J-holomorphic curves in cotangent bundles and the heat flow

A dissertation submitted to
Technische Universität Berlin
for the degree of

Doktor der Naturwissenschaften
– Dr. rer. nat. –

presented by
Joachim Weber
Diplom-Mathematiker Diplom-Physiker
Technische Universität Berlin

accepted on the recommendation of
Prof. Dr. D. Salamon   Prof. Dr. R. Seiler   Prof. Dr. C. Viterbo

Berlin 1999

J-holomorphe Kurven in Kotangentialbündeln und die Wärmeleitungsgleichung

vorgelegt von
Diplom-Mathematiker Diplom-Physiker
Joachim Weber

vom Fachbereich 3 - Mathematik
der Technischen Universität Berlin
zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften
– Dr. rer. nat. –

genehmigte Dissertation

Berichter
Prof. Dr. D. Salamon Prof. Dr. R. Seiler Prof. Dr. C. Viterbo

Tag der wissenschaftlichen Aussprache
2. Juni 1999

Berlin 1999

D 83

For eleo

ABSTRACT. We obtain the gradient flow of the classical action functional on the free loop space ΛM of a closed Riemannian manifold (M, g) (where the Lagrangian is given by kinetic minus potential energy) as an adiabatic limit of the Floer gradient flow of the symplectic action on the free loop space of T^*M (where the Hamiltonian is given by kinetic plus potential energy). The limit is one where the metric on the momentum coordinate converges to zero. There is a natural correspondence between the critical points in both theories (perturbed geodesics) and we prove that their Morse indices equal minus their Conley-Zehnder indices. Nondegeneracy can be achieved by generic choice of an appropriate parameter - the potential energy.

The main result is an existence and uniqueness theorem for perturbed J-holomorphic curves nearby any trajectory of the heat flow between nondegenerate critical points of index difference 1. The proof is by a version of Newton's iteration method. Note that a crucial estimate has been left as a conjecture for $p > 2$. A proof for $p = 2$ is included.

Our result is a major step in establishing the existence of a natural isomorphism between Floer cohomology of the cotangent bundle of M and Morse homology of the classical action functional on ΛM, which in turn represents the singular homology of ΛM.

ZUSAMMENFASSUNG. Wir erhalten den (negativen) Gradientenfluß des klassischen Wirkungsfunktionals auf dem freien Schleifenraum ΛM einer geschlossenen Riemannschen Mannigfaltigkeit (M, g) – betrachtet als parabolisches Randwertproblem – als einen adiabatischen Limes des Floerschen Gradientenflusses auf dem Schleifenraum von T^*M – ein elliptisches Randwertproblem. In diesem Limes wird die vertikale Komponente der induzierten Metrik auf T^*M zu Null skaliert. Die kritischen Punkte in beiden Theorien können mit (gestörten) geschlossenen Geodätischen von (M, g) identifiziert werden und wir beweisen, daß deren Morse Indizes gleich den negativen Conley-Zehnder Indizes der entsprechenden 1-periodischen Orbits des geodätischen Flusses auf T^*M sind. Es stellt sich heraus, dass die Nichtdegeneriertheit der kritischen Punkte durch generische Wahl eines geeigneten Parameters erreicht werden kann.

Das Hauptresultat ist ein Existenz- und Eindeutigkeitstheorem für (gestörte) J-holomorphe Zylinder in T^*M nahe bei jeder Lösung des parabolischen Randwertproblems. Zum Beweis konstruieren wir eine Version der Newtonmethode für einen stetig differenzierbaren Schnitt in einem Banachraumbündel. Eine zentrale Abschätzung im hier relevanten Fall $p > 2$ ist als Vermutung formuliert. Wir geben einen Beweis für $p = 2$.

Unsere Resultate stellen einen wesentlichen Schritt im Beweis der Existenz eines kanonischen Isomorphismus zwischen der Floerkohomologie des Kotangentialbündels T^*M und der Morse Homologie für das klassische Wirkungsfunktional auf ΛM dar. Letztere wiederum repräsentiert die singuläre Homologie von ΛM.

Contents

CHAPTER 1

Introduction and main results

Let us first describe the background and context of this thesis, which originates in a joint research project with Dietmar Salamon. Our aim is to give a proof of theorem 1.0.3 by studying J-holomorphic curves in cotangent bundles and relating them to the heat flow of the underlying closed Riemannian manifold via an adiabatic limit argument. The main point is to establish a bijection between certain moduli spaces. In this thesis we prove injectivity. Further results are the index theorem and transversality for loops.

Let (M, g) be a compact smooth Riemannian manifold of dimension n and without boundary. ΛM denotes the free loop space of M consisting of absolutely continuous maps from S^1 to M. For $V \in C^\infty(S^1 \times M, \mathbb{R})$ and $x \in \Lambda M$ consider the classical action functional

$$\mathcal{I}_V(x) = \int_0^1 \Big(\tfrac{1}{2}|\dot{x}(t)|^2 - V(t, x(t)) \Big)\, dt$$

whose integrand is the Lagrangian $L_V : S^1 \times TM \to \mathbb{R}$. The set $Crit\, \mathcal{I}_V$ of critical points of $\mathcal{I}_V$ are the smooth loops x satisfying

$$-\nabla_t \dot{x} - \nabla V(t, x) = 0 \tag{1}$$

where ∇ is the Levi-Civita connection and ∇V the gradient with respect to the x-variable. For constant V these loops are the closed geodesics. Via the Legendre transformation the solutions of (1) can be interpreted as critical points of the symplectic action functional $\mathcal{A}_V$ on ΛT^*M

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

where $x(t)$ is the basepoint in M of $z(t)$ and the Hamiltonian $H_V : S^1 \times T^*M \to \mathbb{R}$ is given by

$$H_V(t, z) = \tfrac{1}{2}|z|^2 + V(t, x) \tag{2}$$

for $z \in T_x^*M$. A loop z in T^*M is a critical point of $\mathcal{A}_V$ iff the loop x of its basepoints in M solves (1) and $z(t) = g(x(t))\dot{x}(t)$ where in abuse of notation $g : TM \to T^*M$ also denotes the isomorphism provided by the Riemannian metric. For such loops both functionals agree.

Let $f : X \to \mathbb{R}$ be a Morse function on a compact Riemannian manifold X, i.e. its Hessian at any critical point is required to be nondegenerate. If the negative gradient flow of f is Morse-Smale (stable and unstable manifolds intersect transversally) then it gives rise to a Morse-Witten complex [**Wi82**] which is generated by the critical points and graded by the Morse index. The boundary operator is defined by counting the connecting orbits with index difference 1 (modulo 2 in the case of $\mathbb{Z}_2$ coefficients and otherwise with suitable signs which take account of the orientations). Full details may be found in the book by Matthias Schwarz [**Sch93**].

This principle applies equally well to the classical action $\mathcal{I}_V : \Lambda M \to \mathbb{R}$ and the symplectic action $\mathcal{A}_V : \Lambda T^*M \to \mathbb{R}$. In both cases the chain groups C_* are generated by the 1-periodic solutions $x : S^1 \to M$ of (1). Such a solution is nondegenerate as a critical point of $\mathcal{I}_V$ if and only if the corresponding loop $z : S^1 \to T^*M$ with $z = g(x)\dot{x}$ and $z(t) \in T^*_{x(t)}M$ is nondegenerate as a critical point of $\mathcal{A}_V$. Here $g(x) : T_xM \to T^*_xM$ denotes the isomorphism induced by the Riemannian metric. Each periodic solution $x \in Crit\,\mathcal{I}_V$ has finite Morse index $Ind(x)$ as a critical point of $\mathcal{I}_V$ (the number of negative eigenvalues of the Hessian, counted with multiplicity) and for a fixed level a we consider the chain groups

$$C_k^a = \bigoplus_{\substack{x \in Crit\,\mathcal{I}_V,\, \mathcal{I}_V(x) \le a, \\ Ind(x) = k}} \mathbb{Z}_2 x.$$

In theorem 6.2.1 we prove that $\mathcal{I}_V$ and $\mathcal{A}_V$ are Morse functions for generic potential V. We also assume that the L^2-gradient flows of $\mathcal{I}_V$ and $\mathcal{A}_V$ are Morse-Smale. In both cases this should again be achievable by a generic perturbation of the potential V. Under this assumption there are two boundary operators. There is $\partial^M : C_k^a \to C_{k-1}^a$, determined by the set $\mathcal{M}^0(x^-, x^+)$ of negative L^2-gradient flow lines of $\mathcal{I}_V$, i.e. smooth maps $u : \mathbb{R} \times S^1 \to M$ which satisfy

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

and

$$\lim_{s \to \mp\infty} u(s, t) = x^{\mp}(t). \tag{4}$$

The second boundary operator $\delta^F : C_k^a \to C_{k-1}^a$ is determined by the set $\mathcal{M}^\epsilon(x^-, x^+)$ of negative L^2-gradient flow lines of $\mathcal{A}_V$. These are smooth maps $w : \mathbb{R} \times S^1 \to T^*M$ which satisfy $w(s,t) \in T^*_{u(s,t)}M$,

$$\partial_s u - g(u)^{-1}\nabla_t w - \nabla V(t, u) = 0, \qquad \nabla_s w + \epsilon^{-2}\Big(g(u)\partial_t u - w\Big) = 0. \tag{5}$$

and

$$\lim_{s \to \mp\infty} w(s, t) = g(x^{\mp})\partial_t x^{\mp}. \tag{6}$$

As a result there are two homology theories, namely the Floer cohomology of T^*M [**F89b**], and the Morse-Witten homology of the classical action $\mathcal{I}_V$. They are denoted by

$$HF_a^{-*}(T^*M, H_V) = \frac{ker\ \delta^F}{im\ \delta^F}, \qquad HM_*^a(\Lambda M, \mathcal{I}_V) = \frac{ker\ \partial^M}{im\ \partial^M}.$$

As in the finite dimensional case we expect $HM_*^a(\Lambda M, \mathcal{I}_V)$ to be naturally isomorpic to the singular homology of the sublevel set

$$\Lambda^a M = \{x \in \Lambda M | \mathcal{I}_V(x) \leq a\}.$$

CONJECTURE 1.0.1. *There is a natural isomorphism*

$$HM_*^a(\Lambda M, \mathcal{I}_V) \cong H_*(\Lambda^a M, \mathbb{Z}_2).$$

On the other hand the negative L^2-gradient flow of the symplectic action $\mathcal{A}_V$ gives rise to Floer cohomology groups $HF_a^{-*}(T^*M, H_V)$. The main result in this thesis is a major step towards a proof of

CONJECTURE 1.0.2. *There is a natural isomorphism*

$$HF_a^{-*}(T^*M, H_V) \cong HM_*^a(\Lambda M, \mathcal{I}_V)$$

where $H_V : S^1 \times T^*M \to \mathbb{R}$ *is given by (2).*

Its proof relies on a bijection between certain moduli spaces. Injectivity – the implicit function theorem part – is the content of this thesis, while surjectivity – the compactness part – has not been worked out yet.

While working on this project we received a preprint by Claude Viterbo in which he proves by different methods that Floer cohomology of the cotangent bundle and the singular cohomology of the loop space are isomorphic. His proof relates both homology theories to Lisa Traynor's generating function homology [**T94**].

THEOREM 1.0.3 (Viterbo, [**V96**]). *There is an isomorphism*

$$HF_a^{-*}(T^*M) \simeq H_*^{sing}(\Lambda^a M).$$

The above conjectures together give rise to an alternative proof of Viterbo's theorem where, in addition, the isomorphism is natural. If M is not simply connected there is a separate isomorphism for each component of the loop space.

The proof of conjecture 1.0.1 will be analogous to the finite dimensional theorem which asserts that the homology of the Morse-Witten complex on a compact manifold agrees with the singular homology (cf.[**F89a, SZ92**]). The proof of conjecture 1.0.2 will be discussed below. Both results should be extendable to arbitrary coefficient rings if one takes account of the orientations of the moduli spaces of connecting orbits as is done in [**FH93**]. The details of the proof of conjecture 1.0.1 as well as of the orientation problem will be carried out elsewhere.

Our strategy to prove conjecture 1.0.2 is as follows: Both Morse-Witten homology $HM_*^a(\Lambda M, \mathcal{I}_V)$ and Floer cohomology $HF_a^{-*}(T^*M, H_V)$ arise from the same chain complex C_*^a generated by the solutions of (1) and graded by the Morse index. The index theorem 3.0.1 states that this equals minus the Conley-Zehnder index when viewed as a critical point of $\mathcal{A}_V$ and this explains the minus sign in the grading of Floer cohomology. It remains to compare the boundary operators and this will be done using a family of metrics on T^*M which scale the vertical component down to zero. This reduces the problem to the study of an adiabatic limit of a family of elliptic boundary value problems in T^*M approaching a parabolic one – the heat flow equation in M with perturbed closed geodesics as boundary data.

In this thesis the implicit function theorem part of this singular perturbation problem is studied: Given a parabolic solution we will identify a unique elliptic solution nearby. A compactness argument then is needed to establish the existence of a limit element of the elliptic families with sufficiently fast rate of convergence. This will be carried out in future research.

Statement of main results The proofs of the following results are based on conjecture 1.0.6 and work for any $p > 2$ such that $\kappa(p) \in (0,1)$. The conjecture is proven below for $p = 2$, in which case $\kappa(2) = 1/2$. Let exp denote the exponential map of (M, g) and

$$\mathcal{T}^*(X) : T^*_{u_0}M \to T^*_{exp_{u_0}X}M$$

parallel transport of covector fields along the curve $\tau \mapsto exp_{u_0}\tau X$. The moduli spaces $\mathcal{M}^0(x^-, x^+)$ and $\mathcal{M}^\epsilon(x^-, x^+)$ can be interpreted as zero sets of sections $\mathcal{F}_0$ and $\mathcal{F}_\epsilon$ of certain Banach space bundles. If its linearization $\mathcal{D}^0_{u_0}$ at a zero u_0 is onto we call u_0 regular. Define $w_0 = g(u_0)\partial_t u_0$ and denote by $\mathcal{D}^\epsilon_{w_0}$ the linearization of $\mathcal{F}_\epsilon$.

THEOREM 1.0.4. (**Existence**) *Assume Conjecture 1.0.6 below. Let $p > 2$ and choose nondegenerate $x^-, x^+ \in Crit\,\mathcal{I}_V$ as well as a parabolic cylinder $u_0 \in \mathcal{M}^0(x^-, x^+)$ such that $\mathcal{D}^0_{u_0}$ is onto. Then there exist constants $\epsilon_0, c > 0$ such that for any $\epsilon \in (0, \epsilon_0)$ the following is true: There exists an element $Z_\epsilon = (X_\epsilon, Y_\epsilon) \in im\,\mathcal{D}^{\epsilon\,*}_{w_0}$ such that*

$$T_\epsilon(u_0) := \mathcal{T}^*(X_\epsilon)\Big(g(u_0)\partial_t u_0 + Y_\epsilon\Big) \in \mathcal{M}^\epsilon(x^-, x^+)$$

and

$$\boxed{\|Z_\epsilon\|_{1,p,\epsilon} \le c\epsilon^2 \quad , \quad \|Z_\epsilon\|_{\infty,\epsilon} \le c\epsilon^{2-\frac{3}{2p}}.}$$

More precisely, $Z_\epsilon = (X_\epsilon, Y_\epsilon)$ satisfies

$$\boxed{\begin{aligned}
\|X_\epsilon\|_p + \|g(u_0)\nabla_t X_\epsilon - Y_\epsilon\|_p &\le c\,\epsilon^2 \\
\|Y_\epsilon\|_p + \|\nabla_t X_\epsilon\|_p &\le c\,\epsilon^{3/2} \\
\|\nabla_t Y_\epsilon\|_p + \|\nabla_s X_\epsilon\|_p &\le c\,\epsilon \\
\|\nabla_s Y_\epsilon\|_p &\le c\,\epsilon^{\min\{3/2-\kappa_p,1\}} \\
\|X_\epsilon\|_\infty + \|g(u_0)\nabla_t X_\epsilon - Y_\epsilon\|_\infty &\le c\,\epsilon^{2-\frac{3}{2p}} \\
\|Y_\epsilon\|_\infty + \|\nabla_t X_\epsilon\|_\infty &\le c\,\epsilon^{\frac{3}{2}-\frac{3}{2p}} \\
\|\nabla_t Y_\epsilon\|_\infty + \|\nabla_s X_\epsilon\|_\infty &\le c\,\epsilon^{1-\frac{3}{2p}} \\
\|\nabla_s Y_\epsilon\|_\infty &\le c\,\epsilon^{\min\{\frac{3}{2}-\kappa_p,\frac{9}{4}-2\kappa_p,1\}-\frac{3}{2p}}.
\end{aligned}}$$

THEOREM 1.0.5. (**Uniqueness**) *Assume Conjecture 1.0.6 below. Let $p > 2$ and fix nondegenerate $x^-, x^+ \in Crit\, \mathcal{I}_V$ as well as a parabolic cylinder $u_0 \in \mathcal{M}^0(x^-, x^+)$ such that $\mathcal{D}^0_{u_0}$ is onto. Then for any constant $c > 0$ there exists $\epsilon_0 > 0$ such that the following is true for any $\epsilon \in (0, \epsilon_0)$: If $Z = (X, Y) \in im\, \mathcal{D}^{\epsilon\, *}_{w_0}$ with*

$$\mathcal{T}^*(X)\Big(g(u_0)\partial_t u_0 + Y\Big) \in \mathcal{M}^\epsilon(x^-, x^+)$$

and

$$\boxed{\begin{aligned} \|X\|_\infty &\le c\, \epsilon^{\frac{5}{4} - \frac{3}{2p}} \\ \|Y\|_\infty + \|\nabla_t X\|_\infty &\le c\, \epsilon^{\frac{3}{4} - \frac{3}{2p}} \\ \|\nabla_t Y\|_\infty + \|\nabla_s X\|_\infty &\le c\, \epsilon^{\frac{1}{4} - \frac{3}{2p}} \end{aligned}}$$

then $Z = Z_\epsilon$, where Z_ϵ is the element provided by the existence theorem 1.0.4.

Hence we obtain for fixed nondegenerate x^-, $x^+ \in Crit\, \mathcal{I}_V$ of Morse index difference 1 a map

$$T_\epsilon \;:\; \mathcal{M}^0(x^-, x^+) \;\to\; \mathcal{M}^\epsilon(x^-, x^+) \quad , \; \epsilon > 0 \text{ sufficiently small}$$

which associates to every regular solution u_0 of the parabolic boundary value problem a solution $w_\epsilon := T_\epsilon(u_0)$ of the elliptic one. The existence theorem establishes the map T_ϵ and specifies the distance between u_0 and w_ϵ (strictly speaking between $g(u_0)\partial_t u_0$ and w_ϵ in a certain trivialization) to be *quadratic* in ϵ. The uniqueness theorem asserts that T_ϵ is well-defined. Together they show that T_ϵ is injective.

CONJECTURE 1.0.6. *Let $A : \mathbb{R} \times S^1 \to \mathbb{R}^{n \times n}$ be a differentiable family of skew-symmetric matrices such that for $s \to \mp\infty$*

$$A(s,t) \to A^\mp(t) \quad , \quad \partial_s A(s,t) \to 0$$

uniformly in t for skew-symmetric loops $A^\mp$. Then there exists a continuous function $\kappa : [2, \infty) \to \mathbb{R}$ with $\kappa(2) = 1/2$ and such that the following holds: For any $p \ge 2$ there exist $\epsilon_0 = \epsilon_0(p, A) > 0$ and $c = c(p, A) > 0$ such that

$$\begin{aligned} &\|\partial_s \vec{\xi}\|_p + \epsilon \|\partial_s \vec{\eta}\|_p \\ &\qquad \le c\Big(\|\partial_s \vec{\xi} - \partial_t \vec{\eta} - A\vec{\eta}\|_p + \epsilon \|\partial_s \vec{\eta} + \epsilon^{-2}(\partial_t \vec{\xi} + A\vec{\xi} - \vec{\eta})\|_p \\ &\qquad\quad + \epsilon^{-\kappa(p)} \Big(\|\vec{\xi}\|_p + \epsilon \|\vec{\eta}\|_p \Big) \Big) \end{aligned}$$

for $\epsilon \in (0, \epsilon_0)$ and $\vec{\xi}$, $\vec{\eta} \in C_0^\infty(\mathbb{R} \times S^1, \mathbb{R}^n)$. The same holds for ∂_s replaced by $-\partial_s$. We set $\|\cdot\|_p = \|\cdot\|_{L^p(\mathbb{R} \times S^1, \mathbb{R}^n)}$.

Let us discuss its proof for $p = 2$ to see that the skew-symmetry of A is an essential assumption:

$$\begin{aligned}
&\|\partial_s\vec{\xi} - \partial_t\vec{\eta} - A\vec{\eta}\|_2^2 + \epsilon^2\|\partial_s\vec{\eta} + \epsilon^{-2}(\partial_t\vec{\xi} + A\vec{\xi} - \vec{\eta})\|_2^2 \\
&\qquad = \|\partial_s\vec{\xi}\|_2^2 + \|\partial_t\vec{\eta} + A\vec{\eta}\|_2^2 - 2\langle\partial_s\vec{\xi}, \partial_t\vec{\eta} + A\vec{\eta}\rangle \\
&\qquad\quad + \epsilon^2\|\partial_s\vec{\eta}\|_2^2 + \epsilon^{-2}\|\partial_t\vec{\xi} + A\vec{\xi} - \vec{\eta}\|_2^2 + 2\langle\partial_s\vec{\eta}, \partial_t\vec{\xi} + A\vec{\xi} - \vec{\eta}\rangle \\
&\qquad = \|\partial_s\vec{\xi}\|_2^2 + \|\partial_t\vec{\eta} + A\vec{\eta}\|_2^2 - 2\langle\partial_s\vec{\xi}, A\vec{\eta}\rangle \\
&\qquad\quad + \epsilon^2\|\partial_s\vec{\eta}\|_2^2 + \epsilon^{-2}\|\partial_t\vec{\xi} + A\vec{\xi} - \vec{\eta}\|_2^2 + 2\langle\partial_s\vec{\eta}, A\vec{\xi}\rangle \\
&\qquad \geq \epsilon^{-2}\|\partial_t\vec{\xi} + A\vec{\xi} - \vec{\eta}\|_2^2 + \|\partial_t\vec{\eta} + A\vec{\eta}\|_2^2 + \|\partial_s\vec{\xi}\|_2^2 + \epsilon^2\|\partial_s\vec{\eta}\|_2^2 \\
&\qquad\quad - \|\partial_s A\|_\infty \frac{1}{\epsilon}\left(\|\vec{\xi}\|_2^2 + \epsilon^2\|\vec{\eta}\|_2^2\right)
\end{aligned}$$

where we used (partial integration)

$$\langle\partial_s\vec{\eta}, \vec{\eta}\rangle = -\langle\vec{\eta}, \partial_s\vec{\eta}\rangle = -\langle\partial_s\vec{\eta}, \vec{\eta}\rangle$$

and (partial integration and $[\partial_s, \partial_t] = 0$)

$$\begin{aligned}
\langle\partial_s\vec{\xi}, \partial_t\vec{\eta}\rangle &= -\langle\vec{\xi}, \partial_s\partial_t\vec{\eta}\rangle = -\langle\vec{\xi}, \partial_t\partial_s\vec{\eta}\rangle \\
&= \langle\partial_t\vec{\xi}, \partial_s\vec{\eta}\rangle
\end{aligned}$$

as well as (partial integration, $A^T = -A$, its asymptotic behavior and Young's inequality lemma 4.1.7)

$$\begin{aligned}
&-2\langle\partial_s\vec{\xi}, A\vec{\eta}\rangle + 2\langle\partial_s\vec{\eta}, A\vec{\xi}\rangle \\
&\qquad = 2\langle\vec{\xi}, (\partial_s A)\vec{\eta} + A\partial_s\vec{\eta}\rangle - 2\langle A\partial_s\vec{\eta}, \vec{\xi}\rangle \\
&\qquad \geq -\frac{2}{\sqrt{\epsilon}}\|\vec{\xi}\|_2 \, \|\partial_s A\|_\infty \, \sqrt{\epsilon}\|\vec{\eta}\|_2 \\
&\qquad \geq -2\|\partial_s A\|_\infty \left(\frac{1}{2\epsilon}\|\vec{\xi}\|_2^2 + \frac{\epsilon}{2}\|\vec{\eta}\|_2^2\right).
\end{aligned}$$

REMARK 1.0.7. 1) Let

$$\begin{aligned}
f &= \partial_s\vec{\xi} - \partial_t\vec{\eta} - A\vec{\eta} \\
g &= \partial_s\vec{\eta} + \epsilon^{-2}(\partial_t\vec{\xi} + A\vec{\xi} - \vec{\eta})
\end{aligned}$$

then conjecture 1.0.6 implies (add 0)

$$\begin{aligned}
&\epsilon^{-1}\|\partial_t\vec{\xi} + A\vec{\xi} - \vec{\eta}\|_p + \|\partial_t\vec{\eta} + A\vec{\eta}\|_p \\
&\qquad \leq \epsilon(c+1)\Big(\|\partial_s\vec{\xi} - \partial_t\vec{\eta} - A\vec{\eta}\|_p + \epsilon\|\partial_s\vec{\eta} + \epsilon^{-2}(\partial_t\vec{\xi} + A\vec{\xi} - \vec{\eta})\|_p\Big) \\
&\qquad\quad + c\epsilon^{-\kappa(p)}\Big(\|\vec{\xi}\|_p + \epsilon\|\vec{\eta}\|_p\Big).
\end{aligned}$$

2) The conjecture leads to a proof of the fundamental estimate, lemma 4.2.5.
3) If we can prove a modified Calderon-Zygmund estimate of the following

form: for any $p \geq 2$ there exists a constant $c_p > 0$ such that

$$\|\partial_s \vec{\xi}\|_p + \|\partial_s \vec{\eta}\|_p \leq c_p \Big(\|\partial_s \vec{\xi} - \partial_t \vec{\eta}\|_p + \|\partial_s \vec{\eta} + \partial_t \vec{\xi} - \vec{\eta}\|_p \Big)$$

for all $\vec{\xi}, \vec{\eta} \in C_0^\infty(\mathbb{R}^2, \mathbb{R}^n)$, then the same rescaling argument as in the proof of theorem 4.3.2 implies conjecture 1.0.6 with $\kappa \equiv 1$. In the above estimate $\|\cdot\|_p = \|\cdot\|_{L^p(\mathbb{R}^2,\mathbb{R}^n)}$.

Acknowledgements. Having been equipped with two advisors was of great help in many respects and I wish to thank them both. Let me start with the external one, Dietmar Salamon at Warwick University (now ETH Zürich), who greatly supported me with his immense enthusiasm for mathematics and his extraordinary insight and intuition in analysis and geometry. Also I do like to mention his generous special lectures on selected french wines and italian pasta. My internal advisor was Ruedi Seiler at TU Berlin and I am very grateful to him for attracting me to the field of mathematics through his inspiring lecture courses on mathematical physics and his preferences for geometric concepts in the world of physics as well as for his support over many years.

I would also like to mention the beautyful lectures of Udo Simon at TU Berlin on differential geometry which surely influenced the choice of my area of research. Many thanks to Helmut Hofer for inviting me for a term to ETH Zürich during which this research project was inspired. I am particularly grateful to my friend and colleague Kai Cielieak who encouraged me greatly and shared with me some of his insights in the depths of analysis. For hospitality during my stay I thank the Mathematics Institute at Warwick where the major part of this thesis has been carried out. I thank Ana Rita Gaio at Warwick for stimulating discussions on the quadratic estimates and Claude Viterbo for useful comments and critics on a preliminary version of this text.

Among the many people I met on my mathematical Odyssey over the past few years I would like to mention the following and thank them particularly for their friendship: Burkhard Nobbe in Berlin, Casim Abbas at Courant, Mihai Halic in Grenoble and Meike Akveld at ETH. Many thanks and best regards to Meike for carefully proofreading a substantial part of the manuscript.

Last not least I should like to thank the following institutions for partial financial support: European Union *TMR Programme Marie Curie Fellowship*, Graduiertenkolleg *Geometrie und nichtlineare Analysis* Humboldt-Universität Berlin, DAAD german academic exchange service Bonn, SFB 288 *Differentialgeometrie und Quantenphysik* Deutsche Forschungsgemeinschaft.

1.1. Morse theory on the loop space

We shall discuss in more detail the Hessians of the two variational problems, the gradient flow lines, and the linearized operators.

The Hessian. The Hessian of $\mathcal{I}_V$ at a critical point $x \in Crit\, \mathcal{I}_V$ is the perturbed Jacobi operator

$$A_x^0 : W^{2,2}(S^1, x^*TM) \to L^2(S^1, x^*TM)$$

given by

$$A_x^0 \xi = -\nabla_t \nabla_t \xi - R(\xi, \dot{x})\dot{x} - \nabla_\xi \nabla V(t, x).$$

Here $R \in \Omega^2(M, End TM)$ denotes the Riemann curvature tensor. For every $x \in Crit\, \mathcal{I}_V$ the operator A_x^0 has finitely many negative eigenvalues. The number of these, counted with multiplicity, is the Morse index $Ind(x)$.

To compare this with the Hessian of $\mathcal{A}_V$ it is convenient to identify the tangent space $T_z T^*M$ with the direct sum $T_x M \oplus T_x^* M$, where $z \in T_x^* M$, via the isomorphism which takes the derivative $\dot{z}$ of a path $z(t) \in T^*M$ to the pair $(\dot{x}, \nabla_t z)$. With this identification, which is studied in B.1.2, the Hessian of $\mathcal{A}_V$ at a critical point $z : S^1 \to T^*M$ with $z = g(x)\dot{x}$ is the operator

$$A_x^1 : W^{1,2}(S^1, x^*TM \oplus x^*T^*M) \to L^2(S^1, x^*TM \oplus x^*T^*M)$$

given by

$$A_x^1 \begin{pmatrix} \xi \\ \eta \end{pmatrix} = \begin{pmatrix} 0 & -g^{-1} \\ g & 0 \end{pmatrix} \begin{pmatrix} \nabla_t \xi \\ \nabla_t \eta \end{pmatrix} - \begin{pmatrix} R(\xi, \dot{x})\dot{x} + \nabla_\xi \nabla V \\ \eta \end{pmatrix}.$$

This operator is injective iff A_x^0 is injective and thus $\mathcal{I}_V$ is a Morse function iff $\mathcal{A}_V$ is a Morse function. However, while the critical points of $\mathcal{I}_V$ have finite Morse index, in the case of $\mathcal{A}_V$ both the Morse index and the coindex are infinite. But there is a relative Morse index, given by the spectral flow, and it is shown in [**SZ92**] that this agrees with the Conley-Zehnder index. This index can be canonically defined as follows. Choose an orthonormal trivialization of the bundle $x^*TM \to S^1$ and consider the dual trivialization of x^*T^*M (in the nonorientable case choose a trivialization over $[0, 1]$ with suitable boundary conditions). Via the above isomorphism these give rise to a unitary trivialization of the bundle z^*TT^*M where $z = g(x)\dot{x}$. Now the linearized Hamiltonian flow gives rise to a path of symplectic matrices $\Psi_x : [0, 1] \to Sp(2n)$ with $\Psi_x(0) = \mathbb{1}$ and $\det(\mathbb{1} - \Psi_x(1)) \neq 0$ by

$$\partial_t \Psi_x = -J_0 S(t) \Psi_x(t) \quad , \quad \Psi_x(0) = 0 \quad , \quad S(t) = \begin{pmatrix} Q(t) & 0 \\ 0 & \mathbb{1} \end{pmatrix} \tag{7}$$

where $S = S^t$ represents $\nabla H(t, z)$ in the unitary frame. The Conley-Zehnder index of x is defined to be the Conley-Zehnder index of Ψ_x (cf.[**CZ84, SZ92**]). In appendix D we provide an elementary discussion of the Conley-Zehnder index and visualize the concept in the case $n = 1$.

To get an idea of what relation between the Morse index and the Conley-Zehnder index of x to expect consider the simple case where $Q(t) = Q$ is a constant path of matrices with $\|Q\| < 2\pi$. The solution Ψ_x of (7) is then given by $\Psi_x(t) = e^{-tJ_0S}$ and hence $\mu_{CZ}(\Psi_x) = \frac{1}{2}\, sign\ S$ ([**SZ92**] Theorem 3.3 iv), where the *signature of* S is defined to be the number of its negative eigenvalues minus the number of its positive ones. As we derive in appendix A.4 the perturbed Jacobi operator A_x^0 may be represented in an orthonormal frame by the operator

$$I : W^{2,2}(S^1, \mathbb{R}^n) \to L^2(S^1, \mathbb{R}^n) \quad , \quad \vec{\xi} \mapsto -\partial_t\partial_t\vec{\xi} - Q\vec{\xi}.$$

Now observe that the number of negative eigenvalues of I coincides with $n^+(Q)$, the number of positive eigenvalues of Q. This leads to

$$\begin{aligned}\mu_{CZ}(\Psi_x) &= \tfrac{1}{2} sign\ S = n^-(S) - n = n^-(Q) - n \\ &= -n^+(Q) = -n^-(I) = -Ind(x).\end{aligned}$$

In the general case this still holds true and is the content of Theorem 3.0.1. Related questions have been studied previously by Duistermaat [**D76**] and Viterbo [**V90**] (cf. Chapter 3).

The gradient flow lines. Recall that the set $\mathcal{M}^0(x^-, x^+)$ of negative L^2-gradient flow lines of the energy functional $\mathcal{I}_V : \Lambda M \to \mathbb{R}$ are the smooth solutions of 3 and 4. The Morse-Smale condition implies that this is a manifold of dimension

$$\dim\ \mathcal{M}^0(x^-, x^+) = Ind(x^-) - Ind(x^+).$$

There is a free $\mathbb{R}$-action on $\mathcal{M}^0(x^-, x^+)$ and in the 1-dimensional case the quotient $\mathcal{M}^0(x^-, x^+)/\mathbb{R}$ is a finite set. Counting the connecting orbits gives rise to a boundary operator

$$\partial^M : C_k^a \to C_{k-1}^a$$

whose homology is denoted by $HM_*^a(\Lambda M, \mathcal{I}_V)$.

If we identify $T_zT^*M \cong T_xM \oplus T_x^*M$ as above then the negative L^2-gradient flow equation of the symplectic action $\mathcal{A}_V$ can be written in the form

$$\partial_s u - g(u)^{-1}\nabla_t w - \nabla V(t,u) = 0, \qquad \nabla_s w + g(u)\partial_t u - w = 0,$$

where $w : \mathbb{R} \times S^1 \to T^*M$ is smooth and $w(s,t) \in T^*_{u(s,t)}M$. These are the J-holomorphic curves of the title. The limit conditions take the form

$$\lim_{s\to\mp\infty} w(s,t) = g(x^\mp)\partial_t x^\mp$$

where $x^\mp \in Crit\,\mathcal{I}_V$. The set of such w is denoted by $\mathcal{M}^1(x^-, x^+)$. For a generic g and V this is a manifold of dimension

$$\dim\ \mathcal{M}^1(x^-, x^+) = \mu_{CZ}(x^+) - \mu_{CZ}(x^-)$$

where $\mu_{CZ}(x)$ denotes the Conley-Zehnder index of $x \in Crit\,\mathcal{I}_V$. The genericity statement will be subject of future research. Counting the connecting orbits in the case of index difference 1 gives rise to a boundary operator

$$\delta^F : C_k^a \to C_{k-1}^a$$

whose cohomology is denoted by $HF_a^{-*}(T^*M, H_V)$; cf. remark 1.2.1 Note that C_k^a is identified with the Floer cochain group CF_a^{-k} so that $\delta^F : CF_a^{-k} \to CF_a^{-k+1}$ increases the grading given by the Conley-Zehnder index by 1.

The linearized operators. Linearizing the gradient flow equation (3) of the classical action gives rise to the operator $\mathcal{D}_u^0 : C^\infty(\mathbb{R} \times S^1, u^*TM) \to C^\infty(\mathbb{R} \times S^1, u^*TM)$ given by

$$\mathcal{D}_u^0 \xi = \nabla_s \xi - \nabla_t \nabla_t \xi - R(\xi, \partial_t u)\partial_t u - \nabla_\xi \nabla V(t, u) \tag{8}$$

for $\xi \in C^\infty(\mathbb{R} \times S^1, u^*TM)$. This is a Fredholm operator between appropriate Sobolev completions. For example, if we define $\mathcal{H}^p = \mathcal{W}_u^{0,p}$ and $\mathcal{W}^{1,p} = \mathcal{W}_u^{1,p}$ as the completions of $C_0^\infty(\mathbb{R} \times S^1, u^*TM)$ with respect to the norms

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

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

then $\mathcal{D}_u^0 : \mathcal{W}^{1,p} \to \mathcal{H}^p$ is a Fredholm operator with index

$$Ind\, \mathcal{D}_u^0 = Ind(x^-) - Ind(x^+).$$

If this operator is surjective for all $u \in \mathcal{M}^0(x^-, x^+)$ then the implicit function theorem asserts that the space $\mathcal{M}^0(x^-, x^+)$ is a smooth manifold whose tangent space at u is the kernel of $\mathcal{D}_u^0$ and whose dimension therefore equals the Fredholm index of $\mathcal{D}_u^0$. A reference is Theorem A in [**RS93**] where the Fredholm index is expressed via the spectral flow which is the index difference in the case at hand.

Linearizing the gradient flow equation (5) of the symplectic action gives rise to the first order differential operator

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

given by

$$\begin{aligned} \mathcal{D}_w^1 \begin{pmatrix} \xi \\ \eta \end{pmatrix} &= \begin{pmatrix} \nabla_s \xi \\ \nabla_s \eta \end{pmatrix} + \begin{pmatrix} 0 & -g^{-1} \\ g & 0 \end{pmatrix} \begin{pmatrix} \nabla_t \xi \\ \nabla_t \eta \end{pmatrix} \\ &\quad + \begin{pmatrix} -R(\xi, \partial_t u) g^{-1} w - \nabla_\xi \nabla V \\ gR(\xi, \partial_s u) g^{-1} w - \eta \end{pmatrix} \end{aligned}$$

for $\xi \in C_0^\infty(\mathbb{R} \times S^1, u^*TM)$ and $\eta \in C_0^\infty(\mathbb{R} \times S^1, u^*T^*M)$. This formula is calculated in section A.2. The operator $\mathcal{D}_w^1$ is Fredholm for every $w \in \mathcal{M}^1(x^-, x^+)$ and its index is given by

$$Ind\, \mathcal{D}_w^1 = \mu_{CZ}(x^+) - \mu_{CZ}(x^-).$$

We expect that $\mathcal{D}_w^1$ can be made surjective for all solutions of (5) and (6) by a generic perturbation of V. If this is the case then, by the implicit function theorem, the space $\mathcal{M}^1(x^-, x^+)$ is a smooth manifold whose tangent space at $w \in \mathcal{M}^1(x^-, x^+)$ is the kernel of the operator $\mathcal{D}_w^1$. The two moduli spaces of connecting orbits are genuinely different but in the next section we show how to deform the equations (5) into (3) by a change of the metric on T^*M or equivalently by a change of the almost complex structure J compatible with the metric.

1.2. The adiabatic limit

Exploiting the independence of Floer cohomology of the compatible almost complex structure $\Omega = -d\Theta$ on T^*M [**F89b**], [**SZ92**] we may, in the identification $T_zT^*M = T_xM \oplus T_xM$, choose a family

$$J_\epsilon = \begin{pmatrix} 0 & -\epsilon g^{-1} \\ \epsilon^{-1} g & 0 \end{pmatrix}$$

which is compatible with Ω, i.e. $\Omega(\cdot, J_\epsilon \cdot) = G_\epsilon$ is a Riemannian metric on T^*M, which rescales the horizontal component by the factor ϵ^{-1} and the vertical component by the factor ϵ. The space $\mathcal{M}^1(x^-, x^+, J_\epsilon)$ of solutions w with $w(s,t) \in T^*_{u(s,t)}M$ of

$$\partial_s \tilde{u} - \epsilon g(\tilde{u})^{-1} \nabla_t \tilde{w} - \epsilon \nabla V(t, \tilde{u}) = 0, \qquad \nabla_s \tilde{w} + \epsilon^{-1}\Big(g(\tilde{u})\partial_t \tilde{u} - \tilde{w}\Big) = 0.$$

with boundary condition (6) corresponds, via rescaling $w(s,t) = \tilde{w}(\epsilon^{-1}s, t)$, naturally to $\mathcal{M}^\epsilon(x^-, x^+)$ the space of solutions of (6) and

$$\partial_s u - g(u)^{-1} \nabla_t w - \nabla V(t, u) = 0, \qquad \nabla_s w + \epsilon^{-2}\Big(g(u)\partial_t u - w\Big) = 0. \tag{9}$$

Although the spaces $\mathcal{M}^1(x^-, x^+) = \mathcal{M}^1(x^-, x^+, J_1)$ and $\mathcal{M}^1(x^-, x^+, J_\epsilon)$ might be different, the resulting Floer cohomology groups $HF_a^{-*}(T^*M, H_V)$ and $HF_a^{-*}(T^*M, H_V, J_\epsilon)$ are naturally isomorphic and so it suffices to study $\mathcal{M}^\epsilon(x^-, x^+)$ in order to compare the boundary operators.

REMARK 1.2.1. The construction of Floer homology for a *compact* symplectic manifold subject to certain topological constraints in order to deal with the possible presence of J-holomorphic spheres is standard (see [**Sa97**] for a beautyful exposition). Although $(T^*M, \Omega = -d\Theta)$ is not compact, it exhibits two nice features. Firstly, the existence of a global Lagrangian splitting of TT^*M allows for a natural normalization of the Conley-Zehnder index of critical points of the symplectic action. Secondly, the exactness of Ω excludes the existence of nontrivial J-holomorphic spheres and so one may use the integers as coefficient ring and, more importantly, standard bubbling-off analysis leads to uniform C^1-bounds for the solutions of Floer's elliptic boundary value problem. However, first one needs a C^0-bound which in the compact case is trivial and in the present case of T^*M and a Hamiltonian *quadratic at infinity* has been established by Kai Cieliebak [**Ci94**], theorem 5.4. The same bound holds uniformly for all solutions on which the symplectic action takes values in a fixed interval. An essential tool in his proof is lemma 5.3 which says that

$$\{z \in W^{1,2}(S^1, T^*M) \mid \mathcal{A}_V(z) \leq a, \|L^2 - grad\, \mathcal{A}_V(z)\|_{L^2}^2 \leq b\}$$

is bounded in the $W^{1,2}$-norm by a constant $c = c(a,b,V)$. In particular, the lemma gives a uniform C^0-bound for all critical points of action at most a, and so implies – in view of their nondegeneracy – that there is only a finite number of them. Therefore the chain groups C_k^a are well-defined.

We shall study the limit behaviour of the solutions of (9) as $\epsilon \to 0$. This is a singular perturbation problem. Heuristically, one expects the solutions to converge to elements w which satisfy the first equation in (9) and where the second equation is replaced by $w = g(u)\partial_t u$. But these are exactly the solutions of (3). In other words the solutions of the elliptic equation (9) degenerate in the small ϵ limit to the solutions of the parabolic equation (3). Strong evidence for this limit behaviour comes from the energy identity

$$\begin{aligned}\mathcal{E}_\epsilon(w) &= \frac{1}{2}\int_{-\infty}^{\infty}\int_0^1 \left(|\partial_s u|^2 + |g^{-1}\nabla_t w + \nabla V(t,u)|^2\right) dtds \\ &\quad + \frac{1}{2}\int_{-\infty}^{\infty}\int_0^1 \left(\epsilon^2|\nabla_s w|^2 + \epsilon^{-2}|g\partial_t u - w|^2\right) dtds \\ &= \mathcal{I}_V(x^-) - \mathcal{I}_V(x^+)\end{aligned} \tag{10}$$

for the solutions of (9) and (6).

PROOF. Use in the second step that w solves (9), condition (6) in the third one and set $z^\mp = g(x^\mp)\partial_t x^\mp$ to obtain

$$\begin{aligned}\mathcal{E}_\epsilon(w) &= \int_{-\infty}^{\infty}\int_0^1 G_\epsilon\Big(\underbrace{\tfrac{1}{\epsilon}J_\epsilon\partial_t w - \tfrac{1}{\epsilon}{}^{G_\epsilon}\nabla H(t,w)}_{=-\partial_s w}, J_\epsilon\partial_t w - {}^{G_\epsilon}\nabla H(t,w)\Big)\, dtds \\ &= \int_0^1\int_{-\infty}^{\infty} \Omega(\partial_s w, \partial_t w) + \tfrac{d}{ds}H(t,w)\, dsdt \\ &= -\int_{\mathbb{R}\times S^1} w^* d\Theta + \int_0^1 H(t,z^+) - H(t,z^-)\, dt \\ &= \int_{S^1}(z^-)^*\Theta - \int_{S^1}(z^+)^*\Theta + \int_0^1 H(t,z^+) - H(t,z^-)\, dt \\ &= \mathcal{A}_V(z^-) - \mathcal{A}_V(z^+) = \mathcal{I}_V(x^-) - \mathcal{I}_V(x^+).\end{aligned}$$

Note that in the last but one step we used Stokes theorem and the fact $\partial(\mathbb{R}\times S^1) = -(-\infty\times S^1) \sqcup (+\infty \times S^1)$, where the minus sign in front of the first term indicates a change of orientation. □

The proof of conjecture 1.0.2 is based on establishing a bijection between the space $\mathcal{M}^\epsilon(x^-,x^+)$ of solutions of (9) and (6) and the space $\mathcal{M}^0(x^-,x^+)$ for small $\epsilon > 0$. The main idea is to think of the solutions of (3) as approximate solutions of (9) for ϵ small and to use the implicit function theorem to prove the existence of a nearby true solution. This is the content of chapter

2. Here the crucial ingredient is to establish the invertibility of the linearized operator (on the range of its formal adjoint operator)

$$\mathcal{D}_w^\epsilon : W_\epsilon^{1,p}(\mathbb{R} \times S^1, u^*TM \oplus u^*T^*M) \to L_\epsilon^p(\mathbb{R} \times S^1, u^*TM \oplus u^*T^*M)$$

for the ϵ-equation. The formula for this operator is derived in appendix A.2

$$\mathcal{D}_w^\epsilon \begin{pmatrix} \xi \\ \eta \end{pmatrix} = \begin{pmatrix} \nabla_s \xi - g^{-1} \nabla_t \eta - R(\xi, \partial_t u) g^{-1} w - \nabla_\xi \nabla V \\ \nabla_s \eta + \epsilon^{-2} g \nabla_t \xi + g R(\xi, \partial_s u) g^{-1} w - \epsilon^{-2} \eta \end{pmatrix}$$

for $\xi \in C_0^\infty(\mathbb{R} \times S^1, u^*TM)$ and $\eta \in C_0^\infty(\mathbb{R} \times S^1, u^*T^*M)$.

Let us fix a solution u_0 of (3) and define

$$w_0 = g(u_0)\partial_t u_0.$$

For w_0 we must prove that the operator $\mathcal{D}_{w_0}^\epsilon$ is onto for $\epsilon > 0$ sufficiently small and prove an estimate for the right inverse which is independent of ϵ. We will establish this in Theorem 4.4.4 under the assumption that the operator $\mathcal{D}_{u_0}^0$ is onto. To obtain uniform estimates for the inverse with constants independent of ϵ we must work with suitable ϵ-dependent norms. These are suggested by comparing powers of ϵ appearing in the energy identity (10). For compactly supported sections $\xi \in C^\infty(\mathbb{R} \times S^1, u^*TM)$ and $\eta \in C^\infty(\mathbb{R} \times S^1, u^*T^*M)$ define

$$\|(\xi, \eta)\|_{0,p,\epsilon} = \left(\int_{-\infty}^{\infty} \int_0^1 |\xi|^p + \epsilon^p |\eta|^p \, dt ds \right)^{1/p} \tag{11}$$

and

$$\|(\xi, \eta)\|_{1,p,\epsilon}^p = \|(\xi, \eta)\|_{0,p,\epsilon}^p + \|(\nabla_t \xi, \nabla_t \eta)\|_{0,p,\epsilon}^p + \|(\nabla_s \xi, \nabla_s \eta)\|_{0,p,\epsilon}^p . \tag{12}$$

To indicate the presence of these new norms on the standard Sobolev spaces $W^{1,p}$ and L^p we denote them by $W_e^{1,p} ps$ and L_ϵ^p, respectively.

Geometrically, the difference between the operators $\mathcal{D}_u^0$ and $\mathcal{D}_w^\epsilon$ is the difference between configuration space and phase space, or between loops in M and loops in T^*M. Consider the embedding

$$\Lambda M \to \Lambda T^*M : x \mapsto (x, g(x)\dot{x}).$$

The differential of this embedding is given by

$$C^\infty(S^1, x^*TM) \to C^\infty(S^1, x^*TM \oplus x^*T^*M) : \xi \mapsto (\xi, g(x)\nabla_t \xi).$$

To compare the operators $\mathcal{D}_u^0$ and $\mathcal{D}_w^\epsilon$ we must choose a projection onto the image of this embedding (along u). A natural candidate would be the orthogonal projection $\pi_\epsilon^\perp$ with respect to the Hilbert space structure (11). This would be given by $(\xi, \eta) \mapsto (\mathbb{1} - \epsilon^2 \nabla_t \nabla_t)^{-1}(\xi - \epsilon^2 g^{-1} \nabla_t \eta)$. Instead we introduce the projection operator $\pi_\epsilon : W_\epsilon^{1,p}(\mathbb{R} \times S^1, u^*TM \oplus u^*T^*M) \to \mathcal{W}^{1,p}(\mathbb{R} \times S^1, u^*TM)$ given by (cf. section 4.1)

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

We denote by $\iota : \mathcal{W}^{1,p}(\mathbb{R} \times S^1, u^*TM) \to W_\epsilon^{1,p}(\mathbb{R} \times S^1, u^*TM \oplus u^*T^*M)$ the inclusion

$$\iota \xi_0 = (\xi_0, g \nabla_t \xi_0).$$

We refer to chapter 4 for precise statements about the relevant estimates and their proofs.

1.3. Overview

As a matter of fact the order of presentation is essentially reverse to the order in which the project has been worked through.

In Chapter 2 we present the proof of the main results – existence and uniqueness of elliptic cylinders nearby parabolic ones – by carrying out a Newton-type iteration method for the Banach space bundle section $\mathcal{F}_\epsilon$ represented in a suitable trivialization. Key ingredients are the uniform bound on the right inverse of its linearization derived in chapter 4 as well as the quadratic estimates of chapter 5.

Appendix A provides analytical results on the exponential map and parallel transport required to optimize the quadratic estimates. Moreover, the linearized operators and the representation of the section with respect to a local trivialization are calculated.

Chapters 3 on the index theorem and 6 on transversality theory stand on their on and can be read independently.

Appendices B, C and D recollect basic facts about the two variational theories at hand, Newton's iteration method and the topology of the symplectic linear group $Sl(2,\mathbb{R})$, respectively. They may be a starting point for the novice. Certainly they were for me.

CHAPTER 2

The approximation Theorem

This section is at the heart of the thesis as we combine the elliptic estimates obtained in chapter 4 and the quadratic estimates of chapter 5 to carry out the iteration leading to the main theorems 1.0.4 on existence and 1.0.5 on uniqueness of elliptic cylinders nearby parabolic ones.

The strategy is to consider $w_0 := g(u_0)\partial_t u_0$ as an approximate zero of the section $\mathcal{F}_\epsilon$ of a Banach space bundle $\mathcal{E}^p$ over a Banach manifold $\mathcal{P}^{1,p}_{x^-,x^+}$ and then carry out Newton's iteration method in order to find a true zero nearby.

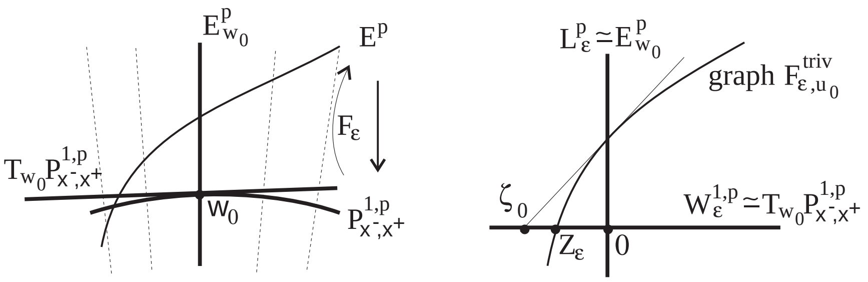

FIGURE 2.1. Local trivialization of $\mathcal{E}^{1,p} \to \mathcal{P}^{1,p}_{x^-,x^+}$ at w_0

To define the Banach manifold we follow [**FHS95**] and fix a number $p > 2$, two loops $x^-, x^+ \in \Lambda M$ and choose trivializations

$$\phi^\mp(t) : \mathbb{R}^n \to T_{x^\mp(t)}M$$

with $\phi^\mp(t+1) = \phi^\mp(t)$. Denote by

$$\mathcal{P} = \mathcal{P}^{1,p}_{x^-,x^+}$$

the space of continuous maps $w : \mathbb{R} \times S^1 \to T^*M$ which satisfy (6), are locally of class $W^{1,p}$, and satisfy ξ^-, $\eta^- \in W^{1,p}([T,\infty) \times S^1, \mathbb{R}^n)$ and ξ^+, $\eta^+ \in W^{1,p}((-\infty, T] \times S^1, \mathbb{R}^n)$, where $\xi^\mp$, $\eta^\mp$ are defined by

$$exp_{x^\mp(t)}\Big(\phi^\mp(t)\xi^\mp(s,t)\Big) = u(s,t)$$

and

$$\mathcal{T}^*(\xi^\mp(s,t))\Big(g(x^\mp(t))\partial_t x^\mp(t) + \eta^\mp(t)\Big) = w(s,t)$$

for $\mp s \geq T$ with T sufficiently large. Of course, the first condition is simply the base part of the second one. Here *exp* denotes the exponential map of (M, g) and

$$\mathcal{T}^*(\xi) : T^*_{u_0}M \to T^*_{exp_{u_0}\xi}M$$

denotes parallel transport of covector fields along the geodesic $\tau \mapsto exp_u \tau\xi$. The space $\mathcal{P}$ is an infinite-dimensional Banach manifold with tangent space

$$T_w\mathcal{P} = W^{1,p}(\mathbb{R} \times S^1, u^*TM \oplus u^*T^*M).$$

The fibre of the Banach space bundle $\mathcal{E}^p \to \mathcal{P}$ over $w \in \mathcal{P}$ is the space

$$\mathcal{E}^p_w = L^p(\mathbb{R} \times S^1, u^*TM \oplus u^*T^*M).$$

A standard reference for Banach bundles and manifolds is Eliasson [**E67**].

As we need to work in a Banach *space* setting, we have to trivialize the Banach bundle $\mathcal{E}^p \to \mathcal{P}^{1,p}_{x^-,x^+}$ locally at w_0 and then study the induced operator $\mathcal{F}^{triv}_{\epsilon,u_0}$ (figure 2.1). To define the local trivialization of $\mathcal{E}^p$ around w_0 we use again the exponential map on (M, g) and the parallel transport

$$\mathcal{T}(X) : T_{u_0}M \to T_{exp_{u_0}X}M$$

along the geodesic $\tau \mapsto exp_u \tau X$. Note that in local coordinates $\mathcal{T}(X)^{T^{-1}} = \mathcal{T}^*(X)$. $\mathcal{F}_\epsilon$ is then represented by the following nonlinear map between Banach spaces

$$\begin{aligned} \mathcal{F}^{triv}_{\epsilon,u_0} \; &: \; W^{1,p}_\epsilon(\mathbb{R} \times S^1, u_0^*TM \oplus u_0^*T^*M) \to L^p_\epsilon(\mathbb{R} \times S^1, u_0^*TM \oplus u_0^*T^*M) \\ Z = (X, Y) &\mapsto \begin{pmatrix} \mathcal{T}(X)^{-1} & 0 \\ 0 & \mathcal{T}(X)^* \end{pmatrix} \mathcal{F}_\epsilon\Big(\mathcal{T}(X)^{*-1}(w_0 + Y)\Big). \end{aligned}$$

The basepoint of $\mathcal{T}(X)^{*-1}(w_0 + Y)$ is given by $exp_{u_0}X$.

As is discussed in great detail in appendix C, the Newton method is an inductive process and there are essentially three ingredients that have to be controlled: A small initial value $\mathcal{F}^{triv}_{\epsilon,u_0}(0)$, a uniformly (in ϵ) bounded right inverse $Q^\epsilon_{w_0}$ of $d\mathcal{F}^{triv}_{\epsilon,u_0}(0) = \mathcal{D}^\epsilon_{w_0}$ and the variation of derivatives $d\mathcal{F}^{triv}_{\epsilon,u_0}(Z) - \mathcal{D}^\epsilon_{w_0}$. The bound on the right inverse is expressed by the *key estimate* for $\mathcal{D}^\epsilon_{w_0}$ on the range of ${\mathcal{D}^\epsilon_{w_0}}^*$ and control on the variation of derivatives is gained by the quadratic estimates in chapter 5.

In appendix A, Theorem A.3.1, it is shown that

$$d\mathcal{F}^{triv}_{\epsilon,u_0}(0, 0) = \mathcal{D}^\epsilon_{w_0}.$$

A right inverse of $\mathcal{D}^\epsilon_{w_0}$ will be defined in section 4.4 by

$$\begin{aligned} \mathcal{Q}^\epsilon_{w_0} \; &: \; L^p_\epsilon(\mathbb{R} \times S^1, u_0^*TM \oplus u_0^*T^*M) \to W^{1,p}_\epsilon(\mathbb{R} \times S^1, u_0^*TM \oplus u_0^*T^*M) \\ &\qquad Z \mapsto {\mathcal{D}^\epsilon_{w_0}}^*(\mathcal{D}^\epsilon_{w_0}{\mathcal{D}^\epsilon_{w_0}}^*)^{-1} Z \end{aligned}$$

where ${\mathcal{D}^\epsilon_{w_0}}^*$ is the formal adjoint of $\mathcal{D}^\epsilon_{w_0}$ with respect to the L^2_ϵ-inner product. As mentioned above, the main tools to estimate the right inverse are the *key estimates*, Theorem 4.4.4, for $\mathcal{D}^{\epsilon}_{w_0}$ on the range of ${\mathcal{D}^{\epsilon}_{w_0}}^*$ – here the surjectivity of $\mathcal{D}_{u_0}$ is used –

$$\begin{aligned} \|\mathcal{D}^{\epsilon\,*}_{w_0}\zeta\|_{1,p,\epsilon} &\le c_3\left(\epsilon\|\mathcal{D}^{\epsilon}_{w_0}\mathcal{D}^{\epsilon\,*}_{w_0}\zeta\|_{0,p,\epsilon} + \|\pi_\epsilon\mathcal{D}^{\epsilon}_{w_0}\mathcal{D}^{\epsilon\,*}_{w_0}\zeta\|_p\right) \\ \|\eta^*\|_p + \|\nabla_t\xi^*\|_p &\le c_3\left(\epsilon^{1/2}\|\mathcal{D}^{\epsilon}_{w_0}\mathcal{D}^{\epsilon\,*}_{w_0}\zeta\|_{0,p,\epsilon} + \epsilon^{-1/2}\|\pi_\epsilon\mathcal{D}^{\epsilon}_{w_0}\mathcal{D}^{\epsilon\,*}_{w_0}\zeta\|_p\right) \end{aligned} \tag{13}$$

where $(\xi^*,\eta^*) = \mathcal{D}^{\epsilon\,*}_{w_0}(\zeta)$, together with the *fundamental estimate* (Lemma 4.2.5)

$$\begin{aligned} &\epsilon^{-1}\|\eta - g(u)\nabla_t\xi\|_p + \|\nabla_s\xi\|_p + \|\nabla_t\eta\|_p + \epsilon\|\nabla_s\eta\|_p \\ &\qquad\le c_4\left(\|\mathcal{D}^{\epsilon\,(*)}_{w_0}(\xi,\eta)\|_{0,p,\epsilon} + \epsilon^{-\kappa_p}\|(\xi,\eta\|_{0,p,\epsilon}\right). \end{aligned} \tag{14}$$

which holds for $\mathcal{D}^{\epsilon}_{w_0}$ as well as for ${\mathcal{D}^{\epsilon}_{w_0}}^*$ uniformly for $\epsilon \in (0,\epsilon_0)$. Here $\kappa : [2,\infty) \to \mathbb{R}$ is a continuous function with $\kappa(2) = 1/2$. Note that due to the nonlinearities we need to choose some $p > 2$. As it turns out in chapter 4, (13) is a consequence of (14) for all $p \ge 2$ with $\kappa_p \in (0,1)$. Let us fix throughout such a $p > 2$ and note that its existence is based on conjecture 1.0.6.

This chapter is devoted to the proof of the following approximation result: Any parabolic cylinder can be approximated by a unique family of elliptic ones.

In section 2.1 we prove the existence part, theorem 1.0.4, by constructing a version of Newton's iteration method for the map $\mathcal{F}^{triv}_{\epsilon,u_0}$. It turns out that the primary step of the induction process will determine the quality of the final estimate. Extensive use of the fundamental estimate at this stage will prove extremely valuable in order to get optimal results.

In section 2.2 we prove uniqueness by combining the estimates obtained from the iteration, the key estimates Theorem 4.4.4 and the quadratic estimates I and II from chapter 5.

2.1. Existence

PROOF. (OF THEOREM 1.0.4 – EXISTENCE) Note that the nondegeneracy of the boundary conditions implies exponential decay of the cylinder u_0 and its derivatives. We may therefore assume that there is a constant $c_0 > 0$ such that $\|\nabla_s \partial_t u_0\|_p + \|\nabla_t \nabla_s \partial_t u_0\|_p < c_0$ for all $u_0 \in \mathcal{M}^0(x^-, x^+)$. By choosing c_0 sufficiently big, this clearly continues to hold uniformly for derivatives up to order 4, let's say. It follows that the cylinder w_0 is indeed an approximate zero of $\mathcal{F}_\epsilon$ in the sense that we can arrange its value being as small as we like by choosing $\epsilon_0 > 0$ sufficiently small

$$\|\mathcal{F}_\epsilon(w_0)\|_{0,p,\epsilon} = \|\mathcal{F}^{triv}_{\epsilon,u_0}(0)\|_{0,p,\epsilon} = \left\| \begin{pmatrix} 0 \\ \nabla_s w_0 \end{pmatrix} \right\|_{0,p,\epsilon} \leq c_0 \epsilon. \tag{15}$$

Now we are in position to start the Newton iteration.

Step $\nu = 0$: Let $Z_0 = 0$ be the initial point and define the correction term ξ_0 by

$$(\xi_0, \eta_0) = \zeta_0 = -\mathcal{Q}_\epsilon \circ \mathcal{F}^{triv}_{\epsilon,u_0}(Z_0) = -\mathcal{Q}_\epsilon \begin{pmatrix} 0 \\ \nabla_s w_0 \end{pmatrix}.$$

This implies

$$\mathcal{D}^\epsilon_{w_0} \zeta_0 = -\mathcal{F}^{triv}_{\epsilon,u_0}(Z_0) = -\begin{pmatrix} 0 \\ \nabla_s w_0 \end{pmatrix} \tag{16}$$

and we define the next starting point to be $Z_1 = Z_0 + \zeta_0$. The estimates for $\zeta_0 = (\xi_0, \eta_0)$ and its derivatives, which we are going to prove in this step, are as follows: there exist constants ϵ_0, c_1 such that for all $\epsilon \in (0, \epsilon_0)$

$$\boxed{\begin{aligned}
\|\xi_0\|_p + \|g(u_0)\nabla_t \xi_0 - \eta_0\|_p &\leq c_1\, \epsilon^2 \\
\|g(u_0)\nabla_t \nabla_t \xi_0 - \nabla_t \eta_0\|_p &\leq c_1\, \epsilon^{\min\{5/2-\kappa_p, 2\}} \\
\|\eta_0\|_p + \|\nabla_t \xi_0\|_p &\leq c_1\, \epsilon^{3/2} \\
\|g(u_0)\nabla_s \nabla_t \xi_0 - \nabla_s \eta_0\|_p &\leq c_1\, \epsilon^{2-\kappa_p} \\
\|\nabla_t \eta_0\|_p + \|\nabla_s \xi_0\|_p + \|\nabla_t \nabla_t \xi_0\|_p &\leq c_1\, \epsilon \\
\|\nabla_s \eta_0\|_p + \|\nabla_t \nabla_s \xi_0\|_p + \|\nabla_t \nabla_t \eta_0\|_p &\leq c_1\, \epsilon^{\min\{3/2-\kappa_p, 1\}} \\
\|\nabla_s \nabla_s \xi_0\|_p + \|\nabla_t \nabla_s \eta_0\|_p &\leq c_1\, \epsilon^{1-\kappa_p} \\
\|\nabla_s \nabla_s \eta_0\|_p &\leq c_1\, \epsilon^{\min\{3/2-2\kappa_p, 1-\kappa_p\}} \\
\|\xi_0\|_\infty + \|g(u_0)\nabla_t \xi_0 - \eta_0\|_\infty &\leq c_1\, \epsilon^{2-\frac{3}{2p}} \\
\|\eta_0\|_\infty + \|\nabla_t \xi_0\|_\infty &\leq c_1\, \epsilon^{\frac{3}{2}-\frac{3}{2p}} \\
\|\nabla_t \eta_0\|_\infty + \|\nabla_s \xi_0\|_\infty &\leq c_1\, \epsilon^{1-\frac{3}{2p}} \\
\|\nabla_s \eta_0\|_\infty &\leq c_1\, \epsilon^{\min\{3/2-\kappa_p, 1\}-\frac{3}{2p}}.
\end{aligned}} \tag{17}$$

Before entering their proof we show how they lead to an estimate for the two components of $\mathcal{F}^{triv}_{\epsilon,u_0}(Z_1)$. Using (16) and the fundamental quadratic estimate theorem 5.1.1 we get

$$\begin{aligned}\underline{\|(\mathcal{F}^{triv}_{\epsilon,u_0}(Z_1))_1\|_p} &= \|(\mathcal{F}^{triv}_{\epsilon,u_0}(\zeta_0) - \mathcal{F}^{triv}_{\epsilon,u_0}(0) - \mathcal{D}^{\epsilon}_{w_0}\zeta_0)_1\|_p \\ &\le c_2 c_1^2 \epsilon^{2-\frac{3}{2p}} \left(\epsilon^2 + \epsilon^{3/2} + \epsilon + \epsilon^{3/2} + \epsilon\right) \\ &\le \underline{3c_1^2 c_2 \epsilon^{3-\frac{3}{2p}}}\end{aligned}$$

and

$$\begin{aligned}\underline{\|(\mathcal{F}^{triv}_{\epsilon,u_0}(Z_1))_2\|_p} &= \|(\mathcal{F}^{triv}_{\epsilon,u_0}(\zeta_0) - \mathcal{F}^{triv}_{\epsilon,u_0}(0) - \mathcal{D}^{\epsilon}_{w_0}\zeta_0)_2\|_p \\ &\le c_2 \epsilon^{-2} c_1^2 \epsilon^{2-\frac{3}{2p}} \left(\epsilon^2 + \epsilon^2\right) \\ &\quad + c_2 c_1^2 \epsilon^{2-\frac{3}{2p}} \left(\epsilon + \epsilon^{3/2} + \epsilon\right) + c_1^2 c_2 \epsilon^{\frac{5}{2}-\frac{3}{2p}} \\ &\le \underline{3c_1^2 c_2 \epsilon^{2-\frac{3}{2p}}}\end{aligned}$$

for $1/c_1 + \epsilon_0 > 0$ sufficiently small. It turns out that we even need partial derivatives of the section evaluated at Z_1. We apply the corresponding fundamental quadratic estimates in theorem 5.1.1 and simply state the final results. Observe that the partial derivatives of the first component of the section are $\epsilon^{1/2}$ better as expected

$$\begin{aligned}\|\nabla_t(\mathcal{F}^{triv}_{\epsilon,u_0}(Z_1))_1\|_p &\le 3c_1^2 c_2 \epsilon^{3-\frac{3}{2p}} \\ \|\nabla_t(\mathcal{F}^{triv}_{\epsilon,u_0}(Z_1))_2\|_p &\le 3c_1^2 c_2 \epsilon^{\frac{3}{2}-\frac{3}{2p}} \\ \|\nabla_s(\mathcal{F}^{triv}_{\epsilon,u_0}(Z_1))_1\|_p &\le 3c_1^2 c_2 \epsilon^{\frac{5}{2}-\frac{3}{2p}} \\ \|\nabla_s(\mathcal{F}^{triv}_{\epsilon,u_0}(Z_1))_2\|_p &\le 3c_1^2 c_2 \epsilon^{1-\frac{3}{2p}}.\end{aligned}$$

Let us now derive the estimates in (17). The key estimates (13) give three of them

$$\begin{aligned}\|\xi_0\|_p &\le \|\zeta_0\|_{1,p,\epsilon} \\ &\le c_3 \left(\epsilon \left\| \begin{pmatrix} 0 \\ \nabla_s w_0 \end{pmatrix} \right\|_{0,p,\epsilon} + \|(\mathbb{1} - \epsilon^{\alpha_p}\nabla_t\nabla_t)^{-1}(0 + \epsilon^2 \nabla_t \nabla_s w_0)\|_p \right) \\ &\le 2c_0 c_3 \epsilon^2\end{aligned}$$

where we applied Lemma 4.2.4; moreover

$$\begin{aligned}&\|\eta_0\|_p + \|\nabla_t \xi_0\|_p \\ &\le c_3 \left(\epsilon^{1/2} \left\| \begin{pmatrix} 0 \\ \nabla_s w_0 \end{pmatrix} \right\|_{0,p,\epsilon} + \epsilon^{-1/2}\|(\mathbb{1} - \epsilon^{\alpha_p}\nabla_t\nabla_t)^{-1}(0 + \epsilon^2 \nabla_t \nabla_s w_0)\|_p \right) \\ &\le 2c_0 c_3 \epsilon^{3/2}.\end{aligned}$$

Using these in the fundamental estimate (14) for $\mathcal{D}^{\epsilon}_{w_0}$ gives another two estimates in (17)

$$\begin{aligned}\epsilon^{-1}\|g(u_0)\nabla_t\xi_0-\eta_0\|_p+\|\nabla_t\eta_0\|_p+\|\nabla_s\xi_0\|_p+\epsilon\|\nabla_s\eta_0\|_p&\\ \leq c_4\left(\left\|\begin{pmatrix}0\\ \nabla_s w_0\end{pmatrix}\right\|_{0,p,\epsilon}+\epsilon^{-\kappa_p}\|\xi_0\|_p+\epsilon^{1-\kappa_p}\|\eta_0\|_p\right)&\\ \leq c_4\left(\epsilon c_0+2c_0c_3\epsilon^{2-\kappa_p}+2c_0c_3\epsilon^{5/2-\kappa_p}\right)&\\ \leq 2c_0c_4\epsilon&\end{aligned}$$

for $\epsilon_0>0$ sufficiently small. It remains to improve the estimate for $\nabla_s\eta_0$ in the L^p-norm as well as to obtain the L^∞-estimates. Note that the standard local Sobolev estimate (of the L^∞- by the $W^{1,p}$-norm for $p>2$) pulls through to the case of cylinders, just as in the proof of Lemma 4.2.6, and so we get the existence of a constant $c_5>0$ such that for all $\xi\in C_0^\infty(\mathbb{R}\times S^1,u_0^*TM)$

$$\|\xi\|_\infty\leq c_5\left(\|\xi\|_p+\|\nabla_t\xi\|_p+\|\nabla_s\xi\|_p\right). \tag{18}$$

On the other hand Lemma 4.2.6 with $\beta_1=1/2$, $\beta_2=1$ tells us that there exists a constant $c_5>0$ such that

$$\|\xi\|_\infty\leq c_5\epsilon^{-\frac{3}{2p}}\left(\|\xi\|_p+\epsilon^{1/2}\|\nabla_t\xi\|_p+\epsilon\|\nabla_s\xi\|_p\right). \tag{19}$$

for all ξ. The unbalanced version only gives $\|\xi_0\|_\infty\leq c\epsilon$, but the balanced one leads to

$$\|\xi_0\|_\infty\leq c_5\epsilon^{-\frac{3}{2p}}\left(\|\xi_0\|_p+\epsilon^{1/2}\|\nabla_t\xi_0\|_p+\epsilon\|\nabla_s\xi_0\|_p\right)\leq 3c_0c_4c_5\epsilon^{2-\frac{3}{2p}}$$

for $\epsilon_0>0$ sufficiently small.

To prove the remaining estimates in (17) it is convenient to work in an orthonormal frame which is parallel with respect to s. We indicate this situation by putting an arrow on top of all objects which would generate a collision with the global notation otherwise. For instance $\vec{\nabla}_t$ denotes ∂_t+A, where $A(s,t)\in so(n,\mathbb{R})$, i.e. $A^T=-A$ pointwise. As we will apply several times the fundamental estimate (14), it is usefull to introduce some notation: Let $\mu=(i,j)\in\mathbb{N}_0\times\mathbb{N}_0$ and define

$$\partial_\mu=\underbrace{\vec{\nabla}_t\cdots\vec{\nabla}_t}_{i\text{ times}}\underbrace{\partial_s\cdots\partial_s}_{j\text{ times}}=(\vec{\nabla}_t)^i(\partial_s)^j.$$

Setting $(\vec{\xi}',\vec{\eta}')=\mathcal{D}_\epsilon(\vec{\xi},\vec{\eta})$ we find (cf. appendix A section A.4)

$$\begin{aligned}\mathcal{D}_\epsilon\begin{pmatrix}\partial_\mu\vec{\xi}\\ \partial_\mu\vec{\eta}\end{pmatrix}&=\begin{pmatrix}\partial_s\partial_\mu\vec{\xi}-\vec{\nabla}_t\partial_\mu\vec{\eta}-S\partial_\mu\vec{\xi}\\ \partial_s\partial_\mu\vec{\eta}+\epsilon^{-2}\left(\vec{\nabla}_t\partial_\mu\vec{\xi}-\partial_\mu\vec{\eta}\right)+B\partial_\mu\vec{\xi}\end{pmatrix}\\ &=\begin{pmatrix}\partial_\mu\vec{\xi}'+[\partial_s,\partial_\mu]\vec{\xi}-[\vec{\nabla}_t,\partial_\mu]\vec{\eta}-[S,\partial_\mu]\vec{\xi}\\ \partial_\mu\vec{\eta}'+[\partial_s,\partial_\mu]\vec{\eta}+\epsilon^{-2}[\vec{\nabla}_t,\partial_\mu]\vec{\xi}+[B,\partial_\mu]\vec{\xi}\end{pmatrix}\end{aligned}$$

so that the fundamental estimate yields

$$
\boxed{
\begin{aligned}
&\epsilon^{-1}\|\vec{\nabla}_t\partial_\mu\vec{\xi}-\partial_\mu\vec{\eta}\|_p+\|\vec{\nabla}_t\partial_\mu\vec{\eta}\|_p+\|\partial_s\partial_\mu\vec{\xi}\|_p+\epsilon\|\partial_s\partial_\mu\vec{\eta}\|_p\\
&\qquad\le c_4\Big(\|\partial_\mu\vec{\xi}\,'+[\partial_s,\partial_\mu]\vec{\xi}-[\vec{\nabla}_t,\partial_\mu]\vec{\eta}-[S,\partial_\mu]\vec{\xi}\|_p\\
&\qquad\quad+\epsilon\|\partial_\mu\vec{\eta}\,'+[\partial_s,\partial_\mu]\vec{\eta}+\epsilon^{-2}[\vec{\nabla}_t,\partial_\mu]\vec{\xi}+[B,\partial_\mu]\vec{\xi}\|_p\\
&\qquad\quad+\epsilon^{-\kappa_p}\|\partial_\mu\vec{\xi}\|_p+\epsilon^{1-\kappa_p}\|\partial_\mu\vec{\eta}\|_p\Big).
\end{aligned}
}
\tag{20}
$$

In what follows we apply this estimate to

$$
\begin{pmatrix}\vec{\xi}\,'\\ \vec{\eta}\,'\end{pmatrix}=\mathcal{D}_\epsilon\begin{pmatrix}\vec{\xi}_0\\ \vec{\eta}_0\end{pmatrix}=\begin{pmatrix}0\\ -\partial_s\vec{w}_0\end{pmatrix}
$$

where the last equality is equation (16).

<u>The case $\mu=(1,0)$:</u> $\partial_\mu=\vec{\nabla}_t$ and hence $[\vec{\nabla}_t,\partial_\mu]=0$ and $[\vec{\partial}_s,\partial_\mu]=\partial_s A$, as was shown in appendix A equation (142). Moreover,

$$
-[S,\vec{\nabla}_t]=-S\vec{\nabla}_t+(\vec{\nabla}_t S)+S\vec{\nabla}_t=(\vec{\nabla}_t S)=(\partial_t S)+AS
$$

and similarly for $[B,\vec{\nabla}_t]$. Equation (20) together with the estimates obtained so far implies

$$
\begin{aligned}
\epsilon^{-1}\|\vec{\nabla}_t\vec{\nabla}_t\vec{\xi}_0-\vec{\nabla}_t\vec{\eta}_0\|_p&+\|\vec{\nabla}_t\vec{\nabla}_t\vec{\eta}_0\|_p+\|\partial_s\vec{\nabla}_t\vec{\xi}_0\|_p+\epsilon\|\partial_s\vec{\nabla}_t\vec{\eta}_0\|_p\\
&\le c_4\Big(0+c_{\partial_s A}\|\vec{\xi}_0\|_p+c_{\vec{\nabla}_t S}\|\vec{\xi}_0\|_p+\epsilon c_0+\epsilon c_{\partial_s A}\|\vec{\eta}_0\|_p\\
&\quad+\epsilon c_{\vec{\nabla}_t B}\|\vec{\xi}_0\|_p+\epsilon^{-\kappa_p}\|\vec{\nabla}_t\vec{\xi}_0\|_p+\epsilon^{1-\kappa_p}\|\vec{\nabla}_t\vec{\eta}_0\|_p\Big)\\
&\le c_4^3\epsilon^{\min\{3/2-\kappa_p,1\}}
\end{aligned}
$$

for $\epsilon_0>0$ sufficiently small. This yields appropriate estimates for $\|\vec{\nabla}_t\vec{\nabla}_t\vec{\xi}_0-\vec{\nabla}_t\vec{\eta}_0\|_p$, $\|\vec{\nabla}_t\partial_s\vec{\xi}_0\|_p$ and – using the L^p-estimate for $\vec{\nabla}_t\vec{\eta}_0$ obtained above – we get $\|\vec{\nabla}_t\vec{\nabla}_t\vec{\xi}_0\|_p\le 3c_0c_4\epsilon$.

<u>The case $\mu=(0,1)$:</u> $\partial_\mu=\partial_s$ and so $[\partial_s,\partial_\mu]=0$, $[\vec{\nabla}_t,\partial_\mu]=-\partial_s A$ and

$$
-[S,\partial_s]=(\partial_s S).
$$

Equation (20) leads to three more estimates in (17)

$$
\begin{aligned}
\epsilon^{-1}\|\vec{\nabla}_t\partial_s\vec{\xi}_0-\partial_s\vec{\eta}_0\|_p&+\|\vec{\nabla}_t\partial_s\vec{\eta}_0\|_p+\|\partial_s\partial_s\vec{\xi}_0\|_p+\epsilon\|\partial_s\partial_s\vec{\eta}_0\|_p\\
&\le c_4\Big(0+c_{\partial_s A}\|\vec{\eta}_0\|_p+c_{\partial_s S}\|\vec{\xi}_0\|_p+\epsilon c_0+\epsilon^{-1}c_{\partial_s A}\|\vec{\xi}_0\|_p\\
&\quad+\epsilon c_{\partial_s B}\|\vec{\xi}_0\|_p+\epsilon^{-\kappa_p}\|\partial_s\vec{\xi}_0\|_p+\epsilon^{1-\kappa_p}\|\partial_s\vec{\eta}_0\|_p\Big)\\
&\le 3c_0c_4\epsilon^{1-\kappa_p}
\end{aligned}
$$

for $\epsilon_0 > 0$ sufficiently small. Here we used the estimate $\|\partial_s\vec{\eta}_0\|_p \leq 2c_0c_4$ obtained before. Observe that we get from this and formerly obtained estimates

$$\begin{aligned}\|\partial_s\vec{\eta}_0\|_p &\leq 3c_0c_4\epsilon^{2-\kappa_p} + \|\partial_s\vec{\nabla}_t\vec{\xi}_0\|_p + \|[\vec{\nabla}_t,\partial_s]\xi_0\|_p \\ &\leq 2c_4^3\epsilon^{\min\{3/2-\kappa_p,1\}}\end{aligned}$$

for $\epsilon_0 > 0$ sufficiently small.

We are now in position to derive some of the L^∞-estimates. The balanced versions are as follows

$$\begin{aligned}&\|\vec{\nabla}_t\vec{\xi}_0 - \vec{\eta}_0\|_\infty \\ &\leq c_5\epsilon^{-\frac{3}{2p}}\left(\|\vec{\nabla}_t\vec{\xi}_0 - \vec{\eta}_0\|_p + \epsilon^{1/2}\|\vec{\nabla}_t\vec{\nabla}_t\vec{\xi}_0 - \vec{\nabla}_t\vec{\eta}_0\|_p + \epsilon\|\partial_s\vec{\nabla}_t\vec{\xi}_0 - \partial_s\vec{\eta}_0\|_p\right) \\ &\leq c_5\epsilon^{-\frac{3}{2p}}\left(2c_0c_4\epsilon^2 + c_4^3\epsilon^{\min\{3-\kappa_p,5/2\}} + 3c_4^4\epsilon^{3-\kappa_p} + c_{\partial_sA}2c_0c_3\epsilon^3\right) \\ &\leq 4c_4^4c_5\epsilon^{2-\frac{3}{2p}}\end{aligned}$$

and

$$\|\vec{\eta}_0\|_\infty \leq c_5\epsilon^{-\frac{3}{2p}}\left(\|\vec{\eta}_0\|_p + \epsilon^{1/2}\|\vec{\nabla}_t\vec{\eta}_0\|_p + \epsilon\|\partial_s\vec{\eta}_0\|_p\right) \leq 3c_4^4c_5\epsilon^{\frac{3}{2}-\frac{3}{2p}}$$

The same estimate holds for $\vec{\nabla}_t\vec{\xi}_0$ in view of the result for $\|\vec{\nabla}_t\vec{\xi}_0 - \vec{\eta}_0\|_\infty$. Moreover,

$$\|\vec{\nabla}_t\vec{\eta}_0\|_\infty + \|\partial_s\vec{\xi}_0\|_\infty \leq c\epsilon^{1-\frac{3}{2p}}.$$

We could do better for p close to 2 using the unbalanced estimate (18):

$$\begin{aligned}\|\vec{\nabla}_t\vec{\xi}_0 - \vec{\eta}_0\|_\infty &\leq 4c_4^4c_5\epsilon^{2-\kappa_p} \\ \|\vec{\eta}_0\|_\infty + \|\vec{\nabla}_t\vec{\xi}_0\|_\infty &\leq 3c_4^4c_5\epsilon^{\min\{3/2-\kappa_p,1\}} \\ \|\vec{\nabla}_t\vec{\eta}_0\|_\infty + \|\partial_s\vec{\xi}_0\|_\infty &\leq c\epsilon^{1-\kappa_p}.\end{aligned}$$

Observe that so far we only get

$$\|\partial_s\vec{\eta}_0\|_\infty \leq c\epsilon^{1-\kappa_p-3/2p},$$

which is *not sufficient*. To get a better estimate consider

<u>The case $\mu = (1,1)$:</u> $\partial_\mu = \vec{\nabla}_t\partial_s$ and so

$$[\vec{\nabla}_t,\partial_\mu] = -(\vec{\nabla}_t\partial_sA) - (\partial_sA)\vec{\nabla}_t \quad , \quad [\partial_s,\partial_\mu] = (\partial_sA)\partial_s$$
$$[S,\partial_\mu] = -(\vec{\nabla}_t\partial_sS) - (\partial_sS)\vec{\nabla}_t - (\vec{\nabla}_tS)\partial_s$$

and the quality of the next estimate is due to the one of $\epsilon^{-\kappa_p}\|\nabla_t\nabla_s\xi_0\|_p$

$$\begin{aligned}&\epsilon^{-1}\|\vec{\nabla}_t\vec{\nabla}_t\partial_s\vec{\xi}_0 - \vec{\nabla}_t\partial_s\vec{\eta}_0\|_p + \|\vec{\nabla}_t\vec{\nabla}_t\partial_s\vec{\eta}_0\|_p \\ &\qquad + \|\partial_s\vec{\nabla}_t\partial_s\vec{\xi}_0\|_p + \epsilon\|\partial_s\vec{\nabla}_t\partial_s\vec{\eta}_0\|_p \leq c\epsilon^{\min\{3/2-2\kappa_p,1-\kappa_p\}}.\end{aligned}$$

The case $\mu = (0,2)$: $\partial_\mu = \partial_s\partial_s$ and so $[\partial_s, \partial_\mu] = 0$,

$$[\vec{\nabla}_t, \partial_\mu] = -2(\partial_s A)\partial_s - (\partial_s\partial_s A) \quad , \quad [S, \partial_\mu] = -2(\partial_s S)\partial_s - (\partial_s\partial_s S),$$

therefore

$$\begin{aligned}\epsilon^{-1}\|\vec{\nabla}_t\partial_s\partial_s\vec{\xi}_0 - \partial_s\partial_s\vec{\eta}_0\|_p + \|\vec{\nabla}_t\partial_s\partial_s\vec{\eta}_0\|_p \\ + \|\partial_s\partial_s\partial_s\vec{\xi}_0\|_p + \epsilon\|\partial_s\partial_s\partial_s\vec{\eta}_0\|_p \leq c\epsilon^{\min\{1-2\kappa_p,0\}}.\end{aligned}$$

The latter implies, using a result from case $\mu = (1,1)$,

$$\begin{aligned}\|\partial_s\partial_s\vec{\eta}_0\|_p &\leq 4c_4^5\epsilon^{\min\{2-2\kappa_p,1\}} + \|\partial_s\vec{\nabla}_t\partial_s\vec{\xi}_0\|_p + \|[\vec{\nabla}_t,\partial_s]\partial_s\vec{\xi}_0\|_p \\ &\leq 4c_4^5\epsilon^{\min\{2-2\kappa_p,1\}} + c_4^4\epsilon^{\min\{3/2-2\kappa_p,1-\kappa_p\}} + 2c_{\partial_s A}c_0c_3\epsilon \\ &\leq 2c_4^4\epsilon^{\min\{3/2-2\kappa_p,1-\kappa_p\}}\end{aligned}$$

which finally gives, using the balanced estimate,

$$\begin{aligned}\|\partial_s\vec{\eta}_0\|_\infty &\leq c\epsilon^{-\frac{3}{2p}}\left(\epsilon^{\min\{3/2-\kappa_p,1\}} + \epsilon^{3/2-\kappa_p} + \epsilon^{\min\{5/2-2\kappa_p,2-\kappa_p\}}\right) \\ &\leq c\epsilon^{\min\{3/2-\kappa_p,1\}-\frac{3}{2p}}.\end{aligned}$$

The unbalanced version leads to

$$\|\partial_s\vec{\eta}_0\|_\infty \leq c\epsilon^{\min\{3/2-2\kappa_p,1-\kappa_p\}}.$$

Choose $c_1 > 0$ sufficiently large to get the desired constant in the estimates.

Induction step $\nu - 1 \Rightarrow \nu$: Let $\nu \in \mathbb{N}$ and suppose we had already constructed $\zeta_0, \ldots, \zeta_{\nu-1}$ and $Z_1, \ldots, Z_\nu$, then define

$$(\xi_\nu, \eta_\nu) = \zeta_\nu = -\mathcal{Q}^\epsilon_{w_0}\circ\mathcal{F}^{triv}_{\epsilon,u_0}(Z_\nu)$$

and

$$Z_{\nu+1} = Z_\nu + \zeta_\nu.$$

This implies

$$\mathcal{D}^\epsilon_{w_0}\zeta_\nu = -\mathcal{F}^{triv}_{\epsilon,u_0}(Z_\nu). \tag{21}$$

The claim is to prove the induction hypothesis (H_ν) under the assumption that $(H_{\nu-1})$ holds.

$$
(H_\nu)\qquad
\begin{aligned}
\|\xi_\nu\|_p + \|g(u_0)\nabla_t\xi_\nu - \eta_\nu\|_p &\le \frac{c_1}{2^\nu}\,\epsilon^2\\
\|\eta_\nu\|_p + \|\nabla_t\xi_\nu\|_p &\le \frac{c_1}{2^\nu}\,\epsilon^{3/2}\\
\|\nabla_t\eta_\nu\|_p + \|\nabla_s\xi_\nu\|_p + \|\nabla_t\nabla_t\xi_\nu\|_p &\le \frac{c_1}{2^\nu}\,\epsilon\\
\|\nabla_s\eta_\nu\|_p + \|\nabla_t\nabla_s\xi_\nu\|_p + \|\nabla_t\nabla_t\eta_\nu\|_p &\le \frac{c_1}{2^\nu}\,\epsilon^{\min\{\frac{3}{2}-\kappa_p,1\}}\\
\|\nabla_s\nabla_s\xi_\nu\|_p + \|\nabla_t\nabla_s\eta_\nu\|_p &\le \frac{c_1}{2^\nu}\,\epsilon^{1-\kappa_p}\\
\|\nabla_s\nabla_s\eta_\nu\|_p &\le \frac{c_1}{2^\nu}\,\epsilon^{\min\{\frac{5}{4}-2\kappa_p,\frac{1}{4}\}}\\
\|\xi_\nu\|_\infty + \|g(u_0)\nabla_t\xi_\nu - \eta_\nu\|_\infty &\le \frac{c_1}{2^\nu}\,\epsilon^{2-\frac{3}{2p}}\\
\|\eta_\nu\|_\infty + \|\nabla_t\xi_\nu\|_\infty &\le \frac{c_1}{2^\nu}\,\epsilon^{\frac{3}{2}-\frac{3}{2p}}\\
\|\nabla_t\eta_\nu\|_\infty + \|\nabla_s\xi_\nu\|_\infty &\le \frac{c_1}{2^\nu}\,\epsilon^{1-\frac{3}{2p}}\\
\|\nabla_s\eta_\nu\|_\infty &\le \frac{c_1}{2^\nu}\,\epsilon^{\min\{\frac{3}{2}-\kappa_p,\frac{9}{4}-2\kappa_p,1\}-\frac{3}{2p}}\\
\|\mathcal{F}^{triv}_{\epsilon,u_0}(Z_{\nu+1})\|_{0,p,\epsilon} &\le \frac{3c_1^2c_2}{2^\nu}\,\epsilon^{3-\frac{3}{2p}}\\
\|\nabla_t\mathcal{F}^{triv}_{\epsilon,u_0}(Z_{\nu+1})\|_{0,p,\epsilon} &\le \frac{3c_1^2c_2}{2^\nu}\,\epsilon^{\frac{5}{2}-\frac{3}{2p}}\\
\|\nabla_s\mathcal{F}^{triv}_{\epsilon,u_0}(Z_{\nu+1})\|_{0,p,\epsilon} &\le \frac{3c_1^2c_2}{2^\nu}\,\epsilon^{2-\frac{3}{2p}}.
\end{aligned}
$$

Here c_1 and c_2 are the constants introduced in step $\nu = 0$. (H_ν) holds uniformly for all $\epsilon \in (0, \epsilon_0)$, where $\epsilon_0 > 0$ has been chosen sufficiently small. As is to be expected we will get significantly better estimates here than in step $\nu = 0$ and as actually was our claim in (H_ν). However, there will appear additional constants and we use the extra powers of ϵ to neutralize them and therefore get (H_ν). It will also be important to use the better results in estimating $\mathcal{F}^{triv}_{\epsilon,u_0}$ via the quadratic estimates derived in chapter 5.

Remark 2.1.1. Observe that the estimates for two terms, namely $\|\nabla_s\eta_\nu\|_\infty$ and $\|\nabla_s\nabla_s\eta_\nu\|_p$ are worse than the ones obtained in step $\nu = 0$. This is not to be expected, but as the former estimate is not used anywhere in this text and the latter one is sufficient for the Newton method we don't make any effort to improve them. However, just in case improvement is needed at a later stage here are two proposals how it should work.

One way is two consider higher derivatives as in step $\nu = 0$, namely the cases $\mu = (1,1)$ and $\mu = (0,2)$, but this leads to the problem of calculating quadratic estimates for the *second* partial derivatives $\nabla_t\nabla_s\mathcal{F}^{triv}_{\epsilon,u_0}$ and $\nabla_s\nabla_s\mathcal{F}^{triv}_{\epsilon,u_0}$.

A more realistic way is to establish a modified version of conjecture 1.0.6 where $\epsilon^{-\kappa_p}(\|\xi\|_p + \epsilon\|\eta\|_p)$ is replaced by $\|\xi\|_p + \|\eta\|_p$. This holds is true for $p = 2$, cf. lemma 4.1.6. Then iterating case $\mu = (0,1)$ should result in the improved estimates.

