 3.2

Fix a constant $p > 1$ and define

$$f(\xi,\eta) := \nabla_s\xi - \nabla_t\eta, \qquad g(\xi,\eta) := \nabla_s\eta + \varepsilon^{-2}(\nabla_t\xi - \eta),$$

for compactly supported vector fields $\zeta = (\xi, \eta) \in \Omega^0(\mathbb{R}\times S^1, u^*TM \oplus u^*TM)$. It suffices to show that

$$\begin{aligned} &\varepsilon^{-1}\left\|\nabla_t\xi - \eta\right\|_p + \left\|\nabla_t\eta\right\|_p + \left\|\nabla_s\xi\right\|_p + \varepsilon\left\|\nabla_s\eta\right\|_p \\ &\le c\left(\left\|f\right\|_p + \varepsilon\left\|g\right\|_p + \left\|\xi\right\|_p + \varepsilon^2\left\|\eta\right\|_p\right) \end{aligned} \tag{127}$$

for some constant $c>0$ independent of ε and (ξ,η). The general case (for $\mathcal{D}^\varepsilon_{u,v}$) then follows easily:

$$\begin{aligned} &\varepsilon^{-1}\left\|\nabla_t\xi - \eta\right\|_p + \left\|\nabla_t\eta\right\|_p + \left\|\nabla_s\xi\right\|_p + \varepsilon\left\|\nabla_s\eta\right\|_p \\ &\le c'\Big(\left\|f - R(\xi,\partial_t u)v - \mathcal{H}_{\mathcal{V}}(u)\xi\right\|_p + \varepsilon\left\|g + R(\xi,\partial_s u)v\right\|_p + \left\|\xi\right\|_p + \varepsilon^2\left\|\eta\right\|_p\Big). \end{aligned}$$

To prove the estimate for the formal adjoint operator $(\mathcal{D}^\varepsilon_{u,v})^*$ apply (127) to the vector fields $\xi(-s,t)$ and $\eta(-s,t)$ and then proceed as above.

To prove (127) we split ζ into two components. Let

$$\pi_\varepsilon(\xi,\eta) := (\mathbb{1} - \varepsilon\nabla_t\nabla_t)^{-1}(\xi - \varepsilon^2\nabla_t\eta), \qquad \iota(\xi) := (\xi, \nabla_t\xi),$$

(compare with equation (139) below) and define

$$\zeta_0 := \begin{pmatrix}\xi_0\\ \eta_0\end{pmatrix} := \iota\pi_\varepsilon\zeta = \begin{pmatrix}(\mathbb{1} - \varepsilon\nabla_t\nabla_t)^{-1}(\xi - \varepsilon^2\nabla_t\eta)\\ \nabla_t(\mathbb{1} - \varepsilon\nabla_t\nabla_t)^{-1}(\xi - \varepsilon^2\nabla_t\eta)\end{pmatrix},$$

$$\zeta_1 := \begin{pmatrix}\xi_1\\ \eta_1\end{pmatrix} := \zeta - \zeta_0 = \begin{pmatrix}(\mathbb{1} - \varepsilon\nabla_t\nabla_t)^{-1}(\varepsilon^2\nabla_t\eta - \varepsilon\nabla_t\nabla_t\xi)\\ (\mathbb{1} - \varepsilon\nabla_t\nabla_t)^{-1}(\eta - \nabla_t\xi + (\varepsilon^2 - \varepsilon)\nabla_t\nabla_t\eta)\end{pmatrix}.$$

Note that $\eta_0 = \nabla_t\xi_0$ and

$$\xi_1 - \varepsilon\nabla_t\eta_1 = (\varepsilon^2 - \varepsilon)\nabla_t\eta. \tag{128}$$

Since f and g are linear, we obtain the splitting $f = f_0 + f_1$ and $g = g_0 + g_1$, where $f_i := f(\xi_i, \eta_i)$ and $g_i := g(\xi_i,\eta_i)$ for $i = 0,1$. Thus

$$f_0 = \nabla_s\xi_0 - \nabla_t\nabla_t\xi_0, \qquad g_0 = \nabla_s\nabla_t\xi_0.$$

Now apply the parabolic estimate of Proposition D.2, with a constant $c_0 > 0$, to ξ_0 and the elliptic estimate of Proposition D.1, with a constant $c_1 > 0$, to (ξ_1, η_1). This gives

$$\begin{aligned} &\varepsilon^{-1}\left\|\nabla_t\xi - \eta\right\|_p + \left\|\nabla_t\eta\right\|_p + \left\|\nabla_s\xi\right\|_p + \varepsilon\left\|\nabla_s\eta\right\|_p \\ &\le \left\|\nabla_t\nabla_t\xi_0\right\|_p + \left\|\nabla_s\xi_0\right\|_p + \varepsilon\left\|\nabla_s\nabla_t\xi_0\right\|_p \\ &\quad + \varepsilon^{-1}\left\|\nabla_t\xi_1 - \eta_1\right\|_p + \left\|\nabla_t\eta_1\right\|_p + \left\|\nabla_s\xi_1\right\|_p + \varepsilon\left\|\nabla_s\eta_1\right\|_p \\ &\le c_0\big(\left\|f_0\right\|_p + \left\|\xi_0\right\|_p\big) + \varepsilon\left\|g_0\right\|_p \\ &\quad + c_1\big(\left\|f_1\right\|_p + \varepsilon\left\|g_1\right\|_p + \varepsilon^{-1}\left\|\xi_1\right\|_p + \left\|\eta_1\right\|_p\big) \\ &\le c_1\big(\left\|f\right\|_p + \varepsilon\left\|g\right\|_p + \varepsilon^{-1}\left\|\xi_1\right\|_p + \left\|\eta_1\right\|_p\big) \\ &\quad + (c_0 + c_1)\left\|f_0\right\|_p + (1 + c_1)\varepsilon\left\|g_0\right\|_p + c_0\left\|\xi_0\right\|_p. \end{aligned} \tag{129}$$

We examine the last five terms on the right individually. For this we shall need the commutator identities

$$[\nabla_s, \nabla_t] = R(\partial_s u, \partial_t u), \tag{130}$$

$$\begin{aligned} [\nabla_s, \nabla_t\nabla_t] &= 2\nabla_t[\nabla_s, \nabla_t] - (\nabla_{\partial_t u} R)(\partial_s u, \partial_t u) \\ &\quad - R(\nabla_t \partial_s u, \partial_t u) + R(\partial_s u, \nabla_t \partial_t u), \end{aligned} \tag{131}$$

$$\begin{aligned} [\nabla_s, (\mathbb{1} - \varepsilon\nabla_t\nabla_t)^{-1}] &= (\mathbb{1} - \varepsilon\nabla_t\nabla_t)^{-1}[\mathbb{1} - \varepsilon\nabla_t\nabla_t, \nabla_s](\mathbb{1} - \varepsilon\nabla_t\nabla_t)^{-1} \\ &= \varepsilon(\mathbb{1} - \varepsilon\nabla_t\nabla_t)^{-1}[\nabla_s, \nabla_t\nabla_t](\mathbb{1} - \varepsilon\nabla_t\nabla_t)^{-1}. \end{aligned} \tag{132}$$