Let us now assume $(H_{\nu-1})$ holds. In what follows we will need again and again the estimates

$$\begin{aligned}\|\mathcal{F}^{triv}_{\epsilon,u_0}(Z_\nu)\|_{0,p,\epsilon} &\le \|(\mathcal{F}^{triv}_{\epsilon,u_0}(Z_\nu))_1\|_p + \epsilon\|(\mathcal{F}^{triv}_{\epsilon,u_0}(Z_\nu))_2\|_p \\ &\le \frac{3c_1^2c_2}{2^{\nu-1}}\,\epsilon^{3-\frac{3}{2p}} + \frac{3c_1^2c_2}{2^{\nu-1}}\,\epsilon^{3-\frac{3}{2p}} \le \frac{6c_1^2c_2}{2^{\nu-1}}\,\epsilon^{3-\frac{3}{2p}}\end{aligned}$$

and

$$\begin{aligned}\|\pi_\epsilon\mathcal{F}^{triv}_{\epsilon,u_0}(Z_\nu)\|_p &\le \|(1\!\!1 - \epsilon\nabla_t\nabla_t)^{-1}\left(\mathcal{F}^{triv}_{\epsilon,u_0}(Z_\nu)_1 - \epsilon^2 g^{-1}\nabla_t\mathcal{F}^{triv}_{\epsilon,u_0}(Z_\nu)_2\right)\|_p \\ &\le \|\mathcal{F}^{triv}_{\epsilon,u_0}(Z_\nu)_1\|_p + 2p\epsilon^{3/2}\|\mathcal{F}^{triv}_{\epsilon,u_0}(Z_\nu)_2\|_p \\ &\le \frac{3c_1^2c_2}{2^{\nu-1}}\,\epsilon^{3-\frac{3}{2p}} + \frac{6pc_1^2c_2}{2^{\nu-1}}\,\epsilon^{\frac{7}{2}-\frac{3}{2p}} \le \frac{4c_1^2c_2}{2^{\nu-1}}\,\epsilon^{3-\frac{3}{2p}}\end{aligned}$$

for $\epsilon_0 > 0$ sufficiently small. We used the induction hypothesis $(H_{\nu-1})$ and Lemma 4.2.4 to estimate the inverse operator. The key estimates (13) then give

$$\begin{aligned}\|\xi_\nu\|_p \le \|\zeta_\nu\|_{1,p,\epsilon} &\le c_3\Big(\epsilon\|\mathcal{F}^{triv}_{\epsilon,u_0}(Z_\nu)\|_{0,p,\epsilon} + \|\pi_\epsilon\mathcal{F}^{triv}_{\epsilon,u_0}(Z_\nu)\|_p\Big) \\ &\le c_3\Big(\frac{6c_1^2c_2}{2^{\nu-1}}\,\epsilon^{4-\frac{3}{2p}} + \frac{4c_1^2c_2}{2^{\nu-1}}\,\epsilon^{3-\frac{3}{2p}}\Big) \\ &\le \frac{5c_1^2c_2c_3}{2^{\nu-1}}\,\epsilon^{3-\frac{3}{2p}}\end{aligned}$$

and

$$\begin{aligned}\|\eta_\nu\|_p + \|\nabla_t\xi_\nu\|_p &\le c_3\Big(\epsilon^{1/2}\|\mathcal{F}^{triv}_{\epsilon,u_0}(Z_\nu)\|_{0,p,\epsilon} + \epsilon^{-1/2}\|\pi_\epsilon\mathcal{F}^{triv}_{\epsilon,u_0}(Z_\nu)\|_p\Big) \\ &\le \frac{5c_1^2c_2c_3}{2^{\nu-1}}\,\epsilon^{\frac{5}{2}-\frac{3}{2p}}\end{aligned}$$

for $\epsilon_0 > 0$ sufficiently small. Using these results the fundamental estimate (14) leads to

$$\begin{aligned}&\epsilon^{-1}\|g(u_0)\nabla_t\xi_\nu - \eta_\nu\|_p + \|\nabla_t\eta_\nu\|_p + \|\nabla_s\xi_\nu\|_p + \epsilon\|\nabla_s\eta_\nu\|_p \\ &\qquad\le c_4\Big(\|\mathcal{F}^{triv}_{\epsilon,u_0}(Z_\nu)\|_{0,p,\epsilon} + \epsilon^{-\kappa_p}\|\xi_\nu\|_p + \epsilon^{1-\kappa_p}\|\eta_\nu\|_p\Big) \\ &\qquad\le \frac{6c_1^2c_2c_3c_4}{2^{\nu-1}}\,\epsilon^{3-\kappa_p-\frac{3}{2p}}\end{aligned}$$

for $\epsilon_0 > 0$ sufficiently small. Choosing $\epsilon_0 > 0$ again smaller, if necessary, gives the desired estimates in (H_ν) for all terms involving L^p-norms other than $\nabla_s\eta_\nu$.

Let us now repeat the procedure from step $\nu = 0$, namely application of the fundamental estimate (14) not simply to (ξ_ν, η_ν), but to $(\nabla_t\xi_\nu, \nabla_t\eta_\nu)$ and $(\nabla_s\xi_\nu, \nabla_s\eta_\nu)$. To do so we are working again in an orthonormal frame which is parallel with respect to the variable s and indicate this by our usual vector notation. In what follows we apply estimate (20) to

$$\begin{pmatrix} \vec{\xi}_\nu' \\ \vec{\eta}_\nu' \end{pmatrix} = \mathcal{D}_\epsilon \begin{pmatrix} \vec{\xi}_\nu \\ \vec{\eta}_\nu \end{pmatrix} = -\vec{\mathcal{F}}^{triv}_{\epsilon,u_0}(\vec{Z}_\nu).$$

The case $\mu = (1,0)$: $\partial_\mu = \vec{\nabla}_t$ and, using the identities $[\vec{\nabla}_t, \partial_\mu] = 0$, $[\vec{\partial}_s, \partial_\mu] = \partial_s A$ and $-[S, \vec{\nabla}_t] = (\partial_t S) + AS$ derived in step $\nu = 0$, as well as equation (20) we get

$$\begin{aligned}
&\epsilon^{-1}\|\vec{\nabla}_t\vec{\nabla}_t\vec{\xi}_\nu - \vec{\nabla}_t\vec{\eta}_\nu\|_p + \|\vec{\nabla}_t\vec{\nabla}_t\vec{\eta}_\nu\|_p + \|\partial_s\vec{\nabla}_t\vec{\xi}_\nu\|_p + \epsilon\|\partial_s\vec{\nabla}_t\vec{\eta}_\nu\|_p \\
&\quad\le c_4\Big(\|(\vec{\nabla}_t\vec{\mathcal{F}}^{triv}_{\epsilon,u_0}(\vec{Z}_\nu))_1\|_p + c_{\partial_s A}\|\vec{\xi}_\nu\|_p + c_{\vec{\nabla}_t S}\|\vec{\xi}_\nu\|_p + \epsilon\|(\vec{\nabla}_t\vec{\mathcal{F}}^{triv}_{\epsilon,u_0}(\vec{Z}_\nu))_2\|_p \\
&\qquad + \epsilon c_{\partial_s A}\|\vec{\eta}_\nu\|_p + \epsilon c_{\vec{\nabla}_t B}\|\vec{\xi}_\nu\|_p + \epsilon^{-\kappa_p}\|\vec{\nabla}_t\vec{\xi}_\nu\|_p + \epsilon^{1-\kappa_p}\|\vec{\nabla}_t\vec{\eta}_\nu\|_p\Big) \\
&\quad\le \frac{7c_1^2c_2c_3c_4}{2^{\nu-1}}\,\epsilon^{\frac{5}{2}-\kappa_p-\frac{3}{2p}}
\end{aligned}$$

where we used $(H_{\nu-1})$, the estimates derived above and $\epsilon_0, 1/c_4 > 0$ have been chosen sufficiently small. Use the estimate for $\|\vec{\nabla}_t\vec{\eta}_\nu\|_p$ and choose again $\epsilon_0 > 0$ sufficiently small, then

$$\|\vec{\nabla}_t\vec{\nabla}_t\vec{\xi}_\nu\|_p \le \frac{7c_1^2c_2c_3c_4}{2^{\nu-1}}\,\epsilon^{3-\kappa_p-\frac{3}{2p}}$$

and

$$\|\vec{\nabla}_t\vec{\nabla}_t\vec{\eta}_\nu\|_p + \|\partial_s\vec{\nabla}_t\vec{\xi}_\nu\|_p \le \frac{7c_1^2c_2c_3c_4}{2^{\nu-1}}\,\epsilon^{\frac{5}{2}-\kappa_p-\frac{3}{2p}}$$

which imply the corresponding estimates in (H_ν) for $\epsilon_0 > 0$ sufficiently small. Moreover, it follows from the *balanced* estimate that

$$\begin{aligned}
\|\vec{\nabla}_t\vec{\xi}_\nu\|_\infty &\le c_5\,\epsilon^{-\frac{3}{2p}}\Big(\|\vec{\nabla}_t\vec{\xi}_\nu\|_p + \epsilon^{\frac{1}{2}}\|\vec{\nabla}_t\vec{\nabla}_t\vec{\xi}_\nu\|_p + \epsilon\|\partial_s\vec{\nabla}_t\vec{\xi}_\nu\|_p\Big) \\
&\le \frac{6c_1^2c_2c_3c_5}{2^{\nu-1}}\,\epsilon^{1-\frac{3}{2p}}\cdot\epsilon^{\frac{3}{2}-\frac{3}{2p}} \le \frac{c_1}{2^\nu}\,\epsilon^{\frac{3}{2}-\frac{3}{2p}}.
\end{aligned}$$

The case $\mu = (0,1)$: $\partial_\mu = \partial_s$ and so $[\partial_s, \partial_\mu] = 0$, $[\vec{\nabla}_t, \partial_\mu] = -\partial_s A$ and $-[S, \partial_s] = (\partial_s S)$. Equation (20) gives

$$\begin{aligned}
&\epsilon^{-1}\|\vec{\nabla}_t\partial_s\vec{\xi}_\nu - \partial_s\vec{\eta}_\nu\|_p + \|\vec{\nabla}_t\partial_s\vec{\eta}_\nu\|_p + \|\partial_s\partial_s\vec{\xi}_\nu\|_p + \epsilon\|\partial_s\partial_s\vec{\eta}_\nu\|_p \\
&\quad\le c_4\Big(\|(\partial_s\vec{\mathcal{F}}^{triv}_{\epsilon,u_0}(\vec{Z}_\nu))_1\|_p + c_{\partial_s A}\|\vec{\eta}_\nu\|_p + c_{\partial_s S}\|\vec{\xi}_\nu\|_p + \epsilon\|(\partial_s\vec{\mathcal{F}}^{triv}_{\epsilon,u_0}(\vec{Z}_\nu))_2\|_p \\
&\qquad + \epsilon^{-1}c_{\partial_s A}\|\vec{\xi}_\nu\|_p + \epsilon c_{\partial_s B}\|\vec{\xi}_\nu\|_p + \epsilon^{-\kappa_p}\|\partial_s\vec{\xi}_\nu\|_p + \epsilon^{1-\kappa_p}\|\partial_s\vec{\eta}_\nu\|_p\Big) \\
&\quad\le \frac{7c_1^2c_2c_3c_4^2}{2^{\nu-1}}\,\epsilon^{\min\{3-2\kappa_p,2\}-\frac{3}{2p}}
\end{aligned}$$

for $\epsilon_0 > 0$ sufficiently small. This implies

$$\|\vec{\nabla}_t \partial_s \vec{\eta}_\nu\|_p + \|\partial_s \partial_s \vec{\xi}_\nu\|_p \leq \frac{7c_1^2 c_2 c_3 c_4^2}{2^{\nu-1}}\, \epsilon^{\min\{3-2\kappa_p,2\}-\frac{3}{2p}}$$

and

$$\|\partial_s \partial_s \vec{\eta}_\nu\|_p \leq \frac{7c_1^2 c_2 c_3 c_4^2}{2^{\nu-1}}\, \epsilon^{\min\{2-2\kappa_p,1\}-\frac{3}{2p}}$$

as well as

$$\begin{aligned} \|\partial_s \vec{\eta}_\nu\|_p &\leq \frac{7c_1^2 c_2 c_3 c_4^2}{2^{\nu-1}}\, \epsilon^{\min\{4-2\kappa_p,3\}-\frac{3}{2p}} + c_{[\nabla_t,\partial_s]}\|\xi_\nu\|_p + \|\partial_s \nabla_t \xi_\nu\|_p \\ &\leq \frac{8c_1^2 c_2 c_3 c_4^2}{2^{\nu-1}}\, \epsilon^{\frac{5}{2}-\kappa_p-\frac{3}{2p}} \end{aligned}$$

which for $\epsilon_0 > 0$ sufficiently small implies the last missing L^p-estimates in (H_ν).

Using these results we can estimate the L^∞-norms for $\epsilon_0, 1/c_4 > 0$ sufficiently small. The balanced version of the L^∞-estimate leads to

$$\begin{aligned} \|\xi_\nu\|_\infty + \|g(u_0)\nabla_t \xi_\nu - \eta_\nu\|_\infty &\leq \frac{6c_1^2 c_2 c_3 c_5}{2^{\nu-1}}\, \epsilon^{2-\frac{3}{2p}}\, \epsilon^{1-\frac{3}{2p}} \leq \frac{c_1}{2^\nu}\, \epsilon^{2-\frac{3}{2p}} \\ \|\eta_\nu\|_\infty &\leq \frac{6c_1^2 c_2 c_3 c_5}{2^{\nu-1}}\, \epsilon^{\frac{3}{2}-\frac{3}{2p}}\, \epsilon^{1-\frac{3}{2p}} \leq \frac{c_1}{2^\nu}\, \epsilon^{\frac{3}{2}-\frac{3}{2p}} \\ \|\nabla_t \eta_\nu\|_\infty + \|\nabla_s \xi_\nu\|_\infty &\leq \frac{7c_1^2 c_2 c_3 c_4 c_5}{2^{\nu-1}}\, \epsilon^{1-\frac{3}{2p}}\, \epsilon^{2-\kappa_p-\frac{3}{2p}} \leq \frac{c_1}{2^\nu}\, \epsilon^{1-\frac{3}{2p}} \end{aligned}$$

and

$$\begin{aligned} \|\nabla_s \eta_\nu\|_\infty &\leq \frac{\tilde{c}}{2^{\nu-1}}\, \epsilon^{\frac{3}{4}-\frac{3}{2p}} \epsilon^{-\frac{3}{2p}} \left(\epsilon^{\frac{7}{4}-\kappa_p} + \epsilon^{\min\{\frac{9}{4}-2\kappa_p,\frac{5}{4}\}}\right) \\ &\leq \frac{c_1}{2^\nu}\, \epsilon^{\min\{\frac{9}{4}-2\kappa_p,\frac{5}{4}\}-\frac{3}{2p}} \leq \frac{c_1}{2^\nu}\, \epsilon^{\min\{\frac{3}{2}-\kappa_p,\frac{9}{4}-2\kappa_p,1\}-\frac{3}{2p}}. \end{aligned}$$

Note that the minimum is taken on by the first, second, third number if κ_p is in the interval $[1/2, 3/4]$, $[3/4, 1)$, $(0, 1/2]$, respectively.

Next we derive estimates for X_ν and Y_ν. Recall that $X_\nu = \sum_{l=0}^{\nu-1} \xi_l$ and $Y_\nu = \sum_{l=0}^{\nu-1} \eta_l$; so it follows from the previous induction steps $(H_0), \ldots, (H_{\nu-1})$

$$\|X_\nu\|_p \leq \sum_{l=0}^{\nu-1} \|\xi_l\|_p \leq c_1\, \epsilon^2 \sum_{l=0}^{\nu-1} 2^{-l} \leq 2c_1\, \epsilon^2.$$

Similarly we get

$$
\begin{aligned}
\|g(u_0)\nabla_t X_\nu - Y_\nu\|_p &\le 2c_1\,\epsilon^2 \\
\|Y_\nu\|_p + \|\nabla_t X_\nu\|_p &\le 2c_1\,\epsilon^{3/2} \\
\|\nabla_t Y_\nu\|_p + \|\nabla_s X_\nu\|_p + \|\nabla_t\nabla_t X_\nu\|_p &\le 2c_1\,\epsilon \\
\|\nabla_s Y_\nu\|_p + \|\nabla_t\nabla_s X_\nu\|_p + \|\nabla_t\nabla_t Y_\nu\|_p &\le 2c_1\,\epsilon^{\min\{\frac{3}{2}-\kappa_p,1\}} \\
\|\nabla_s\nabla_s X_\nu\|_p + \|\nabla_t\nabla_s Y_\nu\|_p &\le 2c_1\,\epsilon^{1-\kappa_p} \\
\|\nabla_s\nabla_s Y_\nu\|_p &\le 2c_1\,\epsilon^{\min\{\frac{5}{4}-2\kappa_p,\frac{1}{4}\}} \\
\|X_\nu\|_\infty + \|g(u_0)\nabla_t X_\nu - Y_\nu\|_\infty &\le 2c_1\,\epsilon^{2-\frac{3}{2p}} \\
\|Y_\nu\|_\infty + \|\nabla_t X_\nu\|_\infty &\le 2c_1\,\epsilon^{\frac{3}{2}-\frac{3}{2p}} \\
\|\nabla_t Y_\nu\|_\infty + \|\nabla_s X_\nu\|_\infty &\le 2c_1\,\epsilon^{1-\frac{3}{2p}} \\
\|\nabla_s Y_\nu\|_\infty &\le 2c_1\,\epsilon^{\min\{\frac{3}{2}-\kappa_p,\frac{9}{4}-2\kappa_p,1\}-\frac{3}{2p}}.
\end{aligned}
\tag{22}
$$

So far we have established all estimates in (H_ν) except for those of the section $\mathcal{F}^{triv}_{\epsilon,u_0}$ and its first partial derivatives. To obtain them we apply the quadratic estimates theorem 5.2.1 and theorem 5.3.1, the estimates derived above – not the ones from (H_ν) ! , equation (21) and the old trick of adding zero to get for $\epsilon_0 > 0$ sufficiently small (appearance of both exponents $3/p$ and $3/2p$ is not a typo)

$$
\begin{aligned}
&\|(\mathcal{F}^{triv}_{\epsilon,u_0}(Z_{\nu+1}))_1\|_p \\
&\le \|(\mathcal{F}^{triv}_{\epsilon,u_0}(Z_\nu+\zeta_\nu) - \mathcal{F}^{triv}_{\epsilon,u_0}(Z_\nu) - d\mathcal{F}^{triv}_{\epsilon,u_0}(Z_\nu)\,\zeta_\nu)_1\|_p \\
&\quad + \|(d\mathcal{F}^{triv}_{\epsilon,u_0}(Z_\nu)\,\zeta_\nu - \mathcal{D}^\epsilon_{w_0}\,\zeta_\nu)_1\|_p \\
&\le c_8\Big(\frac{\epsilon^{3-\frac{3}{p}}}{2^{\nu-1}}\Big)^2\Big(\epsilon^{3/2}+\epsilon+\epsilon\Big) \\
&\quad + c_8\Big(\frac{\epsilon^{3-\frac{3}{p}}}{2^{\nu-1}}\Big)\epsilon^{\frac{5}{2}-\frac{3}{2p}}\Big(\epsilon^{\frac{1}{2}} + \underline{1} + \epsilon^{\frac{7}{2}-\kappa_p-\frac{3}{p}} + \underline{1} + \epsilon^{\frac{7}{2}-\kappa_p-\frac{3}{p}}\Big) \\
&\quad + c_8\Big(\frac{\epsilon^{2-\frac{3}{2p}}}{2^{\nu-1}}\Big)\Big(\epsilon^{2-\frac{3}{2p}}(\epsilon^{3-\frac{3}{2p}} + \epsilon^{3-\kappa_p-\frac{3}{2p}} + \epsilon^{\frac{5}{2}-\frac{3}{2p}} + \epsilon^{3-\kappa_p-\frac{3}{2p}}) + \epsilon^{10-\frac{9}{2p}}\Big) \\
&\quad + c_8\Big(\frac{\epsilon^{3-\frac{3}{p}}}{2^{\nu-1}}\Big)\Big(\underline{\epsilon^{3/2}} + \epsilon\cdot\epsilon^{2-\frac{3}{2p}} + \underline{\epsilon^{3/2}} + \epsilon\cdot\epsilon^{2-\frac{3}{2p}}\Big) \\
&\quad + c_8\Big(\frac{\epsilon^{2-\frac{3}{2p}}}{2^{\nu-1}}\Big)\Big(\epsilon^{3-\frac{3}{2p}} + \underline{\epsilon^{\frac{5}{2}-\frac{3}{2p}}} + \epsilon^{5-\kappa_p-\frac{3}{p}} + \underline{\epsilon^{\frac{5}{2}-\frac{3}{2p}}}\Big) \\
&\le c\,\epsilon^{\frac{3}{2}-\frac{3}{2p}}\,\frac{1}{2^{\nu-1}}\,\epsilon^{3-\frac{3}{2p}} \le \frac{1}{2^\nu}\,\epsilon^{3-\frac{3}{2p}}
\end{aligned}
$$

and

$$
\begin{aligned}
&\|(\mathcal{F}^{triv}_{\epsilon,u_0}(Z_{\nu+1}))_2\|_p \\
&\quad \le \|(\mathcal{F}^{triv}_{\epsilon,u_0}(Z_\nu+\zeta_\nu) - \mathcal{F}^{triv}_{\epsilon,u_0}(Z_\nu) - d\mathcal{F}^{triv}_{\epsilon,u_0}(Z_\nu)\,\zeta_\nu)_2\|_p \\
&\qquad + \|(d\mathcal{F}^{triv}_{\epsilon,u_0}(Z_\nu)\,\zeta_\nu - \mathcal{D}^\epsilon_{w_0}\,\zeta_\nu)_2\|_p \\
&\quad \le c_8\Big(\frac{\epsilon^{3-\frac{3}{p}}}{2^{\nu-1}}\Big)^2\Big(\epsilon^{-2}(\epsilon^2+\epsilon^{3/2})+\epsilon+\epsilon^{3/2}+\epsilon^{\min\{3/2-\kappa_p,1\}}\Big) \\
&\qquad + c_8\Big(\frac{\epsilon^{3-\frac{3}{p}}}{2^{\nu-1}}\Big)\epsilon^{-\frac{3}{2p}}\Big((\epsilon^3+\epsilon^{\frac{5}{2}})\epsilon^{1-\frac{3}{p}}+\epsilon^{3-\kappa_p}+\epsilon^{\frac{5}{2}}\Big) \\
&\qquad + c_8\Big(\frac{\epsilon^{2-\frac{3}{2p}}}{2^{\nu-1}}\Big)\Big(\epsilon^{-2}(\epsilon^{3-\frac{3}{2p}}+\underline{\epsilon^{\frac{5}{2}-\frac{3}{2p}}})\epsilon^{2-\frac{3}{2p}}+\epsilon^{3-\kappa_p-\frac{3}{2p}}\epsilon^{\frac{5}{2}-\frac{3}{p}}\Big) \\
&\qquad + c_8\Big(\frac{\epsilon^{3-\frac{3}{p}}}{2^{\nu-1}}\Big)\Big(\epsilon^{\frac{3}{2}-\frac{3}{2p}}+\epsilon+\epsilon^{\frac{3}{2}}+\epsilon^{\frac{5}{2}-\frac{3}{2p}}\Big) \\
&\qquad + c_8\Big(\frac{\epsilon^{2-\frac{3}{2p}}}{2^{\nu-1}}\Big)\Big(\underline{\epsilon^{-2}\epsilon^{3-\frac{3}{2p}}}+\epsilon^{-2}\epsilon^{\frac{5}{2}-\frac{3}{2p}}\epsilon^{2-\frac{3}{2p}}+\epsilon^{3-\kappa_p-\frac{3}{2p}}+\epsilon^{\frac{5}{2}-\frac{3}{2p}}\Big) \\
&\quad \le c\,\epsilon^{1-\frac{3}{2p}}\,\frac{1}{2^{\nu-1}}\,\epsilon^{2-\frac{3}{2p}} \le \frac{1}{2^\nu}\,\epsilon^{2-\frac{3}{2p}}.
\end{aligned}
$$

The underlined terms expose the worst behavior in terms of powers of ϵ. We used the spare positive powers of ϵ to take care of all constants appearing by choosing $\epsilon_0 > 0$ sufficiently small. Similarly the partial derivatives of the section are estimated by using the corresponding quadratic estimates. As a matter of fact we obtain for fixed $p > 2$, fixed $\kappa_p \in (0,1)$ and choosing $\epsilon_0 > 0$ sufficiently small

$$
\begin{aligned}
\|\nabla_t(\mathcal{F}^{triv}_{\epsilon,u_0}(Z_{\nu+1}))_1\|_p &\le c\,\frac{\epsilon^{\frac{3}{2}-\frac{3}{2p}}}{2^{\nu-1}}\,\epsilon^{\frac{5}{2}-\frac{3}{2p}}+\ldots \le \frac{1}{2^\nu}\,\epsilon^{\frac{5}{2}-\frac{3}{2p}} \\
\|\nabla_t(\mathcal{F}^{triv}_{\epsilon,u_0}(Z_{\nu+1}))_2\|_p &\le c\,\frac{\epsilon^{\frac{3}{2}-\frac{3}{p}}}{2^{\nu-1}}\,\epsilon^{\frac{3}{2}-\frac{3}{2p}}+\ldots \le \frac{1}{2^\nu}\,\epsilon^{\frac{3}{2}-\frac{3}{2p}} \\
\|\nabla_s(\mathcal{F}^{triv}_{\epsilon,u_0}(Z_{\nu+1}))_1\|_p &\le c\,\frac{\epsilon^{\frac{5}{2}-\kappa_p-\frac{3}{2p}}}{2^{\nu-1}}\,\epsilon^{2-\frac{3}{2p}}+\ldots \le \frac{1}{2^\nu}\,\epsilon^{2-\frac{3}{2p}} \\
\|\nabla_s(\mathcal{F}^{triv}_{\epsilon,u_0}(Z_{\nu+1}))_2\|_p &\le c\,\frac{\epsilon^{\frac{5}{2}-\kappa_p\frac{3}{p}}}{2^{\nu-1}}\,\epsilon^{1-\frac{3}{2p}}+\ldots \le \frac{1}{2^\nu}\,\epsilon^{1-\frac{3}{2p}}
\end{aligned}
$$

This concludes the induction step.

It remains to show that $(Z_\nu)_{\nu\in\mathbb{N}_0}$ is a Cauchy sequence in the Banach space $W^{1,p}_\epsilon(\mathbb{R}\times S^1, u_0^*TM \oplus u_0^*T^*M)$ and that $\mathcal{F}^{triv}_{\epsilon,u_0}(Z_\epsilon) = 0$. The estimates for $Z_\epsilon = (X_\epsilon, Y_\epsilon)$ in the statement of the Theorem are then an immediate consequence of the estimates (22) for the elements (X_ν, Y_ν) of the Cauchy sequence. To see that $(Z_\nu)_{\nu\in\mathbb{N}_0}$ is a Cauchy sequence observe that its norm is dominated by a standard Cauchy sequence in $\mathbb{R}$, i.e. using the induction hypothesis (H_ν) and assuming without loss of generality that $\nu > \mu$, we get

$$\|Z_\nu - Z_\mu\|_{1,p,\epsilon} \le \sum_{l=\mu}^{\nu-1} \|\zeta_l\|_{1,p,\epsilon} \le c_1 \epsilon^2 \sum_{l=\mu}^{\nu-1} 2^{-l} \to 0 \quad \text{for } \nu, \mu \to \infty.$$

(H_ν) also implies

$$\lim_{\nu\to\infty} \|\mathcal{F}^{triv}_{\epsilon,u_0}(Z_\nu)\|_{0,p,\epsilon} \le \lim_{\nu\to\infty} \frac{6c_1^2 c_2}{2^{\nu-1}}\, \epsilon^{3-\frac{3}{2p}} = 0$$

uniformly in $\epsilon \in (0, \epsilon_0)$. □

2.2. Uniqueness

PROOF. (OF THEOREM 1.0.5) Set $\tilde{Z} = Z - Z_\epsilon$. As $Z, Z_\epsilon \in im\, \mathcal{D}^{\epsilon\,*}_{w_0}$ we may apply the key estimate Theorem 4.4.4 to obtain

$$\begin{aligned} \|Z - Z_\epsilon\|_{1,p,\epsilon} &\le c\left(\epsilon\|\mathcal{D}^\epsilon_{w_0}\tilde{Z}\|_{0,p,\epsilon} + \|\pi_\epsilon \mathcal{D}^\epsilon_{w_0}\tilde{Z}\|_p\right) \\ &\le c'\left(\|(d\mathcal{F}^{triv}_{\epsilon,u_0}(0)\tilde{Z})_1\|_p + \epsilon^{3/2}\|(d\mathcal{F}^{triv}_{\epsilon,u_0}(0)\tilde{Z})_2\|_p\right) \end{aligned} \tag{23}$$

where we used Lemma 4.2.4 for the π_ϵ-term. To estimate the last two terms add 0 and use

$$\mathcal{F}^{triv}_{\epsilon,u_0}(Z_\epsilon) = 0 = \mathcal{F}^{triv}_{\epsilon,u_0}(Z)$$

to get two differences based at Z_ϵ (better estimates than at $\tilde{Z}$)

$$\begin{aligned} \mathcal{D}^\epsilon_{w_0}\tilde{Z} = d\mathcal{F}^{triv}_{\epsilon,u_0}(0)\tilde{Z} &= -\mathcal{F}^{triv}_{\epsilon,u_0}(Z_\epsilon + \tilde{Z}) + \mathcal{F}^{triv}_{\epsilon,u_0}(Z_\epsilon) + d\mathcal{F}^{triv}_{\epsilon,u_0}(Z_\epsilon)\tilde{Z} \\ &\quad + \left(d\mathcal{F}^{triv}_{\epsilon,u_0}(0) - d\mathcal{F}^{triv}_{\epsilon,u_0}(Z_\epsilon)\right)\tilde{Z}. \end{aligned}$$

Now we are in position to apply the quadratic estimates I and II theorem 5.2.1 and 5.3.1. It turns out to be necessary to place some L^∞- and L^p-norms differently as in the quadratic estimates and therefore we restate them in the form needed here. Moreover, we use the estimates for (X_ϵ, Y_ϵ) obtained in the existence Theorem 1.0.4. For $(\tilde{X}, \tilde{Y})$ the same L^∞-estimates as for (X, Y) hold. We use $\|\tilde{X}\|_p \le 2\|\tilde{Z}\|_{1,p,\epsilon}$ and apply Lemma 4.2.6 to estimate $\|\tilde{X}\|_\infty \le \|\tilde{Z}\|_{\infty,\epsilon} \le c_p\, \epsilon^{-3/p}\, \|\tilde{Z}\|_{1,p,\epsilon}$.

$$\begin{aligned} &\|(d\mathcal{F}^{triv}_{\epsilon,u_0}(0)\tilde{Z})_1\|_p \\ &\le \|(\mathcal{F}^{triv}_{\epsilon,u_0}(Z_\epsilon + \tilde{Z}) - \mathcal{F}^{triv}_{\epsilon,u_0}(Z_\epsilon) - d\mathcal{F}^{triv}_{\epsilon,u_0}(Z_\epsilon)\tilde{Z})_1\|_p \\ &\quad + \|(d\mathcal{F}^{triv}_{\epsilon,u_0}(0)\tilde{Z} - d\mathcal{F}^{triv}_{\epsilon,u_0}(Z_\epsilon)\tilde{Z})_1\|_p \\ &\le c_1\|\tilde{X}\|_\infty\|\tilde{X}\|_\infty\Big(\|\nabla_t X_\epsilon\|_p + \underline{\|\nabla_s X_\epsilon\|_p} + \underline{\|\nabla_t Y_\epsilon\|_p}\Big) \\ &+ c_1\|\tilde{X}\|_p\Big(\|\tilde{X}\|_\infty + \underline{\|\nabla_t\tilde{X}\|_\infty} + \underline{\|\nabla_s\tilde{X}\|_\infty}\|\tilde{X}\|_\infty + \underline{\|\tilde{Y}\|_\infty} + \underline{\|\nabla_t\tilde{Y}\|_\infty}\|\tilde{X}\|_\infty\Big) \\ &+ c_1\|X_\epsilon\|_\infty\Big(\|X_\epsilon\|_\infty\big(\|\tilde{X}\|_p + \|\nabla_s\tilde{X}\|_p + \|\tilde{Y}\|_p + \|\nabla_t\tilde{Y}\|_p\big) + \|\tilde{Y}\|_\infty\|\nabla_t\tilde{X}\|_p\Big) \\ &+ c_1\|\tilde{X}\|_\infty\Big(\|\nabla_t X_\epsilon\|_p + \|\nabla_s X_\epsilon\|_p\|X_\epsilon\|_\infty + \|Y_\epsilon\|_p + \|\nabla_t Y_\epsilon\|_p\|X_\epsilon\|_\infty\Big) \\ &+ c_1\|X_\epsilon\|_\infty\Big(\|\tilde{X}\|_p + \|\nabla_t\tilde{X}\|_p + \|\nabla_s\tilde{X}\|_p\|X_\epsilon\|_\infty + \|\tilde{Y}\|_p\Big) \\ &\le c_2\|\tilde{Z}\|_{1,p,\epsilon}\left(\epsilon^{\frac{9}{4}-\frac{9}{2p}} + \epsilon^{\frac{3}{4}-\frac{3}{2p}} + \epsilon^{\frac{7}{4}-\frac{3}{p}} + \epsilon^{\frac{3}{2}-\frac{3}{p}} + \epsilon^{\frac{5}{4}-\frac{3}{2p}}\right) \\ &\le 2c_2\epsilon^{\frac{3}{4}-\frac{3}{2p}}\|\tilde{Z}\|_{1,p,\epsilon} \le \frac{1}{4c'}\,\|\tilde{Z}\|_{1,p,\epsilon} \end{aligned}$$

and

$$
\begin{aligned}
&\epsilon^{\frac{3}{2}}\|(d\mathcal{F}^{triv}_{\epsilon,u_0}(0)\tilde{Z})_2\|_p \\
&\leq \epsilon^{\frac{3}{2}}\|(\mathcal{F}^{triv}_{\epsilon,u_0}(Z_\epsilon+\tilde{Z})-\mathcal{F}^{triv}_{\epsilon,u_0}(Z_\epsilon)-d\mathcal{F}^{triv}_{\epsilon,u_0}(Z_\epsilon)\tilde{Z})_2\|_p \\
&\quad+\epsilon^{\frac{3}{2}}\|(d\mathcal{F}^{triv}_{\epsilon,u_0}(0)\tilde{Z}-d\mathcal{F}^{triv}_{\epsilon,u_0}(Z_\epsilon)\tilde{Z})_2\|_p \\
&\leq c_1\epsilon^{\frac{3}{2}}\|\tilde{X}\|^2_\infty\Big(\epsilon^{-2}\left(\|X_\epsilon\|_p+\underline{\|\nabla_t X_\epsilon\|_p}\right)+\|\nabla_s X_\epsilon\|_p+\|Y_\epsilon\|_p+\|\nabla_s Y_\epsilon\|_p\Big) \\
&\quad+c_1\epsilon^{\frac{3}{2}}\|\tilde{X}\|_p\Big(\epsilon^{-2}\|\tilde{X}\|_\infty\big(\|\tilde{X}\|_\infty+\underline{\|\nabla_t\tilde{X}\|_\infty}\big)+\|\nabla_s\tilde{X}\|_\infty \\
&\qquad\qquad+\|\tilde{Y}\|_\infty\big(1+\|\nabla_s X_\epsilon\|_\infty\big)\Big) \\
&\quad+c_1\epsilon^{\frac{3}{2}}\|X_\epsilon\|_\infty\Big(\epsilon^{-2}\big(\|\tilde{X}\|_p+\|\nabla_t\tilde{X}\|_p\big)\|X_\epsilon\|_\infty+\|\nabla_s\tilde{X}\|_p\|\tilde{Y}\|_\infty\Big) \\
&\quad+c_1\epsilon^{\frac{3}{2}}\|\tilde{X}\|_\infty\Big(\epsilon^{-2}\|\nabla_t X_\epsilon\|_p\|X_\epsilon\|_\infty+\|\nabla_s X_\epsilon\|_p+\|Y_\epsilon\|_p\Big) \\
&\quad+c_1\epsilon^{\frac{3}{2}}\|X_\epsilon\|_\infty\Big(\epsilon^{-2}\|\tilde{X}\|_p+\epsilon^{-2}\|\nabla_t\tilde{X}\|_p\|X_\epsilon\|_\infty+\|\nabla_s\tilde{X}\|_p+\|\tilde{Y}\|_p\Big) \\
&\leq c_2\|\tilde{Z}\|_{1,p,\epsilon}\left(\epsilon^{\frac{9}{4}-\frac{9}{2p}}+\epsilon^{\frac{3}{2}-\frac{3}{p}}+\epsilon^{\frac{5}{2}-\frac{3}{p}}+\epsilon^{\frac{9}{4}-\frac{3}{2p}}+\epsilon^{\frac{3}{2}-\frac{3}{2p}}\right) \\
&\leq 2c_2(\epsilon^{\frac{9}{4}-\frac{9}{2p}}+\epsilon^{\frac{3}{2}-\frac{3}{p}})\|\tilde{Z}\|_{1,p,\epsilon}\leq\frac{1}{4c'}\|\tilde{Z}\|_{1,p,\epsilon}
\end{aligned}
$$

for $\epsilon_0 > 0$ sufficiently small. We underlined the terms which enforce the assumptions on the L^∞-norms of (X, Y). Insert these estimates in (23) to obtain

$$\|Z - Z_\epsilon\|_{1,p,\epsilon} \leq \frac{1}{2}\|Z - Z_\epsilon\|_{1,p,\epsilon}$$

and so

$$Z = Z_\epsilon.$$

□

CHAPTER 3

The index theorem

For a nondegenerate perturbed closed geodesic $x \in Crit\, \mathcal{I}_V$ we would like to compare its Morse index $Ind(x)$ with the Conley-Zehnder index $\mu_{CZ}(z_x)$, where $z_x = g(x)\partial_t x$ is the corresponding 1-periodic Hamiltonian orbit.

THEOREM 3.0.1. (**Index**) *For any nondegenerate, closed perturbed geodesic $x \in Crit\, \mathcal{I}_V$*

$$Ind(x) = -\mu_{CZ}(z_x).$$

The relation between the Maslov index and the Morse index of a closed geodesic has been studied first, as far as we know, by Duistermaat [**D76**]. In the case of a closed geodesic on a flat torus, i.e. the corresponding Hamiltonian system evolves in $\mathbb{R}^{2n}$, theorem 3.0.1 had been obtained by Claude Viterbo [**V90**] with a slightly different definition of the Conley-Zehnder index (apart from the different normalization which causes a sign difference in the formula): due to the degeneracy of the action functional he considered the Conley-Zehnder index of the linearized Hamiltonian flow on the energy surface restricted to directions normal to the trajectory.

In what follows we will prove the theorem. The main idea is to construct a 2-parameter family of Lagrangian planes where the parameter domain is a square. In view of its contractibility the Maslov index of the loop around the boundary is zero. On the other hand it is additive under catenation of paths, so it remains to identify the Maslov indices of the four obvious subpaths with the quantities in the statement of the theorem. As there are different choices in the literature, let us first state the normalizations which we are going to use.

REMARK 3.0.2. (**Normalizations**) The signature of a symmetric matrix S is defined by

$$sign\ S = n^-(S) - n^+(S)$$

where $n^{\mp}(S)$ is the number of negative respectively positive eigenvalues of S. The Maslov index for Lagrangian planes, the Conley-Zehnder index and the spectral flow are normalized as follows

$$\mu_{Lag}(Gr\ e^{-tJ_0S}, \Delta) = \mu_{CZ}(e^{-tJ_0S}) = \tfrac{1}{2}\, sign\ S$$
$$\mu_{Spec}(\{\arctan t\}_{t\in(-\infty,\infty)}) = 1$$

where the constant symmetric matrix S satisfies $\|S\| < 2\pi$.

Let $\phi : S^1 \times \mathbb{R}^n \to x^*TM$ be an orthonormal trivialization (cf. appendix A.4) and $\{E_1, \ldots, E_n\}$ the associated orthonormal frame of the vector bundle $x^*TM \to S^1$. The (perturbed) Jacobi operator A_x^0 – which represents the Hessian of $\mathcal{I}_V$ at x – is given with respect to this frame by the self adjoint operator

$$\begin{aligned} I : L^2(S^1, \mathbb{R}^n) &\to L^2(S^1, \mathbb{R}^n) \\ \xi &\mapsto -\ddot{\xi} - Q\xi \end{aligned}$$

with dense domain $W^{2,2}(S^1, \mathbb{R}^n)$, where $\dot{\xi}$ denotes $\partial_t \xi$ and $Q : S^1 \to \mathcal{L}(\mathbb{R}^n)$ is a smooth family of symmetric matrices. We know that I has a real and discrete spectrum (appendix B.2.2) with finitely many negative eigenvalues – counted with multiplicities – and x nondegenerate is by definition equivalent to $0 \notin spec\, I$.

Now the linearized flow along the corresponding 1-periodic Hamiltonian orbit $z_x = g(x)\partial_t x$ is represented in the unitary frame

$$\Phi = \begin{pmatrix} \phi & 0 \\ 0 & {\phi^*}^{-1} \end{pmatrix} : S^1 \times \mathbb{R}^{2n} \to x^*TM \oplus x^*T^*M$$

by the smooth path of symplectic matrices determined by

$$\dot{\Psi}(t) = -J_0 S(t) \Psi(t) \quad , \quad \Psi(0) = \mathbb{1} \tag{24}$$

where

$$J_0 = \begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} \quad , \quad S(t) = \begin{pmatrix} Q(t) & 0 \\ 0 & \mathbb{1} \end{pmatrix}.$$

Observe that $0 \notin spec\, I$ is equivalent to $det\,(\mathbb{1} - \Psi(1)) \neq 0$, which in turn means that z is nondegenerate as a critical point of the symplectic action $\mathcal{A}_V$. Therefore the Conley-Zehnder index of the path Ψ is well-defined. Moreover, in between the lines we used the fact shown in remark A.2.2 that the linearized flow along z_x satisfies the linearized equations which, with respect to the unitary frame, take on the form (24).

Let $N = Ind(x) \in \mathbb{N}_0$ and denote the eigenvalues of I by

$$\mu_1 \leq \mu_2 \leq \ldots \leq \mu_N < 0 < \mu_{N+1} \leq \ldots .$$

Pick $\hat{\mu} < \min\{0, \mu_1\}$ and choose a monotonically decreasing cut-off function $\beta \in C^\infty([0,1], [\hat{\mu}, 0])$ which is identically 0 in a small neighborhood of 0 and identically $\hat{\mu}$ near 1, cf. figure 3.1. Let λ_i be determined by $\beta(\lambda_i) = \mu_i$ for $i = 1, \ldots, N$. It will be a crucial point later on (regularity of paths) to choose β in such a way that $\beta'(\lambda_i)$, $i = 1, \ldots, N$, is not an eigenvalue of the symmetric linear operator $Q(1) : \mathbb{R}^n \to R^n$.

Consider the family of self adjoint operators I_λ, $\lambda \in [0,1]$, given by

$$\begin{aligned} I_\lambda : L^2(S^1, \mathbb{R}^n) \supset W^{2,2}(S^1, \mathbb{R}^n) &\to L^2(S^1, \mathbb{R}^n) \\ \xi &\mapsto -\ddot{\xi} - (1-\lambda)Q\xi - \beta(\lambda)\xi. \end{aligned} \tag{25}$$

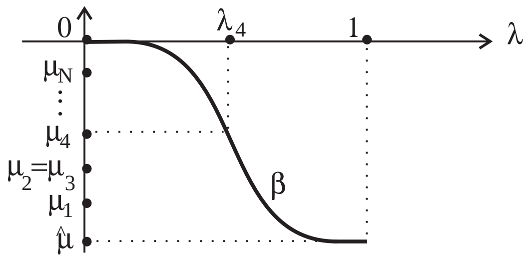

FIGURE 3.1. Cut-off function β and negative spectrum of I

Clearly $I_1 = -\partial_t\partial_t - \hat{\mu}\mathbb{1}$ is a positive operator and so $Ind\, I_1 = 0$. Note that $I_0 = I$. Studying the kernel of I_λ, we observe that $\xi_\lambda \in ker\ I_\lambda$ iff $\xi_\lambda : [0,1] \to \mathbb{R}^n$ solves

$$\begin{aligned} -\ddot{\xi}_\lambda - (1-\lambda)Q\xi_\lambda - \beta(\lambda)\xi_\lambda = 0 \\ \xi_\lambda(0) = \xi_\lambda(1) \quad , \quad \dot{\xi}_\lambda(0) = \dot{\xi}_\lambda(1) \end{aligned} \tag{26}$$

where the last two conditions reflect the periodicity of the domain S^1. We may rephrase this in terms of fundamental solutions as follows. For $k = 1, 2$ consider the solutions $\psi_{\lambda,k} : [0,1] \to Mat(n, \mathbb{R})$ of

$$\begin{gathered} -\partial_t\partial_t\psi_{\lambda,k} - (1-\lambda)Q\psi_{\lambda,k} - \beta(\lambda)\psi_{\lambda,k} = 0 \\ \psi_{\lambda,1}(0) = \mathbb{1} \qquad \psi_{\lambda,2}(0) = 0, \\ \dot{\psi}_{\lambda,1}(0) = 0 \qquad \dot{\psi}_{\lambda,2}(0) = \mathbb{1}. \end{gathered}$$

Define $\xi_\lambda^0 = \xi_\lambda(0)$ and $\eta_\lambda^0 = \eta_\lambda(0)$, then

$$\xi_\lambda \in ker\ I_\lambda \qquad \Leftrightarrow \qquad \begin{aligned} \xi_\lambda(t) = \psi_{\lambda,1}(t)\xi_\lambda^0 + \psi_{\lambda,2}(t)\eta_\lambda^0 \\ \xi_\lambda^0 \in ker\ (\mathbb{1} - \psi_{\lambda,1}(1)) \cap ker\ \dot{\psi}_{\lambda,1}(1), \\ \eta_\lambda^0 \in ker\ \psi_{\lambda,2}(1) \cap ker\ (\mathbb{1} - \dot{\psi}_{\lambda,2}(1)). \end{aligned}$$

Setting $\eta_\lambda = \dot{\xi}_\lambda$ equation (26) transforms into the first order system of ODE's

$$\left\{ \begin{aligned} \dot{\xi}_\lambda &= \eta_\lambda \\ \dot{\eta}_\lambda &= \ddot{\xi}_\lambda = -(1-\lambda)Q\xi_\lambda - \beta(\lambda)\xi_\lambda \end{aligned} \right. \tag{27}$$

whose fundamental solution $\Psi_\lambda : [0,1] \to Sp(2n, \mathbb{R})$ is determined by

$$\left\{ \begin{aligned} \partial_t\Psi_\lambda &= -J_0 S_\lambda \Psi_\lambda \\ \Psi_\lambda(0) &= \mathbb{1} \end{aligned} \right. \quad , \; S_\lambda(t) = \begin{pmatrix} (1-\lambda)Q(t) + \beta(\lambda)\mathbb{1} & 0 \\ 0 & \mathbb{1} \end{pmatrix} \tag{28}$$

Note that $\Psi_0 : [0,1] \to Sp(2n, \mathbb{R})$ is precisely the symplectic path obtained by linearizing the Hamiltonian flow on T^*M along the 1-periodic orbit $z_x = g(x)\partial_t x$ and so $\mu_{CZ}(z_x) = \mu_{CZ}(\Psi_0)$ by definition. Moreover, simple matrix multiplication and uniqueness of solutions of ODE's shows that

$$\Psi_\lambda(t) = \begin{pmatrix} \psi_{\lambda,1}(t) & \psi_{\lambda,2}(t) \\ \dot\psi_{\lambda,1}(t) & \dot\psi_{\lambda,2}(t) \end{pmatrix}$$

indeed satisfies (28) and so

$$\xi_{\lambda_i} \in ker\, I_{\lambda_i} \qquad \Leftrightarrow \qquad \begin{pmatrix} \xi_{\lambda_i}(0) \\ \dot\xi_{\lambda_i}(0) \end{pmatrix} \in ker\,(1\!\!1 - \Psi_{\lambda_i}(1)) \tag{29}$$

where $\xi_{\lambda_i}(t) = \psi_{\lambda_i,1}(t)\,\xi_{\lambda_i}(0) + \psi_{\lambda_i,2}(t)\,\dot\xi_{\lambda_i}(0)$.

The 2-parameter family $\Psi_\lambda(t)$ of solutions to (28) gives rise to a 2-parameter family $\Lambda_\lambda(t)$ of Lagrangian subspaces by pointwise taking its graph

$$\Lambda_\lambda(t) := Graph\,\Psi_\lambda(t) \subset (\mathbb{R}^{2n} \times \mathbb{R}^{2n}, -\omega_0 \oplus \omega_0).$$

Here $\omega_0 = dx_i \wedge dy^i$ is the standard symplectic form on $\mathbb{R}^{2n}$ equipped with coordinates $(x_1, \ldots, x_n, y^1, \ldots, y^n)$. As the parameter domain $[0,1]\times[0,1]$ is contractible, the Maslov index μ_{Lag} of the loop Γ of Lagrangian subspaces (relative to the diagonal $\Delta \subset \mathbb{R}^{2n} \times \mathbb{R}^{2n}$) obtained by going around the boundary of the domain clockwise is zero. Let the paths of Lagrangian subspaces γ_i be as indicated in figure 3.2 and $\tilde\Gamma = \gamma_4\gamma_3\gamma_2\gamma_1$ be their composition in the sense of paths (i.e. first follow γ_1, then $\gamma_2 \ldots$).

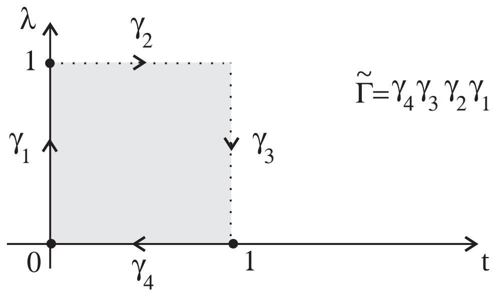

FIGURE 3.2. Contractible loop of Lagrangian subspaces

The Maslov index for paths is additive under composition of paths (cf. [**RS93**] thm. 2.3 CATENATION) and so

$$0 = \mu_{Lag}(\tilde\Gamma, \Delta) = \sum_{i=1}^{4} \mu_{Lag}(\gamma_i, \Delta).$$

LEMMA 3.0.3. *i)* $\mu_{Lag}(\gamma_1, \Delta) = 0$
ii) $\mu_{Lag}(\gamma_2, \Delta) = 0$
iii) $\mu_{Lag}(\gamma_3, \Delta) = -Ind(I)$
iv) $\mu_{Lag}(\gamma_4, \Delta) = -\mu_{CZ}(\Psi_0)$.

The lemma implies theorem 3.0.1. Before proving it we shall briefly recall ([**RS93**]) the definition of the *Maslov index* $\mu_{Lag}(\Lambda, V)$ *of a path of Lagrangian subspaces* Λ of a symplectic vector space with respect to a fixed Lagrangian V. t_i is called a *crossing* if $\Lambda(t_i) \cap V \neq \{0\}$. At t_i there is a quadratic form on $\Lambda(t_i)$: Pick any Lagrangian complement W of $\Lambda(t_i)$ and for $v \in \Lambda(t_i)$ and sufficiently small $\delta > 0$ define $w(\delta) \in W$ by $v + w(\delta) \in \Lambda(t_i + \delta)$. Then

$$\hat{Q}(\partial_t \Lambda(t_i))\, v := \tfrac{d}{d\delta}\big|_{\delta=0} \Omega(v, w(\delta))$$

is a quadratic form on $\Lambda(t_i)$ and independent of the choice of W ([**RS93**], thm. 1.1). The *crossing form at* t_i is the quadratic form on $\Lambda(t_i) \cap V$ defined by

$$\Gamma(\Lambda, V, t_i) = \hat{Q}(\partial_t \Lambda(t_i))|_{\Lambda(t_i)\cap V}$$

and t_i is called a *regular crossing* if the crossing form is nonsingular.

For the special case $\Lambda(t) = Graph\, \Psi(t) \subset (\mathbb{R}^{2n} \times \mathbb{R}^{2n}, -\omega_0 \oplus \omega_0)$, where $\Psi : [0,1] \to Sp(2n, \mathbb{R})$ is a smooth path and $V = \Delta = \{(\zeta, \zeta) \mid \zeta \in \mathbb{R}^{2n}\}$ is the diagonal, we derive an explicit formula for the crossing form at a crossing t_i. Observe first that Ψ determines a smooth path of symmetric matrices S by

$$\partial_t \Psi(t) = -J_0\, S(t)\, \Psi(t)$$

or, equivalently,

$$S(t) = J_0\, \partial_t \Psi(t)\, \Psi(t)^{-1}.$$

For $v \in Graph\, \Psi(t_i) \cap \Delta$, i.e. $v = (\zeta, \zeta) = (\zeta, \Psi(t_i)\zeta)$, and $W := 0 \times \mathbb{R}^{2n}$ we have $w(\delta) = (0, w_2(\delta))$ and the condition $v + w(\delta) \in Graph\, \Psi(t_i + \delta)$ leads to

$$(\zeta, \zeta) + (0, w_2(\delta)) = (\zeta, \Psi(t_i + \delta)\zeta)$$

and so

$$w_2(\delta) = \Psi(t_i + \delta)\zeta - \zeta.$$

This implies

$$\begin{aligned}
\hat{Q}(\partial_t \Lambda(t_i))\, v &= \tfrac{d}{d\delta}\big|_{\delta=0} (-\omega_0 \oplus \omega_0)\Big((\zeta, \zeta), (0, w_2(\delta))\Big) \\
&= \tfrac{d}{d\delta}\big|_{\delta=0} \Big(-\omega_0(\zeta, 0) + \omega_0(\zeta, \Psi(t_i + \delta)\zeta - \zeta)\Big) \\
&= \omega_0(\zeta, \partial_t \Psi(t_i)\zeta) \\
&= -\langle \zeta, S(t_i)\Psi(t_i)\zeta \rangle \\
&= -\langle \zeta, S(t_i)\zeta \rangle,
\end{aligned}$$

where we used $\omega_0(\cdot, J_0 \cdot) = \langle \cdot, \cdot \rangle$. Identifying $(Graph\, \Psi(t_i)) \cap \Delta$ with $ker\, (\mathbb{1} - \Psi(t_i))$ via $(\zeta, \zeta) \mapsto \zeta$ we may write

$$\Gamma(Graph\, \Psi, \Delta, t_i) \cdot = -\langle \cdot, S(t_i) \cdot \rangle\, |_{ker\, \mathbb{1} - \Psi(t_i)}\, .$$

The Maslov index of the path $Graph\,\Psi$ with respect to the diagonal Δ is then given by

$$\begin{aligned}
&\mu_{Lag}(Graph\,\Psi, \Delta)\\
&\quad\overset{\text{def}}{=} -\tfrac{1}{2}\, sign\, \Gamma(Graph\,\Psi, \Delta, 0) - \tfrac{1}{2}\, sign\, \Gamma(Graph\,\Psi, \Delta, 1)\\
&\qquad - \sum_{0<t_i<1} sign\, \Gamma(Graph\,\Psi, \Delta, t_i)\\
&\quad= \tfrac{1}{2}\, sign\, \langle\cdot, S(0)\cdot\rangle\,|_{ker\ \mathbb{1}-\Psi(0)} + \tfrac{1}{2}\, sign\, \langle\cdot, S(1)\cdot\rangle\,|_{ker\ \mathbb{1}-\Psi(1)}\\
&\qquad + \sum_{0<t_i<1} sign\, \langle\cdot, S(t_i)\cdot\rangle\,|_{ker\ \mathbb{1}-\Psi(t_i)}
\end{aligned} \tag{30}$$

where the sums are over all t_i with $det(\mathbb{1} - \Psi(t_i)) = 0$.

PROOF. (of lemma 3.0.3) **ad *i*)** As $\Psi_\lambda(0) = \mathbb{1}$ for $\lambda \in [0,1]$ we get $\gamma_1(\lambda) = \Lambda_\lambda(0) = \Delta$, but the Maslov index for a constant path is zero ([**RS93**], thm. 2.3 ZERO).
ad *ii*) γ_2 is the graph of the path $\Psi_1 = \Psi_1(\cdot)$ and so by formula (30) and the fact that there are no crossings for $t > 0$

$$\mu_{Lag}(\gamma_2, \Delta) = \mu_{Lag}(Graph\,\Psi_1, \Delta) = \tfrac{1}{2}\, sign\, \langle\cdot, S_1(0)\cdot\rangle\,|_{\mathbb{R}^{2n}} = 0$$

where $R^{2n} = ker\,(\mathbb{1} - \Psi_1(0))$. The last step follows because the signature of

$$S_1(0) = \begin{pmatrix} \hat{\mu}\mathbb{1} & 0 \\ 0 & \mathbb{1} \end{pmatrix} \quad , \quad \hat{\mu} < 0$$

is zero.

It remains to show that there are no crossings $t_i > 0$. $S_1(t) = S_1(0)$ for all $t \in [0,1]$ and so we can solve (28) explicitely

$$\begin{aligned}
\Psi_1(t) = e^{-tJ_0S_1(0)} &= exp\begin{pmatrix} 0 & t\,\mathbb{1} \\ -t\,\mathbb{1}\hat{\mu} & 0 \end{pmatrix}\\
&= \begin{pmatrix} \cosh\sqrt{-\hat{\mu}}t\,\mathbb{1} & (-\hat{\mu})^{-1/2}\sinh\sqrt{-\hat{\mu}}t\,\mathbb{1} \\ (-\hat{\mu})^{1/2}\sinh\sqrt{-\hat{\mu}}t\,\mathbb{1} & \cosh\sqrt{-\hat{\mu}}t\,\mathbb{1} \end{pmatrix}.
\end{aligned}$$

Studying its characteristic polynomial, the eigenvalues of $\Psi_1(t)$ (of multiplicity n each) turn out to be

$$\rho_\pm(t) = \cosh\sqrt{-\hat{\mu}}t \pm \sinh\sqrt{-\hat{\mu}}t \quad , \quad t \geq 0$$

so that (cf. figure D.1 in appendix D)

$$\rho_+(t) = 1 = \rho_-(t) \;\Leftrightarrow\; t = 0.$$

Let us remark that as $\Psi_1(t)$ is symplectic, it follows $\rho_+(t) = \rho_-(t)^{-1}$ and this fact is reflected in the key identity for hyperbolic functions $\cosh^2 - \sinh^2 \equiv 1$.
ad *iv*) Note that the Maslov index changes its sign if a path of Lagrangian subspaces is traversed in the opposite direction. Let $\tilde{\gamma}_4(t) = \gamma_4(1-t)$, then

$$\mu_{Lag}(\gamma_4, \Delta) = -\mu_{Lag}(\tilde{\gamma}_4, \Delta) = -\mu_{Lag}(Graph\,\Psi_0, \Delta) = -\mu_{CZ}(\Psi_0)$$

where the last statement is shown in [**RS93**] remark 5.4 .

ad *iii*) The idea is to study the *spectral flow* of the family of selfadjoint operators $\{I_\lambda\}_{\lambda\in[0,1]}$, which is defined to be the number of eigenvalues changing sign from $-$ to $+$ minus the number changing from $+$ to $-$ during the deformation from $I_0 = -\partial_t\partial_t - Q(t) = I$ to $I_1 = -\partial_t\partial_t - \hat{\mu}$. I_1 is positive definite as $\hat{\mu} < 0$ and so

$$\mu_{Spec}(\{I_\lambda\}_{\lambda\in[0,1]}) = \#(- \curvearrowright +) - \#(+ \curvearrowright -) = Ind(I).$$

On the other hand it is a main result in [**RS95**], lemma 4.27, that the spectral flow may also be calculated as the sum of the signatures of certain crossing operators at regular crossings (these terms will be defined later on) and this gives the first equality in

$$\begin{aligned}
\mu_{Spec}(\{I_\lambda\}_{\lambda\in[0,1]}) &= \sum_{0\le\lambda_j\le 1} sign\, \Gamma(\{I_\lambda\}_{\lambda\in[0,1]}, \lambda_j)\ |_{ker\ I_{\lambda_j}} \\
&= \sum_{0\le\lambda_j\le 1} sign\, \langle\cdot, \partial_\lambda I_{\lambda_j}\cdot\rangle_{L^2}\ |_{ker\ I_{\lambda_j}} \\
&= \sum_{0\le\lambda_i\le 1} -sign\, \langle\cdot, \hat{S}_1(\lambda_i)\cdot\rangle\ |_{ker\ (\mathbb{1}-\Psi_{\lambda_i}(1))} \\
&= \sum_{0\le\lambda_i\le 1} sign\, \Gamma(Gr\ \Psi_{.}(1), \Delta, \lambda_i)\ |_{\Delta\cap Gr\Psi_{\lambda_i}(1)} \\
&= \mu_{Lag}(Gr\ \Psi_{.}(1), \Delta) \\
&= \mu_{Lag}(\tilde{\gamma}_3, \Delta) \\
&= -\mu_{Lag}(\gamma_3, \Delta)
\end{aligned}$$

where $\tilde{\gamma}_3(\lambda) = \gamma_3(1-\lambda)$. The fifth equality is by definition and the fourth one is given by formula (30) where the path of symmetric matrices $\hat{S}_t(\lambda)$ is determined for fixed t by the symplectic path $\lambda \mapsto \Psi_\lambda(t)$ by

$$\hat{S}_t(\lambda) = J_0\, \partial_\lambda\Psi_\lambda(t)\ \Psi_\lambda(t)^{-1}.$$

Note that $\hat{S}_0(\lambda) = 0$ for all λ. We will now define the crossing operator for the operator family, thereby establishing equality two, and then finally prove the third equality by showing that the crossings are the same and the corresponding quadratic forms are isomorphic.

Following [**RS95**] we define the *crossing operator*

$$\Gamma(\{I_\lambda\}_{\lambda\in[0,1]}, \lambda_i) = P_{\lambda_i}\ (\partial_\lambda I_{\lambda_i})\ P_{\lambda_i}\ |_{ker\ I_{\lambda_i}}$$

for the family $I_\lambda : W^{2,2}(S^1, \mathbb{R}^n) \to L^2(S^1, \mathbb{R}^n)$, where $P_{\lambda_i} : L^2(S^1, \mathbb{R}^n) \to L^2(S^1, \mathbb{R}^n)$ denotes the orthogonal projection onto the kernel of I_{λ_i}. λ_i is called a *crossing* if I_{λ_i} is not injective and it is *regular* if in addition the crossing operator is nonsingular. Now the second equality follows, because the orthogonal projections are selfadjoint and act as the identity on $ker\ I_{\lambda_i}$.

We had already shown that the crossings in the third equality are the same, namely recall (29)

$$\xi_{\lambda_i} \in ker\, I_{\lambda_i} \;\Leftrightarrow\; \zeta_0 := \begin{pmatrix} \xi_0 \\ \eta_0 \end{pmatrix} := \begin{pmatrix} \xi_{\lambda_i}(0) \\ \dot{\xi}_{\lambda_i}(0) \end{pmatrix} \in ker\,(\mathbb{1} - \Psi_{\lambda_i}(1))$$

where $\xi_{\lambda_i}(t) = \psi_{\lambda_i,1}(t)\,\xi_{\lambda_i}(0) + \psi_{\lambda_i,2}(t)\,\dot{\xi}_{\lambda_i}(0)$. Hence it remains to show that for all $\xi_{\lambda_i} \in ker\, I_{\lambda_i}$

$$\langle \xi_{\lambda_i}, (\partial_\lambda I_{\lambda_i})\, \xi_{\lambda_i}\rangle_{L^2} = -\langle \zeta_0, \hat{S}_1(\lambda_i)\, \zeta_0\rangle.$$

Integrate the identity (obtained by using several times the defining equations for $S_\lambda(t)$ and $\hat{S}_t(\lambda)$)

$$\begin{aligned}
&\partial_t(\Psi_{\lambda_i}(t)^T\, \hat{S}_t(\lambda_i)\, \Psi_{\lambda_i}(t)) \\
&= (\partial_t \Psi_{\lambda_i}(t)^T)\, \hat{S}_t(\lambda_i)\, \Psi_{\lambda_i}(t) + \Psi_{\lambda_i}(t)^T\, \partial_t(J_0 \partial_\lambda \Psi_{\lambda_i}(t)) \\
&= \Psi_{\lambda_i}(t)^T\, S_{\lambda_i}(t)\, J_0\, \hat{S}_t(\lambda_i)\, \Psi_{\lambda_i}(t) + \Psi_{\lambda_i}(t)^T\, J_0\, \partial_\lambda(-J_0\, S_{\lambda_i}(t)\, \Psi_{\lambda_i}(t)) \\
&= -\Psi_{\lambda_i}(t)^T\, S_{\lambda_i}(t)\, \partial_\lambda \Psi_{\lambda_i}(t) + \Psi_{\lambda_i}(t)^T\, (\partial_\lambda S_{\lambda_i}(t))\, \Psi_{\lambda_i}(t) \\
&\quad + \Psi_{\lambda_i}(t)^T\, S_{\lambda_i}(t)\, \partial_\lambda \Psi_{\lambda_i}(t) \\
&= \Psi_{\lambda_i}(t)^T\, (\partial_\lambda S_{\lambda_i}(t))\, \Psi_{\lambda_i}(t)
\end{aligned} \tag{31}$$

over t from 0 to 1 and use $\hat{S}_0(\lambda) \equiv 0$ to get

$$\Psi_{\lambda_i}(t)^T\, \hat{S}_1(\lambda_i)\, \Psi_{\lambda_i}(1) = \int_0^1 \Psi_{\lambda_i}(t)^T\, (\partial_\lambda S_{\lambda_i}(t))\, \Psi_{\lambda_i}(t)\, dt.$$

Now use $\partial_\lambda I_{\lambda_i} = Q - \beta'(\lambda_i)$ and $\xi_{\lambda_i} \in ker I_{\lambda_i}$ to obtain

$$\begin{aligned}
&\langle \xi_{\lambda_i}, (\partial_\lambda I_{\lambda_i})\, \xi_{\lambda_i}\rangle_{L^2} \\
&= \langle \xi_{\lambda_i}, (Q - \beta'(\lambda_i))\, \xi_{\lambda_i}\rangle_{L^2} \\
&= \int_0^1 \langle \psi_{\lambda_i,1}(t)\xi^0_{\lambda_i} + \psi_{\lambda_i,2}(t)\eta^0_{\lambda_i}, (Q(t) - \beta'(\lambda_i))(\psi_{\lambda_i,1}(t)\xi^0_{\lambda_i} + \psi_{\lambda_i,2}(t)\eta^0_{\lambda_i})\rangle dt \\
&= \int_0^1 \left\langle \Psi_{\lambda_i}(t) \begin{pmatrix} \xi^0_{\lambda_i} \\ \eta^0_{\lambda_i} \end{pmatrix}, \begin{pmatrix} Q(t) - \beta'(\lambda_i) & 0 \\ 0 & 0 \end{pmatrix} \Psi_{\lambda_i}(t) \begin{pmatrix} \xi^0_{\lambda_i} \\ \eta^0_{\lambda_i} \end{pmatrix} \right\rangle dt \\
&= \left\langle \begin{pmatrix} \xi^0_{\lambda_i} \\ \eta^0_{\lambda_i} \end{pmatrix}, \left(\int_0^1 \Psi_{\lambda_i}(t)^T\, (-\partial_\lambda S_{\lambda_i}(t))\, \Psi_{\lambda_i}(t)\, dt \right) \begin{pmatrix} \xi^0_{\lambda_i} \\ \eta^0_{\lambda_i} \end{pmatrix} \right\rangle \\
&= -\left\langle \Psi_{\lambda_i}(1) \begin{pmatrix} \xi^0_{\lambda_i} \\ \eta^0_{\lambda_i} \end{pmatrix}, \hat{S}_1(\lambda_i)\, \Psi_{\lambda_i}(1) \begin{pmatrix} \xi^0_{\lambda_i} \\ \eta^0_{\lambda_i} \end{pmatrix} \right\rangle \\
&= -\langle \zeta^0_{\lambda_i}, \hat{S}_1(\lambda_i) \zeta^0_{\lambda_i}\rangle
\end{aligned}$$

where we used formula (3) for $\Psi_{\lambda_i}(t)$, the identity (31) and $\hat{S}_0(\lambda_i) = 0$ in the last but one, as well as (29) in the last step. Our choice of cut-off function (recall $\beta'(\lambda_i) \notin spec\, Q(1)$) guarantees that $\{I_\lambda\}_{\lambda \in [0,1]}$ indeed is a family with regular crossings only and – in view of the isomorphism just shown – this implies that $\lambda \mapsto \Psi_\lambda(1)$ is a regular path, too. More precisely, assume $\xi \in \mathit{ker}\, \partial_\lambda I_{\lambda_i} \cap \mathit{ker}\, I_{\lambda_i}$, then

$$\begin{aligned} Q(t)\xi(t) &= \beta'(\lambda_i)\xi(t) \quad \forall t \in [0,1] \\ \xi(0) = \xi(1) \quad &, \quad \dot{\xi}(0) = \dot{\xi}(1). \end{aligned} \tag{32}$$

Now $\beta'(\lambda_i) \notin \mathit{spec}\, Q(1)$ implies $\xi(1) = 0$. Differentiating (32) with respect to t leads at $t = 1$ to

$$(\beta'(\lambda_i) - Q(1))\dot{\xi}(1) = \dot{Q}(1)\xi(1) = 0$$

and in view of the nonsingularity of $\beta'(\lambda_i) - Q(1)$ we get $\dot{\xi}(1) = 0$. As ξ is also in the kernel of the second order differential operator I_{λ_i} and all its boundary conditions are zero, it follows $\xi = 0$. This proves nondegeneracy of the crossing operator $\Gamma(\{I_\lambda\}_{\lambda \in [0,1]}, \lambda_i)$. □

CHAPTER 4

Elliptic estimates

The main result in this chapter is Theorem 4.4.4 which provides the key estimates for the right inverse of $\mathcal{D}_w^\epsilon$ uniformly in $\epsilon > 0$ sufficiently small.

Due to the nonlinearities we are forced to choose $p > 2$. However, it will be convenient to deal with the case $p = 2$ first in section 4.1 as the combination of the Hilbert space structure with Young's inequality simplifies the estimates enormously. Section 4.2 then generalizes the results of section 4.1 to the case $p > 2$. In section 4.3 we introduce the important technique of rescaling in the proof of the linear estimate Theorem 4.3.2. That way we reduce the proof to the Calderon-Zygmund inequality, the basic elliptic L^p-estimate which holds uniformly for all compactly supported functions on $\mathbb{R}^2$. The formal adjoint operator is introduced in section 4.4 in order to define the right inverse $\mathcal{Q}_w^\epsilon$ of $\mathcal{D}_w^\epsilon$. Finally the main estimate of $\mathcal{D}_w^\epsilon$ on the range of ${\mathcal{D}_w^\epsilon}^*$ is derived.

4.1. Nonstandard estimates for $p = 2$

Pick two smooth loops x, y in M and a smooth cylinder u between them, i.e. $u \in \mathcal{P}_{x,y}(\mathbb{R} \times S^1, M)$, and define $w = g(u)\partial_t u$. Note that the choice of boundary conditions guarantees the boundedness of the linear operators. Our goal in this section will be to show that surjectivity of $\mathcal{D}_u^0$ implies surjectivity of $\mathcal{D}_w^\epsilon$ for all $\epsilon > 0$ sufficiently small. We formulate and prove this in terms of injectivity of the formally adjoint operators ${\mathcal{D}_u^0}^*$ and ${\mathcal{D}_w^\epsilon}^*$. In this section we mainly work in the orthonormal frames introduced in appendix A section A.3, so that for $\vec{\xi}, \vec{\eta} \in C_0^\infty(\mathbb{R} \times S^1, \mathbb{R}^n)$

$$\mathcal{D}_0^* \vec{\xi} = -\partial_s \vec{\xi} - \vec{\nabla}_t \vec{\nabla}_t \vec{\xi} - S\vec{\xi} \tag{33}$$

and

$$\mathcal{D}_\epsilon^* \begin{pmatrix} \vec{\xi} \\ \vec{\eta} \end{pmatrix} = \begin{pmatrix} -\partial_s \vec{\xi} - \vec{\nabla}_t \vec{\eta} - S\vec{\xi} + \epsilon^2 B^* \vec{\eta} \\ -\partial_s \vec{\eta} + \epsilon^{-2}(\vec{\nabla}_t \vec{\xi} - \vec{\eta}) \end{pmatrix} \tag{34}$$

where we have taken the latter adjoint with respect to the ϵ-dependent Hilbert space structure $\langle \cdot, \cdot \rangle_\epsilon$ suggested by (11) with $p = 2$.

Recall the inclusion

$$\iota : C_0^\infty(\mathbb{R} \times S^1, u^*TM) \to C_0^\infty(\mathbb{R} \times S^1, u^*TM \oplus u^*T^*M) \quad ,$$
$$\xi_0 \mapsto (\xi_0, g(u)\nabla_t \xi_0).$$

To compare the operators ${D_u^0}^*$ and ${D_w^\epsilon}^*$ we must choose a projection onto the image of this embedding. A natural candidate would be the orthogonal projection $\pi_\epsilon^\perp$ with respect to the Hilbert space structure $\langle \cdot, \cdot \rangle_\epsilon$.

LEMMA 4.1.1. *The orthogonal projection $\pi_\epsilon^\perp$ from $L^2_\epsilon(\mathbb{R} \times S^1, u^*TM \oplus u^*T^*M)$ to $\iota \circ L^2(\mathbb{R} \times S^1, u^*TM)$ with respect to $\langle \cdot, \cdot \rangle_\epsilon$ is given by*

$$\pi_\epsilon^\perp(\xi, \eta) = (\mathbb{1} - \epsilon^2 \nabla_t \nabla_t)^{-1}(\xi - \epsilon^2 g^{-1}(u) \nabla_t \eta).$$

PROOF. The condition for $\pi_\epsilon^\perp$ to be an orthogonal projection is

$$\langle \iota \pi_\epsilon^\perp(\xi, \eta) - (\xi, \eta), \iota \xi_0 \rangle_\epsilon = 0 \quad , \forall \xi_0 \in L^2(\mathbb{R} \times S^1, u^*TM).$$

Setting $\xi_1 = \pi_\epsilon^\perp(\xi, \eta)$ this leads to

$$\begin{aligned} 0 &= \langle (\xi_1, g \nabla_t \xi_1) - (\xi, \eta), (\xi_0, g \nabla_t \xi_0) \rangle_\epsilon \\ &= \langle \xi_1 - \xi, \xi_0 \rangle + \epsilon^2 \langle \nabla_t \xi_1 - g^{-1} \eta, \nabla_t \xi_0 \rangle \\ &= \langle \xi_0, (\mathbb{1} - \epsilon^2 \nabla_t \nabla_t) \xi_1 - \xi + \epsilon^2 g^{-1} \nabla_t \eta \rangle \end{aligned}$$

for all $\xi_0 \in L^2(\mathbb{R} \times S^1, u^*TM)$, hence nondegeneracy implies

$$\xi_1 = (\mathbb{1} - \epsilon^2 \nabla_t \nabla_t)^{-1}(\xi - \epsilon^2 g^{-1} \nabla_t \eta).$$

□

Instead we define more generally

$$\boxed{\pi_\epsilon(\xi, \eta) = (\mathbb{1} - \epsilon^\alpha \nabla_t \nabla_t)^{-1}(\xi - \epsilon^\beta g^{-1} \nabla_t \eta)} \tag{35}$$

where $\alpha, \beta \in \mathbb{R}$. As we will see during the proof of lemma 4.1.3 and proposition 4.1.2 the right choices are $\beta = 2$ respectively $\alpha = 1$. The significance of these definitions lies in the next proposition and subsequent four lemmata. All the norms in this section are L^2-norms unless otherwise indicated.

PROPOSITION 4.1.2. *Let $u \in \mathcal{P}_{x,y}(\mathbb{R} \times S^1, M)$, where x, y are smooth loops in M, and define $w = g(u)\partial_t u$. Then for every constant $c_0 > 0$ there exist constants $c > 0$, $\epsilon_0 > 0$ such that the following holds. If the injectivity assumption*

$$\|\xi\| + \|\nabla_t \nabla_t \xi\| \leq c_0 \|{\mathcal{D}_u^0}^* \xi\|$$

holds for all $\xi \in C_0^\infty(\mathbb{R} \times S^1, u^*TM)$, then

$$\boxed{\begin{aligned} \|\xi\| &\leq c\Big(\epsilon \|{\mathcal{D}_w^\epsilon}^*(\xi, \eta)\|_{0,2,\epsilon} + \|\pi_\epsilon {\mathcal{D}_w^\epsilon}^*(\xi, \eta)\|\Big) \\ \|\eta\| &\leq c\Big(\epsilon^{1/2} \|{\mathcal{D}_w^\epsilon}^*(\xi, \eta)\|_{0,2,\epsilon} + \|\pi_\epsilon {\mathcal{D}_w^\epsilon}^*(\xi, \eta)\|\Big) \end{aligned}}$$

and therefore

$$\boxed{\begin{aligned} \|\zeta\|_{0,2,\epsilon} &\leq c\Big(\epsilon \|{\mathcal{D}_w^\epsilon}^* \zeta\|_{0,2,\epsilon} + \|\pi_\epsilon {\mathcal{D}_w^\epsilon}^* \zeta\|\Big) \\ \|\zeta\|_{0,2,\epsilon} &\leq c \|{\mathcal{D}_w^\epsilon}^* \zeta\|_{0,2,\epsilon} \end{aligned}}$$

for $\epsilon \in (0, \epsilon_0)$ *and* $\zeta = (\xi, \eta) \in C_0^\infty(\mathbb{R} \times S^1, u^*TM \oplus u^*T^*M)$, *where we set* $\alpha = 1$ *and* $\beta = 2$.

Note that in combining the estimates for ξ and η the extra factor $\epsilon^{1/2}$ in the η-estimate is wasted. For the convergence of the Newton method later on it will be a crucial point to remember this fact and use the estimates for ξ and η separately.

Lemma 4.1.3. *Let* $u \in \mathcal{P}_{x,y}(\mathbb{R} \times S^1, M)$, x, y *smooth loops in* M, $\beta = 2$ *and define* $w = g(u)\partial_t u$. *Then there exists a constant* $c > 0$ *such that*

$$\boxed{\begin{aligned} \|\mathcal{D}_u^{0\,*}\pi_\epsilon\zeta - \pi_\epsilon\mathcal{D}_w^{\epsilon\,*}\zeta\| &\le c\epsilon^{\alpha/2}\|\zeta\|_{0,2,\epsilon} + c\epsilon^{2-\alpha}\|\nabla_t\eta\| \\ &\le c\epsilon^{2-\alpha}\|\mathcal{D}_w^{\epsilon\,(*)}\zeta\|_{0,2,\epsilon} + c(\epsilon^{\alpha/2} + \epsilon^{3/2-\alpha})\|\zeta\|_{0,2,\epsilon} \end{aligned}}$$

for $0 < \epsilon \le 1$, $0 \le \alpha \le 2$ *and* $\zeta = (\xi, \eta) \in C_0^\infty(\mathbb{R} \times S^1, u^*TM \oplus u^*T^*M)$. *The same estimate holds for* $\mathcal{D}_u^0\pi_\epsilon - \pi_\epsilon\mathcal{D}_w^\epsilon$.

In the proof of lemma 4.1.3 it will be a crucial point to set $\beta = 2$ in order to eliminate certain terms. Moreover, mixed estimates like $\|\mathcal{D}_u^{0\,*}\pi_\epsilon\zeta - \pi_\epsilon\mathcal{D}_w^\epsilon\zeta\|$ and $\|\mathcal{D}_u^0\pi_\epsilon\zeta - \pi_\epsilon\mathcal{D}_w^{\epsilon\,*}\zeta\|$ will contain a term of the form $c\|\nabla_s\xi\|$. The absence of any ϵ in this term obstructs getting an injectivity estimate for $\mathcal{D}_w^\epsilon$ – assuming one for $\mathcal{D}_u^{0\,*}$ – along the lines of the proof of proposition 4.1.2. As in our case of interest $\mathcal{D}_w^\epsilon$ is not injective anyway, the best we can hope for is an injectivity estimate for $\mathcal{D}_w^\epsilon$ on the image of $\mathcal{D}_w^{\epsilon\,*}$. This will be the crucial estimate for the Newton method and it will be carried out in section 4.4. From now on we set $\beta = 2$ unless mentioned otherwise.

Lemma 4.1.4. *Let* $u \in C^\infty(\mathbb{R} \times S^1, M)$, *then*

$$\boxed{\begin{aligned} \|\xi - \pi_\epsilon\zeta\| &\le \epsilon^{\alpha/2}\|\eta - g(u)\nabla_t\xi\| + 2\epsilon^\alpha\|\nabla_t\eta\| \\ \|\eta - g(u)\nabla_t\pi_\epsilon\zeta\| &\le \|\eta - g(u)\nabla_t\xi\| + 2\epsilon^{\alpha/2}\|\nabla_t\eta\| \\ \|\zeta - \iota\pi_\epsilon\zeta\|_{0,2,\epsilon} &\le 2\epsilon^{\alpha/2}\|\eta - g(u)\nabla_t\xi\| + 4\epsilon^\alpha\|\nabla_t\eta\| \\ \|\pi_\epsilon\zeta\| &\le \|\iota\pi_\epsilon\zeta\|_{0,2,\epsilon} \le 2\|\zeta\|_{0,2,\epsilon} \end{aligned}}$$

for $0 < \epsilon \le 1$, $0 \le \alpha \le 2$ *and* $\zeta = (\xi, \eta) \in C_0^\infty(\mathbb{R} \times S^1, u^*TM \oplus u^*T^*M)$.

Lemma 4.1.5. *Let* $u \in C^\infty(\mathbb{R} \times S^1, M)$, *then*

$$\boxed{\begin{aligned} \|(\mathbb{1} - \epsilon^\alpha\nabla_t\nabla_t)^{-1}\xi\| &\le \|\xi\| \\ \|(\mathbb{1} - \epsilon^\alpha\nabla_t\nabla_t)^{-1}\epsilon^{\alpha/2}\nabla_t\xi\| &\le \|\xi\| \\ \|(\mathbb{1} - \epsilon^\alpha\nabla_t\nabla_t)^{-1}\epsilon^\alpha\nabla_t\nabla_t\xi\| &\le \|\xi\| \end{aligned}}$$

for $\epsilon > 0$, $\alpha \in \mathbb{R}$ *and* $\xi \in C_0^\infty(\mathbb{R} \times S^1, u^*TM)$.

Motivation for the next lemma comes from the energy identity (10). It turns out to be the fundamental estimate to carry out the Newton iteration in the proof of the main theorem. It also points the way towards the right definition of the norm $\|\cdot\|_{1,p,\epsilon}$ by comparing powers of ϵ appearing.

Lemma 4.1.6. *Let* $u \in \mathcal{P}_{x,y}(\mathbb{R}\times S^1, M)$*,* x, y *smooth loops in* M*, and define* $w = g(u)\partial_t u$*. Then there exists a constant* $c > 0$ *such that*

$$\boxed{\begin{aligned}&\epsilon^{-1}\|\eta - g(u)\nabla_t\xi\| + \|\nabla_s\xi\| + \|\nabla_t\eta\| + \epsilon\|\nabla_s\eta\|\\ &\qquad \le c\left(\|\mathcal{D}_w^{\epsilon}{}^*(\xi,\eta)\|_{0,2,\epsilon} + \frac{1}{\sqrt{\epsilon}}\|(\xi,\eta)\|_{0,2,\epsilon}\right)\end{aligned}}$$

and likewise

$$\boxed{\begin{aligned}&\epsilon^{-1}\|\eta - g(u)\nabla_t\xi\| + \|\nabla_s\xi\| + \|\nabla_t\eta\| + \epsilon\|\nabla_s\eta\|\\ &\qquad \le c\left(\|\mathcal{D}_w^{\epsilon}{}^*(\xi,\eta)\|_{0,2,\epsilon} + \|\xi\| + \|\eta\|\right)\end{aligned}}$$

for $0 < \epsilon \le 1$ *and* $(\xi,\eta) \in C_0^\infty(\mathbb{R}\times S^1, u^*TM \oplus u^*T^*M)$*. The estimate continues to hold for* D_w^ϵ*.*

Proof. (**of Lemma 4.1.6**) We work in the orthonormal frame $\Phi = diag\,(\phi, \phi^{*-1})$ defined in (144). For simplicity we drop the vector notation here, e.g. we will write ξ instead of $\vec{\xi}$. Let $\zeta = (\xi,\eta)$ and $\mathcal{D}_\epsilon^*\zeta = \tilde{\zeta} = (\tilde{\xi},\tilde{\eta})$ so that

$$\tilde{\xi} = -\partial_s\xi - \nabla_t\eta - S\xi + \epsilon^2 B^*\eta \quad , \quad \tilde{\eta} = -\partial_s\eta + \epsilon^{-2}(\nabla_t\xi - \eta).$$

We compute with L^2-norms

$$\begin{aligned}\|D_\epsilon^*\zeta\|_{0,2,\epsilon}^2 &= \|\tilde{\xi}\|^2 + \epsilon^2\|\tilde{\eta}\|^2\\ &= \|-\partial_s\xi - \nabla_t\eta - S\xi + \epsilon^2B^*\eta\|^2 + \epsilon^2\|-\partial_s\eta + \epsilon^{-2}(\nabla_t\xi - \eta)\|^2\\ &= \|\partial_s\xi\|^2 + \|\nabla_t\eta - \epsilon^2B^*\eta + S\xi\|^2 + 2\langle\partial_s\xi, \nabla_t\eta - \epsilon^2B^*\eta + S\xi\rangle\\ &\quad + \epsilon^2\|\partial_s\eta\|^2 + \epsilon^{-2}\|\nabla_t\xi - \eta\|^2 - 2\langle\partial_s\eta, \nabla_t\xi - \eta\rangle.\end{aligned}$$

Now use the identities $\langle\partial_s\eta,\eta\rangle = 0$,

$$2\langle\partial_s\xi, S\xi\rangle = -\langle\xi, \partial_s(S\xi)\rangle + \langle S\partial_s\xi, \xi\rangle = -\langle\xi, (\partial_sS)\xi\rangle \ge -c_{\partial_sS}\|\xi\|^2,$$

$$\begin{aligned}-2\langle\partial_s\xi, \epsilon^2B^*\eta\rangle &= 2\langle\xi, \epsilon^2(\partial_sB^*)\eta\rangle + 2\langle\xi, \epsilon^2B^*\partial_s\eta\rangle\\ &\ge -\frac{1}{\epsilon}\|\xi\|^2 - \epsilon^5 c_{\partial_sB^*}{}^2\|\eta\|^2 - \frac{1}{\epsilon}\|\xi\|^2 - \epsilon^5 c_{B^*}\|\partial_s\eta\|^2\end{aligned}$$

and

$$\begin{aligned}\|\nabla_t\eta + S\xi - \epsilon^2B^*\eta\|^2 &\ge \frac{1}{2}\|\nabla_t\eta\|^2 - 7\left(c_S\|\xi\| + \epsilon^2c_{B^*}\|\eta\|\right)^2\\ &\ge \frac{1}{2}\|\nabla_t\eta\|^2 - 14c_S{}^2\|\xi\|^2 - 14\epsilon^4c_{B^*}{}^2\|\eta\|^2\end{aligned}$$

as well as the crucial estimate

$$\begin{aligned}2\langle\partial_s\xi, \nabla_t\eta\rangle - 2\langle\nabla_t\xi, \partial_s\eta\rangle &= 2\langle\xi, (\nabla_t\partial_s - \partial_s\nabla_t)\eta\rangle = -2\langle\xi, (\partial_sA)\eta\rangle\\ &\ge -\frac{1}{\epsilon}\|\xi\|^2 - 4\epsilon c_{\partial_sA}{}^2\|\eta\|^2\end{aligned}$$

with $\partial_s A = \phi^{-1} R(\partial_s u, \partial_t u)\phi$. Note that for the finiteness of the constants $c_B = \|B\|_\infty = \sup_{(s,t)\in\mathbb{R}\times S^1} \|B(s,t)\|$ we essentially used the boundary conditions for the cylinder u.

$$\begin{aligned}
\|D_\epsilon^* \zeta\|_{0,2,\epsilon}^2 &= \|\partial_s \xi\|^2 + \|\nabla_t \eta - \epsilon^2 B^* \eta + S\xi\|^2 + \epsilon^2 \|\partial_s \eta\|^2 + \epsilon^{-2} \|\nabla_t \xi - \eta\|^2 \\
&\quad - \langle \xi, (\partial_s S)\xi\rangle - 2\langle \partial_s \xi, \epsilon^2 B^* \eta\rangle - 2\langle \xi, (\partial_s A)\eta\rangle \\
&\geq \epsilon^{-2} \|\nabla_t \xi - \eta\|^2 + \|\partial_s \xi\|^2 + \frac{1}{2}\|\nabla_t \eta\|^2 \\
&\quad + (1 - \epsilon^5 c_{B^*})\epsilon^2 \|\partial_s \eta\|^2 - (14 {c_S}^2 \epsilon + c_{\partial_s S}\epsilon + 3)\frac{1}{\epsilon}\|\xi\|^2 \\
&\quad - (14 {c_{B^*}}^2 \epsilon^3 + {c_{\partial_s B^*}}^2 \epsilon^4 + 4 {c_{\partial_s A}}^2)\epsilon \|\eta\|^2
\end{aligned}$$

so for $\epsilon_0 > 0$ sufficiently small the result follows.