By Lemma D.3 and (131), we have

$$\begin{aligned} &\varepsilon^{1/2} \left\|(\mathbb{1} - \varepsilon\nabla_t\nabla_t)^{-1}[\nabla_s, \nabla_t\nabla_t]\xi\right\|_p \\ &\le 2\varepsilon^{1/2} \left\|(\mathbb{1} - \varepsilon\nabla_t\nabla_t)^{-1}\nabla_t[\nabla_s, \nabla_t]\xi\right\|_p + c_1\varepsilon^{1/2} \|\xi\|_p \\ &\le 2\kappa_p \|[\nabla_s, \nabla_t]\xi\|_p + c_1\varepsilon^{1/2} \|\xi\|_p \\ &\le c_2 \|\xi\|_p . \end{aligned} \tag{133}$$

Here we have used the L^∞ bounds on $\partial_s u$, $\partial_t u$, $\nabla_t\partial_t u$, and $\nabla_t\partial_s u$. Now the five relevant terms are estimated as follows.

The term $\|\xi_0\|_p$: By definition,

$$\xi_0 = (\mathbb{1} - \varepsilon\nabla_t\nabla_t)^{-1}(\xi - \varepsilon^2\nabla_t\eta).$$

Hence, by Lemma D.3,

$$\|\xi_0\|_p \le \|\xi\|_p + \varepsilon^2 \|\nabla_t\eta\|_p . \tag{134}$$

The term $\|f_0\|_p$: Consider the identity

$$\begin{aligned} &(\mathbb{1} - \varepsilon\nabla_t\nabla_t)f_0 - f + \varepsilon^2\nabla_t g \\ &= \nabla_s\xi_0 - \varepsilon\nabla_t\nabla_t\nabla_s\xi_0 - \nabla_t\nabla_t\xi_0 + \varepsilon\nabla_t\nabla_t\nabla_t\nabla_t\xi_0 - \nabla_s\xi + \varepsilon^2\nabla_t\nabla_s\eta + \nabla_t\nabla_t\xi \\ &= \varepsilon^2\nabla_t\nabla_t\nabla_t\eta + \varepsilon^2 R(\partial_t u, \partial_s u)\eta + \varepsilon[\nabla_s, \nabla_t\nabla_t]\xi_0. \end{aligned}$$

Apply the operator $(\mathbb{1}-\varepsilon\nabla_t\nabla_t)^{-1}$ to this equation and use Lemma D.3 and (133) to obtain

$$\|f_0\|_p \le \|f\|_p + \kappa_p\varepsilon^{3/2} \|g\|_p + 2\varepsilon \|\nabla_t\eta\|_p + \varepsilon^2 c_3 \|\eta\|_p + \varepsilon^{1/2} c_2 \|\xi_0\|_p , \tag{135}$$

where $c_3 := \|R\|_\infty \|\partial_s u\|_\infty \|\partial_t u\|_\infty$.

The term $\varepsilon\|g_0\|_p$: By (132), we have

$$\begin{aligned} g_0 &= \nabla_s\nabla_t\xi_0 \\ &= (\mathbb{1} - \varepsilon\nabla_t\nabla_t)^{-1} \left(\nabla_t\nabla_s\xi + [\nabla_s, \nabla_t]\xi - \varepsilon^2\nabla_t\nabla_t\nabla_s\eta - \varepsilon^2[\nabla_s, \nabla_t\nabla_t]\eta\right) \\ &\quad + \varepsilon(\mathbb{1} - \varepsilon\nabla_t\nabla_t)^{-1}[\nabla_s, \nabla_t\nabla_t](\mathbb{1} - \varepsilon\nabla_t\nabla_t)^{-1} \left(\nabla_t\xi - \varepsilon^2\nabla_t\nabla_t\eta\right) . \end{aligned}$$

Hence, by Lemma D.3, (130), and (133),

$$\begin{aligned}
\varepsilon\left\|g_0\right\|_p &\le \kappa_p\varepsilon^{1/2}\left\|\nabla_s\xi\right\|_p + c_3\varepsilon\left\|\xi\right\|_p + 2\varepsilon^2\left\|\nabla_s\eta\right\|_p + c_2\varepsilon^{5/2}\left\|\eta\right\|_p \\
&\quad + c_2\varepsilon^{3/2}\left\|(\mathbb{1}-\varepsilon\nabla_t\nabla_t)^{-1}\left(\nabla_t\xi-\varepsilon^2\nabla_t\nabla_t\eta\right)\right\|_p \\
&\le \kappa_p\varepsilon^{1/2}\left\|\nabla_s\xi\right\|_p + 2\varepsilon^2\left\|\nabla_s\eta\right\|_p \\
&\quad + \varepsilon(\kappa_p c_2 + c_3)\left\|\xi\right\|_p + 3c_2\varepsilon^{5/2}\left\|\eta\right\|_p .
\end{aligned}\tag{136}$$

The term $\varepsilon^{-1}\|\xi_1\|_p$: By (128), we have

$$\varepsilon^{-1}\xi_1 = \nabla_t\eta_1 + \varepsilon\nabla_t\eta - \nabla_t\eta = \varepsilon\nabla_t\eta - \nabla_t\eta_0 .$$

Hence

$$\begin{aligned}
\varepsilon^{-1}\|\xi_1\|_p &\le \varepsilon\left\|\nabla_t\eta\right\|_p + \left\|\nabla_t\nabla_t\xi_0\right\|_p \\
&\le \varepsilon\left\|\nabla_t\eta\right\|_p + c_0\left(\left\|f_0\right\|_p + \left\|\xi_0\right\|_p\right).
\end{aligned}\tag{137}$$

In the last step we have used the parabolic estimate of Proposition D.2.
The term $\|\eta_1\|_p$: By definition,

$$\eta_1 = (\mathbb{1}-\varepsilon\nabla_t\nabla_t)^{-1}\left(\eta - \nabla_t\xi + (\varepsilon^2-\varepsilon)\nabla_t\nabla_t\eta\right).$$

Hence, by the triangle inequality and Lemma D.3, we have

$$\begin{aligned}
\|\eta_1\|_p &\le \left\|(\mathbb{1}-\varepsilon\nabla_t\nabla_t)^{-1}(\eta-\nabla_t\xi)\right\|_p + \varepsilon\left\|(\mathbb{1}-\varepsilon\nabla_t\nabla_t)^{-1}\nabla_t\nabla_t\eta\right\|_p \\
&\le \left\|\eta-\nabla_t\xi\right\|_p + \kappa_p\sqrt{\varepsilon}\left\|\nabla_t\eta\right\|_p .
\end{aligned}\tag{138}$$

Insert the five estimates (134-138) into (129) to obtain (15), provided that ε is sufficiently small. This proves Theorem 3.2.