Using Young's inequality in a slightly different manner, we get

$$-2\langle \partial_s \xi, \epsilon^2 B^* \eta\rangle \geq -\epsilon^2 \left(\|\xi\|^2 + c_{\partial_s B^*}^2 \|\eta\|^2\right) - \epsilon^2 \left(\|\xi\|^2 + c_{B^*}^2 \|\partial_s \eta\|^2\right)$$

and

$$-2\langle \xi, (\partial_s A)\eta\rangle \geq -\|\xi\|^2 - c_{\partial_s A}^2 \|\eta\|^2$$

which lead to the alternative form of the estimate we are claiming for. The estimates for D_w^ϵ work similarly, they may be found in [**SW98**]. □

Proof. (**of Lemma 4.1.5**) Let $\xi' = (\mathbb{1} - \epsilon^\alpha \nabla_t \nabla_t)^{-1}\xi$ so that $\xi' - \epsilon^\alpha \nabla_t \nabla_t \xi' = \xi$. Take the L^2-inner product with ξ' and use the Cauchy-Schwarz as well as the Young-inequality to obtain

$$\|\xi'\|^2 + \epsilon^\alpha \|\nabla_t \xi'\|^2 = \langle \xi', \xi\rangle \leq \frac{1}{2}\|\xi'\|^2 + \frac{1}{2}\|\xi\|^2.$$

Hence $\|\xi'\|^2 + 2\|\epsilon^{\alpha/2}\nabla_t \xi'\|^2 \leq \|\xi\|^2$ and this implies the first two inequalities. To prove the third inequality write $\xi'' = (\mathbb{1} - \epsilon^\alpha \nabla_t \nabla_t)^{-1}\nabla_t \xi$ so that $\xi'' - \epsilon^\alpha \nabla_t \nabla_t \xi'' = \nabla_t \xi$. As above take the L^2-inner product with ξ'' to obtain

$$\|\xi''\|^2 + \epsilon^\alpha \|\nabla_t \xi''\|^2 = \langle \xi'', \nabla_t \xi\rangle = -\langle \nabla_t \xi'', \xi\rangle \leq \frac{\epsilon^\alpha}{2}\|\nabla_t \xi''\|^2 + \frac{1}{2\epsilon^\alpha}\|\xi\|^2.$$

Hence $\|\epsilon^\alpha \nabla_t \xi''\|^2 \leq \|\xi\|^2$ and this implies the third inequality. □

Proof. (**of Lemma 4.1.4**) Still working in an orthonormal frame we denote

$$\xi_0 = \pi_\epsilon \zeta = (\mathbb{1} - \epsilon^\alpha \nabla_t \nabla_t)^{-1}(\xi - \epsilon^2 \nabla_t \eta)$$

Then

$$\xi - \xi_0 = \epsilon^\alpha (\mathbb{1} - \epsilon^\alpha \nabla_t \nabla_t)^{-1} \nabla_t (\eta - \nabla_t \xi) + (\epsilon^2 - \epsilon^\alpha)(\mathbb{1} - \epsilon^\alpha \nabla_t \nabla_t)^{-1} \nabla_t \eta$$

and hence, by lemma 4.1.5

$$\|\xi - \xi_0\| \leq \epsilon^{\alpha/2} \|\eta - \nabla_t \xi\| + 2\epsilon^\alpha \|\nabla_t \eta\|$$

Similarly,

$$\eta - \nabla_t\xi_0 = (1\!\!1 - \epsilon^\alpha \nabla_t\nabla_t)^{-1}(\eta - \nabla_t\xi) - (\epsilon^\alpha - \epsilon^2)(1\!\!1 - \epsilon^\alpha\nabla_t\nabla_t)^{-1}\nabla_t\nabla_t\eta$$

and hence, again by lemma 4.1.5,

$$\epsilon\|\eta - \nabla_t\xi_0\| \le \epsilon\|\eta - \nabla_t\xi\| + 2\epsilon^{1+\alpha/2}\|\nabla_t\eta\|.$$

Take the sum of these two inequalities to obtain

$$\|\zeta - \iota\pi_\epsilon\zeta\|_{0,2,\epsilon} \le \|\xi - \xi_0\| + \epsilon\|\eta - \nabla_t\xi_0\| \le 2\epsilon^{\alpha/2}\|\eta - \nabla_t\xi\| + 4\epsilon^\alpha\|\nabla_t\eta\|$$

for $0 < \epsilon \le 1$ and $\alpha \le 2$. To prove the final inequality denote

$$\xi_0 = \pi_\epsilon\zeta = (1\!\!1 - \epsilon^\alpha\nabla_t\nabla_t)^{-1}(\xi - \epsilon^2\nabla_t\eta)$$

and use lemma 4.1.5 to get

$$\|\xi_0\| \le \|\xi\| + \epsilon^{2-\alpha/2}\|\eta\| \quad , \quad \epsilon\|\nabla_t\xi_0\| \le \epsilon^{1-\alpha/2}\|\xi\| + \epsilon^{3-\alpha}\|\eta\|.$$

Square these two inequalities, use (41) with $p = 2$ and take the sum to obtain

$$\begin{aligned}\|\iota\pi_\epsilon\zeta\|_{0,2,\epsilon}^2 &= \|\xi_0\|^2 + \epsilon^2\|\nabla_t\xi_0\|^2 \\ &\le (2 + 2\epsilon^{2-\alpha})\|\xi\|^2 + (2\epsilon^{2-\alpha} + 2\epsilon^{4-2\alpha})\epsilon^2\|\eta\|^2 \\ &\le 4\|\zeta\|_{0,2,\epsilon}^2\end{aligned}$$

for $0 < \epsilon \le 1$ and $\alpha \le 2$. □

Proof. (**of Lemma 4.1.3**) It is convenient to work in an orthonormal frame as specified in the proof of lemma 4.1.6 so that the operators $\mathcal{D}_u^{0\,*}$ and $\mathcal{D}_w^{\epsilon\,*}$ are given by (33) and (34), respectively. We also adapt the notation used there. As above denote

$$\xi_0 = \pi_\epsilon\zeta = (1\!\!1 - \epsilon^\alpha\nabla_t\nabla_t)^{-1}(\xi - \epsilon^\beta\nabla_t\eta)$$

where $\nabla_t = \partial_t + A$ denotes the covariant derivative in the local frame. Then

$$\begin{aligned}\mathcal{D}_0^*\pi_\epsilon\zeta &= -\partial_s\xi_0 - \nabla_t\nabla_t\xi_0 - S\xi_0 \\ &= [-\partial_s - S, (1\!\!1 - \epsilon^\alpha\nabla_t\nabla_t)^{-1}](\xi - \epsilon^\beta\nabla_t\eta) \\ &\quad + (1\!\!1 - \epsilon^\alpha\nabla_t\nabla_t)^{-1}(-\partial_s\xi + \epsilon^\beta\partial_s\nabla_t\eta - \nabla_t\nabla_t\xi \\ &\qquad + \epsilon^\beta\nabla_t\nabla_t\nabla_t\eta - S\xi + \epsilon^\beta S\nabla_t\eta).\end{aligned}$$

As in the proof of lemma 4.1.6, denote $\mathcal{D}_\epsilon^*\zeta = \tilde\zeta = (\tilde\xi, \tilde\eta)$ so that

$$\tilde\xi = -\partial_s\xi - \nabla_t\eta - S\xi + \epsilon^2B^*\eta \quad , \quad \tilde\eta = -\partial_s\eta + \epsilon^{-2}(\nabla_t\xi - \eta).$$

Then

$$\begin{aligned}\pi_\epsilon\mathcal{D}_\epsilon^*\zeta &= (1\!\!1 - \epsilon^\alpha\nabla_t\nabla_t)^{-1}(\tilde\xi - \epsilon^\beta\nabla_t\tilde\eta) \\ &= (1\!\!1 - \epsilon^\alpha\nabla_t\nabla_t)^{-1}(-\partial_s\xi - \nabla_t\eta - S\xi + \epsilon^2B^*\eta \\ &\qquad + \epsilon^\beta\nabla_t\partial_s\eta - \epsilon^{\beta-2}\nabla_t(\nabla_t\xi - \eta)).\end{aligned}$$

Taking the difference we find

$$\begin{aligned}
&\mathcal{D}_0^*\pi_\epsilon\zeta - \pi_\epsilon\mathcal{D}_\epsilon^*\zeta \\
&\qquad = [-\partial_s - S, (1\!\!1 - \epsilon^\alpha\nabla_t\nabla_t)^{-1}](\xi - \epsilon^\beta\nabla_t\eta) \\
&\qquad\quad + \epsilon^\beta(1\!\!1 - \epsilon^\alpha\nabla_t\nabla_t)^{-1}([\partial_s,\nabla_t]\eta + \nabla_t\nabla_t\nabla_t\eta + S\nabla_t\eta - \epsilon^{2-\beta}B^*\eta) \\
&\qquad\quad + (\epsilon^{\beta-2} - 1)(1\!\!1 - \epsilon^\alpha\nabla_t\nabla_t)^{-1}\nabla_t(\nabla_t\xi - \eta).
\end{aligned}$$

From now on we set $\beta = 2$ in order to eliminate the last term. Using $[a,b] = b[b^{-1},a]b$ in the first step, the commutator is given by

$$\begin{aligned}
&[-\partial_s - S, (1\!\!1 - \epsilon^\alpha\nabla_t\nabla_t)^{-1}] \\
&\qquad = (1\!\!1 - \epsilon^\alpha\nabla_t\nabla_t)^{-1}[1\!\!1 - \epsilon^\alpha\nabla_t\nabla_t, -\partial_s - S](1\!\!1 - \epsilon^\alpha\nabla_t\nabla_t)^{-1} \\
&\qquad = \epsilon^\alpha(1\!\!1 - \epsilon^\alpha\nabla_t\nabla_t)^{-1}(-\partial_s A + \nabla_t S)\nabla_t(1\!\!1 - \epsilon^\alpha\nabla_t\nabla_t)^{-1} \\
&\qquad\quad + \epsilon^\alpha(1\!\!1 - \epsilon^\alpha\nabla_t\nabla_t)^{-1}\nabla_t(-\partial_s A + \nabla_t S)(1\!\!1 - \epsilon^\alpha\nabla_t\nabla_t)^{-1}
\end{aligned}$$

where $\partial_s A = [\partial_s, \nabla_t]$. Hence

$$\begin{aligned}
&\mathcal{D}_0^*\pi_\epsilon\zeta - \pi_\epsilon\mathcal{D}_\epsilon^*\zeta \\
&= \epsilon^\alpha(1\!\!1 - \epsilon^\alpha\nabla_t\nabla_t)^{-1}(-\partial_s A + \nabla_t S)\nabla_t(1\!\!1 - \epsilon^\alpha\nabla_t\nabla_t)^{-1}(\xi - \epsilon^2\nabla_t\eta) \\
&\quad + \epsilon^\alpha(1\!\!1 - \epsilon^\alpha\nabla_t\nabla_t)^{-1}\nabla_t(-\partial_s A + \nabla_t S)(1\!\!1 - \epsilon^\alpha\nabla_t\nabla_t)^{-1}(\xi - \epsilon^2\nabla_t\eta) \\
&\quad + \epsilon^2(1\!\!1 - \epsilon^\alpha\nabla_t\nabla_t)^{-1}((\partial_s A)\eta + \nabla_t\nabla_t\nabla_t\eta + S\nabla_t\eta - B^*\eta).
\end{aligned} \tag{36}$$

Inspecting these expressions term by term and using lemma 4.1.5 as well as $\alpha \le 2$ we find

$$\|\mathcal{D}_0^*\pi_\epsilon\zeta - \pi_\epsilon\mathcal{D}_\epsilon^*\zeta\| \le c\epsilon^{2-\alpha}\|\nabla_t\eta\| + c\epsilon^{\alpha/2}\|\zeta\|_{0,2,\epsilon}.$$

The last claim now follows from lemma 4.1.6. The estimate for $\mathcal{D}_u^0\pi_\epsilon - \pi_\epsilon\mathcal{D}_w^\epsilon$ works similarly and may be found in [**SW98**]. It is carried out in the case $p > 2$ in the proof of lemma 4.2.2 below.
Note that the compactification of the cylinder u via the imposed boundary conditions x, y is a crucial point in our proof as it implied the finiteness of certain supremum norms we have used in between the lines, e.g. $\|B^*\|_\infty < \infty$. □

PROOF. (**of Proposition 4.1.2**) First note that the injectivity estimate

$$\|\xi_0\| + \|\nabla_t\nabla_t\xi_0\| \le c_0\|\mathcal{D}_u^{0\,*}\xi_0\|$$

for $\xi_0 \in C_0^\infty(\mathbb{R}\times S^1, u^*TM)$ implies

$$\|\nabla_t\xi_0\| \le \|\nabla_t\nabla_t\xi_0\| + \|\xi_0\| \le c_0\|\mathcal{D}_u^{0\,*}\xi_0\|. \tag{37}$$

This follows from partial integration and Young's inequality

$$\begin{aligned}\|\nabla_t\xi_0\|^2 &= \int_{-\infty}^{\infty}\int_0^1 -\langle\nabla_t\nabla_t\xi_0,\xi_0\rangle\, dtds\\ &\le \int_{-\infty}^{\infty}\int_0^1 \Big(\frac{|\nabla_t\nabla_t\xi_0|^2}{2}+\frac{|\xi_0|^2}{2}\Big)dtds.\end{aligned}$$

Let us restrict α to $[0,2]$ in order to apply lemmata 4.1.3-4.1.6, then

$$\begin{aligned}\|\xi\| &\le \|\xi-\pi_\epsilon\zeta\|+\|\pi_\epsilon\zeta\|\\ &\le \epsilon^{\alpha/2}\|\eta-g(u)\nabla_t\xi\|+2\epsilon^{\alpha}\|\nabla_t\eta\|+c_0\|\mathcal{D}_u^{0\,*}\pi_\epsilon\zeta\|\\ &\le c_1\epsilon^{\alpha}\|\mathcal{D}_w^{\epsilon\,*}\zeta\|_{0,2,\epsilon}+c_1\epsilon^{\alpha-1/2}\|\zeta\|_{0,2,\epsilon}\\ &\quad+c_0\|\mathcal{D}_u^{0\,*}\pi_\epsilon\zeta-\pi_\epsilon\mathcal{D}_w^{\epsilon\,*}\zeta\|+c_0\|\pi_\epsilon\mathcal{D}_w^{\epsilon\,*}\zeta\|\\ &\le c_2(\epsilon^{\alpha}+\epsilon^{2-\alpha})\|\mathcal{D}_w^{\epsilon\,*}\zeta\|_{0,2,\epsilon}+c_0\|\pi_\epsilon\mathcal{D}_w^{\epsilon\,*}\zeta\|\\ &\quad+c_2(\epsilon^{\alpha-1/2}+\epsilon^{3/2-\alpha}+\epsilon^{\alpha/2})\|\zeta\|_{0,2,\epsilon},\end{aligned}$$

where we used lemma 4.1.4 and the injectivity assumption in the second and lemma 4.1.6 for $\mathcal{D}_w^{\epsilon\,*}$ in the third inequality. The last one follows from lemma 4.1.5. We observe that the best estimate is obtained by setting $\alpha=1$. Now for $\epsilon_0>0$ sufficiently small incorporate the ξ-part of ζ into the LHS and obtain

$$\|\xi\|\le c\left(\epsilon\|\mathcal{D}_w^{\epsilon\,*}\zeta\|_{0,2,\epsilon}+\|\pi_\epsilon\mathcal{D}_w^{\epsilon\,*}\zeta\|+\epsilon^{3/2}\|\eta\|\right). \tag{38}$$

Repeating the steps of the ξ-estimate above leads to

$$\begin{aligned}\|\eta\| &\le \|\eta-g(u)\nabla_t\pi_\epsilon\zeta\|+\|\nabla_t\pi_\epsilon\zeta\|\\ &\le \|\eta-g(u)\nabla_t\xi\|+2\epsilon^{\alpha/2}\|\nabla_t\eta\|+c_0\|\mathcal{D}_u^{0\,*}\pi_\epsilon\zeta\|\\ &\le c_1\epsilon^{\alpha/2}\|\mathcal{D}_w^{\epsilon\,*}\zeta\|_{0,2,\epsilon}+c_1\epsilon^{\alpha/2-1/2}\|\zeta\|_{0,2,\epsilon}\\ &\quad+c_0\|\mathcal{D}_u^{0\,*}\pi_\epsilon\zeta-\pi_\epsilon\mathcal{D}_w^{\epsilon\,*}\zeta\|+c_0\|\pi_\epsilon\mathcal{D}_w^{\epsilon\,*}\zeta\|\\ &\le c_2(\epsilon^{\alpha/2}+\epsilon^{2-\alpha})\|\mathcal{D}_w^{\epsilon\,*}\zeta\|_{0,2,\epsilon}+c_0\|\pi_\epsilon\mathcal{D}_w^{\epsilon\,*}\zeta\|\\ &\quad+c_2(\epsilon^{\alpha/2-1/2}+\epsilon^{3/2-\alpha}+\epsilon^{\alpha/2})\|\zeta\|_{0,2,\epsilon}.\end{aligned}$$

For $\epsilon_0>0$ sufficiently small incorporate the η-part of ζ into the LHS and obtain with $\alpha=1$

$$\|\eta\|\le c\left(\epsilon^{1/2}\|\mathcal{D}_w^{\epsilon\,*}\zeta\|_{0,2,\epsilon}+\|\pi_\epsilon\mathcal{D}_w^{\epsilon\,*}\zeta\|+\|\xi\|\right). \tag{39}$$

Inserting (39) into (38) establishes our first claim for $\epsilon_0>0$ sufficiently small and, similarly, inserting (38) into (39) the second one. Together they imply the third claim and lemma 4.1.4 then gives the fourth one. □

We would like to state two inequalities which we have used occasionally during the above proofs.

LEMMA 4.1.7. (**Young-inequality**) *Let* $a, b \geq 0$ *and* $1 < p, q < \infty$ *such that* $\frac{1}{p} + \frac{1}{q} = 1$, *then*

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

PROOF. For one of a, b equal to zero the inequality holds trivially. So assume $a, b > 0$. Taking the logarithm of both sides and using its concavity we get

$$\log(ab) = \log a + \log b = \frac{1}{p}\log a^p + \frac{1}{q}\log b^q \leq \log\left(\frac{a^p}{p} + \frac{b^q}{q}\right).$$

The result follows as the logarithm is strictly increasing. □

Another useful inequality is

$$(a+b)^2 \leq 2a^2 + 2b^2 \quad , \forall a, b \in \mathbb{R} \tag{40}$$

which follows from the binomial identities $(a+b)^2 = a^2 + 2ab + b^2$ and $0 \leq (a-b)^2 = a^2 - 2ab + b^2$. More generally, for $a, b \in \mathbb{R}$ and $1 \leq p < \infty$

$$|a+b|^p \leq (|a| + |b|)^p \leq 2^{p-1}\left(|a|^p + |b|^p\right) \tag{41}$$

which follows from convexity of t^p in $(0, \infty)$ (cf. [**Ru87**] proof of thm. 3.5). Note that for $p \geq 1$ this implies

$$\begin{aligned} \|(\xi,\eta)\|_{0,p,\epsilon} &= \Big(\underbrace{\|\xi\|_{L^p}^p + \epsilon^p\|\eta\|_{L^p}^p}_{\leq(\|\xi\|_{L^p}+\epsilon\|\eta\|_{L^p})^p}\Big)^{1/p} \leq \|\xi\|_{L^p} + \epsilon\|\eta\|_{L^p} \\ \|\xi\|_{L^p} + \epsilon\|\eta\|_{L^p} &\leq 2^{(p-1)/p}\left(\|\xi\|_{L^p}^p + \epsilon^p\|\eta\|_{L^p}^p\right)^{1/p} \leq 2\|(\xi,\eta)\|_{0,p,\epsilon}. \end{aligned} \tag{42}$$

4.2. Nonstandard estimates for $p > 2$

New ideas are required to prove lemma 4.2.4 and lemma 4.2.5. Once these are established the other results follow quite similar as in the case $p = 2$ in the last section.

Throughout $\|\cdot\|_p$ denotes the $L^p(\mathbb{R}\times S^1, u^*TM)$-norm unless mentioned otherwise. As it turns out in the proof of the main result of this section, proposition 4.2.1, the critical exponent κ_p has to be in the range $(0,1)$. Moreover, the projection π_ϵ is given by (35) with constants $\alpha = 1$ and $\beta = 2$. This choice of β is forced by lemma 4.2.2 below, whereas the choice of α optimizes the ξ-estimate in proposition 4.2.1. At this point also the crucial condition $\kappa_p < 1$ arises in order to incorporate certain terms into the left hand side. The following result says that surjectivity of $\mathcal{D}_u^0$ implies surjectivity of $\mathcal{D}_w^\epsilon$ and leads to a uniform bound for its right inverse in section 4.4.

Proposition 4.2.1. *Let $u \in \mathcal{P}_{x,y}(\mathbb{R}\times S^1, M)$, where x, y are smooth loops in M, define $w = g(u)\partial_t u$, $\alpha = 1$, $\beta = 2$ in π_ϵ and assume $\kappa(p) \in (0,1)$. Then for every $p > 2$ and $c_0 > 0$ there exist constants $c_p > 0$, $\epsilon_0 = \epsilon_0(p) > 0$ such that the following holds. If the injectivity assumption*

$$\|\xi\|_p + \|\nabla_t\nabla_t\xi\|_p \le c_0\|{\mathcal{D}_u^0}^*\xi\|_p$$

holds for all $\xi \in C_0^\infty(\mathbb{R}\times S^1, u^*TM)$, then

$$\boxed{\begin{aligned}\|\xi\|_p &\le c_p\Big(\epsilon\|{\mathcal{D}_w^\epsilon}^*(\xi,\eta)\|_{0,p,\epsilon} + \|\pi_\epsilon{\mathcal{D}_w^\epsilon}^*(\xi,\eta)\|_p\Big)\\ \|\eta\|_p &\le c_p\Big(\epsilon^{1/2}\|{\mathcal{D}_w^\epsilon}^*(\xi,\eta)\|_{0,p,\epsilon} + \epsilon^{\min\{1/2-\kappa_p,0\}}\|\pi_\epsilon{\mathcal{D}_w^\epsilon}^*(\xi,\eta)\|_p\Big)\end{aligned}}$$

and therefore

$$\boxed{\begin{aligned}\|\zeta\|_{0,p,\epsilon} &\le c_p\Big(\epsilon\|{\mathcal{D}_w^\epsilon}^*\zeta\|_{0,p,\epsilon} + \|\pi_\epsilon{\mathcal{D}_w^\epsilon}^*\zeta\|_p\Big)\\ \|\zeta\|_{0,p,\epsilon} &\le c_p\|{\mathcal{D}_w^\epsilon}^*\zeta\|_{0,p,\epsilon}\end{aligned}}$$

for $\epsilon \in (0,\epsilon_0)$ and $\zeta = (\xi,\eta) \in C_0^\infty(\mathbb{R}\times S^1, u^*TM\oplus u^*T^*M)$.

Lemma 4.2.2. *Let $u \in \mathcal{P}_{x,y}(\mathbb{R}\times S^1, M)$, x, y smooth loops in M, $\beta = 2$ and define $w = g(u)\partial_t u$. Then for any $p > 2$ there exists a constant $c_p > 0$ such that*

$$\boxed{\begin{aligned}\|{\mathcal{D}_u^0}^*\pi_\epsilon\zeta - \pi_\epsilon{\mathcal{D}_w^\epsilon}^*\zeta\|_p &\le c_p\epsilon^{\alpha_p/2}\|\zeta\|_{0,p,\epsilon} + c_p\epsilon^{2-\alpha_p}\|\nabla_t\eta\|_p\\ &\le c_p\epsilon^{2-\alpha_p}\|{\mathcal{D}_w^\epsilon}^{(*)}\zeta\|_{0,p,\epsilon} + c_p\epsilon^{\min\{2-\alpha_p-\kappa_p,\alpha_p/2\}}\|\zeta\|_{0,p,\epsilon}\end{aligned}}$$

for $0 < \epsilon \le 1$, $0 \le \alpha_p \le 2$ and $\zeta = (\xi,\eta) \in C_0^\infty(\mathbb{R}\times S^1, u^*TM\oplus u^*T^*M)$. The same estimate holds for $\mathcal{D}_u^0\pi_\epsilon - \pi_\epsilon\mathcal{D}_w^\epsilon$.

In the proof of lemma 4.2.2 it turns out that $\beta = 2$ is a natural choice. Moreover, setting $\alpha_p/2 = 2 - \alpha_p - \kappa_p$, or equivalently $\alpha_p = 4/3 - 2\kappa_p/3$ optimizes the last estimate.

LEMMA 4.2.3. *Let $u \in C^\infty(\mathbb{R} \times S^1, M)$, $\beta = 2$ and $p > 2$, then*

$$\boxed{\begin{aligned} \|\xi - \pi_\epsilon\zeta\|_p &\leq 2p\epsilon^{\alpha_p/2}\|\eta - g(u)\nabla_t\xi\|_p + 2\epsilon^{\alpha_p}\|\nabla_t\eta\|_p \\ \|\eta - g(u)\nabla_t\pi_\epsilon\zeta\|_p &\leq \|\eta - g(u)\nabla_t\xi\|_p + 4p\epsilon^{\alpha_p/2}\|\nabla_t\eta\|_p \\ \|\zeta - \iota\pi_\epsilon\zeta\|_{0,p,\epsilon} &\leq 4p\epsilon^{\alpha_p/2}\|\eta - g(u)\nabla_t\xi\|_p + 8p\epsilon^{\alpha_p}\|\nabla_t\eta\|_p \\ \|\pi_\epsilon\zeta\|_p &\leq \|\iota\pi_\epsilon\zeta\|_{0,p,\epsilon} \leq 6p\|\zeta\|_{0,p,\epsilon} \end{aligned}}$$

for $0 < \epsilon \leq 1$, $0 \leq \alpha_p \leq 2$ and $\zeta = (\xi, \eta) \in C_0^\infty(\mathbb{R} \times S^1, u^*TM \oplus u^*T^*M)$.

LEMMA 4.2.4. *Let $u \in C^\infty(\mathbb{R} \times S^1, M)$ and $p > 2$, then*

$$\boxed{\begin{aligned} \|(1\!\!1 - \epsilon^\alpha\nabla_t\nabla_t)^{-1}\xi\|_p &\leq \|\xi\|_p \\ \|(1\!\!1 - \epsilon^\alpha\nabla_t\nabla_t)^{-1}\epsilon^{\alpha/2}\nabla_t\xi\|_p &\leq 2p\|\xi\|_p \\ \|(1\!\!1 - \epsilon^\alpha\nabla_t\nabla_t)^{-1}\epsilon^{\alpha}\nabla_t\nabla_t\xi\|_p &\leq 2\|\xi\|_p \end{aligned}}$$

for $\epsilon > 0$, $\alpha \in \mathbb{R}$ and $\xi \in C_0^\infty(\mathbb{R} \times S^1, u^*TM)$.

LEMMA 4.2.5. *Let $u \in \mathcal{P}_{x,y}(\mathbb{R} \times S^1, M)$, $x, y \in C^\infty(S^1, M)$, and define $w = g(u)\partial_t u$. There exists a continuous function $\kappa : [2, \infty) \to \mathbb{R}$ such that $\kappa(2) = 1/2$ and the following holds. For any $p > 2$ there exists a constant $c_p > 0$ such that*

$$\boxed{\begin{aligned} &\epsilon^{-1}\|\eta - g(u)\nabla_t\xi\|_p + \|\nabla_s\xi\|_p + \|\nabla_t\eta\|_p + \epsilon\|\nabla_s\eta\|_p \\ &\qquad \leq c_p\left(\|\mathcal{D}_w^{\epsilon\,*}(\xi,\eta)\|_{0,p,\epsilon} + \epsilon^{-\kappa_p}\|(\xi,\eta)\|_{0,p,\epsilon}\right) \end{aligned}}$$

for $\epsilon \in (0, 1]$ and $(\xi, \eta) \in C_0^\infty(\mathbb{R} \times S^1, u^*TM \oplus u^*T^*M)$. The estimate also holds for D_w^ϵ.

PROOF. (**of Lemma 4.2.5 assuming Conjecture 1.0.6**) We work in the orthonormal frame $\Phi = diag\,(\phi, \phi^{*-1})$ defined in (144). For simplicity we drop the vector notation here, e.g. we will write ξ instead of $\vec{\xi}$. Let $\zeta = (\xi, \eta)$ and $\mathcal{D}_\epsilon^*\zeta = \tilde{\zeta} = (\tilde{\xi}, \tilde{\eta})$ so that

$$\tilde{\xi} = -\partial_s\xi - \nabla_t\eta - S\xi + \epsilon^2 B^*\eta \quad , \quad \tilde{\eta} = -\partial_s\eta + \epsilon^{-2}(\nabla_t\xi - \eta).$$

Use conjecture 1.0.6, remark 1.0.7 1) as well as addition of 0 to obtain

$$\begin{aligned} &\epsilon^{-1}\|\partial_t\xi + A\xi - \eta\|_p + \|\partial_t\eta + A\eta\|_p + \|\partial_s\xi\|_p + \epsilon\|\partial_s\eta\|_p \\ &\quad \leq c_p\Big(\| - \partial_s\xi - \partial_t\eta - A\eta - S\xi + \epsilon^2 B^*\eta\|_p + \|S\xi\|_p + \epsilon^2\|B^*\eta\|_p \\ &\qquad + \epsilon\| - \partial_s\eta + \epsilon^{-2}(\partial_t\xi + A\xi - \eta)\|_p + \epsilon^{-\kappa(p)}\|\xi\|_p + \epsilon^{1-\kappa(p)}\|\eta\|_p\Big) \\ &\quad \leq 2c_p c_S c_{B^*}\left(\|\mathcal{D}_\epsilon^*(\xi,\eta)\|_{0,p,\epsilon} + \epsilon^{-\kappa(p)}\|(\xi,\eta)\|_{0,p,\epsilon}\right) \end{aligned}$$

where $c_S = \|S\|_\infty$, $C_{B^*} = \|B^*\|_\infty$ and $\|\cdot\|_p = \|\cdot\|_{L^p(\mathbb{R}\times S^1,\mathbb{R}^n)}$. The estimate for D_w^ϵ works similarly. □

The next lemma is independent of the others. It is extremely useful in carrying out the Newton iteration to prove the existence theorem 1.0.4 as well as in the proof of the uniqueness theorem 1.0.5. We will prove it first. Note that for the (∞,ϵ)-norm of (ξ,η) a choice of $\beta_1 = 1$ and $\beta_2 = 1/2$ is natural (cf. proof), whereas to estimate ξ only, other choices can be useful. For instance $\beta_1 = 1/2, \beta_2 = 1$ is the right choice to prove the uniqueness theorem 1.0.5.

LEMMA 4.2.6. *Let $u \in \mathcal{P}_{x,y}(\mathbb{R}\times S^1, M)$, where x, y are smooth loops in M, then for any $p > 2$ there exists a constant $c_p > 0$ such that*

$$\boxed{\begin{aligned} \|(\xi,\eta)\|_{\infty,\epsilon} &\le c_p\, \epsilon^{-3/p}\|(\xi,\eta)\|_{1,p,\epsilon} \\ \|\xi\|_\infty &\le c_p\, \epsilon^{-\frac{\beta_1+\beta_2}{p}}\left(\|\xi\|_p + \epsilon^{\beta_1}\|\nabla_t\xi\|_p + \epsilon^{\beta_2}\|\nabla_s\xi\|_p\right)\end{aligned}}$$

for $\epsilon \in (0,1], \beta_1, \beta_2 > 0$ and $(\xi,\eta) \in C_0^\infty(\mathbb{R}\times S^1, u^*TM \oplus u^*T^*M)$.

PROOF. Let $Z_\epsilon = \mathbb{R}\times S^1_\epsilon = \mathbb{R}\times([0,\epsilon^{-1}]/\{0,\epsilon^{-1}\})$. Rescaling (ξ,η) in the usual way (cf. proof of theorem 4.3.2) leads to

$$\|(\tilde\xi,\tilde\eta)\|_{W^{1,p}(Z_\epsilon)} = \epsilon^{-3/p}\,\|(\xi,\eta)\|_{1,p,\epsilon}$$

and

$$\begin{aligned}\|(\xi,\eta)\|_{\infty,\epsilon} &:= \sup_{(s,t)\in\mathbb{R}\times S^1}|\xi(s,t)| + \epsilon \sup_{(s,t)\in\mathbb{R}\times S^1}|\eta(s,t)| \\ &= \sup_{(\tilde s,\tilde t)\in\mathbb{R}\times S^1_\epsilon}|\tilde\xi(\tilde s,\tilde t)| + \epsilon \sup_{(\tilde s,\tilde t)\in\mathbb{R}\times S^1_\epsilon}|\epsilon^{-1}\tilde\eta(\tilde s,\tilde t)| = \|(\tilde\xi,\tilde\eta)\|_{L^\infty(Z_\epsilon)}.\end{aligned}$$

These two identities will imply the result once we have shown that there exists a constant $c_p > 0$ such that

$$\|(\tilde\xi,\tilde\eta)\|_{L^\infty(Z_\epsilon)} \le c_p\,\|(\tilde\xi,\tilde\eta)\|_{W^{1,p}(Z_\epsilon)} \tag{43}$$

for $\epsilon\in(0,1]$ and $(\tilde\xi,\tilde\eta)\in C_0^\infty(\mathbb{R}\times S^1_\epsilon, \tilde u^*TM \oplus \tilde u^*T^*M)$. Note that $\tilde u(\tilde s,\tilde t) = u(\epsilon^2\tilde s,\epsilon\tilde t)$. This inequality is a consequence of the standard Sobolev estimate for $p > 2$ (cf. [**MS94**] Theorem B.1.4)

$$\|u\|_{L^\infty(\Omega)} \le c_p(\Omega)\,\|u\|_{W^{1,p}(\Omega)}$$

which holds for all $u \in C^\infty(\bar\Omega,\mathbb{R})$, where $\Omega\subset\mathbb{R}^2$ is a bounded domain with smooth boundary. This is shown as follows: Cover Z_ϵ by translating a bounded subset $\Omega\subset\mathbb{R}^2$ with smooth boundary. Denote the (countable) cover by $\{U_i\}_{i\in\mathbb{Z}}$, then

$$\begin{aligned}\|(\tilde\xi,\tilde\eta)\|_{L^\infty(Z_\epsilon)} &\le \sup_{i\in Z}\|(\tilde\xi,\tilde\eta)\|_{L^\infty(U_i)} \\ &\le c_p(\Omega)\sup_{i\in Z}\|(\tilde\xi,\tilde\eta)\|_{W^{1,p}(U_i)} \le c_p(\Omega)\|(\tilde\xi,\tilde\eta)\|_{W^{1,p}(Z_\epsilon)}.\end{aligned}$$

To prove the second assertion pick $\beta_1, \beta_2 > 0$, set

$$Z_{\beta_1} = \mathbb{R} \times ([0, \epsilon^{-\beta_1}]/\{0, \epsilon^{-\beta_1}\})$$

and $\tilde{\xi}(\tilde{s}, \tilde{t}) = \xi(\epsilon^{\beta_2}\tilde{s}, \epsilon^{\beta_1}\tilde{t})$. As above we get

$$\begin{aligned}\|\xi\|_\infty = \|\tilde{\xi}\|_{L^\infty(Z_{\beta_1})} &\le c_p \|\tilde{\xi}\|_{W^{1,p}(Z_{\beta_1})}\\ &\le c_p \epsilon^{-\frac{\beta_1+\beta_2}{p}} \left(\|\xi\|_p + \epsilon^{\beta_1}\|\nabla_t \xi\|_p + \epsilon^{\beta_2}\|\nabla_s \xi\|_p \right).\end{aligned}$$

□

PROOF. (**of Lemma 4.2.4**) The Lemma is proved in three steps. As the s-variable is irrelevant for the estimates to be shown, we restrict in steps 1 and 2 to the 1-dimensional case of the t-variable. Step 2 is a rescaling argument and integrating its result over $s \in \mathbb{R}$ then proves the Lemma.

Step 1 *Let $S_\epsilon = [0, \epsilon^{-1}]/\{0, \epsilon^{-1}\}$ and $\gamma \in C^\infty(S_\epsilon, M)$, then*

$$\begin{aligned}\|(\mathbb{1} - \nabla_{\tilde{t}}\nabla_{\tilde{t}})^{-1}\tilde{\xi}\|_{L^p(S_\epsilon,\gamma^*TM)} &\le \|\tilde{\xi}\|_{L^p(S_\epsilon,\gamma^*TM)}\\ \|(\mathbb{1} - \nabla_{\tilde{t}}\nabla_{\tilde{t}})^{-1}\nabla_{\tilde{t}}\tilde{\xi}\|_{L^p(S_\epsilon,\gamma^*TM)} &\le 2p\|\tilde{\xi}\|_{L^p(S_\epsilon,\gamma^*TM)}\\ \|(\mathbb{1} - \nabla_{\tilde{t}}\nabla_{\tilde{t}})^{-1}\nabla_{\tilde{t}}\nabla_{\tilde{t}}\tilde{\xi}\|_{L^p(S_\epsilon,\gamma^*TM)} &\le 2\|\tilde{\xi}\|_{L^p(S_\epsilon,\gamma^*TM)}\end{aligned}$$

for any $\epsilon > 0$ and any $\tilde{\xi} \in C^\infty(S_\epsilon, \gamma^*TM)$.

PROOF OF STEP 1 Pick $\tilde{\xi} \in C^\infty(S_\epsilon, \gamma^*TM)$ and set $\tilde{\eta} = (\mathbb{1} - \nabla_{\tilde{t}}\nabla_{\tilde{t}})^{-1}\tilde{\xi}$, i.e. $\tilde{\xi} = \tilde{\eta} - \nabla_{\tilde{t}}\nabla_{\tilde{t}}\tilde{\eta}$. This is well-defined as the bounded linear operator

$$\mathbb{1} - \nabla_{\tilde{t}}\nabla_{\tilde{t}} \ : \ W^{2,p}(S_\epsilon, \gamma^*TM) \to L^p(S_\epsilon, \gamma^*TM)$$

is bijective by elliptic regularity and hence admits a bounded inverse by the open mapping theorem. To be shown are inequalities of the form $\|\tilde{\eta}\|_p \le \|\tilde{\xi}\|_p$, etc. We use the short notation $\|\cdot\|_p$ for the $L^p(S_\epsilon, \gamma^*TM)$-norm, whereas $|\tilde{\eta}|$ denotes the function $\langle \tilde{\eta}(t), \tilde{\eta}(t)\rangle^{1/2}$. To start with consider

$$\begin{aligned}\frac{d^2}{d\tilde{t}^2}|\tilde{\eta}|^p &= \frac{d}{d\tilde{t}}\left(p\langle\tilde{\eta},\tilde{\eta}\rangle^{\frac{p}{2}-1}\langle\nabla_{\tilde{t}}\tilde{\eta},\tilde{\eta}\rangle\right)\\ &= \underbrace{p(p-2)|\tilde{\eta}|^{p-4}\langle\nabla_{\tilde{t}}\tilde{\eta},\tilde{\eta}\rangle^2}_{\ge 0} + \underbrace{p|\tilde{\eta}|^{p-2}}_{\ge 0}\Big(\langle\underbrace{\nabla_{\tilde{t}}\nabla_{\tilde{t}}\tilde{\eta}}_{=\tilde{\eta}-\tilde{\xi}},\tilde{\eta}\rangle + \underbrace{|\nabla_{\tilde{t}}\tilde{\eta}|^2}_{\ge 0}\Big)\\ &\ge p|\tilde{\eta}|^p - p|\tilde{\eta}|^{p-2}\langle\tilde{\xi},\tilde{\eta}\rangle \ge p|\tilde{\eta}|^p - p \underbrace{|\tilde{\eta}|^{p-1}}_{=a;p/(p-1)} \underbrace{|\tilde{\xi}|}_{=b;p}\\ &\ge p|\tilde{\eta}|^p - p\left(\frac{|\tilde{\eta}|^p}{p/(p-1)} + \frac{|\tilde{\xi}|^p}{p}\right)\\ &\ge p|\tilde{\eta}|^p - (p-1)|\tilde{\eta}|^p - |\tilde{\xi}|^p = |\tilde{\eta}|^p - |\tilde{\xi}|^p\end{aligned}$$

where in the second inequality we applied Young's inequality. Integrate the result over $\tilde{t} \in [0, \epsilon^{-1}]$ and use the periodicity of the LHS to get $\|\tilde{\eta}\|_p \le \|\tilde{\xi}\|_p$.

The essential tool in the following calculation is again Young's inequality lemma 4.1.7 ; a and b are indicated before each estimate (their exponents follow the semicolons).

$$\begin{aligned}
&\frac{d}{d\tilde t}\Big(\langle\tilde\eta,\nabla_{\tilde t}\tilde\eta\rangle\,|\nabla_{\tilde t}\tilde\eta|^{p-2}\Big)\\
&= |\nabla_{\tilde t}\tilde\eta|^p + \langle\tilde\eta,\underbrace{\nabla_{\tilde t}\nabla_{\tilde t}\tilde\eta}_{=\tilde\eta-\tilde\xi}\rangle|\nabla_{\tilde t}\tilde\eta|^{p-2} + (p-2)\langle\tilde\eta,\nabla_{\tilde t}\tilde\eta\rangle|\nabla_{\tilde t}\tilde\eta|^{p-4}\langle\nabla_{\tilde t}\tilde\eta,\underbrace{\nabla_{\tilde t}\nabla_{\tilde t}\tilde\eta}_{=\tilde\eta-\tilde\xi}\rangle\\
&\ge |\nabla_{\tilde t}\tilde\eta|^p + \frac{|\tilde\eta|^2}{2}\,|\nabla_{\tilde t}\tilde\eta|^{p-2} - \underbrace{|\tilde\xi|^2}_{=a;p/2}\ \underbrace{2^{-1}|\nabla_{\tilde t}\tilde\eta|^{p-2}}_{=b;p/(p-2)}\\
&\quad + \underbrace{(p-2)\langle\tilde\eta,\nabla_{\tilde t}\tilde\eta\rangle^2|\nabla_{\tilde t}\tilde\eta|^{p-4}}_{\ge 0} - |\nabla_{\tilde t}\tilde\eta|^{p-4}\underbrace{|\tilde\eta|\cdot|\nabla_{\tilde t}\tilde\eta|}_{=a;2}\ \underbrace{(p-2)|\tilde\xi|\cdot|\nabla_{\tilde t}\tilde\eta|}_{=b;2}\\
&\ge |\nabla_{\tilde t}\tilde\eta|^p + \frac{|\tilde\eta|^2}{2}\,|\nabla_{\tilde t}\tilde\eta|^{p-2} - \tfrac{2}{p}|\tilde\xi|^p - \tfrac{p-2}{p}\,2^{\frac{-p}{p-2}}|\nabla_{\tilde t}\tilde\eta|^p\\
&\quad - |\nabla_{\tilde t}\tilde\eta|^{p-4}\frac{|\tilde\eta|^2}{2}|\nabla_{\tilde t}\tilde\eta|^{p-2} - (p-2)^2\ \underbrace{|\tilde\xi|^2}_{=a;p/2}\ \underbrace{2^{-1}|\nabla_{\tilde t}\tilde\eta|^{p-2}}_{=b;p/(p-2)}\\
&\ge \Big(1 - \tfrac{p-2}{p}\ \underbrace{2^{\frac{-p}{p-2}}}_{<1/2} - \tfrac{p-2}{p}\ \underbrace{2^{\frac{-p}{p-2}}}_{<1/2}\Big)|\nabla_{\tilde t}\tilde\eta|^p - \Big(\tfrac{2}{p} + \tfrac{2}{p}(p-2)^2\Big)\,|\tilde\xi|^p\\
&\ge \tfrac{2}{p}|\nabla_{\tilde t}\tilde\eta|^p - \tfrac{2}{p}\left(1+(p-2)^2\right)|\tilde\xi|^p.
\end{aligned}$$

Integrate the result over $\tilde t \in [0,\epsilon^{-1}]$ and use the periodicity of the LHS to obtain

$$\|\nabla_{\tilde t}\tilde\eta\|_p{}^p \le \left(1+(p-2)^2\right)\|\tilde\xi\|_p{}^p \le 2p^2\|\tilde\xi\|_p{}^p \le 2^p p^p\|\tilde\xi\|_p{}^p.$$

The last one is easy as we may use the first inequality

$$\|\nabla_{\tilde t}\nabla_{\tilde t}\tilde\eta\|_p = \|\tilde\eta-\tilde\xi\|_p \le \|\tilde\eta\|_p + \|\tilde\xi\|_p \le 2\|\tilde\eta\|_p.$$

Step 2 *Let $\gamma \in C^\infty(S^1,\gamma^*TM)$, then*

$$\begin{aligned}
\|(\mathbb{1}-\epsilon^\alpha\nabla_t\nabla_t)^{-1}\xi\|_{L^p(S^1,\gamma^*TM)} &\le \|\xi\|_{L^p(S^1,\gamma^*TM)}\\
\|(\mathbb{1}-\epsilon^\alpha\nabla_t\nabla_t)^{-1}\epsilon^{\alpha/2}\nabla_t\xi\|_{L^p(S^1,\gamma^*TM)} &\le 2p\|\xi\|_{L^p(S^1,\gamma^*TM)}\\
\|(\mathbb{1}-\epsilon^\alpha\nabla_t\nabla_t)^{-1}\epsilon^\alpha\nabla_t\nabla_t\xi\|_{L^p(S^1,\gamma^*TM)} &\le 2\|\xi\|_{L^p(S^1,\gamma^*TM)}
\end{aligned}$$

for all $\epsilon > 0$, $\alpha > 0$ and $\xi \in C^\infty(S^1,\gamma^*TM)$.

Proof of Step 2 Rescale $\xi \in C^\infty(S^1,\gamma^*TM)$ by

$$\tilde\xi(\tilde t) = \xi(\epsilon^{\alpha/2}\tilde t)$$

where $\tilde t \in [0,\epsilon^{-\alpha/2}]/\{0,\epsilon^{-\alpha/2}\} = S_\epsilon$, then

$$\nabla_{\tilde t}\tilde\xi(\tilde t) = \epsilon^{\alpha/2}\nabla_t\xi(\epsilon^{\alpha/2}\tilde t) \quad , \quad \nabla_{\tilde t}\nabla_{\tilde t}\tilde\xi(\tilde t) = \epsilon^\alpha\nabla_t\nabla_t\xi(\epsilon^{\alpha/2}\tilde t).$$

Use this and Step 1 to get

$$\begin{aligned}
\epsilon^{-\alpha/2}\|\xi\|^p_{L^p(S^1,\gamma^*TM)} &= \int_0^1 |\xi(t)|^p \frac{dt}{\epsilon^{\alpha/2}} = \int_0^{\epsilon^{-\alpha/2}} |\tilde{\xi}(\tilde{t})|^p d\tilde{t} \\
&= \|\tilde{\xi}\|^p_{L^p(S_\epsilon,\tilde{\gamma}^*TM)} \geq \|(\mathbb{1} - \nabla_{\tilde{t}}\nabla_{\tilde{t}})^{-1}\tilde{\xi}\|^p_{L^p(S_\epsilon,\tilde{\gamma}^*TM)} \\
&= \int_0^{\epsilon^{-\alpha/2}} |(\mathbb{1} - \nabla_{\tilde{t}}\nabla_{\tilde{t}})^{-1}\tilde{\xi}(\tilde{t})|^p d\tilde{t} \\
&= \int_0^1 |(\mathbb{1} - \epsilon^\alpha\nabla_t\nabla_t)^{-1}\xi(t)|^p \frac{dt}{\epsilon^{\alpha/2}} \\
&= \epsilon^{-\alpha/2}\|(\mathbb{1} - \epsilon^\alpha\nabla_t\nabla_t)^{-1}\xi\|^p_{L^p(S^1,\gamma^*TM)}.
\end{aligned}$$

The other two estimates work similarly.

Step 3 The Lemma follows by applying Step 2 pointwise to $\xi(s,\cdot) \in C^\infty(S^1, u(s,\cdot)^*TM)$ and integrating over $s \in \mathbb{R}$. □

PROOF. (**of Lemma 4.2.3**) Using Lemma 4.2.4 ($p > 2$) instead of Lemma 4.1.5 ($p = 2$), the proof of Lemma 4.1.4 goes through almost literally. Adopting the notation used there we only indicate the minor differences.

$$\begin{aligned}
\xi - \xi_0 &= (\mathbb{1} - \epsilon^{\alpha_p}\nabla_t\nabla_t)^{-1}\Big(\xi - \epsilon^{\alpha_p}\nabla_t\nabla_t\xi - \xi + \underbrace{\epsilon^{\alpha_p}\nabla_t\eta - \epsilon^{\alpha_p}\nabla_t\eta}_{\text{add } 0} + \epsilon^2\nabla_t\eta\Big) \\
&= -\epsilon^{\alpha_p}(\mathbb{1} - \epsilon^{\alpha_p}\nabla_t\nabla_t)^{-1}\Big(\nabla_t(\nabla_t\xi - \eta) + (1 - \epsilon^{2-\alpha_p})\nabla_t\eta\Big)
\end{aligned}$$

and therefore lemma 4.2.4 implies

$$\|\xi - \xi_0\|_p \leq 2p\epsilon^{\alpha_p/2}\|\nabla_t\xi - \eta\|_p + 2\epsilon^{\alpha_p}\|\nabla_t\eta\|_p$$

where we used $\epsilon \in (0,1]$. Similarly we get

$$\|\eta - \nabla_t\xi_0\|_p \leq \|\nabla_t\xi - \eta\|_p + 4p\epsilon^{\alpha_p/2}\|\nabla_t\eta\|_p.$$

These estimates establish the first three claims. To prove the fourth one we observe that

$$\begin{aligned}
\|\xi_0\|_p &\leq \|\xi\|_p + 2p\epsilon^{2-\alpha_p/2}\|\eta\|_p \\
\epsilon\|\nabla_t\xi_0\|_p &\leq 2p\epsilon^{1-\alpha_p/2}\|\xi\|_p + 2\epsilon^{3-\alpha_p}\|\eta\|_p.
\end{aligned}$$

Use these expressions, $\alpha_p \in [0,2]$ and (42) to get

$$\begin{aligned}
\|\iota\pi_\epsilon\xi\|_{0,p,\epsilon} &= \|\xi_0\|_p + \epsilon\|\nabla_t\xi_0\|_p \\
&\leq (1 + 2p\epsilon^{1-\alpha_p/2})\|\xi\|_p + (2p\epsilon^{2-\alpha_p/2} + 2\epsilon^{3-\alpha_p})\|\eta\|_p \\
&\leq 3p\left(\|\xi\|_p + \epsilon\|\eta\|_p\right) \\
&\leq 6p\|(\xi,\eta)\|_{0,p,\epsilon}.
\end{aligned}$$

□

PROOF. (**of Lemma 4.2.2**) The proof is essentially the same as the one of lemma 4.1.3. The same notation will be used here. We only indicate the minor differences. Denote

$$\xi_0 = \pi_\epsilon \zeta = (1\!\!1 - \epsilon^{\alpha_p} \nabla_t \nabla_t)^{-1} (\xi - \epsilon^{\beta=2} \nabla_t \eta)$$

where $\nabla_t = \partial_t + A$ and $\beta = 2$. That this is the right choice follows from taking the difference below where certain inconvenient terms cancel each other precisely for $\beta = 2$. Observe the opposite sign in front of the ∂_s-terms in

$$\begin{aligned} \mathcal{D}_0 \pi_\epsilon \zeta = \mathcal{D}_0 \xi_0 &= \partial_s \xi_0 - \nabla_t \nabla_t \xi_0 - S \xi_0 \\ &= [\partial_s - S, (1\!\!1 - \epsilon^{\alpha_p} \nabla_t \nabla_t)^{-1}] (\xi - \epsilon^2 \nabla_t \eta) \\ &\quad + (1\!\!1 - \epsilon^{\alpha_p} \nabla_t \nabla_t)^{-1} (\partial_s \xi - \epsilon^2 \partial_s \nabla_t \eta - \nabla_t \nabla_t \xi \\ &\qquad + \epsilon^2 \nabla_t \nabla_t \nabla_t \eta - S\xi + \epsilon^2 S \nabla_t \eta). \end{aligned}$$

and

$$\begin{aligned} \pi_\epsilon \mathcal{D}_\epsilon \zeta = (1\!\!1 - \epsilon^{\alpha_p} \nabla_t \nabla_t)^{-1} (&\partial_s \xi - \nabla_t \eta - S\xi \\ &- \epsilon^2 \nabla_t \partial_s \eta - \nabla_t \nabla_t \xi + \nabla_t \eta - \epsilon^2 \nabla_{t(B\xi)}) \end{aligned}$$

which leads to

$$\begin{aligned} &\mathcal{D}_0 \pi_\epsilon \zeta - \pi_\epsilon \mathcal{D}_\epsilon \zeta \\ &\quad = \epsilon^{\alpha_p} (1\!\!1 - \epsilon^{\alpha_p} \nabla_t \nabla_t)^{-1} (\partial_s A + \nabla_t S) \nabla_t (1\!\!1 - \epsilon^{\alpha_p} \nabla_t \nabla_t)^{-1} (\xi - \epsilon^2 \nabla_t \eta) \\ &\quad\quad + \epsilon^{\alpha_p} (1\!\!1 - \epsilon^{\alpha_p} \nabla_t \nabla_t)^{-1} \nabla_t (\partial_s A + \nabla_t S) (1\!\!1 - \epsilon^{\alpha_p} \nabla_t \nabla_t)^{-1} (\xi - \epsilon^2 \nabla_t \eta) \\ &\quad\quad + (1\!\!1 - \epsilon^{\alpha_p} \nabla_t \nabla_t)^{-1} (\epsilon^2 \nabla_t \nabla_t \nabla_t \eta - \epsilon^2 [\partial_s, \nabla_t] \eta + \epsilon^2 S \nabla_t \eta + \epsilon^2 \nabla_{t(B\xi)}). \end{aligned}$$

Inspecting these expressions term by term and using lemma 4.2.4 as well as $\partial_s A = [\partial_s, \nabla_t]$ we find

$$\begin{aligned} &\|\mathcal{D}_0 \pi_\epsilon \zeta - \pi_\epsilon \mathcal{D}_\epsilon \zeta\|_p \\ &\qquad \le 2pc\epsilon^{\alpha_p/2} \|\xi\|_p + 2c\epsilon^2 \|\eta\|_p + 2pc\epsilon^{\alpha_p/2} \|\xi\|_p + 4p^2 c\epsilon^2 \|\eta\|_p \\ &\qquad\quad + 2\epsilon^{2-\alpha_p} \|\nabla_t \eta\|_p + c\epsilon^2 \|\eta\|_p + c\epsilon^2 \|\nabla_t \eta\|_p + 2p\epsilon^{2-\alpha_p/2} \|B\xi\|_p \\ &\qquad \le 20p^2 c\, \epsilon^{\alpha_p/2} \|\zeta\|_{0,p,\epsilon} + 4c\, \epsilon^{2-\alpha_p} \|\nabla_t \eta\|_p. \end{aligned}$$

Note that the compactification of the cylinder u via the imposed boundary conditions x, y is a crucial point in our proof as it implied the finiteness of certain supremum norms we have used in between the lines, e.g. $\|B\xi\|_p \le \|B\|_\infty \|\xi\|_p$ with $\|B\|_\infty < \infty$. Now use lemma 4.2.5 to obtain the second assertion

$$\begin{aligned} &20p^2 c\epsilon^{\alpha_p/2} \|\zeta\|_{0,p,\epsilon} + 4c\epsilon^{2-\alpha_p} \|\nabla_t \eta\|_p \\ &\qquad \le \tilde{c}\epsilon^{2-\alpha_p} \|\mathcal{D}_w^{\epsilon\,(*)} \zeta\|_{0,p,\epsilon} + \tilde{c} \left(\epsilon^{\alpha_p/2} + \epsilon^{2-\alpha_p-\kappa_p} \right) \|\zeta\|_{0,p,\epsilon} \end{aligned}$$

which is optimal for $\alpha_p/2 = 2 - \alpha_p - \kappa_p$ or $\alpha_p = 4/3 - 2\kappa_p/3$. The estimate for $\mathcal{D}_0^* \pi_\epsilon \zeta - \pi_\epsilon \mathcal{D}_\epsilon^* \zeta$ is quite similar. It is carried out in the case $p = 2$ in the proof of lemma 4.1.3. □

PROOF. (**of Proposition 4.2.1**) The proof is the same as the one of proposition 4.1.2 in the case $p = 2$ and will be sketched below. Just replace lemmata 4.1.3-4.1.6 for $p = 2$ by lemmata 4.2.2-4.2.5 and use

$$\frac{1}{\sqrt{p}}\|\nabla_t\xi\|_p \le \|\xi\|_p + \|\nabla_t\nabla_t\xi\|_p \le c_0\|\mathcal{D}_u^{0\,*}\xi\|_p$$

for any $\xi \in C_0^\infty(\mathbb{R}\times S^1, u^*TM)$. Let us prove the first of these two inequalities for $p > 2$. As in the proof of lemma 4.2.4 consider

$$\begin{aligned}
&\frac{d}{dt}\left(\langle\xi,\nabla_t\xi\rangle|\nabla_t\xi|^{p-2}\right)\\
&\quad= |\nabla_t\xi|^p + \langle\xi,\nabla_t\nabla_t\xi\rangle|\nabla_t\xi|^{p-2}\\
&\qquad+ (p-2)\langle\xi,\nabla_t\xi\rangle|\nabla_t\xi|^{p-4}\langle\nabla_t\xi,\nabla_t\nabla_t\xi\rangle\\
&\quad\ge |\nabla_t\xi|^p - (p-1)\underbrace{|\xi|}_{=a;2}\underbrace{|\nabla_t\nabla_t\xi|}_{=b;2}|\nabla_t\xi|^{p-2}\\
&\quad\ge |\nabla_t\xi|^p - \underbrace{(p-1)|\xi|^2}_{=a;p/2}\underbrace{\tfrac{1}{2}|\nabla_t\xi|^{p-2}}_{=b;\frac{p}{p-2}} - \underbrace{(p-1)|\nabla_t\nabla_t\xi|^2}_{=a;p/2}\underbrace{\tfrac{1}{2}|\nabla_t\xi|^{p-2}}_{=b;\frac{p}{p-2}}\\
&\quad\ge \left(1-\tfrac{p-2}{p}\,2^{p/(2-p)}\right)|\nabla_t\xi|^p - \tfrac{2}{p}(p-1)^{p/2}|\xi|^p\\
&\qquad- \tfrac{2}{p}(p-1)^{p/2}|\nabla_t\nabla_t\xi|^p.
\end{aligned}$$

Integration over $t \in S^1$ and $s \in \mathbb{R}$ eliminates the LHS and we get

$$\|\nabla_t\xi\|_p{}^p \le (p-1)^{p/2}\left(\|\xi\|_p{}^p + \|\nabla_t\nabla_t\xi\|_p{}^p\right).$$

Carrying out the same steps as in the proof of proposition 4.1.2 and assuming $\kappa_p \in (0,1)$ leads for $\epsilon_0 > 0$ sufficiently small to

$$\begin{aligned}
\|\xi\|_p &\le \|\xi - \pi_\epsilon\zeta\|_p + c_0\|\mathcal{D}_u^{0\,*}\pi_\epsilon\zeta\|_p\\
&\le c\epsilon^{\alpha/2}\|\eta - g\nabla_t\xi\|_p + c\epsilon^\alpha\|\nabla_t\eta\|_p\\
&\quad+ c\epsilon^{\alpha/2}\|\zeta\|_{0,p,\epsilon} + c\epsilon^{2-\alpha}\|\nabla_t\eta\|_p + c_0\|\pi_\epsilon\mathcal{D}_w^{\epsilon\,*}\zeta\|_p\\
&\le c\epsilon\|\mathcal{D}_w^{\epsilon\,*}\zeta\|_{0,p,\epsilon} + c\epsilon^{\min\{2-\kappa_p,3/2\}}\|\eta\|_p + c\|\pi_\epsilon\mathcal{D}_w^{\epsilon\,*}\zeta\|_p
\end{aligned}$$

and

$$\|\eta\|_p \le c\epsilon^{1/2}\|\mathcal{D}_w^{\epsilon\,*}\zeta\|_{0,p,\epsilon} + c\epsilon^{1/2-\kappa_p}\|\xi\|_p + c\|\pi_\epsilon\mathcal{D}_w^{\epsilon\,*}\zeta\|_p.$$

Note that only in the ξ-estimate we were forced to require $\alpha = 1$ and κ_p to be strictly less than 1 and there is no way out in view of the $\nabla_t\eta$-terms. Inserting these two estimates into one another, again using $\kappa_p \in (0,1)$ and choosing $\epsilon_0 > 0$ sufficiently small implies

$$\|\xi\|_p \le c\epsilon\|\mathcal{D}_w^{\epsilon\,*}\zeta\|_{0,p,\epsilon} + c\|\pi_\epsilon\mathcal{D}_w^{\epsilon\,*}\zeta\|_p$$

and

$$\|\eta\|_p \le c\epsilon^{1/2}\|\mathcal{D}_w^{\epsilon\,*}\zeta\|_{0,p,\epsilon} + c\epsilon^{\min\{1/2-\kappa_p,0\}}\|\pi_\epsilon\mathcal{D}_w^{\epsilon\,*}\zeta\|_p.$$

Hence our first two claims are established and they imply the third one. The fourth one follows from the third using lemma 4.2.3. ☐

4.3. The standard linear estimate

Throughout this section we simply denote $\mathbb{R}^n$-valued functions for instance by ξ – instead of $\vec{\xi}$ as in other parts of this text. For $\epsilon > 0$ let $S^1_\epsilon = \mathbb{R}/\epsilon^{-1}\mathbb{Z}$ and $Z_\epsilon = \mathbb{R} \times S^1_\epsilon$ with coordinates (s,t).

THEOREM 4.3.1. (**standard linear estimate**) *Let $\bar\partial + T = \partial_s + J_0\partial_t + T$, where $T \in C^\infty(Z_\epsilon, \mathbb{R}^{2n\times 2n})$ and $\lim_{s\to\mp\infty} T(s,t) = T^\mp(t)$ exists uniformly in t. Then for any $1 < p < \infty$ there exists a constant $c_p > 0$ such that*

$$\boxed{\|\zeta\|_{W^{1,p}(Z_\epsilon,\mathbb{R}^{2n})} \le c_p\left(\|(\bar\partial + T)\zeta\|_{L^p(Z_\epsilon,\mathbb{R}^{2n})} + \|\zeta\|_{L^p(Z_\epsilon,\mathbb{R}^{2n})}\right)}$$

for any $\epsilon > 0$ and $\zeta \in W^{1,p}(Z_\epsilon, \mathbb{R}^{2n})$. This continues to hold for a uniformly bounded family $\{T_a\}_{a\in A}$, i.e.

$$\sup_{a\in A}\sup_{(s,t)\in Z_\epsilon} \|T_a(s,t)\|_{\mathcal{L}(\mathbb{R}^{2n})} = \sup_{a\in A}\|T_a\|_\infty < \infty.$$

Using the technique of rescaling the standard linear estimate immediately implies linear estimates for the operators $\mathcal{D}^\epsilon_w$ and $\mathcal{D}^{\epsilon\,*}_w$ uniformly in $\epsilon \in (0,1]$. The former one will be used in the proof of the key estimate of $\mathcal{D}^\epsilon_w$ on the range of $\mathcal{D}^{\epsilon\,*}_w$ (theorem 4.4.4).

THEOREM 4.3.2. (**linear estimate**) *Let $w \in C^\infty(\mathbb{R}\times S^1, T^*M)$ such that $w \to g(y^\mp)\partial_t y^\mp$ uniformly in t for $s \to \mp\infty$ and $y^\mp \in C^\infty(S^1, M)$. Set $u = \tau_M^* w$. Then for any $1 < p < \infty$ there exists a constant $c = c(p) > 0$ such that*

$$\boxed{\|(\xi,\eta)\|_{1,p,\epsilon} \le c\left(\epsilon^2\|\mathcal{D}^\epsilon_w(\xi,\eta)\|_{0,p,\epsilon} + \|(\xi,\eta)\|_{0,p,\epsilon}\right)}$$

for $0 < \epsilon \le 1$ and $(\xi,\eta) \in W^{1,p}_\epsilon(\mathbb{R}\times S^1, u^*TM \oplus u^*T^*M)$. The same holds for $\mathcal{D}^{\epsilon\,*}_w$.

PROOF. We are working in an orthonormal frame, cf. (144), so that the operators $\mathcal{D}^\epsilon_w$ and $\mathcal{D}^{\epsilon\,*}_w$ are given by $\mathcal{D}_\epsilon$ as in (143) and $\mathcal{D}^*_\epsilon$ as in (34) (with S replaced by $C^* = C^T$ strictly speaking). Pick $(\xi,\eta) \in C^\infty_0(\mathbb{R}\times S^1, \mathbb{R}^{2n})$ and $\epsilon \in (0,1]$, then rescale

$$\begin{pmatrix}\tilde\xi(\tilde s,\tilde t)\\ \tilde\eta(\tilde s,\tilde t)\end{pmatrix} = \begin{pmatrix}\xi(\epsilon^2\tilde s,\epsilon\tilde t)\\ \epsilon\eta(\epsilon^2\tilde s,\epsilon\tilde t)\end{pmatrix} \in C^\infty_0(Z_\epsilon, \mathbb{R}^{2n}), \tag{44}$$

$\tilde w(\tilde s,\tilde t) = \epsilon w(\epsilon^2\tilde s,\epsilon\tilde t)$ and $\tilde y^\mp(\tilde t) = y^\mp(\epsilon\tilde t)$. Note that $\tilde w(\tilde s,\tilde t) \to g(\tilde y^\mp(\tilde t))\partial_{\tilde t}\tilde y^\mp(\tilde t)$ for $\tilde s \to \mp\infty$ uniformly in $\tilde t$. Now apply the standard linear estimate theorem 4.3.1 to the operator

$$(\tilde{\bar\partial} + \tilde T_\epsilon)\begin{pmatrix}\tilde\xi\\ \tilde\eta\end{pmatrix} = (\partial_{\tilde s} + J_0\partial_{\tilde t})\begin{pmatrix}\tilde\xi\\ \tilde\eta\end{pmatrix} + \begin{pmatrix}-\tilde C & \tilde A\\ \tilde A + \tilde B & -\mathbb{1}\end{pmatrix}\begin{pmatrix}\tilde\xi\\ \tilde\eta\end{pmatrix}$$

where

$$\tilde A(\tilde s,\tilde t) = \epsilon A(\epsilon^2\tilde s,\epsilon\tilde t)\ ,\ \tilde B(\tilde s,\tilde t) = \epsilon^3 B(\epsilon^2\tilde s,\epsilon\tilde t)\ ,\ \tilde C(\tilde s,\tilde t) = \epsilon^2 C(\epsilon^2\tilde s,\epsilon\tilde t).$$

The families of matrices A, B and C are defined in appendix A section A.4. Indeed

$$\tilde{T}_\epsilon(\tilde{s},\tilde{t}) = \begin{pmatrix} -\epsilon^2 C & \epsilon A \\ \epsilon A + \epsilon^3 B & -\mathbb{1} \end{pmatrix} (\epsilon^2\tilde{s}, \epsilon\tilde{t})$$

is a family of operators uniformly bounded for $\epsilon \in (0,1]$, so that we get a constant $c_p > 0$ (uniformly in ϵ) such that

$$\|(\tilde{\xi},\tilde{\eta})\|_{W^{1,p}(Z_\epsilon,\mathbb{R}^{2n})} \leq c_p\Big(\|(\tilde{\bar{\partial}} + \tilde{T}_\epsilon)(\tilde{\xi},\tilde{\eta})\|_{L^p(Z_\epsilon,\mathbb{R}^{2n})} + \|(\tilde{\xi},\tilde{\eta})\|_{L^p(Z_\epsilon,\mathbb{R}^{2n})}\Big)$$

for all $(\tilde{\xi},\tilde{\eta}) \in W^{1,p}(Z_\epsilon,\mathbb{R}^{2n})$. Recall that as the frame is chosen to be parallel with respect to the variable s, $\vec{\nabla}_{\tilde{s}} = \partial_{\tilde{s}}$ and $\vec{\nabla}_{\tilde{t}}$ is represented by $\partial_{\tilde{t}} + \tilde{A}$, then

$$\begin{aligned}
&\|(\tilde{\xi},\tilde{\eta})\|^p_{L^p(Z_\epsilon,\mathbb{R}^{2n})} + \|\vec{\nabla}_{\tilde{t}}(\tilde{\xi},\tilde{\eta})\|^p_{L^p(Z_\epsilon,\mathbb{R}^{2n})} + \|\vec{\nabla}_{\tilde{s}}(\tilde{\xi},\tilde{\eta})\|^p_{L^p(Z_\epsilon,\mathbb{R}^{2n})} \\
&\quad = \int_{-\infty}^{\infty}\int_0^{\epsilon^{-1}} \Big(|\tilde{\xi}(\tilde{s},\tilde{t})|^p + |\tilde{\eta}(\tilde{s},\tilde{t})|^p \\
&\qquad\qquad + |\vec{\nabla}_{\tilde{t}}\tilde{\xi}(\tilde{s},\tilde{t})|^p + |\vec{\nabla}_{\tilde{t}}\tilde{\eta}(\tilde{s},\tilde{t})|^p \\
&\qquad\qquad + |\partial_{\tilde{s}}\tilde{\xi}(\tilde{s},\tilde{t})|^p + |\partial_{\tilde{s}}\tilde{\eta}(\tilde{s},\tilde{t})|^p \Big)\, d\tilde{t}d\tilde{s} \\
&\quad = \int_{-\infty}^{\infty}\int_0^{1} \Big(|\xi(s,t)|^p + \epsilon^p|\eta(s,t)|^p \\
&\qquad\qquad + \epsilon^p|\vec{\nabla}_t\xi(s,t)|^p + \epsilon^{2p}|\vec{\nabla}_t\eta(s,t)|^p \\
&\qquad\qquad + \epsilon^{2p}|\partial_s\xi(s,t)|^p + \epsilon^{3p}|\partial_s\eta(s,t)|^p \Big)\, \frac{dt}{\epsilon}\frac{ds}{\epsilon^2} \\
&\quad = \frac{1}{\epsilon^3}\, \|(\xi,\eta)\|^p_{1,p,\epsilon}.
\end{aligned}$$

In the above calculation we first used the definition (44) of $(\tilde{\xi},\tilde{\eta})$ as well as the identities

$$\begin{aligned}
\vec{\nabla}_{\tilde{t}}\tilde{\xi}(\tilde{s},\tilde{t}) &= \epsilon\vec{\nabla}_t\xi(\epsilon^2\tilde{s},\epsilon\tilde{t}) & \partial_{\tilde{s}}\tilde{\xi}(\tilde{s},\tilde{t}) &= \epsilon^2\partial_s\xi(\epsilon^2\tilde{s},\epsilon\tilde{t}) \\
\vec{\nabla}_{\tilde{t}}\tilde{\eta}(\tilde{s},\tilde{t}) &= \epsilon\vec{\nabla}_t\epsilon\eta(\epsilon^2\tilde{s},\epsilon\tilde{t}) & \partial_{\tilde{s}}\tilde{\eta}(\tilde{s},\tilde{t}) &= \epsilon^2\partial_s\epsilon\eta(\epsilon^2\tilde{s},\epsilon\tilde{t})
\end{aligned}$$

and then carried out the change of variables $s = \epsilon^2\tilde{s}$, $t = \epsilon\tilde{t}$. Note that this also includes

$$\|(\tilde{\xi},\tilde{\eta})\|^p_{L^p(Z_\epsilon,\mathbb{R}^{2n})} = \frac{1}{\epsilon^3}\, \|(\xi,\eta)\|^p_{0,p,\epsilon}.$$

The next step is to compute

$$\begin{aligned}
&\|(\tilde{\bar{\partial}} + \tilde{T}_\epsilon)(\tilde{\xi},\tilde{\eta})\|^p_{L^p(Z_\epsilon,\mathbb{R}^{2n})} \\
&= \int_{-\infty}^{\infty}\int_0^{\epsilon^{-1}} |\partial_{\tilde{s}}\tilde{\xi} - \partial_{\tilde{t}}\tilde{\eta} + \tilde{A}\tilde{\eta} - \tilde{C}\tilde{\xi}|^p \\
&\qquad\qquad + |\partial_{\tilde{s}}\tilde{\eta} + \partial_{\tilde{t}}\tilde{\xi} + \tilde{A}\tilde{\xi} + \tilde{B}\tilde{\xi} - \tilde{\eta}|^p \, d\tilde{t} d\tilde{s} \\
&= \int_{-\infty}^{\infty}\int_0^{1} |\epsilon^2\partial_s\xi - \epsilon^2\partial_t\eta + \epsilon^2 A\eta - \epsilon^2 C\xi|^p \\
&\qquad\qquad + |\epsilon^3\partial_s\eta + \epsilon\partial_t\xi + \epsilon A\xi + \epsilon^3 B\xi - \epsilon\eta|^p \, \frac{dt}{\epsilon}\frac{ds}{\epsilon^2} \\
&= \frac{\epsilon^{2p}}{\epsilon^3} \, \|\mathcal{D}_\epsilon(\xi,\eta)\|^p_{0,p,\epsilon}.
\end{aligned}$$

Using the results obtained above one gets

$$\begin{aligned}
&\frac{1}{\epsilon^{3/p}} \, \|(\xi,\eta)\|_{1,p,\epsilon} \\
&= \left(\|(\tilde{\xi},\tilde{\eta})\|^p_{L^p(Z_\epsilon,\mathbb{R}^{2n})} + \|\vec{\nabla}_{\tilde{t}}(\tilde{\xi},\tilde{\eta})\|^p_{L^p(Z_\epsilon,\mathbb{R}^{2n})} + \|\vec{\nabla}_{\tilde{s}}(\tilde{\xi},\tilde{\eta})\|^p_{L^p(Z_\epsilon,\mathbb{R}^{2n})}\right)^{1/p} \\
&\le 2c_A\|(\tilde{\xi},\tilde{\eta})\|_{W^{1,p}(Z_\epsilon,\mathbb{R}^{2n})} \\
&\le 2c_A c_p\left(\|(\tilde{\bar{\partial}} + \tilde{T}_\epsilon)(\tilde{\xi},\tilde{\eta})\|_{L^p(Z_\epsilon,\mathbb{R}^{2n})} + \|(\tilde{\xi},\tilde{\eta})\|_{L^p(Z_\epsilon,\mathbb{R}^{2n})}\right) \\
&= \frac{2c_A c_p}{\epsilon^{3/p}} \left(\epsilon^2\|\mathcal{D}_\epsilon(\xi,\eta)\|_{0,p,\epsilon} + \|(\xi,\eta)\|_{0,p,\epsilon}\right)
\end{aligned}$$

which proves the first claim. The estimate for $\mathcal{D}_w^{\epsilon\,*}$ is obtained the same way. □

The proof of the standard linear estimate, theorem 4.3.1, will involve a partition of unity argument and the following consequence of the Calderon-Zygmund inequality

PROPOSITION 4.3.3. *For $1 < p < \infty$ there exists a constant $c_p > 0$ such that*

$$\|\nabla\zeta\|_{L^p(\mathbb{R}^2,\mathbb{R}^{2n})} \le c_p \, \|\bar{\partial}\zeta\|_{L^p(\mathbb{R}^2,\mathbb{R}^{2n})}$$

for $\zeta \in C_0^\infty(\mathbb{R}^2,\mathbb{R}^{2n})$ and $\bar{\partial} = \partial_s + J_0\partial_t$.

Following the exposition in [**MS95**], appendix B.2, we briefly recall the *fundamental solution* of Laplace's equation

$$K(x) = \frac{1}{2\pi}\log|x| \quad , \quad x \in \mathbb{R}^2 \setminus \{0\}.$$

Denote $x = (s,t)$ and $K_j = \partial_j K$ for $j = s,t$. Every $u \in C_0^\infty(\mathbb{R}^2,\mathbb{R})$ satisfies

$$u = K * \triangle u \quad , \quad \partial_j u = K_j * \triangle u \tag{45}$$

with $\triangle = \partial_s\partial_s + \partial_t\partial_t$. For every $f \in C_0^\infty(\mathbb{R}^2, \mathbb{R})$ one has

$$\triangle(K * f) = f \quad , \quad \triangle(K_j * f) = \partial_j f. \tag{46}$$

Hence $u := K * f$ for $f \in C_0^\infty(\mathbb{R}^2, \mathbb{R})$ solves the *inhomogeneous Laplace equation* $\triangle u = f$. Let $\Omega \subset \mathbb{R}^2$ be open and $f \in L^1_{loc}(\Omega)$ (i.e. $f \in L^1(Q)$ for any compact set $Q \subset \Omega$), then a function $u \in L^1_{loc}(\Omega)$ is called a *weak solution* of $\triangle u = f$ if

$$\int_\Omega u(x)\, \triangle\phi(x)\, dx = \int_\Omega f(x)\, \phi(x)\, dx \quad , \quad \forall \phi \in C_0^\infty(\Omega).$$

Note that the last identity determines u up to a set of measure zero, cf. Folgerung 2.11 in [**Al92**].

THEOREM 4.3.4. (**Calderon-Zygmund inequality**), [**CaZ52**] *For any* $1 < p < \infty$ *there exists a constant* $c = c(p) > 0$ *such that*

$$\|\nabla(K_j * f)\|_{L^p} \leq c\, \|f\|_{L^p}$$

for $f \in C_0^\infty(\mathbb{R}^2, \mathbb{R})$ *and* $j = s, t$.

Now we identify $\mathbb{C}$ with $\mathbb{R}^2$ by $z = x + iy \mapsto (x, y)$. Multiplication with i is then represented by the matrix

$$J_0 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

We define the first order differential operators

$$\partial_{\bar z} = \frac{1}{2}(\partial_x + J_0\partial_y) \quad , \quad \partial_z = \frac{1}{2}(\partial_x - J_0\partial_y).$$

The fundamental solution of the *Cauchy Riemann operator* $\partial_{\bar z}$ is given by $N(z) = 1/\pi z$, i.e. if $f \in C_0^\infty(\mathbb{R}^2, \mathbb{R}^2)$ then $u = N * f$ solves $\partial_{\bar z} u = f$ ([**MS95**], lemma B.3.1). Moreover, for $u, f \in L^p(\mathbb{R}^2, \mathbb{R}^2)$ with compact support it holds ([**MS95**], lemma B.3.2): u is a weak solution of $\partial_{\bar z} u = f$, iff $u = N * f$.
Here $u \in L^1_{loc}(\Omega)$ is called a *weak solution* of $\partial_{\bar z} u = f$ for $f \in L^1_{loc}(\Omega)$ if

$$\int_\Omega \langle \partial_z \phi, u\rangle + \int_\Omega \langle \phi, f\rangle = 0 \quad , \quad \forall \phi \in C_0^\infty(\Omega, \mathbb{R}^2).$$

$\langle \cdot, \cdot \rangle$ denotes the euclidean inner product on $\mathbb{R}^2$. A straightforward calculation shows that $\triangle = 4\partial_z\partial_{\bar z}$ and $N = 4\partial_z K$.

PROOF. (OF PROPOSITION 4.3.3) It suffices to consider the case $n = 1$. For $u \in C_0^\infty(\mathbb{R}^2, \mathbb{R}^2)$ we define $f = \partial_{\bar z} u$. As mentioned above

$$u = N * f = 4\partial_z K * f = 2(\partial_s - J_0\partial_t)K * f$$

and hence

$$\begin{aligned}\|\nabla u\|_{L^p}^p &= \|\partial_s u\|_{L^p}^p + \|\partial_t u\|_{L^p}^p \\ &\leq 2^{p(p-1)}\big(\|\partial_s(\partial_s K * f)\|_{L^p}^p + \|J_0\partial_s(\partial_t K * f)\|_{L^p}^p \\ &\quad + \|\partial_t(\partial_s K * f)\|_{L^p}^p + \|J_0\partial_t(\partial_t K * f)\|_{L^p}^p\big) \\ &= 2^{p(p-1)}\big(\|\nabla(\partial_s K * f)\|_{L^p}^p + \|\nabla(\partial_t K * f)\|_{L^p}^p\big) \\ &\leq 2^{2p}c^p\|f\|_{L^p}^p = (4c)^p\|\bar\partial u\|_{L^p}^p.\end{aligned}$$

The last inequality uses the Calderon-Zygmund inequality. Moreover, we used that J_0 is constant and leaves the norm invariant. □

Proof. (of theorem 4.3.1) **Step 1** First we construct an open cover of $Z_1 = \mathbb{R}\times S^1$ by two open sets and a subordinate partition of unity (which is constant in the s-direction). Identifying $S^1 = [0,1]/\{0,1\}$ we define an open cover of S^1 by

$$\hat U_1^t = (1/8, 7/8)\ ,\ \hat U_2^t = [0, 3/8) \cup (5/8, 1].$$

Let $\hat\rho_1(t)$ be a smooth compactly supported function on $\hat U_1$ which takes values in $[0,1]$ and is identically 1 on $[3/8, 5/8]$. Define $\hat\rho_2(t) = 1 - \hat\rho_1(t)$. We extend the above trivially to the s-direction

$$\begin{aligned}U_1 = \mathbb{R}\times\hat U_1\ &,\ U_2 = \mathbb{R}\times\hat U_2\\ \rho_1(s,t) = \hat\rho_1(t)\ &,\ \rho_2(s,t) = \hat\rho_2(t).\end{aligned}$$

Step 2 For $\epsilon > 0$ we get a covering of Z_ϵ and a subordinate partition of unity by rescaling:

$$\begin{aligned}\tilde U_1 = \mathbb{R}\times(\epsilon^{-1}/8, \epsilon^{-1}7/8)\ ,\ \tilde U_2 = \mathbb{R}\times\Big([0,\epsilon^{-1}3/8)\cup(\epsilon^{-1}5/8,\epsilon^{-1}]\Big)\\ \tilde\rho_1(\tilde s,\tilde t) = \rho_1(\tilde s,\epsilon\tilde t)\ ,\ \tilde\rho_2(\tilde s,\tilde t) = \rho_2(\tilde s,\epsilon\tilde t).\end{aligned}$$

Note that for $i = 1,2$

$$\nabla\tilde\rho_i(\tilde s,\tilde t) = \begin{pmatrix}\partial_{\tilde s}\tilde\rho_i(\tilde s,\tilde t)\\ \partial_{\tilde t}\tilde\rho_i(\tilde s,\tilde t)\end{pmatrix} = \begin{pmatrix}0\\ \epsilon\ \partial_t\rho_i(\tilde s,\epsilon\tilde t)\end{pmatrix} = \epsilon\ \nabla\rho_i(\tilde s,\epsilon\tilde t)$$

and

$$\begin{aligned}\|\bar\partial\tilde\rho_i(\tilde s,\tilde t)\| &= \|\begin{pmatrix}\partial_{\tilde s}\tilde\rho_i & -\partial_{\tilde t}\tilde\rho_i\\ \partial_{\tilde t}\tilde\rho_i & \partial_{\tilde s}\tilde\rho_i\end{pmatrix}(\tilde s,\tilde t)\|\\ &= \epsilon\,\|\begin{pmatrix}0 & -\partial_t\rho_i\\ \partial_t\rho_i & 0\end{pmatrix}(\tilde s,\epsilon\tilde t)\| \leq \epsilon b_i \leq \epsilon b,\end{aligned}$$

where

$$b_i = \max_{t\in supp\,\rho_i}\|\begin{pmatrix}0 & -\partial_t\rho_i(t)\\ \partial_t\rho_i(t) & 0\end{pmatrix}\| < \infty \quad,\quad b = \max\{1, b_1, b_2\}.$$

Step 3 We may restrict to $\zeta \in C_0^\infty(Z_\epsilon,\mathbb{R}^{2n})$ as this space is dense in $W^{1,p}(Z_\epsilon,\mathbb{R}^{2n})$. This nicely combines with ρ_1 and ρ_2 having compact support with respect to the variable t and so results in $\rho_1\zeta$ and $\rho_2\zeta$ having compact support. Therefore we may apply the consequence proposition 4.3.3 of the Calderon-Zygmund inequality.

$$
\begin{aligned}
\|\nabla\zeta\|_{L^p(Z_\epsilon,\mathbb{R}^{2n})} &\leq \sum_{i=1}^{2} \|\nabla(\tilde{\rho}_i\zeta)\|_{L^p(\tilde{U}_i,\mathbb{R}^{2n})} \\
&\leq c_p \sum_{i=1}^{2} \|\bar{\partial}(\tilde{\rho}_i\zeta)\|_{L^p(\tilde{U}_i,\mathbb{R}^{2n})} \\
&\leq c_p \sum_{i=1}^{2} \left(\|(\bar{\partial}\tilde{\rho}_i)\zeta\|_{L^p(supp\ \tilde{\rho}_i\zeta,\mathbb{R}^{2n})} + \|\tilde{\rho}_i(\bar{\partial}\zeta)\|_{L^p(supp\ \tilde{\rho}_i\zeta,\mathbb{R}^{2n})}\right) \\
&\leq c_p \left(\sum_{i=1}^{2} 1^q\right)^{1/q} \epsilon b \left(\sum_{i=1}^{2} \|\zeta\|^p_{L^p(supp\ \tilde{\rho}_i\zeta,\mathbb{R}^{2n})}\right)^{1/p} \\
&\quad + c_p \left(\sum_{i=1}^{2} 1^q\right)^{1/q} \left(\sum_{i=1}^{2} \|\bar{\partial}\zeta\|^p_{L^p(supp\ \tilde{\rho}_i\zeta,\mathbb{R}^{2n})}\right)^{1/p} \\
&\leq 2c_p \left(\|\bar{\partial}\zeta\|_{L^p(Z_\epsilon,\mathbb{R}^{2n})} + \epsilon b\|\zeta\|_{L^p(Z_\epsilon,\mathbb{R}^{2n})}\right).
\end{aligned}
$$

In the fourth estimate we used the discrete Hölder inequality with $1/q + 1/p = 1$ and in the last one the additivity of integrals.

Step 4 Note that

$$
\|\mathcal{D}\zeta\|_{L^p(Z_\epsilon,\mathbb{R}^{2n})} \geq \|\bar{\partial}\zeta\|_{L^p(Z_\epsilon,\mathbb{R}^{2n})} - \|T\|_\infty \|\zeta\|_{L^p(Z_\epsilon,\mathbb{R}^{2n})}
$$

hence using Step 3 we get

$$
\begin{aligned}
\|\zeta\|_{W^{1,p}(Z_\epsilon,\mathbb{R}^{2n})} &= \|\zeta\|_{L^p(Z_\epsilon,\mathbb{R}^{2n})} + \|\nabla\zeta\|_{L^p(Z_\epsilon,\mathbb{R}^{2n})} \\
&\leq (1+\epsilon\tilde{c}_p)\|\zeta\|_{L^p(Z_\epsilon,\mathbb{R}^{2n})} + \tilde{c}_p\|\bar{\partial}\zeta\|_{L^p(Z_\epsilon,\mathbb{R}^{2n})} \\
&\leq c_p\big(\|\mathcal{D}\zeta\|_{L^p(Z_\epsilon,\mathbb{R}^{2n})} + \|\zeta\|_{L^p(Z_\epsilon,\mathbb{R}^{2n})}\big).
\end{aligned}
$$

In the last inequality we used $\|T\|_\infty < \infty$, which follows from the assumption on the asymptotic uniform convergence of T. The argument also works in the case of a uniformly bounded family. □

4.4. An estimate for the right inverse

In this section we derive the key estimates for $\mathcal{D}_w^\epsilon$ on the range of ${\mathcal{D}_w^\epsilon}^*$. In other words these give rise to bounds, uniformly in $\epsilon \in (0, \epsilon_0)$, for the right inverse $\mathcal{Q}_w^\epsilon$ of $\mathcal{D}_w^\epsilon$.

We start with the definition of the *formal adjoint* ${\mathcal{D}_w^\epsilon}^*$ of $\mathcal{D}_w^\epsilon$ and some considerations concerning their kernels and cokernels. Let $u \in \mathcal{P}_{x^-,x^+}(\mathbb{R} \times S^1, M)$ and $x^-,\ x^+ \in \mathit{Crit}\, \mathcal{I}_V$, then $w := g(u)\partial_t u$ is a cylinder in T^*M and $z^\mp := g(x^\mp)\partial_t x^\mp$ are its hamiltonian boundary conditions. If they are nondegenerate as critical points of the symplectic action, then for $1 < p < \infty$ the operator $\mathcal{D}_w^\epsilon$ is a Fredholm operator, cf. [**RS95**] theorem A for a general exposition or [**Sa97**] theorem 2.2 for the case $\epsilon = 1$. Recall that with respect to an orthonormal frame which is parallel in the s-direction (cf. appendix A section A.4) this operator is represented by

$$\mathcal{D}_\epsilon : W_\epsilon^{1,p}(\mathbb{R} \times S^1, \mathbb{R}^{2n}) \to L_\epsilon^p(\mathbb{R} \times S^1, \mathbb{R}^{2n})$$
$$\begin{pmatrix} \vec{\xi} \\ \vec{\eta} \end{pmatrix} \mapsto \begin{pmatrix} \partial_s \vec{\xi} - \vec{\nabla}_t \vec{\eta} - Q\vec{\xi} \\ \partial_s \vec{\eta} + \epsilon^{-2}(\vec{\nabla}_t \vec{\xi} - \vec{\eta}) + B\vec{\xi} \end{pmatrix}.$$

That $\mathcal{D}_\epsilon$ is bounded follows from the *B.L.T. theorem* ([**RS1**] theorem I.7): As $\mathcal{D}_\epsilon : (C_0^\infty(\mathbb{R} \times S^1, \mathbb{R}^{2n}), \|\cdot\|_{1,p,\epsilon}) \to L_\epsilon^p(\mathbb{R} \times S^1, \mathbb{R}^{2n})$ is bounded and the target space is complete, $\mathcal{D}_\epsilon$ extends to the completion $W_\epsilon^{1,p}(\mathbb{R}\times S^1, \mathbb{R}^{2n})$ of the domain with the same bound. Therefore we may assume without loss of generality that $\mathcal{D}_\epsilon$ acts on $C_0^\infty(\mathbb{R} \times S^1, \mathbb{R}^{2n})$.

Definition 4.4.1. Let $1/p + 1/q = 1$ and $p > 1$, then the *formal adjoint* ${\mathcal{D}_w^\epsilon}^* : W_\epsilon^{1,q}(\mathbb{R} \times S^1, u^*TM \oplus u^*T^*M) \to L_\epsilon^q(\mathbb{R} \times S^1, u^*TM \oplus u^*T^*M)$ is defined by

$$(47)\quad \langle \zeta', \mathcal{D}_w^\epsilon \zeta \rangle_\epsilon = \langle {\mathcal{D}_w^\epsilon}^* \zeta', \zeta \rangle_\epsilon \quad , \quad \forall \zeta, \zeta' \in C_0^\infty(\mathbb{R} \times S^1, u^*TM \oplus u^*T^*M),$$

where for $\epsilon \in (0, 1]$

$$\langle \cdot, \cdot \rangle_\epsilon : C_0^\infty(\mathbb{R} \times S^1, u^*TM \oplus u^*T^*M)^{\times 2} \to \mathbb{R}$$
$$\left(\begin{pmatrix} \xi_1 \\ \eta_1 \end{pmatrix}, \begin{pmatrix} \xi_2 \\ \eta_2 \end{pmatrix} \right) \mapsto \langle \xi_1, \xi_2 \rangle_{L^2} + \epsilon^2 \langle \eta_1, \eta_2 \rangle_{L^2}.$$

Existence of ${\mathcal{D}_w^\epsilon}^* : C_0^\infty \to C_0^\infty \subset L_\epsilon^q$ follows by explicit calculation via partial integration and uniqueness by the nondegeneracy of $\langle \cdot, \cdot \rangle_\epsilon$; then extend ${\mathcal{D}_w^\epsilon}^*$ to $W_\epsilon^{1,q}$ (by B.L.T.). For $(\vec{\xi}, \vec{\eta}) \in C_0^\infty(\mathbb{R} \times S^1, \mathbb{R}^{2n})$ we get

$$\mathcal{D}_\epsilon^* \begin{pmatrix} \vec{\xi} \\ \vec{\eta} \end{pmatrix} = \begin{pmatrix} -\partial_s \vec{\xi} - \vec{\nabla}_t \vec{\eta} - Q\vec{\xi} + \epsilon^2 B^* \vec{\eta} \\ -\partial_s \vec{\eta} + \epsilon^{-2}(\vec{\nabla}_t \vec{\xi} - \vec{\eta}) \end{pmatrix}.$$

Note that as $1/p + 1/q = 1$ the dual space of L_ϵ^p may be identified with L_ϵ^q and the duality pairing is given by the bilinear form $\langle \cdot, \cdot \rangle_\epsilon$ (cf. [**Al92**] Satz 4.13). The cokernel of $\mathcal{D}_w^\epsilon$ is now defined as the annihilator of $\mathit{ran}\, \mathcal{D}_w^\epsilon$.

DEFINITION 4.4.2.

$$\begin{aligned} Coker\, \mathcal{D}_w^\epsilon &= (ran\, \mathcal{D}_w^\epsilon)^\perp \\ &= \{\zeta' \in (L_\epsilon^p)^* \simeq L_\epsilon^q \mid 0 = \langle \zeta', \mathcal{D}_w^\epsilon \zeta \rangle_\epsilon, \forall \zeta \in W_\epsilon^{1,p}\} \\ Coker\, \mathcal{D}_w^{\epsilon\,*} &= (ran\, \mathcal{D}_w^{\epsilon\,*})^\perp \\ &= \{\zeta' \in (L_\epsilon^q)^* \simeq L_\epsilon^p \mid 0 = \langle \mathcal{D}_w^{\epsilon\,*} \zeta', \zeta \rangle_\epsilon, \forall \zeta' \in W_\epsilon^{1,p}\}. \end{aligned}$$

Elliptic regularity implies that for $\zeta' \in Coker\, \mathcal{D}_w^\epsilon$, i.e. a priori $\zeta' \in L_\epsilon^q$, indeed $\zeta' \in W_\epsilon^{1,q}$ and $\mathcal{D}_w^{\epsilon\,*}\zeta' = 0$ (cf. [**MS94**] exercise B.3.5). This shows $Coker\, \mathcal{D}_w^\epsilon \subset Ker\, \mathcal{D}_w^{\epsilon\,*}$. The opposite inclusion follows from the definition of the formal adjoint; the same for $Coker\, \mathcal{D}_w^{\epsilon\,*}$ and $Ker\, \mathcal{D}_w^\epsilon$. This proves

LEMMA 4.4.3.

$$\begin{aligned} Coker\, \mathcal{D}_w^\epsilon &= Ker\, \mathcal{D}_w^{\epsilon\,*} \\ Coker\, \mathcal{D}_w^{\epsilon\,*} &= Ker\, \mathcal{D}_w^\epsilon. \end{aligned}$$

Similar conclusions hold for the operator

$$\begin{aligned} \mathcal{D}_0 : \mathcal{W}^{1,p} &\to \mathcal{H}^p \\ \vec{\xi} &\mapsto \partial_s \vec{\xi} - \vec{\nabla}_t \vec{\nabla}_t \vec{\xi} - Q\vec{\xi}. \end{aligned}$$

Assume now that $\mathcal{D}_w^\epsilon$ is onto, then we can define a right inverse

$$\begin{aligned} \mathcal{Q}_w^\epsilon : L_\epsilon^p(\mathbb{R} \times S^1, u^*TM \oplus u^*T^*M) &\to W_\epsilon^{1,p}(\mathbb{R} \times S^1, u^*TM \oplus u^*T^*M) \\ \zeta &\mapsto \mathcal{D}_w^{\epsilon\,*} \left(\mathcal{D}_w^\epsilon \mathcal{D}_w^{\epsilon\,*}\right)^{-1} \zeta. \end{aligned}$$

Being the composition of two bounded operators, $\mathcal{Q}_w^\epsilon$ is itself a bounded operator. To see the boundedness of $((\mathcal{D}_w^\epsilon \mathcal{D}_w^{\epsilon\,*})$ consider

$$W_\epsilon^{1,q} \supset W_\epsilon^{2,p} \xrightarrow{\mathcal{D}_w^{\epsilon\,*}} W_\epsilon^{1,p} \xrightarrow{\mathcal{D}_w^\epsilon} L_\epsilon^p.$$

Note that the inclusion holds because of $p > 2$. As $Ker\, \mathcal{D}_w^\epsilon = Coker\, \mathcal{D}_w^{\epsilon\,*}$ the operator $\mathcal{D}_w^\epsilon \mathcal{D}_w^{\epsilon\,*}$ is a bounded bijection from $W_\epsilon^{2,p}$ onto L_ϵ^p, hence it has a bounded inverse by the open mapping theorem (cf. [**RS1**] theorem III.10 and III.11). The crucial point is to get a bound for $\mathcal{Q}_w^\epsilon$, which is independent of $\epsilon \in (0, \epsilon_0)$. This and more refined estimates are the content of the next theorem. Note that throughout this subsection the projection π_ϵ is given by

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

and we need the assumption $\kappa_p \in (0,1)$ on the critical exponent.

THEOREM 4.4.4. *Let $u \in \mathcal{P}_{x,y}(\mathbb{R} \times S^1, M)$, where x, y are smooth loops in M, and define $w = g(u)\partial_t u$ and $(\xi^*, \eta^*) = \mathcal{D}_w^{\epsilon\,*}(\xi, \eta)$. Assume that D_u^0 is onto and $\kappa_p \in (0,1)$. Then for every $p > 2$ there exist constants $c_p > 0$,*

$\epsilon_0 = \epsilon_0(p) > 0$ such that $\mathcal{D}_w^\epsilon$ is onto,

$$\boxed{\begin{gathered}\|\xi^*\|_p \le c_p\Big(\epsilon\|\mathcal{D}_w^\epsilon{\mathcal{D}_w^\epsilon}^*\zeta\|_{0,p,\epsilon} + \|\pi_\epsilon\mathcal{D}_w^\epsilon{\mathcal{D}_w^\epsilon}^*\zeta\|_p\Big)\\ \|\eta^*\|_p + \|\nabla_t\xi^*\|_p \le c_p\Big(\epsilon^{1/2}\|\mathcal{D}_w^\epsilon{\mathcal{D}_w^\epsilon}^*\zeta\|_{0,p,\epsilon} + \epsilon^{-1/2}\|\pi_\epsilon\mathcal{D}_w^\epsilon{\mathcal{D}_w^\epsilon}^*\zeta\|_p\Big)\end{gathered}}$$

and

$$\boxed{\begin{gathered}\|{\mathcal{D}_w^\epsilon}^*\zeta\|_{1,p,\epsilon} \le c_p\Big(\epsilon\|\mathcal{D}_w^\epsilon{\mathcal{D}_w^\epsilon}^*\zeta\|_{0,p,\epsilon} + \|\pi_\epsilon\mathcal{D}_w^\epsilon{\mathcal{D}_w^\epsilon}^*\zeta\|_p\Big)\\ \|{\mathcal{D}_w^\epsilon}^*\zeta\|_{1,p,\epsilon} \le c_p\|\mathcal{D}_w^\epsilon{\mathcal{D}_w^\epsilon}^*\zeta\|_{0,p,\epsilon}\end{gathered}}$$

for any $\epsilon \in (0, \epsilon_0)$ and $\zeta \in C_0^\infty(\mathbb{R}\times S^1, u^*TM \oplus u^*T^*M)$.

To prove the theorem we need

LEMMA 4.4.5. *Let $u \in \mathcal{P}_{x,y}(\mathbb{R}\times S^1, M)$, where x, y are smooth loops in M, and define $w = g(u)\partial_t u$. Then for every $p > 2$ there exists a constant $c_p > 0$ such that if $\mathcal{D}_u^0$ is onto, then*

$$\boxed{\|\pi_\epsilon{\mathcal{D}_w^\epsilon}^*\zeta\|_p \le c_p\left(\|\pi_\epsilon{\mathcal{D}_w^\epsilon}^*\zeta - {\mathcal{D}_u^0}^*\pi_\epsilon\zeta\|_p + \|D_u^0\pi_\epsilon{\mathcal{D}_w^\epsilon}^*\zeta\|_p\right)}$$

for all $\epsilon > 0$ and $\zeta \in C_0^\infty(\mathbb{R}\times S^1, u^*TM \oplus u^*T^*M)$.