The difference of $\mathcal{D}^0$ and $\mathcal{D}^\varepsilon$

Geometrically, the difference between the operators $\mathcal{D}_u^0$ and $\mathcal{D}_{u,v}^\varepsilon$ is the difference between configuration space and phase space, or between loops in M and loops in $T^*M \cong TM$. Consider the embedding

$$\mathcal{L}M \to \mathcal{L}TM : x \mapsto (x, \dot{x}).$$

The differential of this embedding is given by

$$\Omega^0(S^1, x^*TM) \to \Omega^0(S^1, x^*TM \oplus x^*TM) : \xi \mapsto (\xi, \nabla_t\xi).$$

To compare the operators $\mathcal{D}_u^0$ and $\mathcal{D}_u^\varepsilon := \mathcal{D}_{u,\partial_t u}^\varepsilon$ we must choose a projection onto the image of this embedding (along u). At first glance it might seem natural to choose the orthogonal projection with respect to the inner product determined by the $(0,2,\varepsilon)$-Hilbert space structure. This is given by

$$(\xi,\eta) \mapsto (\mathbb{1}-\varepsilon^\alpha\nabla_t\nabla_t)^{-1}(\xi - \varepsilon^\beta\nabla_t\eta)$$

with $\alpha = \beta = 2$. Instead we introduce the projection operator

$$\pi_\varepsilon : L^p(S^1, u^*TM) \times L^p(S^1, u^*TM) \to W^{1,p}(S^1, u^*TM)$$

given by

$$\pi_\varepsilon(\xi, \eta) := (\mathbb{1} - \varepsilon \nabla_t \nabla_t)^{-1}(\xi - \varepsilon^2 \nabla_t \eta). \tag{139}$$

The reason for this choice becomes visible in the proof of Proposition D.5 below, which requires $\beta = 2$. Moreover, the estimates in Step 1 of the proof of Theorem 3.3 are optimized for $\alpha = 1$. We denote by $\iota : W^{1,p}(\mathbb{R} \times S^1, u^*TM) \to L^p(S^1, u^*TM) \times L^p(S^1, u^*TM)$ the inclusion

$$\iota\xi_0 := (\xi_0, \nabla_t \xi_0). \tag{140}$$

The significance of these definitions lies in the next proposition and lemma. The proofs rely on Lemma D.3.

Proposition D.5. *Let $u \in C^\infty(\mathbb{R} \times S^1, M)$ be a smooth map such that the derivatives $\partial_s u, \partial_t u, \nabla_t \partial_s u, \nabla_t \partial_t u, \nabla_t \nabla_t \partial_t u$ are bounded. Then, for every $p > 1$, there exists a constant $c > 0$ such that*

$$\left\| \mathcal{D}_u^0 \pi_\varepsilon \zeta - \pi_\varepsilon \mathcal{D}_u^\varepsilon \zeta \right\|_p \le c\varepsilon^{1/2} \left\| \xi \right\|_p + c\varepsilon^2 \left\| \eta \right\|_p + c\varepsilon \left\| \nabla_t \eta \right\|_p$$

for $\varepsilon \in (0, 1]$ and compactly supported $\zeta = (\xi, \eta) \in \Omega^0(\mathbb{R} \times S^1, u^*TM \oplus u^*TM)$. The same estimate holds for $(\mathcal{D}_u^0)^* \pi_\varepsilon - \pi_\varepsilon (\mathcal{D}_u^\varepsilon)^*$. Moreover, the constant c is invariant under s-shifts of u.

Lemma D.6. *For $u \in C^\infty(\mathbb{R} \times S^1, M)$, $p > 1$, κ_p as in (125), and $0 < \varepsilon \le 1$,*

$$\begin{aligned}
\left\| \xi - \pi_\varepsilon \zeta \right\|_p &\le \kappa_p \varepsilon^{1/2} \left\| \nabla_t \xi - \eta \right\|_p + \varepsilon \left\| \nabla_t \eta \right\|_p \\
\left\| \eta - \nabla_t \pi_\varepsilon \zeta \right\|_p &\le \left\| \nabla_t \xi - \eta \right\|_p + \kappa_p \varepsilon^{1/2} \left\| \nabla_t \eta \right\|_p \\
\left\| \zeta - \iota \pi_\varepsilon \zeta \right\|_{0,p,\varepsilon} &\le 2\kappa_p \varepsilon^{1/2} \left\| \nabla_t \xi - \eta \right\|_p + 2\kappa_p \varepsilon \left\| \nabla_t \eta \right\|_p \\
\left\| \pi_\varepsilon \zeta \right\|_p &\le \left\| \iota \pi_\varepsilon \zeta \right\|_{0,p,\varepsilon} \le 2\kappa_p \left\| \zeta \right\|_{0,p,\varepsilon}
\end{aligned}$$

for every compactly supported $\zeta = (\xi, \eta) \in \Omega^0(\mathbb{R} \times S^1, u^*TM \oplus u^*TM)$.

Proof. Denote

$$\xi_0 := \pi_\varepsilon \zeta = (\mathbb{1} - \varepsilon \nabla_t \nabla_t)^{-1}(\xi - \varepsilon^2 \nabla_t \eta).$$

Then

$$\xi - \xi_0 = \varepsilon(\mathbb{1} - \varepsilon \nabla_t \nabla_t)^{-1} \nabla_t(\eta - \nabla_t \xi) + (\varepsilon^2 - \varepsilon)(\mathbb{1} - \varepsilon \nabla_t \nabla_t)^{-1} \nabla_t \eta$$

and hence, by Lemma D.3,

$$\left\| \xi - \xi_0 \right\|_p \le \kappa_p \varepsilon^{1/2} \left\| \nabla_t \xi - \eta \right\|_p + \varepsilon \left\| \nabla_t \eta \right\|_p$$

Similarly,

$$\eta - \nabla_t \xi_0 = (\mathbb{1} - \varepsilon \nabla_t \nabla_t)^{-1}(\eta - \nabla_t \xi) + (\varepsilon^2 - \varepsilon)(\mathbb{1} - \varepsilon \nabla_t \nabla_t)^{-1} \nabla_t \nabla_t \eta$$

and hence, again by Lemma D.3,

$$\varepsilon \|\eta - \nabla_t\xi_0\|_p \le \varepsilon \|\nabla_t\xi - \eta\|_p + \kappa_p\varepsilon^{3/2} \|\nabla_t\eta\|_p .$$

Take the sum of these two inequalities to obtain

$$\begin{aligned}\|\zeta - \iota\pi_\varepsilon\zeta\|_{0,p,\varepsilon} &\le \|\xi - \xi_0\|_p + \varepsilon \|\eta - \nabla_t\xi_0\|_p \\ &\le 2\kappa_p\varepsilon^{1/2} \|\nabla_t\xi - \eta\|_p + 2\kappa_p\varepsilon \|\nabla_t\eta\|_p\end{aligned}$$

for $0 < \varepsilon \le 1$. Moreover, using Lemma D.3 the formula for ξ_0 gives

$$\|\xi_0\|_p \le \|\xi\|_p + \kappa_p\varepsilon^{3/2} \|\eta\|_p , \qquad \varepsilon \|\nabla_t\xi_0\|_p \le \kappa_p\varepsilon^{1/2} \|\xi\|_p + 2\varepsilon^2 \|\eta\|_p .$$

Take these two inequalities to the power p and take the sum to obtain

$$\begin{aligned}\|\iota\pi_\varepsilon\zeta\|_{0,p,\varepsilon}^p &= \|\xi_0\|_p^p + \varepsilon^p \|\nabla_t\xi_0\|_p^p \\ &\le (1 + \kappa_p^p\varepsilon^{p/2}) \|\xi\|_p^p + (\kappa_p^p\varepsilon^{p/2} + 2^p\varepsilon^p)\varepsilon^p \|\eta\|_p^p \\ &\le (2\kappa_p)^p \|\zeta\|_{0,p,\varepsilon}^p\end{aligned}$$

for $0 < \varepsilon \le 1$. This proves Lemma D.6. $\square$

Proof of Proposition D.5. As above, denote

$$\xi_0 := \pi_\varepsilon\zeta = (\mathbb{1} - \varepsilon\nabla_t\nabla_t)^{-1}(\xi - \varepsilon^2\nabla_t\eta).$$

Then

$$\begin{aligned}\mathcal{D}_u^0\pi_\varepsilon\zeta &= \nabla_s\xi_0 - \nabla_t\nabla_t\xi_0 - R(\xi_0, \partial_t u)\partial_t u - \mathcal{H}_\mathcal{V}(u)\xi_0 \\ &= (\mathbb{1} - \varepsilon\nabla_t\nabla_t)^{-1} \left(\nabla_s\xi - \varepsilon^2\nabla_s\nabla_t\eta - \nabla_t\nabla_t\xi + \varepsilon^2\nabla_t\nabla_t\nabla_t\eta\right) \\ &\quad + \varepsilon(\mathbb{1} - \varepsilon\nabla_t\nabla_t)^{-1}[\nabla_s, \nabla_t\nabla_t]\xi_0 \\ &\quad + R\big((\mathbb{1} - \varepsilon\nabla_t\nabla_t)^{-1}\varepsilon^2\nabla_t\eta \,,\, \partial_t u\big)\partial_t u + \mathcal{H}_\mathcal{V}(u)(\mathbb{1} - \varepsilon\nabla_t\nabla_t)^{-1}\varepsilon^2\nabla_t\eta \\ &\quad - R\big((\mathbb{1} - \varepsilon\nabla_t\nabla_t)^{-1}\xi \,,\, \partial_t u\big)\partial_t u - \mathcal{H}_\mathcal{V}(u)(\mathbb{1} - \varepsilon\nabla_t\nabla_t)^{-1}\xi.\end{aligned}$$

Denote $\zeta' := (\xi', \eta') := \mathcal{D}_u^\varepsilon\zeta$, then

$$\begin{aligned}\pi_\varepsilon\mathcal{D}_u^\varepsilon\zeta &= (\mathbb{1} - \varepsilon\nabla_t\nabla_t)^{-1}(\xi' - \varepsilon^2\nabla_t\eta') \\ &= (\mathbb{1} - \varepsilon\nabla_t\nabla_t)^{-1}\big(\nabla_s\xi - R(\xi, \partial_t u)\partial_t u - \mathcal{H}_\mathcal{V}(u)\xi \\ &\quad - \varepsilon^2\nabla_t\nabla_s\eta - \varepsilon^2\nabla_t\big(R(\xi, \partial_s u)\partial_t u\big) - \nabla_t\nabla_t\xi\big).\end{aligned}$$

Taking the difference we find

$$\begin{aligned}&\mathcal{D}_u^0\pi_\varepsilon\zeta - \pi_\varepsilon\mathcal{D}_u^\varepsilon\zeta \\ &= (\mathbb{1} - \varepsilon\nabla_t\nabla_t)^{-1} \left(-\varepsilon^2[\nabla_s, \nabla_t]\eta + \varepsilon^2\nabla_t\nabla_t\nabla_t\eta + \varepsilon^2\nabla_t\big(R(\xi, \partial_s u)\partial_t u\big)\right) \\ &\quad + \varepsilon(\mathbb{1} - \varepsilon\nabla_t\nabla_t)^{-1}[\nabla_s, \nabla_t\nabla_t]\xi_0 \\ &\quad + R\big((\mathbb{1} - \varepsilon\nabla_t\nabla_t)^{-1}\varepsilon^2\nabla_t\eta \,,\, \partial_t u\big)\partial_t u + \mathcal{H}_\mathcal{V}(u)(\mathbb{1} - \varepsilon\nabla_t\nabla_t)^{-1}\varepsilon^2\nabla_t\eta \\ &\quad + (\mathbb{1} - \varepsilon\nabla_t\nabla_t)^{-1}R(\xi, \partial_t u)\partial_t u - R\big((\mathbb{1} - \varepsilon\nabla_t\nabla_t)^{-1}\xi \,,\, \partial_t u\big)\partial_t u \\ &\quad + (\mathbb{1} - \varepsilon\nabla_t\nabla_t)^{-1}\mathcal{H}_\mathcal{V}(u)\xi - \mathcal{H}_\mathcal{V}(u)(\mathbb{1} - \varepsilon\nabla_t\nabla_t)^{-1}\xi.\end{aligned} \tag{141}$$