PROOF. (**of lemma 4.4.5**) The proof is standard (cf. [**DS94**] lemma 4.5) and consists of three steps. Choose $q > 1$ such that $1/q + 1/p = 1$.
Step 1 Surjectivity of $\mathcal{D}_u^0$ *implies that there exists a constant c_0 such that for $\xi \in \mathcal{W}^{1,q}$*

$$\|\xi\|_q \le c_0\|{\mathcal{D}_u^0}^*\xi\|_q.$$

This fact may be found for instance in [**Br83**] theorem II.19 for unbounded operators and in [**Ru87**] theorem 4.13 for bounded ones.
Step 2 *There is a constant $c_1 > 0$ such that for $\tilde{\xi} \in \mathcal{W}^{1,p}$*

$$\|{\mathcal{D}_u^0}^*\tilde{\xi}\|_p \le c_1 \sup_{\xi\in\mathcal{W}^{1,q}} \frac{\langle{\mathcal{D}_u^0}^*\tilde{\xi}, {\mathcal{D}_u^0}^*\xi}{\|{\mathcal{D}_u^0}^*\xi\|_q}.$$

As $dim\ Ker\ \mathcal{D}_u^0 < \infty$ we can find a basis $\{e_1, \dots, e_m\}$ of $Ker\ \mathcal{D}_u^0$, which is orthonormal with respect to the L^2-inner product. Choose an element $\hat{\xi} \in L^q$ such that

$$\langle\hat{\xi}, {\mathcal{D}_u^0}^*\tilde{\xi}\rangle = \|{\mathcal{D}_u^0}^*\tilde{\xi}\|_p \quad , \quad \|\hat{\xi}\|_q = 1.$$

That this choice is possible is a consequence of the Hahn-Banach theorem for linear functionals (cf. [**Al92**] Folgerung 4.4). Since $\mathcal{D}_u^0$ is onto there exists a unique $\xi_0 \in \mathcal{W}^{1,q}$ such that

$$\hat{\xi} = {\mathcal{D}_u^0}^*\xi_0 + \sum_{j=1}^m \langle\hat{\xi}, e_j\rangle e_j.$$

This follows from the decomposition $L^q = ran\,\mathcal{D}_u^{0\,*} \oplus Ker\,\mathcal{D}_u^0$ and the fact that $\mathcal{D}_u^0$ onto implies $\mathcal{D}_u^{0\,*}$ injective and so

$$\mathcal{D}_u^{0\,*} : \mathcal{W}^{1,q} \to ran\,\mathcal{D}_u^{0\,*}$$

is a bijection. Now

$$\begin{aligned}
\|\mathcal{D}_u^{0\,*}\tilde{\xi}\|_p &= \langle \hat{\xi}, \mathcal{D}_u^{0\,*}\tilde{\xi}\rangle = \langle \mathcal{D}_u^{0\,*}\xi_0, \mathcal{D}_u^{0\,*}\tilde{\xi}\rangle \\
&= \|\hat{\xi} - \sum_{j=1}^m \langle \hat{\xi}, e_j\rangle e_j\|_q \; \frac{\langle \mathcal{D}_u^{0\,*}\xi_0, \mathcal{D}_u^{0\,*}\tilde{\xi}\rangle}{\|\mathcal{D}_u^{0\,*}\xi_0\|_q} \\
&\le \left(1 + \sum_{j=1}^m \|e_j\|_p \|e_j\|_q\right) \; \frac{\langle \mathcal{D}_u^{0\,*}\xi_0, \mathcal{D}_u^{0\,*}\tilde{\xi}\rangle}{\|\mathcal{D}_u^{0\,*}\xi_0\|_q} \\
&\le c_1 \sup_{\xi \in \mathcal{W}^{1,q}} \; \frac{\langle \mathcal{D}_u^{0\,*}\xi, \mathcal{D}_u^{0\,*}\tilde{\xi}\rangle}{\|\mathcal{D}_u^{0\,*}\xi\|_q}.
\end{aligned}$$

Step 3 We prove the claim. For all $\zeta \in W_\epsilon^{1,p}(\mathbb{R} \times S^1, u^*TM \oplus u^*T^*M)$ and $\xi_0 \in \mathcal{W}^{1,q}$

$$\begin{aligned}
\frac{\langle \mathcal{D}_u^{0\,*}\xi_0, \mathcal{D}_u^{0\,*}\pi_\epsilon\zeta\rangle}{\|\mathcal{D}_u^{0\,*}\xi_0\|_q} &= \frac{\langle \mathcal{D}_u^{0\,*}\xi_0, \mathcal{D}_u^{0\,*}\pi_\epsilon\zeta - \pi_\epsilon\mathcal{D}_w^{\epsilon\,*}\zeta\rangle}{\|\mathcal{D}_u^{0\,*}\xi_0\|_q} + \frac{\langle \xi_0, \mathcal{D}_u^0\pi_\epsilon\mathcal{D}_w^{\epsilon\,*}\zeta\rangle}{\|\mathcal{D}_u^{0\,*}\xi_0\|_q} \\
&\le \|\mathcal{D}_u^{0\,*}\pi_\epsilon\zeta - \pi_\epsilon\mathcal{D}_w^{\epsilon\,*}\zeta\|_p + c_0\|\mathcal{D}_u^0\pi_\epsilon\mathcal{D}_w^{\epsilon\,*}\zeta\|_p
\end{aligned}$$

where in the last estimate we applied step 1. Use step 2 and the former estimate to get

$$\begin{aligned}
\|\mathcal{D}_u^{0\,*}\pi_\epsilon\zeta\|_p &\le c_1 \sup_{\xi_0 \in \mathcal{W}^{1,q}} \frac{\langle \mathcal{D}_u^{0\,*}\xi_0, \mathcal{D}_u^{0\,*}\pi_\epsilon\zeta\rangle}{\|\mathcal{D}_u^{0\,*}\xi_0\|_q} \\
&\le c_1\|\mathcal{D}_u^{0\,*}\pi_\epsilon\zeta - \pi_\epsilon\mathcal{D}_w^{\epsilon\,*}\zeta\|_p + c_0c_1\|\mathcal{D}_u^0\pi_\epsilon\mathcal{D}_w^{\epsilon\,*}\zeta\|_p.
\end{aligned}$$

This implies

$$\begin{aligned}
\|\pi_\epsilon\mathcal{D}_w^{\epsilon\,*}\zeta\|_p &\le \|\pi_\epsilon\mathcal{D}_w^{\epsilon\,*}\zeta - \mathcal{D}_u^{0\,*}\pi_\epsilon\zeta\|_p + \|\mathcal{D}_u^{0\,*}\pi_\epsilon\zeta\|_p \\
&\le (1 + c_1)\|\mathcal{D}_u^{0\,*}\pi_\epsilon\zeta - \pi_\epsilon\mathcal{D}_w^{\epsilon\,*}\zeta\|_p + c_0c_1\|\mathcal{D}_u^0\pi_\epsilon\mathcal{D}_w^{\epsilon\,*}\zeta\|_p.
\end{aligned}$$

□

PROOF. (**of theorem 4.4.4**) Translating surjectivity of $\mathcal{D}_w^\epsilon$ into injectivity of $\mathcal{D}_w^{\epsilon\,*}$, the linear estimate theorem 4.3.2 for $\mathcal{D}_w^{\epsilon\,*}$ together with proposition 4.2.1 lead to the injectivity estimate

$$\|\zeta\|_{1,p,\epsilon} \le c\left(\|\mathcal{D}_w^{\epsilon\,*}\zeta\|_{0,p,\epsilon} + \|\zeta\|_{0,p,\epsilon}\right) \le \tilde{c}\|\mathcal{D}_w^{\epsilon\,*}\zeta\|_{0,p,\epsilon}.$$

Throughout we will use the assumption $\kappa_p \in (0,1)$ in between the lines. Apply the linear estimate for $\mathcal{D}_w^\epsilon$ theorem 4.3.2 to $\mathcal{D}_w^{\epsilon\,*}\zeta$

$$\|\mathcal{D}_w^{\epsilon\,*}\zeta\|_{1,p,\epsilon} \le c_0\left(\epsilon^2\|\mathcal{D}_w^\epsilon\mathcal{D}_w^{\epsilon\,*}\zeta\|_{0,p,\epsilon} + \|\mathcal{D}_w^{\epsilon\,*}\zeta\|_{0,p,\epsilon}\right),$$

then it remains to estimate the last term. With $\zeta^* = (\xi^*, \eta^*) = \mathcal{D}_w^{\epsilon\,*}\zeta$ we get

$$\begin{aligned}\|\mathcal{D}_w^{\epsilon\,*}\zeta\|_{0,p,\epsilon} &\le \|\mathcal{D}_w^{\epsilon\,*}\zeta - \iota\pi_\epsilon\mathcal{D}_w^{\epsilon\,*}\zeta\|_{0,p,\epsilon} + \|\iota\pi_\epsilon\mathcal{D}_w^{\epsilon\,*}\zeta\|_{0,p,\epsilon}\\ &\le c_1\epsilon\|\mathcal{D}_w^\epsilon\mathcal{D}_w^{\epsilon\,*}\zeta\|_{0,p,\epsilon} + c_1\epsilon^{\min\{1-\kappa_p,1/2\}}\|\mathcal{D}_w^{\epsilon\,*}\zeta\|_{0,p,\epsilon}\\ &\quad + \|\pi_\epsilon\mathcal{D}_w^{\epsilon\,*}\zeta\|_p + \epsilon\|\nabla_t\pi_\epsilon(\xi^*,\eta^*)\|_p\\ &\le c_1\epsilon\|\mathcal{D}_w^\epsilon\mathcal{D}_w^{\epsilon\,*}\zeta\|_{0,p,\epsilon} + c_1\epsilon^{\min\{1-\kappa_p,1/2\}}\|\mathcal{D}_w^{\epsilon\,*}\zeta\|_{0,p,\epsilon}\\ &\quad + \|\pi_\epsilon\mathcal{D}_w^{\epsilon\,*}\zeta\|_p + c_2\epsilon^{1/2}\|\xi^*\|_p + c_2\epsilon^2\|\eta^*\|_p\\ &\le c_1\epsilon\|\mathcal{D}_w^\epsilon\zeta^*\|_{0,p,\epsilon} + (c_1+c_2)\epsilon^{\min\{1-\kappa_p,1/2\}}\|\zeta^*\|_{0,p,\epsilon}\\ &\quad + \|\pi_\epsilon\zeta^*\|_p.\end{aligned}$$

In the 2^{nd} inequality we applied lemma 4.2.2 to $\mathcal{D}_w^\epsilon$. The definition of π_ϵ and lemma 4.2.4 imply the 3^{rd} inequality. For $\epsilon_0 > 0$ sufficiently small this implies

$$\|\mathcal{D}_w^{\epsilon\,*}\zeta\|_{0,p,\epsilon} \le c\epsilon\|\mathcal{D}_w^\epsilon\mathcal{D}_w^{\epsilon\,*}\zeta\|_{0,p,\epsilon} + \|\pi_\epsilon\mathcal{D}_w^{\epsilon\,*}\zeta\|_p.$$

Use lemma 4.4.5 to estimate the remaining term

$$\begin{aligned}\|\pi_\epsilon\mathcal{D}_w^{\epsilon\,*}\zeta\|_p &\le c_3\|\pi_\epsilon\mathcal{D}_w^{\epsilon\,*}\zeta - \mathcal{D}_u^{0\,*}\pi_\epsilon\zeta\|_p + c_3\|\mathcal{D}_u^0\pi_\epsilon(\mathcal{D}_w^{\epsilon\,*}\zeta)\|_p\\ &\le c_3c_4\epsilon\|\mathcal{D}_w^{\epsilon\,*}\zeta\|_{0,p,\epsilon} + c_3c_4\epsilon^{\min\{1-\kappa_p,1/2\}}\|\zeta\|_{0,p,\epsilon}\\ &\quad + c_3\|\mathcal{D}_u^0\pi_\epsilon(\mathcal{D}_w^{\epsilon\,*}\zeta) - \pi_\epsilon\mathcal{D}_w^\epsilon(\mathcal{D}_w^{\epsilon\,*}\zeta)\|_p\\ &\quad + c_3\|\pi_\epsilon\mathcal{D}_w^\epsilon(\mathcal{D}_w^{\epsilon\,*}\zeta)\|_p\\ &\le 2c_3c_4\epsilon^{\min\{1-\kappa_p,1/2\}}\|\mathcal{D}_w^{\epsilon\,*}\zeta\|_{0,p,\epsilon} + c_3\|\pi_\epsilon\mathcal{D}_w^\epsilon(\mathcal{D}_w^{\epsilon\,*}\zeta)\|_p\\ &\quad + c_3c_4\epsilon\|\mathcal{D}_w^\epsilon\mathcal{D}_w^{\epsilon\,*}\zeta\|_{0,p,\epsilon} + c_3c_4c_5\epsilon^{\min\{1-\kappa_p,1/2\}}\|\pi_\epsilon\mathcal{D}_w^{\epsilon\,*}\zeta\|_p\end{aligned}$$

where the 2^{nd} inequality follows from lemma 4.2.2 and the third from proposition 4.2.1 and again lemma 4.2.2. Choosing $\epsilon_0 > 0$ sufficiently small we may incorporate the last term into the left hand side and combine the result with the former estimate to obtain

$$\begin{aligned}\|\xi^*\|_p &\le \|\mathcal{D}_w^{\epsilon\,*}\zeta\|_{0,p,\epsilon}\\ &\le c\epsilon\|\mathcal{D}_w^\epsilon\mathcal{D}_w^{\epsilon\,*}\zeta\|_{0,p,\epsilon} + c\|\pi_\epsilon\mathcal{D}_w^\epsilon\mathcal{D}_w^{\epsilon\,*}\zeta\|_p\end{aligned}\tag{48}$$

and therefore

$$\|\mathcal{D}_w^{\epsilon\,*}\zeta\|_{1,p,\epsilon} \le c\epsilon\|\mathcal{D}_w^\epsilon\mathcal{D}_w^{\epsilon\,*}\zeta\|_{0,p,\epsilon} + c\|\pi_\epsilon\mathcal{D}_w^\epsilon\mathcal{D}_w^{\epsilon\,*}\zeta\|_p.$$

This proves the first and third assertions. The fourth one then follows from the third one by lemma 4.2.3.

To prove the estimate for $\|\eta^*\|_p$ apply lemma 4.2.3 as well as the definition of π_ϵ and lemma 4.2.4 in the 2^{nd} inequality to obtain

$$\begin{aligned}
\|\eta^*\|_p &\leq \|\eta^* - g(u)\nabla_t \pi_\epsilon \zeta^*\|_p + \|\nabla_t \pi_\epsilon \zeta^*\|_p \\
&\leq c_5 \|\eta^* - g(u)\nabla_t \xi^*\|_p + c_5 \epsilon^{1/2} \|\nabla_t \eta^*\|_p \\
&\quad + (c_6/2)\epsilon^{-1/2}\|\xi^*\|_p + (c_6/2)\epsilon \|\eta^*\|_p \\
&\leq c_7 \epsilon^{1/2} \|\mathcal{D}_w^\epsilon \zeta^*\|_{0,p,\epsilon} + c_7(\epsilon^{1/2-\kappa_p} + \epsilon^{-1/2})\|\xi^*\|_p \\
&\quad + c_7(\epsilon^{3/2-\kappa_p} + \epsilon)\|\eta^*\|_p \\
&\leq c_8 \epsilon^{1/2} \|\mathcal{D}_w^\epsilon \zeta^*\|_{0,p,\epsilon} + c_8 \epsilon^{-1/2} \|\pi_\epsilon \mathcal{D}_w^\epsilon \zeta^*\|_p \\
&\quad + c_8(\epsilon^{3/2-\kappa_p} + \epsilon)\|\eta^*\|_p.
\end{aligned}$$

Here we used lemma 4.2.5 for $\mathcal{D}_w^\epsilon$ in the 3^{rd} inequality and estimate (48) for ξ^* in the 4^{th} inequality. Now incorporate the η^*-term into the left hand side for $\epsilon_0 > 0$ sufficiently small. Moreover, using lemma 4.2.5 for $\mathcal{D}_w^\epsilon$ in the 2^{nd} and the above estimates for ξ^* and η^* in the 3^{rd} inequality we get

$$\begin{aligned}
\|\nabla_t \xi^*\|_p &\leq \|\nabla_t \xi^* - g(u)^{-1}\eta^*\|_p + \|\eta^*\|_p \\
&\leq c_9 \epsilon \|\mathcal{D}_w^\epsilon \zeta^*\|_{0,p,\epsilon} + c_9 \epsilon^{1-\kappa_p}\|\xi^*\|_p + c_9\|\eta^*\|_p \\
&\leq c_{10}\epsilon^{1/2}\|\mathcal{D}_w^\epsilon \zeta^*\|_{0,p,\epsilon} + c_{10}\epsilon^{-1/2}\|\pi_\epsilon \mathcal{D}_w^\epsilon \zeta^*\|_p.
\end{aligned}$$

□

CHAPTER 5

Quadratic estimates

Deriving the quadratic estimates, although mainly a tedious technical process, provides an essential ingredient to carry out the Newton type method in chapter 1. The quality of the quadratic estimates obtained here determines the qualitative results, with respect to powers of ϵ, on the size of the existence and uniqueness neighborhoods of the zeroes detected by the iteration process. In order to get optimal results we use elements of Riemannian geometry summarized in section A.1.

In section 5.1 we calculate the fundamental quadratic estimate needed in the initial step of the Newton method. It turns out to be necessary to preserve the combination $g\nabla_t\xi - \eta$ in the term with coefficient ϵ^{-2}, as well as to do the estimates for both components of $\mathcal{F}^{triv}_{\epsilon,u_0}(\zeta) - \mathcal{F}^{triv}_{\epsilon,u_0}(0) - d\mathcal{F}^{triv}_{\epsilon,u_0}(0)\circ\zeta$ *separately*.

This holds true also in sections 5.2 and 5.3 where we get the estimates I and II applied in the induction step of the iteration. Note that in this whole chapter – due to the nonlinearities – we heavily rely on the results of the geometric analysis in section A.1 about the exponential and parallel transport maps.

As a first step we will produce pointwise estimates in local coordinates and integration will then lead to estimates with respect to L^p- and L^∞-norms, where we shall throw the former ones on terms involving derivatives. So let us construct a new norm$\|\cdot\|'_{0,p,\epsilon}$ on $C_0^\infty(\mathbb{R}\times S^1, u^*TM \oplus u^*T^*M)$, $u \in \mathcal{P}^\infty_{x,y}$, using a cover of M by local coordinate charts. This new norm, although dependent on the choice of the cover, is equivalent to the norm $\|\cdot\|_{0,p,\epsilon}$ introduced previously.

Let ι_M be the injectivity radius of the exponential map of the Riemannian manifold (M,g). As M is compact it follows $\iota_M > 0$. Let $\{U_i\}_{i=1}^N$ be a cover of M by open sets, which has the following technical properties: U_i is contained for any $i \in \{1,\dots,N\}$ in some coordinate chart (V_i,φ_i) of M and $exp_q\xi$ and $exp_q\eta$ are contained in the same coordinate chart for all $q \in U_i$ and $\xi \in T_qM$, $\eta \in T_q^*M$ of norm less than a constant

$$\iota'_M \in (0,\iota_M). \tag{49}$$

Clearly ι'_M may be choosen independently of i. (A cover as described above can be constructed for instance via a partition of unity subordinate to a finite coordinate cover of M. The interiors of the supports of the partition of unity serve as the cover, if we choose ι'_M sufficiently small.)

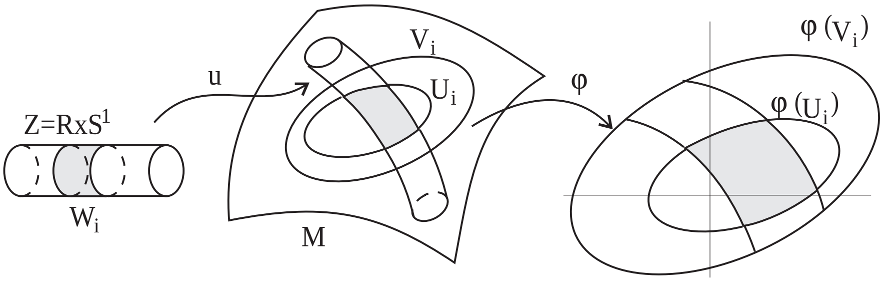

FIGURE 5.1. Construction of the norm $\|\cdot\|'_{0,p,\epsilon}$

DEFINITION 5.0.6. Using the same notation as above define $W_i = u^{-1}(u(\mathbb{R}\times S^1)\cap U_i) \subset \mathbb{R}\times S^1$. For $\epsilon > 0$ and $(\xi,\eta) \in C_0^\infty(\mathbb{R}\times S^1, u^*TM \oplus u^*T^*M)$ with $|\xi(s,t)|, |\eta(s,t)| < \iota'_M$ for all $(s,t) \in \mathbb{R}\times S^1$ define

$$\|(\xi,\eta)\|'_{0,p,\epsilon}{}^p = \sum_{i=1}^N \int_{W_i} \left(|\vec{\xi}(s,t)|^p + \epsilon^p|\vec{\eta}(s,t)|^p\right)\, dtds$$

where $\vec{\xi}$ resp. $\vec{\eta}$ denote the representatives of ξ resp. η in the local coordinates on U_i.

LEMMA 5.0.7. *$\|\cdot\|'_{0,p,\epsilon}$ and $\|\cdot\|_{0,p,\epsilon}$ are equivalent norms on $C_0^\infty(\mathbb{R}\times S^1, u^*TM \oplus u^*T^*M)$, i.e. there exists a constant $C > 0$ such that*

$$\frac{1}{C}\|(\xi,\eta)\|'_{0,p,\epsilon} \le \|(\xi,\eta)\|_{0,p,\epsilon} \le C\|(\xi,\eta)\|'_{0,p,\epsilon}$$

for all $(\xi,\eta) \in C_0^\infty(\mathbb{R}\times S^1, u^*TM \oplus u^*T^*M)$.

PROOF. Let ν be the maximal number of sets W_i having nonempty intersection, then $C = \nu^{1/p}$ will do the job. □

Hence we may choose to work in local coordinates to get estimates for the derivatives of $\mathcal{F}^{triv}_{\epsilon,u}$. Although these derivatives do not have in general an invariant meaning, the above lemma justifies this local approach, which allows us to use simply calculus on $\mathbb{R}^n$.

REMARK 5.0.8. The quadratic estimates with respect to L^p- and L^∞-norms are an immediate consequence of the corresponding pointwise estimates (by integrating them). By the equivalence of norms it suffices to prove the quadratic estimates in the norm $\|\cdot\|'_{0,p,\epsilon}$ and therefore we may restrict to a coordinate chart $(U_i \subset V_i, \varphi_i)$ as described above.
Let us consider for instance the following term of pointwise estimate II, where $|\cdot|$ denotes $|\cdot|_{T_{u(s,t)}M}$

$$\begin{aligned}|\vec{T}_1(s,t) + \vec{T}_2(s,t)| + \epsilon|\vec{T}_3(s,t) + \vec{T}_4(s,t)| \\ \le c(\vec{X}(s,t))\, |\vec{\xi}(s,t)| \cdot |\partial_s\vec{X}(s,t)|.\end{aligned}$$

Take this inequality to the power p and integrate (s,t) over

$$W_i = u^{-1}(u(\mathbb{R}\times S^1)\cap U_i)$$

to obtain

$$\begin{aligned}
\|\vec{T}_1+\vec{T}_2\|^p_{L^p(W_i,\mathbb{R}^n)} &+ \epsilon^p\|\vec{T}_3+\vec{T}_4\|^p_{L^p(W_i,\mathbb{R}^n)}\\
&\le \int_{W_i} c(\vec{X}(s,t))^p\,|\vec{\xi}(s,t)|^p\cdot|\partial_s\vec{X}(s,t)|^p\,dtds\\
&\le c_{p,i}\int_{W_i}|\vec{\xi}(s,t)|^p\Big(|\vec{\nabla}_s\vec{X}(s,t)|^p+|\vec{X}(s,t)|^p\Big)dtds\\
&\le c_{p,i}\|\xi\|^p_{L^\infty(W_i,\mathbb{R}^n)}\Big(\|\vec{\nabla}_s\vec{X}\|^p_{L^p(W_i,\mathbb{R}^n)}+\|\vec{X}\|^p_{L^p(W_i,\mathbb{R}^n)}\Big)
\end{aligned}$$

where we used in the 2^{nd} inequality that

$$\begin{aligned}
|\partial_s\vec{X}(s,t)|^p &\le 2^{p-1}\left|\big(\partial_s X^k(s,t)+\Gamma^k_{ij}|_{u(s,t)}\partial_s u^i(s,t)\,X^j(s,t)\big)\partial_k\right|^p\\
&\quad+2^{p-1}\left|\Gamma^k_{ij}|_{u(s,t)}\partial_s u^i(s,t)\,X^j(s,t)\big)\partial_k\right|^p\\
&\le const(i,p)\Big(|\vec{\nabla}_s\vec{X}(s,t)|^p+|\vec{X}(s,t)|^p\Big).
\end{aligned}$$

We observe that any partial derivative in the pointwise estimates gives rise to a covariant derivative plus a corresponding zero order term in the quadratic estimates. Moreover, the distribution of L^p- and L^∞-norms to the factors of products is clearly motivated by the intention to optimize the estimates in the Newton iteration. The strategy will be to throw L^p-norms on the terms involving derivatives.

The next lemma is the major technical tool in the proof of the pointwise estimates.

Lemma 5.0.9. *Let $f\in C^2(\mathbb{R}^n,\mathbb{R}^n)$, $n\in\mathbb{N}$. Then for any $\delta>0$ there exists $c_\delta\in C^0(\mathbb{R}^n,\mathbb{R}^+)$ such that*

$$\begin{aligned}
&i)\quad |f(X+\xi)-f(X)|\le c_\delta(\xi)\,|\xi|\\
&ii)\quad |f(X+\xi)-f(X)-df(X)\circ\xi|\le c_\delta(\xi)\,|\xi|^2
\end{aligned}$$

for all X with $|X|\le\delta$ and all $\xi\in\mathbb{R}^n$.

Proof. We only prove $ii)$ as $i)$ follows quite similarly. It suffices to prove the estimate for a component f^i of f, i.e.

$$|a^i|_{\mathbb{R}}\overset{def}{=}|f^i(X+\xi)-f^i(X)-df^i(X)\circ\xi|_{\mathbb{R}}\le c_{\delta,i}(\xi)\,|\xi|^2_{\mathbb{R}^n}.$$

The general result may then be obtained as follows

$$\begin{aligned}
|(a^1,\dots,a^n)| &= \Big(\sum_{i=1}^n|a^i|^2_{\mathbb{R}}\Big)^{1/2}\\
&\le \Big(\sum_{i=1}^n c_{\delta,i}(\xi)^2\Big)^{1/2}|\xi|^2_{\mathbb{R}^n}\overset{def}{=}c_\delta(\xi)\,|\xi|^2_{\mathbb{R}^n}.
\end{aligned}$$

Now

$$\begin{aligned}
&|f^i(X+\xi) - f^i(X) - df^i(X)\circ\xi|_{\mathbb{R}} \\
&\qquad = \left| \int_0^1 df^i(X+\tau\xi)\circ\xi - df^i(X)\circ\xi \, d\tau \right|_{\mathbb{R}} \\
&\qquad = \left| \int_0^1 \int_0^\tau d^2 f^i(X+\sigma\xi)\circ(\xi,\xi)\, d\sigma\, d\tau \right|_{\mathbb{R}} \\
&\qquad \le \int_0^1 \int_0^\tau \|d^2 f^i(X+\sigma\xi)\|_{\mathcal{L}(\mathbb{R}^n\times\mathbb{R}^n,\mathbb{R})}\, d\sigma\, d\tau \cdot |\xi|^2_{\mathbb{R}^n} \\
&\qquad \le \frac{1}{2} \sup_{0\le\tau\le 1} \|d^2 f^i(X+\tau\xi)\|_{\mathcal{L}(\mathbb{R}^n\times\mathbb{R}^n,\mathbb{R})} \cdot |\xi|^2_{\mathbb{R}^n} \\
&\qquad \le \frac{1}{2} \sup_{|X|\le\delta} \sup_{0\le\tau\le 1} \|d^2 f^i(X+\tau\xi)\|_{\mathcal{L}(\mathbb{R}^n\times\mathbb{R}^n,\mathbb{R})} \cdot |\xi|^2_{\mathbb{R}^n} \\
&\qquad \overset{def}{=} c_{\delta,i}(\xi)\, |\xi|^2_{\mathbb{R}^n}.
\end{aligned}$$

□

REMARK 5.0.10. Throughout the subsequent proofs we drop the arrows (indicating $\mathbb{R}^n$-valued functions) and the argument (s,t) from our notation, so that for instance ξ denotes $\vec{\xi}(s,t) = (\xi^1(s,t),\dots,\xi^n(s,t))$. Define

$$|\xi| := |\xi|_{T_{u_0}M} = g_{ij}|_{u_0(s,t)}\xi^i(s,t)\xi^j(s,t).$$

and

$$\begin{aligned}
w_0 &= g(u_0)\partial_t u_0 \\
a(u_0,\xi) &= exp_{u_0}\xi \qquad (50)\\
b(u_0,\xi,\eta) &= \mathcal{T}^{-1}(\xi)^*(w_0+\eta),
\end{aligned}$$

then

$$\begin{aligned}
\partial_t a(u_0,\xi) &= \partial_1 a(u_0,\xi)\circ\partial_t u_0 + \partial_2 a(u_0,\xi)\circ\partial_t\xi \qquad (51)\\
\partial_t b(u_0,\xi,\eta) &= d\mathcal{T}^{-1}|^*_\xi(\partial_t\xi, w_0+\eta) + \mathcal{T}^{-1}(\xi)^*\circ(\partial_t w_0 + \partial_t\eta)
\end{aligned}$$

and similarly for ∂_s. The notation $d\mathcal{T}^{-1}|^*_\xi(\partial_t\xi, w_0+\eta)$ means that $d\mathcal{T}^{-1}|^*_\xi(\cdot,\cdot)$ is bilinear and is given by $(d\mathcal{T}^{-1}|_\xi\circ\partial_t\xi)^*\circ(w_0+\eta)$, strictly speaking. Moreover,

$$\begin{aligned}
\tfrac{d}{d\tau}\Big|_0 \partial_s a(u, X+\tau\xi) &= \partial_s(\partial_2 a(u,X)\circ\xi) \qquad (52)\\
&= \partial_1\partial_2 a(u,X)\circ(\xi,\partial_s u) + \partial_2\partial_2 a(u,X)\circ(\xi,\partial_s X) \\
&\quad + \partial_2 a(u,X)\circ\partial_s\xi.
\end{aligned}$$

5.1. The fundamental quadratic estimate

THEOREM 5.1.1. *(**Fundamental quadratic estimate**) Let $p > 2$ and u_0 be an element of the moduli space $\mathcal{M}^0(x^-, x^+)$, where x^-, $x^+ \in Crit\,\mathcal{I}_V$, and denote*

$$\mathcal{F}^{triv}_{\epsilon,u_0}(\zeta) - \mathcal{F}^{triv}_{\epsilon,u_0}(0) - d\mathcal{F}^{triv}_{\epsilon,u_0}(0)\circ\zeta = \begin{pmatrix} F_1 \\ F_2 \end{pmatrix}.$$

Let $\alpha = t$ or $\alpha = s$. Then there exists a constant $c_p > 0$ such that

$$\begin{aligned} \|F_1\|_p &\le c_p\|\xi\|_\infty\Big(\|\xi\|_p + \|\nabla_t\xi\|_p + \underline{\|\nabla_s\xi\|_p} + \|\eta\|_p + \underline{\|\nabla_t\eta\|_p}\Big) \\ \|F_2\|_p &\le \frac{c_p}{\epsilon^2}\,\|\xi\|_\infty\Big(\underline{\|g\nabla_t\xi - \eta\|_p} + \underline{\|\xi\|_p}\Big) \\ &\quad + c_p\,\|\xi\|_\infty\Big(\|\nabla_s\xi\|_p + \|\eta\|_p + \|\nabla_s\eta\|_p\Big) + c_p\,\|\eta\|_\infty\|\nabla_s\xi\|_p \end{aligned}$$

and

$$\begin{aligned} &\|\nabla_\alpha F_1\|_p \\ &\quad\le c_p\,\|\xi\|_\infty\Big(\|\xi\|_p + \|\nabla_t\xi\|_p + \|\nabla_s\xi\|_p + \|\eta\|_p + \|\nabla_t\eta\|_p + \underline{\|\nabla_\alpha\eta\|_p} \\ &\qquad\quad + \underline{\|\nabla_\alpha\nabla_t\xi\|_p} + \|\nabla_\alpha\nabla_s\xi\|_p\|\xi\|_\infty\Big) \\ &\qquad + c_p\,\|\eta\|_\infty\Big(\|\nabla_t\xi\|_p + \underline{\|\nabla_\alpha\xi\|_p}\Big) + c_p\,\underline{\|\nabla_t\xi\|_\infty\|\nabla_\alpha\xi\|_p} \\ &\|\nabla_\alpha F_2\|_p \\ &\quad\le \frac{c_p}{\epsilon^2}\,\|\xi\|_\infty\Big(\|\nabla_\alpha\nabla_t\xi - g^{-1}\nabla^*_\alpha\eta\|_p + \|\xi\|_p + \|\nabla_t\xi\|_p + \|\eta\|_p \\ &\qquad\quad + \underline{\|\nabla_\alpha\xi\|_p} + \|\xi\|_\infty(\|\nabla_\alpha\eta\|_p + \|\nabla_\alpha\nabla_t\xi\|_p)\Big) \\ &\qquad + \frac{c_p}{\epsilon^2}\,\|\nabla_t\xi - g^{-1}\eta\|_p\Big(\|\xi\|_\infty + \underline{\|\nabla_\alpha\xi\|_\infty}\Big) \\ &\qquad + c_p\|\xi\|_\infty\Big(\|\nabla_s\xi\|_p + \|\nabla_s\eta\|_p + \|\nabla_\alpha\eta\|_p + \|\nabla_\alpha\nabla_s\xi\|_p + \|\nabla_\alpha\nabla_s\eta\|_p\Big) \\ &\qquad + c_p\|\eta\|_\infty\Big(\|\nabla_\alpha\xi\|_p + \|\nabla_s\xi\|_p + \|\nabla_\alpha\nabla_s\xi\|_p\Big) \\ &\qquad + c_p\|\nabla_\alpha\xi\|_\infty\Big(\|\nabla_s\xi\|_p + \|\nabla_s\eta\|_p\Big) + c_p\|\nabla_\alpha\eta\|_p\|\nabla_s\xi\|_\infty \end{aligned}$$

for $\epsilon \in (0,1]$ and $\zeta = (\xi,\eta) \in C_0^\infty(\mathbb{R}\times S^1, u_0^*TM \oplus u_0^*T^*M)$ with $\|\xi\|_\infty \le \iota'_M/2$ and $\|\eta\|_\infty + \|\nabla_t\xi\|_\infty + \|\nabla_\alpha\xi\|_\infty \le \sqrt{c_p}/2$.

The theorem follows from the pointwise estimate 5.1.3 via integration as described in remark 5.0.8. Note that the condition on $\|\xi\|_\infty$ is necessary in order for the local constructions to be well defined. The other L^∞-conditions only serve to simplify the expressions. We underlined the terms which determine the rate of convergence in the Newton method.

REMARK 5.1.2. The constant c appearing in lemma 5.1.3 depends linearly on $\partial_t u_0$, $\partial_s u_0$ and in the estimate for $\nabla_\alpha F_i$ even on $\partial_\alpha\partial_t u_0$. In order to derive theorem 5.1.1 via integration as described in remark 5.0.8 we need continuity of all these partial derivatives of u_0 and some information on their behavior for $s \to \mp\infty$. By definition of $\mathcal{M}^0(x^-, x^+)$ any element u_0 is C^∞ anyway and there are prescribed boundary conditions $x^\mp$; however, at some point one needs to justify this definition by working out an appropriate regularity theory for the solutions of the parabolic PDE. This will be carried out elsewhere.

LEMMA 5.1.3. (**Fundamental pointwise estimate**) *Let u_0 be an element of the moduli space $\mathcal{M}^0(x^-, x^+)$, where x^-, $x^+ \in Crit\,\mathcal{I}_V$, and denote*

$$\mathcal{F}^{triv}_{\epsilon,u_0}(\zeta) - \mathcal{F}^{triv}_{\epsilon,u_0}(0) - d\mathcal{F}^{triv}_{\epsilon,u_0}(0)\circ\zeta = \begin{pmatrix} F_1 \\ F_2 \end{pmatrix}$$

for $\zeta = (\xi, \eta) \in C_0^\infty(\mathbb{R} \times S^1, u_0^*TM \oplus u_0^*T^*M)$ with $\|\xi\|_\infty \le \iota'_M/2$, where ι'_M is the constant introduced in (49). Then in a local coordinate chart $(U_i \subset V_i, \varphi_i)$ as in figure 5.1 the following pointwise estimates hold: There exists a continuous function $c = c(\vec{\xi}) \ge 0$ – also depending continuously on $\vec{u}_0$ and its partial derivatives of first and second order – such that

$$\begin{aligned} |\vec{F}_1| &\le c\,|\vec{\xi}|\Big(|\vec{\xi}| + |\partial_t\vec{\xi}| + |\partial_s\vec{\xi}| + |\vec{\eta}| + |\partial_t\vec{\eta}|\Big) + c\,|\vec{\eta}|\cdot|\partial_t\vec{\xi}|\cdot|\vec{\xi}| \\ |\vec{F}_2| &\le c\,\epsilon^{-2}|\vec{\xi}|\Big(|\vec{g}\,\vec{\nabla}_t\vec{\xi} - \vec{\eta}| + |\vec{\xi}|\,(|\vec{\eta}| + \|R(\cdot,\dot{u}_0)\cdot\|)\Big) \\ &\quad + c\,|\vec{\xi}|\Big(|\vec{\xi}| + |\partial_s\vec{\xi}| + |\vec{\eta}| + |\partial_s\vec{\eta}|\Big) + c\,|\vec{\eta}|\cdot|\partial_s\vec{\xi}| \end{aligned}$$

and for $\alpha = t$ or $\alpha = s$

$$\begin{aligned} |\vec{\nabla}_\alpha\vec{F}_1| &\le c\,|\vec{\xi}|\Big(|\vec{\xi}| + |\partial_t\vec{\xi}| + |\partial_\alpha\vec{\xi}| + |\partial_s\vec{\xi}|\big(1 + |\partial_\alpha\vec{\xi}|\big) + |\vec{\eta}| + |\partial_t\vec{\eta}|\big(1 + |\partial_\alpha\vec{\xi}|\big) \\ &\qquad + |\partial_\alpha\vec{\eta}|\big(1 + |\partial_t\vec{\xi}|\big) + |\partial_\alpha\partial_t\vec{\xi}|\big(1 + |\vec{\eta}|\big) + |\partial_\alpha\partial_s\vec{\xi}|\cdot|\vec{\xi}|\Big) \\ &\quad + c\,|\vec{\eta}|\cdot|\partial_\alpha\vec{\xi}|\Big(1 + |\partial_t\vec{\xi}|\Big) + c\,|\partial_t\vec{\xi}|\Big(|\vec{\eta}| + |\partial_\alpha\vec{\xi}|\Big) \\ |\vec{\nabla}_\alpha\vec{F}_2| &\le c\epsilon^{-2}|\xi|\Big(|\vec{\nabla}_\alpha\vec{\nabla}_t\xi - g^{-1}\vec{\nabla}^*_\alpha\eta| + |\eta| + |\partial_t\xi| + |\partial_\alpha\xi|\big(1 + |\eta| + |\partial_t\xi|\big)\Big) \\ &\quad + c\epsilon^{-2}|\vec{\nabla}_t\xi - g^{-1}\eta|\Big(|\xi| + |\partial_\alpha\xi|\Big) + c\epsilon^{-2}|\xi|^2\Big(1 + |\partial_\alpha\eta| + |\partial_\alpha\partial_t\xi|\Big) \\ &\quad + c|\xi|\Big(|\partial_s\xi| + |\partial_\alpha\eta| + |\partial_s\eta| + |\partial_\alpha\partial_s\xi| + |\partial_\alpha\partial_s\eta|\Big) + c|\partial_s\xi|\cdot|\partial_\alpha\eta| \\ &\quad + c|\eta|\Big(|\partial_\alpha\xi| + |\partial_s\xi| + |\partial_\alpha\partial_s\xi|\Big) + c|\partial_\alpha\xi|\Big(|\partial_s\xi|\big(1 + |\eta|\big) + |\partial_s\eta|\Big) \end{aligned}$$

for $\epsilon \in (0,1]$, where $|\cdot| = |\cdot|_{T_{u_0(s,t)}M}$ and an arrow on top of an object indicates that it is represented in local coordinates and evaluated at (s,t).

Recall the definition of $\mathcal{F}^{triv}_{\epsilon,u}$: Pick a smooth cylinder $u \in \mathcal{P}_{x,y}$, x, y smooth loops in M, and set

$$w = g(u)\partial_t u,$$

then

$$\mathcal{F}^{triv}_{\epsilon,u}(X,Y) = \begin{pmatrix} \mathcal{T}^{-1}(X) & 0 \\ 0 & \mathcal{T}(X)^* \end{pmatrix} \circ \mathcal{F}_\epsilon \begin{pmatrix} a(u,X) \\ b(u,X,Y) \end{pmatrix}$$
$$= \begin{pmatrix} \mathcal{T}^{-1}(X)\big(\partial_s a - g^{-1}(a)\nabla_t^* b - \nabla V_t(a)\big) \\ \mathcal{T}(X)^*\big(\nabla_s^* b + \epsilon^{-2} g(a)\partial_t a - \epsilon^{-2} b\big) \end{pmatrix}$$

where

$$a(u,X) = exp_u X$$
$$b(u,X,Y) = \mathcal{T}^{-1}(X)^*(w+Y)$$

and $\mathcal{T}(X)$ is parallel transport along the geodesic

$$[0,1] \to U_i \subset V_i \subset M$$
$$\sigma \mapsto exp_{u(s,t)}\sigma X(s,t).$$

Note that $\mathcal{T}(X)$ was denoted $\mathcal{T}(1,0)$ in appendix A and so in local coordinates

$$(\mathcal{T}(X)^*)_l^k = (\mathcal{T}^*(0,1))_l^k = (\mathcal{T}(1,0))_l^k = \mathcal{T}(X)_l^k.$$

Moreover, we define $\mathcal{T}^{-1}(X) = \mathcal{T}(X)^{-1}$.

In the following proof we will use several times lemma 5.0.9 as well as the notation introduced in remark 5.0.10 below.

Proof. (**of lemma 5.1.3 – fundamental pointwise estimate**) Let us denote

$$\mathcal{F}^{triv}_{\epsilon,u_0}(\xi,\eta) = \begin{pmatrix} \mathcal{T}^{-1}(\xi)\; f_1(\xi,\eta) \\ \mathcal{T}(\xi)^*\; f_2(\xi,\eta) \end{pmatrix}$$

where

$$f_1(\xi,\eta) = \partial_s a(u_0,\xi) - g^{-1}|_{a(u_0,\xi)}\nabla_t^* b(u_0,\xi,\eta) - \nabla V_t|_{a(u_0,\xi)}$$
$$f_2(\xi,\eta) = \nabla_s^* b(u_0,\xi,\eta) + \epsilon^{-2} g|_{a(u_0,\xi)}\partial_t a(u_0,\xi) - \epsilon^{-2} b(u_0,\xi,\eta).$$

Estimates for F_1 and F_2

The estimate for the first component F_1 is less subtle than the one for F_2 as there are no terms containing factors ϵ^{-2}. For F_1 we may simply apply the corresponding estimate of lemma 5.2.2 (pointwise estimate I) with $X=0$ and $Y=0$. A more delicate analysis is required in order to estimate

$$\begin{aligned} |F_2| &= |\mathcal{T}(\xi)^*\, f_2(\xi,\eta) - f_2(0,0) - \tfrac{d}{d\tau}\big|_0(\mathcal{T}(\tau\xi)^*\, f_2(\tau\xi,\tau\eta))| \\ &= |\mathcal{T}(\xi)^*\, f_2(\xi,\eta) - \nabla_s^*(g|_{u_0}\partial_t u_0) - d\mathcal{T}|_0^*(\xi, \nabla_s^*(g|_{u_0}\partial_t u_0)) \\ &\quad - \partial_1 f_2(0,0)\xi - \partial_2 f_2(0,0)\eta| \\ &= \Big|\Big(\mathcal{T}(\xi)^* - \mathbb{1} - d\mathcal{T}|_0^*\,\xi\Big)\, f_2(\xi,\eta) \\ &\quad + \Big(f_2(\xi,\eta) - \nabla_s^*(g|_{u_0}\partial_t u_0) - \partial_1 f_2(0,0)\,\xi - \partial_2 f_2(0,0)\,\eta\Big) \\ &\quad + d\mathcal{T}|_0^*\Big(\xi, f_2(\xi,\eta) - \nabla_s^*(g|_{u_0}\partial_t u_0)\Big)\Big| \\ &= |IV + V + VI|. \end{aligned} \tag{53}$$

Note that the last equality defines terms IV, V and VI. Moreover, we have added twice zero in the steps above. As the ϵ^{-2}-terms are the worst ones in order to get estimates with highest possible powers of ϵ and, on the other hand, we have strong estimates for $g(u_0)\nabla_t\xi - \eta$, the *crucial point* in all what follows is to keep those differences together rather than treating $\nabla_t\xi$ and η separately. Use formula (72) for $\nabla_s^* b(u_0,\xi,\eta)$ to obtain

(54)

$$\begin{aligned} VI = d\mathcal{T}|_0^*\Big(&\xi\ ,\ d\mathcal{T}^{-1}|_\xi^*(\partial_s\xi, g|_{u_0}\partial_t u_0 + \eta) + \mathcal{T}^{-1}(\xi)^*\big(\partial_s(g|_{u_0}\partial_t u_0) + \partial_s\eta\big) \\ &- \Gamma|_{a(u_0,\xi)}\Big(\partial_1 a(u_0,\xi)\ \partial_s u_0 + \partial_2 a(u_0,\xi)\ \partial_s\xi\ ,\ \mathcal{T}^{-1}(\xi)^*(g|_{u_0}\partial_t u_0 + \eta)\Big) \\ &+ \epsilon^{-2} g|_{a(u_0,\xi)}\big(\partial_1 a(u_0,\xi)\ \partial_t u_0 + \partial_2 a(u_0,\xi)\ \partial_t\xi\big) \\ &- \epsilon^{-2}\mathcal{T}^{-1}(\xi)^*(g|_{u_0}\partial_t u_0 + \eta) - \partial_s(g|_{u_0}\partial_t u_0) + \Gamma|_{u_0}(\partial_s u_0\ ,\ g|_{u_0}\partial_t u_0)\Big) \end{aligned}$$

and

(55)

$$\begin{aligned} IV = \Big(\mathcal{T}(\xi)^* - \mathcal{T}(0)^* - d\mathcal{T}|_0^*\xi\Big)\Big(&d\mathcal{T}^{-1}|_\xi^*(\partial_s\xi\ ,\ g|_{u_0}\partial_t u_0 + \eta) \\ &+ \mathcal{T}^{-1}(\xi)^*\big(\partial_s(g|_{u_0}\partial_t u_0) + \partial_s\eta\big) - \epsilon^{-2}\mathcal{T}^{-1}(\xi)^*(g|_{u_0}\partial_t u_0 + \eta) \\ &- \Gamma|_{a(u_0,\xi)}\Big(\partial_1 a(u_0,\xi)\partial_s u + \partial_2 a(u_0,\xi)\partial_s\xi\ ,\ \mathcal{T}^{-1}(\xi)^*(g|_{u_0}\partial_t u_0 + \eta)\Big) \\ &+ \epsilon^{-2} g|_{a(u_0,\xi)}\big(\partial_1 a(u_0,\xi)\partial_t u_0 + \partial_2 a(u_0,\xi)\partial_t\xi\big)\Big) \end{aligned}$$

as well as (using $\partial_1 f_2(0,0)\xi = \frac{d}{d\tau}\big|_0 f_2(\tau\xi,0)$ and similarly for $\partial_2 f_2(0,0)\eta$)

(56)

$$\begin{aligned} V = &\Big(d\mathcal{T}^{-1}|_\xi^* - d\mathcal{T}^{-1}|_0^*\Big)(\partial_s\xi\ ,\ g|_{u_0}\partial_t u_0) + d\mathcal{T}^{-1}|_\xi^*(\partial_s\xi\ ,\ \eta) \\ &+ \Big(\mathcal{T}^{-1}(\xi)^* - 1\!\!1 - d\mathcal{T}^{-1}|_0^*\xi\Big)\partial_s(g|_{u_0}\partial_t u_0) + \Big(\mathcal{T}^{-1}(\xi)^* - 1\!\!1\Big)\partial_s\eta \\ &- \Big(\Gamma|_{a(u_0,\xi)}\big(\partial_1 a(u_0,\xi)\partial_s u_0\ ,\ \mathcal{T}^{-1}(\xi)^* g|_{u_0}\partial_t u_0\big) - \Gamma|_{u_0}\big(\partial_s u_0\ ,\ g|_{u_0}\partial_t u_0\big) \\ &- \tfrac{d}{d\tau}\big|_0 \Gamma|_{a(u_0,\tau\xi)}\big(\partial_1 a(u_0,\tau\xi)\partial_s u_0\ ,\ \mathcal{T}^{-1}(\tau\xi)^* g|_{u_0}\partial_t u_0\big)\Big) \\ &- \Big(\Gamma|_{a(u_0,\xi)}\big(\partial_2 a(u_0,\xi)\,\cdot\ ,\ \mathcal{T}^{-1}(\xi)^* g|_{u_0}\partial_t u_0\big) - \Gamma|_{u_0}\big(\cdot\ ,\ g|_{u_0}\partial_t u_0\big)\Big)\partial_s\xi \\ &- \Big(\Gamma|_{a(u_0,\xi)}\big(\partial_1 a(u_0,\xi)\partial_s u_0\ ,\ \mathcal{T}^{-1}(\xi)^*\,\cdot\big) - \Gamma|_{u_0}\big(\partial_s u_0\ ,\ \cdot\big)\Big)\eta \\ &- \Gamma|_{a(u_0,\xi)}\big(\partial_2 a(u_0,\xi)\partial_s\xi\ ,\ \mathcal{T}^{-1}(\xi)^*\eta\big) \\ &+ \epsilon^{-2}\Big(g|_{a(u_0,\xi)}\big(\partial_1 a(u_0,\xi)\partial_t u_0 + \partial_2 a(u_0,\xi)\partial_t\xi\big) - \mathcal{T}^{-1}(\xi)^*\big(g|_{u_0}\partial_t u_0 + \eta\big) \\ &- dg|_{u_0}\big(\xi\ ,\ \partial_t u_0\big) - g|_{u_0}\partial_t\xi + d\mathcal{T}^{-1}|_0^*\big(\xi\ ,\ g|_{u_0}\partial_t u_0\big) + \eta\Big). \end{aligned}$$

Let us consider for the moment the special case of the standard flat metric $g = 1\!\!1$. Then the ϵ^{-2}-terms in IV and VI reduce to $\epsilon^{-2}(\partial_t\xi - \eta)$, whereas the one in V vanishes. Distributing the L^∞-norms on $\partial_t\xi - \eta$, the ϵ^{-2}-terms in IV, V and VI contribute as follows (note the extra ϵ coming in from the $(0,p,\epsilon)$-norm)

$$\epsilon^{-1}\|\xi\|_p^2 \cdot \|\partial_t\xi - \eta\|_\infty + 0 + \epsilon^{-1}\|\xi\|_p \cdot \|\partial_t\xi - \eta\|_\infty.$$

Inserting (ξ_0, η_0) from step 1 of the Newton iteration and using the estimates derived there, the above term is less than a constant times ϵ^3 and that is exactly what we are heading for. Unfortunately the subsequent calculations do *not* result in the worst term being $c|\xi|^3$ (which would give ϵ^3 indeed), but $c|\xi|^2\|R(\cdot, \partial_t u_0)\cdot\|$ (which gives ϵ^2 only!). On the other hand the occurrence of curvature terms in the nonflat case should not be too surprising and the result specializes in the flat case to the one derived above. Moreover, it will turn out later that this a priori *bad* term will not lead to worse estimates in the Newton method below, if we only use the quadratic estimates separately for each component.

Back to the general case we now derive a corresponding estimate in the realm of Riemannian geometry. Rewrite the ϵ^{-2}-terms in VI (and IV)

$$\begin{aligned}
&\big(g|_{a(u_0,\xi)}\partial_2 a(u_0,\xi) - g|_{u_0}\big)\Big(\partial_t\xi + \Gamma_{u_0}\big(\partial_t u_0, \xi\big) - g^{-1}|_{u_0}\eta\Big)\\
&+ \big(g|_{a(u_0,\xi)}\partial_2 a(u_0,\xi) - \mathcal{T}^{-1}(\xi)^* g|_{u_0}\big)g^{-1}|_{u_0}\eta\\
&+ \Big(g|_{u_0}\partial_t\xi + g|_{u_0}\Gamma_{u_0}\big(\partial_t u_0, \xi\big) - \eta\Big)\\
&+ \bigg(g|_{a(u_0,\xi)}\Big(\partial_1 a(u_0,\xi)\cdot - \partial_2 a(u_0,\xi)\Gamma|_{u_0}\big(\cdot, \xi\big)\Big) - \mathcal{T}^{-1}(\xi)^* g|_{u_0}\cdot\bigg)\partial_t u_0.
\end{aligned} \tag{57}$$

We'd like to estimate the second term in the sum by $c\,|\xi|^2|\eta|$ and the last one by $c\,|\xi|^2$. Call the last one $h(\xi)$, then we have to show $h(0) = 0$ and $dh(0)\xi = 0$. Both will follow from the discussion of the exponential map and the parallel transport carried out in appendix A section A.1. $h(0) = 0$ is obvious and

$$\begin{aligned}
dh(0)\xi &= \tfrac{d}{d\tau}\big|_0 h(\tau\xi)\\
&= dg|_{u_0}\Big(\xi\,,\ \partial_t u_0 - \Gamma|_{u_0}\big(\partial_t u_0, 0\big)\Big)\\
&\quad + g|_{u_0}\Big(\partial_2\partial_1 a(u_0,0)\big(\partial_t u_0, \xi\big) - \partial_2\partial_2 a(u_0,0)\big(\xi, \Gamma|_{u_0}(\partial_t u_0, 0)\big)\\
&\qquad\qquad - \partial_2 a(u_0,0)\Gamma|_{u_0}\big(\partial_t u_0, \xi\big)\Big)\\
&\quad - d\mathcal{T}^{-1}|_0^*\big(\xi, g|_{u_0}\partial_t u_0\big)\\
&= dg|_{u_0}\big(\xi, \partial_t u_0\big) - g|_{u_0}\Gamma|_{u_0}\big(\partial_t u_0, \xi\big) - d\mathcal{T}^{-1}|_0^*\big(\xi, g|_{u_0}\partial_t u_0\big)\\
&= 0
\end{aligned} \tag{58}$$

where we used the results on the derivatives of the exponential map in proposition A.1.2 and the last equality uses lemma A.1.11 – both in appendix A section A.1.
Let $k(\xi)$ be the second term in the sum in equation (57), then we compute $k(0) = 0$ and

$$\begin{aligned} dk(0)\xi &= \tfrac{d}{d\tau}\big|_0 k(\tau\xi) \\ &= dg|_{u_0}\big(\xi, g^{-1}|_{u_0}\eta\big) + g|_{u_0}\partial_2\partial_2 a(u_0,0)\big(\xi, g^{-1}|_{u_0}\eta\big) \\ &\quad - d\mathcal{T}^{-1}|_0^*\big(\xi,\eta\big) \\ &= 0. \end{aligned} \tag{59}$$

The last equality is again due to proposition A.1.2 *iv*) and lemma A.1.11 – both in appendix A section A.1. Summing up, the absolute value of (57) may be estimated by

$$c\Big(|\xi|\cdot|\nabla_t\xi - g^{-1}\eta| + |\xi|^2\cdot|\eta| + |\nabla_t\xi - g^{-1}\eta| + |\xi|^2\Big). \tag{60}$$

This leads to

$$\boxed{\begin{aligned} |IV| \le c|\xi|^2\epsilon^{-2}&\Big(|\nabla_t\xi - g^{-1}\eta| + |\xi|^2\cdot|\eta| + |\xi|^2\Big) \\ &+ c|\xi|^2\Big(c + |\partial_s\xi| + |\partial_s\xi|\cdot|\eta| + |\eta| + |\partial_s\eta|\Big). \end{aligned}} \tag{61}$$

In order to estimate VI we have to rewrite the terms without an ϵ^{-2} in front:

$$\begin{aligned} & d\mathcal{T}|_0^*\Bigg(\xi\,,\, d\mathcal{T}^{-1}|_\xi^*\big(\partial_s\xi, g(u_0)\partial_t u_0 + \eta\big) \\ &+ \big(\mathcal{T}^{-1}(\xi)^* - \mathbb{1}\big)\partial_s\big(g(u_0)\partial_t u_0\big) + \mathcal{T}^{-1}(\xi)^*\partial_s\eta \\ &- \Big(\Gamma_{a(u_0,\xi)}\big(\partial_1 a(u_0,\xi)\,\cdot\,,\ \mathcal{T}^{-1}(\xi)^*\,\cdot\big) - \Gamma|_{u_0}\big(\cdot,\cdot\big)\Big)\big(\partial_s u_0, g(u_0)\partial_t u_0\big) \\ &- \Gamma_{a(u_0,\xi)}\big(\partial_1 a(u_0,\xi)\partial_s u_0\ ,\ \mathcal{T}^{-1}(\xi)^*\eta\big) \\ &- \Gamma_{a(u_0,\xi)}\Big(\partial_2 a(u_0,\xi)\partial_s\xi\ ,\ \mathcal{T}^{-1}(\xi)^*\big(g(u_0)\partial_t u_0 + \eta\big)\Big)\Bigg) \\ &\le c|\xi|\Big(|\xi| + |\partial_s\xi| + |\partial_s\xi|\cdot|\eta| + |\eta| + |\partial_s\eta|\Big) \end{aligned}$$

and therefore

$$\boxed{\begin{aligned} |VI| \le c|\xi|\epsilon^{-2}&\Big(|\nabla_t\xi - g^{-1}\eta| + |\xi|^2\cdot|\eta| + |\xi|^2\Big) \\ &+ c|\xi|\Big(|\xi| + |\partial_s\xi| + |\partial_s\xi|\cdot|\eta| + |\eta| + |\partial_s\eta|\Big). \end{aligned}} \tag{62}$$

Rewrite the ϵ^{-2}-terms in term V and add twice zero to obtain

(63)

$$\begin{aligned}
&\Big(g|_{a(u_0,\xi)}\partial_2 a(u_0,\xi) - g|_{u_0}\Big)\Big(\partial_t\xi + \Gamma|_{u_0}\big(\partial_t u_0,\xi\big) - g^{-1}|_{u_0}\eta\Big)\\
&+\Big(g|_{a(u_0,\xi)}\partial_2 a(u_0,\xi) - \mathcal{T}^{-1}(\xi)^* g|_{u_0}\Big)\, g^{-1}|_{u_0}\eta\\
&+\Big(g|_{a(u_0,\xi)}\partial_1 a(u_0,\xi)\partial_t u_0 - \mathcal{T}^{-1}(\xi)^* g|_{u_0}\partial_t u_0 + d\mathcal{T}^{-1}|_0^*\big(\xi, g|_{u_0}\partial_t u_0\big)\\
&\quad - dg|_{u_0}\big(\xi,\partial_t u_0\big) + g|_{u_0}\Gamma|_{u_0}\big(\partial_t u_0,\xi\big) - g|_{a(u_0,\xi)}\partial_2 a(u_0,\xi)\Gamma|_{u_0}\big(\partial_t u_0,\xi\big)\Big).
\end{aligned}$$

The first term in the sum gives $c\,|\xi|\cdot|\nabla_t\xi - g^{-1}\eta|$ and the second term is identical to the one in (57), which we had estimated previously by $c|\xi|^2|\eta|$, so it remains to estimate the third term in the sum by $c\,|\xi|^3$. Call it again $h(\xi)$ and observe that terms 3, 4 and 5 cancel due to lemma A.1.11 in appendix A section A.1. So it coincides with the function $h(\xi)$ considered above and we know $h(0) = 0$ and $dh(0)\,\xi = 0$. Let us compute also the second derivative of h as it was great if it vanished (the term then would contribute $c|\xi|^3$). It turns out to be important to keep track of indices from now on. Derivatives of $a(u_0,\xi)$ with respect to the first variable will be denoted by $\partial/\partial u^i$ and with respect to the second one by $\partial/\partial x^j$.

$$\begin{aligned}
d^2h_j(0)\big(\xi,\xi\big) &= \tfrac{d^2}{d\tau^2}\Big|_0 h_j(\tau\xi)\\
&= \tfrac{d^2}{d\tau^2}\Big|_0\Bigg(g_{jl}|_{a(u_0,\tau\xi)}\,\frac{\partial a(u_0,\tau\xi)^l}{\partial u^i}\,\partial_t u_0^i - \mathcal{T}^*(\tau,0)_j^k\, g_{ki}|_{u_0}\partial_t u_0^i\\
&\quad - g_{jl}|_{a(u_0,\tau\xi)}\,\frac{\partial a(u_0,\tau\xi)^l}{\partial x^i}\,\Gamma^i_{rs}|_{u_0}\,\partial_t u_0^r\,\tau\xi^s\Bigg)\\
&= \tfrac{d}{d\tau}\Big|_0\Bigg(\frac{\partial g_{jl}}{\partial u^s}|_{a(u_0,\tau\xi)}\,\frac{\partial a(u_0,\tau\xi)^s}{\partial x^r}\,\xi^r\,\frac{\partial a(u_0,\tau\xi)^l}{\partial u^i}\,\partial_t u_0^i\\
&\quad + g_{jl}|_{a(u_0,\tau\xi)}\,\frac{\partial^2 a(u_0,\tau\xi)^l}{\partial x^s\partial u^i}\,\partial_t u_0^i\,\xi^s - \partial_\tau\mathcal{T}^*(\tau,0)_j^k\, g_{ki}|_{u_0}\partial_t u_0^i\\
&\quad - \frac{\partial g_{jl}}{\partial u^m}|_{a(u_0,\tau\xi)}\,\frac{\partial a(u_0,\tau\xi)^m}{\partial x^n}\,\xi^n\,\frac{\partial a(u_0,\tau\xi)^l}{\partial x^i}\,\Gamma^i_{rs}|_{u_0}\,\partial_t u_0^r\,\tau\xi^s\\
&\quad - g_{jl}|_{a(u_0,\tau\xi)}\,\frac{\partial^2 a(u_0,\tau\xi)^l}{\partial x^m\partial x^i}\,\xi^m\,\Gamma^i_{rs}|_{u_0}\,\partial_t u_0^r\,\tau\xi^s\\
&\quad - g_{jl}|_{a(u_0,\tau\xi)}\,\frac{\partial a(u_0,\tau\xi)^l}{\partial x^i}\,\Gamma^i_{rs}|_{u_0}\,\partial_t u_0^r\,\xi^s\Bigg)\\
&= \frac{\partial^2 g_{jl}}{\partial u^m\partial u^s}|_{u_0}\,\xi^m\xi^s\partial_t u_0^l + \frac{\partial g_{jl}}{\partial u^s}|_{u_0}\,\frac{\partial^2 a(u_0,0)^s}{\partial x^m\partial x^r}\,\xi^m\xi^r\partial_t u_0^l\\
&\quad + \frac{\partial g_{jl}}{\partial u^s}|_{u_0}\,\xi^s\,\frac{\partial^2 a(u_0,0)^l}{\partial x^m\partial u^i}\,\xi^m\partial_t u_0^i + \frac{\partial g_{jl}}{\partial u^r}|_{u_0}\,\xi^r\,\frac{\partial^2 a(u_0,0)^l}{\partial x^s\partial u^i}\,\xi^s\partial_t u_0^i\\
&\quad + g_{jl}|_{u_0}\,\frac{\partial^3 a(u_0,0)^l}{\partial x^r\partial x^s\partial u^i}\,\xi^r\xi^s\partial_t u_0^i - \tfrac{d^2}{d\tau^2}\Big|_0\mathcal{T}^*(\tau,0)_j^k\, g_{ki}|_{u_0}\partial_t u_0^i
\end{aligned} \tag{64}$$

$$
\begin{aligned}
&- 0 - 0 - 0 - \frac{\partial g_{jl}}{\partial u^m}|_{u_0}\, \xi^m\, \Gamma^l_{rs}|_{u_0}\, \xi^s \partial_t u_0^r \\
&- 0 - 0 + g_{jl}|_{u_0}\, \Gamma^l_{mi}|_{u_0}\, \xi^m\, \Gamma^i_{rs}|_{u_0}\, \xi^s \partial_t u_0^r \\
&- \frac{\partial g_{jl}}{\partial u^m}|_{u_0}\, \xi^m\, \Gamma^l_{rs}|_{u_0}\, \xi^s \partial_t u_0^r - g_{jl}|_{u_0}\, \frac{\partial^2 a(u_0,0)^l}{\partial x^m \partial x^i}\, \xi^m\, \Gamma^i_{rs}|_{u_0}\, \xi^s \partial_t u_0^r \\
&= g_{ji}|_{u_0}\, \xi^s \xi^r \partial_t u_0^l \left(\frac{\partial \Gamma^i_{lr}}{\partial u^s} - \frac{\partial \Gamma^i_{sr}}{\partial u^l} + \Gamma^i_{sm}\, \Gamma^m_{lr} - \Gamma^i_{ml}\, \Gamma^m_{sr} \right)|_{u_0} \\
&= g_{ji}|_{u_0}\, R^i_{rsl}|_{u_0}\, \xi^r \xi^s \partial_t u_0^l \\
&= g_{ji}|_{u_0}\, (R(\xi, \partial_t u_0)\xi)^i|_{u_0}
\end{aligned}
$$

which is not zero in general! Here the first four equalities are a straightforward calculation using the definition of $h_j(\tau\xi)$ and the product and chain rules for derivatives. The fifth equality uses results from appendix A section A.1, namely proposition A.1.2 on the derivatives of the exponential map and lemma A.1.11 on the second derivative of the parallel transport. Note that in the sum preceding the fifth equality sign terms 3 and 4 are zero. Moreover, terms $1, 2, 6, (10 + 14), 15$ are identified with terms $2, 3, 1, 4, 7$, respectively, in the formula of lemma A.1.11. The remaining terms are 5 and 13 here and terms 5 and 6 in the lemma. Now use formula (98) for the curvature tensor in terms of Christoffel symbols and their derivatives to get the last but one equality.

The above estimates yield

$$
\boxed{
\begin{aligned}
|V| \le c\epsilon^{-2}|\xi| \Big(|\nabla_t \xi - g^{-1}\eta| + |\xi| \cdot |\eta| + |\xi| \cdot \|R(\cdot, \partial_t u_0) \cdot \| \Big) \\
+ c|\xi| \Big(|\xi| + |\partial_s \xi| + |\eta| + |\partial_s \eta| \Big) + c|\eta| \cdot |\partial_s \xi|.
\end{aligned}
}
$$

All together this gives finally our claim for the second component

$$
\begin{aligned}
|F_2| &= |IV + V + VI| \\
&\le c\epsilon^{-2}|\xi| \Big(|\nabla_t \xi - g^{-1}\eta| + |\xi| \cdot |\eta| + |\xi| \cdot \|R(\cdot, \partial_t u_0) \cdot \| \Big) \\
&\quad + c|\xi| \Big(|\xi| + |\partial_s \xi| \big(1 + |\eta|\big) + |\eta| + |\partial_s \eta| \Big) + c|\eta| \cdot |\partial_s \xi|.
\end{aligned}
$$

Estimates for $\nabla_\alpha F_1$ and $\nabla^*_\alpha F_2$

To estimate $\nabla_\alpha F_1$ we may simply apply the corresponding estimate of lemma 5.2.2 (pointwise estimate I) with $X = 0$ and $Y = 0$. Moreover, as

$$
\nabla^*_\alpha F_2 = \partial_\alpha F_2 - \Gamma|_{u_0}(\partial_\alpha u_0, F_2),
$$

we may use the estimate for F_2 derived above for the second term and it remains to consider $\partial_\alpha F_2$. Using the notation introduced in (53) we obtain

$$
|\partial_\alpha F_2| \le |\partial_\alpha IV| + |\partial_\alpha V| + |\partial_\alpha VI|
$$

and proceed by treating each term separately.

Let us begin with term IV given by equation (55) as a product of two factors which we denote by IV_1 and IV_2. Note that it turns out to be sufficient to estimate the ϵ^{-2}-terms individually. We obtain

$$
\begin{aligned}
|\partial_\alpha IV| &\le |\partial_\alpha IV_1| \cdot |IV_2| + |IV_1| \cdot |\partial_\alpha IV_2| \\
&\le |d\mathcal{T}(\xi)^* \, \partial_\alpha \xi - d\mathcal{T}(0)^* \, \partial_\alpha \xi| \cdot |IV_2| \\
&\quad + |\mathcal{T}(\xi)^* - \mathcal{T}(0)^* - d\mathcal{T}(0)^* \xi| \cdot |\partial_\alpha IV_2| \\
&\le c\,|\xi| \cdot |\partial_\alpha \xi| \cdot |IV_2| + c\,|\xi|^2 \cdot |\partial_\alpha IV_2| \\
&\le c\,|\xi| \cdot |\partial_\alpha \xi| \Big(1 + |\eta| + |\partial_s \xi|\,(1 + |\eta|) + |\partial_s \eta|\Big) \\
&\quad + c\,\underline{\epsilon^{-2}\,|\xi| \cdot |\partial_\alpha \xi|} \Big(\underline{1} + |\partial_t \xi| + |\eta|\Big) \\
&\quad + c\,|\xi|^2 \Bigg(1 + |\eta|\,(1 + |\partial_s \xi|) + |\partial_s \xi|\,(1 + |\partial_\alpha \eta|) \\
&\quad + |\partial_\alpha \xi|\,(|\partial_s \xi| + |\partial_s \eta| + |\partial_s \xi| \cdot |\eta|) + |\partial_\alpha \partial_s \xi|\,(1 + |\eta|) + |\partial_\alpha \partial_s \eta|\Bigg) \\
&\quad + c\,\epsilon^{-2}\,|\xi|^2 \Bigg(1 + |\partial_t \xi| + |\partial_\alpha \eta| + |\partial_\alpha \xi|\,(1 + |\eta| + |\partial_t \xi|) + |\partial_\alpha \partial_t \xi|\Bigg)
\end{aligned}
$$

where we underlined the worst term with respect to the Newton method.

Next we deal with term VI given by equation (54) and define new ϵ-independent functions VI_1 and VI_2 by the identity

$$VI = d\mathcal{T}|_0^*(\xi, VI_1 + \epsilon^{-2} VI_2).$$

We obtain

$$|\partial_\alpha VI| \le c\,|\partial_\alpha \xi| \left(|VI_1| + \epsilon^{-2}|VI_2|\right) + c\,|\xi| \left(|\partial_\alpha VI_1| + \epsilon^{-2}|\partial_\alpha VI_2|\right).$$

Note that $VI_1 + \epsilon^{-2} VI_2$ satisfies estimate (62) with the common term $|\xi|$ replaced by $|\partial_\alpha \xi|$.

$$
\begin{aligned}
&|\,\partial_\alpha \xi\,| \left(|VI_1| + \epsilon^{-2}|VI_2|\right) \\
&\qquad \le c|\partial_\alpha \xi| \epsilon^{-2} \Big(|\nabla_t \xi - g^{-1}\eta| + |\xi|^2 \cdot |\eta| + |\xi|^2\Big) \\
&\qquad\quad + c|\partial_\alpha \xi| \Big(|\xi| + |\partial_s \xi| + |\partial_s \xi| \cdot |\eta| + |\eta| + |\partial_s \eta|\Big).
\end{aligned} \tag{65}
$$

Use product and chain rule to calculate $\partial_\alpha VI_1$ and then estimate each term containing ξ, η or partial derivatives thereof individually and pair the remaining terms appropriately to obtain

$$
\begin{aligned}
|\xi| \cdot |\partial_\alpha VI_1| \le c\,|\xi| \Bigg(&|\eta|\,(1 + |\partial_\alpha \xi| + |\partial_s \xi| + |\partial_\alpha \xi| \cdot |\partial_s \xi|) \\
&+ |\partial_s \xi|\,(1 + |\partial_\alpha \eta|) + |\partial_\alpha \xi|\,(1 + |\partial_s \eta|) \\
&+ |\partial_\alpha \eta| + |\partial_\alpha \partial_s \xi|\,(1 + |\eta|) + |\partial_\alpha \partial_s \eta|\Bigg).
\end{aligned} \tag{66}
$$

Let

$$\partial_\alpha V I_2 = k_1 + k_2 + k_3$$

where

$$\begin{aligned}
k_1 &= dg|_{a(u_0,\xi)}\Big(\partial_2 a(u_0,\xi)\,\partial_\alpha\xi, \partial_1 a(u_0,\xi)\,\partial_t u_0 + \partial_2 a(u_0,\xi)\,\partial_t\xi\Big)\\
&\quad + dg|_{a(u_0,\xi)}\Big(\partial_1 a(u_0,\xi)\,\partial_\alpha u_0 + \partial_2 a(u_0,\xi)\,\partial_\alpha\xi, +\partial_2 a(u_0,\xi)\,\partial_t\xi\Big)\\
&\quad + g|_{a(u_0,\xi)}\Big(\partial_2\partial_1 a(u_0,\xi)\,(\partial_\alpha\xi,\partial_t u_0) + \partial_1\partial_2 a(u_0,\xi)\,(\partial_\alpha u_0,\partial_t\xi)\\
&\quad + \partial_2\partial_2 a(u_0,\xi)\,(\partial_\alpha\xi,\partial_t\xi)\Big) - d\mathcal{T}^{-1}|_\xi^*\,(\partial_\alpha\xi, g\partial_t u_0 + \eta)\\
k_2 &= dg|_{a(u_0,\xi)}(\partial_1 a(u_0,\xi)\,\partial_\alpha u_0, \partial_1 a(u_0,\xi)\,\partial_t u_0)\\
&\quad + g|_{a(u_0,\xi)}\,\partial_1\partial_1 a(u_0,\xi)\,(\partial_\alpha u_0,\partial_t u_0) - \mathcal{T}^{-1}(\xi)^*\,dg|_{u_0}\,(\partial_\alpha u_0,\partial_t u_0)
\end{aligned}$$

and

$$\begin{aligned}
k_3 &= g|_{a(u_0,\xi)}\Big(\partial_1 a(u_0,\xi)\,\partial_\alpha\partial_t u_0 + \partial_2 a(u_0,\xi)\,\partial_\alpha\partial_t\xi\Big)\\
&\quad - \mathcal{T}^{-1}(\xi)^*\,\Big(g\partial_\alpha\partial_t u_0 + \partial_\alpha\eta\Big).
\end{aligned}$$

Estimate k_1 term by term to obtain

$$\epsilon^{-2}\,|\xi|\cdot|k_1| \le c\,\epsilon^{-2}\,|\xi|\,\Big(|\partial_t\xi| + |\partial_\alpha\xi|\,(1 + |\partial_t\xi| + |\eta|)\Big).$$

Consider k_2 as a function of ξ, then proposition A.1.2 *iii)* shows that $k_2(0) = 0$ and we obtain

$$\epsilon^{-2}\,|\xi|\cdot|k_2(\xi)| \le c\,\epsilon^{-2}\,|\xi|^2.$$

k_3 is harder to deal with. We rewrite it in a form similar to (57)

$$\begin{aligned}
k_3 &= \big(g|_{a(u_0,\xi)}\partial_2 a(u_0,\xi) - g|_{u_0}\big)\Big(\nabla_\alpha\nabla_t\xi - g^{-1}|_{u_0}\nabla_\alpha^*\eta\Big)\\
&\quad + g\Big(\nabla_\alpha\nabla_t\xi - g^{-1}|_{u_0}\nabla_\alpha^*\eta\Big)\\
&\quad + \big(g|_{a(u_0,\xi)}\partial_2 a(u_0,\xi) - \mathcal{T}^{-1}(\xi)^* g|_{u_0}\big)g^{-1}|_{u_0}\partial_\alpha\eta\\
&\quad + \Bigg(g|_{a(u_0,\xi)}\Big(\partial_1 a(u_0,\xi)\,\partial_\alpha\partial_t u_0 - \partial_2 a(u_0,\xi)\Gamma|_{u_0}(\partial_\alpha\partial_t u_0,\xi)\Big)\\
&\quad - \mathcal{T}^{-1}(\xi)^*\,g|_{u_0}\,\partial_\alpha\partial_t u_0\Bigg)\\
&\quad - g|_{a(u_0,\xi)}\,\partial_2 a(u_0,\xi)\,\Bigg(d\Gamma|_{\partial_\alpha u_0}(\partial_t u_0,\xi) + \Gamma|_{u_0}(\partial_t u_0,\partial_\alpha\xi)\\
&\quad + \Gamma|_{u_0}(\partial_\alpha u_0,\partial_t\xi) + \Gamma|_{u_0}\Big(\partial_\alpha u_0, \Gamma|_{u_0}(\partial_t u_0,\xi)\Big) - \Gamma|_{u_0}(\partial_\alpha u_0,\eta)\Bigg).
\end{aligned}$$

The first line is estimated by

$$c|\xi|\cdot|\nabla_\alpha\nabla_t\xi - g^{-1}\nabla_\alpha^*\eta|,$$

the second one by $c|\nabla_\alpha \nabla_t \xi - g^{-1}\nabla_\alpha^* \eta|$ and the third one had been calculated before in (59) and yields $c|\xi|^2|\partial_\alpha \eta|$. The fourth line contributes $c|\xi|^2$ as we know from (58). Estimate the remaining terms individually and get $c\left(|\xi| + |\partial_t \xi| + |\eta| + |\partial_\alpha \xi|\right)$ as upper bound. This implies

$$\begin{aligned}\epsilon^{-2}\,|\xi|\cdot|k_3| \leq c\,\epsilon^{-2}\,|\xi|\Big(&|\nabla_\alpha \nabla_t \xi - g^{-1}|_{u_0}\nabla_\alpha^* \eta|(1+|\xi|) + |\xi|^2|\partial_\alpha \eta| \\ &+ |\xi| + |\partial_t \xi| + |\eta| + |\partial_\alpha \xi|\Big)\end{aligned}$$

and we finally obtain

$$\begin{aligned}\epsilon^{-2}\,|\xi|\cdot|\partial_\alpha V I_2| \leq c\,\epsilon^{-2}\,|\xi|\Big(&|\nabla_\alpha \nabla_t \xi - g^{-1}\nabla_\alpha^* \eta|(1+|\xi|) + |\xi| \\ &+ |\eta| + |\partial_t \xi| + |\partial_\alpha \xi|(1 + |\partial_t \xi| + |\eta|) + |\xi|^2|\partial_\alpha \eta|\Big).\end{aligned} \tag{67}$$

Term $\partial_\alpha V$ is the hardest one. Recall that V is given by equation (56) as a sum of nine terms h_i, $i \in \{1, \dots, 9\}$, (here we consider everything inside outmost brackets as being one term). So we may write

$$\partial_\alpha V = \partial_\alpha h_1 + \ldots + \partial_\alpha h_8 + \epsilon^{-2}\partial_\alpha h_9$$

and remark that only the last term contains a power of ϵ. We only state the results of estimating the terms with $i = 1, \dots, 8$ as the calculations are rather lenghty but straightforward (given the explicit estimates done so far). A guiding principle is to group together terms with the same linear factors. Here are the results: Note that the constant c depends on u_0 and its partial derivatives up to second order, all evaluated at the point (s,t)

$$\begin{aligned}
\frac{1}{c}\,|\partial_\alpha h_1| &\leq |\partial_\alpha \xi|\cdot|\partial_s \xi| + |\xi|\left(|\partial_\alpha \partial_s \xi| + |\partial_s \xi|\right) \\
\frac{1}{c}\,|\partial_\alpha h_2| &\leq |\partial_\alpha \xi|\cdot|\partial_s \xi|\cdot|\eta| + |\partial_\alpha \partial_s \xi|\cdot|\eta| + |\partial_s \xi|\cdot|\partial_\alpha \eta| \\
\frac{1}{c}\,|\partial_\alpha h_3| &\leq |\xi|\left(|\xi| + |\partial_\alpha \xi|\right) \\
\frac{1}{c}\,|\partial_\alpha h_4| &\leq |\partial_s \eta|\left(|\xi| + |\partial_\alpha \xi|\right) \\
\frac{1}{c}\,|\partial_\alpha h_5| &\leq |\xi|\left(|\xi| + |\partial_\alpha \xi|\right) \\
\frac{1}{c}\,|\partial_\alpha h_6| &\leq |\xi|\cdot|\partial_\alpha \partial_s \xi| + |\partial_s \xi|\left(|\xi| + |\partial_\alpha \xi|\right) \\
\frac{1}{c}\,|\partial_\alpha h_7| &\leq |\eta|\left(|\xi| + |\partial_\alpha \xi|\right) \\
\frac{1}{c}\,|\partial_\alpha h_8| &\leq |\eta|\left(|\partial_s \xi| + |\partial_s \xi|\cdot|\partial_\alpha \xi| + |\partial_\alpha \partial_s \xi|\right) + |\partial_s \xi|\cdot|\partial_\alpha \eta|.
\end{aligned}$$

Now hit h_9 with ∂_α and rewrite the resulting avalanche of terms as follows

$$
\begin{aligned}
&\partial_\alpha h_9 = \\
&\quad \Big(dg|_{a(u_0,\xi)}(\partial_1 a(u_0,\xi)\,\partial_\alpha u_0, \partial_2 a(u_0,\xi)\,\cdot) - dg|_{u_0}(\partial_\alpha u_0,\cdot)\Big)\ (\nabla_t\xi - g^{-1}\eta) \\
&\quad + dg|_{a(u_0,\xi)}\big(\partial_2 a(u_0,\xi)\,\partial_\alpha\xi, \nabla_t\xi - g^{-1}\eta\big) \\
&\quad + g|_{a(u_0,\xi)}\,\partial_2\partial_2 a(u_0,\xi)\,\big(\partial_\alpha\xi, \nabla_t\xi - g^{-1}\eta\big) \\
&\quad + \Big(g|_{a(u_0,\xi)}\partial_1\partial_2 a(u_0,\xi)(\partial_\alpha u_0,\cdot) - g\partial_1\partial_2 a(u_0,0)(\partial_\alpha u_0,\cdot)\Big)(\nabla_t\xi - g^{-1}\eta) \\
&\quad + \big(g|_{a(u_0,\xi)} - g|_{u_0}\big)\ (\nabla_\alpha\nabla_t\xi - g^{-1}\nabla_\alpha^*\eta) \\
&\quad - \big(g|_{a(u_0,\xi)} - g\big)\Big(\Gamma\big(\partial_\alpha u_0, \partial_t\xi\big) + g^{-1}\Gamma\big(\partial_\alpha u_0,\eta\big) + \Gamma\big(\partial_\alpha u_0, \Gamma(\partial_t u_0,\xi)\big)\Big) \\
&\quad + \Big(dg|_{a(u_0,\xi)}\big(\partial_1 a(u_0,\xi)\,\partial_\alpha u_0, \partial_2 a(u_0,\xi)\,g^{-1}\eta\big) \\
&\qquad + g|_{a(u_0,\xi)}\,\partial_1\partial_2 a(u_0,\xi)\,\big(\partial_\alpha u_0, g^{-1}\eta\big) - \mathcal{T}^{-1}(\xi)^*\,dg|_{u_0}\big(\partial_\alpha u_0, g^{-1}\eta\big)\Big) \\
&\quad + \Big(dg|_{a(u_0,\xi)}\big(\partial_2 a(u_0,\xi)\,\partial_\alpha\xi, \partial_2 a(u_0,\xi)\,g^{-1}\eta\big) \\
&\qquad + g|_{a(u_0,\xi)}\,\partial_2\partial_2 a(u_0,\xi)\,\big(\partial_\alpha\xi, g^{-1}\eta\big) - d\mathcal{T}^{-1}|_\xi^*\ \big(\partial_\alpha\xi,\eta\big)\Big) \\
&\quad + \Big(g|_{a(u_0,\xi)}\,\partial_2 a(u_0,\xi) - \mathcal{T}^{-1}(\xi)^* g|_{u_0}\Big)\big(dg^{-1}|_{u_0}(\partial_\alpha u_0,\eta) + g^{-1}\partial_\alpha\eta\big) \\
&\quad + \Big(dg|_{a(u_0,\xi)}\big(\partial_2 a(u_0,\xi)\,\partial_\alpha\xi, \partial_1 a(u_0,\xi)\,\partial_t u_0\big) - d\mathcal{T}^{-1}|_\xi^*\big(\partial_\alpha\xi, g\partial_t u_0\big) \\
&\qquad + g|_{a(u_0,\xi)}\,\partial_2\partial_1 a(u_0,\xi)\,\big(\partial_\alpha\xi,\partial_t u_0\big) - g|_{a(u_0,\xi)}\,\partial_2 a(u_0,\xi)\,\Gamma\big(\partial_t u_0,\partial_\alpha\xi\big)\Big) \\
&\qquad + d\mathcal{T}^{-1}|_0^*\big(\partial_\alpha\xi, g\partial_t u_0\big) - dg|_{u_0}\big(\partial_\alpha\xi, g\partial_t u_0\big) + g\Gamma\big(\partial_t u_0,\partial_\alpha\xi\big) \\
&\quad + \Big(g|_{a(u_0,\xi)}\,\partial_1 a(u_0,\xi)\,\partial_\alpha\partial_t u_0 - \mathcal{T}^{-1}(\xi)^*\ g\,\partial_\alpha\partial_t u_0 + d\mathcal{T}^{-1}|_0^*\big(\xi, g\partial_\alpha\partial_t u_0\big) \\
&\qquad - dg|_{u_0}\ \big(\xi,\partial_\alpha\partial_t u_0\big) + g\Gamma\big(\partial_\alpha\partial_t u_0,\xi\big) - g|_{a(u_0,\xi)}\,\partial_2 a(u_0,\xi)\,\Gamma\big(\partial_\alpha\partial_t u_0,\xi\big)\Big) \\
&\quad + \Big(dg|_{u_0}\ \big(\partial_\alpha u_0, \Gamma(\partial_t u_0,\xi)\big) - g|_{a(u_0,\xi)}\,\partial_2 a(u_0,\xi)\,d\Gamma|_{u_0}\big(\partial_\alpha u_0,\partial_t u_0,\xi\big) \\
&\qquad + g d\Gamma|_{u_0}\big(\partial_\alpha u_0,\partial_t u_0,\xi\big) - g|_{a(u_0,\xi)}\,\partial_1\partial_2 a(u_0,\xi)\,\big(\partial_\alpha u_0,\Gamma(\partial_t u_0,\xi)\big) \\
&\qquad - dg|_{a(u_0,\xi)}\big(\partial_1 a(u_0,\xi)\,\partial_\alpha u_0, \partial_2 a(u_0,\xi)\,\Gamma(\partial_t u_0,\xi)\big)\Big) \\
&\quad - \Big(dg|_{a(u_0,\xi)}\big(\partial_2 a(u_0,\xi)\,\partial_\alpha\xi, \partial_2 a(u_0,\xi)\,\Gamma(\partial_t u_0,\xi)\big) \\
&\qquad + g|_{a(u_0,\xi)}\,\partial_2\partial_2 a(u_0,\xi)\,\big(\partial_\alpha\xi, \Gamma(\partial_t u_0,\xi)\big)\Big) \\
&\quad + \big(-\mathcal{T}^{-1}(\xi)^* + \mathbb{1} + d\mathcal{T}^{-1}|_0^*\,\xi\big)\ dg|_{u_0}\ \big(\partial_\alpha u_0,\partial_t u_0\big) \\
&\quad + \Big(dg|_{a(u_0,\xi)}\big(\partial_1 a(u_0,\xi)\,\partial_\alpha u_0, \partial_1 a(u_0,\xi)\,\partial_t u_0\big) - dg|_{u_0}\ \big(\partial_\alpha u_0,\partial_t u_0\big) \\
&\qquad + g|_{a(u_0,\xi)}\,\partial_1\partial_1 a(u_0,\xi)\,\big(\partial_\alpha u_0,\partial_t u_0\big) - d^2 g|_{u_0}\big(\partial_\alpha u_0,\xi,\partial_t u_0\big)\Big).
\end{aligned}
$$

We obtain termwise

$$\begin{aligned}
&\frac{1}{c\epsilon^2}\,|\partial_\alpha h_9| \\
&\le \Big(|\nabla_t\xi - g^{-1}\eta|\big(|\xi| + |\partial_\alpha\xi| + |\partial_\alpha\xi| + |\xi| + |\xi|\big) + |\xi|\big(|\partial_t\xi| + |\eta| + |\xi|\big)\Big) \\
&\quad + \Big(|\xi|\big(|\eta| + |\partial_\alpha\xi|\cdot|\eta|\big) + |\xi|^2\big(|\eta| + |\partial_\alpha\eta|\big)\Big) \\
&\quad + |\xi|\cdot|\partial_\alpha\xi| + |\xi|^2 + |\xi|^2 + |\xi|\cdot|\partial_\alpha\xi| + |\xi|^2 + |\xi|^2.
\end{aligned}$$

All in all we get

$$\begin{aligned}
|\partial_\alpha V| \le{}& c\epsilon^{-2}|\nabla_t\xi - g^{-1}\eta|\cdot\big(|\xi| + |\partial_\alpha\xi|\big) + c\epsilon^{-2}|\xi|^2\big(|\eta| + |\partial_\alpha\eta|\big) \\
&+ c\epsilon^{-2}|\xi|\Big(|\xi| + |\eta| + |\partial_t\xi| + |\eta|\cdot|\partial_\alpha\xi| + |\partial_\alpha\xi|\Big) \\
&+ c|\xi|\big(|\xi| + |\eta| + |\partial_\alpha\xi| + |\partial_s\xi| + |\partial_s\eta| + |\partial_\alpha\partial_s\xi|\big) + c|\partial_s\xi|\cdot|\partial_\alpha\eta| \\
&+ c|\partial_\alpha\xi|\big(|\eta| + |\partial_s\xi| + |\partial_s\eta| + |\eta|\cdot|\partial_s\xi|\big) + c|\eta|\big(|\partial_s\xi| + |\partial_\alpha\partial_s\xi|\big).
\end{aligned}$$

This estimate for $\partial_\alpha V$, the one for $\partial_\alpha IV$ and estimates (65),(66),(67) for $\partial_\alpha VI$ together give the claimed estimate for the second component $\nabla_\alpha^* F_2$. $\square$

5.2. Quadratic Estimate I

The following quadratic estimate is an essential qualitative ingredient to carry out the induction step in the Newton method. The theorem follows from the pointwise estimate 5.2.2 via integration as described in remark 5.0.8. Note that the conditions on $\|\xi\|_\infty$ and $\|X\|_\infty$ are necessary in order for the local constructions to be well defined. The other L^∞-conditions only serve to simplify the expressions. Terms involving s-derivatives have not been dropped in simplifying the estimates for the components F_1 and F_2. This is of importance in the uniqueness part of the Newton method because these terms may appear with negative powers of ϵ. We underlined the worst terms with respect to rate of convergence in the existence part of the Newton method. Moreover, the theorem actually holds for any cylinder with appropriate smoothness and asymptotic convergence properties.