To finish the proof it remains to inspect the L^p norm of this expression line by line. Using Lemma D.3, we obtain for the first line

$$\begin{aligned}&\left\|(\mathbb{1}-\varepsilon\nabla_t\nabla_t)^{-1}\left(-\varepsilon^2[\nabla_s,\nabla_t]\eta+\varepsilon^2\nabla_t\nabla_t\nabla_t\eta+\varepsilon^2\nabla_t\big(R(\xi,\partial_s u)\partial_t u\big)\right)\right\|_p\\ &\le \varepsilon^2\left\|R\right\|_\infty\left\|\partial_s u\right\|_\infty\left\|\partial_t u\right\|_\infty\left\|\eta\right\|_p+2\varepsilon\left\|\nabla_t\eta\right\|_p\\ &\quad+\kappa_p\varepsilon^{3/2}\left\|R\right\|_\infty\left\|\partial_s u\right\|_\infty\left\|\partial_t u\right\|_\infty\left\|\xi\right\|_p.\end{aligned}\tag{142}$$

Application of (133) with constant $C_1 := C$ results in an estimate for the second line in (141), namely

$$\left\|\varepsilon(\mathbb{1}-\varepsilon\nabla_t\nabla_t)^{-1}[\nabla_s,\nabla_t\nabla_t]\xi_0\right\|_p\le\varepsilon^{1/2}C_1\left\|\xi\right\|_p+\varepsilon^{5/2}C_1\left\|\nabla_t\eta\right\|_p.\tag{143}$$

Lemma D.3 yields for line three in (141)

$$\begin{aligned}&\left\|R\big((\mathbb{1}-\varepsilon\nabla_t\nabla_t)^{-1}\varepsilon^2\nabla_t\eta\,,\,\partial_t u\big)\partial_t u+\mathcal{H}_{\mathcal{V}}(u)(\mathbb{1}-\varepsilon\nabla_t\nabla_t)^{-1}\varepsilon^2\nabla_t\eta\right\|_p\\ &\le\big(\left\|R\right\|_\infty\left\|\partial_t u\right\|_\infty^2+C\big)\varepsilon^2\left\|\nabla_t\eta\right\|_p,\end{aligned}\tag{144}$$

where C is the constant in $(V1)$. Let us temporarily denote

$$T := \mathbb{1}-\varepsilon\nabla_t\nabla_t.$$

Then the penultimate line in (141) has the form $[T^{-1},\Phi]=T^{-1}[\Phi,T]T^{-1}$ where the endomorphism $\Phi : u^*TM\to u^*TM$ is given by $\Phi\xi=R(\xi,\partial_t u)\partial_t u$. This term can be expressed in the form

$$[T^{-1},\Phi]\xi=\varepsilon T^{-1}\left((\nabla_t\nabla_t\Phi)T^{-1}\xi+2(\nabla_t\Phi)T^{-1}\nabla_t\xi\right)$$

and hence

$$\left\|[T^{-1},\Phi]\xi\right\|_p\le\varepsilon^{1/2}\kappa_p C\left\|\xi\right\|_p.$$

Thus

$$\begin{aligned}&\left\|(\mathbb{1}-\varepsilon\nabla_t\nabla_t)^{-1}R(\xi,\partial_t u)\partial_t u-R\big((\mathbb{1}-\varepsilon\nabla_t\nabla_t)^{-1}\xi,\partial_t u\big)\partial_t u\right\|_p\\ &\le\varepsilon^{1/2}\kappa_p C_2\left\|\xi\right\|_p,\end{aligned}\tag{145}$$

where C_2 depends on $\left\|R\right\|_{C^2}$ $\left\|\partial_t u\right\|_\infty$, $\left\|\nabla_t\partial_t u\right\|_\infty$, and $\left\|\nabla_t\nabla_t\partial_t u\right\|_\infty$. Similarly,

$$\left\|(\mathbb{1}-\varepsilon\nabla_t\nabla_t)^{-1}\mathcal{H}_{\mathcal{V}}(u)\xi-\mathcal{H}_{\mathcal{V}}(u)(\mathbb{1}-\varepsilon\nabla_t\nabla_t)^{-1}\xi\right\|_p\le\varepsilon^{1/2}\kappa_p C_3\left\|\xi\right\|_p,\tag{146}$$

where C_3 depends on the constants in $(V1-V3)$ and on $\left\|\partial_t u\right\|_\infty$ and $\left\|\nabla_t\partial_t u\right\|_\infty$. The estimates (142-146) together give the desired L^p bound for (141) and this proves the first claim of Proposition D.5. The estimate for $(\mathcal{D}_u^0)^*\pi_\varepsilon\zeta-(\pi_\varepsilon\mathcal{D}_u^\varepsilon)^*\zeta$ follows analoguously. Since all constants appearing in the proof depend on L^∞ norms of derivatives of u, they are invariant under s-shifts of u. This completes the proof of Proposition D.5. □

The next lemma establishes the relevant estimates for the operator $\mathcal{D}_u^0$ and its adjoint in the Morse–Smale case, i.e. when $\mathcal{D}_u^0$ is onto.

Lemma D.7. *Let $\mathcal{V} : \mathcal{L}M \to \mathbb{R}$ be a perturbation that satisfies $(V0 - V4)$. Assume $\mathcal{S}_\mathcal{V}$ is Morse-Smale and let $u \in \mathcal{M}^0(x^-, x^+; \mathcal{V})$. Then, for every $p > 1$, there is a constant $c > 0$ such that*

$$\|\eta\|_p + \|\nabla_s\eta\|_p + \|\nabla_t\nabla_t\eta\|_p \le c\left\|(\mathcal{D}_u^0)^*\eta\right\|_p$$

and

$$\|\xi\|_p + \|\nabla_s\xi\|_p + \|\nabla_t\nabla_t\xi\|_p \le c\left(\left\|\xi - (\mathcal{D}_u^0)^*\eta\right\|_p + \left\|\mathcal{D}_u^0\xi\right\|_p\right)$$

for all compactly supported vector fields $\xi, \eta \in \Omega^0(\mathbb{R} \times S^1, u^*TM)$.