THEOREM 5.2.1. (**Quadratic estimate I**) *Let $p > 2$ and u be an element of the moduli space $\mathcal{M}^0(x^-, x^+)$, where x^-, $x^+ \in \mathrm{Crit}\, \mathcal{I}_V$, define $w = g(u)\partial_t u$ and denote*

$$\mathcal{F}^{triv}_{\epsilon,u}(Z+\zeta) - \mathcal{F}^{triv}_{\epsilon,u}(Z) - d\mathcal{F}^{triv}_{\epsilon,u}(Z)\circ\zeta = \begin{pmatrix} F_1 \\ F_2 \end{pmatrix}.$$

Then $\exists$ a constant $c_p > 0$ such that for $Z = (X, Y) \in C_0^\infty(\mathbb{R}\times S^1, u^*TM \oplus u^*T^*M)$ with $\|X\|_\infty < \iota'_M/2$ and $\|Y\|_\infty + \|\nabla_t X\|_\infty + \|\nabla_s X\|_\infty \le \sqrt{c_p}/2$

$$\begin{aligned}
\|F_1\|_p \le{}& c_p\|\xi\|_\infty^2\Big(\|\nabla_t X\|_p + \|\nabla_s X\|_p + \|\nabla_t Y\|_p\Big)\\
&+ c_p\|\xi\|_\infty\Big(\|\xi\|_p + \underline{\|\nabla_t\xi\|_p} + \|\nabla_s\xi\|_p\|\xi\|_\infty + \underline{\|\eta\|_p} + \|\nabla_t\eta\|_p\|\xi\|_\infty\Big)\\
&+ c_p\|X\|_\infty^2\Big(\|\xi\|_p + \|\nabla_s\xi\|_p + \|\eta\|_p + \|\nabla_t\eta\|_p\Big)\\
&+ c_p\|X\|_\infty\|\eta\|_\infty\|\nabla_t\xi\|_p\\
\|F_2\|_p \le{}& c_p\|\xi\|_\infty^2\Big(\epsilon^{-2}\big(\|X\|_p + \|\nabla_t X\|_p\big) + \|\nabla_s X\|_p + \|Y\|_p + \|\nabla_s Y\|_p\Big)\\
&+ c_p\|\xi\|_\infty\Big(\frac{\|\xi\|_\infty}{\epsilon^2}\big(\|\xi\|_p + \|\nabla_t\xi\|_p\big) + \|\nabla_s\xi\|_p + \|\eta\|_p\big(1 + \|\nabla_s X\|_\infty\big)\Big)\\
&+ c_p\|X\|_\infty\Big(\epsilon^{-2}\big(\|\xi\|_p + \underline{\|\nabla_t\xi\|_p}\big)\|X\|_\infty + \|\nabla_s\xi\|_p\|\eta\|_\infty\Big)
\end{aligned}$$

and

$$\begin{aligned}
\|\nabla_\alpha F_1\|_p \le{}& c_p\|\xi\|_\infty^2\Big(\|\nabla_t X\|_p + \|Y\|_p + \|\nabla_s X\|_p + \|\nabla_t Y\|_p + \|\nabla_\alpha Y\|_p\\
&\qquad + \|\nabla_\alpha\nabla_t X\|_p + \|\nabla_\alpha\nabla_t Y\|_p + \|\nabla_\alpha\nabla_s X\|_p\Big)\\
&+ c_p\|\xi\|_\infty\Big(\|\xi\|_p + \|\eta\|_p + \|\nabla_t\xi\|_p + \|\nabla_t\eta\|_p + \|\nabla_\alpha\xi\|_p\\
&\qquad + \|\nabla_s\xi\|_p + \|\nabla_\alpha\eta\|_p + \|\nabla_\alpha\nabla_t\xi\|_p + \|\nabla_\alpha\nabla_s\xi\|_p\|\xi\|_\infty\Big)
\end{aligned}$$

$$
\begin{aligned}
&+ c_p\|\xi\|_\infty\Big(\|\nabla_t X\|_p + \|\nabla_\alpha X\|_p + \|\nabla_t Y\|_p\big(\|X\|_\infty + \|\nabla_t X\|_\infty\big) \\
&\qquad + \|\nabla_\alpha Y\|_p\|\nabla_t\xi\|_\infty + \|\nabla_\alpha\nabla_t X\|_p\|\eta\|_\infty\Big) \\
&+ c_p\|X\|_\infty^2\Big(\|\eta\|_p + \|\nabla_t\eta\|_p + \|\nabla_\alpha\nabla_s\xi\|_p\Big) \\
&+ c_p\|X\|_\infty\Big(\|X\|_p + \|\xi\|_p + \|\nabla_\alpha\xi\|_p + \|\nabla_s\xi\|_p + \|\nabla_\alpha\eta\|_p \\
&\qquad + \|\nabla_\alpha\nabla_s\xi\|_p\|\xi\|_\infty + \|\nabla_\alpha\nabla_t\xi\|_p\|\eta\|_\infty\Big) \\
&+ c_p\|\eta\|_\infty\Big(\|\nabla_t\xi\|_p + \|\nabla_\alpha\xi\|_p + \|\nabla_\alpha\nabla_t\xi\|_p\|\xi\|_\infty\Big) \\
&+ c_p\|\nabla_\alpha\xi\|_p\Big(\underline{\|\nabla_t X\|_\infty} + \|\nabla_t\xi\|_\infty\Big)
\end{aligned}
$$

and

$$
\begin{aligned}
&\|\nabla_\alpha F_2\|_p \\
&\quad\leq c_p\|\xi\|_\infty^2\bigg(\epsilon^{-2}\Big(\|X\|_p + \|\nabla_t X\|_p + \|\nabla_\alpha X\|_p + \|\nabla_\alpha\nabla_t X\|_p\Big) \\
&\qquad\quad + \|\nabla_s X\|_p + \|\nabla_\alpha Y\|_p + \|\nabla_s Y\|_p + \|\nabla_\alpha\nabla_s X\|_p + \|\nabla_\alpha\nabla_s Y\|_p\bigg) \\
&\quad + c_p\|\xi\|_\infty\bigg(\epsilon^{-2}\Big(\|\xi\|_p + \|\nabla_t\xi\|_p + \|\nabla_\alpha\xi\|_p + \underline{\|\nabla_\alpha\nabla_t\xi\|_p}\Big) \\
&\qquad\quad + \|\nabla_s\xi\|_p + \|\eta\|_p + \|\nabla_\alpha\eta\|_p + \|\nabla_s\eta\|_p + \|\nabla_\alpha\nabla_s\xi\|_p\bigg) \\
&\quad + c_p\|\xi\|_\infty\bigg(\|X\|_p + \|\nabla_\alpha X\|_p + \|\nabla_\alpha Y\|_p\|\nabla_s\xi\|_\infty \\
&\qquad\quad + \|\nabla_s Y\|_p\Big(\|\nabla_\alpha X\|_\infty + \|\nabla_\alpha\xi\|_\infty\Big)\bigg) \\
&\quad + c_p\|X\|_\infty^2\epsilon^{-2}\bigg(\|\xi\|_p + \|\nabla_t\xi\|_p + \|\nabla_\alpha\xi\|_p + \|\nabla_\alpha\nabla_t\xi\|_p\bigg) \\
&\quad + c_p\|X\|_\infty\bigg(\epsilon^{-2}\Big(\|\xi\|_p + \|\nabla_t\xi\|_p\Big) + \|\nabla_\alpha\xi\|_p + \|\nabla_\alpha\eta\|_p\|\nabla_s\xi\|_\infty \\
&\qquad\quad + \|\nabla_s Y\|_p\|\xi\|_\infty + \|\nabla_\alpha\nabla_s\xi\|_p\|\eta\|_\infty\bigg) \\
&\quad + c_p\|\nabla_\alpha\xi\|_\infty\bigg(\epsilon^{-2}\Big(\|\xi\|_p + \underline{\|\nabla_t\xi\|_p}\Big) + \|\nabla_s\xi\|_p + \|\nabla_s X\|_p\bigg) \\
&\quad + c_p\|\eta\|_\infty\bigg(\|\nabla_\alpha\xi\|_p + \|\nabla_s\xi\|_p + \|\nabla_s X\|_p\|\xi\|_\infty + \|\nabla_\alpha\nabla_s X\|_p\|\xi\|_\infty\bigg)
\end{aligned}
$$

for $\epsilon \in (0,1]$ *and* $\zeta = (\xi,\eta) \in C_0^\infty(\mathbb{R}\times S^1, u^*TM \oplus u^*T^*M)$ *with* $\|\xi\|_\infty < \iota'_M/2$ *and* $\|\eta\|_\infty + \|\nabla_t\xi\|_\infty + \|\nabla_s\xi\|_\infty \leq \sqrt{c_p}/2$.

LEMMA 5.2.2. (**Pointwise estimate I**) *Let $u \in \mathcal{P}_{x,y}$, x, y smooth loops in M, and denote*

$$\mathcal{F}^{triv}_{\epsilon,u}(Z+\zeta) - \mathcal{F}^{triv}_{\epsilon,u}(Z) - d\mathcal{F}^{triv}_{\epsilon,u}(Z)\circ\zeta = \begin{pmatrix} F_1 \\ F_2 \end{pmatrix} = \begin{pmatrix} T_1+T_2+T_3 \\ T_4+T_5+T_6 \end{pmatrix}$$

for $Z=(X,Y)$, $\zeta=(\xi,\eta)\in C^\infty_0(\mathbb{R}\times S^1, u^*TM\oplus u^*T^*M)$ with $\|\xi\|_\infty$, $\|X\|_\infty \le \iota'_M/2$. Then in a local coordinate chart $(U_i\subset V_i,\varphi_i)$ as above the following pointwise estimates hold: There exists a continuous function $c(\vec{\xi})\ge 0$ such that

$$\begin{aligned}
|\vec{F}_1| &\le c(\vec{\xi})\,|\vec{\xi}|^2\Big(1+|\partial_t\vec{X}|\big(1+|\vec{Y}|\big)+|\partial_s\vec{X}|+|\vec{Y}|+|\partial_t\vec{Y}|+|\partial_t\vec{\eta}|\Big)\\
&\quad + c(\vec{\xi})\,|\vec{\xi}|\Big(|\partial_t\vec{\xi}|\big(1+|\vec{\eta}|+|\vec{Y}|\big)+|\partial_s\vec{\xi}|\cdot|\vec{\xi}|+|\vec{\eta}|\big(1+|\partial_t\vec{X}|\big)\Big)\\
&\quad + c(\vec{\xi})\,|\vec{X}|\Big(|\partial_t\vec{\xi}|\cdot|\vec{\eta}|+|\partial_s\vec{\xi}|\cdot|\vec{X}|+|\partial_t\vec{\eta}|\cdot|\vec{X}|\Big)\\
|\vec{F}_2| &\le c(\vec{\xi})\,|\vec{\xi}|^2\Big(\epsilon^{-2}\big(1+|\partial_t\vec{X}|\big)+|\partial_s\vec{X}|\big(1+|\vec{Y}|\big)+|\vec{Y}|+|\partial_s\vec{Y}|\Big)\\
&\quad + c(\vec{\xi})\,|\vec{\xi}|\Big(\epsilon^{-2}|\partial_t\vec{\xi}|\cdot|\vec{\xi}|+|\partial_s\vec{\xi}|\big(1+|\vec{\eta}|+|\vec{Y}|\big)+|\vec{\eta}|\big(1+|\partial_s\vec{X}|\big)\Big)\\
&\quad + c(\vec{\xi})\,|\vec{X}|\Big(\epsilon^{-2}|\partial_t\vec{\xi}|\cdot|\vec{X}|+|\partial_s\vec{\xi}|\cdot|\vec{\eta}|\Big)\\
|\vec{\nabla}_\alpha\vec{F}_1| &\le c\big((79)+(82)+(81)+|\vec{F}_1|\big)\\
|\vec{\nabla}_\alpha\vec{F}_2| &\le c\big((83)+(80)+(78)+|\vec{F}_2|\big)
\end{aligned}$$

for $\epsilon\in(0,1]$, where $|\cdot| = |\cdot|_{T_{u(s,t)}M}$ and an arrow on top of an object indicates that it is represented in local coordinates and is evaluated at (s,t).

PROOF. (**of lemma 5.2.2 – pointwise estimate I**) Throughout the proof we use the notation introduced in remark 5.0.10.

Estimates for F_1 and F_2

The term T_6 : Use the definition (50) of b to get

$$\begin{aligned}
\epsilon^2 T_6 &= -\mathcal{T}(X+\xi)^* b(u, X+\xi, Y+\eta) + \mathcal{T}(X)^* b(u,X,Y)\\
&\quad + \tfrac{d}{d\tau}\Big|_0\Big(\mathcal{T}(X+\tau\xi)^* b(u, X+\tau\xi, Y+\tau\eta)\Big)\\
&= -\eta + \tfrac{d}{d\tau}\Big|_0 (w+Y+\tau\eta),
\end{aligned}$$

i.e.

(68) $$\underline{T_6 = 0}.$$

The term T_3 :

$$\begin{aligned}
T_3 &= -\mathcal{T}^{-1}(X+\xi)\circ\nabla V_t|_{a(u,X+\xi)} + \mathcal{T}^{-1}(X)\circ\nabla V_t|_{a(u,X)} \\
&\quad + \tfrac{d}{d\tau}\big|_0\Big(\mathcal{T}^{-1}(X+\tau\xi)\circ\nabla V_t|_{a(u,X+\tau\xi)}\Big) \\
&= -\Big(\mathcal{T}^{-1}(X+\xi) - \mathcal{T}^{-1}(X) - d\mathcal{T}^{-1}|_X\circ\xi\Big)\circ\nabla V_t|_{a(u,X)} \\
&\quad - \mathcal{T}^{-1}(X+\xi)\circ\Big(\nabla V_t|_{a(u,X+\xi)} - \nabla V_t|_{a(u,X)} - d(\nabla V_t|_{a(u,X)})\circ\xi\Big) \\
&\quad - \Big(\mathcal{T}^{-1}(X+\xi) - \mathcal{T}^{-1}(X)\Big)\circ d(\nabla V_t|_{a(u,X)})\circ\xi \\
&= I + II + III
\end{aligned}$$

where in the 2^{nd} equality we have added twice zero (terms $1+5$ as well as $6+7$). Now use lemma 5.0.9 to get

$$\begin{aligned}
|I| &\le \big\|\mathcal{T}^{-1}(X+\xi) - \mathcal{T}^{-1}(X) - d\mathcal{T}^{-1}|_X\circ\xi\big\| \cdot \big|\nabla V_t|_{a(u,X)}\big| \\
&\le c(X,\xi)\cdot|\xi|^2\cdot c(X) \\
|II| &\le \big\|\mathcal{T}^{-1}(X+\xi)\big\| \cdot \big|\nabla V_t|_{a(u,X+\xi)} - \nabla V_t|_{a(u,X)} - d(\nabla V_t|_{a(u,X)})\circ\xi\big| \\
&\le c(X,\xi)\cdot\tilde{c}(X,\xi)\cdot|\xi|^2 \\
|III| &\le \big\|\mathcal{T}^{-1}(X+\xi) - \mathcal{T}^{-1}(X)\big\| \cdot \big\|d(\nabla V_t)|_{a(u,X)}\circ\partial_2 a(u,X)\circ\xi\big\| \\
&\le c(X,\xi)\cdot|\xi|\cdot c(X)\cdot|\xi|.
\end{aligned}$$

As the constants depend continuously on X and $\|X\|_\infty \le \iota'_M/2$ it follows

$$\underline{|T_3| \le c\,|\xi|^2}. \tag{69}$$

The term T_1 :

$$\begin{aligned}
T_1 &= \mathcal{T}^{-1}(X+\xi)\circ\partial_s a(u,X+\xi) - \mathcal{T}^{-1}(X)\circ\partial_s a(u,X) \\
&\quad - \tfrac{d}{d\tau}\big|_0\Big(\mathcal{T}^{-1}(X+\tau\xi)\circ\partial_s a(u,X+\tau\xi)\Big) \\
&= \Big(\mathcal{T}^{-1}(X+\xi) - \mathcal{T}^{-1}(X) - d\mathcal{T}^{-1}|_X\circ\xi\Big)\circ\partial_s a(u,X) \\
&\quad + \mathcal{T}^{-1}(X+\xi)\circ\Big(\partial_s a(u,X+\xi) - \partial_s a(u,X) - \partial_s(\partial_2 a(u,X)\circ\xi)\Big) \\
&\quad + \Big(\mathcal{T}^{-1}(X+\xi) - \mathcal{T}^{-1}(X)\Big)\circ\partial_s(\partial_2 a(u,X)\circ\xi) \\
&= I + II + III
\end{aligned}$$

where in the 2^{nd} equality we again added twice zero (terms $1+5$ and $6+7$). Using lemma 5.0.9, equations (51) and (52), adding zero and moving a term in III containing $\partial_s\xi$ to II, we obtain

$$
\begin{aligned}
&|I|\\
&\quad\le c(X,\xi)\,|\xi|^2\cdot|\partial_1 a(u,X)\circ\partial_s u+\partial_2 a(u,X)\circ\partial_s X|\\
&\quad\le c(X,\xi)\,|\xi|^2\,c(X)\Big(c+|\partial_s X|\Big)\\
&|III-\mathcal{T}^{-1}(X+\xi)\partial_2 a(u,X)\partial_s\xi+\partial_s\xi|\\
&\quad\le\|\mathcal{T}^{-1}(X+\xi)-\mathcal{T}^{-1}(X)\|\cdot|\partial_1\partial_2 a(u,X)\circ(\partial_s u,\xi)+\partial_2\partial_2 a(u,X)\circ(\partial_s X,\xi)|\\
&\qquad+|-\mathcal{T}^{-1}(X)\partial_2 a(u,X)\circ\partial_s\xi+\partial_s\xi|\\
&\quad\le c(X,\xi)\Big(|\xi|^2+|\xi|^2|\partial_s X|+|X|^2|\partial_s\xi|\Big)\\
&|II+\mathcal{T}^{-1}(X+\xi)\partial_2 a(u,X)\partial_s\xi-\partial_s\xi|\\
&\quad\le\Big|\mathcal{T}^{-1}(X+\xi)\Big(\partial_1 a(u,X+\xi)-\partial_1 a(u,X)-\partial_2\partial_1 a(u,X)\xi\Big)(\partial_s u+\partial_s X)\Big|\\
&\qquad+|\mathcal{T}^{-1}(X+\xi)\partial_2 a(u,X+\xi)\circ\partial_s\xi-\partial_s\xi|\\
&\quad\le c(X,\xi)\Big(|\xi|^2+|\xi|^2\cdot|\partial_s X|+|X+\xi|^2\cdot|\partial_s\xi|\Big)
\end{aligned}
$$

where we applied results derived in appendix A to

$$h(X)=\mathcal{T}^{-1}(X)\partial_2 a(u,X)-1\!\!1.$$

Namely, we get $h(0)=0$ and $dh(0)X=d\mathcal{T}^{-1}|_0(X,\cdot)+\partial_2\partial_2 a(u,0)(X,\cdot)=0$. The assumption on $\|X\|_\infty$ implies

$$\underline{|T_1|\le c\Big(|\xi|^2+|\xi|^2|\partial_s X|+|\xi|^2|\partial_s\xi|+|X|^2|\partial_s\xi|\Big)}.\tag{70}$$

The term T_5 : Again add twice zero in the 2^{nd} equality (as above) to get

$$
\begin{aligned}
\epsilon^2\,T_5&=\mathcal{T}(X+\xi)^*\circ g|_{a(u,X+\xi)}\circ\partial_t a(u,X+\xi)-\mathcal{T}(X)^*\circ g|_{a(u,X)}\circ\partial_t a(u,X)\\
&\qquad-\tfrac{d}{d\tau}\big|_0\Big(\mathcal{T}(X+\tau\xi)^*\circ g|_{a(u,X+\tau\xi)}\circ\partial_t a(u,X+\tau\xi)\Big)\\
&=\Big(\mathcal{T}(X+\xi)^*\circ g|_{a(u,X+\xi)}-\mathcal{T}(X)^*\circ g|_{a(u,X)}\\
&\qquad-d(\mathcal{T}(X)^*\circ g|_{a(u,X)})\circ\xi\Big)\circ\partial_t a(u,X)\\
&\quad+\mathcal{T}(X+\xi)^*g|_{a(u,X+\xi)}\Big(\partial_t a(u,X+\xi)-\partial_t a(u,X)-\partial_t(\partial_2 a(u,X)\circ\xi)\Big)\\
&\quad+\Big(\mathcal{T}(X+\xi)^*\circ g|_{a(u,X+\xi)}-\mathcal{T}(X)^*\circ g|_{a(u,X)}\Big)\circ\partial_t(\partial_2 a(u,X)\circ\xi)\\
&=I+II+III
\end{aligned}
$$

Using lemma 5.0.9, equations (51) and (52), adding zero and moving a term in III containing $\partial_t\xi$ to II, we obtain

$$\begin{aligned}
&|I|\\
&\quad\le c(X,\xi)\,|\xi|^2\,|\partial_1 a(u,X)\circ\partial_t u + \partial_2 a(u,X)\circ\partial_t X|\\
&\quad\le c(X,\xi)\,|\xi|^2\,c(X)\Big(c+|\partial_t X|\Big)\\
&|III - \big(\mathcal{T}(X+\xi)^* g|_{a(u,X+\xi)}\partial_2 a(u,X) - g|_u\big)\,\partial_t\xi|\\
&\quad\le |\mathcal{T}(X+\xi)^* g|_{a(u,X+\xi)} - \mathcal{T}(X)^* g|_{a(u,X)}|\\
&\qquad\cdot|\partial_1\partial_2 a(u,X)\circ(\partial_t u,\xi) + \partial_2\partial_2 a(u,X)\circ(\partial_t X,\xi)|\\
&\qquad+|\,\big(-\mathcal{T}(X)^* g|_{a(u,X)}\partial_2 a(u,X) + g|_u\big)\,\partial_t\xi|\\
&\quad\le c(X,\xi)\Big(|\xi|^2 + |\xi|^2\cdot|\partial_t X| + |X|^2|\partial_t\xi|\Big)\\
&|II + \big(\mathcal{T}(X+\xi)^* g|_{a(u,X+\xi)}\partial_2 a(u,X) - g|_u\big)\,\partial_t\xi|\\
&\quad\le \Big|\mathcal{T}(X+\xi)^* g|_{a(u,X+\xi)}\Big(\partial_1 a(u,X+\xi) - \partial_1 a(u,X) - \partial_2\partial_1 a(u,X)\xi\Big)\partial_t u\Big|\\
&\qquad+\Big|\mathcal{T}(X+\xi)^* g|_{a(u,X+\xi)}\Big(\partial_2 a(u,X+\xi) - \partial_2 a(u,X) - \partial_2\partial_2 a(u,X)\xi\Big)\partial_t X\Big|\\
&\qquad+\Big|\mathcal{T}(X+\xi)^* g|_{a(u,X+\xi)}\partial_2 a(u,X+\xi)\partial_t\xi - g|u\partial_t\xi\Big|\\
&\quad\le c(X,\xi)\Big(|\xi|^2 + |\xi|^2\cdot|\partial_t X| + |X+\xi|^2\cdot|\partial_t\xi|\Big)
\end{aligned}$$

where we used a result derived in the proof of the fundamental quadratic estimate theorem 5.1.1, namely (59) which states that

$$\mathcal{T}(X)^* g|_{a(u,X)}\partial_2 a(u,X) - g|u$$

is of order $|X|^2$. The assumption on $\|X\|_\infty$ implies

$$|T_5| \le c\,\epsilon^{-2}\Big(|\xi|^2 + |\xi|^2|\partial_t X| + |\xi|^2|\partial_t\xi| + |X|^2|\partial_t\xi|\Big). \tag{71}$$

The term T_4 : Add twice zero in the 2^{nd} equality (terms $1+5$ and $6+7$) to get

$$\begin{aligned}
T_4 &= \mathcal{T}(X+\xi)^*\circ\nabla_s^* b(u,X+\xi,Y+\eta) - \mathcal{T}(X)^*\circ\nabla_s^* b(u,X,Y)\\
&\quad - \tfrac{d}{d\tau}\Big|_0\Big(\mathcal{T}(X+\tau\xi)^*\circ\nabla_s^* b(u,X+\tau\xi,Y+\tau\eta)\Big)\\
&= \Big(\mathcal{T}(X+\xi)^* - \mathcal{T}(X)^* - d(\mathcal{T}|_X^*)\circ\xi\Big)\circ\nabla_s^* b(u,X,Y)\\
&\quad + \mathcal{T}(X+\xi)^*\circ\Big(\nabla_s^* b(u,X+\xi,Y+\eta) - \nabla_s^* b(u,X,Y)\\
&\qquad\qquad - \tfrac{d}{d\tau}\Big|_0(\nabla_s^* b(u,X+\tau\xi,Y+\tau\eta))\Big)\\
&\quad + \Big(\mathcal{T}(X+\xi)^* - \mathcal{T}(X)^*\Big)\circ\tfrac{d}{d\tau}\Big|_0\nabla_s^* b(u,X+\tau\xi,Y+\tau\eta)\\
&= I + II + III
\end{aligned}$$

We need to compute

$$\begin{aligned}
&-\nabla_s^* b(u,X,Y)\\
&\quad= -\Big(\partial_s\cdot -\Gamma|_{a(u,X)}(\partial_s a(u,X),\cdot)\Big)\circ\mathcal{T}^{-1}(X)^*\circ(w+Y)\\
&\quad= -d\mathcal{T}^{-1}|_X^*\circ(\partial_s X, w+Y) - \mathcal{T}^{-1}(X)^*\circ(\partial_s w+\partial_s Y)\\
&\qquad + \Gamma|_{a(u,X)}\circ\Big(\partial_1 a(u,X)\circ\partial_s u + \partial_2 a(u,X)\circ\partial_s X\,,\\
&\qquad\qquad \mathcal{T}^{-1}(X)^*\circ(w+Y)\Big)
\end{aligned}\tag{72}$$

and

$$\begin{aligned}
&\nabla_s^* b(u,X+\xi,Y+\eta)\\
&\quad= d\mathcal{T}^{-1}|_{X+\xi}^*\circ(\partial_s X+\partial_s\xi, w+Y+\eta)\\
&\qquad + \mathcal{T}^{-1}(X+\xi)^*\circ(\partial_s w+\partial_s Y+\partial_s\eta)\\
&\qquad - \Gamma|_{a(u,X+\xi)}\circ\Big(\partial_1 a(u,X+\xi)\circ\partial_s u + \partial_2 a(u,X+\xi)\circ(\partial_s X+\partial_s\xi)\,,\\
&\qquad\qquad \mathcal{T}^{-1}(X+\xi)^*\circ(w+Y+\eta)\Big)
\end{aligned}\tag{73}$$

as well as

(74)

$$\begin{aligned}
&\tfrac{d}{d\tau}\big|_0 \nabla_s^* b(u,X+\tau\xi,Y+\tau\eta)\\
&= \tfrac{d}{d\tau}\big|_0\Big(\partial_s\big(\mathcal{T}^{-1}(X+\tau\xi)^*\circ(w+Y+\tau\eta)\big)\\
&\quad - \Gamma|_{a(u,X+\tau\xi)}\big(\partial_s a(u,X+\tau\xi), \mathcal{T}^{-1}(X+\tau\xi)^*\circ(w+Y+\tau\eta)\big)\Big)\\
&= d^2\mathcal{T}^{-1}|_X^*\circ(\partial_s X,\xi,w+Y) + d\mathcal{T}^{-1}|_X^*\circ(\partial_s\xi,w+Y)\\
&\quad + d\mathcal{T}^{-1}|_X^*\circ(\xi,\partial_s w+\partial_s Y) + d\mathcal{T}^{-1}|_X^*\circ(\partial_s X,\eta) + \mathcal{T}^{-1}(X)^*\circ\partial_s\eta\\
&\quad - d\Gamma|_{a(u,X)}\Big(\partial_2 a(u,X)\circ\xi, \partial_1 a(u,X)\circ\partial_s u + \partial_2 a(u,X)\circ\partial_s X,\\
&\qquad\qquad \mathcal{T}^{-1}|_X^*\circ(w+Y)\Big)\\
&\quad - \Gamma|_{a(u,X)}\Big(\partial_1\partial_2 a(u,X)\circ(\partial_s u,\xi) + \partial_2\partial_2 a(u,X)\circ(\partial_s X,\xi)\\
&\qquad\qquad + \partial_2 a(u,X)\circ\partial_s\xi\,,\ \mathcal{T}^{-1}|_X^*\circ(w+Y)\Big)\\
&\quad - \Gamma|_{a(u,X)}\Big(\partial_1 a(u,X)\partial_s u + \partial_2 a(u,X)\partial_s X, d\mathcal{T}^{-1}|_X^*(\xi,w+Y) + \mathcal{T}^{-1}(X)^*\eta\Big)
\end{aligned}$$

where we used $\frac{d}{d\tau}\big|_0\partial_s\cdot = \partial_s\frac{d}{d\tau}\big|_0$ as well as equations (51) and (52). Using lemma 5.0.9 and equation (72) we can estimate I

$$\begin{aligned}
|I| &\le \|\mathcal{T}(X+\xi)^* - \mathcal{T}(X)^* - d\mathcal{T}|_X^*\circ\xi\|\cdot|\nabla_s^* b(u,X,Y)\|\\
&\le c\,|\xi|^2\Big(1+|\partial_s X|+|\partial_s X|\cdot|Y|+|Y|+|\partial_s Y|\Big)
\end{aligned}$$

where we used $\|X\|_\infty \le \iota'_M/2$. In what follows we transfer the terms containing $\partial_s\eta$ from III to II so that they will disappear. Equation (74) leads to

$$\begin{aligned}
&|III - \mathcal{T}(X+\xi)^*\mathcal{T}^{-1}(X)^*\partial_s\eta + \partial_s\eta| \\
&\le \|\mathcal{T}(X+\xi)^* - \mathcal{T}(X)^*\| \cdot \left|\tfrac{d}{d\tau}\big|_0 \nabla_s^* b(u, X+\tau\xi, Y+\tau\eta)\right\| \\
&\le c(X,\xi)\ |\xi|\ c(X)\Big(|\xi|\cdot|\partial_s X| + |\xi|\cdot|\partial_s X|\cdot|Y| \\
&\quad + |\partial_s\xi| + |Y|\cdot|\partial_s\xi| + |\xi| + |\xi|\cdot|\partial_s Y| + |\eta|\cdot|\partial_s X| + 0 \\
&\quad + |\xi| + |\xi|\cdot|Y| + |\partial_s X|\cdot|\xi| + |\partial_s X|\cdot|\xi|\cdot|Y| + |\partial_s\xi| + |\partial_s\xi|\cdot|Y| \\
&\quad + |\xi| + |\xi|\cdot|Y| + |\partial_s X|\cdot|\xi| + |\partial_s X|\cdot|\xi|\cdot|Y| + |\eta| + |\partial_s X|\cdot|\eta|\Big) \\
&\le c\,|\xi|\Big(|\xi| + |\partial_s\xi| + |\xi|\cdot|\partial_s X| + |\xi|\cdot|\partial_s Y| + |\eta| + 0 \\
&\quad + |\eta|\cdot|\partial_s X| + |\xi|\cdot|Y| + |\partial_s\xi|\cdot|Y| + |\xi|\cdot|Y|\cdot|\partial_s X|\Big)
\end{aligned}$$

where in the last step we used $\|X\|_\infty \le \iota'_M/2$. Now we estimate term II using equations (73), (72) and (74) as well as lemma 5.0.9

$$\begin{aligned}
&|II + \mathcal{T}(X+\xi)^*\mathcal{T}^{-1}(X)^*\partial_s\eta - \partial_s\eta| \\
&\le c(X,\xi)\ \Big|\nabla_s^* b(u, X+\xi, Y+\eta) - \nabla_s^* b(u,X,Y) \\
&\qquad\qquad - \tfrac{d}{d\tau}\big|_0 \nabla_s^* b(u, X+\tau\xi, Y+\tau\eta)\Big| \\
&\le c(X,\xi)\Bigg(0 + \left\|\mathcal{T}^{-1}(X+\xi)^* - \mathcal{T}^{-1}(X)^* - d\mathcal{T}^{-1}|_X^*\circ\xi\right\| \cdot |\partial_s w + \partial_s Y| \\
&\quad + \left\|d\mathcal{T}^{-1}|_{X+\xi}^* - d\mathcal{T}^{-1}|_X^* - d^2\mathcal{T}^{-1}|_X^*\circ\xi\right\| \cdot |\partial_s X|\cdot|w+Y| \\
&\quad + \left\|d\mathcal{T}^{-1}|_{X+\xi}^* - d\mathcal{T}^{-1}|_X^*\right\| \cdot |\partial_s X|\cdot|\eta| \\
&\quad + \left\|d\mathcal{T}^{-1}|_{X+\xi}^* - d\mathcal{T}^{-1}|_X^*\right\| \cdot |\partial_s \xi|\cdot|w+Y| \\
&\quad + \Big|\Gamma|_{a(u,X)}\Big(\partial_1 a(u,X)\circ\partial_s u + \partial_2 a(u,X)\circ\partial_s X, \mathcal{T}^{-1}|_X^*(w+Y)\Big) \\
&\quad - \Gamma|_{a(u,X+\xi)}\Big(\partial_1 a(u,X+\xi)\circ\partial_s u + \partial_2 a(u,X+\xi)\circ\partial_s X, \mathcal{T}^{-1}|_{X+\xi}^*(w+Y)\Big) \\
&\quad + d\Big(\Gamma|_{a(u,X)}\big(\partial_1 a(u,X)\circ\partial_s u + \partial_2 a(u,X)\circ\partial_s X\ ,\ \mathcal{T}^{-1}|_X^*(w+Y)\big)\Big)\circ\xi\Big| \\
&\quad + \Big|-\Gamma|_{a(u,X+\xi)}\Big(\partial_1 a(u,X+\xi)\circ\partial_s u + \partial_2 a(u,X+\xi)\circ\partial_s X, \mathcal{T}^{-1}|_{X+\xi}^*\circ\eta\Big) \\
&\quad + \Gamma|_{a(u,X)}\Big(\partial_1 a(u,X)\circ\partial_s u + \partial_2 a(u,X)\circ\partial_s X, \mathcal{T}^{-1}|_X^*\circ\eta\Big)\Big| \\
&\quad + \Big|-\Gamma|_{a(u,X+\xi)}\Big(\partial_2 a(u,X+\xi)\circ\partial_s\xi, \mathcal{T}^{-1}|_{X+\xi}^*\circ(w+Y)\Big) \\
&\quad + \Gamma|_{a(u,X)}\Big(\partial_2 a(u,X)\circ\partial_s\xi, \mathcal{T}^{-1}|_X^*\circ(w+Y)\Big)\Big| \\
&\quad + \Big\|d\mathcal{T}^{-1}|_{X+\xi}^*(\partial_s\xi,\eta) - \Gamma|_{a(u,X+\xi)}\Big(\partial_2 a(u,X+\xi)\circ\partial_s\xi, \mathcal{T}^{-1}|_{X+\xi}^*\circ\eta\Big)\Big\|\Bigg)
\end{aligned}$$

where the last term in the sum is less or equal to

$$c(X,\xi)\,|X+\xi|\cdot|\partial_s\xi|\cdot|\eta|\leq c(X,\xi)\,|\eta|\cdot|\partial_s\xi|\,(|\xi|+|X|)\,.$$

Inspecting the estimate above term by term leads to

$$\begin{aligned}|II|\leq c(X,\xi)\Big(&0+|\xi|^2\big(1+|\partial_sY|\big)+|\xi|^2|\partial_sX|\big(1+|Y|\big)\\&+|\xi|\cdot|\partial_sX|\cdot|\eta|+|\xi|\cdot|\partial_s\xi|\big(1+|Y|\big)\\&+|\xi|^2\big(1+|Y|+|\partial_sX|+|\partial_sX|\cdot|Y|\big)+|\xi|\cdot|\eta|\big(1+|\partial_sX|\big)\\&+|\xi|\cdot|\partial_s\xi|\big(1+|Y|\big)+(|\xi|+|X|)\,|\partial_s\xi|\cdot|\eta|\Big)\\\leq c\,|\xi|\Big(&|\partial_s\xi|+|\xi|\cdot(1+|\partial_sX|+|Y|+|\partial_sY|)+|\eta|\,(1+|\partial_sX|)\\&+0+|\xi|\cdot|\partial_sX|\cdot|Y|+|\partial_s\xi|\,(1+|Y|)\Big)+c\,|\eta|\cdot|\partial_s\xi|\Big(|\xi|+|X|\Big)\end{aligned}$$

where we used $\|X\|_\infty\leq\iota'_M/2$. The estimates for I, II and III give

$$\boxed{\begin{aligned}|T_4|\leq c\,|\xi|\bigg(&|\xi|\Big(1+|\partial_sX|+|\partial_sX|\cdot|Y|+|Y|+|\partial_sY|\Big)+|\partial_s\xi|\\&+|\partial_s\xi|\cdot|Y|+|\eta|\,(1+|\partial_sX|)\bigg)+c\,|\eta|\cdot|\partial_s\xi|\Big(|\xi|+|X|\Big).\end{aligned}}\tag{75}$$

The term T_2 : This term is quite similar to T_4, we therefore only give the results of the calculations. Add twice zero in the 2^{nd} equality (terms $1+5$ and $6+7$) to get

$$\begin{aligned}T_2&=-\mathcal{T}^{-1}(X+\xi)\circ g^{-1}|_{a(u,X+\xi)}\circ\nabla_t^*b(u,X+\xi,Y+\eta)\\&\quad+\mathcal{T}^{-1}(X)\circ g^{-1}|_{a(u,X)}\circ\nabla_t^*b(u,X,Y)\\&\quad+d\mathcal{T}^{-1}|_X\circ\Big(\xi\;,\;g^{-1}|_{a(u,X)}\circ\nabla_t^*b(u,X,Y)\Big)\\&\quad+\mathcal{T}^{-1}(X)\circ\tfrac{d}{d\tau}\big|_0\Big(g^{-1}|_{a(u,X+\tau\xi)}\circ\nabla_t^*b(u,X+\tau\xi,Y+\tau\eta)\Big)\\&=-\Big(\mathcal{T}^{-1}(X+\xi)-\mathcal{T}^{-1}(X)-d\mathcal{T}^{-1}|_X\circ\xi\Big)\circ g^{-1}|_{a(u,X)}\circ\nabla_t^*b(u,X,Y)\\&\quad-\mathcal{T}^{-1}(X+\xi)\circ\\&\qquad\Big(g^{-1}|_{a(u,X+\xi)}\circ\nabla_t^*b(u,X+\xi,Y+\eta)-g^{-1}|_{a(u,X)}\circ\nabla_t^*b(u,X,Y)\\&\qquad-\tfrac{d}{d\tau}\big|_0(g^{-1}|_{a(u,X+\tau\xi)}\circ\nabla_t^*b(u,X+\tau\xi,Y+\tau\eta))\Big)\\&\quad-\Big(\mathcal{T}^{-1}|_{X+\xi}-\mathcal{T}^{-1}|_X\Big)\circ\tfrac{d}{d\tau}\big|_0\Big(g^{-1}|_{a(u,X+\tau\xi)}\nabla_t^*b(u,X+\tau\xi,Y+\tau\eta)\Big)\\&=I+II+III.\end{aligned}$$

Using equation (72) with s replaced by t together with lemma 5.0.9 and the fact that g^{-1} is an isometry we get

$$\begin{aligned}|I| &\le c(X,\xi)\,|\xi|^2\,|\nabla_t^* b(u,X,Y)\| \\ &\le c(X,\xi)\,|\xi|^2\,c(X)\Big(|\partial_t X| + |\partial_t X|\cdot|Y| + 1 + |\partial_t Y| + 1 + |Y| \\ &\qquad + |\partial_t X| + |\partial_t X|\cdot|Y|\Big) \\ &\le c\,|\xi|^2\Big(1 + |\partial_t X| + |\partial_t X|\cdot|Y| + |Y| + |\partial_t Y|\Big)\end{aligned}$$

where we used $\|X\|_\infty \le \iota'_M/2$. To estimate term III we need to calculate

$$\begin{aligned}&\tfrac{d}{d\tau}\big|_0\Big(g^{-1}|_{a(u,X+\tau\xi)}\nabla_t^* b(u,X+\tau\xi,Y+\tau\eta)\Big) \\ &\quad = dg^{-1}|_{a(u,X)}\circ\Big(\partial_2 a(u,X)\circ\xi,\ \nabla_t^* b(u,X,Y)\Big) \\ &\qquad + g^{-1}|_{a(u,X)}\circ\tfrac{d}{d\tau}\big|_0\nabla_t^* b(u,X+\tau\xi,Y+\tau\eta).\end{aligned} \tag{76}$$

Now replace the terms involving b by equations (72), (74) with s replaced by t and transfer two terms from III to II, then

$$\begin{aligned}&|III + \big(\mathcal{T}^{-1}(X+\xi) - \mathcal{T}^{-1}(X)\big)\,g^{-1}|_{a(u,X)}\mathcal{T}^{-1}(X)^*\partial_t\eta| \\ &\quad\le c\,|\xi|\Big(|\xi|\,(1 + |\partial_t X| + |\partial_t X|\cdot|Y| + |Y| + |\partial_t Y|) \\ &\qquad + |\partial_t\xi| + |\partial_t\xi|\cdot|Y| + |\eta| + |\eta|\cdot|\partial_t X| + 0\Big)\end{aligned}$$

where we used $\|X\|_\infty \le \iota'_M/2$. To estimate term II insert equations (73) and (72) with s replaced by t as well as equation (76) to get exactly the same estimate as for term II in T_4 just with s replaced by t (the extra metric term present here does not contribute additional terms to the final estimate)

$$\begin{aligned}&|II - \big(\mathcal{T}^{-1}(X+\xi) - \mathcal{T}^{-1}(X)\big)\,g^{-1}|_{a(u,X)}\mathcal{T}^{-1}(X)^*\partial_t\eta| \\ &\quad\le c\,|\xi|\Big(|\xi|\,(1 + |\partial_t X| + |\partial_t X|\cdot|Y| + |Y| + |\partial_t Y|) + |\partial_t\eta|\cdot|\xi| \\ &\qquad + |\partial_t\xi|\,(1+|Y|) + |\eta|\,(1+|\partial_t X|)\Big) + c\,|X|^2|\partial_t\eta| \\ &\qquad + c\,|\eta|\cdot|\partial_t\xi|\Big(|\xi| + |X|\Big)\end{aligned}$$

where we used $\|X\|_\infty \le \iota'_M/2$. Note that the terms containing $\partial_t\eta$ appeared as follows:

$$\begin{aligned}&\Big(\mathcal{T}^{-1}(X)g^{-1}|_{a(u,X)}\mathcal{T}^{-1}(X)^*\partial_t\eta - g^{-1}|_u\partial_t\eta\Big) \\ &+ \Big(-\mathcal{T}^{-1}(X+\xi)g^{-1}|_{a(u,X+\xi)}\mathcal{T}^{-1}(X+\xi)^*\partial_t\eta + g^{-1}|_u\partial_t\eta\Big) \\ &= h_1(X) + h_2(X).\end{aligned}$$

Now use that $h_1(0) = 0$ and $dh_1(0)X = 0$ and so h_1 contributes $|X|^2|\partial_t\eta|$ and h_2 contributes $|X+\xi|^2|\partial_t\eta|$. These results together imply

$$\boxed{\begin{aligned}|T_2| \le c\,|\xi|\Big(&|\xi|\,(1+|\partial_t X| + |\partial_t X|\cdot|Y| + |Y| + |\partial_t Y|) + |\partial_t\eta|\cdot|\xi| \\ &+ |\partial_t\xi|\,(1+|\eta|+|Y|) + |\eta|\,(1+|\partial_t X|)\Big) \\ &+ c\,|X|\Big(|\partial_t\xi|\cdot|\eta| + |\partial_t\eta|\cdot|X|\Big).\end{aligned}} \tag{77}$$

Estimates for $\nabla_\alpha F_1$ and $\nabla^*_\alpha F_2$

Using the same notation as above for $F_1 = T_1 + T_2 + T_3$ and $F_2 = T_4 + T_5 + T_6$ we obtain

$$|\ \nabla_\alpha F_1\ | = |\ \partial_\alpha F_1 + \Gamma(u)(\partial_\alpha u, F_1)\ | \le |\ \partial_\alpha(T_1+T_2+T_3)\ | + c\ |\ F_1\ |$$

and similarly for $\nabla_\alpha F_2$. We derive estimates for the individual terms $\partial_\alpha T_i$, however in some cases interactions between them have to be taken into account. Moreover, we only indicate the main steps of the calculations as they involve the *same* techniques as in the case of the F_i's considered above. The difference is that the extra partial derivative ∂_α considered here blows up the number of terms involved by a great factor and full details would require some extra thirty pages – of local calculations.

The term $\partial_\alpha T_6$: As we derived above $T_6 = 0$ and so

$$\underline{\partial_\alpha T_6 = 0}. \tag{78}$$

The term $\partial_\alpha T_1$:

$$\begin{aligned}\partial_\alpha T_1 &= \partial_\alpha\Big(\big(\mathcal{T}^{-1}(X+\xi) - \mathcal{T}^{-1}(X) - d\mathcal{T}^{-1}|_X\circ\xi\big)\circ\partial_s a(u,X)\Big) \\ &\quad + \partial_\alpha\Big(\mathcal{T}^{-1}(X+\xi)\circ\big(\partial_s a(u,X+\xi) - \partial_s a(u,X) - \partial_s(\partial_2 a(u,X)\circ\xi)\big)\Big) \\ &\quad + \partial_\alpha\Big(\big(\mathcal{T}^{-1}(X+\xi) - \mathcal{T}^{-1}(X)\big)\circ\partial_s(\partial_2 a(u,X)\circ\xi)\Big) \\ &=: \partial_\alpha I + \partial_\alpha II + \partial_\alpha III\end{aligned}$$

and straightforward calculation leads to

$$\begin{aligned}|\partial_\alpha I| \le c(u,X,\xi,\partial_s u,\partial_\alpha u)\,|\xi|\Big(&|\partial_\alpha\xi|(1+|\partial_s X|) \\ &+ |\xi|\big(1+|\partial_\alpha X| + |\partial_s X|(1+|\partial_\alpha X|)\big)\Big).\end{aligned}$$

Moving a term from III to II we get

$$\begin{aligned}&|\partial_\alpha II + \partial_\alpha\big(\mathcal{T}^{-1}(X+\xi)\partial_2 a(u,X)\partial_s\xi - \partial_s\xi\big)| \\ &\le c(u,X,\xi,\partial_\alpha u,\partial_s u)\Big(|\xi|^2\big(|\partial_\alpha\xi| + |\partial_\alpha\partial_s\xi| + |\partial_\alpha X|(1+|\partial_s X|) + |\partial_\alpha\partial_s X|\big)\end{aligned}$$

$$+|\xi|\big(|\xi|(1+|\partial_\alpha\partial_s u|)+|\partial_s\xi|(1+|\partial_\alpha\xi|+|\partial_\alpha X|)+|\partial_s X||\partial_\alpha\xi|+|\partial_\alpha X|\big)$$
$$+|X|\big(|\partial_\alpha X|+|\partial_\alpha\xi|+|\partial_s\xi|(1+|\partial_\alpha\xi|+|\partial_\alpha X|)+|\xi|\cdot|\partial_\alpha\partial_s\xi|\big)$$
$$+|X|^2|\partial_\alpha\partial_s\xi|\Big)$$

and

$$\begin{aligned}|\partial_\alpha III-\partial_\alpha\big(\mathcal{T}^{-1}(X+\xi)\partial_2 a(u,X)\partial_s\xi-\partial_s\xi|\big)&\\ \le c(u,X,\xi,\partial_\alpha u,\partial_s u)\Big(&|\xi|^2\big(|\partial_\alpha\xi|+|\partial_\alpha\partial_s\xi|(1+|\partial_\alpha X|)\\ &+|\partial_\alpha X|(1+|\partial_s X|)+|\partial_s X|+|\partial_\alpha\partial_s X|\big)\\ &+|\xi|\big(|\xi||\partial_\alpha\partial_s u|+|\partial_\alpha\xi|\cdot|\partial_s X|\big)\\ &+|X|\cdot|\partial_s\xi|(1+|\partial_\alpha X|)+|X|^2|\partial_\alpha\partial_s\xi|\Big).\end{aligned}$$

Altogether we obtain

(79)

$$\begin{aligned}\underline{|\partial_\alpha T_1|}\le c(u,X,\xi,\partial_\alpha u,\partial_s u,\partial_\alpha\partial_s u)\Big(&|\xi|^2\big(|\partial_\alpha\partial_s\xi|+|\partial_\alpha X|(1+|\partial_s X|)\\ &+|\partial_\alpha\partial_s X|\big)+|\xi|\big(|\xi|+|\partial_\alpha\xi|(1+|\partial_s X|)+|\partial_s\xi|(1+|\partial_\alpha\xi|+|\partial_\alpha X|)\big)\\ &+|X|\big(|\partial_\alpha X|+|\partial_\alpha\xi|+|\partial_s\xi|(1+|\partial_\alpha\xi|+|\partial_\alpha X|)+|\xi|\cdot|\partial_\alpha\partial_s\xi|\big)\\ &+|X|^2|\partial_\alpha\partial_s\xi|\Big).\end{aligned}$$

The term $\boldsymbol{\partial_\alpha T_5}$: A similar calculation as for term $\partial_\alpha T_1$ leads to

$$\begin{aligned}\underline{|\partial_\alpha T_5|}\le\epsilon^{-2}c(u,X,\xi,\partial_t u,\partial_\alpha u,\partial_\alpha\partial_t u)\Bigg(&|\partial_\alpha\xi|\cdot|\partial_t\xi|(1+|\partial_t X|)\\ &+|\xi|\Big(|\xi|+|\partial_t\xi|(1+|\partial_\alpha X|)+|\partial_\alpha\xi|(1+|\partial_t X|)+|\partial_\alpha\partial_t\xi|\Big)\\ &+|X|\Big(|\partial_t\xi|(1+|\partial_\alpha X|)+(|X|+|\xi|)|\partial_\alpha\partial_t\xi|\Big)\\ &+|\xi|^2(1+|\partial_t X|)\Big(|\partial_\alpha X|+|\partial_t X|(1+|\partial_\alpha X|)+|\partial_\alpha\partial_t X|\Big)\Bigg).\end{aligned}\tag{80}$$

The term $\boldsymbol{\partial_\alpha T_3}$:

$$\begin{aligned}\partial_\alpha T_3=&-\partial_\alpha\Big(\big(\mathcal{T}^{-1}(X+\xi)-\mathcal{T}^{-1}(X)-d\mathcal{T}^{-1}|_X\circ\xi\big)\circ\nabla V_t|_{a(u,X)}\Big)\\ &-\partial_\alpha\Big(\mathcal{T}^{-1}(X+\xi)\circ\big(\nabla V_t|_{a(u,X+\xi)}-\nabla V_t|_{a(u,X)}-d(\nabla V_t|_{a(u,X)})\circ\xi\big)\Big)\\ &-\partial_\alpha\Big(\big(\mathcal{T}^{-1}(X+\xi)-\mathcal{T}^{-1}(X)\big)\circ d(\nabla V_t|_{a(u,X)})\circ\xi\Big).\end{aligned}$$

Now compactness of M implies that $|\nabla V_t(p)|$ is uniformly bounded for all $(t,p)\in S^1\times M$. Similarly such a uniform bound exists for $|(\partial_\alpha\nabla V_t)(p)|$ in case $\alpha=t$. For $\alpha=s$ this expression is zero anyway since the potential

V does not depend explicitely on s. Now straightforward application of product and chain rule leads to

$$\underline{|\partial_\alpha T_3|} \leq c(u,X,\xi,\partial_\alpha u)|\xi|\Big(|\xi| + |\partial_\alpha \xi| + |\partial_\alpha X|\Big). \tag{81}$$

The term $\boldsymbol{\partial_\alpha T_2}$: Note that $\partial_\alpha T_4$ is quite similar to $\partial_\alpha T_2$ except for the missing metric term g^{-1} in front of it and partial s-derivatives instead of partial t-derivatives. In what follows we underline all terms which vanish in case of $\partial_\alpha T_4$. Let us now get started

$$\begin{aligned}
T_2 = \partial_\alpha\Bigg(& -\mathcal{T}^{-1}(X+\xi)g^{-1}|_{a(u,X+\xi)}\nabla_t^* b(u,X+\xi,Y+\eta) \\
& + \mathcal{T}^{-1}(X)g^{-1}|_{a(u,X)}\nabla_t^* b(u,X,Y) \\
& + \tfrac{d}{d\tau}\big|_0\Big(\mathcal{T}^{-1}(X+\tau\xi)g^{-1}|_{a(u,X+\tau\xi)}\nabla_t^* b(u,X+\tau\xi,Y+\tau\eta)\Big)\Bigg) \\
=: & \ \partial_\alpha I + \partial_\alpha II + \partial_\alpha III
\end{aligned}$$

where

$$\begin{aligned}
\partial_\alpha I &= -\partial_\alpha\Bigg(\big(\mathcal{T}^{-1}(X+\xi) - \mathcal{T}^{-1}(X) - d\mathcal{T}^{-1}|_X\xi\big)g^{-1}|_{a(u,X)}\nabla_t^* b(u,X,Y)\Bigg) \\
\partial_\alpha II &= -\partial_\alpha\Bigg(\mathcal{T}^{-1}(X+\xi)\Big(g^{-1}|_{a(u,X+\xi)}\nabla_t^* b(u,X+\xi,Y+\eta) \\
&\quad - g^{-1}|_{a(u,X)}\nabla_t^* b(u,X,Y) \\
&\quad - \tfrac{d}{d\tau}\big|_0(g^{-1}|_{a(u,X+\tau\xi)}\nabla_t^* b(u,X+\tau\xi,Y+\tau\eta))\Big) \\
&\quad + \big(\mathcal{T}^{-1}(X+\xi) - \mathcal{T}^{-1}(X)\big)g^{-1}|_{a(u,X)}\mathcal{T}^{-1}(X)^*\partial_t\eta\Bigg)
\end{aligned}$$

and

$$\begin{aligned}
\partial_\alpha III = -\partial_\alpha\Bigg(& \big(\mathcal{T}^{-1}|_{X+\xi} - \mathcal{T}^{-1}|_X\big) \\
& \tfrac{d}{d\tau}\big|_0\big(g^{-1}|_{a(u,X+\tau\xi)}\nabla_t^* b(u,X+\tau\xi,Y+\tau\eta)\big) \\
& - \big(\mathcal{T}^{-1}(X+\xi) - \mathcal{T}^{-1}(X)\big)g^{-1}|_{a(u,X)}\mathcal{T}^{-1}(X)^*\partial_t\eta\Bigg).
\end{aligned}$$

The first term $\partial_\alpha I$ clearly is the easiest one, with $w = g(u)\partial_t u$ and

$$c = c(u,X,\xi;\partial_t u,\partial_\alpha u,w,\partial_t w,\partial_\alpha w,\partial_\alpha\partial_t w)$$

we arrive at

$$\begin{aligned}
\frac{|\partial_\alpha I|}{c} \leq & \ |\xi|\cdot|\partial_\alpha\xi|\big(1+|Y|+|\partial_t X|(1+|Y|)+|\partial_t Y|\big) \\
& + |\xi|^2(1+|Y|)\big(1+|\partial_t X|+|\partial_\alpha X|(1+|\partial_t X|)+|\partial_\alpha\partial_t X|\big) \\
& + |\xi|^2\big(|\partial_t Y|(\underline{1}+|\partial_\alpha X|)+|\partial_\alpha Y|(1+|\partial_t X|)+|\partial_\alpha\partial_t Y|\big).
\end{aligned}$$

The term $\partial_\alpha III$ should be rewritten as follows

$$
\begin{aligned}
|\partial_\alpha III| = \Big| &-\big(d\mathcal{T}^{-1}|_{X+\xi}(\partial_\alpha X + \partial_\alpha \xi) - d\mathcal{T}^{-1}|_X \partial_\alpha X\big)\circ \\
&\quad dg^{-1}|_{a(u,X)}(\partial_2 a|_{u,X}\xi, \nabla_t^* b|_{u,X,Y})\Big| \\
&+ \Big| -\big(\mathcal{T}^{-1}(X+\xi) - \mathcal{T}^{-1}(X)\big)\circ \\
&\quad \Big(d^2 g^{-1}|_{a(u,X)}(\partial_\alpha a|_{u,X}, \partial_2 a|_{u,X}\xi, \nabla_t^* b|_{u,X,Y}) \\
&\quad + dg^{-1}|_{a(u,X)}\big(\partial_\alpha(\partial_2 a|_{u,X}\xi), \nabla_t^* b|_{u,X,Y}\big) \\
&\quad + dg^{-1}|_{a(u,X)}(\partial_2 a|_{u,X}\xi, \partial_\alpha \nabla_t^* b|_{u,X,Y}) \\
&\quad + dg^{-1}|_{a(u,X)}(\partial_\alpha a|_{u,X}, \tfrac{d}{d\tau}\big|_0 \nabla_t^* b|_{u,X+\tau\xi,Y+\tau\eta}) \\
&\quad - dg^{-1}|_{a(u,X)}(\partial_\alpha a|_{u,X}, \mathcal{T}^{-1}(X)^*\partial_t \eta)\Big)\Big| \\
&+ \Big| -\big(d\mathcal{T}^{-1}|_{X+\xi}(\partial_\alpha X + \partial_\alpha \xi) - d\mathcal{T}^{-1}|_X \partial_\alpha X\big)\circ \\
&\quad \big(g^{-1}|_{a(u,X)} \tfrac{d}{d\tau}\big|_0 \nabla_t^* b(u, X+\tau\xi, Y+\tau\eta) \\
&\qquad - g^{-1}|_{a(u,X)}\mathcal{T}^{-1}(X)^*\partial_t\eta\big) \\
&\quad - \big(\mathcal{T}^{-1}(X+\xi) - \mathcal{T}^{-1}(X)\big)\circ \\
&\quad \big(g^{-1}|_{a(u,X)}\partial_\alpha \tfrac{d}{d\tau}\big|_0 \nabla_t^* b(u, X+\tau\xi, Y+\tau\eta) \\
&\qquad - g^{-1}|_{a(u,X)}(d\mathcal{T}^{-1}|_X^*(\partial_\alpha X, \partial_t\eta) + \mathcal{T}^{-1}(X)^*\partial_\alpha\partial_t\eta)\big)\Big|
\end{aligned}
$$

in order to obtain

$$
\begin{aligned}
\frac{|\partial_\alpha III|}{c} \le\ & |\xi|^2(1+|Y|)\Big(|\partial_t X|(1+|\partial_\alpha X|) + |\partial_\alpha X|(1+|\partial_t Y|) + |\partial_\alpha\partial_t X| \\
& + |\partial_\alpha\partial_t Y|\Big) + \underline{|\xi|^2}\Big(|\partial_t Y| + |\partial_\alpha Y|(1+|\partial_t X|)\Big) \\
& + |\xi|(1+|Y|)\Big(|\xi| + |\partial_t\xi|(1+|\partial_\alpha X|) + |\partial_\alpha\xi|(1+|\partial_t X|) \\
& + |\partial_\alpha\partial_t\xi|\Big) \\
& + |\xi|\Big(|\partial_t\xi|(1+|\partial_\alpha Y|) + |\partial_\alpha\xi|\cdot|\partial_t Y| + |\eta| \\
& + |\eta|\big(|\partial_t X|(1+|\partial_\alpha X|) + |\partial_\alpha X| + |\partial_\alpha\partial_t X|\big) + |\partial_\alpha\eta|(1+|\partial_t X|)\Big) \\
& + |\partial_\alpha\xi|\Big(|\partial_t\xi|(1+|Y|) + |\eta|(1+|\partial_t X|)\Big).
\end{aligned}
$$

The term of greatest complexity is $\partial_\alpha II$, which we rewrite as follows

$$\partial_\alpha II = a_1 + a_2 + a_3 + a_4$$

where

$$
\begin{aligned}
a_1 = &-d\mathcal{T}^{-1}|_{X+\xi}(\partial_\alpha X + \partial_\alpha \xi)\Big(g^{-1}|_{a(u,X+\xi)}\nabla_t^* b|_{u,X+\xi,Y+\eta} \\
&- g^{-1}|_{a(u,X)}\nabla_t^* b|_{u,X,Y} - dg^{-1}|_{a(u,X)}(\partial_2 a(u,X)\xi, \nabla_t^* b|_{u,X,Y}) \\
&- g^{-1}|_{a(u,X)}\tfrac{d}{d\tau}\big|_0 \nabla_t^* b|_{u,X+\tau\xi,Y+\tau\eta} + g^{-1}|_{a(u,X)}\mathcal{T}^{-1}(X)^*\partial_t\eta \\
&- g^{-1}|_{a(u,X+\xi)}\mathcal{T}^{-1}(X+\xi)^*\partial_t\eta\Big) \\
a_2 = &-\mathcal{T}^{-1}(X+\xi)\Big(g^{-1}|_{a(u,X+\xi)}\partial_\alpha\nabla_t^* b|_{u,X+\xi,Y+\eta} - g^{-1}|_{a(u,X)}\partial_\alpha\nabla_t^* b|_{u,X,Y} \\
&- dg^{-1}|_{a(u,X)}\big(\partial_2 a(u,X)\xi, \partial_\alpha\nabla_t^* b|_{u,X,Y}\big) \\
&- g^{-1}|_{a(u,X)}\partial_\alpha \tfrac{d}{d\tau}\big|_0 \nabla_t^* b|_{u,X+\tau\xi,Y+\tau\eta} - g^{-1}|_{a(u,X+\xi)}d\mathcal{T}^{-1}|_{X+\xi}^*(\partial_\alpha\xi, \partial_t\eta) \\
&- g^{-1}|_{a(u,X+\xi)}\mathcal{T}^{-1}(X+\xi)^*\partial_\alpha\partial_t\eta + g^{-1}|_{a(u,X)}\mathcal{T}^{-1}(X)^*\partial_\alpha\partial_t\eta\Big) \\
a_3 = &-\mathcal{T}^{-1}(X+\xi)\Big(dg^{-1}|_{a(u,X+\xi)}\big(\partial_\alpha a(u,X+\xi), \nabla_t^* b|_{u,X+\xi,Y+\eta}\big) \\
&- dg^{-1}|_{a(u,X)}\big(\partial_\alpha a(u,X), \nabla_t^* b|_{u,X,Y}\big) \\
&- d^2 g^{-1}|_{a(u,X)}(\partial_\alpha a(u,X), \partial_2 a(u,X)\xi, \nabla_t^* b|_{u,X,Y}) \\
&- dg^{-1}|_{a(u,X)}\big(\partial_\alpha(\partial_2 a(u,X)\xi), \nabla_t^* b|_{u,X,Y}\big) \\
&- dg^{-1}|_{a(u,X)}\big(\partial_\alpha a(u,X), \tfrac{d}{d\tau}\big|_0 \nabla_t^* b|_{u,X+\tau\xi,Y+\tau\eta}\big) \\
&- dg^{-1}|_{a(u,X+\xi)}\big(\partial_2 a(u,X+\xi)\partial_\alpha\xi, \mathcal{T}^{-1}(X+\xi)^*\partial_t\eta\big)\Big).
\end{aligned}
$$

and

$$
\begin{aligned}
|a_1| = \Big| &d\mathcal{T}^{-1}|_X(\partial_\alpha X, g^{-1}|_{a(u,X)}\mathcal{T}^{-1}(X)^*\partial_t\eta) \\
&- d\mathcal{T}^{-1}|_{X+\xi}(\partial_\alpha X, g^{-1}|_{a(u,X+\xi)}\mathcal{T}^{-1}(X+\xi)^*\partial_t\eta) \\
&- \big(\mathcal{T}^{-1}(X+\xi) - \mathcal{T}^{-1}(X)\big)\circ \\
&\quad \Big(dg^{-1}|_{a(u,X)}\big(\partial_t a(u,X), \mathcal{T}^{-1}|_X^*\partial_t\eta\big) + g^{-1}|_{a(u,X)}d\mathcal{T}^{-1}|_X^*(\partial_\alpha X, \partial_t\eta)\Big) \\
&+ \Big(\mathcal{T}^{-1}|_X g^{-1}|_{a(u,X)}\mathcal{T}^{-1}|_X^* - \mathcal{T}^{-1}|_{X+\xi}g^{-1}|_{a(u,X+\xi)}\mathcal{T}^{-1}|_{X+\xi}^*\Big)\partial_\alpha\partial_t\eta \\
&- d\mathcal{T}^{-1}|X+\xi\big(\partial_\alpha\xi, g^{-1}|_{a(u,X+\xi)}\mathcal{T}^{-1}(X+\xi)^*\partial_t\eta\big) \\
&- \mathcal{T}^{-1}(X+\xi)g^{-1}|_{a(u,X+\xi)}d\mathcal{T}^{-1}|X+\xi^*(\partial_\alpha\xi, \partial_t\eta) \\
&- \mathcal{T}^{-1}(X+\xi)dg^{-1}|_{a(u,X+\xi)}\big(\partial_2 a(u,X+\xi)\partial_\alpha\xi, \mathcal{T}^{-1}(X+\xi)^*\partial_t\eta)\Big| \\
\le{}& c|\xi|\cdot|\partial_t\eta|\big(\underline{1} + |\partial_\alpha X| + |\partial_\alpha\xi|\big) + c|X|\cdot|\partial_t\eta|\cdot|\partial_\alpha\xi|.
\end{aligned}
$$

Moreover, we obtain

$$
\begin{aligned}
|a_1| \le c|\xi|^2\big(|\partial_\alpha X| + |\partial_\alpha \xi|\big)&\Big(1 + |\partial_t X| + |Y|(1 + |\partial_t X|) + |\partial_t Y|\Big)\\
&+ c|\xi|\big(|\partial_\alpha X| + |\partial_\alpha \xi|\big)\Big(|\eta|(1 + |\partial_t X| + |\partial_t \xi|) + |\partial_t \xi|(1 + |Y|)\Big)\\
&+ c|\eta|\big(|\partial_\alpha X| + |\partial_\alpha \xi|\big)|\partial_t \xi| \cdot |X|
\end{aligned}
$$

and

$$
\begin{aligned}
\frac{|a_2|}{c} \le |\xi|^2&\Big(|\partial_\alpha Y|(1 + |\partial_t X|) + |\partial_\alpha \partial_t X|(1 + |Y|) + |\partial_\alpha \partial_t Y|\Big)\\
&+ |\xi|(1 + |Y|)\Big(|\partial_t \xi|(1 + |\partial_\alpha X|) + (|\xi| + |\partial_\alpha \xi| + |\partial_\alpha X|)(1 + |\partial_t X|)\Big)\\
&+ |\xi|\Big(|\partial_t \xi|(|\partial_\alpha \eta| + |\partial_\alpha Y|) + |\partial_\alpha \xi| \cdot |\partial_t Y| + |\eta|(1 + |\partial_\alpha X|)(1 + |\partial_t X|)\\
&\quad + |\xi| + |\partial_\alpha X|\big(1 + |\partial_t \eta| + |\partial_t Y|(1 + \underline{|\partial_t X|})\big) + |\partial_\alpha \eta|(1 + |\partial_t X|)\Big)\\
&+ |\partial_t \xi|\Big(|\eta|(1 + |\partial_\alpha X| + |\partial_\alpha \xi|) + |\partial_\alpha \xi|(1 + |Y|) + |\partial_\alpha \eta| \cdot |X|\Big)\\
&+ |\eta|\Big(|\partial_\alpha \xi|(1 + |\partial_t X|) + |\partial_\alpha \partial_t \xi|(|X| + |\xi|) + |\partial_\alpha \partial_t X| \cdot |\xi|\Big)\\
&+ |\partial_t X| \cdot |\partial_\alpha \xi|(1 + |\eta| + |Y|).
\end{aligned}
$$

The estimate for a_3 turns out to be

$$
\begin{aligned}
\frac{|a_3|}{c} \le \underline{|\xi|}(1 + |Y|)&\Big(|\xi|(1 + |\partial_t X|)(1 + |\partial_\alpha X|) + |\partial_t \xi|(1 + |\partial_\alpha X|)\\
&\quad + |\partial_\alpha \xi|(1 + |\partial_t X|) + |\eta|(1 + |\partial_t X|) + |\partial_t \eta|\Big)\\
&+ \underline{|\xi|} \cdot |\partial_\alpha \xi| \cdot |\partial_t Y| + |\eta| \cdot |\partial_\alpha \xi|(1 + |\partial_t X|)\\
&+ \underline{|\eta|} \cdot |\partial_t \xi|\big(1 + |\partial_\alpha X| + |\partial_\alpha \xi|(1 + |Y|)\big).
\end{aligned}
$$

Altogether these estimates lead to

$$
\begin{aligned}
\underline{|\partial_\alpha T_2|} \le c|\xi|^2&\Big(|\partial_t Y| + |\partial_\alpha Y|(1 + |\partial_t X|)\Big)\\
&+ c|\partial_\alpha \xi|(1 + |Y|)\Big(|\partial_t \xi| + |\partial_t X|\Big)\\
&+ c|\xi|^2(1 + |Y|)\Big(|\partial_t X| + |\partial_\alpha X| \cdot |\partial_t Y| + |\partial_\alpha \partial_t X| + |\partial_\alpha \partial_t Y|\Big)\\
&+ c|\xi|\Big(|\partial_\alpha X|\big(|\partial_t \xi| + |\partial_t \eta| + |\partial_t Y|(1 + |\partial_t X|)\big)\\
&\quad + |\partial_\alpha \eta|\big(1 + |\partial_t X| + |\partial_t \xi|\big) + |\partial_t \xi| \cdot |\partial_\alpha Y| + |\partial_\alpha \xi| \cdot |\partial_t Y|\\
&\quad + |\partial_t \eta|\big(|\partial_t X| + |\partial_\alpha \xi|(1 + |X|)\big) + |\partial_\alpha \xi| \cdot |\partial_t \xi|\\
&\quad + |\eta|\big(|\partial_\alpha X|(1 + |\partial_t X|) + |\partial_\alpha \partial_t \xi| + |\partial_\alpha \partial_t X|\big)\Big)
\end{aligned}
\tag{82}
$$

$$
\begin{aligned}
&+ c|\xi|(1+|Y|)\Big(|\xi| + |\partial_t\xi|(1+|\partial_\alpha X|) + |\partial_\alpha\xi| + |\partial_\alpha\partial_t\xi| \\
&\quad + |\partial_\alpha X|(1+|\partial_t X|) + |\partial_t\eta|\Big) \\
&+ c|\xi|\cdot|\eta|(1+|Y|)(1+|\partial_s X|) + |X|\Big(|\partial_t\xi|\cdot|\partial_\alpha\eta| + |\partial_\alpha\partial_t\xi|\cdot|\eta|\Big) \\
&+ c|\eta|\Big(|\partial_t\xi|(1+|\partial_\alpha X| + |\partial_\alpha\xi|) + |\partial_\alpha\xi|(1+|\partial_t X|)\Big).
\end{aligned}
$$

The term $\boldsymbol{\partial_\alpha T_4}$: We drop all the underlined terms in the estimate for $\partial_\alpha T_2$ (they come from the extra metric term present there) and replace all partial derivatives ∂_t by ∂_s to obtain

$$
\begin{aligned}
\underline{|\partial_\alpha T_4|} \le\ & c|\xi|^2|\partial_\alpha Y|(1+|\partial_s X|) + |\xi|^2(1+|Y|)\Big(|\partial_s X| \\
& \quad + |\partial_\alpha X|\cdot|\partial_s Y| + |\partial_\alpha\partial_s X| + |\partial_\alpha\partial_s Y|\Big) \\
& + c|\xi|\Big(|\partial_\alpha X|(|\partial_s\xi| + |\partial_s\eta| + |\partial_s Y|) + |\partial_\alpha\xi|\cdot|\partial_s Y| \\
& \quad + |\partial_\alpha\eta|\big(1+|\partial_s X| + |\partial_s\xi|\big) + |\partial_\alpha\xi|\cdot|\partial_s\xi| \\
& \quad + |\partial_s\xi|\cdot|\partial_\alpha Y| + |\partial_s\eta|\big(|\partial_s X| + |\partial_\alpha\xi|(1+|X|)\big) \\
& \quad + |\eta|\big(|\partial_\alpha X|(1+|\partial_s X|) + |\partial_\alpha\partial_s\xi| + |\partial_\alpha\partial_s X|\big)\Big) \\
& + c|\xi|(1+|Y|)\Big(|\xi| + |\partial_s\xi|(1+|\partial_\alpha X|) + |\partial_\alpha\xi| \\
& \quad + |\partial_\alpha\partial_s\xi| + |\partial_\alpha X|(1+|\partial_s X|)\Big) \\
& + c|\xi|\cdot|\eta|(1+|\partial_s X|) + |X|\big(|\partial_s\xi|\cdot|\partial_\alpha\eta| + |\partial_\alpha\partial_s\xi|\cdot|\eta|\big) \\
& + c|\partial_\alpha\xi|(1+|Y|)\big(|\partial_s\xi| + |\partial_s X|\big) \\
& + c|\eta|\big(|\partial_s\xi|(1+|\partial_\alpha X| + |\partial_\alpha\xi|) + |\partial_\alpha\xi|(1+|\partial_s X|)\big).
\end{aligned}
\tag{83}
$$

□

5.3. Quadratic Estimate II

The following quadratic estimate is an essential qualitative ingredient to carry out the induction step in the Newton method.
The theorem follows from the pointwise estimate 5.2.2 via integration as described in remark 5.0.8. Note that the conditions on $\|\xi\|_\infty$ and $\|X\|_\infty$ are necessary in order for the local constructions to be well defined. The other L^∞-conditions only serve to simplify the expressions. We underlined the worst terms with respect to rate of convergence in the existence part of the Newton method. Moreover, the theorem actually holds for any cylinder with appropriate smoothness and asymptotic convergence properties.

THEOREM 5.3.1. (**Quadratic estimate II**) *Let $p > 2$ and u be an element of the moduli space $\mathcal{M}^0(x^-, x^+)$, where x^-, $x^+ \in Crit\, \mathcal{I}_V$, define $w = g(u)\partial_t u$ and denote*

$$d\mathcal{F}^{triv}_{\epsilon,u}(Z)\circ\zeta - d\mathcal{F}^{triv}_{\epsilon,u}(0)\circ\zeta = \begin{pmatrix} F_1 \\ F_2 \end{pmatrix}.$$

Then there exists a constant $c_p > 0$ such that for $Z = (X, Y) \in C_0^\infty(\mathbb{R} \times S^1, u^*TM \oplus u^*T^*M)$ with $\|X\|_\infty < \iota'_M/2$ and $\|Y\|_\infty + \|\nabla_t X\|_\infty \le \sqrt{c_p}/2$

$$\begin{aligned}
\|F_1\|_p &\le c_p\|\xi\|_\infty\Big(\underline{\|\nabla_t X\|_p} + \|\nabla_s X\|_p\|X\|_\infty + \underline{\|Y\|_p} + \|\nabla_t Y\|_p\|X\|_\infty\Big)\\
&\quad + c_p\|X\|_\infty\Big(\|\xi\|_p + \underline{\|\nabla_t\xi\|_p} + \|\nabla_s\xi\|_p\|X\|_\infty + \underline{\|\eta\|_p}\Big)\\
\|F_2\|_p &\le c_p\|\xi\|_\infty\Big(\epsilon^{-2}\|\nabla_t X\|_p\|X\|_\infty + \|\nabla_s X\|_p + \|Y\|_p\Big)\\
&\quad + c_p\|X\|_\infty\Big(\epsilon^{-2}\big(\underline{\|\xi\|_p} + \|\nabla_t\xi\|_p\|X\|_\infty\big) + \|\nabla_s\xi\|_p + \|\eta\|_p\Big)
\end{aligned}$$

and

$$\begin{aligned}
&\|\nabla_\alpha F_1\|_p\\
&\le c_p\|\xi\|_\infty\Big(\|X\|_p + \|\nabla_t X\|_p + \|\nabla_s X\|_p + \|Y\|_p + \|\nabla_\alpha Y\|_p\\
&\qquad + \big(\|\nabla_t Y\|_p + \|\nabla_s Y\|_p\big)\|\nabla_\alpha X\|_\infty + \|\nabla_\alpha\nabla_t X\|_p + \|\nabla_\alpha\nabla_s X\|_p\|X\|_\infty\Big)\\
&\quad + c_p\|\xi\|_\infty\Big(\|\xi\|_p + \|\nabla_t\xi\|_p + \|\nabla_\alpha\xi\|_p + \|\nabla_\alpha\nabla_t\xi\|_p\Big)\\
&\quad + c_p\|X\|_\infty\Big(\|\nabla_t\xi\|_p + \|\nabla_s\xi\|_p + \|\eta\|_p + \|\nabla_\alpha\eta\|_p + \|\nabla_\alpha Y\|_p\|\nabla_t\xi\|_\infty\\
&\qquad + \|\nabla_\alpha\nabla_t\xi\|_p + \|\nabla_\alpha\nabla_s\xi\|_p\|X\|_\infty\Big)\\
&\quad + c_p\|\eta\|_\infty\Big(\|\nabla_t X\|_p + \|\nabla_\alpha X\|_p + \|\nabla_\alpha\nabla_t X\|_p\|X\|_\infty\Big)\\
&\quad + c_p\|\nabla_\alpha\xi\|_\infty\Big(\|Y\|_p + \|\nabla_t X\|_p\Big) + c_p\|\nabla_t\xi\|_\infty\Big(\|Y\|_p + \|\nabla_\alpha X\|_p\Big)
\end{aligned}$$

as well as

$$
\begin{aligned}
&\|\nabla_\alpha F_2\|_p \\
&\le c_p\epsilon^{-2}\|\xi\|_\infty\Big(\|\xi\|_p+\|\nabla_t\xi\|_p+\|X\|_p+\underline{\|\nabla_\alpha X\|_p}+\|\nabla_\alpha\nabla_t X\|_p\|X\|_\infty\Big)\\
&\quad+c_p\|\xi\|_\infty\Big(\|Y\|_p+\|\nabla_s X\|_p+\|\nabla_\alpha Y\|_p\big(1+\|\nabla_s X\|_\infty\big)\Big)\\
&\quad+c_p\epsilon^{-2}\|X\|_\infty\Big(\|\nabla_t\xi\|_p\big(1+\|\nabla_\alpha X\|_\infty\big)+\underline{\|\nabla_\alpha\xi\|_p}+\|\nabla_\alpha\nabla_t\xi\|_p\|\xi\|_\infty\Big)\\
&\quad+c_p\|X\|_\infty\Big(\|\eta\|_p+\|\nabla_s\xi\|_p+\|\nabla_\alpha Y\|_p\|\nabla_s\xi\|_\infty+\|\nabla_\alpha\eta\|_p\big(1+\|\nabla_s X\|_\infty\big)\Big)\\
&\quad+c_p\|\nabla_\alpha\nabla_s X\|_p\Big(\|X\|_\infty\|\eta\|_\infty+\|\xi\|_\infty\Big)+c_p\|\nabla_\alpha\nabla_s\xi\|_p\Big(\|X\|_\infty+\|\xi\|_\infty\Big)\\
&\quad+c_p\|\eta\|_\infty\|\nabla_s X\|_p+c_p\|\nabla_s\xi\|_\infty\big(\|Y\|_p+\|\nabla_\alpha X\|_p\big)\\
&\quad+c_p\|\nabla_\alpha\xi\|_\infty\big(\|Y\|_p+\|\nabla_s X\|_p\big)
\end{aligned}
$$

for $\epsilon\in(0,1]$ *and* $\zeta=(\xi,\eta)\in C_0^\infty(\mathbb{R}\times S^1,u^*TM\oplus u^*T^*M)$ *with* $\|\xi\|_\infty<\iota'_M/2$. *and* $\|\eta\|_\infty\le\sqrt{c_p}/2$.

Next we state and prove the pointwise estimate.

LEMMA 5.3.2. (**Pointwise estimate II**) *Let* $u\in\mathcal{P}_{x,y}$, x, y *smooth loops in* M, *and denote*

$$d\mathcal{F}^{triv}_{\epsilon,u}(Z)\circ\zeta-d\mathcal{F}^{triv}_{\epsilon,u}(0)\circ\zeta=\begin{pmatrix}T_1+T_2\\T_3+T_4\end{pmatrix}$$

for $Z=(X,Y)$, $\zeta=(\xi,\eta)\in C_0^\infty(\mathbb{R}\times S^1,u^*TM\oplus u^*T^*M)$. *Assume* $\|X\|_\infty\le\iota'_M/2$, *then in a local coordinate chart* $(U_i\subset V_i,\varphi_i)$ *as above the following pointwise estimates hold: There exists a constant* $c>0$ *such that*

$$
\begin{aligned}
&|\vec T_1+\vec T_2|\\
&\le c\,|\vec\xi|\Big(|\vec X|+|\partial_t\vec X|\big(1+|\vec X|\big)\big(1+|\vec Y|\big)+|\partial_s\vec X|\cdot|\vec X|+|\vec Y|+|\partial_t\vec Y|\cdot|\vec X|\Big)\\
&\quad+c\,|\vec X|\Big(|\partial_t\vec\xi|\big(1+|\vec Y|\big)+|\partial_s\vec\xi|\cdot|\vec X|+|\eta|\big(1+|\partial_t\vec X|\big)\Big)
\end{aligned}
$$

and

$$
\begin{aligned}
|\vec T_3+\vec T_4|\le c\,|\vec\xi|&\Big(|\partial_s\vec X|\big(1+|\vec Y|\big)+|\vec Y|\Big)\\
&+c\,|\vec X|\Big(\epsilon^{-2}|\vec\xi|\big(1+|\partial_t\vec X|\big)+\epsilon^{-2}|\partial_t\vec\xi|\cdot|\vec X|\\
&\qquad+|\partial_s\vec\xi|\big(1+|\vec Y|\big)+|\vec\eta|\big(1+|\partial_s\vec X|\big)\Big).
\end{aligned}
$$

for $\epsilon\in(0,1]$, *where* $|\cdot|=|\cdot|_{T_{u(s,t)}M}$ *and an arrow on top of an object indicates that it is represented in local coordinates and is evaluated at the point* (s,t). *The estimates for covariant derivatives of* $\vec T_1+\vec T_2$ *and* $\vec T_3+\vec T_4$ *are as in theorem 5.3.1.*

PROOF. (**of lemma 5.3.2 – pointwise estimate II**) Throughout the proof we use the notation introduced in remark 5.0.10. The main point is again that we are working in a local coordinate chart and therefore may use simply calculus in $\mathbb{R}^n$. Let us first specify the terms T_i. Set

$$w = g(u)\partial_t u$$

and denote

$$\begin{aligned} f_1(u,X,Y) &= \partial_s a(u,X) - g^{-1}|_{a(u,X)}\nabla_t^* b(u,X,Y) - \nabla V_t|_{a(u,X)} \\ f_2(u,X,Y) &= \nabla_s^* b(u,X,Y) + \epsilon^{-2}\big(g|_{a(u,X)}\partial_t a(u,X) - b(u,X,Y)\big) \end{aligned}$$

then

$$\begin{aligned} T_1 &= d\mathcal{T}^{-1}|_X \circ (\xi, f_1(u,X,Y)) - d\mathcal{T}^{-1}|_0 \circ (\xi, f_1(u,0,0)) \\ T_2 &= \mathcal{T}^{-1}(X) \circ \tfrac{d}{d\tau}\Big|_0 f_1(u, X+\tau\xi, Y+\tau\eta) - \tfrac{d}{d\tau}\Big|_0 f_1(u,\tau\xi,\tau\eta) \\ T_3 &= d\mathcal{T}|_X^* \circ (\xi, f_2(u,X,Y)) - d\mathcal{T}|_0^* \circ (\xi, f_2(u,0,0)) \\ T_4 &= \mathcal{T}(X)^* \circ \tfrac{d}{d\tau}\Big|_0 f_2(u, X+\tau\xi, Y+\tau\eta) - \tfrac{d}{d\tau}\Big|_0 f_2(u,\tau\xi,\tau\eta). \end{aligned}$$

The term T_3 : Note that we have already a linear factor ξ.

$$|T_3| = \|d\mathcal{T}|_X^*(\xi, f_2(u,X,Y)) - d\mathcal{T}|_0^*(\xi, f_2(u,0,0))\| = \|I + II + III\|.$$

Use (51), $\partial_2 a(u,0) = 1\!\!1$ and lemma 5.0.9 in the following two estimates. Moreover, the performance of the quadratic estimates will be crucially improved, if we consider certain terms in term II of T_3 and T_4 together (similarly for III and I). In order to do so, we subtract a term here and add it to II respectively III and I of T_4. Call the modified terms $\widetilde{II}$, $\widetilde{III}$ and $\widetilde{I}$.

$$\begin{aligned} \epsilon^2\|\widetilde{II}\| &= \epsilon^2\|II - d\mathcal{T}|_X^*(\xi, g|_{a(u,X)}\partial_2 a(u,X)\circ\partial_t X)\| \\ &= \big\|d\mathcal{T}|_X^*(\xi, g|_{a(u,X)}\partial_t a(u,X)) - d\mathcal{T}|_0^*(\xi, g(u)\partial_t u) \\ &\qquad - d\mathcal{T}|_X^*(\xi, g|_{a(u,X)}\partial_2 a(u,X)\circ\partial_t X)\big\| \\ &= \big\|d\mathcal{T}|_X^*(\xi, g|_{a(u,X)}\partial_1 a(u,X)\circ\partial_t u) - d\mathcal{T}|_0^*(\xi, g|_u\partial_t u)\big\| \\ &\le c(X)\,|X|\cdot|\xi| \end{aligned}$$

and

$$\begin{aligned} \epsilon^2\|\widetilde{III}\| &= \epsilon^2\|III + d\mathcal{T}|_X^*(\xi, \mathcal{T}^{-1}(X)^*Y)\| \\ &= \big\|-d\mathcal{T}|_X^*(\xi, \mathcal{T}^{-1}(X)^*\circ(w+Y)) + d\mathcal{T}|_0^*(\xi, w) \\ &\qquad + d\mathcal{T}|_X^*(\xi, \mathcal{T}^{-1}(X)^*Y)\big\| \\ &= \big\|-d\mathcal{T}|_X^*(\xi, \mathcal{T}^{-1}(X)^*w) + d\mathcal{T}|_0^*(\xi, \mathcal{T}^{-1}(0)^*\circ w)\big\| \\ &\le c(X)\,|X|\cdot|\xi|. \end{aligned}$$

Use formula (72) for $\nabla_s^* b(u,X,Y)$ to get

$$\begin{aligned}
\|\widetilde{I}\| &= \|I + \mathcal{T}(X)^* d\mathcal{T}^{-1}|_X^*(\xi, \partial_s Y)\| \\
&\leq \left\| d\mathcal{T}|_X^* \circ \nabla_s^* b(u,X,Y) - d\mathcal{T}|_0^* \circ \nabla_s^* w - d\mathcal{T}^{-1}|_0^* \partial_s Y \right\| \cdot |\xi| \\
&\leq \left\| d\mathcal{T}|_X^* \circ \mathcal{T}^{-1}(X)^* \circ \partial_s w - d\mathcal{T}|_0^* \circ \partial_s w \right\| \cdot |\xi| \\
&\quad + \Big\| -d\mathcal{T}|_X^* \circ \Gamma|_{a(u,X)} \big(\partial_1 a(u,X) \circ \partial_s u, \mathcal{T}^{-1}(X)^* w\big) \\
&\qquad + d\mathcal{T}|_0^* \circ \Gamma|_{a(u,0)} \big(\partial_1 a(u,0) \circ \partial_s u, \mathcal{T}^{-1}(0)^* w\big) \Big\| \cdot |\xi| \\
&\quad + \left\| d\mathcal{T}|_X^* \circ d\mathcal{T}^{-1}|_X^* \circ \big(\partial_s X, w + Y\big) \right\| \cdot |\xi| \\
&\quad + \left\| d\mathcal{T}|_X^* \circ \mathcal{T}^{-1}(X)^* \circ \partial_s Y + \mathcal{T}^{-1}(X)^* d\mathcal{T}^{-1}|_X^* \partial_s Y \right\| \cdot |\xi| \quad (= 0) \\
&\quad + \left\| -d\mathcal{T}|_X^* \circ \Gamma|_{a(u,X)} \big(\partial_2 a(u,X) \circ \partial_s X, \mathcal{T}^{-1}(X)^* w\big) \right\| \cdot |\xi| \\
&\quad + \left\| -d\mathcal{T}|_X^* \Gamma|_{a(u,X)} \big(\partial_1 a(u,X) \partial_s u + \partial_2 a(u,X) \partial_s X, \mathcal{T}^{-1}|_X^* Y\big) \right\| \cdot |\xi| \\
&\leq c(X)|\xi| \Big(|X| + |X| + |\partial_s X| \big(1 + |Y|\big) \\
&\quad + |X||\partial_s Y| + |\partial_s X| + |Y| \big(1 + |\partial_s X|\big) \Big) \\
&\leq c(X)|\xi| \Big(|X| + |\partial_s X| \big(1 + |Y|\big) + |Y| \Big).
\end{aligned}$$

Using the assumption $\|X\|_\infty \leq \iota_M'/2$ these estimates imply

$$\underline{|\widetilde{I} + \widetilde{II} + \widetilde{III}| \leq c\,|\xi| \Big(\epsilon^{-2}|X| + |\partial_s X|\,(1 + |Y|) + |Y| \Big).} \tag{84}$$

The term T_1 :

$$\begin{aligned}
|T_1| &= \left| d\mathcal{T}^{-1}|_X \circ (\xi, f_1(u,X,Y)) - d\mathcal{T}^{-1}|_0 \circ (\xi, f_1(u,0,0)) \right| \\
&= |I + II + III|.
\end{aligned}$$

Use lemma 5.0.9 to get

$$|III| = \left| -d\mathcal{T}^{-1}|_X (\xi, \nabla V_t|_{a(u,X)}) + d\mathcal{T}^{-1}|_0 (\xi, \nabla V_t|_{a(u,0)}) \right| \leq c(X)\,|X|\,|\xi|.$$

I respectively II work step by step the same as II respectively I in term T_3. We therefore only state the results. Note that we are moving a term from I here to I in T_2 and one from II here to II_2 in T_2.

$$\|I - d\mathcal{T}^{-1}|_X (\xi, \partial_2 a(u,X) \partial_s X)\| \leq c(X)\,|X|\,|\xi|$$

$$\begin{aligned}
\|II + d\mathcal{T}^{-1}|_X (\xi, g^{-1}|_{a(u,X)} \mathcal{T}^{-1}(X)^* \partial_t Y)\| \leq c(X)|\xi| \Big(&|X| + |Y| \\
&+ |X| \cdot |\partial_t X| \big(1 + |Y|\big) \Big)
\end{aligned}$$

and, using $\|X\|_\infty \leq \iota'_M/2$, we get

$$|T_1| \leq c\,|\xi|\Big(|X|\Big(1+|\partial_t X|\big(1+|Y|\big)\Big)+|Y|\Big). \tag{85}$$

The term T_2 :

$$\begin{aligned}|T_2| &= \Big|\mathcal{T}^{-1}(X)\tfrac{d}{d\tau}\big|_0\Big(\partial_s a(u,X+\tau\xi)\\ &\quad - g^{-1}|_{a(u,X+\tau\xi)}\nabla_t^* b(u,X+\tau\xi,Y+\tau\eta) - \nabla V_t|_{a(u,X+\tau\xi)}\Big)\\ &\quad - \tfrac{d}{d\tau}\big|_0\Big(\partial_s a(u,\tau\xi) - g^{-1}|_{a(u,\tau\xi)}\nabla_t^* b(u,\tau\xi,\tau\eta) - \nabla V_t|_{a(u,\tau\xi)}\Big)\Big|\\ &= |I+II+III|.\end{aligned}$$

Applying lemma 5.0.9 to III, I and (52) with t replaced by s to I we get

$$\begin{aligned}|III| &= \big|-\mathcal{T}^{-1}(X)\circ d\nabla V_t|_{a(u,X)}\circ\partial_2 a(u,X)\circ\xi + d\nabla V_t|_{a(u,0)}\circ\partial_2 a(u,0)\circ\xi\big|\\ &\leq c(X)\,|X|\cdot|\xi|.\end{aligned}$$

Carrying over the disturbing term from I in T_1 we get

$$\begin{aligned}&|I + d\mathcal{T}^{-1}|_X(\xi,\partial_2 a(u,X)\partial_s X)|\\ &= \big|\mathcal{T}^{-1}(X)\circ\partial_s\big(\partial_2 a(u,X)\circ\xi\big) - \partial_s\big(\partial_2 a(u,0)\circ\xi\big) + d\mathcal{T}^{-1}|_X(\xi,\partial_2 a(u,X)\partial_s X)\big|\\ &= \Big|\mathcal{T}^{-1}(X)\Big(\partial_1\partial_2 a(u,X)\circ(\xi,\partial_s u) + \partial_2\partial_2 a(u,X)\circ(\xi,\partial_s X) + \partial_2 a(u,X)\circ\partial_s\xi\Big)\\ &\qquad - \mathcal{T}^{-1}(0)\Big(\partial_1\partial_2 a(u,0)\circ(\xi,\partial_s u) + \partial_2 a(u,0)\circ\partial_s\xi\Big)\\ &\qquad + d\mathcal{T}^{-1}|_X(\xi,\partial_2 a(u,X)\partial_s X)\Big|\\ &\leq c(X)\Big(|X|\cdot|\xi| + |X|\cdot|\xi|\cdot|\partial_s X| + \|R(X,\partial_s\xi)X\|\cdot|\partial_s\xi|\Big)\end{aligned}$$

where the curvature in the last term of the sum arises as follows: Let

$$k(X) = \mathcal{T}^{-1}(X)\partial_2 a(u,X)\partial_s\xi - \partial_s\xi,$$

then $k(0)=0$ and

$$\begin{aligned}dk(0)X &= d\mathcal{T}^{-1}|_0(X,\partial_S\xi) + \partial_2\partial_2 a(u,0)(X,\partial_s\xi) = 0\\ d^2k(0)(X,X) &= R(X,\partial_s\xi)X.\end{aligned}$$

Here we used results from appendix A. Similarly for the term involving $\partial_s X$. Let $h(u,X) = -\mathcal{T}^{-1}(X)\circ g^{-1}|_{a(u,X)}$, then together with the term carried over from II in T_1 we get

$$
\begin{aligned}
&|II - d\mathcal{T}^{-1}|_X(\xi, g^{-1}|_{a(u,X)}\mathcal{T}^{-1}(X)^*\partial_t Y)| \\
&\quad = \big|-\mathcal{T}^{-1}(X)\circ \tfrac{d}{d\tau}\big|_0 \big(g^{-1}|_{a(u,X+\tau\xi)}\nabla_t^* b(u, X+\tau\xi, Y+\tau\eta)\big) \\
&\qquad + \tfrac{d}{d\tau}\big|_0 \big(g^{-1}|_{a(u,\tau\xi)}\nabla_t^* b(u,\tau\xi,\tau\eta)\big) - d\mathcal{T}^{-1}|_X(\xi, g^{-1}|_{a(u,X)}\mathcal{T}^{-1}(X)^*\partial_t Y)\big| \\
&\quad \le \big|-\mathcal{T}^{-1}(X) dg^{-1}|_{a(u,X)}\big(\partial_2 a(u,X)\xi, \nabla_t^* b(u,X,Y)\big) + dg^{-1}|_u\big(\xi, \nabla_t^* w\big)\big| \\
&\qquad + \big|h(u,X)\circ \tfrac{d}{d\tau}\big|_0 \nabla_t^* b(u, X+\tau\xi, Y+\tau\eta) \\
&\qquad - h(u,0)\circ \tfrac{d}{d\tau}\big|_0 \nabla_t^* b(u,\tau\xi,\tau\eta) - d\mathcal{T}^{-1}|_X(\xi, g^{-1}|_{a(u,X)}\mathcal{T}^{-1}(X)^*\partial_t Y)\big| \\
&\quad = |II_1 + \mathcal{T}^{-1}(X) dg^{-1}|_{a(u,X)}(\partial_2 a(u,X)\xi, \mathcal{T}^{-1}(X)^*\partial_t Y)| \\
&\qquad + |II_2 - \mathcal{T}^{-1}(X) dg^{-1}|_{a(u,X)}(\partial_2 a(u,X)\xi, \mathcal{T}^{-1}(X)^*\partial_t Y) \\
&\qquad - d\mathcal{T}^{-1}|_X(\xi, g^{-1}|_{a(u,X)}\mathcal{T}^{-1}(X)^*\partial_t Y)|
\end{aligned}
$$

where we moved a term from II_2 to II_1. Estimate exactly like I in term T_3 with s replaced by t:

$$
\begin{aligned}
&|II_1 + \mathcal{T}^{-1}(X) dg^{-1}|_{a(u,X)}(\partial_2 a(u,X)\xi, \mathcal{T}^{-1}(X)^*\partial_t Y)| \\
&\qquad \le c(X)\ |\xi| \Big(|X| \big(1 + |\partial_t X| \big(1+|Y|\big)\big) + |Y| \Big).
\end{aligned} \tag{86}
$$

Use formula (74) with s replaced by t for $\frac{d}{d\tau}\big|_0 \nabla_t^* b(u, X+\tau\xi, Y+\tau\eta)$ to get

(87)

$$
\begin{aligned}
|\widetilde{II_2}| &= |II_2 - \mathcal{T}^{-1}(X) dg^{-1}|_{a(u,X)}(\partial_2 a(u,X)\xi, \mathcal{T}^{-1}(X)^*\partial_t Y) \\
&\qquad - d\mathcal{T}^{-1}|_X(\xi, g^{-1}|_{a(u,X)}\mathcal{T}^{-1}(X)^*\partial_t Y)| \\
&\le \big|h(u,X)\circ d^2\mathcal{T}^{-1}|_X^*(\partial_t X, \xi, w+Y)\big| \\
&+ \big|h(u,X)\circ d\mathcal{T}^{-1}|_X^*(\partial_t\xi, w) - h(u,0)\circ d\mathcal{T}^{-1}|_0^*(\partial_t\xi, w)\big| \\
&+ \big|h(u,X)\circ d\mathcal{T}^{-1}|_X^*(\partial_t\xi, Y) - h(u,X)\circ d\mathcal{T}^{-1}|_X^*(\partial_t\xi, Y)\big| \\
&+ \big|h(u,X)\circ d\mathcal{T}^{-1}|_X^*(\xi, \partial_t w) - h(u,0)\circ d\mathcal{T}^{-1}|_0^*(\xi, \partial_t w)\big| \\
&+ \big|h(u,X) d\mathcal{T}^{-1}|_X^*(\xi, \partial_t Y) - \mathcal{T}^{-1}|_X dg^{-1}|_{a(u,X)}(\partial_2 a(u,X)\xi, \mathcal{T}^{-1}(X)^*\partial_t Y) \\
&\qquad - d\mathcal{T}^{-1}|_X(\xi, g^{-1}|_{a(u,X)}\mathcal{T}^{-1}(X)^*\partial_t Y)\big| \\
&+ \big|h(u,X)\circ d\mathcal{T}^{-1}|_X^*(\partial_t X, \eta) - h(u,X)\circ d\mathcal{T}^{-1}|_X^*(\partial_t X, \eta)\big| \\
&+ \big|h(u,X)\circ \mathcal{T}^{-1}(X)^*\circ\partial_t\eta - h(u,0)\circ\mathcal{T}^{-1}(0)^*\circ\partial_t\eta\big| \quad (= 0 \text{ by lemma A.1.11}) \\
&+ \big|-h(u,X)\circ\Gamma|_{a(u,X)}\big(\partial_2 a(u,X)\circ\partial_t\xi, \mathcal{T}^{-1}(X)^* w\big) \\
&\qquad + h(u,0)\circ\Gamma|_u\big(\partial_t\xi, w\big)\big| \\
&+ \big|-h(u,X)\Gamma|_{a(u,X)}\big(\partial_2 a(u,X)\partial_t\xi, \mathcal{T}^{-1}(X)^* Y\big) + h(u,X) d\mathcal{T}^{-1}|_X^*(\partial_t\xi, Y)\big| \\
&+ \big|-h(u,X)\circ\Gamma|_{a(u,X)}\big(\partial_1 a(u,X)\circ\partial_t u, d\mathcal{T}^{-1}|_X^*(\xi, w)\big) \\
&\qquad + h(u,0)\circ\Gamma|_u\big(\partial_t u, d\mathcal{T}^{-1}|_0^*(\xi, w)\big)\big|
\end{aligned}
$$

$$
\begin{aligned}
&+ \left|-h(u,X)\circ\Gamma|_{a(u,X)}\left(\partial_1 a(u,X)\circ\partial_t u, d\mathcal{T}^{-1}|_X^*(\xi,Y)\right)\right| \\
&+ \left|-h(u,X)\circ\Gamma|_{a(u,X)}\left(\partial_2 a(u,X)\circ\partial_t X, d\mathcal{T}^{-1}|_X^*(\xi,w+Y)\right)\right| \\
&+ \left|-h(u,X)\Gamma|_{a(u,X)}\left(\partial_1 a(u,X)\partial_t u, \mathcal{T}^{-1}(X)^*\eta\right) + h(u,0)\Gamma|_u\left(\partial_t u,\eta\right)\right| \\
&+ \left|-h(u,X)\Gamma|_{a(u,X)}\left(\partial_2 a(u,X)\circ\partial_t X, \mathcal{T}^{-1}|_X^*\eta\right) + h(u,X)d\mathcal{T}^{-1}|_X^*(\partial_t X,\eta)\right| \\
&+ \left|-h(u,X)\circ\Gamma|_{a(u,X)}\left(\partial_1\partial_2 a(u,X)\circ(\xi,\partial_t u), \mathcal{T}^{-1}(X)^* w\right)\right. \\
&\qquad \left. + h(u,0)\circ\Gamma|_u\left(\partial_1\partial_2 a(u,0)\circ(\xi,\partial_t u), w\right)\right| \\
&+ \left|-h(u,X)\circ\Gamma|_{a(u,X)}\left(\partial_1\partial_2 a(u,X)\circ(\xi,\partial_t u), \mathcal{T}^{-1}(X)^* Y\right)\right| \\
&+ \left|-h(u,X)\circ\Gamma|_{a(u,X)}\left(\partial_2\partial_2 a(u,X)\circ(\xi,\partial_t X), \mathcal{T}^{-1}(X)^* Y\right)\right| \\
&+ \left|-h(u,X)\circ d\Gamma|_{a(u,X)}\left(\partial_2 a(u,X)\circ\xi, \partial_1 a(u,X)\circ\partial_t u, \mathcal{T}^{-1}(X)^* w\right)\right. \\
&\qquad \left. + h(u,0)\circ d\Gamma|_u\left(\xi,\partial_t u, w\right)\right| \\
&+ \left|-h(u,X)\circ d\Gamma|_{a(u,X)}\left(\partial_2 a(u,X)\circ\xi, \partial_1 a(u,X)\circ\partial_t u, \mathcal{T}^{-1}(X)^* Y\right)\right| \\
&+ \left|-h(u,X) d\Gamma|_{a(u,X)}\left(\partial_2 a(u,X)\xi, \partial_2 a(u,X)\partial_t X, \mathcal{T}^{-1}(X)^*(w+Y)\right)\right|
\end{aligned}
$$

where we combined terms 3 and 9 in the sum so that term 3 results in a zero contribution and – calling term 9 $k(X)$, we get $k(0) = 0$ (use results from appendix A) – term 9 in one of order $c(X)\,|X| \cdot |\partial_t\xi| \cdot |Y|$. Moreover, in term 5 we profited from transferring the inconvenient term between II_1 and II_2. Similarly we combined terms 6 and 14 so that the former one contributes zero and the latter $c(X)\ |X| \cdot |\partial_t X| \cdot |\eta|$. Note also that the terms carried over to II_2 appear in term 5 of the sum. Denoting that term by $k_1(X)$ one readily computes using lemma A.1.11 that $k_1(0) = 0$ and so we get a contribution $c(X)\ |\xi| \cdot |X| \cdot |\partial_t Y|$. We get (keeping the order of terms)

$$
\begin{aligned}
|\widetilde{II}_2| \le c(X)\Big(& |\partial_t X| \cdot |\xi|(1+|Y|) + |X| \cdot |\partial_t\xi| + 0 + |X| \cdot |\xi| \\
& + |\xi| \cdot |X| \cdot |\partial_t Y| + 0 + 0 + |X| \cdot |\partial_t\xi| + |\partial_t\xi| \cdot |Y| \cdot |X| \\
& + |X| \cdot |\xi| + |\xi| \cdot |Y| + |\partial_t X| \cdot |\xi|(1+|Y|) + |X| \cdot |\eta| \\
& + |\partial_t X| \cdot |\eta| \cdot |X| + |X| \cdot |\xi| + |\xi| \cdot |Y| \\
& + |\xi| \cdot |\partial_t X|(1+|Y|) + |X| \cdot |\xi| + |\xi| \cdot |Y| + |\xi| \cdot |\partial_t X|(1+|Y|)\Big) \\
\le c(X)\ |\xi| & \Big(|X| + |\partial_t X|\ (1+|Y|) + |Y| + |\partial_t Y| \cdot |X|\Big) \\
& + c(X)\ |X|\Big(|\partial_t\xi|\ (1+|Y|) + |\eta|\ (1+|\partial_t X|)\Big).
\end{aligned}
$$

These estimates together with the assumption $\|X\|_\infty \le \iota'_M/2$ finally give

$$
\boxed{
\begin{aligned}
|T_2| \le c\,|\xi| & \Big(|X| + |\partial_t X|\ (1+|Y|) + |X| \cdot |\partial_s X| + |Y| + |\partial_t Y| \cdot |X|\Big) \\
& + c\,|X|\Big(|\partial_t\xi|\ (1+|Y|) + |\partial_s\xi| \cdot |X| + |\eta| \cdot (1+|\partial_t X|)\Big).
\end{aligned}
} \tag{88}
$$

The term T_4 :

$$
\begin{aligned}
|T_4| \le \Big|\mathcal{T}(X)^* \tfrac{d}{d\tau}\big|_0 \Big(&\nabla_s^* b(u, X+\tau\xi, Y+\tau\eta) + \epsilon^{-2} g|_{a(u,X+\tau\xi)} \partial_t a(u, X+\tau\xi) \\
&- \epsilon^{-2} b(u, X+\tau\xi, Y+\tau\eta)\Big) \\
&- \tfrac{d}{d\tau}\big|_0 \Big(\nabla_s^* b(u,\tau\xi,\tau\eta) + \epsilon^{-2} g|_{a(u,\tau\xi)} \partial_t a(u,\tau\xi) - \epsilon^{-2} b(u,\tau\xi,\tau\eta)\Big)\Big| \\
= |I + II + III|.&
\end{aligned}
$$

Recall that according to our modification of terms I, II and III of T_3 we have to add here what we subtracted there. Denote the modified versions of I, II respectively III here by $\widetilde{I}$, $\widetilde{II}$ respectively $\widetilde{III}$. We may use result (87) with t replaced by s to estimate I

$$
\begin{aligned}
&|I - \mathcal{T}(X)^* d\mathcal{T}^{-1}|_X^* (\xi, \partial_s Y)| \\
&= \Big|\mathcal{T}|_X^* \tfrac{d}{d\tau}\big|_0 \nabla_s^* b(u, X+\tau\xi, Y+\tau\eta) - \tfrac{d}{d\tau}\big|_0 \nabla_s^* b(u,\tau\xi,\tau\eta) \\
&\quad - \mathcal{T}(X)^* d\mathcal{T}^{-1}|_X^* (\xi, \partial_s Y)\Big| \\
&\le c(X)\, |\xi| \Big(|X| + |\partial_s X|\,(1+|Y|) + |Y|\Big) \\
&\quad + c(X)\, |X| \Big(|\partial_s \xi|\,(1+|Y|) + |\eta|\,(1+|\partial_s X|)\Big).
\end{aligned}
$$

Moreover, we get

$$
\begin{aligned}
\epsilon^2 |\widetilde{III}| &= \epsilon^2 |III - d\mathcal{T}|_X^* (\xi, \mathcal{T}^{-1}(X)^* Y)| \\
&= \Big| -\mathcal{T}(X)^* \tfrac{d}{d\tau}\big|_0 \Big(\mathcal{T}^{-1}(X+\tau\xi)^* \circ (w + Y + \tau\eta)\Big) \\
&\qquad + \tfrac{d}{d\tau}\big|_0 \Big(\mathcal{T}^{-1}(\tau\xi)^* \circ (w+\tau\eta)\Big) - d\mathcal{T}|_X^* (\xi, \mathcal{T}^{-1}(X)^* Y)\Big| \\
&\le \big| -\mathcal{T}(X)^* \circ d\mathcal{T}^{-1}|_X^* \circ (\xi, w) + d\mathcal{T}^{-1}|_0^* \circ (\xi, w)\big| \\
&\quad + \big| -\mathcal{T}(X)^* \circ d\mathcal{T}^{-1}|_X^* \circ (\xi, Y) - d\mathcal{T}|_X^* (\xi, \mathcal{T}^{-1}(X)^* Y)\big| \\
&\quad + \big| -\mathcal{T}(X)^* \circ \mathcal{T}^{-1}(X)^* \circ \eta + \eta\big| \\
&\le c(X)\, |X| \cdot |\xi|.
\end{aligned}
$$

To see that the second term in the sum vanishes apply the formula $dA^{-1} = -A^{-1} \circ dA \circ A^{-1}$ which holds for any smooth family of matrices $A : \mathbb{R}^n \to \mathcal{L}(\mathbb{R}^n)$. Moreover,

$$
\begin{aligned}
\epsilon^2 |\widetilde{II}| &= |II + d\mathcal{T}|_X^* (\xi, g|_{a(u,X)} \partial_2 a(u,X) \partial_t X)| \\
&= \big|\mathcal{T}(X)^* \tfrac{d}{d\tau}\big|_0 \big(g|_{a(u,X+\tau\xi)} \partial_t a(u, X+\tau\xi)\big) \\
&\quad - \tfrac{d}{d\tau}\big|_0 \big(g|_{a(u,\tau\xi)} \partial_t a(u,\tau\xi)\big) + d\mathcal{T}|_X^* (\xi, g|_{a(u,X)} \partial_2 a(u,X) \partial_t X)\big|
\end{aligned}
$$

$$\begin{aligned}
&\leq \big|\mathcal{T}(X)^*\circ dg|_{a(u,X)}\big(\partial_2 a(u,X)\circ\xi, \partial_1 a(u,X)\circ\partial_t u\big) - dg|_u\big(\xi, \partial_t u\big) \\
&\quad +\mathcal{T}(X)^*\circ g|_{a(u,X)}\circ\partial_1\partial_2 a(u,X)(\partial_t u, \xi) - g|_u\circ\partial_1\partial_2 a(u,0)(\partial_t u, \xi)\big| \\
&\quad + \big|\mathcal{T}(X)^*\circ dg|_{a(u,X)}(\partial_2 a(u,X)\,\xi, \partial_2 a(u,X)\,\partial_t X) \\
&\quad + \mathcal{T}(X)^*\circ g|_{a(u,X)}\circ(\partial_2\partial_2 a(u,X)\,(\partial_t X, \xi)) \\
&\quad +d\mathcal{T}|_X^*(\xi, g|_{a(u,X)}\partial_2 a(u,X)\partial_t X)\big| \\
&\leq c(X)\Big(|X|\cdot|\xi| + |X|^2\cdot|\partial_t\xi| + |X|\cdot|\partial_t X|\cdot|\xi|\Big).
\end{aligned}$$

Clearly in these estimate we used lemma 5.0.9 in combination with the results of analyzing the relation between exponential maps and parallel transport in appendix A section A.1. Using the assumption $\|X\|_\infty \leq \iota'_M/2$ and the estimates for $\widetilde{I}, \widetilde{II}$ and $\widetilde{III}$ gives

$$\boxed{\begin{aligned}
|\widetilde{I} + \widetilde{II} + \widetilde{III}| \leq c\,|X|\Big(&\epsilon^{-2}|\xi|\,(1 + |\partial_t X|) + \epsilon^{-2}|\partial_t\xi|\cdot|X| \\
&+ |\partial_s\xi|\,(1 + |Y|) + |\eta|\,(1 + |\partial_s X|)\Big) \\
&+ c\,|\xi|\Big(+|\partial_s X|\,(1 + |Y|) + |Y|\Big).
\end{aligned}} \tag{89}$$

We skip the proof of the estimates for $\nabla_\alpha T_1 + \nabla_\alpha T_2$ and $\nabla_\alpha T_3 + \nabla_\alpha T_4$ because the way to proceed is similar as in quadratic estimate I above. □

CHAPTER 6

Transversality

In section 6.1 we discuss transversality theory from a fairly general point of view and isolate the input which varies case by case from the mechanisms that are intrinsic to the general theory. As it will turn out, conditions (F) and (A) below, as well as the degree of differentiability of the section involved, is the external input to be checked in a particular situation.