Proof. By Theorem A.4, the operators $\mathcal{D}_u^0$ and $(\mathcal{D}_u^0)^*$ are Fredholm. Since $\mathcal{S}_\mathcal{V}$ is Morse–Smale, the operator $\mathcal{D}_u^0$ is onto and $(\mathcal{D}_u^0)^*$ is injective. Moreover, the operator

$$\mathcal{W}_u^p \to \mathcal{L}_u^p \oplus \mathcal{L}_u^p/\mathrm{im}\,(\mathcal{D}_u^0)^* : \xi \mapsto (\mathcal{D}_u^0\xi, [\xi])$$

is also an injective Fredholm operator. Hence the estimates follow from the open mapping theorem. □

Proof of Theorem 3.3

Fix a constant $p > 1$. Then the L^∞ norms of $\partial_s u$, $\partial_t u$ and $\nabla_t\partial_t u$ are finite by Theorem A.1 and $\|\nabla_t\partial_s u\|_\infty$ is finite by Theorem A.2. Use the parabolic equations for u to conclude that $\|\nabla_t\nabla_t\partial_t u\|_\infty$ is finite as well. Hence we are in a position to apply Theorem 3.2 and Proposition D.5. We prove the estimate in two steps.

Step 1. *There are positive constants $c_1 = c_1(p)$ and $\varepsilon_0 = \varepsilon_0(p)$ such that*

$$\|\zeta\|_{0,p,\varepsilon} \le \|\xi\|_p + \varepsilon^{1/2}\|\eta\|_p \le c_1\left(\varepsilon\left\|(\mathcal{D}_u^\varepsilon)^*\zeta\right\|_{0,p,\varepsilon} + \left\|\pi_\varepsilon(\mathcal{D}_u^\varepsilon)^*\zeta\right\|_p\right) \tag{147}$$

for every $\varepsilon \in (0, \varepsilon_0)$ and every compactly supported vector field $\zeta = (\xi, \eta) \in \Omega^0(\mathbb{R} \times S^1, u^*TM \oplus u^*TM)$.

By Lemmata D.4 and D.7, there exists a constant $c_2 = c_2(p) > 0$ such that

$$\|\xi\|_p + \|\nabla_s\xi\|_p + \|\nabla_t\xi\|_p + \|\nabla_t\nabla_t\xi\|_p \le c_2\left\|(\mathcal{D}_u^0)^*\xi\right\|_p \tag{148}$$

for every compactly supported $\xi \in \Omega^0(\mathbb{R} \times S^1, u^*TM)$. Hence

$$\begin{aligned}
\|\xi\|_p &\le \|\xi - \pi_\varepsilon\zeta\|_p + \|\pi_\varepsilon\zeta\|_p \\
&\le \|\xi - \pi_\varepsilon\zeta\|_p + c_2\left\|(\mathcal{D}_u^0)^*\pi_\varepsilon\zeta\right\|_p \\
&\le \|\xi - \pi_\varepsilon\zeta\|_p + c_2\left\|(\mathcal{D}_u^0)^*\pi_\varepsilon\zeta - \pi_\varepsilon(\mathcal{D}_u^\varepsilon)^*\zeta\right\|_p + c_2\left\|\pi_\varepsilon(\mathcal{D}_u^\varepsilon)^*\zeta\right\|_p \\
&\le (\kappa_p + c_2c_3)\varepsilon\left(\varepsilon^{-1}\|\nabla_t\xi - \eta\|_p + \|\nabla_t\eta\|_p\right) + c_2\left\|\pi_\varepsilon(\mathcal{D}_u^\varepsilon)^*\zeta\right\|_p \\
&\quad + c_2c_3\left(\varepsilon^{1/2}\|\xi\|_p + \varepsilon^2\|\eta\|_p\right) \\
&\le (\kappa_p + c_2c_3)c_4\varepsilon\left\|(\mathcal{D}_u^\varepsilon)^*\zeta\right\|_{0,p,\varepsilon} + c_2\left\|\pi_\varepsilon(\mathcal{D}_u^\varepsilon)^*\zeta\right\|_p \\
&\quad + (c_2c_3 + \kappa_pc_4 + c_2c_3c_4)\left(\varepsilon^{1/2}\|\xi\|_p + \varepsilon^2\|\eta\|_p\right)
\end{aligned}$$

In the fourth step we have used Lemma D.6 and Proposition D.5 with a constant $c_3 = c_3(p) > 0$. The final step follows from Theorem 3.2 for the formal adjoint operator with a constant $c_4 = c_4(p) > 0$. Choose $\varepsilon_0 > 0$ so small that

$$(c_2c_3 + \kappa_p c_4 + c_2c_3c_4)\varepsilon_0{}^{1/2} < \frac{1}{2}. \tag{149}$$

Then we can incorporate the term $\|\xi\|_p$ into the left hand side and obtain

$$\|\xi\|_p \le 2(\kappa_p + c_2c_3)c_4\varepsilon \left\|(\mathcal{D}_u^\varepsilon)^*\zeta\right\|_{0,p,\varepsilon} + 2c_2 \left\|\pi_\varepsilon(\mathcal{D}_u^\varepsilon)^*\zeta\right\|_p + \varepsilon^{3/2}\|\eta\|_p . \tag{150}$$

Similarly,

$$\begin{aligned}
\|\eta\|_p &\le \|\eta - \nabla_t\pi_\varepsilon\zeta\|_p + \|\nabla_t\pi_\varepsilon\zeta\|_p \\
&\le \|\eta - \nabla_t\pi_\varepsilon\zeta\|_p + c_2\left\|(\mathcal{D}_u^0)^*\pi_\varepsilon\zeta\right\|_p \\
&\le (\kappa_p + c_2c_3\varepsilon^{1/2})c_4\varepsilon^{1/2}\left\|(\mathcal{D}_u^\varepsilon)^*\zeta\right\|_{0,p,\varepsilon} + c_2\left\|\pi_\varepsilon(\mathcal{D}_u^\varepsilon)^*\zeta\right\|_p \\
&\quad + (c_2c_3 + \kappa_p c_4 + c_2c_3c_4\varepsilon^{1/2})\left(\varepsilon^{1/2}\|\xi\|_p + \varepsilon^2\|\eta\|_p\right).
\end{aligned}$$

Use (149) again to obtain

$$\|\eta\|_p \le 2(\kappa_p + c_2c_3)c_4\varepsilon^{1/2}\left\|(\mathcal{D}_u^\varepsilon)^*\zeta\right\|_{0,p,\varepsilon} + 2c_2\left\|\pi_\varepsilon(\mathcal{D}_u^\varepsilon)^*\zeta\right\|_p + \|\xi\|_p . \tag{151}$$

The assertion of Step 1 now follows from (151) and (150).