Verifying these conditions for the classical action $\mathcal{I}_V$ and the symplectic action $\mathcal{A}_V$ in section 6.2 leads to the transversality theorem for loops 6.2.1 which states, roughly speaking, that both functionals are Morse functions for generic potential V.

6.1. Thom-Smale transversality

Let $\mathcal{A}, \mathcal{B}$ be smooth Banach manifolds modeled on separable Banach spaces and such that they admit a countable atlas. $\mathcal{E} \to \mathcal{A} \times \mathcal{B}$ denotes a smooth Banach space bundle and $\mathcal{F}$ a section of $\mathcal{E}$ of class C^k, $k \geq 1$. Assume $\mathcal{F}$ has the following two properties

(F) $\quad d\mathcal{F}_b(a) : T_a\mathcal{A} \to \mathcal{E}_{(a,b)}$ *is Fredholm for all* $a \in \mathcal{F}_b^{-1}(0)$ *and* $d\mathcal{F}_a(b)$ *is bounded.*

(S) $\quad d\mathcal{F}(a,b) : T_a\mathcal{A} \times T_b\mathcal{B} \to \mathcal{E}_{(a,b;0)}$ *is surjective for all* $(a,b) \in \mathcal{F}^{-1}(0)$.

Note that $\mathcal{F}_b : \mathcal{A} \to \mathcal{E}$ is defined by $a \mapsto \mathcal{F}(a,b)$ and similarly for $\mathcal{F}_a$. Moreover, the above differentials are to be understood as the ordinary differential followed by projection onto the fibres of $\mathcal{E}$. Here we use the natural splitting of the tangent space of $\mathcal{E}$ at the zero section

$$T_{(a,b;0)}\mathcal{E} \simeq T_{(a,b)}(\mathcal{A} \times \mathcal{B}) \oplus \mathcal{E}_{(a,b)}.$$

Condition (F), lemma 6.1.5 $i)$ below and

$$d\mathcal{F}(a,b) = d\mathcal{F}_b(a) \oplus d\mathcal{F}_a(b)$$

imply that $d\mathcal{F}(a,b)$ has a closed range. In order to verify (S) it is therefore sufficient to prove that *Ran* $d\mathcal{F}(a,b)$ is dense. It is a consequence [**Br83**, corollaire I.8] of the Hahn-Banach theorem that this is equivalent to the triviality of its annihilator $(Ran\, d\mathcal{F}(a,b))^\perp$

(A) $\quad \{v^* \in \mathcal{E}^*_{(a,b)} \mid v^*(v) = 0 \;\forall v \in Ran\, d\mathcal{F}(a,b)\} = \{0\} \;\; \forall (a,b) \in \mathcal{F}^{-1}(0).$

Plug the Fredholm (F) and surjectivity (S) properties of $\mathcal{F}$ into lemma 6.1.5 to obtain that $d\mathcal{F}(a,b)$ admits a right inverse for all $(a,b) \in \mathcal{F}^{-1}(0)$; in other words 0 is a regular value of $\mathcal{F}$. The implicit function theorem C.3.4 then implies that the *universal moduli space*

$$X = \mathcal{F}^{-1}(0)$$

is a C^k-Banach manifold. It is locally at (a,b) modeled on the separable Banach space $ker\, d\mathcal{F}(a,b)$ and admits a countable atlas. Define the projection onto the second factor

$$\pi : X \to \mathcal{B}$$

and observe that for any $(a,b) \in \mathcal{F}^{-1}(0)$

$$d\pi(a,b) : T_{(a,b)}X = ker\ (d\mathcal{F}_b(a) \oplus d\mathcal{F}_a(b)) \to T_b\mathcal{B} \quad , \quad (A,B) \mapsto B$$

is a Fredholm operator by lemma 6.1.5 *ii)* with

$$Ind\, d\pi(a,b) = Ind\, d\mathcal{F}_b(a).$$

Hence π is a Fredholm map of class C^k between Banach manifolds and so we can apply the Sard/Smale theorem 6.1.4 to its representations with respect to the countably many coordinate charts of X and $\mathcal{B}$ and obtain that the set of regular values of π is of the second category in $\mathcal{B}$ for

$$k \geq \max\{1, Ind\, d\mathcal{F}_b(a) + 1\}.$$

Definition 6.1.1. A subset of a complete metric space is said to be of the *second category in the sense of Baire* if it contains a countable intersection of open and dense sets.

Recall that by Baire's category theorem, every set of the second category is dense; for references cf. [**RS1**, Notes to section III.5]. The crucial step is now to observe that

Lemma 6.1.2.

$$\{\textit{regular values of}\ \pi\} = \{b \in \mathcal{B} \mid d\mathcal{F}_b(a)\ \textit{onto}\ \forall a \in \mathcal{F}_b^{-1}(0)\}.$$

Proof. Because π is a Fredholm map, the kernel of its linearization automatically admits a topological complement (lemma C.2.2) and so we have

$$\begin{aligned} & b \text{ regular value of } \pi \\ & \Leftrightarrow d\pi(a,b) \text{ onto } \forall (a,b) \in \pi^{-1}(b) \\ & \Leftrightarrow d\pi(a,b) \text{ onto } \forall a \in \mathcal{F}_b^{-1}(0) \\ & \Leftrightarrow \forall a \in \mathcal{F}_b^{-1}(0)\ \forall B \in T_b\mathcal{B}\ \exists A \in T_a\mathcal{A} \text{ with } (A,B) \in T_{(a,b)}X \\ & \qquad \text{such that } d\pi(a,b)\begin{pmatrix} A \\ B \end{pmatrix} = B. \end{aligned} \tag{90}$$

On the other hand

$$b \in \{b' \in \mathcal{B} \mid d\mathcal{F}_{b'}(a) \text{ onto } \forall a \in \mathcal{F}_{b'}^{-1}(0)\} \tag{91}$$
$$\Leftrightarrow \forall a \in \mathcal{F}_b^{-1}(0)\ \forall e \in \mathcal{E}_{(a,b)}\ \exists \hat{A} \in T_a\mathcal{A}\ :\ d\mathcal{F}_b(a)\ \hat{A} = e.$$

(91) ⇒ (90): Let $a \in \mathcal{F}_b^{-1}(0)$, pick $B \in T_b\mathcal{B}$ and define $e = -d\mathcal{F}_a(b)\ B$. By (91) $\exists A \in T_a\mathcal{A}$ such that

$$d\mathcal{F}_b(a)\ A = e = -d\mathcal{F}_a(b)\ B.$$

(90) ⇒ (91): Here surjectivity of $d\mathcal{F}(a,b)$ for $(a,b) \in \mathcal{F}^{-1}(0)$ enters: $\forall e \in \mathcal{E}_{(a,b)}\ \exists (A', B') \in T_a\mathcal{A} \times T_b\mathcal{B}$:

$$d\mathcal{F}(a,b) \begin{pmatrix} A' \\ B' \end{pmatrix} = d\mathcal{F}_b(a)\ A' + d\mathcal{F}_a(b)B' = e.$$

Let now $a \in \mathcal{F}_b^{-1}(0)$ and pick $e \in \mathcal{E}_{(a,b)}$, then by surjectivity of $d\mathcal{F}(a,b)$ there exists (A', B') such that

$$d\mathcal{F}_b(a)\ A' + d\mathcal{F}_a(b)B' = e.$$

For this B' there exists by (90) an element A such that $(A, B') \in T_{(a,b)}X$; i.e.

$$d\mathcal{F}_b(a)\ A + d\mathcal{F}_a(b)\ B' = 0.$$

$\hat{A}$ as required in (91) is now obtained by setting $\hat{A} = A' - A$:

$$d\mathcal{F}_b(A)\ (A' - A) = e - d\mathcal{F}_a(b)B' + d\mathcal{F}_a(b)B' = e.$$

□

For the sake of completeness we recall the theorem of Sard/Smale [**SS73**] as well as two technical lemmata used above.

Definition 6.1.3. A metric space is called *separable* if it admits a dense sequence.

Theorem 6.1.4. (**Sard/Smale**) *Let X and Y be separable Banach spaces and $U \subset X$ be an open set. Suppose that $f : U \to Y$ is a C^∞-smooth Fredholm map. Then the set*

$$Y_{reg} = \{y \in Y \mid Ran\, df(x) = Y \text{ for all } x \in U \text{ with } f(x) = y\}$$

of regular values of f is of the second category in the sense of Baire. More precisely, this continues to hold if f is of class C^k with $k \geq \max\{1, Ind f + 1\}$.

For a proof using a local Kuranishi model we refer to [**Sa96**] theorem B.13. The next lemma may be found there, too (lemma B.5).

Lemma 6.1.5. *Let X, Y, Z be Banach spaces. Assume $D : X \to Y$ is a Fredholm operator and $L : Z \to Y$ is bounded and linear, then*

i) the bounded linear operator $D \oplus L : X \oplus Z \to Y$ has a closed range with a finite-dimensional complement.

ii) if $D \oplus L$ is onto, then ker $(D \oplus L)$ admits a topological complement (i.e. $D \oplus L$ admits a right inverse). Moreover, the projection on the second factor

$$\Pi : ker\,(D \oplus L) \to Z$$

is a Fredholm operator with

$$ker\,\Pi \simeq ker\,D \quad , \quad coker\,\Pi \simeq coker\,D$$

and hence $Ind\,\Pi = Ind\,D$.

LEMMA 6.1.6. *Let $(X, \|\cdot\|)$ be a normed linear space, then*

i) if $V \subset X$ is a closed and $W \subset X$ a finite dimensional subspace, then $V + W$ is a closed subspace of X.

ii) if $V \subset X$ is a closed subspace with finite dimensional complement and $W \subset X$ is any subspace, then $V + W$ is a closed subspace of X.

PROOF. (**of lemma 6.1.6**) *i)* Assume $W \neq \{0\}$ and $V \cap W = \{0\}$. Otherwise consider $V + W' = V + W$, where $W' = W \setminus (V \cap W)$ and $dim\,W' < \infty$, $W' \cap V = \{0\}$. Let $\zeta_\nu = v_\nu + w_\nu \in V + W$ be such that $\zeta_\nu \to \zeta \in X$ for $\nu \to \infty$. We have to show $\zeta \in V + W$. If $\zeta \in V$ we are done, so assume the contrary. It follows $\lim_{\nu\to\infty} w_\nu \neq 0$ since otherwise $V \ni v_\nu = \zeta_\nu - w_\nu \to \zeta$ for $\nu \to \infty$. Because V is closed we get to the contradiction $\zeta \in V$.
Compactness of the unit ball B_W in W leads to the existence of a subsequence such that

$$\frac{w_{\nu_k}}{|w_{\nu_k}|} \stackrel{k\to\infty}{\longrightarrow} w \in B_W.$$

We conclude that $\{|w_{\nu_k}| : k \in \mathbb{N}\}$ is bounded, because otherwise a further subsequence (same notation) converges to $+\infty$ and so

$$V \ni \frac{v_{\nu_k}}{|w_{\nu_k}|} = \frac{\zeta_{\nu_k}}{|w_{\nu_k}|} - \frac{w_{\nu_k}}{|w_{\nu_k}|} \stackrel{k\to\infty}{\longrightarrow} 0 - w.$$

Hence $w \in V \cap W = \{0\}$ – a contradiction to $w \in B_W$. Taking a further subsequence if necessary we may assume $|w_{\nu_k}| \to \omega_0 \in \mathbb{R} \setminus \{0\}$ for $k \to \infty$. Closedness of V implies

$$V \ni \frac{v_{\nu_k}}{|w_{\nu_k}|} = \frac{\zeta_{\nu_k}}{|w_{\nu_k}|} - \frac{w_{\nu_k}}{|w_{\nu_k}|} \stackrel{k\to\infty}{\longrightarrow} \frac{\zeta}{\omega_0} - w \in V$$

and so

$$\zeta - \omega_0 w = \omega_0 \left(\frac{\zeta}{\omega_0} - w \right) \in V.$$

Therefore

$$\zeta = (\zeta - \omega_0 w) + \omega_0 w \in V + W.$$

ii) Writing

$$V + W = V + \left(\frac{X}{V} \cap W \right)$$

reduces the problem to part $i)$ because X/V is finite dimensional and so is $(X/V) \cap W$. □

PROOF. (**of lemma 6.1.5**) $i)$ As $Ran D$ is closed with finite dimensional complement lemma 6.1.6 $ii)$ applies and yields closedness of

$$Ran\,(D \oplus L) = Ran\,D + Ran\,L.$$

Because $Ran\,D \subset Ran\,(D \oplus L)$ we obtain

$$\frac{Y}{Ran\,(D \oplus L)} \subset \frac{Y}{Ran\,D}$$

where the latter is finite dimensional.
$ii)$ As D is Fredholm $dim\,ker\,D < \infty$ and so we can choose a topological complement X_1 by lemma C.2.2. $Ran\,D$ closed with finite dimensional complement $coker\,D$ implies that we can write $Y = Ran\,D \oplus coker\,D$. Because $D \oplus L : X \oplus Z \to Y$ is surjective it follows $coker\,D \subset Ran\,L$ and so we can choose a basis $\{Le_1, \ldots, le_N\}$ of $coker\,D$, where $\{e_1, \ldots, e_N\}$ is a set of linearly independent elements of Z.
Our claim is that $W := Ran\,T$ is the required topological complement of $ker\,(D \oplus L)$, where the linear map T is defined as follows (actually T is a right inverse of D)

$$T : Ran\,D \oplus coker\,D \to ker\,D \oplus X_1 \oplus Z$$
$$(y_1, y_2) \mapsto (0, x_1, \sum_{\nu=1}^{N} \lambda_\nu e_\nu).$$

Here x_1 is determined uniquely by $y_1 = Dx_1$ and $y_2 = \sum_{\nu=1}^{N} \lambda_\nu L e_\nu$.
W closed: $W = X_1 + Span(e_1, \ldots, e_N)$ and so lemma 6.1.6 $i)$ applies.
$W \cap ker\,(D \oplus L) = \{0\}$: Let $(0, x_1, \sum_{\nu=1}^{N} \lambda_\nu e_\nu) \in W \cap ker\,(D \oplus L)$, then

$$(0,0) = (D \oplus L)\,(0, x_1, \sum_{\nu=1}^{N} \lambda_\nu e_\nu) = (Dx_1, \sum_{\nu=1}^{N} \lambda_\nu L e_\nu)$$

and so $x_1 = 0$ (because D is injective on X_1) and $\lambda_\nu = 0$ for all $\lambda \in \{1, \ldots, N\}$ because $\{Le_1, \ldots, Le_N\}$ is a basis.
$W + \ker\,(D \oplus L) = X \oplus Z$: $\subset$ is trivial. To see $\supset$ pick $(x, z) \in X \oplus Z$, write $x = (x_0, x_1) \in ker\,D \oplus X_1$ and $Lz = y' + \sum_{\nu=1}^{N} \lambda_\nu L e_\nu \in Ran\,D \oplus coker\,D$. Note that $y' = Dx'$ for a unique $x' \in X_1$. Now

$$(x, z) = (x_0, x_1, z) = (0, x_1 + x', \sum_{\nu=1}^{N} \lambda_\nu e_\nu)$$
$$+ (x_0, -x', z - \sum_{\nu=1}^{N} \lambda_\nu e_\nu) \in W + ker\,(D \oplus L).$$

It remains to check

$$(D \oplus L)\,(x_0, -x', z - \sum_{\nu=1}^{N} \lambda_\nu e_\nu) = 0 - Dx' + Lz - \sum_{\nu=1}^{N} \lambda_\nu L e_\nu = 0$$

because $Dx' = y'$.

The above proves that W is a topological complement of $ker\,(D \oplus L)$. Now assume $(x, z) \in ker\,(D \oplus L)$, then

$$\begin{aligned}(x,z) \in ker\,\Pi &\Leftrightarrow 0 = \Pi(x,z) = z \\ &\Leftrightarrow Dx = 0 \\ &\Leftrightarrow x \in ker\,D.\end{aligned}$$

This proves $ker\,\Pi = ker\,D \oplus 0$. Define

$$L^{-1}(Ran\,D) := \{z \in Z \mid Lz = Dx \quad \text{for some} \quad x \in X\},$$

then $Ran\,\Pi = L^{-1}(Ran\,D)$ and this set is closed: $Ran\,D$ is closed and so is its preimage under the continuous map L. Finally we obtain

$$coker\,D \simeq \frac{Y}{Ran\,D} \simeq \frac{Ran\,L}{Ran\,D \cap Ran\,L} \simeq \frac{Z}{L^{-1}(Ran\,D)} \simeq \frac{Z}{Ran\,\Pi} \simeq coker\,\Pi$$

where in the second step we used $D \oplus L$ onto and so

$$\frac{Y}{Ran\,D} = \frac{Ran\,D + Ran\,L}{Ran\,D} \simeq \frac{Ran\,L}{Ran\,D \cap Ran\,L}.$$

To see the third step observe that

$$L^{-1}(Ran\,D) \simeq \frac{Z}{ker\,L}$$

so that

$$\frac{Z}{L^{-1}(Ran\,D)} \simeq \frac{Z}{ker\,L \oplus (Ran\,D \cap Ran\,L)} \simeq \frac{Ran\,L}{Ran\,D \cap Ran\,L}.$$

□

6.2. Transversality for loops

In this section we prove that the classical action $\mathcal{I}_V$ and, equivalently, the symplectic action $\mathcal{A}_V$ are Morse functions for generic potential $V \in C^k(M \times S^1, \mathbb{R})$. Throughout let us fix an integer $k \geq 2$.

Recall that $\Lambda M = W^{1,2}(S^1, M)$ and

$$\Lambda^a M = \{x \in \Lambda M \mid \mathcal{I}_V(x) < a\}.$$

Note that the definition of $\Lambda^a M$ depends on the choice of V. As the critical points of $\mathcal{I}_V$ and $\mathcal{A}_V$ are canonically identified we denote them by $Crit$. Let $Crit_a := Crit \cap \Lambda^a M$.

On the other hand, in the language of the previous section, we have a smooth Banach space bundle $\mathcal{E}$ over the smooth Banach manifold $\mathcal{A} \times \mathcal{B} = W^{2,2}(S^1, M) \times C^k(M \times S^1, \mathbb{R})$ together with a C^{k-1}-section

$$\mathcal{F}(x, V) = -\nabla_t \dot{x} - \nabla V_t(x).$$

The fibre of $\mathcal{E}$ at (x, V) is given by $\mathcal{E}_{(x,V)} = L^2(x^*TM)$. Smoothness of $\mathcal{E}$ and $\mathcal{A} \times \mathcal{B}$ comes from the smoothness of (M, g), whereas $V \in C^k$ implies $\nabla V \in C^{k-1}$ and therefore $\mathcal{F}$ is only $k-1$ times continuously differentiable. $\mathcal{A}$ is modeled on the separable Banach space $W^{2,2}(S^1, \mathbb{R}^n)$ and admits a countable atlas whereas $\mathcal{B} = (C^k(M \times S^1, \mathbb{R}), \|\cdot\|_{C^k})$ is a separable Banach space itself. The differential of $\mathcal{F}$ at a zero (x, V) followed by projection onto the fibre of $\mathcal{E}$ is given by the bounded linear operator

$$\begin{aligned} d\mathcal{F}(x, V) : W^{2,2}(S^1, M) \times C^k(M \times S^1, \mathbb{R}) &\to L^2(x^*TM) \\ (\xi, \dot{V}) &\mapsto d\mathcal{F}_V(x)\xi + d\mathcal{F}_x(V)\dot{V} \end{aligned}$$

where $\mathcal{F}_V(x) = \mathcal{F}(x, V)$ and $\mathcal{F}_x(V) = \mathcal{F}(x, V)$.

The relation between both formulations of the problem, namely analyzing critical points of a functional or zeroes of a section, is as follows

$$\begin{aligned} Crit &= \{x \in \Lambda M \mid d\mathcal{I}_V(x)\xi = 0, \forall \xi \in W^{1,2}(x^*TM)\} \\ &= \{x \in W^{2,2}(S^1, M) \mid \mathcal{F}(x, V) = 0\} \\ &= \{x \in C^{k+1}(S^1, M) \mid -\nabla_t \dot{x} - \nabla V_t(x) = 0\}, \end{aligned}$$

where $V \in C^k(M \times S^1, \mathbb{R})$ with $k \geq 2$. We are ready to state the main theorem of this chapter.

THEOREM 6.2.1. (**Transversality for loops**) *Fix an integer $k \geq 2$.*

i) The functionals $\mathcal{I}_V$ and $\mathcal{A}_V$ are Morse functions for any $V \in \mathcal{V}^k_{reg}$, where

$$\mathcal{V}^k_{reg} := \{V \in C^k(M \times S^1, \mathbb{R}) \mid d\mathcal{F}_V(x) \text{ onto } \forall x \in Crit\}$$

is a subset of the Banach space $(C^k(M \times S^1, \mathbb{R}), \|\cdot\|_{C^k})$ of the second category in the sense of Baire.

ii) Fix $a \in \mathbb{R}$, then the restricted functionals $\mathcal{I}_V : \Lambda^a M \to \mathbb{R}$ and $\mathcal{A}_V : \Lambda^a T^*M \to \mathbb{R}$ are Morse functions for any $V \in \mathcal{V}_{reg}^{k,a}$, where

$$\mathcal{V}_{reg}^{k,a} := \{V \in C^k(M \times S^1, \mathbb{R}) \mid d\mathcal{F}_V(x) \text{ onto } \forall x \in Crit_a\}$$

is an open and dense subset of $(C^k(M \times S^1, \mathbb{R}), \|\cdot\|_{C^k})$.

iii) Fix $a \in \mathbb{R}$, then the restricted functionals $\mathcal{I}_V : \Lambda^a M \to \mathbb{R}$ and $\mathcal{A}_V : \Lambda^a T^*M \to \mathbb{R}$ are Morse functions for any $V \in \mathcal{V}_{reg}^{a}$, where

$$\mathcal{V}_{reg}^{a} := \{V \in C^\infty(M \times S^1, \mathbb{R}) \mid d\mathcal{F}_V(x) \text{ onto } \forall x \in Crit_a\}$$

is an open and dense subset of the complete metric space $(C^\infty(M \times S^1, \mathbb{R}), d)$ with

$$d(V_1, V_2) := \sum_{k=0}^{\infty} \frac{1}{2^k} \frac{\|V_1 - V_2\|_{C^k}}{1 + \|V_1 - V_2\|_{C^k}}.$$

iv) The functionals $\mathcal{I}_V$ and $\mathcal{A}_V$ are Morse functions for any $V \in \mathcal{V}_{reg}$, where

$$\mathcal{V}_{reg} := \{V \in C^\infty(M \times S^1, \mathbb{R}) \mid d\mathcal{F}_V(x) \text{ onto } \forall x \in Crit\}$$

is a subset of $(C^\infty(M \times S^1, \mathbb{R}), d)$ of the second category in the sense of Baire.

It suffices to prove the theorem for $\mathcal{I}_V$, because its Hessian at a critical point is nondegenerate iff the corresponding Hessian of $\mathcal{A}_V$ is.

Transversality in the C^k-category. Recall that

$$d\mathcal{F}_V(x)\xi = -\nabla_t \nabla_t \xi - R(\xi, \dot{x})\dot{x} - \nabla_\xi \nabla V_t(x)$$

is the perturbed Jacobi operator analyzed in appendix B.2. We called *dim ker* $d\mathcal{F}_V(x)$ nullity, which turned out to be finite (Morse index theorem B.2.8). $d\mathcal{F}_V(x)$ proved to be selfadjoint and so

$$ker\, d\mathcal{F}_V(x) \simeq coker\, d\mathcal{F}_V(x),$$

which means that the operator is Fredholm and its Fredholm index is zero. Note that this is true at all zeroes (x, V) of $\mathcal{F}$. The operator $d\mathcal{F}_x(V)\dot{V} = -\nabla \dot{V}(x)$ is bounded

$$\|d\mathcal{F}_x(V)\dot{V}\|_{L^2}^2 = \int_0^1 \langle \nabla \dot{V}_t(x), \nabla \dot{V}_t(x) \rangle \, dt \le |S^1| \cdot \|\dot{V}\|_{C^1}^2 \le |S^1| \cdot \|\dot{V}\|_{C^k}^2$$

and so we have verified the Fredholm assumption (F) in section 6.1. Assume for now the surjectivity assumption (S) was true, too.

Let us summarize the results of the general theory of section 6.1. The universal moduli space $X = \mathcal{F}^{-1}(0)$ is a Banach manifold of class C^{k-1} modeled on a separable Banach space and it admits a countable atlas. Moreover, the projection onto the second factor $\pi : X \to C^k(M \times S^1, \mathbb{R})$ is a C^{k-1}-Fredholm map with $Ind\, d\pi(x, V) = Ind\, d\mathcal{F}_V(x) = 0$. The condition

$k-1 \geq \max\{1, Ind\, \pi\}$ in the Sard/Smale theorem reflects precisely our assumption $k \geq 2$. Now the set of regular values of π is of the second category in the sense of Baire and by lemma 6.1.2 coincides with

$$\{V \in C^k(M \times S^1, \mathbb{R}) \mid d\mathcal{F}_V(x) \text{ onto}, \forall x \in \mathcal{F}_V^{-1}(0)\} = \mathcal{V}^k_{reg}.$$

So to prove part $i)$ of the theorem it remains to verify surjectivity (S) which is a consequence of condition (A): $\forall (x, V) \in \mathcal{F}^{-1}(0)$

$$\{\eta \in L^2(x^*TM) \mid \langle \eta, \xi \rangle_{L^2} = 0\; \forall \xi \in Ran\; d\mathcal{F}(x, V)\} = \{0\}.$$

PROOF. (Condition (A) holds) We have to show that

$$\langle \eta, d\mathcal{F}_V(x)\xi \rangle = 0 \quad \forall \xi \in W^{2,2}(x^*TM)$$

and

$$\langle \eta, d\mathcal{F}_x(V)\dot{V} \rangle = 0 \quad \forall \dot{V} \in C^k(M \times S^1, \mathbb{R})$$

together imply $\eta = 0$. The first condition says that $\eta \in ker\; d\mathcal{F}_V(x)$. So it satisfies a second order ODE with coefficients of class C^{k-2} and therefore $\eta \in C^k(x^*TM)$. Now assume by contradiction that there is $t_0 \in S^1$ such that $\eta(t_0) \neq 0$. In five steps we are going to construct $\dot{V}_t \in C^\infty$ such that

$$\langle \eta, \nabla \dot{V}_t(x) \rangle_{L^2} \neq 0.$$

As our construction will be local, we may choose geodesic normal coordinates $\vec{\xi} = (\xi_1, \dots, \xi^n)$ around $x_0 = x(t_0)$. Let ι denote the injectivity radius of (M, g). The piece of the loop $x(t)$ which lies inside the coordinate patch is represented by $\vec{\xi}(t) \in \mathbb{R}^n$ via

$$x(t) = exp_{x_0}\, \vec{\xi}(t).$$

Clearly $\vec{\xi}(t_0) = 0$. An arrow indicates quantities represented in our local coordinates. $\langle \cdot, \cdot \rangle$ denotes the euclidean inner product on $\mathbb{R}^n$ and $\mid \cdot \mid$ the associated norm.

STEP 1 Because $x(t)$ is continuous, we may choose a constant $\delta_1 > 0$ sufficiently small such that

$$\mid \vec{\xi}(t) \mid \leq \iota/2 \quad , \quad \forall t \in [t_0 - \delta_1, t_0 + \delta_1].$$

STEP 2 Because η is continuous and $\eta(t_0) \neq 0$, we may choose a constant $\delta_2 > 0$ sufficiently small such that

$$\langle \vec{\eta}(t), \vec{\eta}(t_0) \rangle > 0 \quad , \quad \forall t \in [t_0 - \delta_2, t_0 + \delta_2].$$

STEP 3 Set $\delta = \min\{\delta_1, \delta_2\}$ and choose a cut-off function $\gamma \in C^\infty(\mathbb{R}, [0,1])$ such that

$$\gamma(t) = \begin{cases} 1 & , t \in [\frac{t_0 - \delta}{2}, \frac{t_0 + \delta}{2}] \\ 0 & , t \notin [t_0 - \delta, t_0 + \delta]. \end{cases}$$

Step 4 Choose a cut-off function $\beta \in C^\infty(\mathbb{R}, [0,1])$ such that

$$\beta(\mid \vec{\xi} \mid^2) = \begin{cases} 1 & , \mid \vec{\xi} \mid^2 \leq \iota^2/2 \\ 0 & , \mid \vec{\xi} \mid^2 \geq \iota^2. \end{cases}$$

Step 5 We are ready to define $\dot{V}_t$

$$\dot{V}_t(exp_{x_0}\vec{\xi}) = \begin{cases} \gamma(t)\, \beta(\mid \vec{\xi} \mid^2)\, \langle \vec{\eta}(t_0), \vec{\xi} \rangle & , \mid \vec{\xi} \mid^2 < \iota^2 \\ 0 & , \text{else}. \end{cases}$$

Putting everything together we get

$$\begin{aligned} \langle \eta, \nabla \dot{V}_t(x) \rangle_{L^2} &= \int_0^1 g\left(\eta(t), \nabla \dot{V}_t(x(t))\right) dt \\ &= \int_0^1 d\dot{V}_t(x(t)) \circ \eta(t)\, dt \\ &= \int_{\{t: \mid \vec{\xi}(t) \mid < \iota\}} \frac{\partial \dot{V}_t}{\partial \xi^j} \mid_{exp_{x_0}\vec{\xi}(t)} \; \eta^j(t)\, dt \\ &= \int_{t_0-\delta}^{t_0+\delta} \Big(2\gamma(t)\, \beta'(\mid \vec{\xi} \mid^2)\, \langle \vec{\xi}(t), \vec{\eta}(t) \rangle\, \langle \vec{\eta}(t_0), \vec{\xi}(t) \rangle \\ &\qquad + \gamma(t)\, \beta(\mid \vec{\xi} \mid^2)\, \langle \vec{\eta}(t_0), \vec{\eta}(t) \rangle \Big) dt \\ &= \int_{t_0-\delta}^{t_0+\delta} \gamma(t)\, \langle \vec{\eta}(t_0), \vec{\eta}(t)\, dt > 0. \end{aligned}$$

The third equality follows from the definition of $\dot{V}_t$ (Step 5), and the fourth one from Step 3 ($supp\, \gamma$) as well as a straight forward calculation. In the fifth equality we used that for $t \in [t_0 - \delta, t_0 + \delta]$ Step 1 implies $\mid \vec{\xi}(t) \mid^2 \leq \iota^2/2$ and therefore, by Step 4, $\beta' \equiv 0$ and $\beta \equiv 1$. Step 2 gives the final strict inequality. □

To prove part *ii)* of the theorem we observe that

$$\mathcal{V}^k_{reg} \subset \mathcal{V}^{k,a}_{reg}, \tag{92}$$

where the former space is of the second category in $(C^k(M \times S^1, \mathbb{R}), \|\cdot\|_{C^k})$ and hence, according to definition 6.1.1, the bigger space $\mathcal{V}^{k,a}_{reg}$ is, too.

Openess of $\mathcal{V}^{k,a}_{reg}$ in $(C^k(M \times S^1, \mathbb{R}), \|\cdot\|_{C^k})$ is harder and relies on compactness of $Crit_a$ for regular V: Pick $V \in \mathcal{V}^{k,a}_{reg}$. This means that $d\mathcal{F}_V(x)$ is onto for all $x \in Crit_a$. Equivalenty, 0 is a regular value of $\mathcal{F}_V$ on $W^{2,2}(S^1, M) \cap \Lambda^a M$, so that by the implicit function theorem $Crit_a$ is a manifold of dimension $Ind\, d\mathcal{F}_V(x) = 0$. In view of the a-priori action bound a we derived in remark 1.2.1 compactness of $Crit_a$. Because $X^a = \mathcal{F}^{-1}(0) \cap \Lambda^a M$ is open in X, restriction of the projection π yields a Fredholm map of class C^{k-1}

$$\pi_a : X^a \to C^k(M \times S^1, \mathbb{R}).$$

Now

$$\pi_a^{-1}(V) = \{(x_j, V) \mid \{x_1, \ldots, x_N\} = Crit_a\}$$

consists of finitely many elements $p_j = (x_j, V)$ around each of which we may find an open neighbourhood U_j in X^a such that

$$(x', V') \in U_j \quad \Rightarrow \quad d\mathcal{F}_{V'}(x') \text{ onto.}$$

Continuity of π_a then allows to find an open neighbourhood W of V in $C^k(M \times S^1, \mathbb{R})$ such that $\pi_a^{-1}(W) \subset \bigcup_{j=1}^N U_j$ and therefore $W \subset \mathcal{V}_{reg}^{k;a}$.
It remains to prove openess of the surjectivity property of $d\mathcal{F}_V(x_j)$ in X^a, which allowed us to choose the U_j's above. Note that for (x', V') near (x_j, V) the operators $d\mathcal{F}_V(x_j)$ and $d\mathcal{F}_{V'}(x')$ differ (after representing them with respect to a common trivialization) by a bounded operator. In view of the subsequent lemma we get

$$dim\ ker\ d\mathcal{F}_{V'}(x') \leq dim\ ker\ d\mathcal{F}_V(x_j) = 0,$$

where the last equation follows from the surjectivity and selfadjointness of $d\mathcal{F}_V(x_j)$. For the same reason we obtain

$$0 = dim\ ker\ d\mathcal{F}_{V'}(x') = dim\ coker\ d\mathcal{F}_{V'}(x').$$

LEMMA 6.2.2. *Let X, Y be Banach spaces and $D : X \to Y$ be a Fredholm operator. Then there exists an $\epsilon > 0$ such that for any linear map $L : X \to Y$ with $\|L\| < \epsilon$*

$$dim\ ker\ (D + L) \leq dim\ ker\ D.$$

PROOF. Following [**BB85**, I.5.C ex.9], let X_1 be a topological complement of $ker\ D$. We prove

$$ker\ (D + L) \cap X_1 = \{0\},$$

which implies our claim in view of $ker D \oplus X_1 = X$. Because $\tilde{D} : X_1 \to Ran D$ is a bounded bijection between Banach spaces, it has a bounded inverse $\tilde{D}^{-1}$ by the open mapping theorem. Let $x \in X_1 \cap ker\ (D + L)$, then $x = -\tilde{D}Lx$ and so

$$\|x\|_X = \|\tilde{D}^{-1}Lx\|_X \leq \|\tilde{D}^{-1}\| \cdot \|L\| \cdot \|x\|_X \leq c\epsilon\|x\|_X.$$

For $0 < \epsilon < \|\tilde{D}^{-1}\|$ it follows $x = 0$. □

Transversality in the C^∞-category. To prove part *iii*) of the transversality theorem let us start with density of $\mathcal{V}_{reg}^a$ in (C^∞, d): Given any $V \in C^\infty(M \times S^1, \mathbb{R})$ we have to construct a sequence $V'_k \in C^\infty(M \times S^1, \mathbb{R})$ such that

$$\forall \epsilon > 0 \quad \exists k_0 \in \mathbb{N} \quad \forall k > k_0 \quad : \quad d(V, V'_k) < \epsilon.$$

The idea will be to approximate V by regular V_k's in the C^k-topology and then approximate V_k by smooth regular elements V'_k in the C^k-topology.

Finally we make use of the observation that in order to control the metric d we essentially have to control only finitely many C^k-norms in its series, because the strong weights $1/2^k$ take care of all the other ones.

STEP 1 Because $\mathcal{V}_{reg}^{k,a}$ is dense in $(C^k(M \times S^1, \mathbb{R}), \|\cdot\|_{C^k})$ for any integer $k \geq 2$, we can find $V_k \in C^k(M \times S^1, \mathbb{R})$ with $\|V - V_k\|_{C^k} < 1/(2k)$. For $k = 0, 1$ let us define $V_0 = V_1 = V$.

STEP 2 Because $\mathcal{V}_{reg}^{k,a}$ is open in $(C^k(M \times S^1, \mathbb{R}), \|\cdot\|_{C^k})$ for any integer $k \geq 2$, we can choose $0 < \epsilon_k < 1/(2k)$ sufficiently small such that $B_{\epsilon_k}(V_k)$, the open ϵ_k-ball around V_k, is contained in $\mathcal{V}_{reg}^{k,a}$.

STEP 3 Because $M \times S^1$ is compact, $C^\infty(M \times S^1, \mathbb{R})$ is dense in $(C^k(M \times S^1, \mathbb{R}), \|\cdot\|_{C^k})$ for $k \in \mathbb{N}_0$, cf. [**Hi76**, theorem 2.6]. Hence we can find $V_k' \in C^\infty \cap B_{\epsilon_k}(V_k)$ for $k \geq 2$. For $k = 0, 1$ we define $V_0' = V_1' = V$.

Now pick $\epsilon > 0$ and choose $\nu_0 \in \mathbb{N}$ sufficiently large such that $f(\nu_0) = \sum_{\nu=\nu_0+1}^{\infty} 2^{-\nu} < \epsilon/2$. Choose $k_0 > \max\{\nu_0, 4/\epsilon\}$ and observe that by Steps 1,2 and 3 for $k \geq 2$

$$\|V - V_k'\|_{C^k} \leq \|V - V_k\|_{C^k} + \|V_k - V_k'\|_{C^k} \leq \frac{1}{2k} + \epsilon_k \leq \frac{1}{k}.$$

Note that this implies $\|V - V_k'\|_{C^\nu} \leq \|V - V_k\|_{C^k} \leq \frac{1}{k}$ for any $2 \leq \nu \leq k$. We get for any $k > k_0$

$$\begin{aligned} d(V, V_k') &= \sum_{\nu=0}^{\nu_0} \frac{1}{2^\nu} \frac{\overbrace{\|V - V_k'\|_{C^k}}^{\leq 1/k}}{1 + \|V - V_k'\|_{C^k}} + \sum_{\nu=\nu_0+1}^{\infty} \frac{1}{2^\nu} \overbrace{\frac{\|V - V_k'\|_{C^k}}{1 + \|V - V_k'\|_{C^k}}}^{\leq 1} \\ &\leq \frac{2}{k} + \frac{\epsilon}{2} < \epsilon. \end{aligned}$$

Openess of $\mathcal{V}_{reg}^a$ is easier: Pick $V \in \mathcal{V}_{reg}^a$ and set $k = 2$. Exploiting openess of $\mathcal{V}_{reg}^{2,a}$ in $(C^2(M \times S^1, \mathbb{R}), \|\cdot\|_{C^2})$ we are able to choose a constant $\epsilon_0 > 0$ such that for any V'' of class C^2 with $\|V - V''\|_{C^2} < \epsilon_0$ it follows $V'' \in \mathcal{V}_{reg}^{2,a}$. Now define

$$\epsilon = \frac{1}{4} \frac{\epsilon_0}{1 + \epsilon_0}.$$

Let V' of class C^∞ be such that $d(V, V') < \epsilon$. Therefore each term in the series on the left hand side has to be strictly smaller then ϵ, in particular the second one

$$\frac{1}{2^2} \frac{\|V - V'\|_{C^2}}{1 + \|V - V'\|_{C^2}} < \epsilon = \frac{1}{4} \frac{\epsilon_0}{1 + \epsilon_0}.$$

But this is equivalent to

$$\|V - V'\|_{C^2} < \epsilon_0$$

and therefore $V' \in \mathcal{V}_{reg}^{2,a}$. Finally $\mathcal{V}_{reg}^a = \mathcal{V}_{reg}^{2,a} \cap C^\infty(M \times S^1, \mathbb{R})$ implies $V' \in \mathcal{V}_{reg}^a$.

We prove part *iv*) of the transversality theorem: As $\mathcal{V}^a_{reg}$ is open and dense in (C^∞, d) it is of the second category in the sense of Baire. The identity

$$\mathcal{V}_{reg} = \bigcap_{a=0}^{\infty} \mathcal{V}^a_{reg}$$

implies the claim, because by Baire's category theorem any countable intersection of sets of the second category is again of the second category.

APPENDIX A

Linearization and trivialization of operators

In section A.1 foundations are laid to carry out the geometric analysis in the main body of this text; particularly to get optimal quadratic estimates in chapter 5. The crucial nonstandard results (at the level of textbooks) are formulae for derivatives of the exponential map and the parallel transport of vector and covector fields.

Section A.2 contains a detailled calculation in local coordinates of the linearization $\mathcal{D}^{\epsilon}_{\tilde{w}}$ of the ϵ-dependent nonlinear elliptic equations at a solution $\tilde{w}$. In other words we linearize the Banach bundle section

$$\mathcal{F}_\epsilon : \mathcal{E}^p \to \mathcal{P}^{1,p}_{x^-,x^+}$$

at a zero $\tilde{w} \in \mathcal{F}_\epsilon^{-1}(0)$. Somewhat implicitely contained is a formula for $\mathcal{D}^0_{\tilde{u}}$ the linearization of the parabolic operator $\mathcal{F}_0$ at a zero $\tilde{u}$.

Using results from section A.1 about the parallel transport we derive in section A.3 a formula for the derivative at $(0,0)$ of $\mathcal{F}^{triv}_{\epsilon,w}$ – the representative of $\mathcal{F}_\epsilon$ in a local trivialization of the Banach bundle at *any* w. This will be done intrinsically. It turns out that the formula for $d\mathcal{F}^{triv}_{\epsilon,w}(0,0)$ coincides with the one for $\mathcal{D}^{\epsilon}_{\tilde{w}}$.

Finally section A.4 provides simpler formulae for the linear operators $\mathcal{D}^0_{\tilde{u}}$ and $\mathcal{D}^{\epsilon}_{\tilde{w}}$ in orthogonal respectively unitary frames.

A.1. Some Riemannian geometry

We recall fundamental concepts in Riemannian geometry, such as Levi-Civita connection and curvature tensor (A.1.1), exponential map (A.1.2) and parallel transport (A.1.3). Throughout let (M^n, g) denote a smooth Riemannian manifold of dimension n and $\Gamma(TM)$ the set of smooth sections of TM – in other words smooth vector fields on M.

A.1.1. Levi-Civita connection and curvature tensor. The *Levi-Civita connection* ∇ on the tangent bundle $TM \to M$ with respect to the metric g is the uniquely determined connection on TM which satisfies the conditions of being *torsion free*

$$T(X,Y) = \nabla_X Y - \nabla_Y X - [X,Y] = 0 \quad , \forall X,Y \in \Gamma(TM), \tag{93}$$

where

$$[X,Y] = XY - YX,$$

and being *compatible with the metric* (i.e. a *Riemannian connection*)

$$Xg(Y,Z) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z) \quad , \forall X, Y, Z \in \Gamma(TM). \tag{94}$$

Let ∇^* denote the associated connection on T^*M defined by

$$(\nabla^*_X \eta)(Y) = X(\eta(Y)) - \eta(\nabla_X Y) \quad , \forall X, Y \in \Gamma(TM), \eta \in \Gamma(T^*M).$$

If $u \in C^\infty(\mathbb{R} \times S^1, M)$ there are induced connections on the vector bundles u^*TM and u^*T^*M. By abuse of notation we use the same symbols ∇ and ∇^* for these induced connections and occasionally we will even use simply the symbol ∇ to denote any one of them.

Let $\{x^1, \dots, x^n\}$ be a system of local coordinates on an open subset $U \subset M$ and $\vec{x} = (x^1, \dots, x^n)$ represent a point $x \in M$. Then we have the natural bases $\{\partial_k = \partial/\partial x^k\}_{k=1}^n$ of the tangent space and $\{dx^j\}_{j=1}^n$ of the cotangent space to U at x. Expressing X, Y and η with respect to these bases as $X = X^i \partial_i$, $Y = Y^j \partial_j$ and $\eta = \eta_j dx^j$ we get

$$\nabla_x Y = X^i \left(\frac{\partial Y^k}{\partial x^i} + \Gamma^k_{ij}(\vec{x})\, Y^j \right) \partial_k \quad , \quad \text{where } \Gamma^k_{ij} = dx^k \left(\nabla_{\partial_i} \partial_j \right)$$

$$\nabla^*_X \eta = X^i \left(\frac{\partial \eta_j}{\partial x^i} - \Gamma^k_{ij}(\vec{x})\, \eta_k \right) dx^j \quad , \quad \text{where } -\Gamma^k_{ij} = \left(\nabla^*_{\partial_i} dx^k \right) \partial_j .$$

The quantities Γ^k_{ij} are called *Christoffel symbols* and they satisfy

$$\Gamma^k_{ij}(x) = \tfrac{1}{2} g^{kl}(x) \left(\frac{\partial g_{il}(x)}{\partial x^j} + \frac{\partial g_{jl}(x)}{\partial x^i} - \frac{\partial g_{ij}(x)}{\partial x^l} \right). \tag{95}$$

The *curvature tensor* R is a skew-symmetric bilinear form $R_q : T_qM \times T_qM \to End(T_qM)$ defined by

$$R(X,Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z \tag{96}$$

for $X, Y, Z \in \Gamma(TM)$. The riemannian condition on ∇ (compatibility with the metric) implies $R(X,Y) \in \underline{so}(TM)$ for all X and Y (cf. [**Sa96**] se. 2.1). Note that there is no agreement in the mathematical literature concerning the sign in the definition of the curvature tensor. In local coordinates

$$(R(X,Y)Z)^m = R^m_{kij}(x) X^i Y^j Z^k \tag{97}$$

where

$$R^l_{kij}(x) = \frac{\partial \Gamma^l_{jk}(x)}{\partial x^i} - \frac{\partial \Gamma^l_{ik}(x)}{\partial x^j} + \Gamma^l_{i\nu}(x)\Gamma^\nu_{jk}(x) - \Gamma^l_{j\nu}(x)\Gamma^\nu_{ik}(x). \tag{98}$$

Moreover there is the following consequence of the first Bianchi identity

$$g(R(X,Y)Z, V) = g(R(Z,V)X, Y)\ . \tag{99}$$

A.1.2. Exponential map. On a Riemannian manifold (M^n, g) one has furthermore the concept of an *exponential map*, which is defined as follows: Let $u \in M$ and $\xi \in T_uM$, then define

$$exp_u\xi = \gamma(1)$$

where $\gamma(t)$ is the (unique) solution of the 2^{nd} order initial value problem

$$\gamma(0) = u \quad , \quad \partial_t\gamma(0) = \xi \quad , \quad \nabla_t\partial_t\gamma \equiv 0,$$

i.e. $\gamma(t)$ is the *geodesic* emanating from u in direction ξ. Clearly $exp_u 0 = u$. Before studying the derivatives of the exponential map let us consider two initial value problems which are closely related: Fix $\tau \geq 0$, then

$$(*_\tau) \quad \begin{aligned} \gamma_{\tau,u}(s=0) &= u \\ \partial_s\gamma_{\tau,u}(0) &= \tau\xi \\ \nabla_s\partial_s\gamma_{\tau,u} &\equiv 0 \end{aligned} \qquad (*_1) \quad \begin{aligned} \gamma_{1,u}(t=0) &= u \\ \partial_t\gamma_{1,u}(0) &= \xi \\ \nabla_t\partial_t\gamma_{1,u} &\equiv 0 \end{aligned}$$

where $s, t \in [0, \infty)$ for compact M.

LEMMA A.1.1. *Let $\gamma_{\tau,u}$ be the solution to $(*_\tau)$ and $\gamma_{1,u}$ be the one to $(*_1)$, then $\gamma_{\tau,u}(s) = \gamma_{1,u}(s\tau)$ for all $s \geq 0$.*

PROOF. Uniqueness of the solution of $(*_\tau)$ implies that it suffices to show that $f(s) := \gamma_{1,u}(s\tau)$ solves $(*_\tau)$:

$$\begin{aligned} f(0) &= \gamma_{1,u}(0) = u \\ \partial_s f(0) &= \partial_s\gamma_{1,u}(s\tau)|_{s=0} = \partial_t\gamma_{1,u}(0) \cdot \tau = \tau\xi \\ \nabla_s\partial_s f(s) &= \tau^2\nabla_t\partial_t\gamma_{1,u}(t) = \tau^2 \cdot 0 = 0. \end{aligned}$$

□

This lemma is the crucial ingredient in calculating the derivatives of the exponential map, which we will also denote by

$$a(u, \xi) = exp_u\xi.$$

Note that *ii)* in the following proposition, more precisely $dexp_u(0) = id_{T_uM}$, implies that there exists a constant $\iota_u > 0$ such that $exp_u : T_uM \supset B_{\iota_u}(0) \to exp_u(B_{\iota_u}(0)) \subset M$ is a diffeomorphism (inverse function theorem C.3.2 in appendix C). ι_u is called *injectivity radius at u*. If M is compact, then there exists $\iota > 0$ such that exp_u is injective on the ball $B_\iota(0) \subset T_uM$ for all $u \in M$. ι is called *injectivity radius of M*.

PROPOSITION A.1.2. *In local coordinates $\vec{u} = (u^1, \dots, u^n)$ on M the following statements are true for any $i, j, k, l \in \{1, \dots, n\}$*

$$\begin{aligned}
&i) && a(\vec{u},0)^k = u^k\\
&ii) && \partial_1 a(\vec{u},0)^k_i = \delta^k_i = \partial_2 a(\vec{u},0)^k_i\\
&iii) && \partial_1\partial_1 a(\vec{u},0)^k_{ij} = \partial_1\partial_2 a(\vec{u},0)^k_{ij} = \partial_2\partial_1 a(\vec{u},0)^k_{ij} = 0\\
&iv) && \partial_2\partial_2 a(\vec{u},0)^k_{ij} = -\Gamma^k_{ij}(\vec{u})\\
&v) && \partial_2\partial_2\partial_1 a(\vec{u},0)^k_{lij} = -\frac{\partial\Gamma^k_{ij}(\vec{u})}{\partial u^l}\\
&vi) && \partial_2\partial_1\partial_1 a(\vec{u},0)^k_{lij} = 0.
\end{aligned}$$

PROOF. *i)* has been shown above. To prove *ii)* use the definition of $exp_u\tau\xi = \gamma_\tau(1)$ as well as lemma A.1.1 to compute

$$\begin{aligned}
\partial_2 a(\vec{u},0)^k_i\,\xi^i &= \tfrac{d}{d\tau}\big|_{\tau=0}\, a(\vec{u},\tau\vec{\xi})^k = \tfrac{d}{d\tau}\big|_{\tau=0}\,(exp_{\vec{u}}\tau\vec{\xi})^k = \tfrac{d}{d\tau}\big|_{\tau=0}\,\gamma_{\tau,\vec{u}}(1)^k\\
&= \tfrac{d}{d\tau}\big|_{\tau=0}\,\gamma_{1,\vec{u}}(\tau)^k = \partial_t\gamma_{1,\vec{u}}(0)^k = \xi^k
\end{aligned}$$

and

$$\partial_1 a(\vec{u},0)^k_i\, x^i = \tfrac{d}{d\tau}\big|_{\tau=0}\, a(\vec{u}+\tau\vec{x},0)^k = \tfrac{d}{d\tau}\big|_{\tau=0}\,(u^k+\tau x^k) = x^k.$$

Similarly we obtain the first statement in *iii)*

$$\partial_1\partial_1 a(\vec{u},0)^k_{ij}\, x^i x^j = \tfrac{d^2}{d\tau^2}\Big|_{\tau=0}\, a(\vec{u}+\tau\vec{x},0)^k = \tfrac{d^2}{d\tau^2}\Big|_{\tau=0}\,(u^k+\tau x^k) = 0.$$

Commutativity of partial derivatives in $\mathbb{R}^n$ implies that to prove the other two statements it suffices to show

$$\begin{aligned}
\partial_1\partial_2 a(\vec{u},0)^k_{ij}\,\xi^i x^j &= \tfrac{d}{d\mu}\Big|_{\mu=0}\,\tfrac{d}{d\tau}\big|_{\tau=0}\Big(exp_{\vec{u}+\mu\vec{x}}\tau\vec{\xi}\Big)^k\\
&= \tfrac{d}{d\mu}\Big|_{\mu=0}\,\tfrac{d}{d\tau}\big|_{\tau=0}\,\gamma_{\tau,\vec{u}+\mu\vec{x}}(1)^k = \tfrac{d}{d\mu}\Big|_{\mu=0}\,\tfrac{d}{d\tau}\big|_{\tau=0}\,\gamma_{1,\vec{u}+\mu\vec{x}}(\tau)^k\\
&= \tfrac{d}{d\mu}\Big|_{\mu=0}\,\partial_t\gamma_{1,\vec{u}+\mu\vec{x}}(0)^k = \tfrac{d}{d\mu}\Big|_{\mu=0}\,\xi^k = 0.
\end{aligned}$$

We prove *iv)*

$$\begin{aligned}
\partial_2\partial_2 a(\vec{u},0)^k_{ij}\,\xi^i\xi^j &= \tfrac{d^2}{d\tau^2}\Big|_{\tau=0}\Big(exp_{\vec{u}}\tau\vec{\xi}\Big)^k = \tfrac{d^2}{d\tau^2}\Big|_{\tau=0}\,\gamma_{\tau,\vec{u}}(1)^k\\
&= \tfrac{d^2}{d\tau^2}\Big|_{\tau=0}\,\gamma_{1,\vec{u}}(\tau)^k = \partial_\tau\partial_\tau\gamma_{1,\vec{u}}(0)^k\\
&= -\Gamma^k_{ij}(\vec{u})\,\partial_\tau\gamma_{1,\vec{u}}(0)^i\,\partial_\tau\gamma_{1,\vec{u}}(0)^j = -\Gamma^k_{ij}(\vec{u})\,\xi^i\xi^j
\end{aligned}$$

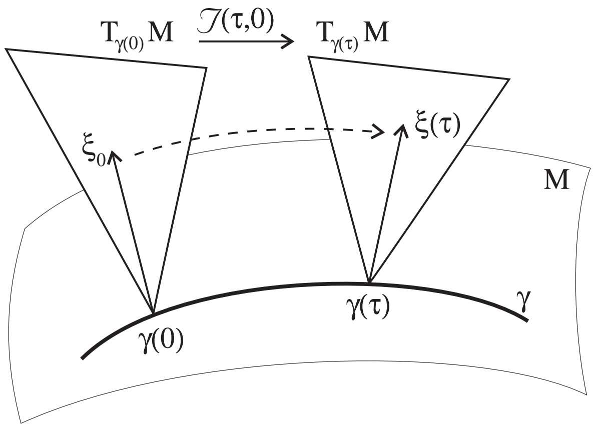

FIGURE A.1. Parallel transport of vector fields along curves

where we used that $\gamma_{1,\vec{u}}(\tau)$ satisfies $\nabla_\tau \partial_\tau \gamma_{1,\vec{u}(\tau)} = 0 \ \forall \tau \geq 0$ as well as $\partial_\tau \gamma_{1,\vec{u}}(0)^i = \xi^i$. Finally use *iv)* to get *v)*

$$\begin{aligned}
\partial_2\partial_2\partial_1 a(\vec{u},0)^k_{lij}\, x^l \xi^i \xi^j &= \partial_1\partial_2\partial_2 a(\vec{u},0)^k_{ijl}\, \xi^i\xi^j x^l \\
&= \frac{d}{d\mu}\Big|_{\mu=0} \frac{d^2}{d\tau^2}\Big|_{\tau=0} a(\vec{u}+\mu\vec{x},\tau\vec{\xi})^k = \frac{d}{d\mu}\Big|_{\mu=0} - \Gamma^k_{ij}|_{\vec{u}+\mu\vec{x}}\, \xi^i\xi^j \\
&= -\frac{d\Gamma^k_{ij}}{du^l}|_{\vec{u}}\, x^l \xi^i \xi^j
\end{aligned}$$

and *ii)* to get *vi)*

$$\begin{aligned}
\partial_2\partial_1\partial_1 a(\vec{u},0)^k_{ijl}\, x^i x^j \xi^l &= \partial_1\partial_1\partial_2 a(\vec{u},0)^k_{lij}\, \xi^l x^i x^j \\
&= \frac{d^2}{d\mu^2}\Big|_{\mu=0} \frac{d}{d\tau}\Big|_{\tau=0} a(\vec{u}+\mu\vec{x},\tau\vec{\xi})^k \\
&= \frac{d^2}{d\mu^2}\Big|_{\mu=0} \partial_2 a(\vec{u}+\mu\vec{x},0)^k_l\, \xi^l = \frac{d^2}{d\mu^2}\Big|_{\mu=0} \delta^k_l\, \xi^l = 0.
\end{aligned}$$

□

A.1.3. Parallel transport. In this subsection we follow closely the exposition in [**St88**], Teil 1. Let $\gamma \in C^\infty(\mathbb{R}, M)$ be a curve and $\xi_0 \in T_{\gamma(0)}M$, then we define the *parallel transport of the vector* ξ_0 *along* γ to be the linear map

$$\begin{aligned}
\mathcal{T}(\tau,0) : T_{\gamma(0)}M &\to T_{\gamma(\tau)}M \\
\xi_0 &\mapsto \xi(\tau)
\end{aligned} \tag{100}$$

(cf. figure A.1), where the vector field $\xi(\tau)$ along γ is defined by the initial value problem

$$\nabla_\tau \xi = 0\ ,\ \xi(0) = \xi_0. \tag{101}$$

In local coordinates we have a linear system of n first order ode's with n

initial values given

$$\partial_\tau \xi^k(\tau) + \Gamma^k_{ij}(\gamma(\tau))\partial_\tau \gamma^i(\tau)\xi^j(\tau) = 0\ ,\ \xi^k(0) = \xi^k_0\ ,\ k = 1,\dots,n. \tag{102}$$

By the existence and uniqueness theorem for ode's $\mathcal{T}(c,b)\mathcal{T}(b,a) = \mathcal{T}(c,a)$ and $\mathcal{T}(a,a) = id$, $a,b,c \in \mathbb{R}$. Being a linear map we may represent $\mathcal{T}(\tau,0)$ with respect to the basis of coordinate vector fields $\{\partial_1,\dots,\partial_n\}$ by

$$\mathcal{T}(\tau,0)^k_j\, \xi^j_0 = \xi^k(\tau)\ ,\ k = 1,\dots,n. \tag{103}$$

Lemma A.1.3. *$\frac{d}{d\tau}\big|_{\tau=0} \mathcal{T}(\tau,0)^k_j = -\Gamma^k_{ij}(\gamma(0))\ \partial_\tau\gamma^i(0)$ for all $k,j = 1,\dots,n$ and more generally*

$$\boxed{\partial_\tau \mathcal{T}(\tau,0)^k_j = -\Gamma^k_{il}(\gamma(\tau))\ \partial_\tau\gamma^i(\tau)\ \mathcal{T}(\tau,0)^l_j.}$$

Proof. $\xi(\tau) = \mathcal{T}(\tau,0)\,\xi_0$ is parallel along γ, i.e.

$$\begin{aligned} 0 &= \partial_\tau \xi^k(\tau) + \Gamma^k_{ij}(\gamma(\tau))\ \partial_\tau\gamma^i(\tau)\ \xi^j(\tau) \\ &= \partial_\tau\big(\mathcal{T}(\tau,0)^k_l\ \xi^l_0\big) + \Gamma^k_{ij}(\gamma(\tau))\ \partial_\tau\gamma^i(\tau)\big(\mathcal{T}(\tau,0)^j_l\ \xi^l_0\big)\ . \end{aligned}$$

Setting $\tau = 0$ and using $\mathcal{T}(0,0)^i_l = \delta^i_l$ the first statement follows. □

Lemma A.1.4. *$\frac{d}{d\tau}\big|_{\tau=0} \mathcal{T}(0,\tau)^k_j = -\frac{d}{d\tau}\big|_{\tau=0} \mathcal{T}(\tau,0)^k_j$.*

Proof. Apply $\frac{d}{d\tau}\big|_{\tau=0}$ to the identity $\mathcal{T}(\tau,0)^k_j\circ\mathcal{T}(0,\tau)^j_l = \delta^k_l$. □

Proposition A.1.5. *Let ξ be a vector field along γ, then*

$$\boxed{\tfrac{d}{d\tau}\big|_{\tau=0} \mathcal{T}(0,\tau)\ \xi(\tau) = (\nabla_\tau\xi)|_{\tau=0}\,.}$$

Proof. The k^{th} component of the LHS equals

$$\begin{aligned} \Big(\tfrac{d}{d\tau}\big|_{\tau=0} \mathcal{T}(0,\tau)^k_j\Big)\ \xi^j(0) + \delta^k_j\ \partial_\tau\xi^j(0) &= \Gamma^k_{ij}(\gamma(0))\ \partial_\tau\gamma^i(0)\xi^j(0) + \partial_\tau\xi^k(0) \\ &= (\nabla_\tau\xi)^k(0). \end{aligned}$$

The first equality holds by Lemma (A.1.3) and Lemma (A.1.4). □

The *parallel transport of the covector* $\eta^0 \in T^*_{\gamma(0)}M$ *along* γ is defined to be the linear map

$$\begin{aligned} \mathcal{T}^*(\tau,0) : T^*_{\gamma(0)}M &\to T^*_{\gamma(\tau)}M \\ \eta^0 &\mapsto \eta(\tau) \end{aligned} \tag{104}$$

where the covector field $\eta(\tau)$ along γ is defined by

$$\nabla^*_\tau \eta = 0\ ,\ \eta(0) = \eta^0\ , \tag{105}$$

or in local coordinates for $j = 1,\dots,n$

$$\partial_\tau\eta_j(\tau) - \Gamma^k_{ij}(\gamma(\tau))\ \partial_\tau\gamma^i(\tau)\ \eta_k(\tau) = 0 \quad , \quad \eta_j(0) = \eta^0_j. \tag{106}$$

Again $\mathcal{T}^*(c,b)\mathcal{T}^*(b,a) = \mathcal{T}^*(c,a)$, $\mathcal{T}^*(a,a) = id$ for $a,b,c \in \mathbb{R}$ and

$$\mathcal{T}^*(\tau,0)^k_j\ \eta^0_k = \eta_j(\tau)\ ,\ j = 1,\dots,n. \tag{107}$$

LEMMA A.1.6. *Let η, ξ be parallel (co)vector fields along γ, then*

$$\tfrac{d}{d\tau}\langle \eta(\tau), \xi(\tau)\rangle = 0$$

where $\langle\cdot,\cdot\rangle$ denotes evaluation of covectors on vectors.

PROOF. The left hand side equals

$$\begin{aligned}&\partial_\tau \eta_j(\tau)\, \xi^j(\tau) + \eta_k(\tau)\, \partial_\tau \xi^k(\tau)\\ &= \Gamma^k_{ij}(\gamma(\tau))\, \partial_\tau\gamma^i(\tau)\, \eta_k(\tau)\, \xi^j(\tau) - \eta_k(\tau)\, \Gamma^k_{ij}(\gamma(\tau))\, \partial_\tau\gamma^i(\tau)\, \xi^j(\tau)\end{aligned}$$

The second equality holds by using the assumption on η, ξ to be parallel, cf. (102), (106). □

LEMMA A.1.7. $\mathcal{T}^*(\tau,0) = \mathcal{T}(0,\tau)^*$, $\mathcal{T}^*(\tau,0)^m_k = \mathcal{T}(0,\tau)^m_k$.

PROOF. Lemma A.1.6 implies

$$\begin{aligned}\xi_0^k\eta_k^0 &= \langle \xi_0, \eta^0\rangle = \langle \xi(\tau), \eta(\tau)\rangle = \langle \mathcal{T}(\tau,0)\, \xi_0, \mathcal{T}^*(\tau,0)\, \eta^0\rangle\\ &= \mathcal{T}(\tau,0)^j_k \xi_0^k\, \mathcal{T}^*(\tau,0)^l_j \eta^0_l\end{aligned}$$

for all $\xi_0 \in T_{\gamma(0)}M$ and $\eta^0 \in T^*_{\gamma(0)}M$. This is equivalent to $\mathcal{T}^*(\tau,0)^*\mathcal{T}(\tau,0) = \mathbb{1}$ or $\mathcal{T}^*(\tau,0)^l_j\mathcal{T}(\tau,0)^j_k = \delta^l_k$. □

LEMMA A.1.8. $\boxed{\tfrac{d}{d\tau}\big|_{\tau=0} \mathcal{T}^*(\tau,0)^k_j = \Gamma^k_{ij}(\gamma(0))\, \partial_\tau\gamma^i(0)}$ $\forall k,j = 1,\dots,n$.

PROOF. To the LHS apply first Lemma A.1.7, then use Lemma A.1.4 and Lemma A.1.3. □

PROPOSITION A.1.9. *Let η be a covector field along γ, then*

$$\tfrac{d}{d\tau}\big|_{\tau=0} \mathcal{T}^*(0,\tau)\, \eta(\tau) = (\nabla^*_\tau \eta)|_{\tau=0}\,.$$

PROOF. Use lemma A.1.7 to get

$$\begin{aligned}\tfrac{d}{d\tau}\big|_{\tau=0}\Big(\mathcal{T}^*(0,\tau)^k_j\, \eta_k(\tau)\Big) &= \tfrac{d}{d\tau}\big|_{\tau=0}\Big(\mathcal{T}(\tau,0)^k_j\, \eta_k(\tau)\Big)\\ &= -\Gamma^k_{ij}(\gamma(0))\, \partial_\tau\gamma^i(0)\, \eta_k(0) + \delta^k_j\, \partial_\tau\eta_k(0)\\ &= (\nabla^*_\tau\eta)_j(0).\end{aligned}$$

The second equality follows by the product rule and lemma A.1.3. □

LEMMA A.1.10. *Let θ be a curve in $T^*_{\gamma(0)}M$ such that $\theta(0) = v$ and $\tfrac{d}{d\tau}\big|_{\tau=0}\theta(\tau) = \eta^0$, γ and $\mathcal{T}^*$ as above, then*

$$\tfrac{D}{d\tau}\big|_{\tau=0} \mathcal{T}^*(\tau,0)\, \theta(\tau) = \eta^0.$$

PROOF. The j^{th} component of the LHS is given by

$$\begin{aligned}&\tfrac{d}{d\tau}\big|_{\tau=0}\Big(\mathcal{T}^*(\tau,0)^k_j\, \theta_k(\tau)\Big) - \Gamma^k_{ij}(\gamma(0))\, \delta^l_k\, \theta_l(0)\, \partial_\tau\gamma^i(0)\\ &= \Big(\tfrac{d}{d\tau}\big|_{\tau=0} \mathcal{T}^*(\tau,0)^k_j\Big)\, v_k + \delta^k_j\, \eta^0_k - \Gamma^k_{ij}(0)\, v_k\, \xi^i = \eta^0_j\end{aligned}$$

where the last step follows using Lemma A.1.8. □

The next result is essential to derive the fundamental quadratic estimate in chapter 5 section 5.1. Note that until the end of this subsection the curves along which we transport (co)vectors are geodesics of the form $\gamma_\xi(\tau) = a(u, \tau\xi) = exp_u \tau\xi$.

LEMMA A.1.11. *In local coordinates* $\{u^1, \dots, u^n\}$ *and for* $\tau \geq 0$, $\xi \in T_uM$ *and* $k, l = 1, \dots, n$ *it holds*

$$\boxed{\mathcal{T}^*(\tau,0)_l^k = g_{lr}|_{a(u,\tau\xi)}\, \mathcal{T}(\tau,0)_s^r\, g^{sk}(u)} \tag{108}$$

and

$$\begin{aligned} 0 = &- \partial_\tau \mathcal{T}^*(\tau,0)_l^k + \frac{\partial g_{lr}}{\partial u^j}|_{a(u,\tau\xi)}\, (\partial_2 a(u,\tau\xi)\,\xi)^j\, \mathcal{T}(\tau,0)_s^r\, g^{sk}(u) \\ &- g_{lr}|_{a(u,\tau\xi)}\, \Gamma_{ij}^r|_{a(u,\tau\xi)}\, (\partial_2 a(u,\tau\xi)\,\xi)^i\, \mathcal{T}(\tau,0)_s^j\, g^{sk}(u) \end{aligned}$$

which for $\tau = 0$ *reduces to*

$$0 = - \tfrac{d}{d\tau}\Big|_{\tau=0} \mathcal{T}^*(\tau,0)_l^k + \frac{\partial g_{lr}}{\partial u^j}(u)\,\xi^j g^{rk}(u) - g_{lr}(u)\,\Gamma_{ij}^r(u)\,\xi^i g^{jk}(u).$$

Moreover,

$$\begin{aligned} 0 = &- \tfrac{d^2}{d\tau^2}\Big|_{\tau=0} \mathcal{T}^*(\tau,0)_l^k + \frac{\partial^2 g_{lr}}{\partial u^i \partial u^j}(u)\,\xi^i\xi^j g^{rk}(u) \\ &- \frac{\partial g_{lr}}{\partial u^j}(u)\,\Gamma_{is}^j(u)\,\xi^i\xi^s\, g^{rk}(u) - 2\frac{\partial g_{lr}}{\partial u^j}(u)\,\xi^j\,\Gamma_{is}^r(u)\,\xi^i\, g^{sk}(u) \\ &- g_{lr}(u)\,\frac{\partial \Gamma_{ij}^r(u)}{\partial u^s}\,\xi^s\xi^i\, g^{jk}(u) + g_{lr}(u)\,\Gamma_{ij}^r(u)\,\Gamma_{sm}^i(u)\,\xi^s\xi^m\, g^{jk}(u) \\ &+ g_{lr}(u)\,\Gamma_{ij}^r(u)\,\xi^i\,\Gamma_{ms}^j(u)\,\xi^m\, g^{sk}(u). \end{aligned}$$

PROOF. Assume that (108) holds and take the derivative with respect to τ, then the second statement follows using the product and chain rules as well as lemma A.1.3 in the last term of the sum. Evaluation at $\tau = 0$ gives the third statement. The last one follows by taking another τ-derivative of statement 2 and evaluating at $\tau = 0$. We use again lemma A.1.3 as well as proposition A.1.2 *iv*).

To prove (108), let $\eta^0 \in T_u^*M$ and set

$$\begin{aligned} \eta(\tau) &= \mathcal{T}^*(\tau,0)_l^k\, \eta_k^0\, du^l \\ Y(\tau) &= g_{lr}|_{a(u,\tau\xi)}\, \mathcal{T}(\tau,0)_s^r\, g^{sk}(u)\, \eta_k^0\, du^l. \end{aligned}$$

Both are covector fields along the curve $\tau \mapsto a(u,\tau\xi) = exp_u\tau\xi$ and $\eta(0) = \eta^0 = Y(0)$. By definition $\eta(\tau)$ satisfies $\nabla_\tau \eta \equiv 0$ and $\eta \equiv Y$ will follow once we have shown $\nabla_\tau Y \equiv 0$ (uniqueness of solution to initial value problem of a system of 1^{st} order ODE's). Consider the j^{th} component of $\nabla_\tau Y$ and use the definition of Y in the second equality to get

$$\begin{aligned}
(\nabla_\tau Y)_j(\tau) &= \partial_\tau Y_j(\tau) - \Gamma^k_{ij}|_{a(u,\tau\xi)}\, (\partial_\tau a(u,\tau\xi))^i\, Y_k(\tau) \\
&= \frac{\partial g_{jr}}{\partial u^l}|_{a(u,\tau\xi)}\, (\partial_\tau a(u,\tau\xi))^l\, \mathcal{T}(\tau,0)^r_s\, g^{sk}(u)\, \eta^0_k \\
&\quad + g_{jr}|_{a(u,\tau\xi)}\, (\partial_\tau \mathcal{T}(\tau,0)^r_s)\, g^{sk}(u)\, \eta^0_k \\
&\quad - \Gamma i j^k|_{a(u,\tau\xi)}\, (\partial_\tau a(u,\tau\xi))^i \left(g_{kr}|_{a(u,\tau\xi)}\, \mathcal{T}(\tau,0)^r_s\, g^{sl}(u)\, \eta^0_l \right) \\
&= \left\{ \frac{\partial g_{jr}}{\partial u^i} - g_{jl}\Gamma^l_{ir} - g_{lr}\Gamma^l_{ij} \right\} |_{a(u,\tau\xi)}\, (\partial_\tau a(u,\tau\xi))^i\, \mathcal{T}(\tau,0)^r_s\, g^{sk}(u)\, \eta^0_k \\
&= 0
\end{aligned}$$

where in the third equality we replaced $\partial_\tau \mathcal{T}(\tau,0)^r_s$ according to lemma A.1.3 and renamed several indices. The fourth equality follows as the term in brackets (the covariant derivative of the metric tensor) is zero, which may be shown by direct calculation using formula (95) for the Christoffel symbols in terms of derivatives of the metric. □

A.2. Linearization at a zero

Let $u \in C^\infty(\mathbb{R} \times S^1, M)$ be a smooth cylinder in M and $w \in \Gamma(u^*T^*M)$ a smooth covector field along u. The coordinates on $\mathbb{R} \times S^1$ are denoted by (s,t). Recall the nonlinear equations

$$\boxed{\begin{aligned} &i)\ \partial_s u - g^{-1}(u)\nabla_t^* w - {}^g\nabla V(t,u) = 0\ , \\ &ii)\ \epsilon^2 \nabla_s^* w + g(u)\partial_t u - w = 0\ . \end{aligned}} \tag{109}$$

Our goal is to linearize these equations. To do so we represent them in local coordinates and linearize the local expression. In order to get back to intrinsic quantities it will be crucial to *use the nonlinear equations* during the linearization procedure. One way to think of this is interpreting (109) as a zero of a section $\mathcal{F}_\epsilon$ in a Banach bundle and then linearizing the section at this zero.