Step 2 *We prove the theorem.*

Let $\varepsilon \in (0, \varepsilon_0)$. By (122) for the formal adjoint operator (with a constant $c_5 > 0$), we obtain

$$\begin{aligned}
\|\zeta\|_{1,p,\varepsilon} &\le c_5\varepsilon^2\left\|(\mathcal{D}_u^\varepsilon)^*\zeta\right\|_{0,p,\varepsilon} + c_5\|\zeta\|_{0,p,\varepsilon} \\
&\le c_5(\varepsilon^2 + c_1\varepsilon + 2\kappa_p c_1)\left\|(\mathcal{D}_u^\varepsilon)^*\zeta\right\|_{0,p,\varepsilon}
\end{aligned} \tag{152}$$

Here we have also used the estimate (147) of Step 1 and Lemma D.6. It follows that $(\mathcal{D}_u^\varepsilon)^*$ is injective and hence $\mathcal{D}_u^\varepsilon$ is onto.

Let $\zeta = (\xi, \eta) \in \Omega^0(\mathbb{R} \times S^1, u^*TM \oplus u^*TM)$ be compactly supported and denote

$$\zeta^* := (\xi^*, \eta^*) := (\mathcal{D}_u^\varepsilon)^*\zeta.$$

Recall that c_6 is the constant of Lemma D.7 and c_3 is the constant of Proposition D.5. By Lemma D.7, with $\xi = \pi_\varepsilon\zeta^*$ and $\eta = \pi_\varepsilon\zeta$, we have

$$\begin{aligned}
\|\pi_\varepsilon\zeta^*\|_p &\le c_6\left\|\pi_\varepsilon\zeta^* - (\mathcal{D}_u^0)^*\pi_\varepsilon\zeta\right\|_p + c_6\left\|\mathcal{D}_u^0\pi_\varepsilon\zeta^*\right\|_p \\
&\le c_6\left\|\pi_\varepsilon(\mathcal{D}_u^\varepsilon)^*\zeta - (\mathcal{D}_u^0)^*\pi_\varepsilon\zeta\right\|_p + c_6\left\|\mathcal{D}_u^0\pi_\varepsilon\zeta^* - \pi_\varepsilon\mathcal{D}_u^\varepsilon\zeta^*\right\|_p \\
&\quad + c_6\left\|\pi_\varepsilon\mathcal{D}_u^\varepsilon\zeta^*\right\|_p \\
&\le c_3c_6\left(\varepsilon^{1/2}\|\xi\|_p + \varepsilon^2\|\eta\|_p + \varepsilon\|\nabla_t\eta\|_p\right) + c_6\left\|\pi_\varepsilon\mathcal{D}_u^\varepsilon\zeta^*\right\|_p \\
&\quad + c_3c_6\left(\varepsilon^{1/2}\|\xi^*\|_p + \varepsilon^2\|\eta^*\|_p + \varepsilon\|\nabla_t\eta^*\|_p\right) \\
&\le 2c_3c_6(1 + c_4\varepsilon^{1/2})\varepsilon^{1/2}\|\zeta\|_{0,p,\varepsilon} + c_6\left\|\pi_\varepsilon\mathcal{D}_u^\varepsilon\zeta^*\right\|_p \\
&\quad + 3c_3c_6(1 + c_4\varepsilon^{1/2})\varepsilon^{1/2}\|\zeta^*\|_{0,p,\varepsilon} + c_3c_4c_6\varepsilon\left\|\mathcal{D}_u^\varepsilon\zeta^*\right\|_{0,p,\varepsilon} \\
&\le c_7\varepsilon^{1/2}\|\zeta^*\|_{0,p,\varepsilon} + c_3c_4c_6\varepsilon\left\|\mathcal{D}_u^\varepsilon\zeta^*\right\|_{0,p,\varepsilon} + c_6\left\|\pi_\varepsilon\mathcal{D}_u^\varepsilon\zeta^*\right\|_p .
\end{aligned} \tag{153}$$

The fourth step follows by applying Theorem 3.2 twice, with the constant c_4, namely for the operator $(\mathcal{D}_u^\varepsilon)^*$ to deal with the term $\nabla_t\eta$, and for the operator $\mathcal{D}_u^\varepsilon$ to deal with the term $\nabla_t\eta^*$. The final step follows from (152).

Now it follows from Lemma D.6 that

$$\begin{aligned}
\|\zeta^*\|_{0,p,\varepsilon} &\le \|\zeta^* - \iota\pi_\varepsilon\zeta^*\|_{0,p,\varepsilon} + \|\iota\pi_\varepsilon\zeta^*\|_{0,p,\varepsilon} \\
&\le 2\kappa_p\varepsilon\left(\varepsilon^{-1}\|\nabla_t\xi^* - \eta^*\|_p + \|\nabla_t\eta^*\|_p\right) + \|\pi_\varepsilon\zeta^*\|_p + \varepsilon\|\nabla_t\pi_\varepsilon\zeta^*\|_p \\
&\le 2\kappa_p c_4\varepsilon\|\mathcal{D}_u^\varepsilon\zeta^*\|_{0,p,\varepsilon} + \|\pi_\varepsilon\zeta^*\|_p + (2\kappa_p + 4\kappa_p c_4)\varepsilon^{1/2}\|\zeta^*\|_{0,p,\varepsilon} \\
&\le c_4(2\kappa_p + c_3c_6)\varepsilon\|\mathcal{D}_u^\varepsilon\zeta^*\|_{0,p,\varepsilon} + (c_7 + 2\kappa_p + 4\kappa_p c_4)\varepsilon^{1/2}\|\zeta^*\|_{0,p,\varepsilon} \\
&\quad + c_6\|\pi_\varepsilon\mathcal{D}_u^\varepsilon\zeta^*\|_p.
\end{aligned}$$

The third step follows from Theorem 3.2 for the operator $\mathcal{D}_u^\varepsilon$ and Lemma D.3. The final step uses (153). Choosing $\varepsilon_0 > 0$ sufficiently small, we obtain

$$\|\xi^*\|_p \le \|\zeta^*\|_{0,p,\varepsilon} \le 2c_4(2\kappa_p + c_3c_6)\varepsilon\|\mathcal{D}_u^\varepsilon\zeta^*\|_{0,p,\varepsilon} + 2c_6\|\pi_\varepsilon\mathcal{D}_u^\varepsilon\zeta^*\|_p. \tag{154}$$

By (122), we have

$$\|\zeta^*\|_{1,p,\varepsilon} \le c_5\left(\varepsilon^2\|\mathcal{D}_u^\varepsilon\zeta^*\|_{0,p,\varepsilon} + \|\zeta^*\|_{0,p,\varepsilon}\right).$$

Combining this with (154) we obtain (17).