Let now u be represented in local coordinates by $\vec{u} = (u^1, \ldots, u^n)$ and $w \in T_u^*M$ in the hereby induced coordinates on T^*M by $(\vec{u}, \vec{v}) = (u^1, \ldots, u^n; v_1, \ldots, v_n)$. With respect to these coordinates the equations take the form

$$\boxed{\begin{aligned} &i)\ \partial_s u^k - g^{kl}(\vec{u})\left(\partial_t v_l - \Gamma_{jl}^\lambda(\vec{u})(\partial_t u^j)v_\lambda\right) - g^{kl}(\vec{u})\frac{\partial V(t,\vec{u})}{\partial u^l} = 0\ , \\ &ii)\ \epsilon^2\left(\partial_s v_k - \Gamma_{ik}^j(\vec{u})(\partial_s u^i)v_j\right) + g_{kl}(\vec{u})\partial_t u^l - v_k = 0\ , \end{aligned}} \tag{110}$$

for $k = 1, \ldots, n$. Assume that $(\vec{u}, \vec{v})$ solves (110). Let $(\vec{u}_\tau, \vec{v}_\tau)$, $\tau \in (-\delta, \delta)$, $\delta > 0$ small, be a variation of $(\vec{u}, \vec{v})$, that is it satisfies

$$\begin{aligned} (\vec{u}_0, \vec{v}_0) &= (\vec{u}, \vec{v})\ , \\ \tfrac{d}{d\tau}(\vec{u}_\tau, \vec{v}_\tau)\ |_{\tau=0} &= (\vec{\xi}, \vec{y})\ . \end{aligned} \tag{111}$$

Note that varying $\vec{u}$ gives the vector $\vec{\xi}$, but varying $\vec{v}$ will not give us a covector, it just gives a local quantity $\vec{y}$ without intrinsic meaning. Later on we will replace y_l by $\eta_l + \Gamma_{lk}^j(u)\xi^k v_j$ in order to interpret the linearization of (110) as a section of $u^*TM \oplus u^*T^*M$.
Replacing $(\vec{u}, \vec{v})$ in (110) by the variation $(\vec{u}_\tau, \vec{v}_\tau)$ and applying $\frac{d}{d\tau}\ |_{\tau=0}$ we get (from now on we simply write u instead of $\vec{u}$)

$$\begin{aligned} &i)\ \partial_s \xi^k - \frac{\partial g^{kl}(u)}{\partial u^i}\xi^i\left(\partial_t v_l - \Gamma_{jl}^\lambda(u)(\partial_t u^j)v_\lambda\right) \\ &\quad - g^{kl}(u)\Big(\partial_t y_l - \frac{\partial \Gamma_{jl}^\lambda|_u}{\partial u^i}\xi^i(\partial_t u^j)v_\lambda - \Gamma_{jl}^\lambda|_u(\partial_t \xi^j)v_\lambda - \Gamma_{jl}^\lambda|_u(\partial_t u^j)y_\lambda\Big) \\ &\quad - \frac{\partial g^{kl}(u)}{\partial u^i}\xi^i\frac{\partial V(t,u)}{\partial u^l} - g^{kl}(u)\frac{\partial^2 V(t,u)}{\partial u^l \partial u^i}\xi^i = 0\ , \end{aligned} \tag{112}$$

$$ii)\ \epsilon^2\partial_s y_k - \epsilon^2\frac{\partial\Gamma^j_{ik}(u)}{\partial u^l}\xi^l(\partial_s u^i)v_j - \epsilon^2\Gamma^j_{ik}(u)(\partial_s\xi^i)v_j - \epsilon^2\Gamma^j_{ik}(u)(\partial_s u^i)y_j$$
$$+\frac{\partial g_{kl}(u)}{\partial u^r}\xi^r\partial_t u^l + g_{kl}(u)\partial_t\xi^l - y_k = 0\ .$$

Now we define the component of a covector (using that we have fixed a solution $(\vec{u},\vec{v})$ of (110))

(113)
$$\eta_l = y_l - \Gamma^\lambda_{il}(u)\xi^i v_\lambda$$

and replace y_l in (112) by this new quantity. Moreover we are going to use the identities

(114)
$$\frac{\partial g^{kl}(u)}{\partial u^i} = -g^{k\nu}\frac{\partial g_{\nu\mu}(u)}{\partial u^i}g^{\mu l}$$

(which follows by applying the operator $\frac{\partial}{\partial u^i}$ to the identity $g^{k\nu}g_{\nu\mu}=\delta^k_\mu$) as well as

(115)
$$\partial_t y_l = \partial_t\eta_l + \frac{\partial\Gamma^\lambda_{il}(u)}{\partial u^\nu}(\partial_t u^\nu)\xi^i v_\lambda + \Gamma^\lambda_{il}(u)(\partial_t\xi^i)v_\lambda + \Gamma^\lambda_{il}(u)\xi^i\partial_t v_\lambda$$

(apply ∂_t to (113)) to get

(116)
$$i)\ \partial_s\xi^k + g^{k\nu}(u)\frac{\partial g^{k\nu}(u)}{\partial u^i}g^{\mu l}(u)\xi^i\left(\partial_t v_l - \Gamma^\lambda_{jl}(u)\partial_t u^j v_\lambda\right)$$
$$- g^{kl}(u)\left(\partial_t\eta_l + \frac{\partial\Gamma^\lambda_{il}(u)}{\partial u^\nu}(\partial_t u^\nu)\xi^i v_\lambda + \Gamma^\lambda_{il}(u)(\partial_t\xi^i)v_\lambda + \Gamma^\lambda_{il}(u)\xi^i(\partial_t v_\lambda)\right)$$
$$+ g^{kl}(u)\left(\frac{\partial\Gamma^\lambda_{jl}(u)}{\partial u^i}\xi^i(\partial_t u^j)v_\lambda + \Gamma^\lambda_{jl}(u)(\partial_t\xi^j)v_\lambda\right)$$
$$+ g^{kl}(u)\left(\Gamma^\lambda_{jl}(u)(\partial_t u^j)\eta_\lambda + \Gamma^\lambda_{jl}(u)(\partial_t u^j)\Gamma^s_{i\lambda}(u)\xi^i v_s\right)$$
$$- \frac{\partial g^{kl}(u)}{\partial u^i}\xi^i\frac{\partial V(t,u)}{\partial u^l} - g^{kl}(u)\frac{\partial^2 V(t,u)}{\partial u^l\partial u^i}\xi^i$$
$$+ \Gamma^k_{ij}(u)\xi^i\left(\partial_s u^j - g^{jl}(u)\partial_t v_l + g^{jl}(u)\Gamma^\lambda_{sl}(u)(\partial_t u^s)v_\lambda - g^{jl}(u)\frac{\partial V(t,u)}{\partial u^l}\right)$$
$$= 0\ ,$$
$$ii)\ \epsilon^2\left(\partial_s\eta_k + \frac{\partial\Gamma^\lambda_{ik}(u)}{\partial u^\nu}(\partial_s u^\nu)\xi^i v_\lambda + \Gamma^\lambda_{ik}(u)(\partial_s\xi^i)v_\lambda + \Gamma^\lambda_{ik}(u)\xi^i(\partial_s v_\lambda)\right)$$
$$- \epsilon^2\frac{\partial\Gamma^j_{ik}(u)}{\partial u^l}\xi^l(\partial_s u^i)v_j - \epsilon^2\Gamma^j_{ik}(u)(\partial_s\xi^i)v_j$$

$$- \epsilon^2 \left(\Gamma^j_{ik}(u)(\partial_s u^i)\eta_j + \Gamma^j_{ik}(u)(\partial_s u^i)\Gamma^\lambda_{ij}(u)\xi^i v_\lambda \right)$$
$$+ \frac{\partial g_{kl}(u)}{\partial u^r}\xi^r(\partial_t u^l) + g_{kl}(u)\partial_t \xi^l - \left(\eta_k - \Gamma^\lambda_{ik}(u)\xi^i v_\lambda \right) = 0 \; .$$

Note that in (116)i) we have added in the last but one line the nonlinear equation (110)i) as this is zero (here we have to use the fact that we are linearizing at a zero of our nonlinear map). The additional terms coming in are essential to identify the global expressions $\nabla_s \xi$ and $-\nabla_\xi {}^g\nabla V(t,u)$:

$$\underline{(\nabla_s \xi)^k} = \partial_s \xi^k + \Gamma^k_{ji}(u)\xi^i \partial_s u^j = (1^{st} + 14^{th}) \text{ term in (116)i)} \; , \tag{117}$$

$$\begin{aligned} \underline{-(\nabla \xi^g \nabla V(t,u))^k} &= -\xi^i \frac{\partial ({}^g\nabla V)^k(t,u)}{\partial u^i} - \Gamma^k_{ij}(u)\xi^i ({}^g\nabla V)^j(t,u) \\ &\qquad , \text{where } ({}^g\nabla V)^k(t,u) = g^{kl}(u)\frac{\partial V(t,u)}{\partial u^l} \\ &= -\xi^i \frac{\partial g^{kl}(u)}{\partial u^i}\frac{\partial V(t,u)}{\partial u^l} - \xi^i g^{kl}(u)\frac{\partial^2 V(t,u)}{\partial u^l \partial u^i} \\ &\quad - \Gamma^k_{ij}(u)\xi^i g^{jl}(u)\frac{\partial V(t,u)}{\partial u^l} \\ &= (12^{th} + 13^{th} + 17^{th}(\text{last})) \text{ term in (116)i)} \; . \end{aligned} \tag{118}$$

Moreover term 6 and 9 of (116)i) cancel and

$$\begin{aligned} \underline{-g^{-1}(u)\nabla^*_t \eta} &= -g^{kl}(u)\left(\partial_t \eta_l - \Gamma^\lambda_{jl}(u)(\partial_t u^j)\eta_\lambda \right) \\ &= (4^{th} + 10^{th}) \text{ term in (116)i)} \; . \end{aligned} \tag{119}$$

Subtracting from equation (116)i) the terms involved in (117)-(119) its lhs reduces to

$$\begin{aligned} & g^{k\nu}(u)\frac{\partial g_{\nu\mu}(u)}{\partial u^i} g^{\mu l}(u)\xi^i \partial_t v_l - g^{k\nu}(u)\frac{\partial g_{\nu\mu}(u)}{\partial u^i} g^{\mu l}(u)\xi^i \Gamma^\lambda_{jl}(u)(\partial_t u^j) v_\lambda \\ & - g^{kl}(u)\frac{\partial \Gamma^\lambda_{il}(u)}{\partial u^\nu}(\partial_t u^\nu)\xi^i v_\lambda - g^{kl}(u)\Gamma^\lambda_{il}(u)\xi^i \partial_t v_\lambda \\ & + g^{kl}(u)\frac{\partial \Gamma^\lambda_{jl}(u)}{\partial u^i}\xi^i(\partial_t u^j)v_\lambda + g^{kl}(u)\Gamma^\lambda_{jl}(u)(\partial_t u^j)\Gamma^s_{i\lambda}(u)\xi^i v_s \\ & - g^{jl}(u)\Gamma^k_{ij}(u)\xi^i \partial_t v_l + g^{jl}(u)\Gamma^k_{ij}(u)\Gamma^\lambda_{sl}(u)\xi^i(\partial_t u^s)v_\lambda \end{aligned} \tag{120}$$

and we are going to show that this equals the curvature tensor defined in equation (96)

$$\left(R(\partial_t u, \xi) g^{-1} w \right)^l = R^l_{kij}(u)(\partial_t u^i)\xi^j g^{k\lambda}(u) v_\lambda \; . \tag{121}$$

First of all we observe that the sum of terms 1,4 and 7 in (120) is zero as was to be expected, because they are the only ones containing a factor $\partial_t v_l$. That their sum is zero may be seen as follows: Replace the Christoffel symbols according to (95), then rename the appropriate summation indices. So (120) equals

$$
\begin{aligned}
&v_\lambda \xi^j(\partial_t u^k) g^{ls}(u)\left(-\frac{\partial g_{s\mu}(u)}{\partial u^j} g^{\mu\nu}(u)\Gamma^\lambda_{k\nu}(u) - \frac{\partial \Gamma^\lambda_{js}(u)}{\partial u^k} + \frac{\partial \Gamma^\lambda_{ks}(u)}{\partial u^j}\right)\\
&+v_\lambda \xi^j(\partial_t u^k)\left(g^{ls}(u)\Gamma^r_{ks}(u)\Gamma^\lambda_{jr}(u) + g^{s\nu}(u)\Gamma^l_{js}(u)\Gamma^\lambda_{k\nu}(u)\right)\\
&=v_\lambda \xi^j(\partial_t u^k) g^{ls}(u)\left(-2\Gamma^\nu_{sj}(u)\Gamma^\lambda_{k\nu}(u) + g^{\nu\mu}(u)\frac{\partial g_{j\mu}(u)}{\partial u^s}\Gamma^\lambda_{k\nu}(u)\right)\\
&\quad + v_\lambda \xi^j(\partial_t u^k) g^{ls}(u)\left(-g^{\nu\mu}(u)\frac{\partial g_{sj}(u)}{\partial u^\mu}\Gamma^\lambda_{k\nu}(u) - \frac{\partial \Gamma^\lambda_{js}(u)}{\partial u^k} + \frac{\partial \Gamma^\lambda_{ks}(u)}{\partial u^j}\right)\\
&\quad + v_\lambda \xi^j(\partial_t u^k)\left(+g^{ls}(u)\Gamma^\nu_{ks}(u)\Gamma^\lambda_{j\nu}(u) + g^{s\nu}(u)\Gamma^l_{js}(u)\Gamma^\lambda_{k\nu}(u)\right),
\end{aligned}
\tag{122}
$$

here we replaced

$$
-g^{\mu\nu}(u)\frac{\partial g_{s\mu}(u)}{\partial u^j} = -2\Gamma^\nu_{sj}(u) + g^{\mu\nu}(u)\frac{\partial g_{j\mu}(u)}{\partial u^s} - g^{\mu\nu}(u)\frac{\partial g_{sj}(u)}{\partial u^\mu}
\tag{123}
$$

(coming from (95)) in the first term of (122). We also need to compute the sum of terms 2,3,6 of the right hand side of (122)

$$
\begin{aligned}
&g^{ls}(u)g^{\nu\mu}(u)\frac{\partial g_{j\mu}(u)}{\partial u^s} - g^{ls}(u)g^{\nu\mu}(u)\frac{\partial g_{sj}(u)}{\partial u^\mu} + g^{s\nu}(u)\Gamma^l_{js}(u)\\
&=g^{ls}(u)g^{\nu\mu}(u)\frac{\partial g_{j\mu}(u)}{\partial u^s} - g^{ls}(u)g^{\nu\mu}(u)\frac{\partial g_{sj}(u)}{\partial u^\mu} + \tfrac{1}{2}g^{s\nu}(u)g^{lr}(u)\frac{\partial g_{sr}(u)}{\partial u^j}\\
&\quad + \tfrac{1}{2}g^{s\nu}(u)g^{lr}(u)\frac{\partial g_{jr}(u)}{\partial u^s} - \tfrac{1}{2}g^{s\nu}(u)g^{lr}(u)\frac{\partial g_{sj}(u)}{\partial u^r}\\
&=\tfrac{1}{2}g^{lr}(u)g^{\nu\mu}(u)\left(+\frac{\partial g_{\mu r}(u)}{\partial u^j} + \frac{\partial g_{\mu j}(u)}{\partial u^r} - \frac{\partial g_{jr}(u)}{\partial u^\mu}\right)\\
&=g^{lr}(u)\Gamma^\nu_{jr}(u)
\end{aligned}
\tag{124}
$$

In the first equality we replaced the Christoffel symbol according to (95), the second equality follows by renaming indices, the third one again uses (95). We proceed by replacing these terms in the right hand side of (122) and get

$$
\begin{aligned}
&v_\lambda \xi^j(\partial_t u^k) g^{ls}(u)\left(-\Gamma^\lambda_{k\nu}\Gamma^\nu_{js} + \Gamma^\lambda_{j\nu}\Gamma^\nu_{ks} + \frac{\partial \Gamma^\lambda_{ks}}{\partial u^j} - \frac{\partial \Gamma^\lambda_{js}}{\partial u^k}\right)\\
&=v_\lambda \xi^j(\partial_t u^k) g^{ls}(u) R^\lambda_{sjk}(u)\\
&=v_\lambda \xi^j(\partial_t u^k) g^{ls}(u) g^{\lambda i}(u) R_{isjk}(u)\\
&=-\, v_\lambda \xi^j(\partial_t u^k) g^{\lambda i}(u) g^{ls}(u) R_{sijk}(u)\\
&=-\left(R(u)(\xi^j\partial_{u^j},(\partial_t u^k)\partial_{u^k}) g^{i\lambda}(u) v_\lambda \partial_{u^i}\right)^l\\
&=\underline{-\left(R(u)(\xi,\partial_t u) g^{-1} w\right)^l}\,.
\end{aligned}
\tag{125}
$$

Summarizing the results obtained so far gives the first component of the linearized equation

$$\boxed{\nabla_s\xi - g^{-1}\nabla_t^*\eta - R(\xi,\partial_t u)g^{-1}w - \nabla_\xi{}^g\nabla V(t,u) = 0\ .} \tag{126}$$

Now we are going to analyze the second equation (116)ii). Terms three and six cancel each other. The first and seventh term together give the covariant derivative

$$\underline{\epsilon^2\nabla_s^*\eta} = \epsilon^2\left(\partial_s\eta_k - \Gamma_{ik}^j(u)(\partial_s u^i)\eta_j\right)du^k\ , \tag{127}$$

term 11 clearly is

$$\underline{-\eta} = -\eta_k du^k\ . \tag{128}$$

Ignoring terms $1,3,6,7,11$ and replacing $\epsilon^2\partial_s v_\lambda$ in term 4 by the second nonlinear equation (110)ii) the lhs of (116)ii) reduces to

$$\begin{aligned}
&\epsilon^2\frac{\partial\Gamma_{ik}^\lambda(u)}{\partial u^r}(\partial_s u^r)\xi^i v_\lambda\\
&+\xi^i\left(\epsilon^2\Gamma_{r\nu}^\lambda(u)\Gamma_{ik}^\nu(u)(\partial_s u^r)v_\lambda - g_{\lambda r}(u)\Gamma_{ik}^\lambda(u)\partial_t u^r + \Gamma_{ik}^\lambda(u)v_\lambda\right)\\
&-\epsilon^2\xi^i v_\lambda\left(\frac{\partial\Gamma_{rk}^\lambda(u)}{\partial u^i}\partial_s u^r + (\partial_s u^r)\Gamma_{ij}^\lambda(u)\Gamma_{rk}^j(u)\right)\\
&+\frac{\partial g_{kr}(u)}{\partial u^i}\xi^i\partial_t u^r + g_{kl}(u)\partial_t\xi^l - \Gamma_{ik}^\lambda(u)\xi^i v_\lambda\\
&=\epsilon^2\xi^i v_\lambda(\partial_s u^r)\left(\frac{\partial\Gamma_{ik}^\lambda(u)}{\partial u^r} - \frac{\partial\Gamma_{rk}^\lambda(u)}{\partial u^i} + \Gamma_{r\nu}^\lambda(u)\Gamma_{ik}^\nu(u) - \Gamma_{i\nu}^\lambda(u)\Gamma_{rk}^\lambda(u)\right)\\
&\quad+\xi^i(\partial_t u^r)\left(\frac{\partial g_{kr}(u)}{\partial u^i} - g_{\lambda r}(u)\Gamma_{ik}^\lambda(u)\right) + g_{kl}(u)\partial_t\xi^l\\
&=-\epsilon^2 R_{kir}^\lambda(u)\xi^i(\partial_s u^r)v_\lambda + \xi^i(\partial_t u^r)g_{ks}(u)\Gamma_{ri}^s(u) + g_{ks}(u)\partial_t\xi^s\\
&=\epsilon^2 g_{kj}(u)R_{lir}^j(u)\xi^i(\partial_s u^r)(g^{lj}(u)v_\lambda) + g_{ks}(u)\left(\partial_t\xi^s + \Gamma_{ri}^s(u)(\partial_t u^r)\xi^i\right)\\
&=\underline{\epsilon^2 g_{kj}(u)\left(R(\xi,\partial_s u)g^{-1}w\right)^j + g_{ks}(u)(\nabla_t\xi)^s}\ .
\end{aligned} \tag{129}$$

The first equality follows just by putting the terms in appropriate order and observing that terms 4 and 9 cancel each other. In the second equality we use the identity (98) for the curvature tensor as well as the following fact

$$\begin{aligned}
&\frac{\partial g_{kr}(u)}{\partial u^i} - g_{\lambda r}(u)\Gamma_{ik}^\lambda(u)\\
&=\frac{\partial g_{kr}(u)}{\partial u^i} - g_{\lambda r}(u)\tfrac{1}{2}g^{\lambda l}(u)\left(\frac{\partial g_{kl}(u)}{\partial u^i} + \frac{\partial g_{li}(u)}{\partial u^k} - \frac{\partial g_{ki}(u)}{\partial u^l}\right)\\
&=\tfrac{1}{2}\left(\frac{\partial g_{kr}(u)}{\partial u^i} + \frac{\partial g_{ki}(u)}{\partial u^r} - \frac{\partial g_{ri}(u)}{\partial u^k}\right)\\
&=g_{ks}(u)\Gamma_{ri}^s(u)\ ,
\end{aligned} \tag{130}$$

where we used (95) in the second and fourth equality. The third equality is induced by

$$-R^{\lambda}_{kir}(u) = -g^{\lambda l}(u)R_{lkir}(u) = g^{\lambda l}(u)R_{klir}(u) \\ = g^{\lambda l}(u)g_{kj}(u)R^{j}_{lir}(u) \tag{131}$$

Now the second linearized equation is given by

$$\boxed{\epsilon^2\nabla^*_s\eta + \epsilon^2 gR(\xi,\partial_s u)g^{-1}w + g\nabla_t\xi - \eta = 0\ .} \tag{132}$$

Setting ϵ in the nonlinear equations (110) and in the linear ones (126),(132) formally zero we get

$$\begin{aligned} &i)\partial_s u - \nabla_t\partial_t u - {}^g\nabla V(t,u) = 0 \\ &ii)v = g(u)\partial_t u \end{aligned} \tag{133}$$

and therefore

$$\begin{aligned} &i)\nabla_s\xi - \nabla_t\nabla_t\xi - R(\xi,\partial_t u)\partial_t u - \nabla_\xi{}^g\nabla V(t,u) = 0 \\ &ii)\eta = g(u)\nabla_t\xi. \end{aligned} \tag{134}$$

The riemannian geometer observes immediately the occurence of the geodesic curvature in (133) $i)$ and the Jacobi equation in (134) $i)$. In many textbooks one term of the Jacobi equation differs in sign. This is a consequence of the nonexistence of a standard sign convention in the definition of the curvature tensor. Note that in the second term in (134) $i)$ we used (setting $X = \partial_t u$)

Lemma A.2.1. *Let $X,\xi \in \Gamma(TM)$, $\eta \in \Gamma(T^*M)$, $g : TM \to T^*M$ the metric isomorphism, ∇ respectively ∇^* the Levi-Civita connection on TM respectively T^*M, then*

$$\begin{aligned} &i)\ g^{-1}\circ\nabla^*_X(g\xi) = \nabla_X\xi\ , \\ &ii)\ \nabla^*_X\eta = g\circ\nabla_X(g^{-1}\eta)\ . \end{aligned}$$

Proof. ad $i)$

$$\begin{aligned} \left(g^{-1}(u)\circ\nabla^*_X(g\xi)\right)^s &= g^{si}(u)\left(X^j\frac{\partial(g_{ik}(u)\xi^k)}{\partial u^j} - \Gamma^k_{ji}(u)X^j(g_{lk}(u)\xi^k)\right) \\ &= X^j\frac{\partial\xi^s}{\partial u^j} + X^j\xi^k g^{si}(u)\left(\frac{\partial g_{ik}(u)}{\partial u^j} - g_{lk}(u)\Gamma^l_{ji}(u)\right) \\ &= X^j\frac{\partial\xi^s}{\partial u^j} + X^j\xi^k\Gamma^s_{jk}(u) \\ &= \nabla_X\xi\ , \end{aligned}$$

where in the last but one equality we used (95).
ad $ii)$ Apply g from the left to $i)$ and set $\xi = g^{-1}\eta$. □

Remark A.2.2. (*Linearized flow satisfies linearized equations*) Ignoring the terms involving s-derivatives and setting $\epsilon = 1$, the nonlinear equations

(109) reduce to $\dot{z} = X_{H_t}(z) = -J_g \nabla H_t(z)$ for a smooth loop z in T^*M. As was shown above, linearizing this equation at a solution z leads to

$$\begin{aligned} -g^{-1}(x)\nabla_t^*\eta - R(\xi, \dot{x})\dot{x} - \nabla_\xi \nabla V_t(x) &= 0 \\ g(x)\nabla_t \xi - \eta &= 0 \end{aligned} \tag{135}$$

where $(\xi, \eta) \in \Gamma(x^*TM \oplus x^*T^*M)$ and $x = \tau_M^* z$. On the other hand X_{H_t} gives rise to the time-t-map φ_t on T^*M and $z(t) = \varphi_t z_0$ for $z_0 = z(0)$ and $\dot{z} = X_{H_t}(z)$. The crucial fact is that its linearization $d\varphi_t(z_0) : T_{z_0}T^*M \to T_{\varphi_t z_0}T^*M$ along the solution z satisfies the linearized equations (135) for any initial condition (ξ_0, η_0). To prove this we work in natural coordinates and pick a particular variation of $\vec{z} = (\vec{x}, \vec{y})$ in (111). Namely, let $\vec{z}(0) = \vec{z}_0 = (\vec{x}_0, \vec{y}_0)$ and

$$(\vec{x}^{\,\tau}, \vec{y}^{\,\tau}) = \vec{\varphi}_t(\vec{x}_0 + \tau\vec{\xi}_0, \vec{y}_0 + \tau\vec{\eta}_0')$$

then

$$(\vec{x}^{\,0}, \vec{y}^{\,0}) = \vec{\varphi}_t(\vec{x}_0, \vec{y}_0) = (\vec{x}, \vec{y})$$

$$\tfrac{d}{d\tau}\Big|_{\tau=0} (\vec{x}^{\,\tau}, \vec{y}^{\,\tau}) = d\vec{\varphi}_t \,|_{(\vec{x}_0, \vec{y}_0)} (\vec{\xi}_0, \vec{\eta}_0').$$

Setting $(\vec{\xi}, \vec{\eta}') = \frac{d}{d\tau}\,|_{\tau=0}\, (\vec{x}^{\,\tau}, \vec{y}^{\,\tau})$ we obtain a variation of the form (111) and the result follows.

A.3. Linearization in a local trivialization

Recall that the section $\mathcal{F}_\epsilon : \mathcal{P}^{1,p}_{x^-,x^+} \to \mathcal{E}^p$ is given for a smooth cylinder w in T^*M by

$$\boxed{\mathcal{F}_\epsilon(w) = \begin{pmatrix} \partial_s u - g^{-1}(u)\nabla_t^* w - \nabla V(t,u) \\ \nabla_s^* w + \epsilon^{-2} g(u)\partial_t u - \epsilon^{-2} w \end{pmatrix}}$$

where

$$u(s,t) = \tau_M^* w(s,t)$$

and $\tau_M^* : T^*M \to M$. As we have seen in the former section its linearization $\mathcal{D}^\epsilon_{w_0} = d\mathcal{F}_\epsilon(w_0)$ is well-defined at a zero $w_0 \in \mathcal{F}_\epsilon^{-1}(0) = \mathcal{M}^\epsilon(x^-,x^+)$ and it is a linear operator on the space $C_0^\infty(\mathbb{R}\times S^1, u^*TM \oplus u^*T^*M)$.

Locally trivializing the bundle of maps, the section $\mathcal{F}_\epsilon$ induces a nonlinear map between linear spaces: Fix any smooth cylinder w, then define

$$\boxed{\begin{gathered} \mathcal{F}^{triv}_{\epsilon,w} : C_0^\infty(\mathbb{R}\times S^1, u^*TM \oplus u^*T^*M) \circlearrowleft \\ \begin{pmatrix} \xi \\ \eta \end{pmatrix} \mapsto \begin{pmatrix} \mathcal{T}(0,1) & 0 \\ 0 & \mathcal{T}^*(0,1) \end{pmatrix} \circ \mathcal{F}_\epsilon \begin{pmatrix} exp_u \xi \\ \mathcal{T}^*(1,0)\,(w+\eta) \end{pmatrix} \end{gathered}} \tag{136}$$

where $\mathcal{T}(1,0)$ respectively $\mathcal{T}^*(1,0)$ denotes parallel transport of covector respectively vector fields along the geodesic $\gamma_\xi : [0,1] \to M$, $\tau \mapsto exp_{u(s,t)}\tau\xi(s,t)$. Note that, strictly speaking, $\mathcal{F}^{triv}_{\epsilon,u,v}$ is defined only for $(\xi,\eta) \in C_0^\infty(\mathbb{R}\times S^1, u^*TM \oplus u^*T^*M)$ such that $\sup_{(s,t)\in\mathbb{R}\times S^1} |\xi(s,t)| \le \iota$, where $\iota > 0$ denotes the injectivity radius of M.

Theorem A.3.1.

$$\boxed{d\mathcal{F}^{triv}_{\epsilon,w}(0,0)\begin{pmatrix} \xi \\ \eta \end{pmatrix} = \begin{pmatrix} \nabla_s\xi - g^{-1}\nabla_t^*\eta - R(\xi,\partial_t u)g^{-1}w - \nabla_\xi\nabla V(t,u) \\ \nabla_s^*\eta + gR(\xi,\partial_s u)g^{-1}w + \epsilon^{-2}g\nabla_t\xi - \epsilon^{-2}\eta \end{pmatrix}}$$

This means that for any smooth cylinder w the linearization of $\mathcal{F}^{triv}_{\epsilon,w}$ at 0 coincides formally with the linearization $\mathcal{D}^\epsilon_{w_0}$ of $\mathcal{F}_\epsilon$ at a zero $w_0 \in \mathcal{F}_\epsilon^{-1}(0)$. Prior to proving the theorem we are going to introduce the concept of two-parameter maps.

Two-parameter maps (cf. [**O'N**] Ch.4)

Let $D \subset \mathbb{R}^2$ be an open set such that horizontal or vertical lines intersect D in intervals (or not at all). A *two-parameter map* is a smooth map

$$f : D \to M, \tag{137}$$

the *s-parameter curve* $t = t_0$ of f is $s \mapsto f(s,t_0)$,

the *t-parameter curve* $s = s_0$ of f is $t \mapsto f(s_0,t)$.

The partial derivatives

$$\partial_s f(s,t) = df(s,t)\,\partial_s \ , \ \partial_t f(s,t) = df(s,t)\,\partial_t$$

are *vector fields on* f, i.e. $\tau_M \circ \partial_s f = f$, where $\tau_M : TM \to M$ denotes the tangent bundle. If f lies in the domain of a coordinate system $(u^1, \dots, u^n)$, i.e. $f = (f^1, \dots, f^n)$

$$\partial_s f(s,t) = \frac{\partial f^i}{\partial s}(s,t)\, \partial_i \ , \ \partial_t f(s,t) = \frac{\partial f^i}{\partial t}(s,t)\, \partial_i \ .$$

If Z is a smooth vector field on f, we denote by

$$\frac{D}{ds} Z = \nabla_s Z = \nabla_{\partial_s f} Z = \left(\partial_s Z^k + \Gamma^k_{ij}(f) \partial_s f^i Z^j\right) \partial_k \tag{138}$$

the *(partial) covariant derivative of* Z *along* s*-parameter curves.* $\frac{D}{dt} Z$ is defined analoguously.

LEMMA A.3.2. ([**O'N**] Prop. 4.44) (1) *If* f *is a two-parameter map into a (semi-) riemannian manifold* M *equipped with the Levi-Civita connection* ∇, *then* $\nabla_s \partial_t f = \nabla_t \partial_s f$.
(2) *If* Z *is a vector field on* f, *then* $\nabla_s \nabla_t Z - \nabla_t \nabla_s Z = R(\partial_s f, \partial_t f) Z$.
(3) *If* η *is a covector field on* f, *then* $\nabla^*_s \nabla^*_t \eta - \nabla^*_t \nabla^*_s \eta = g \circ R(\partial_s f, \partial_t f) g^{-1} \eta$.

PROOF. ad (1): In local coordinates we have

$$\nabla_s \partial_t f = \left(\partial_s \partial_t f^k + \Gamma^k_{ij}(f) \partial_s f^i \partial_t f^j\right) \partial_k$$
$$\nabla_t \partial_s f = \left(\partial_t \partial_s f^k + \Gamma^k_{ij}(f) \partial_t f^i \partial_s f^j\right) \partial_k$$

The result follows by commuting the partial derivatives in the first term and the symmetry of Γ^k_{ij} in the lower indices.
ad (2):

$$\begin{aligned}
&(\nabla_s \nabla_t Z - \nabla_t \nabla_s Z)^k \\
&= \partial_s (\nabla_t Z)^k + \Gamma^k_{ij}(f) (\nabla_t Z)^j \partial_s f^i - \partial_t (\nabla_s Z)^k - \Gamma^k_{ij}(f) (\nabla_s Z)^j \partial_t f^i \\
&= \partial_s \left(\partial_t Z^k + \Gamma^k_{mn}(f) Z^n \partial_t f^m\right) + \Gamma^k_{ij}(f) \left(\partial_t Z^j + \Gamma^j_{mn}(f) Z^n \partial_t f^m\right) \partial_s f^i \\
&\quad - \partial_t \left(\partial_s Z^k + \Gamma^k_{mn}(f) Z^n \partial_s f^m\right) - \Gamma^k_{ij}(f) \left(\partial_s Z^j + \Gamma^j_{mn}(f) Z^n \partial_s f^m\right) \partial_t f^i \\
&= \partial_s f^i \partial_t f^j Z^m \left(\frac{\partial \Gamma^k_{jm}}{\partial u^i}(f) - \frac{\partial \Gamma^k_{im}}{\partial u^j}(f) + \Gamma^k_{i\nu}(f) \Gamma^\nu_{jm}(f) - \Gamma^k_{j\nu}(f) \Gamma^\nu_{im}(f)\right) \\
&= \partial_s f^i \partial_t f^j Z^m R^k_{mij} = (R(\partial_s f, \partial_t f) Z)^k \ .
\end{aligned}$$

We got the first two equalities by expressing the (partial) covariant derivative in local coordinates as in (138). The third equality follows by carrying out the partial derivatives with respect to s and t, using the product and chain rule; note that Γ^k_{ij} depends on $f(s,t)$. Now use the local expression (98) for the curvature tensor.

ad (3): Define $\xi = g^{-1}(f)\eta$ and apply (2) to get

$$\begin{aligned} g^{-1}(\nabla_s^*\nabla_t^*\eta - \nabla_s^*\nabla_t^*\eta) &= \nabla_s\nabla_t\xi - \nabla_s\nabla_t\xi \\ &= R(\partial_s f, \partial_t f)\xi \\ &= R(\partial_s f, \partial_t f)\, g^{-1}\eta. \end{aligned}$$

To pull g^{-1} through the covariant derivatives we used Lemma A.2.1. Now apply g from the left. □

PROOF. (**of Theorem A.3.1**) Consider the two-parameter maps given by $f(s,\tau) = exp_{u(s)}\tau\xi(s)$ (here we assume $t = const$) and $h(t,\tau) = exp_{u(t)}\tau\xi(t)$ (here we assume $s = const$). As $dexp_u(0) = id$ it follows $\partial_\tau f(s,0) = \xi(s)$ and $\partial_\tau h(t,0) = \xi(t)$. Fixing s and t we define the geodesic $\gamma_\xi(\tau) = exp_u\tau\xi$; note that $\gamma_\xi(0) = u$ and $\partial_\tau\gamma_\xi(0) = \xi$. $\mathcal{T}_\xi^*(1,0)$ respectively $\mathcal{T}_\xi(1,0)$ denotes parallel transport of (co)vector fields along γ_ξ from $\gamma_\xi(0) = u$ to $\gamma_\xi(1) = exp_u\xi$. We observe that

$$\begin{aligned} & d\mathcal{F}_{\epsilon,w}^{triv}(0,0)\begin{pmatrix}\xi\\ \eta\end{pmatrix} \\ &= \frac{d}{d\tau}\Big|_{\tau=0} \mathcal{F}_{\epsilon,w}^{triv}\begin{pmatrix}\tau\xi\\ \tau\eta\end{pmatrix} \\ &= \frac{d}{d\tau}\Big|_{\tau=0} \begin{pmatrix}\mathcal{T}_{\tau\xi}(0,1) & 0\\ 0 & \mathcal{T}_{\tau\xi}^*(0,1)\end{pmatrix} \mathcal{F}_\epsilon\begin{pmatrix}exp_u\tau\xi\\ \mathcal{T}_{\tau\xi}^*(1,0)\,(w+\tau\eta)\end{pmatrix} \\ &= \frac{d}{d\tau}\Big|_{\tau=0} \begin{pmatrix}\mathcal{T}_{\xi}(0,\tau) & 0\\ 0 & \mathcal{T}_{\xi}^*(0,\tau)\end{pmatrix} \mathcal{F}_\epsilon\begin{pmatrix}exp_u\tau\xi\\ \mathcal{T}_{\xi}^*(\tau,0)\,(w+\tau\eta)\end{pmatrix} \end{aligned}$$

The last equality follows because parallel transport along a curve does not depend on the parametrization of the curve, i.e. $\mathcal{T}_{\tau\xi}(0,1) = \mathcal{T}_\xi(0,\tau)$ and similarly for $\mathcal{T}^*$.

1^st term: $Z(s,\tau) = \partial_s(exp_{u(s)}\tau\xi(s))$ is a vector field on f. Using proposition A.1.5 in the first and lemma A.3.2 (1) in the third equality we get

$$\begin{aligned} \tfrac{d}{d\tau}\big|_{\tau=0}\,\mathcal{T}(0,\tau)\,\partial_s(exp_u\tau\xi) &= (\nabla_\tau Z)|_{\tau=0} \\ &= (\nabla_\tau\partial_s f)|_{\tau=0} \\ &= (\nabla_s\partial_\tau f)|_{\tau=0} \\ &= \nabla_s\xi. \end{aligned}$$

2^nd term: $Z(t,\tau) = -g^{-1}(exp_u\tau\xi)\,\nabla_t^*\left(\mathcal{T}^*(\tau,0)\,(w+\tau\eta)\right)$ is a vector field and $\theta(t,\tau) = \mathcal{T}^*(\tau,0)\,(w+\tau\eta)$ is a covector field on h. Using proposition A.1.5 in the first, Lemma A.2.1 in the third and Lemma A.3.2 (3) in the fourth equality we get

$$\begin{aligned} -\tfrac{d}{d\tau}\Big|_{\tau=0}\mathcal{T}(0,\tau)\; g^{-1}(exp_u\tau\xi)\;\nabla^*_t\Big(\mathcal{T}^*(\tau,0)\;(w+\tau\eta)\Big) &= (\nabla_\tau Z)|_{\tau=0}\\ &= -\nabla_\tau\Big(g^{-1}(h)\;\nabla^*_t\big(\mathcal{T}^*(\tau,0)\;(w+\tau\eta)\big)\Big)\Big|_{\tau=0}\\ &= -g^{-1}(u)\;\;\nabla^*_\tau\nabla^*_t\theta|_{\tau=0}\\ &= -g^{-1}\;\big(\nabla^*_t\nabla^*_\tau\theta + g\,R(\partial_\tau h,\partial_t h)\;g^{-1}\theta\big)|_{\tau=0}\\ &= -g^{-1}\nabla^*_t\eta - R(\xi,\partial_t u)\;g^{-1}w. \end{aligned}$$

To get the last equality we employed Lemma A.1.10.

3rd term: $\nabla V(\gamma(\tau)) = g^{kl}(\tau)\,\frac{\partial V}{\partial u^l}(\gamma(\tau))\,\partial_k$ is a vector field along γ. Using proposition A.1.5 in the first equality we get

$$\begin{aligned} \tfrac{d}{d\tau}\Big|_{\tau=0}\mathcal{T}(0,\tau)(-1)\nabla V(exp_u\tau\xi) &= -\nabla_\tau\nabla V(\gamma(\tau))|_{\tau=0}\\ &= -\tfrac{\partial}{\partial\tau}\Big|_{\tau=0}(\nabla V)^k(\gamma(\tau))\;\partial_k - \Gamma^k_{ij}(\gamma(0))\;\partial_\tau\gamma^i(0)\;(\nabla V)^j(0)\;\partial_k\\ &= -\left(\frac{\partial(\nabla V)^k}{\partial u^j}(0)\;\partial_\tau\gamma^j(0) + \Gamma^k_{ij}(u)\;\xi^i\;(\nabla V)^j(0)\right)\partial_k\\ &= -\nabla_\xi\nabla V. \end{aligned}$$

4th term: $\theta_1(s,\tau) = \nabla^*_s\left(\mathcal{T}^*(\tau,0)\;(w+\tau\eta)\right)$ and $\theta_2(s,\tau) = \mathcal{T}^*(\tau,0)\,(w+\tau\eta)$ are covector fields on h. Using proposition A.1.9 in the first, Lemma A.3.2 (3) in the third and Lemma A.1.10 in the fourth equality we get

$$\begin{aligned} \tfrac{d}{d\tau}\Big|_{\tau=0}\mathcal{T}^*(0,\tau)\;\nabla^*_s\Big(\mathcal{T}^*(\tau,0)\;(w+\tau\eta)\Big) &= (\nabla^*_\tau\theta_1)|_{\tau=0}\\ &= (\nabla^*_\tau\nabla^*_s\theta_2)|_{\tau=0}\\ &= \big(\nabla^*_s\nabla^*_\tau\theta_2 + g\,R(\partial_\tau h,\partial_s h)\;g^{-1}\theta_2\big)\big|_{\tau=0}\\ &= \nabla^*_s\eta + g\,R(\xi,\partial_s u)\;g^{-1}w. \end{aligned}$$

5th term: $\theta(t,\tau) = g(exp_{u(t)}\tau\xi(t))\;\partial_t(exp_{u(t)}\tau\xi(t))$ is a covector field on h. Using proposition A.1.9 in the first, Lemma A.2.1 in the third and Lemma A.3.2 (1) in the fourth equality we get

$$\begin{aligned} \tfrac{d}{d\tau}\Big|_{\tau=0}\mathcal{T}^*(0,\tau)\;\epsilon^{-2}g(exp_u\tau\xi)\;\partial_t(exp_u\tau\xi) &= \epsilon^{-2}(\nabla^*_\tau\theta)\big|_{\tau=0}\\ &= \epsilon^{-2}\nabla^*_\tau\big(g(h)\;\partial_t h\big)\big|_{\tau=0} = \epsilon^{-2}g(u)\;(\nabla_\tau\partial_t h)\big|_{\tau=0}\\ &= \epsilon^{-2}g(u)\;(\nabla_t\partial_\tau h)\big|_{\tau=0} = \epsilon^{-2}g(u)\;\nabla_t\xi. \end{aligned}$$

6th term: $\tfrac{d}{d\tau}\Big|_{\tau=0}\mathcal{T}^*(0,\tau)\;(-1)\mathcal{T}^*(\tau,0)\epsilon^{-2}(w+\tau\eta) = -\epsilon^{-2}\eta.$ □

A.4. The linear operators represented in frames

Our first claim is to represent the linear operator

$$\begin{aligned} \mathcal{D}_u \;:\; C_0^\infty(\mathbb{R}\times S^1, u^*TM) &\to C_0^\infty(\mathbb{R}\times S^1, u^*TM) \\ \xi &\mapsto \nabla_s\xi - \nabla_t\nabla_t\xi - R(\xi,\dot u)\dot u - \nabla_\xi\nabla V(t,u) \end{aligned} \tag{139}$$

where $u \in C^\infty(\mathbb{R}\times S^1, M)$ with $u \mapsto x^\mp \in Crit\,\mathcal{I}_V$ for $s \to \mp\infty$ uniformly in t, as an operator acting on $\mathbb{R}^n$-valued functions

$$\mathcal{D}_0 \;:\; C_0^\infty(\mathbb{R}\times S^1, \mathbb{R}^n) \to C_0^\infty(\mathbb{R}\times S^1, \mathbb{R}^n) \quad , \quad \vec\xi \mapsto \mathcal{D}_0\vec\xi.$$

If the Riemannian manifold (M,g) is orientable, then there exists an *orthogonal trivialization*

$$\begin{aligned} \phi \;:\; (\mathbb{R}\times S^1)\times\mathbb{R}^n &\to u^*TM \\ (s,t,\vec\xi) &\mapsto (s,t;\phi(s,t)\vec\xi). \end{aligned} \tag{140}$$

Orthogonality here means that $\phi(s,t)^*g = \langle\cdot,\cdot\rangle$. Let $\{e_1,\ldots,e_n\}$ be the standard orthonormal basis of $\mathbb{R}^n$ and define

$$Z_i(s,t) = \phi(s,t)e_i \quad , \; i = 1,\ldots,n$$

then $\{Z_1(s,t),\ldots,Z_n(s,t)\}$ is an orthonormal basis of $T_{u(s,t)}M$. In the nonorientable case we construct an orthogonal trivialization over $[0,1]$ with boundary condition as discussed briefly at the end of subsection B.1.8 in appendix B.

REMARK A.4.1. (**Existence of ϕ**) We first construct ϕ for a fixed value s_0 of s and then extend it to $s \in \mathbb{R}$ via parallel transport of the $Z_i(s_0,t)$ along curves $s \mapsto u(s,t)$ or by the same argument as in the proof of lemma B.1.13. Actually we prefer the parallel transport method in order to get rid of terms $\nabla_s Z_i$ in later computations. Cover S^1 by finitely many intervals $\{I_i\}_{i=1}^N$ over which orthogonal trivializations $\phi_i(t) : \mathbb{R}^n \to T_{u(s_0,t)}M$ exist. On $I_i \cap I_j = (t_i,t_j)$ we patch ϕ and ϕ_j as follows: choose any smooth map $\psi_{ij} : \mathbb{R} \to SO(n,\mathbb{R})$ such that

$$\psi_{ij}(t) = \begin{cases} \mathbb{1} & , t \text{ near } t_i \\ \phi_i^{-1}(t_j)\circ\phi_j(t_j) & , t \text{ near } t_j \end{cases}$$

then define for $t \in I_i \cap I_j$

$$\phi_{ji}(t) = \begin{cases} \phi_i(t) & , t \le t_i \text{ and } t \in I_i \\ \phi_i(t)\circ\psi_{ij}(t) & , t \in (t_i,t_j) \\ \phi_j(t) & , t \ge t_j \text{ and } t \in I_j. \end{cases}$$

Note that this construction works as the orientability of M allows us to reduce the structure group of the riemannian vector bundle from $O(n,\mathbb{R})$ to $SO(n,\mathbb{R})$, which is connected.

With respect to the orthogonal trivialization ϕ of u^*TM (parallel with respect to s) the covariant derivatives ∇_s and ∇_t are represented by

$$\vec{\nabla}_s\vec{\xi} = \phi^{-1}\nabla_s(\phi\vec{\xi}) = \partial_s\vec{\xi} \quad , \quad \vec{\nabla}_t\vec{\xi} = \phi^{-1}\nabla_t(\phi\vec{\xi}) = \partial_t\vec{\xi} + A\vec{\xi} \quad , \quad \vec{\xi} = \xi^i e_i$$

where the *connection potential* $A \in C^\infty(\mathbb{R} \times S^1, \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n))$ is defined by

$$A(s,t)e_i = \phi^{-1}(s,t)\nabla_t\phi(s,t)e_i. \tag{141}$$

Note that $A(s,t)^*e^i = -A(s,t)e_i$, which follows from the formula for intrinsic partial integration in proposition B.2.2. Moreover, it turns out that

$$[\partial_s, \vec{\nabla}_t] = \partial_s A = \phi^{-1}R(\partial_s u, \partial_t u)\phi. \tag{142}$$

Indeed

$$[\partial_s, \vec{\nabla}_t]\vec{\xi} = \partial_s\partial_t\vec{\xi} + (\partial_s A)\vec{\xi} + A\partial_s\vec{\xi} - \partial_t\partial_s\vec{\xi} - A\partial_s\vec{\xi} = (\partial_s A)\vec{\xi}$$

and using $[\partial_s u, \partial_t u] = 0$ (cf. lemma A.3.2) we get

$$\begin{aligned}\phi^{-1}R(\partial_s u, \partial_t u)\phi &= \phi^{-1}\nabla_s\phi\phi^{-1}\nabla_t\phi - \phi^{-1}\nabla_t\phi\phi^{-1}\nabla_s\phi \\ &= \partial_s(\partial_t + A) - (\partial_t + A)\partial_s = \partial_s A.\end{aligned}$$

LEMMA A.4.2. *$\mathcal{D}_u$ as in (139) is represented by*

$$\boxed{\mathcal{D}_0\vec{\xi} = \phi^{-1}\mathcal{D}_u(\phi\vec{\xi}) = \partial_s\vec{\xi} - \vec{\nabla}_t\vec{\nabla}_t\vec{\xi} - Q\vec{\xi}},$$

where $Q \in C^\infty(\mathbb{R} \times S^1, \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n))$ with $Q(s,t)^* = Q(s,t)$ is given by

$$Qe_i = \phi^{-1}R(Z_i, \dot{u})\dot{u} - \phi^{-1}\nabla_{Z_i}\nabla V(t,u).$$

PROOF.

$$\begin{aligned}\mathcal{D}_0\vec{\xi} &= \phi^{-1}\mathcal{D}_u(\phi\vec{\xi}) \\ &= \phi^{-1}\left(\nabla_s(\xi^i Z_i) - \nabla_t\nabla_t(\xi^i Z_i) - R(\xi^i Z_i, \dot{u})\dot{u} - \nabla_{\xi^i Z_i}\nabla V(t,u)\right) \\ &= \phi^{-1}\big((\partial_s\xi^i)Z_i + \xi^i\nabla_s Z_i - (\partial_t\partial_t\xi^i)Z_i - 2(\partial_t\xi^i)\nabla_t Z_i - \xi^i\nabla_t\nabla_t Z_i \\ &\qquad - \xi^i R(Z_i, \dot{u})\dot{u} - \xi^i\nabla_{Z_i}\nabla V(t,u)\big) \\ &= \partial_s\vec{\xi} - \vec{\nabla}_t\vec{\nabla}_t\vec{\xi} - \xi^i\phi^{-1}\big(R(Z_i, \dot{u})\dot{u} - \nabla_{Z_i}\nabla V(t,u)\big)\end{aligned}$$

where we used several times equation (141). The symmetry of the first summand of Q may be seen using the antisymmetry properties of the curvature tensor

$$\begin{aligned}g(R(Z_i, \dot{u})\dot{u}, Z_k) &\stackrel{(99)}{=} g(R(\dot{u}, Z_k)Z_i, \dot{u}) \\ &= g(R(Z_k, \dot{u})\dot{u}, Z_i).\end{aligned}$$

For the second summand we exploit that the Levi-Civita connection is torsionfree as well as its compatibility with the metric

$$\begin{aligned} g(\nabla_{Z_i}\nabla V, Z_k) &=^{(94)} Z_i g(\nabla V, Z_k) - g(\nabla V, \nabla_{Z_i} Z_k) \\ &= Z_i Z_k V - (\nabla_{Z_i} Z_k) V \\ &=^{(93)} Z_k Z_i V - (\nabla_{Z_k} Z_i) V \\ &= g(\nabla_{Z_k}\nabla V, Z_i). \end{aligned}$$

□

As usual setting $u = \tau_M^* w$ our second claim will be to show that the operator

$$\begin{aligned} \mathcal{D}_w : C_0^\infty(\mathbb{R}\times S^1, u^*TM \oplus u^*T^*M) &\to C_0^\infty(\mathbb{R}\times S^1, u^*TM \oplus u^*T^*M) \\ \begin{pmatrix}\xi\\ \eta\end{pmatrix} &\mapsto \begin{pmatrix}\nabla_s\xi - g^{-1}\nabla_t\eta - R(\xi,\dot u)g^{-1}w - \nabla_\xi\nabla V(t,u)\\ \nabla_s\eta + gR(\xi,\partial_s u)g^{-1}w + \epsilon^{-2}(g\nabla_t\xi - \eta),\end{pmatrix} \end{aligned}$$

where $w \in C^\infty(\mathbb{R}\times S^1, T^*M)$ with $w \to g(x^\mp)\partial_t x^\mp \in Crit\,\mathcal{A}_V$ for $s \to \mp\infty$ uniformly in t, may be represented by an operator on $\mathbb{R}^{2n}$-valued functions

$$\boxed{\begin{aligned} \mathcal{D}_\epsilon : C_0^\infty(\mathbb{R}\times S^1,\mathbb{R}^{2n}) &\to C_0^\infty(\mathbb{R}\times S^1,\mathbb{R}^{2n}) \\ \begin{pmatrix}\vec\xi\\ \vec\eta\end{pmatrix} &\mapsto \begin{pmatrix}\partial_s\vec\xi - \vec\nabla_t\vec\eta\\ \partial_s\vec\eta + \epsilon^{-2}(\vec\nabla_t\vec\xi - \vec\eta)\end{pmatrix} - \begin{pmatrix}C\vec\xi\\ -B\vec\xi\end{pmatrix}\end{aligned}} \tag{143}$$

where

$$Ce_i = \phi^{-1}R(Z_i,\dot u)g^{-1}w + \phi^{-1}\nabla_{Z_i}\nabla V(t,u)$$

is asymptotically symmetric as $C(s,t) \to S^\mp(t)$ for $s \to \mp\infty$ and

$$Be_i = (\phi^*)gR(Z_i,\partial_s u)g^{-1}w \to 0 \quad \text{for } s \to \mp\infty.$$

Note that for $w = g(u)\partial_t u$

$$B^*e_j = (\phi^*)gR(Z_j,\partial_t u)\partial_s u.$$

To derive (143) we pick an orthogonal trivialization ϕ of u^*TM as before, then we define

$$\Phi = \begin{pmatrix}\phi & 0\\ 0 & \phi^{*-1}\end{pmatrix} : (\mathbb{R}\times S^1)\times(\mathbb{R}^n\times\mathbb{R}^n) \to u^*TM \oplus u^*T^*M. \tag{144}$$

Φ is a unitary trivialization with

$$\Phi(s,t)\begin{pmatrix}e_i\\ e^j\end{pmatrix} = \begin{pmatrix}Z_i(s,t)\\ Z^j(s,t)\end{pmatrix}$$

where e^j is the dual of e_j and $Z^j(s,t)$ the dual of $Z_j(s,t)$ under the natural identification of the vector space with its dual space via the metric $\langle\cdot,\cdot\rangle$

respectively g. We define

$$\begin{aligned}
&\mathcal{D}_\epsilon \begin{pmatrix} \vec{\xi} \\ \vec{\eta} \end{pmatrix} = \Phi^{-1}\mathcal{D}_w \begin{pmatrix} \phi\,\vec{\xi} \\ \phi^{*-1}\vec{\eta} \end{pmatrix} = \Phi^{-1}\mathcal{D}_w \begin{pmatrix} \xi^i Z_i \\ \eta_j Z^j \end{pmatrix} \\
&= \begin{pmatrix} \phi^{-1}\big(\nabla_s(\xi^i Z_i) - g^{-1}\nabla_t(\eta_j Z^j) - R(\xi^i Z_i, \dot{u})g^{-1}w - \nabla_{\xi^i Z_i}\nabla V(t,u)\big) \\ \phi^*\big(\nabla_s(\eta_j Z^j) + \epsilon^{-2} g\nabla_t(\xi^i Z_i) + R(\xi^i Z_i, \partial_s u)g^{-1}w - \epsilon^{-2}\eta_j Z^j\big) \end{pmatrix} \\
&= \begin{pmatrix} \partial_s\vec{\xi} - \vec{\nabla}_t\vec{\eta} - \xi^i\phi^{-1}\big(R(Z_i,\dot{u})g^{-1}w + \nabla_{Z_i}\nabla V(t,u)\big) \\ \partial_s\vec{\eta} + \epsilon^{-2}(\vec{\nabla}_t\vec{\xi} - \vec{\eta}) + \xi^i\phi^* g R(Z_i, \partial_s u) g^{-1} w \end{pmatrix}
\end{aligned}$$

where we used $(\phi^*)g\phi = \mathbb{1}$ and our frames being parallel with respect to s. Note that in the special case $w = g(u)\partial_t u$ we get

$$\boxed{\mathcal{D}_\epsilon \begin{pmatrix} \vec{\xi} \\ \vec{\eta} \end{pmatrix} = \begin{pmatrix} \partial_s\vec{\xi} - \vec{\nabla}_t\vec{\eta} \\ \partial_s\vec{\eta} + \epsilon^{-2}(\vec{\nabla}_t\vec{\xi} - \vec{\eta}) \end{pmatrix} - \begin{pmatrix} Q\vec{\xi} \\ -B\vec{\xi} \end{pmatrix}}.$$

APPENDIX B

Two variational problems – basic facts

We recall standard facts from the variational theories of the symplectic action functional in section B.1 and the classical action functional in section B.2 (usually called energy functional in Riemannian geometry). Their variational formulae are derived and we discuss the assignment of integers (indices) to their critical points. In contrast to the canonically defined Morse index of a critical point of the classical action, the Conley-Zehnder index of a critical point of the symplectic action usually involves noncanonical choices (of unitary trivializations) in its construction. These are related to the nontriviality of the first Chern class. However, here the first Chern class of the restriction of the tangent bundle $TT^*M \to T^*M$ to any closed submanifold of T^*M vanishes and as a consequence the Conley-Zehnder index can be constructed canonically. Moreover, we give an alternative construction in this context – avoiding first Chern classes – by exploiting the existence of a global Lagrangian splitting of TT^*M. Throughout we will use the following terminology for projections: For any manifold N let τ_N denote its tangent bundle projection and τ_N^* its cotangent bundle projection.

B.1. The symplectic action functional

Let (M^n, g) be a *closed* (i.e. compact and without boundary), smooth Riemannian manifold of dimension n. After discussing the construction of natural coordinates on TT^*M in B.1.1 we introduce certain canonical structures on the $2n$-dimensional manifold T^*M. The Levi-Civita connection of (M, g), which we view in B.1.2 as a bundle morphism $K : TT^*M \to T^*M$, called *connection map* defines a *horizontal subbundle* T^hT^*M; the kernel of the linearized cotangent projection $T\tau_M^*$ – by T we denote the tangent map – defines the *vertical subbundle* T^vT^*M. So we have a natural splitting $TT^*M \cong Ker\ K \oplus Ker\ T\tau_M^* =: T^hT^*M \oplus T^vT^*M$. These subbundles may be identified via $T\tau_M^*|_{T^hT^*M}$ with TM and via $K|_{T^vT^*M}$ with T^*M, respectively; the latter isomorphism however depends on the choice of coordinates. On the other hand – setting $q = \tau_M^* p$ – we have a natural isomorphism between fibers $\Theta(p) = (T\tau_M^*(p)\cdot, K(p)\cdot) : T_pT^*M \to T_qM \oplus T_q^*M$, which we can view as a change of fiber coordinates. The *Liouville form* θ – in natural coordinates (q^i, p_j) given by $p_i dq^i$ – and the *canonical symplectic structure* $\Omega = -d\theta$ *on* T^*M are introduced in B.1.3. In B.1.4 we introduce a Riemannian metric G on T^*M, whose pullback under $\theta(p)^{-1}$ is given by the product metric $g \oplus g^*$. Moreover, the metric g on M leads canonically to an almost complex structure $J \in \Gamma(End\, TT^*M)$, i.e. $J^2 = -id$. (G, J, Ω) are compatible in the sense that $G(\cdot,\cdot) = \Omega(\cdot, J\cdot)$. As we will see in B.1.5 a Hamiltonian function $H : \mathbb{R}/\mathbb{Z} \times T^*M \to \mathbb{R}$ gives rise to the Hamiltonian vector field X_H defined by the identity $dH(\cdot) = \Omega(X_H, \cdot)$. It turns out in B.1.6 that the 1-periodic Hamiltonian orbits $\mathcal{P}er\,(H)$ of X_H are exactly the critical points of the symplectic action functional on the free loop space of T^*M

$$\begin{aligned} \mathcal{A}_V \ : \ C^\infty(S^1, T^*M) = C^\infty(S^1, T^*M) \quad &\to \quad \mathbb{R} \\ z \quad &\mapsto \quad \int_{S^1} z^*\Theta - \int_0^1 H(t, z(t))\, dt. \end{aligned}$$

Assuming that the set $\mathcal{P}er\,(H) = Crit\,\mathcal{A}_V$ is discrete we are going to define a map

$$\mu_{CZ} \ : \ Crit\,\mathcal{A}_V \ \to \mathbb{Z}$$

called *Conley-Zehnder index*. The naturality of this index will be discussed in great detail (cf. introduction to this appendix).

B.1.1. Natural coordinates on TT^*M. First we introduce local coordinates $(u^1, \ldots, u^n)$, short notation (u^i), on M and compute the transformation formulae under a change of coordinates of the coordinate vector and covector fields, the coordinate functions of vectors and covectors, the matrix-valued function representing the metric and the Christoffel symbols of the Levi-Civita connection associated to g.

Next we introduce natural local coordinates (u^i, v_j) for $p \in T^*M$ by applying the cotangent functor T^* to the charts of M. The $(v_1, \ldots, v_n)$ indeed transform like coordinate functions of a covector and may therefore be interpreted as the fibre part, $(u_1, \ldots, u_n)$ representing the corresponding base point $q \in M \subset T^*M$. This procedure however fails in the next iteration step, namely applying the tangent functor T to the natural charts of T^*M to get local coordinates (u^i, v_j, ξ^k, y_l) of TT^*M: the ξ^k transform as coordinate functions of a vector, but the y_l do not transform as coordinate functions of a covector. In fact their transformation formula under a change of coordinates involves the ξ^k as well as the v_j. There are also second derivatives of the transition maps involved, just as in the case of the transition maps for the Christoffel symbols. This leads us to the definition

$$\eta_l = y_l - \Gamma^j_{lk}(u)\xi^k v_j \tag{145}$$

and we will show that η_l indeed transforms like a coordinate function of a covector. In fact we have a natural isomorphism of vector spaces

$$\theta(p) : T_pT^*M \to T_qM \oplus T^*_qM \quad , \quad q = \tau^*_M p \tag{146}$$

given locally by

$$(u^i, v_j; \xi^k, y_l) \mapsto (u^i; \xi^k, y_l - \Gamma^j_{lk}(u)\xi^k v_j)\ . \tag{147}$$

Note the implicit occurence of the metric via the Christoffel symbols. If $p(t)$ denotes a smooth path in T^*M and $q(t) = \tau_M^* p(t)$, then this isomorphism takes $\partial_t p(t) \in T_{p(t)}T^*M$ to the element $(\partial_t q(t), \nabla_t^* p(t)) \in T_qM \oplus T_q^*M$, where ∇^* denotes the Levi-Civita connection on M acting on covector fields.

Let $(U, \phi), (U, \psi)$ be two coordinate charts of the manifold M^n and denote by $(u^i), (\tilde{u}^i)$ the corresponding coordinate functions

$$\begin{aligned} \phi : U &\to \mathbb{R}^n \ , & \psi : U &\to \mathbb{R}^n \ , \\ q &\mapsto (u^1, \ldots, u^n) \ , & q &\mapsto (\tilde{u}^1, \ldots, \tilde{u}^n) \ . \end{aligned} \tag{148}$$

$f(\tilde{u}) = \phi \circ \psi^{-1}(\tilde{u}) = u$ denotes the transition map of the coordinates. Associated to $(u^i), (\tilde{u}^i)$ are the coordinate vector and covector fields $\{\partial_{u^i}\}, \{\partial_{\tilde{u}^i}\}$ and $\{du^i\}, \{d\tilde{u}^i\}$. They transform as follows (throughout this text we use Einstein's summation convention)

LEMMA B.1.1. *For $i, k = 1, \ldots, n$ we have*

$$\begin{aligned} i)\ \partial_{u^i} &= \frac{\partial f^{-1 j}(u)}{\partial u^i} \partial_{\tilde{u}^j} \ , \ \partial_{\tilde{u}^k} = \frac{\partial f^j(\tilde{u})}{\partial \tilde{u}^k} \partial_{u^j} \ , \\ ii)\ du^i &= \frac{\partial f^i(\tilde{u})}{\partial \tilde{u}^j} d\tilde{u}^j \ , \ d\tilde{u}^k = \frac{\partial f^{-1 k}(u)}{\partial u^j} du^j \ . \end{aligned}$$

PROOF. Part two may be proven by the chain rule: As $u^i = f^i(\tilde{u})$ we get $du^i = \frac{\partial f^i(\tilde{u})}{\partial \tilde{u}^j} d\tilde{u}^j$. The proof of part one uses this fact as follows: $\{\partial_{u^i}\}, \{\partial_{\tilde{u}^j}\}$ are bases of T_qM, hence they are related by a linear transformation $\partial_{u^k} = A_k^l \partial_{\tilde{u}^l}$. Now

$$\delta_k^i = du^i(\partial_{u^k}) = \left(\frac{\partial f^i(\tilde{u})}{\partial \tilde{u}^j} d\tilde{u}^j \right) (A_k^l \partial_{\tilde{u}^l}) = \frac{\partial f^i(\tilde{u})}{\partial \tilde{u}^j} A_k^j$$

and therefore $A_k^j = \frac{\partial f^{-1 j}(\tilde{u})}{\partial \tilde{u}^k}$. □

Any element $\xi \in T_qM$ respectively $\eta \in T_q^*M$ may be written as $\xi = \xi^i(u)\partial_{u^i} = \tilde{\xi}^j(\tilde{u})\partial_{\tilde{u}^j}$ respectively $\eta = \eta_i(u)du^i = \tilde{\eta}_j(\tilde{u})d\tilde{u}^j$.

LEMMA B.1.2. *For $i, k = 1, \ldots, n$ we have*

$$\begin{aligned} i)\ \xi^i &= \frac{\partial f^i(\tilde{u})}{\partial \tilde{u}^j} \tilde{\xi}^j \ , \ \tilde{\xi}^k = \frac{\partial f^{-1 k}(u)}{\partial u^j} \xi^j \ , \\ ii)\ \xi_i &= \frac{\partial f^{-1 j}(u)}{\partial u^i} \tilde{\xi}_j \ , \ \tilde{\xi}_k = \frac{\partial f^j(\tilde{u})}{\partial \tilde{u}^k} \xi_j \ . \end{aligned}$$

PROOF.

$$\text{ad } i)\ \xi^i \partial_{u^i} = \xi = \tilde{\xi}^j \partial_{\tilde{u}^j} \overset{\substack{\text{Lemma} \\ B.1.1i)}}{=} \tilde{\xi}^j \frac{\partial f^k(\tilde{u})}{\partial \tilde{u}^j} \partial_{u^k} \ .$$

$$\text{ad } ii)\ \xi_i du^i = \xi = \tilde{\xi}_j d\tilde{u}^j \overset{\substack{\text{Lemma} \\ B.1.1ii)}}{=} \tilde{\xi}_j \frac{\partial f^{-1 j}(u)}{\partial u^k} du^k \ .$$

□

Let $g_{ij}(u) = g_{\phi^{-1}(u)}(\partial_{u^i}, \partial_{u^j})$, $i, j = 1, \dots, n$, be the matrix-valued function on $\phi(U)$ representing the metric g. By $\Gamma^k_{ij}(u)$ we denote the Christoffel symbols of the Levi-Civita connection associated to g. g^* denotes the dual metric. Its local representative $g^{kl}(u)$ is related to $g_{ij}(u)$ by $g^{kl}(u)g_{li}(u) = \delta^k_i$, i.e. it is the inverse matrix of (g_{ij}).

LEMMA B.1.3. *For $i, j, k = 1, \dots, n$ we have*

$$
\begin{aligned}
i)\ g_{ij}(u) &= \frac{\partial {f^{-1}}^k(u)}{\partial u^i}\frac{\partial {f^{-1}}^l(u)}{\partial u^j}\tilde g_{kl}(\tilde u)\ ,\\
g^{ij}(u) &= \frac{\partial f^i(\tilde u)}{\partial \tilde u^k}\frac{\partial f^j(\tilde u)}{\partial \tilde u^l}\tilde g^{kl}(\tilde u)\ ,\\
ii)\ \Gamma^k_{ij}(u) &= \frac{\partial f^k(\tilde u)}{\partial \tilde u^s}\left(\frac{\partial^2 {f^{-1}}^s(u)}{\partial u^i \partial u^j} + \tilde\Gamma^s_{rl}(u)\frac{\partial {f^{-1}}^l(u)}{\partial u^j}\frac{\partial {f^{-1}}^r(u)}{\partial u^i}\right)\ ,\\
\tilde\Gamma^k_{ij}(\tilde u) &= \frac{\partial {f^{-1}}^k(u)}{\partial u^s}\left(\frac{\partial^2 f^s(\tilde u)}{\partial \tilde u^i \partial \tilde u^j} + \Gamma^s_{rl}(u)\frac{\partial f^l(\tilde u)}{\partial \tilde u^j}\frac{\partial f^r(\tilde u)}{\partial \tilde u^i}\right)\ .
\end{aligned}
$$

PROOF. ad *i)* $\xi^i_1\xi^j_2 g_{ij}(u) = g(\xi_1, \xi_2) = \tilde\xi^k_1\tilde\xi^l_2\tilde g_{kl}(\tilde u)$, now use Lemma B.1.2i), similarly for the dual metric.
ad *ii)* Write the Christoffel symbol as sum of derivatives of the metric (95), then use i) and the product rule to calculate this derivatives. Replace also $g^{kl}(u)$ using i). We get a formula involving only coordinates $\tilde u$. Again using (95) it may be simplified to the form stated in ii). □

Starting with a chart (ϕ, U) of M, there is a natural way to get a chart (Φ, T^*U) of T^*M:

$$
\begin{aligned}
\Phi := T^*\phi : T^*U &\to \phi(U) \times \mathbb{R}^n \\
p &\mapsto (\phi(q), d\phi(q)^{*-1}p) \quad , \quad q := \tau^*_M p.
\end{aligned}
\tag{149}
$$

The coordinates are denoted by $(u^i, v_j) = \Phi(p)$. The v_j are exactly the coordinate functions of the fibre part of $p \in T^*M$, i.e. are components of a covector. This may be seen by investigating the transformation behavior under a change of coordinates: Let (Ψ, T^*U) be the chart coming from (ψ, U), then

$$
\begin{aligned}
(u, v) &= \Phi\circ\Psi^{-1}(\tilde u, \tilde v) = \Phi(\psi^{-1}(\tilde u), d\psi(q)^*\tilde v) \\
&= (\phi\circ\psi^{-1}(\tilde u), d\phi(q)^{*-1}\circ d\psi(q)^*\tilde v)\ .
\end{aligned}
\tag{150}
$$

Applying a similar procedure to (Φ, T^*U) we get a chart $(T\Phi, TT^*U)$ of TT^*M with coordinates $(u^1, \dots, u^n, v_1, \dots, v_n; \zeta^1, \dots, \zeta^{2n})$. Denoting the last $2n$ variables by $(\xi^1, \dots, \xi^n, y_1, \dots, y_n)$ we will see that the ξ^k transform as coordinate functions of a vector, but the y_l do not transform as coordinate functions of a covector:

$$
\begin{aligned}
T\Phi : TT^*U &\to (\phi(U) \times \mathbb{R}^n) \times T(\phi(U) \times \mathbb{R}^n) \\
\zeta &\mapsto (\phi(q), d\phi(q)^{*-1}p; d\Psi(p)\zeta)\ ,
\end{aligned}
\tag{151}
$$

where $p = \tau_{T^*M}\zeta$ and $q = \tau_M^* p$. The transformation behavior is as follows

$$
\begin{aligned}
(u, v; \xi, y) &= T\Phi(T\Psi)^{-1}(\tilde{u}, \tilde{v}; \tilde{\xi}, \tilde{y}) \\
&= T\Phi\left(\psi^{-1}(\tilde{u}), d\psi(q)^*\tilde{v}; d\Psi(p)^{-1}\begin{pmatrix}\tilde{\xi}\\ \tilde{y}\end{pmatrix}\right) \\
&= \left(f(\tilde{u}), df(\tilde{u})^{*-1}\tilde{v}; d\Phi(p)\circ d\Psi(p)^{-1}\begin{pmatrix}\tilde{\xi}\\ \tilde{y}\end{pmatrix}\right) \\
&= \left(f(\tilde{u}), df(\tilde{u})^{*-1}\tilde{v}; d(f(\tilde{u}), df(\tilde{u})^{*-1}\tilde{v})\begin{pmatrix}\tilde{\xi}\\ \tilde{y}\end{pmatrix}\right) .
\end{aligned} \tag{152}
$$

The last term may be expressed as follows

$$
\begin{aligned}
& d_{\tilde{u},\tilde{v}}(f(\tilde{u}), d_{\tilde{u}}f(\tilde{u})^{*-1}\tilde{v})\begin{pmatrix}\tilde{\xi}\\ \tilde{y}\end{pmatrix} \\
&= \left(d_{\tilde{u}}f(\tilde{u})\tilde{\xi}, d_u(d_{\tilde{u}}f(\tilde{u}))^{*-1}(d_{\tilde{u}}f(\tilde{u})^{-1}\tilde{\xi}, \tilde{v}) + d_{\tilde{u}}f(\tilde{u})^{*-1}\tilde{y}\right) ,
\end{aligned} \tag{153}
$$

using indices the last expression reads

$$
y_l = \frac{\partial^2 {f^{-1}}^i(u)}{\partial u^l \partial u^s}\frac{\partial {f^{-1}}^s(u)}{\partial u^k}\tilde{\xi}^k\tilde{v}_i + \frac{\partial {f^{-1}}^k(u)}{\partial u^l}\tilde{y}_k \ , \ l = 1, \dots, n \ , \tag{154}
$$

or

$$
\tilde{y}_m = -\frac{\partial f^r(\tilde{u})}{\partial \tilde{u}^m}\frac{\partial^2 {f^{-1}}^i(u)}{\partial u^r \partial u^s}\frac{\partial {f^{-1}}^s(u)}{\partial u^k}\tilde{\xi}^k\tilde{v}_i + \frac{\partial f^r(\tilde{u})}{\partial \tilde{u}^m}y_r \ . \tag{155}
$$

(The notation $d_{\tilde{u},\tilde{v}}$ indicates that this is the differential with respect to variables $\tilde{u}$ and $\tilde{v}$).
Hence we see that ξ transforms like a vector, but y involves second derivatives in its transformation formula. As we met a similar expression in the transformation law of the Christoffel symbols, we might be tempted to define the quantity

$$
\eta_l = y_l - \Gamma^j_{lk}(u)\xi^k v_j \ . \tag{156}
$$

Lemma B.1.4. *η_l transforms like a coordinate function of a covector, i.e.*

$$
\eta_l = \frac{\partial {f^{-1}}^l(u)}{\partial u^k}\tilde{\eta}_k \ , k = 1, \dots, n \ .
$$

Proof. In the first equality of the next calculation we are going to use (156), in the second one we use the transformation formulas Lemma B.1.2, Lemma B.1.3ii) and (154) for $\tilde{\xi}^k, \tilde{v}_j, \tilde{\Gamma}^k_{ij}(\tilde{u})$ and y_i:

$$\begin{aligned}
\frac{\partial {f^{-1}}^i(u)}{\partial u^l}\tilde{\eta}_i =& \frac{\partial {f^{-1}}^i(u)}{\partial u^l}\left(y_i - \tilde{\Gamma}^k_{ij}(\tilde{u})\tilde{\xi}^j\tilde{v}_k\right) \\
=& \frac{\partial {f^{-1}}^i(u)}{\partial u^l}\left(\frac{\partial f^r(\tilde{u})}{\partial \tilde{u}^i}y_r - \frac{\partial f^r(\tilde{u})}{\partial \tilde{u}^i}\frac{\partial^2 {f^{-1}}^k(u)}{\partial u^r \partial u^\nu}\frac{\partial f^\nu(\tilde{u})}{\partial \tilde{u}^j}\tilde{\xi}^j\tilde{v}_k\right) \\
&- \frac{\partial {f^{-1}}^i(u)}{\partial u^l}\tilde{\xi}^j\tilde{v}_k\frac{\partial {f^{-1}}^k(u)}{\partial u^\mu}\left(\frac{\partial^2 f^\mu(\tilde{u})}{\partial \tilde{u}^i\partial \tilde{u}^j} + \Gamma^\mu_{\nu s}(u)\frac{\partial f^s(\tilde{u})}{\partial \tilde{u}^j}\frac{\partial f^\nu(\tilde{u})}{\partial \tilde{u}^i}\right) \\
=& y_l - \tilde{\xi}^j\tilde{v}_k\left(\frac{\partial^2 {f^{-1}}^k(u)}{\partial u^l\partial u^\nu}\frac{\partial f^\nu(\tilde{u})}{\partial \tilde{u}^j} + \frac{\partial^2 f^\mu(\tilde{u})}{\partial \tilde{u}^i\partial \tilde{u}^j}\frac{\partial {f^{-1}}^i(u)}{\partial u^l}\frac{\partial {f^{-1}}^k(u)}{\partial u^\mu}\right) \\
&- \Gamma^j_{lk}(u)\xi^k v_j \\
=& \eta_l \ .
\end{aligned}$$

It follows from applying ∂_{u^l} to the identity $\frac{\partial {f^{-1}}^i(u)}{\partial u^j}\frac{\partial f^j(\tilde{u})}{\partial \tilde{u}^k} = \delta^i_k$ that the second summand in the last but one equation is zero. □

B.1.2. Natural splitting of TT^*M. Let $\mathcal{X}(M)$ denote the *set of smooth vector fields on* M. In Riemannian geometry the *Levi-Civita connection* $\nabla = {}^g\nabla$ is defined as the unique map $\mathcal{X}(M)\times\mathcal{X}(M)\to\mathcal{X}(M)$ such that $\forall f\in C^\infty(M)\ \forall X,Y,Z\in\mathcal{X}(M)$ one has

1. $C^\infty(M)$-linearity in the first component
2. $\mathbb{R}$-linearity in the second component
3. (Leibniz rule) $\nabla_X(fY) = (Xf)Y + f\nabla_X Y$
4. (torsion free) $\nabla_X Y - \nabla_Y X + [X,Y] = 0$
5. (compatibility with metric) $\nabla_Z g(X,Y) = g(\nabla_Z X, Y) + g(X, \nabla_Z Y)$.

Following [**E67**] and [**Kl**] we redefine this concept: First let us recall the short notation $(u^i; v_j)$ for $(u^1,\dots,u^n; v_1,\dots,v_n)$ and similarly for vectors with another number of components. Let the vector bundle morphism $K: TT^*M \to T^*M$ in natural local coordinates (cf. subsection B.1.1) be given by

$$K(u^i, v_j; \xi^k, y_l) = (u^i; y_l - \Gamma^k_{li}\xi^i v_k). \tag{157}$$

K is called *connection map*. The *Christoffel symbols* Γ^k_{ij} of ∇ are defined by

$$\nabla_{\partial_{u^i}}\partial_{u^j} = \Gamma^k_{ij}\partial_{u^k}\ ,\ \forall i,j = 1,\dots,n$$

where $(u^1,\dots,u^n)$ are local coordinates on M. They may be expressed in terms of derivatives of the metric as in equation (95). The relation between the two concepts is

$$\nabla^*_X\alpha = K\circ T\alpha(X) \quad , \quad X,Y\in\mathcal{X}(M) \quad , \quad \omega\in\Omega^1(M)\ .$$

Note that T denotes the *tangent map*. For instance for a 1-form $\alpha: M \to T^*M$ we get $T\alpha: TM\to TT^*M$ or in coordinates $(u^i;\xi^k)\mapsto$

$(u^i, \alpha_j(u); \xi^k, (\partial_{u^k}\alpha_j)\xi^k)$. Let τ_N denote the *projection of the tangent bundle of any manifold* N and τ_N^* the *projection of its cotangent bundle*. Locally we have

$$\begin{aligned} \tau_M^* \ : \ T^*M \ &\to M \\ p &\mapsto q \\ (u^i; v_j) &\mapsto (u^i) \end{aligned} \tag{158}$$

$$\begin{aligned} \tau_{T^*M}^* \ : \ TT^*M \ &\to T^*M \\ \zeta &\mapsto p \\ (u^i, v_j; \xi^k, y_l) &\mapsto (u^i; v_j). \end{aligned} \tag{159}$$

We define the *horizontal* respectively *vertical subbundle of* TT^*M by

$$\begin{aligned} \tau_{T^*M}^h \ &: \ T^hT^*M = Ker\ K \ \to \ T^*M, \\ \tau_{T^*M}^v \ &: \ T^vT^*M = Ker\ T\tau_M^* \ \to \ T^*M. \end{aligned}$$

In local coordinates

$$\begin{aligned} T_p^hT^*M &\cong \{(u^i, v_j; \xi^k, \Gamma_{li}^k \xi^i v_k) | (\xi^1, \dots, \xi^n) \in \mathbb{R}^n\}, \\ T_p^vT^*M &\cong \{(u^i, v_j; 0, y_l) | (y_1, \dots, y_n) \in \mathbb{R}^n\}, \end{aligned}$$

where $q = \tau_M^* p$ is represented in the local coordinate $\varphi : M \supset U \to \mathbb{R}^n$ by $\varphi(q) = (u^1, \dots, u^n)$ and $p = v_j du^j$ by $(u^i; v_j)$. It can be seen from these local expressions that T^hT^*M and T^vT^*M intersect in the zero section $\mathcal{O}_{T^*M}$ of $T(T^*M)$. As their ranks are n and the rank of $T(T^*M)$ is $2n$ it follows that TT^*M is isomorphic to the direct sum

$$TT^*M \cong T^hT^*M \oplus T^vT^*M.$$

Unfortunately the 2^{nd} of the vector space isomorphisms

$$\begin{aligned} T\tau_M^*(p) \ : \ T_p^hT^*M \ &\to T_qM \\ (u^i, v_j; \xi^k, \Gamma_{lj}^k \xi^j v_k) &\mapsto (u^i; \xi^k) \\ K(p) \ : \ T_p^vT^*M \ &\to T_q^*M \\ (u^i, v_j; 0, y_l) &\mapsto (u^i; y_l) \end{aligned} \tag{160}$$

depends on the choice of coordinates as y_l does not transform as a coordinate function of a covector (cf. subsection B.1.1). So they do not provide a natural isomorphism between $T_p^hT^*M \oplus T_p^vT^*M$ and $T_qM \oplus T_q^*M$. However there is a natural vector bundle isomorphism Θ between $\tau_{T^*M} : TT^*M \to T^*M$ and the pull-back bundle $pr_1 : (\tau_M^*)^*(TM \oplus T^*M) \to T^*M$ which, restricted to a fibre T_pT^*M, is in natural coordinates given by

$$\begin{aligned} \Theta(p) \ : \ T_pT^*M \ &\to (\{p\} \oplus T_qM \oplus T_q^*M \qquad , \ q = \tau_M^* p \\ (u^i, v_j; \xi^k, y_l) &\mapsto (u^i, v_j; \xi^k, \eta_l := y_l - \Gamma_{li}^j \xi^i v_j). \end{aligned} \tag{161}$$

Somewhat sloppy we will consider $\Theta(p)$ to be a fibrewise isomorphism between T_pT^*M and $T_qM \oplus T_q^*M$. Later on we will express various sections of bundles over T^*M in these new coordinates (ξ^k, η_l) on $T_qM \oplus T_q^*M$.

B.1.3. Liouville form and symplectic structure. The map

$$\begin{aligned} \Theta \ : \ TT^*M \ &\to \ \mathbb{R} \\ \zeta \ &\mapsto \ (\tau_{T^*M}\, \zeta)\, (T\tau_M^*\, \zeta) \end{aligned} \tag{162}$$

is a 1-form on T^*M. Θ is called *Liouville form.* Linearity follows from its expression in natural local coordinates

$$\Theta(u^i, v_j; \xi^k, y_l) = (u^i; v_j du^j)\, (u^i; \xi^k \partial_{u^k}) = v_j \xi^k du^j(\partial_{u^k}) = v_j \xi^j,$$

hence

$$\begin{aligned} \Theta(u^i, v_j) = v_j\, du^j \ : \ T_pT^*M \ &\to \ \mathbb{R} \\ \xi^k \partial_{u^k} + y_l \partial_{v_l} \ &\mapsto \ v_j \xi^j. \end{aligned} \tag{163}$$

Note that using the metric g the form Θ may be written in the form

$$\Theta(\zeta) = g^*(\tau_{T^*M}\zeta, g{\circ}T\tau_M^*\zeta),$$

where g^* denotes the dual metric of g and in abuse of notation we denote by the same symbol g also the metric isomorphism $TM \to T^*M : \xi \mapsto g(\xi, \cdot)$, which in local coordinates corresponds to lowering of indices.

Definition B.1.5. The *natural symplectic form* Ω on T^*M is defined by $\Omega = -d\Theta$.

Locally we get $\Omega(u^i, v_j)(\cdot, \cdot) = (du^i \wedge dv_i)(\cdot, \cdot)$, hence Ω is clearly closed. Nondegeneracy of Ω, i.e. $\Omega(\zeta, \tilde{\zeta}) = 0\ \forall \tilde{\zeta} \in TT^*M \ \Rightarrow \ \zeta = 0$, may be proven using the fact that $\{du^i, dv_i\}_{i=1}^n$ is a basis of T_pT^*M.

Lemma B.1.6.
i) $\Omega(\zeta, \tilde{\zeta}) = g^*(g{\circ}T\tau_M^*\, \zeta, K\, \tilde{\zeta}) - g^*(K\, \zeta, g{\circ}T\tau_M^*\, \tilde{\zeta})\ \ \forall \zeta, \tilde{\zeta} \in TT^*M$
ii)Ω *is represented under the isomorphism* $T_pT^*M \cong T_qM \oplus T_qM$ *by*

$$\Omega((\xi, \eta), (\tilde{\xi}, \tilde{\eta})) = \tilde{\eta}(\xi) - \eta(\tilde{\xi}) \quad , \quad \forall (\xi, \eta), (\tilde{\xi}, \tilde{\eta}) \in T_qM \oplus T_q^*M.$$

Proof. $\Omega(\zeta, \tilde{\zeta})$ is locally given at (u^i, v_j) by

$$\begin{aligned} (du^i \wedge dv_i)(\xi^k \partial_{u^k} + y_l \partial_{v_l}, \tilde{\xi}^k \partial_{u^k} + \tilde{y}_l \partial_{v_l}) &= \xi^k \tilde{y}_k - y_l \tilde{\xi}^l \\ &= \xi^k \tilde{y}_k - y_l \tilde{\xi}^l - \xi^k \Gamma_{ki}^s \tilde{\xi}^i v_s + \tilde{\xi}^l \Gamma_{li}^s \xi^i v_s \\ &= \xi^k (\tilde{y}_k - \Gamma_{ki}^s \tilde{\xi}^i v_s) - \tilde{\xi}^l (y_l - \Gamma_{li}^s \xi^i v_s) \\ &= \xi^k \tilde{\eta}_k - \tilde{\xi}^l \eta_l = \tilde{\eta}(\xi) - \eta(\tilde{\xi}) \end{aligned}$$

where we added 0 in the 2^{nd} equality (using $\Gamma_{ij}^k = \Gamma_{ji}^k$) and used the definition (145) of the covector η in the 4^{th} equality. This proves *ii)*; to prove *i)* we compute its RHS locally

$$\begin{aligned} g^{ik} g_{kl} \xi^l (\tilde{y}_i - \Gamma^l_{ik} \tilde{\xi}^k v_l) - g^{ik} (y_k - \Gamma^m_{kl} \xi^l v_m) g_{ir} \tilde{\xi}^r \\ = \xi^i \tilde{y}_i - \xi^i \Gamma^l_{ik} \tilde{\xi}^k v_l - \tilde{\xi}^k y_k + \tilde{\xi}^k \Gamma^m_{kl} \xi^l v_m \\ = \xi^i \tilde{y}_i - y_k \tilde{\xi}^k \end{aligned}$$

which equals $\Omega(\zeta, \tilde{\zeta})$ as we have seen above. In the second equality we used again $\Gamma^k_{ij} = \Gamma^k_{ji}$. □

In view of *ii*) we see that Ω is represented on $T_q M \oplus T_q^* M$ by

$$\Omega\left(\begin{pmatrix}\xi\\ \eta\end{pmatrix}, \begin{pmatrix}\tilde{\xi}\\ \tilde{\eta}\end{pmatrix}\right) = \left(\begin{pmatrix}0 & -\mathbb{1}\\ \mathbb{1} & 0\end{pmatrix}\begin{pmatrix}\xi\\ \eta\end{pmatrix}\right)^T \begin{pmatrix}\tilde{\xi}\\ \tilde{\eta}\end{pmatrix}. \tag{164}$$

B.1.4. Riemannian metric and almost complex structure. The map

$$\begin{aligned} G \,:\, TT^*M \oplus TT^*M &\to \mathbb{R} \\ (\zeta, \tilde{\zeta}) &\mapsto g^*(K\zeta, K\tilde{\zeta}) + g(T\tau_M^*\zeta, T\tau_M^*\tilde{\zeta}) \end{aligned}$$

defines a *Riemannian metric on* T^*M: symmetry follows from symmetry of the metrics g and g^*; the same for positive definiteness. Nondegeneracy of G follows from the fact that the subbundles T^hT^*M and T^vT^*M are orthogonal with respect to G; hence it follows from nondegeneracy of g and g^*. Note that under the natural isomorphism Θ in equation (161) the metric G is represented by

$$\begin{aligned} G|_{(q,p)} = \begin{pmatrix} g & 0 \\ 0 & g^{-1}\end{pmatrix} :\, (T_qM \oplus T_q^*M)^{\times 2} &\to \mathbb{R} \\ \left(\begin{pmatrix}\xi\\ \eta\end{pmatrix}, \begin{pmatrix}\tilde{\xi}\\ \tilde{\eta}\end{pmatrix}\right) &\mapsto \left(\begin{pmatrix}g & 0\\ 0 & g^{-1}\end{pmatrix}\begin{pmatrix}\xi\\ \eta\end{pmatrix}\right)^T \begin{pmatrix}\tilde{\xi}\\ \tilde{\eta}\end{pmatrix}. \end{aligned}$$

An *almost complex structure* $J \in \Gamma(End\, TT^*M)$ is determined by the identities

$$\begin{aligned} i)(g \circ T\tau_M^*) \circ J &= -K \\ ii) K \circ J &= (g \circ T\tau_M^*). \end{aligned} \tag{165}$$

Applying J from the right to *i*) respectively *ii*) and then using *ii*) respectively *i*) shows that $J^2 = -id$. The almost complex structure J is represented on $T_qM \oplus T_q^*M$ by

$$J(p) = \begin{pmatrix} 0 & -g^{-1}(q) \\ g(q) & 0 \end{pmatrix} \,:\, T_qM \oplus T_q^*M \to T_qM \oplus T_q^*M, \tag{166}$$

where $q = \tau_M^* p$. This coincides with the natural almost complex structure on the direct sum of a vector space with its dual space.

Compatibility of (Ω, G, J)

We call Ω and J *compatible* if $\Omega(\cdot, J\cdot)$ defines a Riemannian metric, which is the case - in fact it coincides with G. Let $\zeta, \tilde{\zeta} \in TT^*M$, then using the defining equations (165) for J

$$\begin{aligned}\Omega(\zeta, J\tilde{\zeta}) &= g^*(g\circ T\tau_M^*\zeta, K\circ J\tilde{\zeta}) - g^*(K\zeta, g\circ T\tau_M^*\circ J\tilde{\zeta})\\ &= g(T\tau_M^*\zeta, T\tau_M^*\tilde{\zeta}) + g^*(K\zeta, K\tilde{\zeta})\\ &= G(\zeta, \tilde{\zeta}).\end{aligned}$$

On $T_qM \oplus T_q^*M$ this proof reduces to matrix multiplication

$$J^T\Omega = \begin{pmatrix} 0 & g \\ -g^{-1} & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} g & 0 \\ 0 & g^{-1} \end{pmatrix} = G.$$

Note that compatibility immediately implies that J is an isometry with respect to G

$$G(J\zeta, J\tilde{\zeta}) = \Omega(J\zeta, J^2\tilde{\zeta}) = -\Omega(J\zeta, \tilde{\zeta}) = G(\zeta, \tilde{\zeta}),$$

hence

$$G(J\zeta, \tilde{\zeta}) = G(J^2\zeta, J\tilde{\zeta}) = \Omega(J^2\zeta, J^2\tilde{\zeta}) = \Omega(\zeta, \tilde{\zeta})$$

and

$$\Omega(J\zeta, J\tilde{\zeta}) = G(J\zeta, \tilde{\zeta}) = G(\tilde{\zeta}, J\zeta) = \Omega(\tilde{\zeta}, \zeta) = -\Omega(\zeta, \tilde{\zeta}).$$

B.1.5. Hamiltonian functions and vector fields. Consider (time-1-periodic) *Hamiltonian functions* of the form *kinetic + potential energy*, i.e. $H : S^1 \times T^*M \to \mathbb{R}$

$$H(t, p) = \tfrac{1}{2} g^*(p, p) + V(t, q) \quad , \quad q = \tau_M^* p \tag{167}$$

where $V \in C^\infty(\mathbb{R}/\mathbb{Z} \times M, \mathbb{R})$ is the potential energy. The Legendre condition in natural local coordinates $(u^i; v_j)$ of T^*M reads

$$0 \neq det \left(\frac{\partial^2 H}{\partial v_i \partial v_j}\right)_{i,j=1}^n = det\ (g^{ij})_{i,j=1}^n$$

and is satisfied due to the nondegeneracy condition imposed on the metric g. Hence the Legendre transform of H is defined globally and we get the *Lagrangian*

$$L(t, \dot{q}) = p(\dot{q}) - H(t, p) \quad , \quad q = \tau_M^* p. \tag{168}$$

$p = v_j du^j$ may be obtained as a function of $\dot{q} = (\dot{q})^i \partial_{u^i}$ from

$$(\dot{q})^i = \frac{\partial H}{\partial v_i} = g^{ik} v_k,$$

i.e. by lowering of indices. Therefore

$$\begin{aligned}L(t, \dot{q}) &= v_i(\dot{q})^i - \tfrac{1}{2} g^{ik} v_i v_j - V(t, q)\\ &= \tfrac{1}{2} g(\dot{q}, \dot{q}) - V(t, q).\end{aligned}$$

Associated to H and the symplectic structure Ω is the *Hamiltonian vector field* X_H which - in view of nondegeneracy of Ω - is uniquely defined by the identity

$$dH(\cdot) = \Omega(X_H, \cdot). \tag{169}$$

For $\zeta \in TT^*M$

$$\begin{aligned}\Omega(X_H, \zeta) &= dH(\zeta) = G(\nabla H, \zeta) = G(J\nabla H, J\zeta) \\ &= \Omega(J\nabla H, J^2\zeta) = \Omega(-J\nabla H, \zeta)\end{aligned}$$

and therefore

$$X_H = -J\nabla H. \tag{170}$$

In natural fiber coordinates (161) we have

$$\begin{aligned}{}^G\nabla H(t,p) &= \begin{pmatrix} {}^g\nabla V(t,q) \\ p \end{pmatrix}, q = \tau_M^* p \\ X_{H_t}(p) = -J(p)\nabla H(t,p) &= \begin{pmatrix} g^{-1}(q)p \\ -g(q)^g\nabla V(t,q) \end{pmatrix}.\end{aligned} \tag{171}$$

The former identity can be seen as follows: let $\zeta = (\xi^k, y_l) \in T_{(u,v)}T^*M$ and denote by (ξ^k, η_l) the corresponding element of $T_uM \oplus T_u^*M$; the element corresponding to $\nabla H_t(p)$ we denote by (a,b), then

$$dH_t(p)\,\zeta = G(\nabla H, \zeta) = g(a,\xi) + g^*(b,\eta).$$

We compute now the LHS in local coordinates

$$\begin{aligned}&\frac{\partial H}{\partial u^k}(u^i, v_j)\,\xi^k + \frac{\partial H}{\partial v_l}(u^i, v_j)\,y_l \\ &= \frac{\partial V}{\partial u^k}(t, u^i)\,\xi^k + \tfrac{1}{2}\frac{\partial g^{ij}(u)}{\partial u^k}\,v_i v_j \xi^k + g^{ij}(u)\,v_i y_j \\ &= g({}^g\nabla V(t,u), \xi) + \tfrac{1}{2}\frac{\partial g^{ij}(u)}{\partial u^k}\,v_i v_j \xi^k + g^{ij}(u)\,v_i\eta_j + g^{ij}(u)\,v_i\Gamma^s_{jr}(u)\xi^r v_s \\ &= g({}^g\nabla V(t,u), \xi) + g^*(v,\eta).\end{aligned}$$

In the third equality we replaced y_j by $\eta_j + \Gamma^s_{jr}v_s\xi^r$, the last equality follows as the $2^{nd} + 4^{th}$ term is zero (use equation (95)). The statement now follows from the nondegeneracy of g and the orthogonality of the splitting of TT^*M.

X_H generates the *1-parameter group of diffeomorphisms* $\varphi_t : T^*M \to T^*M$, $t \in \mathbb{R}$, defined by

$$\frac{d}{dt}\varphi_t = X_{H_t}\circ\varphi_t\ ,\ \varphi_0 = id\ ,$$

i.e. if $x \in C^\infty(\mathbb{R}, T^*M)$ is a solution of the initial value problem

$$\begin{cases}\dot{x}(t) = X_{H_t}(x(t)) \\ x(0) = x_0\end{cases} \tag{172}$$

then $\varphi_t x_0 = x(t)$. $\mathcal{P}er\,(H)$ denotes the set of 1-periodic solutions of (172). We denote by $Symp(T^*M,\Omega)$ the set of symplectomorphisms, i.e. the set of diffeomorphisms φ of T^*M which preserve the symplectic structure, i.e. $\varphi^*\Omega = \Omega$.

LEMMA B.1.7. $\varphi_t \in Symp(T^*M,\Omega)$ $\forall t \in \mathbb{R}$.

PROOF. Clearly $\varphi_0 = id$ preserves Ω, the idea is now to show that preservation of Ω is constant in t

$$\begin{aligned} \left.\frac{d}{dt}\right|_{t=0} \varphi^*_{t+t_0}\Omega = L_{X_{H_t}}\Omega &= d\iota_{X_{H_{t_0}}}\Omega + \iota_{X_{H_{t_0}}} d\Omega \\ &= ddH_{t_0} = 0. \end{aligned}$$

The first equality is just the definition of the *Lie-derivative* L of a differential form in the direction of a vector field. Then we used Cartan's formula $L_X = d\iota_X + \iota_X d$ (cf. [**AM78**] theorem 2.4.13 *iv*)), the fact that $d\Omega = 0$ and the definition (169) of the Hamiltonian vector field X_{H_t}. □

B.1.6. First variation formula. Let the *(perturbed) symplectic action functional* be defined on the space of free, smooth loops in T^*M by

$$\begin{aligned} \mathcal{A}_V \;:\; C^\infty(S^1,T^*M) &\to \mathbb{R} \\ z &\mapsto \int_{S^1} z^*\Theta - \int_0^1 H(t,z(t))\,dt \end{aligned} \tag{173}$$

where Θ is the Liouville form (162) and H is a 1-periodic Hamiltonian as in (167). Let

$$Crit\,\mathcal{A}_V = \{z \in C^\infty(S^1,T^*M) \mid d\mathcal{A}_V(z) = 0 \text{ on } \Gamma(z^*TT^*M)\}. \tag{174}$$

The next result says that $z \in \mathcal{P}er(H) \Rightarrow z \in Crit\,\mathcal{A}_V$.