We prove (16). By the triangle inequality and Lemmata D.6 and D.3, we have

$$\begin{aligned}
\|\eta^*\|_p &\le \|\eta^* - \nabla_t\pi_\varepsilon\zeta^*\|_{0,p,\varepsilon} + \left\|\nabla_t(\mathbb{1} - \varepsilon\nabla_t\nabla_t)^{-1}(\xi^* - \varepsilon^2\nabla_t\eta^*)\right\|_p \\
&\le \kappa_p\varepsilon^{1/2}\left(\varepsilon^{-1}\|\nabla_t\xi^* - \eta^*\|_p + \|\nabla_t\eta^*\|_p\right) + \kappa_p\varepsilon^{-1/2}\|\xi^*\|_p + 2\varepsilon\|\eta^*\|_p \\
&\le \kappa_p c_4\varepsilon^{1/2}\|\mathcal{D}_u^\varepsilon\zeta^*\|_{0,p,\varepsilon} + 2\kappa_p(1 + c_4\varepsilon)\varepsilon^{-1/2}\|\zeta^*\|_{0,p,\varepsilon}
\end{aligned}$$

The last step follows from Theorem 3.2 for the operator $\mathcal{D}_u^\varepsilon$. Similarly,

$$\begin{aligned}
\|\nabla_t\xi^*\|_p &\le \|\nabla_t\xi^* - \eta^*\|_p + \|\eta^*\|_p \\
&\le c_5\varepsilon\|\mathcal{D}_u^\varepsilon\zeta^*\|_{0,p,\varepsilon} + c_5\varepsilon\|\xi^*\|_p + (1 + c_5\varepsilon^3)\|\eta^*\|_p.
\end{aligned}$$

Combining the last two estimates with (154) proves (16). Since all constants appearing in the proof depend on L^∞ norms of derivatives of u, they are invariant under s-shifts of u. This proves Theorem 3.3.

Acknowledgement. Thanks to Katrin Wehrheim for pointing out to us the work of Marcinkiewicz and Mihlin and to Tom Ilmanen for providing the idea for the proof of Lemma B.2. Thanks to the referee for many helpful suggestions for improving the exposition.

References

[1] A. Abbondandolo and M. Schwarz, On the Floer homology of cotangent bundles, Preprint, August 2004. arxiv:math.SG/0408280, to appear in *Comm. Pure Appl. Math.*

[2] K. Cieliebak, Pseudo-holomorphic curves and periodic orbits on cotangent bundles, *J. Math. Pures Appl.* **73** (1994), 251–278.

[3] T. Davies, The Yang-Mills functional over Riemann surfaces and the loop group, PhD thesis, University of Warwick, 1996.

[4] S.K. Donaldson, *Floer Homology Groups in Yang–Mills Theory*, Cambridge University Press, 2002.

[5] S. Dostoglou and D.A. Salamon, Self-dual instantons and holomorphic curves, *Annals of Mathematics* **139** (1994), 581–640.

[6] A. Floer, A relative Morse index for the symplectic action, *Comm. Pure Appl. Math.* **41** (1988), 393–407.

[7] A. Floer, Wittens complex and infinite dimensional Morse theory, *J. Diff. Geom.* **30** (1989), 207–221.

[8] A. Floer, Symplectic fixed points and holomorphic spheres, *Comm. Math. Phys.* **120** (1989), 575–611.

[9] A. Floer and H. Hofer, Coherent orientations for periodic orbit problems in symplectic geometry, *Math. Zeit.* **212** (1993), 13–38. *Comm. Math. Phys.* **120** (1989), 575–611.

[10] A. Floer, H. Hofer, and D. Salamon, Transversality in elliptic Morse theory for the symplectic action, *Duke Math. Journal* **80** (1996), 251–292.

[11] M. Gromov Pseudo holomorphic curves in symplectic manifolds, *Invent. Math.* **82** (1985), 307–347.

[12] M. Gruber, Harnack inequalities for solutions of general second order parabolic equations and estimates of their Hölder constants, *Math. Z.* **185** (1984), 23–43.

[13] D. Gilbarg and N.S. Trudinger, *Elliptic partial differential equations of second order*, Grundlehren der mathematischen Wissenschaften **224**, Springer-Verlag 1977, third printing 1998.

[14] G.M. Lieberman, *Second order parabolic differential equations*, World Scientific, Singapore, 1996.

[15] O.A. Ladyženskaja, V.A. Solonnikov and N.N. Ural'ceva, *Linear and quasilinear equations of parabolic type*, Translations of Mathematical Monographs **23**, American Mathematical Society, 1968.

[16] D. McDuff and D.A. Salamon, *J-holomorphic curves and Symplectic Topology*, Colloquium Publications, Vol. **52**, American Mathematical Society, Providence, Rhode Island, 2004.

[17] J. Milnor, *Lectures on the h-Cobordism Theorem*, Princeton University Press, 1965.

[18] J. Milnor, *Morse Theory*, Princeton University Press, 1969.

[19] J.W. Robbin and D.A. Salamon, The spectral flow and the Maslov index, *Bull. London Math. Soc.* **27** (1995), 1–33.

[20] J.W. Robbin and D.A. Salamon, Asymptotic behaviour of holomorphic strips, *Annales de l'Institute Henri Poincaré - Analyse Nonlinéaire* **18** (2001), 573–612.

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

[22] D.A. Salamon, Lectures on Floer Homology, In *Symplectic Geometry and Topology*, edited by Y. Eliashberg and L. Traynor, IAS/Park City Mathematics Series, Vol **7**, 1999, pp 143–230.

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

[24] E. Stein, *Singular Integrals and Differentiability Properties of Functions*, Princeton University Press, 1970.

[25] C. Viterbo, Functors and computations in Floer homology with applications, Part II, Preprint, 1996, revised 2003.

[26] J. Weber, *Der Morse–Witten Komplex*, Diploma thesis, TU Berlin, 1993.

[27] J. Weber, *J-holomorphic curves in cotangent bundles and the heat flow*, PhD thesis, TU Berlin, 1999.

[28] J. Weber, Perturbed closed geodesics are periodic orbits: Index and transversality, *Math. Z.* **241** (2002), 45–81.

[29] J. Weber, The Morse-Witten complex via dynamical systems, Preprint, Universität München, November 2004. arxiv:math.GT/0411465, to appear in *Expositiones Mathematicae*.

[30] J. Weber, Noncontractible periodic orbits in cotangent bundles and Floer homology, Preprint, ETHZ, April 2004. arxiv:math.SG/0410609, to appear in *Duke Math. Journal*.

[31] J. Weber, The heat flow and the homology of the loop space, In preparation.