PROPOSITION B.1.8. *For $z \in C^\infty(S^1,T^*M)$ and $\zeta \in \Gamma(z^*TT^*M)$*

$$d\mathcal{A}_V(z)\,\zeta = \int_0^1 \Omega(\dot z(t) - X_{H_t}(z(t)),\zeta(t))\,dt.$$

PROOF. (**intrinsic**) Denote by *Exp* the exponential map of the Levi-Civita connection ${}^G\nabla$ of (T^*M,G). Consider the two-parameter map, cf. (137),

$$A(t,\tau) = Exp_{z(t)}\tau\zeta(t).$$

$A(\tau) = A(\cdot,\tau)$ can be interpreted as a path in $C^\infty(S^1,T^*M)$ with $A(0) = z$ and $\left.\frac{d}{d\tau}\right|_{\tau=0} A(\tau) = \zeta$. We define

$$f(\tau) = \int_0^1 \Theta(\partial_t A(t,\tau))\,dt = \int_{S^1} A(\tau)^*\Theta$$

and compute

$$\begin{aligned} f(0) - f(\tau) &= \int_{S^1} A(0)^*\Theta - A(\tau)^*\Theta \\ &= \int_{S^1\times[0,\tau]} dA(\tau)^*\Theta = -\int_{S^1\times[0,\tau]} A(\tau)^*\Omega \\ &= -\int_0^\tau \int_0^1 \Omega(\partial_t A(t,s), \partial_s A(t,s))\, dt\, ds \\ &=: \int_0^\tau k(s)\, ds =: K(\tau). \end{aligned}$$

In the second equality the induced orientation on the boundary of $S^1 \times [0,\tau]$ has to be used. Note that $K(0) = 0$. Now

$$\begin{aligned} d\mathcal{A}_V(z)\,\zeta &= \frac{d}{d\tau}\Big|_{\tau=0} \mathcal{A}_V(A(\tau)) \\ &= \frac{d}{d\tau}\Big|_{\tau=0} \int_{S^1} A(\tau)^*\Theta - \int_0^1 H(t, A(t,\tau))\, dt \\ &= \lim_{\tau\to\infty} \tfrac{f(\tau)-f(0)}{\tau} - \int_0^1 dH_t(z(t))\,\zeta(t)\, dt \\ &= \lim_{\tau\to\infty} \tfrac{K(0)-K(\tau)}{\tau} - \int_0^1 \Omega(X_{H_t}(z), \zeta)\, dt \\ &= -k(0) - \int_0^1 \Omega(X_{H_t}(z), \zeta)\, dt \\ &= \int_0^1 \Omega(\dot z, \zeta)\, dt - \int_0^1 \Omega(X_{H_t}(z), \zeta)\, dt \\ &= \int_0^1 \Omega(\dot z(t) - X_{H_t}(z(t)), \zeta(t))\, dt \end{aligned}$$

where we have used the definition (169) of X_{H_t} in the third equality. □

PROOF. (**local coordinates**) We set $z = (u^i, v_j)$ and $\zeta = (u^i, v_j; \xi^k, y_l)$, everything depending on $t \in \mathbb{R}/\mathbb{Z}$. Now we pick smooth maps $z_\tau = (u^i_\tau, v^\tau_j)$ such that $(u^i_0, v^0_j) = (u^i, v_j)$ and $\frac{d}{d\tau}\big|_{\tau=0} (u^k_\tau, v^\tau_l) = (\xi^k, y_l)$.

$$\begin{aligned} d\mathcal{A}_V(z)\,\zeta &= \frac{d}{d\tau}\Big|_{\tau=0} \mathcal{A}_V(u^i_\tau, v^\tau_j) \\ &= \frac{d}{d\tau}\Big|_{\tau=0} \int_0^1 v^\tau_j \dot u^j_\tau - \frac{1}{2} g^{ij}(u_\tau) v^\tau_i v^\tau_j - V(t, u^i_\tau)\, dt \\ &= \int_0^1 y_j(\dot u^j - \frac{1}{2} g^{ij}(u) v_i) + v_j \left(\dot\xi^j - \frac{1}{2}\frac{\partial g^{ij}(u)}{\partial u^l}\xi^l v_i - \frac{1}{2} g^{ij}(u) y_i\right) \\ &\quad - \frac{\partial V(t,u)}{\partial u^l}\xi^l\, dt. \end{aligned}$$

We replace y_l by $\eta_l + \Gamma^k_{li} v_k \xi^i$ (η_l is an intrinsic object, namely the component of a covector, cf. subsection B.1.1) and get

$$\int_0^1 \eta_j \dot{u}^j + \Gamma^l_{jk}(u) v_l \xi^k \dot{u}^j - \frac{1}{2} g^{ij}(u) v_i \eta_j - \frac{1}{2} g^{ij}(u) v_i \Gamma^l_{jk}(u) v_l \xi^k - \dot{v}_j \xi^j$$
$$-\frac{1}{2} v_j \frac{\partial g^{ij}(u)}{\partial u^l} \xi^l v_i - \frac{1}{2} g^{ij}(u) v_j \eta_i - \frac{1}{2} g^{ij}(u) v_j \Gamma^l_{ik}(u) v_l \xi^k - \frac{\partial V(t,u)}{\partial u^l} \xi^l \, dt.$$

Denoting by $\langle \cdot, \cdot \rangle$ the pairing between a covector and a vector, the 1^{st} term equals $\langle \eta, \partial_t u \rangle$, the $2^{nd}+5^{th}$ equals $-\langle \nabla^*_t v, \xi \rangle$, the $3^{rd}+7^{th}$ equals $-\langle \eta, g^{-1} v \rangle$, the 9^{th} equals $-\langle dV(t,u), \xi \rangle$ and the $4^{th} + 6^{th} + 8^{th}$ equals 0. The last statement can be seen by using (95). Hence we have with $x = \tau^*_M z$

$$\begin{aligned}
&\int_0^1 \langle \eta, \partial_t x \rangle - \langle \eta, g^{-1}(x) z \rangle - \langle dV(t,x), \xi \rangle - \langle \nabla^*_t z, \xi \rangle \, dt \\
&= \int_0^1 \langle \eta, \partial_t x - g^{-1}(x) z \rangle + \langle -\xi, \nabla^*_t z + g(x)^g \nabla V(t,x) \rangle \, dt \\
&= \int_0^1 - \left(\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix} \right)^T \begin{pmatrix} \partial_t x - g^{-1}(x) z \\ \nabla^*_t z + g(x)^g \nabla V(t,x) \end{pmatrix} dt \\
&= \int_0^1 -\Omega(\zeta(t), \dot{z}(t) - X_{H_t}(z(t))) \, dt
\end{aligned}$$

where in the last equality we used the facts that $\dot{z} \in TT^*M$ is represented by $(\partial_t x, \nabla^*_t z)$ on $T_x M \oplus T^*_x M$ and $X_{H_t}(z)$ by $-J(z)\nabla H_t(z) = (g^{-1}(x) z, -g(x)^g \nabla V_t(x))$, as well as equations (164) and (171). $\square$

Let us mention that one can extend the definition of $\mathcal{A}_V$ to the free loop space ΛT^*M of T^*M – the completion of $C^\infty(S^1, T^*M)$ with respect to the Sobolev norm

$$\|z\|^2_{1,2} = \|z\|^2_{L^2} + \|\partial_t z\|^2_{L^2}.$$

The norms on the RHS are defined via a cover of T^*M by finitely may (natural) coordinate charts (i.e. induced by a finite coordinate cover of M as explained in subsection B.1.1). Therefore this norm depends on the cover, but any two such norms are equivalent. As the elements of ΛT^*M are almost everywhere differentiable, the definition (173) of $\mathcal{A}_V$ still makes sense. Regularity theory techniques are now required to show $Per\, H = Crit\, \mathcal{A}_V$.

B.1.7. First Chern class. We are going to show that the first Chern class of the bundle $\pi : E = TM \oplus T^*M \to M$ is zero. Hence we need to introduce a complex structure on E. Then we define a connection $\hat{\nabla}$ on E and compute the trace of its curvature 2-form $F^{\hat{\nabla}}$, which turns out to be zero. The result now follows from Chern-Weil theory. Functoriality of the Chern class implies that for any closed submanifold $N \subset T^*M$ we have $c_1(T_N T^*M) = 0$.

Starting with a chart (ϕ, U) of M we get a local trivialization

$$\begin{aligned}\Phi \,:\, \pi^{-1}(U) = T_U M \oplus T_U^* M &\rightarrow U \times \mathbb{R}^n \times \mathbb{R}^n \\ (\xi, \eta) &\mapsto (q, d\phi(q)\xi, {d\phi(q)^*}^{-1}\eta)\end{aligned}$$

where $q = \pi(\xi, \eta)$ denotes the base point. Let the *canonical almost complex structure on* $\mathbb{R}^{2n}$ be defined by

$$J_0 = \begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} \,:\, \mathbb{R}^{2n} \rightarrow \mathbb{R}^{2n},$$

i.e. ${J_0}^2 = -\mathbb{1}$. Identifying the complex vector spaces $(\mathbb{R}^{2n}, J_0)$ and $(\mathbb{C}^n, i)$ via the isomorphism

$$\begin{aligned}I \,:\, (\mathbb{R}^{2n}, J_0) &\rightarrow (\mathbb{C}^n, i) \\ (x, y) &\mapsto x + iy =: z,\end{aligned}$$

the complex linear group $GL(n, \mathbb{C})$ is identified with

$$\begin{aligned}GL_{\mathbb{C}}(2n, \mathbb{R}) &= \{A \in GL(2n, \mathbb{R}) \mid J_0 A = A J_0\} \\ &= \{A \in GL(2n, \mathbb{R}) \mid A \text{ is of the form} \begin{pmatrix} X & -Y \\ Y & X \end{pmatrix}\}.\end{aligned}$$

We observe that $\Phi(q)$ respects the complex structures

$$J(q) = \begin{pmatrix} 0 & -g^{-1}(q) \\ g(q) & 0 \end{pmatrix} \,:\, T_q M \oplus T_q^* M \rightarrow T_q M \oplus T_q^* M$$

and J_0 on $\mathbb{R}^{2n}$, i.e. $\Phi(q) \circ J(q) = J_0 \circ \Phi(q)$ hence is a complex isomorphism. Let $(\tilde{\phi}, U)$ be another chart for $U \subset M$ and $f(\tilde{u}) = \phi \circ \tilde{\phi}^{-1}(\tilde{u}) = u$ the transition map, cf. subsection B.1.1. The transition map for E

$$\Phi(q) \circ \tilde{\Phi}(q)^{-1} = \begin{pmatrix} d\phi(q) \circ d\tilde{\phi}^{-1}(\tilde{\phi}q) & 0 \\ 0 & {d\phi(q)^*}^{-1} \circ d\tilde{\phi}(q)^* \end{pmatrix}$$

indeed takes values in $GL_{\mathbb{C}}(2n, \mathbb{R})$. We define a connection $\hat{\nabla}$ on E by

$$\hat{\nabla}_X \begin{pmatrix} \xi \\ \eta \end{pmatrix} = \begin{pmatrix} \nabla_X \xi \\ \nabla_X^* \eta \end{pmatrix}, \quad X, \xi \in \mathcal{X}(M), \eta \in \Omega^1(M) \tag{175}$$

where ∇ respectively ∇^* denotes the Levi-Civita connection of (M, g) acting on vector respectively covector fields. In local coordinates we have

$$\hat{\nabla} = d + \hat{A} = \begin{pmatrix} d & 0 \\ 0 & d \end{pmatrix} + \begin{pmatrix} \Gamma_i du^i & 0 \\ 0 & -\Gamma_j du^j \end{pmatrix}$$

where $\Gamma_i du^i \in \Omega^1(M, End\, \mathbb{R}^n)$ is the connection potential of ∇, i.e. Γ_{ij}^k are the Christoffel symbols. The curvature 2-form is given by

$$F^{\hat{\nabla}} = \hat{\nabla}^2 = d^2 + d\hat{A} - \hat{A}d + \hat{A}d + \hat{A} \wedge \hat{A} = d\hat{A} + \hat{A} \wedge \hat{A}$$

which locally (in the following we use the bundle chart $d\phi \oplus d\phi^{*^{-1}}$) reads

$$\begin{aligned} F^{\hat\nabla}_{ij}\, du^i \wedge du^j &= \begin{pmatrix} \sum\limits_{i,j} \frac{\partial \Gamma_i}{\partial u^j} du^j \wedge du^i & 0 \\ 0 & \sum\limits_{i,j} -\frac{\partial \Gamma_j}{\partial u^i} du^i \wedge du^j \end{pmatrix} \\ &+ \begin{pmatrix} \sum\limits_{i,j} \Gamma_i \Gamma_j du^i \wedge du^j & 0 \\ 0 & \sum\limits_{i,j} \Gamma_i \Gamma_j du^i \wedge du^j \end{pmatrix} \\ &= \sum_{i<j} \begin{pmatrix} \left(\frac{\partial \Gamma_j}{\partial u^i} - \frac{\partial \Gamma_i}{\partial u^j}\right) + [\Gamma_i, \Gamma_j] & 0 \\ 0 & -\left(\frac{\partial \Gamma_j}{\partial u^i} - \frac{\partial \Gamma_i}{\partial u^j}\right) + [\Gamma_i, \Gamma_j] \end{pmatrix} du^i \wedge du^j . \end{aligned}$$

Note that $F^{\hat\nabla} \in \Omega^2(M, End\, \mathbb{R}^{2n})$ and it transforms according to the adjoint action of $GL(2n, \mathbb{R})$ on $End\, \mathbb{R}^{2n}$, hence taking the trace of $F^{\hat\nabla}$ in local coordinates is independent of the choice of these coordinates, therefore $tr\, F^{\hat\nabla} \in \Omega^2(M)$. Now the trace of a commutator vanishes and $tr\, A = tr\, A^t$, hence

$$tr\, F^{\hat\nabla} = 0 \in \Omega^2(M).$$

Moreover Chern-Weil theory asserts that the cohomology class $[tr\, F^{\hat\nabla}]$ is independent of the choice of connection (cf. [**Jo91**] Lemma 1.4.2 or [**Sa96**] section 1.4). The *first Chern class of* E is now given by

$$\boxed{c_1(TM \oplus T^*M) \stackrel{def}{=} \frac{i}{2\pi}[tr\, F^{\hat\nabla}] = 0 \in H^2_{dR}(M, \mathbb{Z}).} \tag{176}$$

THEOREM B.1.9. *Let N be a closed submanifold of T^*M, then it follows* $c_1(T_N T^*M) = 0$.

PROOF. VERSION A Let $i : N \hookrightarrow T^*M$ denote the inclusion and $\tau^*_M|_N = \tau^*_M \circ i : N \to M$ the restriction of the cotangent projection to N, then

$$c_1(T_N T^*M) = c_1((\tau^*_M|_N)^* E) = i^*(\tau^*_M)^*(c_1(E)) = 0.$$

□

PROOF. VERSION B Let $\tilde{E} = T_N T^*M$, then according to [**GH78**] chapter3 section3

$$c_1(\tilde{E}) = -c_1(\tilde{E}^*) = -c_1(\Lambda^n \tilde{E}^*)$$

where $\Lambda^n \tilde{E}^*$ denotes the canonical (complex) line bundle. Now $c_1(\Lambda^n \tilde{E}^*) = 0$ if and only if there exists a nonvanishing section s of $\Lambda^n \tilde{E}^*$. This is due to the fact that the first Chern class of a complex line bundle equals the Poincare dual of the zero set of a generic section. Let (ξ^k, η_l) denote the canonical fiber coordinates of $\tilde{E}$ (cf. subsection B.1.1) and define

$$z_k = \xi^k + i\, \eta_k \ , \ k = 1, \dots , n \ ,$$

then

$$dz_1 \wedge \dots \wedge dz_n$$

is a nonvanishing and globally defined section of $\Lambda^n \tilde{E}^*$ (as the z_i transform under a change of coordinates by multiplication with an element of $Sl(n, \mathbb{C})$).
$\square$

In the case where N is a closed Riemann surface $c_1(T_N T^*M) = 0$ is equivalent to $T_N T^*M$ being symplectically trivial, i.e. there exists a symplectic bundle isomorphism $\Phi : N \times \mathbb{R}^{2n} \to T_N T^*M$. This statement is a consequence of theorem B.1.14 in the next subsection.

We would like to mention the following facts: We have a hermitian structure H on $TM \oplus T^*M$ given by

$$H(\cdot,\cdot) = G(\cdot,\cdot) + i\Omega(\cdot,\cdot),$$

which is complex anti-linear in the first and complex linear in the second variable. The connection $\hat{\nabla}$ is *Riemannian* ($\hat{\nabla} G \equiv 0$)

$$\hat{\nabla}_X G(\cdot,\cdot) = G(\hat{\nabla}_X \cdot, \cdot) + G(\cdot, \hat{\nabla}_X \cdot),$$

unitary ($\hat{\nabla} J \equiv 0$)

$$(\hat{\nabla}_X J)\begin{pmatrix}\xi\\ \eta\end{pmatrix} = \hat{\nabla}_X\left(J\begin{pmatrix}\xi\\ \eta\end{pmatrix}\right) - J\hat{\nabla}_X\begin{pmatrix}\xi\\ \eta\end{pmatrix} = 0$$

and hence *hermitian* ($\hat{\nabla} H \equiv 0$)

$$\hat{\nabla}_X H(\cdot,\cdot) = H(\hat{\nabla}_X \cdot, \cdot) + H(\cdot, \hat{\nabla}_X \cdot).$$

B.1.8. Conley-Zehnder index of periodic Hamiltonian orbits. We are going to recall the definitions of the Maslov index of a loop of symplectic matrices and of the Conley-Zehnder index of a path of symplectic matrices starting at the identity and ending at a matrix which doesn't contain 1 in its spectrum. Then we introduce the first Chern number of a symplectic vector bundle E over a closed Riemann surface Σ and use the result of the former subsection $c_1(E) = 0$ for $\Sigma \subset T^*M$ to construct a well-defined map $\mu_{CZ} : \mathcal{P}er\ (H) \to \mathbb{Z}$.

Maslov index of loops in $Sp(2n, \mathbb{R})$

Let $Mat(2n, \mathbb{R})$ denote the set of $2n$ by $2n$ matrices with real entries. The symplectic linear group is given by

$$\begin{aligned} Sp(2n,\mathbb{R}) &= \{A \in Mat(2n,\mathbb{R}) \mid A^t J_0 A = J_0\} \\ &= \{A \in Mat(2n,\mathbb{R}) \mid A^* \omega_0 = \omega_0\} \end{aligned}$$

where the *standard symplectic structure* ω_0 on $\mathbb{R}^{2n}$ is in coordinates $(x^1, \ldots, x^n, y_1, \ldots, y_n)$ given by $\omega_0 = dx^j \wedge dy_j$ and the *standard complex structure* J_0 by

$$J_0 = \begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}.$$

It is known that $\pi_1(U(n,\mathbb{C})) \simeq \mathbb{Z}$ (cf. [**MS95**] Prop. 2.21); an isomorphism of fundamental groups is induced by the determinant map $det : U(n,\mathbb{C}) \to S^1$. Define

$$U(2n,\mathbb{R}) = \{\begin{pmatrix} X & -Y \\ Y & X \end{pmatrix} \in Gl(2n,\mathbb{R}) \mid X^tY = Y^tX, X^tX + Y^tY = \mathbb{1}\}$$

Then for any element of $U(2n, \mathbb{1}R)$ the transposed complex conjugate of $X + iY$ equals its inverse and so is an element of $U(n,\mathbb{C})$. Moreover,

$$U(2n,\mathbb{R}) = Sp(2n,\mathbb{R}) \cap O(2n,\mathbb{R})$$

and we have an isomorphism

$$\begin{aligned} I^* \ : \ U(n,\mathbb{C}) &\to U(2n,\mathbb{R}) \\ X + iY &\mapsto \begin{pmatrix} X & -Y \\ Y & X \end{pmatrix}. \end{aligned}$$

Now $U(2n,\mathbb{R})$ is a *strong deformation retract* of $Sp(2n,\mathbb{R})$ as there exists a homotopy

$$\begin{aligned} r(t) \ : \ Sp(2n,\mathbb{R}) &\to Sp(2n,\mathbb{R}) \\ A &\mapsto (AA^t)^{-t/2}A \end{aligned}$$

such that $r(0) = \mathbb{1}$, $r(1)(Sp(2n,\mathbb{R})) = U(2n,\mathbb{R})$ and $r(t)A = A$ $\forall t \in [0,1]$ $\forall A \in U(2n,\mathbb{R})$. According to [**StZi**], Satz 5.1.20, $i : U(2n,\mathbb{R}) \hookrightarrow Sp(2n,\mathbb{R})$ induces an isomorphism of fundamental groups. Now consider the map

$$\rho \ : \ Sp(2n,\mathbb{R}) \xrightarrow{r(1)} U(2n,\mathbb{R}) \xrightarrow{I^*} U(n,\mathbb{C}) \xrightarrow{det} S^1 \tag{177}$$

and define the *Maslov index of a symplectic loop* $\gamma : \mathbb{R}/\mathbb{Z} \to Sp(2n,\mathbb{R})$ by

$$\mu_{Symp}(\gamma) = deg\ \rho\circ\gamma. \tag{178}$$

μ_{Symp} provides an explicit isomorphism $\pi_1(Sp(2n,\mathbb{R})) \to \mathbb{Z}$. Moreover it is the unique functor mentioned in the next theorem.

THEOREM B.1.10. *([**MS95**] thm. 2.27) There exists a unique functor μ_{Symp}, called* Maslov index, *which assigns an integer $\mu_{Symp}(\gamma)$ to every loop of symplectic matrices $\gamma : \mathbb{R}/\mathbb{Z} \to Sp(2n,\mathbb{R})$ and satisfies the following axioms*

(homotopy) *Two loops in $Sp(2n,\mathbb{R})$ are homotopic if and only if they have the same Maslov index.*

(product) *For any two loops $\gamma_1, \gamma_2 : \mathbb{R}/\mathbb{Z} \to Sp(2n,\mathbb{R})$ we have*

$$\mu_{Symp}(\gamma_1\circ\gamma_2) = \mu_{Symp}(\gamma_1) + \mu_{Symp}(\gamma_2).$$

In particular the constant loop $\gamma(t) \equiv \mathbb{1}$ has Maslov index 0.

(direct sum) *If $n = n' + n''$ identify $Sp(2n',\mathbb{R}) \oplus Sp(2n'',\mathbb{R})$ in the obvious way with a subgroup of $Sp(2n,\mathbb{R})$. Then*

$$\mu_{Symp}(\gamma' \oplus \gamma'') = \mu_{Symp}(\gamma') + \mu_{Symp}(\gamma'').$$

(normalization) *The loop $\gamma : \mathbb{R}/\mathbb{Z} \to S^1 \simeq U(1,\mathbb{C}) \subset Sp(2,\mathbb{R})$ defined by $\gamma(t) = e^{2\pi it}$ has Maslov index 1.*

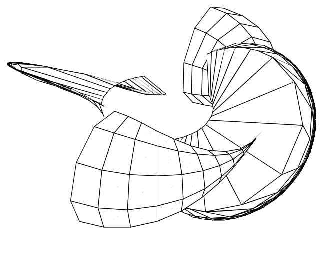

FIGURE B.1. The Maslov cycles $\mathcal{C}_+$ and $\overline{Sp_1}(2,\mathbb{R})$

REMARK B.1.11. For a generic path γ we can interpret $2\mu_{Symp}(\gamma)$ as the intersection number of the loop γ with the *Maslov cycle* (cf. [**MS95**] section 2.2)

$$\overline{Sp_1}(2n,\mathbb{R}) = \bigcup_{k=1}^{n} Sp_k(2n,\mathbb{R})$$

where $Sp_k(2n,\mathbb{R})$ consists of all

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \in Sp(2n,\mathbb{R})$$

such that $rank\, B = n-k$. Each stratum $Sp_k(2n,\mathbb{R})$ is a submanifold of $Sp(2n,\mathbb{R})$ of codimension $k(k+1)/2$ consisting of two connected components, cf. [**RS93**] section 4.

As is explained in great detail in Appendix D we can identify $Sp(2,\mathbb{R})$ with the interior of the full 2-torus (hereby solving an exercise in section 8.5.3 of [**Ar88**]). The image of $\overline{Sp_1}(2,\mathbb{R})$ under this homeomorphism is shown in figure B.1 – the two surfaces to the left and right (actually due to numerical approximation this is only a sketch of the image). The double-trumpet like surface there corresponds to $\mathcal{C}_+$, the set of all symplectic matrices whose spectrum contains 1. The singular point of $\mathcal{C}_+$ represents the identity. Clearly we could also interpret $2\mu_{Symp}(\gamma)$ as intersection number of a generic loop with $\mathcal{C}_+$. $\mathcal{C}_+$ and $\overline{Sp_1}(2,\mathbb{R})$ touch each other in a curve through the identity.

Conley-Zehnder index of paths in $Sp(2n,\mathbb{R})$

We briefly recall the definition of the *Conley-Zehnder index* $\mu_{CZ}(\gamma)$ of a path $\gamma : [0,1] \to Sp(2n,\mathbb{R})$ with $\gamma(0) = \mathbb{1}$ and $\gamma(1) \in Sp^*_{2n} = Sp(2n,\mathbb{R}) \setminus \mathcal{C}_+$. More details may be found in Appendix D. This index was introduced in 1984 by Conley and Zehnder [**CZ84**]. Extend γ by connecting $\gamma(1)$ within Sp^*_{2n} (which has two connected components $Sp^*_{2n,\pm}$) to one of the two reference matrices $W^\pm \in Sp^*_{2n,\pm}$

$$W^+ = -\mathbb{1} \quad \text{or} \quad W^- = diag(2, \underbrace{-1,\dots,-1}_{(n-1)\ times}, 1/2, \underbrace{-1,\dots,-1}_{(n-1)\ times}).$$

Let $\tilde{\gamma} : [0,2] \to Sp(2n,\mathbb{R})$ denote the extended path and let $\tilde{u} = r(1)(\tilde{\gamma})$ be the corresponding path in $U(2n,\mathbb{R})$. Setting $det\, \tilde{u}(t) = e^{i\alpha(t)}$ we define

$$\mu_{CZ}(\gamma) := \frac{\alpha(2) - \alpha(0)}{\pi} \in \mathbb{Z}. \tag{179}$$

Note that μ_{CZ} is independent of the choice of the extension and invariant under homotopies that keep the initial point $\mathbb{1}$ fixed and vary the endpoint only within Sp^*_{2n}. For an interpretation as an intersection number see the former paragraph (figure B.1) and Appendix D.

Trivializations

The following facts without proofs – including the construction of c_1 – are taken from [**MS95**] section 2.6. Let E be a real vector bundle over an l-dimensional manifold N and ω a smooth nondegenerate section of $E^* \wedge E^*$, i.e. on each fibre E_q we have a symplectic bilinear form ω_q varying smoothly with $q \in N$. (E, ω) is called a *symplectic vector bundle* over N. A *complex structure* on a vector bundle $E \to N$ is an automorphism J of E such that $J^2 = -id$. J is called *compatible with* ω if J_q is compatible with ω_q for all $q \in N$, i.e. $\omega_q(J_q\cdot, J_q\cdot) = \omega_q(\cdot,\cdot)$ and $\omega_q(v, J_q v) > 0$ for all nonzero $v \in E_q$. Let $\mathcal{J}(E,\omega)$ denote the *space of complex structures compatible with* ω. $\mathcal{J}(E,\omega)$ is nonempty and contractible ([**MS95**] prop. 2.61). For any compatible pair (J,ω) the bilinear form $g_J(\cdot,\cdot) = \omega(\cdot, J\cdot)$ is symmetric, nondegenerate and positive definite. A triple (ω, J, g) with these properties is called a *hermitian structure on* E. E is called a *hermitian vector bundle.* A *trivialization* of a bundle E is an isomorphism from E to the trivial bundle which preserves the structure under consideration. As two symplectic vector bundles (E_1, ω_1) and (E_2, ω_2) are isomorphic (i.e. $\exists$ vector bundle morphism $\Psi : E_1 \to E_2$ such that $\Psi^*\omega_2 = \omega_1$) if and only if their underlying complex bundles are isomorphic ([**MS95**] thm. 2.60), the notions of symplectic and complex trivialization are essentially the same. We therefore combine them by defining a *unitary trivialization* of a hermitian vector bundle $E \to N$ to be a smooth map

$$\begin{aligned} \Phi \;:\; N \times \mathbb{R}^{2n} &\to E \\ (q,\zeta) &\mapsto \Phi(q)\zeta, \end{aligned}$$

where $\Phi(q) : \mathbb{R}^{2n} \to E_q$ is linear, which pulls back ω, J and g to the standard structures on $\mathbb{R}^{2n}$:

$$\Phi^* J = J_0 \;,\; \Phi^*\omega = \omega_0 \;,\; \Phi^* g = g_0.$$

PROPOSITION B.1.12. *([**MS95**] prop. 2.64) A hermitian vector bundle $E \to \Sigma$ over a compact Riemann surface Σ with nonempty boundary $\partial\Sigma$ admits a unitary trivialization.*

Let $c : [0,1] \to N$ be a smooth curve and $E \to N$ be a hermitian vector bundle. Given any unitary isomorphisms $\Phi_0 : \mathbb{R}^{2n} \to E_{c(0)}$ and

$\Phi_1 : \mathbb{R}^{2n} \to E_{c(1)}$ at the endpoints of c, there exists a unitary trivialization $\Phi(t) : \mathbb{R}^{2n} \to E_{c(t)}$ of c^*E which extends the ones at the endpoints ([**MS95**] Lemma 2.63). This is a consequence of the pathwise connectedness of $U(n,\mathbb{C})$, i.e. any two points are *homotopic.*
Replacing the curve by a cylinder, a corresponding result clearly cannot be expected to hold: The reason is that two loops in $U(n,\mathbb{C})$ are *not homotopic* in general. Moreover, the first Chern number of any symplectic vector bundle $E \to \Sigma^2$ (as defined below by cutting Σ in pieces) could be arranged to be zero.

LEMMA B.1.13. *Let $Z = \{(r,\vartheta) \in [0,1] \times \mathbb{R}/2\pi\mathbb{Z}\}$ be the cylinder in polar coordinates and $E \to N$ be a hermitian vector bundle over an l-dimensional manifold N. Given a smooth map $\gamma : Z \to N$ and a unitary trivialization $\Phi_0 : \mathbb{R}/2\pi\mathbb{Z} \times \mathbb{R}^{2n} \to \gamma_0^*E$ over one end of Z, where $\gamma_r(\vartheta) = \gamma(r,\vartheta)$, there exists a unitary trivialization $\Phi : Z \times \mathbb{R}^{2n} \to \gamma^*E$ which extends Φ_0.*

PROOF. We give a parametrized version of the proof of Lemma 2.63 in [**MS95**]. We first extend the trivialization Φ_0 to a small neighborhood $[0,\epsilon] \times \mathbb{R}/2\pi\mathbb{Z}$ of $0 \times \mathbb{R}/2\pi\mathbb{Z}$: For $\zeta \in \mathbb{R}^{2n}$ let $\Phi_0(\vartheta)\zeta = \sum_j s_{j0}(\vartheta)\zeta^j$, where $\{s_{j0}(\vartheta)\}_{j=1}^{2n}$ is the given symplectic, orthogonal frame of $E_{\gamma(0,\vartheta)}$. We have to construct $2n$ sections s_j of E which satisfy $g(s_j,s_k) = \delta_{jk}$, $\omega(s_j,s_{j+n}) = 1$ and $\omega(s_j,s_k) = 0$ for all other values of j and k. $\{s_j(r,\vartheta))\}_{j=1}^{2n}$ is called a *unitary basis* of $E_{\gamma(r,\vartheta)}$. Choose a Riemannian connection ∇ on E and consider the parallel transport $\tilde{s}_j(r,\vartheta)$ of $s_{j0}(\vartheta)$ along the curve $r \mapsto \gamma(r,\vartheta)$ for fixed ϑ. For small $r(\vartheta)$ and fixed ϑ the first n vectors $\tilde{s}_1(r,\vartheta),\dots,\tilde{s}_n(r,\vartheta)$ will be linearly independent over $\mathbb{C}$. As $\mathbb{R}/2\pi\mathbb{Z}$ is compact this also holds for all $\vartheta \in \mathbb{R}/2\pi\mathbb{Z}$ (choose r smaller than $\min_\vartheta r(\vartheta)$). Now apply Gram-Schmidt over the complex numbers to obtain a unitary basis

$$s_k(r,\vartheta) := \frac{\tilde{s}_k}{|\tilde{s}_k|} - \sum_{J=1}^{k-1} \frac{g(s_j,\tilde{s}_k)}{|\tilde{s}_k|}\, s_j - \sum_{J=1}^{k-1} \frac{\omega(s_j,\tilde{s}_k)}{|\tilde{s}_k|}\, Js_j \quad , \quad s_{k+n} := Js_k$$

where $k = 1,\dots,n$. Cover the annulus Z by finitely many annuli $Z_k = [a_k,b_k] \times S^1$, $k \in \{1,\dots,m\}$, over which such a unitary trivialization $\tilde{\Phi}_k$ exists. We may assume that $\tilde{\Phi}_0(1,\vartheta)$ coincides with the given one $\Phi_0(\vartheta)$ and $Z_k \cap Z_{k'} \neq \emptyset$ if and only if $|k-k'| = 1$. Starting at $k=0$ we apply the following procedure successively to all trivializations $\tilde{\Phi}_k$: Pick two adjacent trivializations $\tilde{\Phi}_k$ and $\tilde{\Phi}_{k+1} : Z_{k+1} \times \mathbb{R}^{2n} \to \gamma_{k+1}^*E$ (where $\gamma_k = \gamma|_{Z_k}$) and consider the transition map $\tilde{\Psi}_{k,k+1} := \tilde{\Phi}_k^{-1} \circ \tilde{\Phi}_{k+1} : Z_k \cap Z_{k+1} \to Sp(2n,\mathbb{R})$. We can associate a Maslov index to $\tilde{\Psi}_{k,k+1}$ by picking any $\gamma_{k,k+1} : S^1 \to Z_k \cap Z_{k+1}$ generating $\pi_1(Z_k \cap Z_{k+1}) = \mathbb{Z}$ and setting $\mu_{Symp}(\tilde{\Psi}_{k,k+1}) := \mu_{Symp}(\tilde{\Psi}_{k,k+1} \circ \gamma_{k,k+1})$. Now pick a loop B in $Sp(2n,\mathbb{R})$ with $\mu_{Symp}(B) = \mu_{Symp}(\tilde{\Psi}_{k,k+1})$ and replace $\tilde{\Phi}_{k+1}(r,\vartheta)$ by $\Phi_{k+1}(r,\vartheta) := \tilde{\Phi}_{k+1}(r,\vartheta) \circ B(\vartheta)^{-1}$, then

$$\Psi_{k,k+1} := \tilde{\Phi}_k^{-1} \circ \Phi_{k+1} = \tilde{\Phi}_k^{-1} \circ \tilde{\Phi}_{k+1} \circ B^{-1}.$$

Clearly $\mu_{Symp}(\Psi_{k,k+1}) = \mu_{Symp}(\tilde{\Psi}_{k,k+1}) - \mu_{Symp}(B) = 0$ and we can define a map

$$\begin{aligned} A_{k,k+1} \; : \; [a_{k+1}, b_k] \times \mathbb{R}/2\pi\mathbb{Z} \quad &\to \quad Sp(2n, \mathbb{R}) \\ (r, \vartheta) \quad &\mapsto \quad \begin{cases} 1\!\!1 & \text{, for } r \text{ near } a_{k+1}, \\ \Psi_{k,k+1}(r,\vartheta) & \text{, for } r \text{ near } b_k. \end{cases} \end{aligned}$$

Then $\Phi(r,\vartheta) := \tilde{\Phi}_k(r,\vartheta) \circ A(r,\vartheta)$ agrees with $\tilde{\Phi}_k$ near a_{k+1} and with Φ_{k+1} near b_k. This way we get a unitary trivialization over $Z_k \cup Z_{k+1} = [a_k, b_{k+1}]$. For $k = m-1$ we end up this way with a trivialization $\Phi_1 : \mathbb{R}/2\pi\mathbb{Z} \times \mathbb{R}^{2n} \to \gamma_1^* E$ over the second boundary component. We see that Φ_1 is determined by Φ_0 up to homotopy: we can arrange to end up with any $\Phi_1 \circ B$ where $\mu_{Symp}(B) = 0$. □

First Chern number

Since the set of isomorphism classes of symplectic and complex vector bundles coincide, these are both characterized by the Chern classes. We are only interested in the first Chern class c_1, which is an element of the integral 2-dimensional cohomology of the base manifold. For bundles over 2-dimensional bases, c_1 is completely described by the *first Chern number*, which is the value taken by c_1 on the fundamental 2-cycle of the base.

THEOREM B.1.14. *(*[**MS95**] *thm. 2.67) There exists a unique functor c_1, called the* first Chern number*, which assigns an integer $c_1(E) \in \mathbb{Z}$ to every symplectic vector bundle E over a compact oriented Riemann surface Σ without boundary and satisfies the following axioms*

(naturality) $(E, \omega) \simeq (E', \omega') \Leftrightarrow rk\, E = rk\, E'$ *and* $c_1(E) = c_1(E')$.

(functoriality) *For any smooth map $\phi : \Sigma' \to \Sigma$ of oriented Riemann surfaces and any symplectic vector bundle $E \to \Sigma$*

$$c_1(\phi^* E) = deg\,(\phi) \cdot c_1(E).$$

(additivity) *For any two symplectic vector bundles $E_1 \to \Sigma$ and $E_2 \to \Sigma$*

$$c_1(E_1 \oplus E_2) = c_1(E_1 \otimes E_2) = c_1(E_1) + c_1(E_2).$$

(normalization) $c_1(T\Sigma) = 2 - 2g$*, where g is the genus of Σ.*

REMARK B.1.15. ([**MS95**] remark 2.68) If E is a symplectic vector bundle over any manifold N then the first Chern number assigns an integer $c_1(f^*E)$ to every smooth map $f : \Sigma \to N$ where Σ is a compact oriented Riemann surface without boundary. The axioms imply that this integer depends only on the homology class of f. Thus the first Chern number generalizes to an integral cohomology class

$$c_1(E) \in H^2(N, \mathbb{Z})$$

which is called *first Chern class*.

Theorem B.1.14 is proven by explicitly defining c_1: let Σ be a compact oriented Riemann surface without boundary, choose a splitting

$$\Sigma = \Sigma_1 \cup_C \Sigma_2$$

such that $\partial\Sigma_1 = \partial\Sigma_2 = C$. Orient the 1-manifold C as the boundary of Σ_1: a vector $v \in T_qC$ is positively oriented if $\{\nu(q), v\}$ is a positively oriented basis of $T_q\Sigma$ where $\nu : C \to T\Sigma$ is a normal vector field along C which points out of Σ_1.
Let E be the symplectic vector bundle over Σ and choose symplectic trivializations for $k = 1, 2$

$$\begin{array}{rcl} \Sigma_k \times \mathbb{R}^{2n} & \to & E \\ (q, \zeta) & \mapsto & \Phi_k(q)\zeta \end{array}$$

of E over Σ_1 and Σ_2. By proposition B.1.12 they exist as $\partial\Sigma_k \neq \emptyset$ for $k = 1, 2$.
The *overlap map* is defined by

$$\begin{array}{rcl} \Psi \ : \ C & \to & Sp(2n, \mathbb{R}) \\ q & \mapsto & \Phi_1(q)^{-1} \circ \Phi_2(q). \end{array}$$

Using the map (177) $\rho = det{\circ}I^*{\circ}r(1) : Sp(2n, \mathbb{R}) \to S^1$ we define $c_1(E)$ to be the degree of $\rho{\circ}\Psi : C \to S^1$

$$c_1(E) = deg\ \rho{\circ}\Psi = \sum_{j=1}^{l} \mu_{Symp}(\Psi{\circ}\gamma_j). \tag{180}$$

I.e. $c_1(E)$ equals the sum of the Maslov indices of the loops $\Psi{\circ}\gamma_j : \mathbb{R}/\mathbb{Z} \to Sp(2n, \mathbb{R})$, where l is the number of components of C and each component is parametrized by a loop $\gamma_j : \mathbb{R}/\mathbb{Z} \to C$ such that $\dot{\gamma}_j(t)$ is positively oriented.

Conley-Zehnder index of periodic orbits

We present two methods of canonically associating an integer to any nondegenerate 1-periodic Hamiltonian orbit $z \in Per\ H$. Although the first method is very much along the lines of the standard one in Floer theory, we need to use the second construction in the main part of this thesis as it allows for an comparison between the Fredholm operators $\mathcal{D}_u^0$ and $\mathcal{D}_w^\epsilon$ appearing there.

Method 1 Let $z \in \mathcal{P}er\ (H)$ be a 1-periodic integral trajectory of the Hamiltonian vector field X_{H_t} on T^*M. We identify $S^1 \cong \mathbb{R}/\mathbb{Z}$ and denote the local coordinate by t. Let $\Lambda_z T^*M$ be the connected component of the free loop space containing z. Pick a reference loop $z_0 \in \Lambda_z T^*M$ and a unitary trivialization

$$\Phi_{z_0} \ : \ S^1 \times \mathbb{R}^{2n} \ \to \ z_0^* TT^*M$$

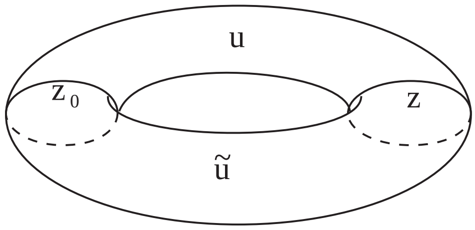

FIGURE B.2. The reference loop z_0 and $z \in \mathcal{P}er\ (H)$

(this exists as $U(n, \mathbb{C})$ is connected). Now choose a homotopy u between z_0 and z (the homotopy parameter is denoted by $s \in [0,1]$) and extend the trivialization of $z_0^* TT^*M$ to u^*TT^*M by Lemma B.1.13. We get an induced trivialization

$$\Phi_z \ : \ S^1 \times \mathbb{R}^{2n} \ \to \ z^*TT^*M.$$

Denote by φ_t the time-t-map induced by X_{H_t}, i.e. $z(t) = \varphi_t z^0$ where $z^0 = z(0)$, then define

$$\begin{aligned} A \ : \ [0,1] \ &\to Sp(2n, \mathbb{R}) \\ t \ &\mapsto \ \Phi_z(t)^{-1} \circ d\varphi_t \mid_{z^0} \ \circ \Phi_x(0). \end{aligned} \tag{181}$$

Clearly $A(0) = \mathbb{1}$ and $det\ (A(1) - \mathbb{1}) = det\ (d\varphi_1(z^0) - \mathbb{1}) \neq 0$; hence we may define the *Conley-Zehnder index of the Hamiltonian orbit* z by

$$\boxed{\mu_{CZ}(z) := \mu_{CZ}(A).} \tag{182}$$

First of all we show that this is independent of the choice of the homotopy u. Here the crucial point is our result that the first Chern number is zero. Let $\tilde{u}$ be another homotopy between z_0 and z with corresponding unitary trivializations $\tilde{\Phi}_{z_0} = \Phi_{z_0}$ and $\tilde{\Phi}_z$. u and $\tilde{u}$ fit together to form a torus $\mathbb{T}^2 = u \#_{z_0,z} \tilde{u} \subset T^*M$ (figure B.2). Now

$$\begin{aligned} 0 = c_1(T_{\mathbb{T}^2} T^*M) &= \mu_{Symp}(\Phi_{x_0}^{-1} \circ \tilde{\Phi}_{z_0}) - \mu_{Symp}(\Phi_z^{-1} \circ \tilde{\Phi}_z) \\ &= \mu_{Symp}(\mathbb{1}) - \mu_{Symp}(\Phi_z^{-1} \circ \tilde{\Phi}_z) = -\mu_{Symp}(\Psi_{z,u,\tilde{u}}) \end{aligned}$$

where the minus sign comes in as we need to run through z in the opposite direction; the last equality is just the definition of the transition map $\Psi_{z,u,\tilde{u}} := \Phi_z^{-1} \circ \tilde{\Phi}_z$. Using this we get $\tilde{\Phi}_z = \Phi_z \Psi_{z,u,\tilde{u}}$ and therefore

$$\begin{aligned} \mu_{CZ}(\tilde{A}) &= \mu_{CZ}(\tilde{\Phi}_z(t)^{-1}\ d\varphi_t z^{(0)}\ \tilde{\Phi}_z(0)) \\ &= \mu_{CZ}(\Psi_{z,u,\tilde{u}}(t)^{-1}\ \Phi_z(t)^{-1}\ d\varphi_t z^{(0)}\ \Phi_z(0)\ \Psi_{z,u,\tilde{u}}(0)) \\ &= 2\mu_{Symp}(\Psi_{z,u,\tilde{u}}^{-1}) + \mu_{CZ}(\Phi_z(t)^{-1}\ d\varphi_t z^{(0)}\ \Phi_z(0)) + 0 = \mu_{CZ}(A). \end{aligned}$$

Note that in the third equality we used the formula

$$\mu_{CZ}(\Psi \circ A) = 2\mu_{Symp}(\Psi) + \mu_{CZ}(A)$$

from [**DS94b**] (property LOOP) for the composition of a path A and a loop Ψ in $Sp(2n,\mathbb{R})$.

We observe that our definition of the Conley-Zehnder index depends on the choice of the trivialization Φ_{z_0}: These trivializations are characterized up to homotopy $\pi_1(U(n,\mathbb{C}))$: the transition map between two trivializations represents an element of the fundamental group. Our construction is independent of the reference loop as long as we extend Φ_{z_0} along a homotopy u from z_0 to a new reference loop $\tilde{z}_0$. Differences of Conley-Zehnder indices of periodic orbits in the same connected component of the loop space, however, are well-defined in any case. Therefore our construction suffices to express canonically, as required in standard Floer theory, a certain Fredholm index as the difference of the Conley-Zehnder indices of two homotopic elements of $Per\ H$.

We remark that in the special case of a contractible periodic solution z, our construction indeed reduces to the standard one described in [**SZ92**] section 5: span in a disc $u : D^2 \to T^*M$ and trivialize $u^*TT^*M \to D^2$. Any two such unitary trivializations $\Phi, \tilde{\Phi}$ restricted to $S^1 = \partial D^2$ are homotopic as $\tilde{\Phi}^{-1} \circ \Phi : S^1 \to U(2n,\mathbb{R})$ is smoothly homotopic to the identity (as the first Chern class is zero). So we have a natural trivialization at any periodic orbit.

METHOD 2 In the main part of this thesis we would like to compare the Conley-Zehnder index of $z \in \mathcal{P}er(H)$ with the Morse index of the underlying perturbed geodesic. As the latter is a well-defined integer, there should be a natural choice for the trivialization Φ_{z_0}. Now to trivialize the Jacobi-operator (187) we need to assume that M is *orientable.* Let $z \in \Lambda T^*M$ and $x \in \Lambda M$ be its base component. Choose any orthogonal trivialization

$$\phi_x \ : \ S^1 \times \mathbb{R}^n \ \to \ x^*TM$$

which exists, precisely because M is orientable and $SO(n,\mathbb{R})$ is connected. Let $\phi_x^{*\,-1}$ be the dual trivialization, then define

$$\Phi_{z_0}(t) = \begin{pmatrix} \phi_x & 0 \\ 0 & \phi_x^{*-1} \end{pmatrix} \ : \ \mathbb{R}^n \oplus \mathbb{R}^{n\,*} \ \to \ x^*TM \oplus x^*T^*M$$

and $A \in C^0([0,1], Sp(2n,\mathbb{R}))$ as in (181). Another orthogonal trivialization $\tilde{\phi}_x$ leads to transition maps $\psi_x(t) = \phi_x^{-1}(t)\, \tilde{\phi}_x(t) \in O(n,\mathbb{R})$, hence

$$\tilde{\Phi}_{z_0} = \begin{pmatrix} \tilde{\phi}_x & 0 \\ 0 & \tilde{\phi}_x^{*-1} \end{pmatrix} = \begin{pmatrix} \phi_x & 0 \\ 0 & \phi_x^{*-1} \end{pmatrix} \begin{pmatrix} \psi_x & 0 \\ 0 & \psi_x \end{pmatrix} = \Phi_{z_0} \circ \Psi.$$

Now as before $\mu_{CZ}(\tilde{A}) = \mu_{CZ}(A)$ because the transition map $\Psi(t) \in Sp(2n,\mathbb{R}) \cap O(2n,\mathbb{R})$ has block diagonal form: therefore $\Psi(t)$ lies entirely in the stratum $Sp_n(2n,\mathbb{R})$ and so $\mu_{Symp}(\Psi) = 0$ by [**RS93**] theorem 4.1 (property ZERO). Note that we have reduced the structure group of $x^*TM \oplus x^*T^*M$ to $O(n,\mathbb{R})$. The underlying principle is the splitting in two Lagrangian subbundles.

In the *nonorientable case* we proceed as above on those connected components of ΛM consisting of loops x such that x^*TM admits an orthonormal trivialization. On the others we may find an orthonormal trivialization ϕ_x over $[0,1]$ with boundary condition $\phi_x(0)^{-1} \circ \phi_x(1) = diag(-1, 1, \dots, 1)$.

B.2. The classical action functional

Let (M,g) be a closed Riemannian manifold. We introduce the energy functional of Riemannian geometry and call it *classical action functional* throughout this text. The first and second variational formulae are derived in B.2.1. Its critical points are the (perturbed) closed geodesics and its Hessian at a critical point gives rise to the (perturbed) Jacobi operator. This is a selfadjoint operator on an appropriate L^2-Hilbert space; the dimension of the largest negative definite subspace is finite as we will se in B.2.2 and is called the Morse index of the critical point.

Let $\Lambda M = W^{1,2}(\mathbb{R}/\mathbb{Z}, M)$ be the *free loop space of* M. For $\gamma \in C^\infty(S^1, M)$ we denote by $\Gamma(\gamma^*TM)$ the smooth sections of the vector bundle γ^*TM and $W^{1,2}(\gamma^*TM)$ denotes, for $\gamma \in \Lambda M$, the completion of $\Gamma(\gamma^*TM)$ with respect to the norm $\|\cdot\|_{1,2}$.

DEFINITION B.2.1. Let $V \in C^\infty(\mathbb{R}/\mathbb{Z} \times M, \mathbb{R})$ and

$$\begin{aligned} \mathcal{I}_V : \Lambda M &\to \mathbb{R} \\ \gamma &\mapsto \int_0^1 \tfrac{1}{2} g(\dot\gamma(t), \dot\gamma(t)) - V(t, \gamma(t))\, dt\ . \end{aligned}$$

We call $\mathcal{I}_V$ the *classical action functional* (in Riemannian geometry $\mathcal{I}_0$ is called *energy functional*). Its integrand $L : \mathbb{R}/\mathbb{Z} \times TM \to \mathbb{R}$ is called the *Lagrangian*; the Lagrangian is the Legendre transform of the Hamiltonian $H : \mathbb{R}/\mathbb{Z} \times T^*M \to \mathbb{R}$.

On $W^{1,2}(\gamma^*TM)$ we have two inner products, namely the L^2*-inner product*

$$\langle \xi_1, \xi_2 \rangle_{0,2} = \int_0^1 g(\xi_1(t), \xi_2(t))\, dt \tag{183}$$

and the $W^{1,2}$*-inner product*

$$\langle \xi_1, \xi_2 \rangle_{1,2} = \int_0^1 g(\xi_1(t), \xi_2(t)) + g(\nabla_t \xi_1(t), \nabla_t \xi_2(t))\, dt\ . \tag{184}$$

PROPOSITION B.2.2. (**partial integration**) *Let* $\xi_1, \xi_2 \in \Gamma(\gamma^*TM)$, *then*

$$\langle \nabla_t \xi_1, \xi_2 \rangle_{0,2} = \langle \xi_1, -\nabla_t \xi_2 \rangle_{0,2}\ .$$

PROOF. In natural local coordinates we have (we drop the argument t in our notation)

$$\langle \nabla_t \xi_1, \xi_2 \rangle_{0,2} = \int_0^1 g_{ij}(\gamma) \left(\dot\xi_1^i + \Gamma^i_{kl}(\gamma) \xi_1^k \dot\gamma^l \right) \xi_2^j\, dt\ ,$$

$$
\begin{aligned}
\langle \xi_1, -\nabla_t \xi_2 \rangle_{0,2} &= \int_0^1 -g_{ij}(\gamma)\xi_1^i \left(\dot{\xi}_2^j + \Gamma_{kl}^j(\gamma)\xi_2^k \dot{\gamma}^l \right) \, dt \\
&= \int_0^1 \left(\frac{\partial g_{ij}(\gamma)}{\partial \gamma^r} \dot{\gamma}^r \xi_1^i \xi_2^j + g_{ij}(\gamma)\dot{\xi}_1^i \xi_2^j - g_{ij}(\gamma)\xi_1^i \Gamma_{kl}^j(\gamma)\xi_2^k \dot{\gamma}^l \right) \, dt.
\end{aligned}
$$

The second term is fine, using (95) twice we observe that the first and third together equal

$$
\int_0^1 \xi_1^k \xi_2^j \dot{\gamma}^l g_{ij}(\gamma) \Gamma_{lk}^i(\gamma) \, dt \ ,
$$

hence we are done. □

B.2.1. First and second variation formulae.

LEMMA B.2.3. (**1. variation formula**) *Let* $\gamma \in C^\infty(S^1, M)$, $\xi \in \Gamma(\gamma^* TM)$, *then*

$$
d\mathcal{I}_V(\gamma)\, \xi = \int_0^1 g\left(-\nabla_t \partial_t \gamma(t) - {}^g\nabla V(t, \gamma(t)), \xi(t)\right) \, dt \ .
$$

PROOF. Let γ_τ, $\tau \in (-\epsilon, \epsilon)$, $\epsilon > 0$ small be a *variation of* γ, i.e. $\gamma(t,\tau) = \gamma_\tau(t)$ is differentiable, $\gamma_0 = \gamma$ and $\frac{d}{d\tau}\gamma_\tau \mid_{\tau=0} = \xi$ (cf. [**Jo95**], Section 4.1), then the LHS is defined to be

$$
\begin{aligned}
\frac{d}{d\tau}\mathcal{I}_V(\gamma_\tau)\Big|_{\tau=0} &= \frac{d}{d\tau}\Big|_{\tau=0} \int_0^1 \tfrac{1}{2} g(\dot{\gamma}_\tau(t), \dot{\gamma}_\tau(t)) - V(t, \gamma_\tau(t)) \, dt \\
&= \int_0^1 \frac{d}{d\tau}\left(\tfrac{1}{2} g_{ij}(\gamma_\tau)\dot{\gamma}_\tau^i \dot{\gamma}_\tau^j\right)\Big|_{\tau=0} - \frac{\partial V}{\partial \gamma^i}\xi^i \, dt \\
&= \int_0^1 \tfrac{1}{2}\frac{\partial g_{ij}(\gamma)}{\partial \gamma^l}\xi^l \dot{\gamma}^i \dot{\gamma}^j + g_{ij}(\gamma)\dot{\xi}^i \dot{\gamma}^j - g_{ls}(\gamma) g^{il}(\gamma)\frac{\partial V}{\partial \gamma^i}\xi^s \, dt \\
&= \int_0^1 g_{ik}(\gamma)\Gamma_{jl}^k(\gamma)\xi^l \dot{\gamma}^i \dot{\gamma}^j + g_{ji}(\gamma)\dot{\xi}^j \dot{\gamma}^i - g_{ls}(\gamma)({}^g\nabla V(t,\gamma))^l \xi^s \, dt \\
&= \int_0^1 g(\dot{\gamma}, \nabla_t \xi) - g({}^g\nabla V(\gamma), \xi) \, dt \\
&= \langle -\nabla_t \partial_t \gamma - {}^g\nabla V(t,\gamma), \xi \rangle_{0,2}
\end{aligned}
$$

where in the 4^{th} equality it is easier to proceed in the opposite direction using (95) and renaming of indices (as there is a symmetry in indices i, j). The last equality follows from Proposition B.2.2 (partial integration). □

Note that without using proposition B.2.2 our result was

(185) $$ d\mathcal{I}_V(\gamma)\, \xi = \langle \partial_t \gamma, \nabla_t \xi \rangle_{0,2} + \langle -{}^g\nabla V(t,\gamma), \xi \rangle_{0,2} \ . $$

This continues to hold if $\gamma \in \Lambda M$ and $\xi \in W^{1,2}(\gamma^* TM)$.

LEMMA B.2.4. *(cf. [**Jo95**], lemma 7.2.1)*

$$\gamma \in \Lambda M \ \text{with}\ d\mathcal{I}_V(\gamma) \equiv 0 \ \text{on}\ W^{1,2}(\gamma^*TM)$$
$$\Longleftrightarrow \gamma \in C^\infty(S^1, M) \ \text{and}\ -\nabla_t \partial_t \gamma - {}^g\nabla V(t,\gamma) \equiv 0 \ .$$

DEFINITION B.2.5. i) $\gamma \in C^\infty(S^1, M)$ with $\nabla_t \partial_t \gamma \equiv 0$ is called a *closed geodesic*. We call $\gamma \in C^\infty(S^1, M)$ with

$$-\nabla_t \partial_t \gamma - {}^g\nabla V(t,\gamma) \equiv 0$$

a *perturbed closed geodesic*.
ii) Let $\gamma \in \Lambda M$, then we define the *L^2-gradient of $\mathcal{I}_V$ at γ* by

$$d\mathcal{I}_V(\gamma)\,\xi = \langle L^2\text{-}grad\,\mathcal{I}_V(\gamma), \xi\rangle_{0,2} \quad , \quad \forall \xi \in W^{1,2}(\gamma^*TM).$$

iii) Finally define $Crit\,\mathcal{I}_V = \{\gamma \in \Lambda M \mid d\mathcal{I}_V(\gamma) \equiv 0 \text{ on } W^{1,2}(\gamma^*TM)\}$.

Hence if $\gamma \in C^\infty(S^1, M)$ we have

$$L^2\text{-}grad\,\mathcal{I}_V(\gamma) = -\nabla_t \partial_t \gamma - {}^g\nabla V(t,\gamma) \ . \tag{186}$$

LEMMA B.2.6. (**2. variation formula**) *For any loop $\gamma \in C^\infty(S^1, M)$ with L^2-grad $\mathcal{I}_V(\gamma) = 0$ and any $\xi_1, \xi_2 \in \Gamma(\gamma^*TM)$ it holds*

$$\begin{aligned} &Hess\,\mathcal{I}_{V}(\gamma)\,(\xi_1,\xi_2) \\ &:= d^2\mathcal{I}_{V}(\gamma)\,(\xi_1,\xi_2) \\ &= \int_0^1 g\left(-\nabla_t\nabla_t\xi_1 - R(\xi_1,\dot\gamma)\dot\gamma - \nabla_{\xi_1}{}^g\nabla V(t,\gamma), \xi_2\right)\,dt\ . \end{aligned}$$

PROOF. Let γ_τ be a variation of γ as above,

$$\begin{aligned} d^2\mathcal{I}_V(\gamma)(\xi_1,\xi_2) &= \frac{d}{d\tau} d\mathcal{I}_{V}(\gamma_\tau)\xi_1\Big|_{\tau=0} \\ &= \int_0^1 \frac{d}{d\tau} g(\gamma_\tau)(-\nabla_t\partial_t\gamma_\tau - {}^g\nabla V(t,\gamma_\tau),\xi_1)\Big|_{\tau=0}\,dt\ . \end{aligned}$$

Now define $\mathcal{F}(\gamma_\tau) = -\nabla_t\partial_t\gamma_\tau - {}^g\nabla V(t\gamma_\tau)$, hence we have to linearize the map $\mathcal{F}$ at a zero (and therefore we can neglect the term involving $\frac{d}{d\tau}g(\gamma_\tau)\,|_{\tau=0}$). For simplicity of notation let $\xi = \xi_2$, then

$$\begin{aligned} -(d\mathcal{F}(\gamma)\xi)^k &= -\frac{d}{d\tau}\mathcal{F}(\gamma_\tau)^k\Big|_{\tau=0} \\ &= \frac{d}{d\tau}\Big|_{\tau=0}\left(\ddot\gamma_\tau^k + \Gamma_{ij}^k(\gamma_\tau)\dot\gamma_\tau^i\dot\gamma_\tau^j + g^{kj}(\gamma_\tau)\frac{\partial V(t,\gamma_\tau)}{\partial\gamma^j}\right) \\ &= \ddot\xi^k + \frac{\Gamma_{ij}^k(\gamma)}{\partial\gamma^l}\xi^l\dot\gamma^i\dot\gamma^j + 2\Gamma_{ij}^k(\gamma)\dot\xi^i\dot\gamma^j + \frac{\partial g^{kj}(\gamma)}{\partial\gamma^l}\xi^l\frac{\partial V(t,\gamma)}{\partial\gamma^j} \\ &\quad + g^{kj}(\gamma)\frac{\partial^2 V(t,\gamma)}{\partial\gamma^l\partial\gamma^j}\xi^l\ . \end{aligned}$$

As we already know where we would like to end up, we are now going backwards: Using $({}^g\nabla V(\gamma))^k = g^{kl}(\gamma)\frac{\partial V(\gamma)}{\partial \gamma^l}$

$$\begin{aligned}
&(\nabla_t\nabla_t\xi)^k + (R(\xi,\dot\gamma)\dot\gamma)^k + (\nabla_\xi {}^g\nabla V(t,\gamma))^k \\
&= \partial_t(\nabla_t\xi)^k + \Gamma^k_{ij}\dot\gamma^i(\nabla_t\xi)^j + R^k_{lij}\xi^i\dot\gamma^j\dot\gamma^l + \xi^i\frac{\partial g^{kl}}{\partial\gamma^i}\frac{\partial V}{\partial \gamma^l} \\
&\quad + \xi^i g^{kl}\frac{\partial^2 V}{\partial\gamma^l\partial\gamma^i} + g^{jl}\Gamma^k_{ij}\xi^i\frac{\partial V}{\partial\gamma^l} \\
&= \Big(\ddot\xi^k + \frac{\partial\Gamma^k_{ij}}{\partial\gamma^l}\dot\gamma^l\dot\gamma^i\xi^j + \Gamma^k_{ij}(-\Gamma^i_{rs}\dot\gamma^r\dot\gamma^s - g^{il}\frac{\partial V}{\partial\gamma^l})\xi^j + \Gamma^k_{ij}\dot\gamma^i\dot\xi^j\Big) \\
&\quad + \Big(\Gamma^k_{ij}\dot\gamma^i\dot\xi^j + \Gamma^k_{ij}\Gamma^j_{rs}\dot\gamma^i\dot\gamma^r\xi^s\Big) \\
&\quad + \Big(\frac{\partial\Gamma^k_{jl}}{\partial\gamma^i} - \frac{\partial\Gamma^k_{il}}{\partial\gamma^j} + \Gamma^k_{i\nu}\Gamma^\nu_{jl} - \Gamma^k_{j\nu}\Gamma^\nu_{il}\Big)\xi^i\dot\gamma^j\dot\gamma^l \\
&\quad + \xi^i\frac{\partial g^{kl}}{\partial\gamma^i}\frac{\partial V}{\partial\gamma^l} + \xi^i g^{kl}\frac{\partial^2 V}{\partial\gamma^l\partial\gamma^i} + g^{jl}\Gamma^k_{ij}\xi^i\frac{\partial V}{\partial\gamma^l}\ .
\end{aligned}$$

In the last equality we used $\mathcal{F}(\gamma) = 0$. Now terms $1, (5+6), 8, 12, 13$ are the ones we are looking for. The remaining terms cancel in pairs: $(2+9), (3+10), (4+14), (7+11)$. ☐

Proposition B.2.7. *The bilinear form*

$$d^2\mathcal{I}_V(\gamma)\,(\cdot,\cdot) : \Gamma(\gamma^*TM)\times\Gamma(\gamma^*TM)\to\mathbb{R}$$

is symmetric.

Proof. We have to check that $\forall \xi_1,\xi_2\in\Gamma(\gamma^*TM)$

$$\begin{aligned}
&\int_0^1 g(-\nabla_t\nabla_t\xi_1 - R(\xi_1,\dot\gamma)\dot\gamma - \nabla_{\xi_1}{}^g\nabla V(t,\gamma),\xi_2)\,dt \\
&= \int_0^1 g(\xi_1, -\nabla_t\nabla_t\xi_2 - R(\xi_2,\dot\gamma)\dot\gamma - \nabla_{\xi_2}{}^g\nabla V(t,\gamma),\xi_2)\,dt\ .
\end{aligned}$$

The first term is fine which follows from partially integrating twice, Proposition B.2.2. The curvature term is also fine

$$-g(R(\xi_1,\dot\gamma)\dot\gamma,\xi_2) = -g(R(\dot\gamma,\xi_2)\xi_1,\dot\gamma) = -g(R(\xi_2,\dot\gamma)\dot\gamma,\xi_2)\ .$$

The first equality uses the identity (99), in the second we get two minus signs. It remains to check the potential term

$$\begin{aligned}
g(\nabla_{\xi_1}{}^g\nabla V(t,\gamma),\xi_2) &= \xi_1\xi_2 V(t,\gamma) - (\nabla_{\xi_1}\xi_2)V(t,\gamma) \\
&= \xi_2\xi_1 V(t,\gamma) - (\nabla_{\xi_2}\xi_1)V(t,\gamma) \\
&= g(\nabla\xi_2{}^g\nabla V(t,\gamma),\xi_1)\ .
\end{aligned}$$

The second equality follows from the connection being torsion free, we also used $g({}^g\nabla V(t,\gamma),\xi) = dV(t,\gamma)\xi = \xi V(t,\gamma)$ several times. In local coordinates the integrand is

$$
\begin{aligned}
&g_{ir}\xi_1^l\Big(\frac{\partial}{\partial\gamma^l}(g^{ij}\frac{\partial V}{\partial\gamma^j}) + \Gamma^i_{ls}g^{sj}\frac{\partial V}{\partial\gamma^j}\Big)\xi_2^r\\
&= \Big(\frac{\partial^2 V}{\partial\gamma^l\partial\gamma^r} - \Gamma^j_{lr}\frac{\partial V}{\partial\gamma^j}\Big)\xi_1^l\xi_2^r\ ,
\end{aligned}
$$

where we applied (95) as well as (114). The first factor is symmetric in l and r. □

B.2.2. Morse index. We may define a linear map $A_\gamma : \Gamma(\gamma^*TM) \to \Gamma(\gamma^*TM)$ – called the *perturbed Jacobi operator* – by setting $\forall \xi_1,\xi_2 \in \Gamma(\gamma^*TM)$

$$
\begin{aligned}
d^2\mathcal{I}_V(\gamma)\,(\xi_1,\xi_2) &= \langle -\nabla_t\nabla_t\xi_1 - R(\xi_1,\dot\gamma)\dot\gamma - \nabla_{\xi_1}{}^g\nabla V(t,\gamma),\xi_2\rangle_{0,2}\\
&=: \langle A_\gamma\xi_1,\xi_2\rangle_{0,2}\ .
\end{aligned}
\tag{187}
$$

Symmetry of $d^2\mathcal{I}_V(\gamma)$ implies symmetry of A_γ with respect to the L^2-inner product $\langle\cdot,\cdot\rangle_{0,2}$. Now we may take the closure of $\Gamma(\gamma^*TM)$ with respect to the L^2-inner product (completion, cf. [**RS1**], thm I.3) and denote it by $L^2(\gamma^*TM)$.

As A_γ is symmetric on $\Gamma(\gamma^*TM)$ and this is a dense subset of $L^2(\gamma^*TM)$, A_γ is closable ([**RS1**], section VIII.2). The closure is denoted by $\overline{A_\gamma}$ and it is defined on $D(\overline{A_\gamma}) \subset L^2(\gamma^*TM)$. One may think of $\overline{A_\gamma}$ as the operator whose graph is the closure of

$$
graph\,A_\gamma = \{(\xi, A_\gamma\,\xi) \mid \xi\in\Gamma(\gamma^*TM)\} \subset L^2(\gamma^*TM)\times L^2(\gamma^*TM) \tag{188}
$$

with respect to the norm on $L^2(\gamma^*TM)\times L^2(\gamma^*TM)$ given by

$$
\|(\xi_1,\xi_2)\|_{L^2(\gamma^*TM)\times L^2(\gamma^*TM)} = \|\xi_1\|_2 + \|\xi_2\|_2\ . \tag{189}
$$

As $\|(\xi,A_\gamma\xi)\|_{L^2(\gamma^*TM)\times L^2(\gamma^*TM)}$ is equivalent to the $W^{2,2}$-norm of ξ, we observe that A_γ is a bounded operator

$$
\overline{A_\gamma} : D(\overline{A_\gamma}) = W^{2,2}(\gamma^*TM) \to L^2(\gamma^*TM)\ . \tag{190}
$$

As it suffices to check properties of a closed operator on a dense subset, we see that $\overline{A_\gamma}$ is symmetric. Moreover, as the first term $-\nabla_t\nabla_t$ of A_γ is a Laplacian, which is known to be selfadjoint on $L^2(\gamma^*TM)$ with dense domain $W^{2,2}(\gamma^*TM)$, and the curvature and potential terms (being bounded operators) are $-\nabla_t\nabla_t$-bounded with relative bound 0, the Kato-Rellich theorem (cf. [**RS2**], theorem X.12) implies that $\overline{A_\gamma}$ is selfadjoint. Hence $\overline{A_\gamma}$ has real eigenvalues. In what follows we denote $\overline{A_\gamma}$ for simplicity by A_γ.

THEOREM B.2.8. (**Morse index theorem**) *Let $\gamma\in Crit\,\mathcal{I}_V$, then the dimension of the largest subspace of $W^{2,2}(\gamma^*TM)$ on which $d^2\mathcal{I}_V(\gamma)\,(\cdot,\cdot)$ is negative definite is finite. We call this number $Ind\,(\gamma)$ the* Morse index of γ. *The dimension of the largest subspace on which* $d^2\mathcal{I}_V(\gamma)\,(\cdot,\cdot)$ *vanishes* *Null* (γ) *is called* nullity of γ *and is also finite.*

PROOF. The following beautyful and surprisingly direct proof may be found in [**Jo95**], Lemma 4.3.2. Assume by contradiction that there is an infinite dimensional subspace W of the domain $W^{2,2}(\gamma^*TM)$ on which $d^2\mathcal{I}_V(\gamma)(\cdot,\cdot)$ is negative semidefinite. Recall that our Hilbert space is $L^2(\gamma^*TM)$ and $W^{2,2}(\gamma^*TM)$ is the dense domain on which $d^2\mathcal{I}_V(\gamma)(\cdot,\cdot)$ and A_γ are well-defined. Now let $\{X_i\}_{i\in\mathbb{N}}$ be a set of elements of W such that

$$\langle X_i, X_j\rangle_{0,2} = \delta_{ij}\ .$$

Using our assumption we get

$$\begin{aligned} 0 &\geq d^2\mathcal{I}_V(\gamma)(X_n,X_n) \\ &= \langle \nabla_t X_n, \nabla_t X_n\rangle_{0,2} \\ &\quad + \langle -R(X_n,\dot\gamma)\dot\gamma, X_n\rangle_{0,2} + \langle -\nabla_{X_n}{}^g\nabla V(t,\gamma), X_n\rangle_{0,2} \end{aligned}$$

and hence

$$\begin{aligned} \|\nabla_t X_n\|_{0,2} &\leq \mid \langle R(X_n,\dot\gamma)\dot\gamma, X_n\rangle_{0,2} \mid + \mid \langle \nabla_{X_n}{}^g\nabla V(t,\gamma), X_n\rangle_{0,2} \mid \\ &\leq \|R\|\cdot\|\dot\gamma\|_{0,2}\cdot\underbrace{\|X_n\|_{0,2}}_{=1} + \|D^g\nabla V(t,\gamma)\|\cdot\underbrace{\|X_n\|_{0,2}^2}_{=1} = const\ . \end{aligned}$$

Here $\|R\|$ is the norm of the linear operator $R(X_n,\dot\gamma) : T_{\gamma(t)}M \to T_{\gamma(t)}M$ integrated over the compact manifold M, similar for $\|D^g\nabla V(t,\gamma)\|$. Now

$$\|X_n\|_{1,2}^2 = \|X_n\|_{0,2}^2 + \|\nabla_t X_n\|_{0,2}^2 \leq 1 + const\ ,$$

hence by Rellich's theorem a subsequence converges in $L^2(\gamma^*TM)$. On the other hand this is impossible as we started with an L^2-orthonormal sequence. □

APPENDIX C

A version of Newton's iteration method

We study a version of Newton's iteration method of finding a zero of a continuously differentiable map f given suitable a-priori data. The difference to the original method (figure C.1) is that we linearize f only once, at the starting point x_0 of the iteration process (figure C.2).

In section C.1 we begin with the case of real-valued functions in order to get familiar with the method in a simple setting. Section C.2 deals with the general case of maps between Banach spaces. Here a new subtlety arises: Namely the condition on the linearization D of f at x_0 to admit a right inverse. This turns out to be equivalent to surjectivity of D and the existence of a topological complement of *ker* D, which is satisfied for instance by any surjective Fredholm operator. In section C.3 we apply theorem C.2.9 on the Newton method to prove the inverse function theorem C.3.2 as well as the implicit function theorem C.3.4. The latter will, in the regular case, give the manifold properties and the dimension formulae for the moduli spaces under investigation $\mathcal{M}^0(x^-, x^+)$ and $\mathcal{M}^1(x^-, x^+)$ of parabolic respectively elliptic boundary value problems.

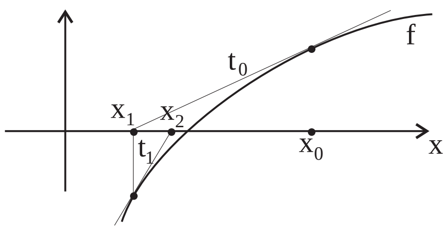

FIGURE C.1. The Newton method

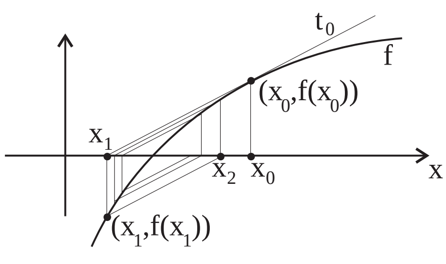

FIGURE C.2. The modified Newton method

C.1. Real-valued functions

We describe the Newton method of finding a zero of a real-valued continuously differentiable function f: Start with a point x_0 such that $f(x_0)$ is sufficiently small, $f'(x_0) \neq 0$ and the first derivatives of f do not vary too much in a neighbourhood of x_0. The idea is to find a zero of f nearby x_0 by solving the linear equation for ξ_0

$$t_0(\xi_0) = f(x_0) + f'(x_0)\xi_0 = 0$$

and then define the new starting point $x_1 = x_0 + \xi_0$. Now iterate this procedure by solving the equation $t_\nu(\xi_\nu) = f(x_\nu) + f'(x_\nu)\xi_\nu = 0$ for $\nu = 1, 2, 3, \ldots$, where $x_\nu = x_{\nu-1} + \xi_{\nu-1}$ (figure C.1). The assumptions on f guarantee convergence of the sequence $(x_\nu)_{\nu=0}^{\infty}$ to a point x_∞ with $f(x_\infty) = 0$.

In order to use throughout the same estimate $|f'(x_0)|^{-1} \leq c$, we slightly modify the method: We only use parallel translates of the tangent t_0. After finding the zero x_1 of t_0, we draw a line parallel to t_0 through the point $(x_1, f(x_1))$ and determine its zero x_2 and so on (figure C.2).

THEOREM C.1.1. *Let $f \in C^1(\mathbb{R}, \mathbb{R})$ and $x_0 \in \mathbb{R}$ such that there exist constants $\delta, c > 0$ with*

$$|f(x_0)| < \frac{\delta}{2c} \quad , \quad \frac{1}{|f'(x_0)|} \leq c \quad , \quad |f'(x) - f'(x_0)| < \frac{1}{2c}$$

whenever $|x - x_0| < \delta$. Then there exists a unique $\hat{x}$ with $|\hat{x} - x_0| \leq \delta$ and $f(\hat{x}) = 0$.

PROOF. The equation for the tangent t_0 in a new coordinate system with origin at x_0 is given by

$$t_0(y) = my + b = f'(x_0)y + f(x_0) \qquad \text{, its zero by} \quad \xi_0 = -\frac{f(x_0)}{f'(x_0)}.$$

With respect to the coordinate system x the zero of t_0 is given by $x_1 = x_0 + \xi_0$. The assumptions imply

$$\begin{aligned}
|\xi_0| &= \frac{|f(x_0)|}{|f'(x_0)|} \leq c|f(x_0)| \\
|\xi_0| &< \frac{\delta}{2} \\
|f(x_1)| &= |f(x_0 + \xi_0) \overbrace{-f(x_0) - f'(x_0)\xi_0}^{=0}| \\
&\leq \int_0^{\xi_0} |f'(x_0 + t) - f'(x_0)|\, dt \;<\; |\xi_0| \frac{1}{2c}.
\end{aligned}$$

Now proceed inductively by defining

$$\xi_\nu = -\frac{f(x_\nu)}{f'(x_0)} \qquad \text{and} \qquad x_{\nu+1} = x_\nu + \xi_\nu \ , \ \nu = 1, 2, 3, \ldots \ . \tag{191}$$

We have to show that

$$\begin{aligned} i) &\quad |\xi_\nu| \le c|f(x_\nu)| \\ ii) &\quad |\xi_\nu| < \frac{1}{2}|\xi_{\nu-1}| \\ iii) &\quad |f(x_{\nu+1})| < \frac{1}{2c}|\xi_\nu| \end{aligned}$$

imply the corresponding statements with ν replaced by $\nu+1$

$$\text{ad } i) \quad |\xi_{\nu+1}| \overset{(191)}{=} \frac{|f(x_{\nu+1})|}{|f'(x_0)|} \le c|f(x_{\nu+1})|$$

$$\text{ad } ii) \quad |\xi_{\nu+1}| \le c|f(x_{\nu+1})| \overset{iii)}{<} \frac{1}{2}|\xi_\nu|$$

$$\begin{aligned} \text{ad } iii) \quad |f(x_{\nu+2})| &= |f(x_{\nu+1}+\xi_{\nu+1}) \overbrace{-f(x_{\nu+1}) - f'(x_0)\xi_{\nu+1}}^{\overset{(191)}{=0}}| \\ &\le \int_0^{\xi_{\nu+1}} |f'(x_{\nu+1}+t) - f'(x_0)|\, dt \;<\; \frac{1}{2c}|\xi_{\nu+1}|. \end{aligned}$$

Note that $ii)$ implies $|x_\nu - x_0| < \delta$, $\forall \nu \in \mathbb{N}$, and $\hat{x} = \lim_{\nu\to\infty} x_\nu = x_0 + \sum_{k=0}^{\infty} \xi_k$ exists as we have absolute summability of the series

$$\sum_{\nu=0}^{N} \xi_\nu \le \sum_{\nu=0}^{N} |\xi_\nu| < \frac{\delta}{2}\sum_{\nu=0}^{N} \frac{1}{2^\nu} \overset{N\to\infty}{\longrightarrow} \frac{\delta}{2}\cdot\frac{1}{1-1/2} = \delta.$$

We get

$$|\hat{x} - x_0| = \lim_{\nu\to\infty} |x_\nu - x_0| = \lim_{\nu\to\infty} \left|\sum_{k=0}^{\nu} \xi_k\right| \le \delta$$

and

$$|f(\hat{x})| = \lim_{\nu\to\infty} |f(x_\nu)| \overset{iii)}{\le} \lim_{\nu\to\infty} \frac{1}{2c}\cdot|\xi_{\nu-1}| \le \lim_{\nu\to\infty} \frac{\delta}{4c}\cdot\frac{1}{2^{\nu-1}} = 0.$$

This proves existence. Uniqueness follows from the mean value theorem and the estimate

$$|f'(x)| \ge |f'(x_0)| - |f'(x) - f'(x_0)| \ge \frac{1}{2c} > 0$$

for all x with $|x - x_0| \le \delta$: assume there was another point $\tilde{x}$ with $|\tilde{x} - x_0| \le \delta$ and $f(\tilde{x}) = 0$, then there is x' between $\tilde{x}$ and $\hat{x}$ with $f'(x') = 0$ (mean value theorem), which contradicts the above estimate. □

C.2. Banach space-valued maps

We define the topological complement of a closed subspace of a Banach space and the right inverse of a bounded linear operator between Banach spaces. The existence of a right inverse T of a bounded linear operator D is equivalent to the existence of a topological complement of $ker\, D$ and surjectivity of D. It turns out that any surjective Fredholm operator admits a right inverse. Next we state a replacement of the *mean value theorem* needed in the uniqueness part of the proof of the main theorem. This theorem asserts the existence of a unique zero of a continuously differentiable map f between Banach spaces X and Y nearby an approximate zero x_0 under suitable conditions. A crucial condition is the existence of a right inverse of $df(x_0)$.

DEFINITION C.2.1. Let $W \subset X$ be a closed subspace of a Banach space X. A subspace $L \subset X$ is called *the topological complement of* W, if L is closed, $W \cap L = \{0\}$ and $W \oplus L = X$.

LEMMA C.2.2. *Any finite dimensional subspace V of a Banach space X admits a topological complement L.*

PROOF. (cf. [**Br83**], se. II.4) Let $\{e_1, \dots, e_n\}$ be a basis of V and write $v \in V$ as $v = \sum_{i=1}^{n} v^i e_i$. For $i = 1, \dots, n$ define continuous linear functionals $\varphi_i(v) = v^i$ and apply Hahn-Banach to extend φ_i to a continuous linear functional $\tilde{\varphi}_i : X \to \mathbb{R}$. Now we set

$$L = \bigcap_{i=1}^{n} (\tilde{\varphi}_i)^{-1}(0).$$

Clearly L is closed, $L \cap V = \{0\}$ and $L \oplus V \subset X$. It remains to show $X \subset L \oplus V$. Let $x \in X$ and write $x = x_L + x_V + z$, where $z \notin V$ and $z \notin L$. The latter implies there exists i with $(\tilde{\varphi}_i)^{-1}(z) = c \neq 0$, hence $z = c \cdot e_i \in V$, a contradiction. □

DEFINITION C.2.3. Let $D : X \to Y$ be a bounded linear operator between Banach spaces. $T : Y \to X$ is called *right inverse of* D, if $DT = id_Y$ and T is a bounded linear operator.

PROPOSITION C.2.4. *Let X, Y be Banach spaces. A bounded linear operator $D : X \to Y$ admits a right inverse T, if and only if D is surjective and $ker\, D$ admits a topological complement in X.*

PROOF. "$\Leftarrow$" Let X_1 denote the topological complement of $ker\, D$ in X, then $D_1 = D|_{X_1} : X_1 \to Y$ is a bijective bounded linear operator. Its inverse $(D_1)^{-1} : Y \to X_1$ is bounded by the open mapping theorem. We define $T = i \circ (D_1)^{-1}$, where $i : X_1 \to X$ is the inclusion.
"$\Rightarrow$" Let $T \in L(Y, X)$ be the right inverse of D. $DT = id_Y$ implies D surjective and T injective. $X_1 := im\, T$ is the required topological complement of $ker\, D$ in X:

i) im T closed: Let $(x_\nu)_{\nu\in\mathbb{N}} \subset X_1 = im\,T$ be a Cauchy sequence in X, then $x_\nu = Ty_\nu$ for a unique element y_ν. Now $(y_\nu)_{\nu\in\mathbb{N}}$ is a Cauchy sequence in Y:

$$\|y_\nu - y_\mu\|_Y \overset{DT=id_Y}{=} \|DT(y_\nu - y_\mu)\|_Y \leq \|D\| \cdot \|x_\nu - x_\mu\|_X.$$

Therefore $y_\nu \overset{\nu\to\infty}{\longrightarrow} y$ and $Ty_\nu \overset{\nu\to\infty}{\longrightarrow} Ty$ as T is continuous. On the other hand $Ty_\nu = x_\nu \overset{\nu\to\infty}{\longrightarrow} x$, so $x = Ty$.
ii) $Ker\,D \cap im T = \{0\}$: Let $x \in ker\,D \cap imT$, then $x = Ty$ and $y = DTy = Dx = 0$, hence $x = 0$.
iii) $ker\ D \oplus im\ T = X$: Let $\hat{D} : X/ker\ D \to Y$ denote the bijective linear operator induced by D. The composition $T\hat{D} : X/ker\,D \to imT$ is bijective. □

An immediate consequence is

COROLLARY C.2.5. *Any surjective Fredholm operator has a right inverse.*

DEFINITION C.2.6. Let X, Y be Banach spaces and $U \subset X$ open. A function $f : X \supset U \to Y$ is called *differentiable at* $x \in U$, if there is an element $f'(x) \in L(X, Y)$ such that

$$\|f(x+h) - f(x) - f'(x)h\| = o(\|h\|) \quad \text{as} \quad \|h\| \to 0.$$

The latter symbol means that

$$\frac{\|f(x+h) - f(x) - f'(x)h\|}{\|h\|} \to 0 \quad \text{as} \quad \|h\| \to 0.$$

If the LHS is only bounded we write $O(\|h\|)$. o and O are called *Landau symbols.*

Let now $I \subset \mathbb{R}$ be an open interval and $\gamma : I \to X$ be a curve in the Banach space X. The following Lemma is a *generalization of the mean value theorem* for real-valued functions.

LEMMA C.2.7. *i) If* $\gamma : I \to X$ *is differentiable at every point in* I*, then*

$$\|\gamma(s) - \gamma(t)\| \leq |s - t| \cdot \sup_{\tau\in[0,1]} \|\gamma'(t + \tau(s - t))\| \quad , \forall s, t \in I.$$

ii) If $\gamma : I \to X$ *is continuous in* $[t, s]$*, differentiable in* (t, s) *and* $v \in X$*, then*

$$\|\gamma(s) - \gamma(t) - v(s - t)\| \leq |s - t| \cdot \sup_{\tau\in(0,1)} \|\gamma'(t + \tau(s - t)) - v\|.$$

PROOF. ([**HöI**], thm. 1.1.1.) i) Fix s, $t \in I$ and let $\delta > 0$, $M_\delta = \delta + \sup_{\tau\in[0,1]} \|\gamma'(t + \tau(s - t))\|$ and set

$$E_\delta = \{\tau \in [0, 1] \mid \|\gamma(t + \tau(s - t)) - \gamma(t)\| \leq M_\delta \cdot \tau \cdot |t - s|\}.$$

γ continuous implies E_δ closed and therefore compact. As $0 \in E_\delta$, E_δ is nonempty and so has a largest element $\tau_{\max}$ (take the maximum of the continuous function $[0,1] \supset E_\delta \to \mathbb{R} : \tau \mapsto \tau$). If $\tau > \tau_{\max}$ and $\tau - \tau_{\max}$ is sufficiently small, we have

$$\begin{aligned}
&\|\gamma(t+\tau(s-t)) - \gamma(t)\| \\
&\le \|\gamma(t+\tau(s-t)) - \gamma(t+\tau_{\max}(s-t))\| + \|\gamma(t+\tau_{\max}(s-t)) - \gamma(t)\| \\
&\le M_\delta \cdot |(\tau-\tau_{\max})(s-t)| + M_\delta \cdot \tau_{\max} \cdot |s-t| = M_\delta \cdot \tau|s-t|.
\end{aligned}$$

The second inequality follows from the differentiability of γ and the fact that $\tau_{\max} \in E_\delta$. Hence $\tau \in E_\delta$, so $\tau_{\max} = 1$. As this holds for any $\delta > 0$ the result follows.
ii) We obtain the estimate in i) with supremum for $\tau \in (0,1)$ as a limit of i) applied to smaller closed intervals. If $v \in X$, application of i) to $\Gamma(t) = \gamma(t) - tv$ gives the result. □

COROLLARY C.2.8. *Let $f : X \to Y$ be a map between Banach spaces, differentiable on the line segment $[x,y] = \{x + \tau(y-x) \mid \tau \in [0,1]\}$. Then for any $S \in L(X,Y)$ it holds*

$$\|f(y) - f(x) - S(y-x)\| \le \|y-x\| \cdot \sup_{\tau\in(0,1)} \|f'(x+\tau(y-x)) - S\|.$$

PROOF. The curve $\gamma(\tau) = f(x+\tau(y-x)) - S\tau(y-x)$ is differentiable in τ on $[0,1]$ with derivative

$$\gamma'(\tau) = f'(x+\tau(y-x))\,(y-x) - S(y-x).$$

Now apply Lemma C.2.7 *ii)* with $s=1$, $t=0$ and $v = S(y-x)$. Then use boundedness of S and f' on $[x,y]$. □

THEOREM C.2.9. (**Newton method**) *Let $f : X \to Y$ be a continuously differentiable map between Banach spaces. Suppose $D = df(x_0)$ is onto with right inverse T and there exist constants $\delta, c > 0$ such that*

$$\|f(x_0)\| \le \frac{\delta}{2c} \quad , \quad \|T\| \le c \quad , \quad \|df(x) - D\| \le \frac{1}{2c} \tag{192}$$

whenever $\|x - x_0\| \le \delta$. Then there exists a unique $\tilde{x} \in X$ with $f(\tilde{x}) = 0$, $\|\tilde{x} - x_0\| \le \delta$ and $\tilde{x} - x_0 \in im\, T$.

PROOF. The first step of Newton's iteration is to define

$$\xi_0 = -Tf(x_0) \qquad , \qquad x_1 = x_0 + \xi_0. \tag{193}$$

Using the assumptions we estimate

$$\begin{aligned}
\|\xi_0\| &= \|Tf(x_0)\| \le c\,\|f(x_0)\| \\
\|\xi_0\| &\overset{(192)}{\le} \frac{\delta}{2} \\
\|f(x_1)\| &= \|f(x_0+\xi_0) \underbrace{-f(x_0) - D\xi_0}_{\overset{(192)}{=}0}\| \le \frac{1}{2c}\,\|\xi_0\|.
\end{aligned}$$

To get the last estimate we applied Corollary C.2.8 with $S = D$, $y = x_0 + \xi_0$ and $x = x_0$:

$$\begin{aligned} \|f(x_0+\xi_0) - f(x_0) - D\xi_0\| &\le \|\xi_0\| \sup_{\tau\in(0,1)} \|df(x_0+\tau\xi_0) - D\| \\ &\overset{(192)}{\le} \frac{1}{2c}\|\xi_0\|. \end{aligned}$$

For $\nu = 1, 2, 3, \dots$ we define inductively

$$\xi_\nu = -Tf(x_\nu) \quad , \quad x_{\nu+1} = x_\nu + \xi_\nu. \tag{194}$$

Assuming that for any $\nu \in \mathbb{N}$

$$\begin{aligned} i)&\quad \|\xi_\nu\| \le c\,\|f(x_\nu)\| \\ ii)&\quad \|\xi_\nu\| \le \frac{1}{2}\|\xi_{\nu-1}\| \\ iii)&\quad \|f(x_{\nu+1})\| \le \frac{1}{2c}\|\xi_\nu\| \end{aligned}$$

we have to show that this implies the corresponding statement with ν replaced by $\nu + 1$:

$$\begin{aligned} \text{ad } i)&\quad \|\xi_{\nu+1}\| \overset{(194)}{=} \|Tf(x_{\nu+1})\| \le c\,\|f(x_{\nu+1})\| \\ \text{ad } ii)&\quad \|\xi_{\nu+1}\| \overset{iii)}{\le} \frac{1}{2}\|\xi_\nu\| \\ \text{ad } iii)&\quad \|f(x_{\nu+2})\| = \|f(x_{\nu+1}+\xi_{\nu+1}) \underbrace{-f(x_{\nu+1}) - D\xi_{\nu+1}}_{\overset{(194)}{=0}}\| \le \frac{1}{2c}\|\xi_{\nu+1}\|. \end{aligned}$$

In the last step we again applied Corollary C.2.8 with $S = D$, $y = x_{\nu+1}+\xi_{\nu+1}$ and $x = x_{\nu+1}$. Note that in estimating the supremum term by $1/2c$ we use the fact

$$\|x_{\nu+1} + \xi_{\nu+1} - x_0\| = \left\|\sum_{k=0}^{\nu+1} \xi_k\right\| \le \sum_{k=0}^{\nu+1} \|\xi_k\| \le \frac{\delta}{2} \sum_{k=0}^{\nu+1} \frac{1}{2^k} < \delta.$$

This estimate also shows the absolute summability of the series $\sum_{k=0}^{\infty} \xi_k$ and so the following limit exists

$$\tilde{x} = \lim_{\nu\to\infty} x_\nu = x_0 + \lim_{\nu\to\infty} \sum_{k=0}^{\nu} \xi_k.$$

Clearly $\|\tilde{x} - x_0\| \le \delta$ and

$$\|f(\tilde{x})\| = \lim_{\nu\to\infty} \|f(x_{\nu+1})\| \overset{iii)}{\le} \lim_{\nu\to\infty} \frac{1}{2c}\|\xi_\nu\| \overset{ii)}{\le} \frac{\delta}{4c} \lim_{\nu\to\infty} \frac{1}{2^\nu} = 0,$$

hence $f(\tilde{x}) = 0$. As $\xi_\nu \in im\,T$ and $im\,T$ is closed, it follows $\tilde{x} - x_0 \in im\,T$. This proves existence.

To prove uniqueness assume there exists another $\hat{x} \in X$ with $\|x_0 - \hat{x}\| \leq \delta$, $f(\hat{x}) = 0$ and $x_0 - \hat{x} \in im\, T$, i.e. there exist $\tilde{y}$, $\hat{y} \in Y$ such that the following hold

$$x_0 - \tilde{x} = T\tilde{y} \quad , \quad x_0 - \hat{x} = T\hat{y}.$$

Now use the assumptions and apply corollary C.2.8 with $S = D$, $y = \tilde{x}$ and $x = \hat{x}$ to get

$$\begin{aligned} \|\tilde{y} - \hat{y}\| &= \|f(\tilde{x}) - f(\hat{x}) - D(\tilde{x} - \hat{x})\| \\ &\leq \|\tilde{x} - \hat{x}\| \sup_{\tau \in (0,1)} \|df|_{\hat{x}+\tau(\tilde{x}-\hat{x})} - D\| \\ &\leq \|T(\hat{y} - \tilde{y})\| \frac{1}{2c} \\ &\leq \frac{1}{2} \|\hat{y} - \tilde{y})\|. \end{aligned}$$

Therefore $\|\tilde{y} - \hat{y}\| = 0$, i.e. $\tilde{y} = \hat{y}$, and so $0 = T\tilde{y} - T\hat{y} = \hat{x} - \tilde{x}$. □

Remark C.2.10. The uniqueness part in the above proof may also be demonstrated by using basic quadratic estimates as in lemma 5.0.9 in chapter 5 instead of corollary C.2.8. Assume there are elements $\tilde{x}$, $\hat{x} \in X$ with

$$\|x_0 - \tilde{x}\| + \|x_0 - \hat{x}\| \leq \delta$$
$$f(\tilde{x}) = 0 = f(\hat{x})$$

and

$$x_0 - \tilde{x} = T\tilde{y} \quad , \quad x_0 - \hat{x} = T\hat{y}$$

for some $\tilde{y}$, $\hat{y} \in Y$, then

$$\begin{aligned} \|\tilde{x} - \hat{x}\| &= \|T(\hat{y} - \tilde{y})\| \leq c\,\|\hat{y} - \tilde{y}\| = c\,\|D(\tilde{x} - \hat{x})\| \\ &= c\,\|f(\hat{x}) - f(\tilde{x}) - D(\hat{x} - \tilde{x})\| \\ &\leq c\,\|f(\tilde{x} + (\hat{x} - \tilde{x})) - f(\tilde{x}) - df|_{\tilde{x}}\,(\hat{x} - \tilde{x})\| \\ &\quad + c\,\|(df|_{\tilde{x}} - D)\,(\hat{x} - \tilde{x})\| \\ &\leq c \cdot c(\hat{x} - \tilde{x})\,\|\hat{x} - \tilde{x}\|^2 + c\,\frac{1}{2c}\,\|\hat{x} - \tilde{x}\| \\ &= (c \cdot c(\hat{x} - \tilde{x})\,2\delta + 1/2)\;\|\hat{x} - \tilde{x}\|. \end{aligned}$$

Here $c(\hat{x} - \tilde{x})$ is the continuous function appearing in lemma 5.0.9 in chapter 5, hence for $\delta > 0$ sufficiently small it follows $\|\hat{x} - \tilde{x}\| = 0$.

C.3. Inverse and implicit function theorem

As an application of the Newton method we prove the inverse function theorem and as a consequence the implicit function theorem. The latter states that the preimage of a regular value of a smooth map between Banach spaces is a smooth manifold. Indeed it can be locally represented as the graph of a smooth function (figure C.4).

Definition C.3.1. We say that a bounded linear operator D between Banach spaces X and Y is *invertible*, if there exists a bounded linear operator $T : Y \to X$ such that $D \circ T = id_Y$ and $T \circ D = id_X$. In this case we denote T by D^{-1}. D invertible is equivalent to D bijective.

Theorem C.3.2. (**Inverse function theorem**) *Let $f \in C^k(X,Y)$, $k \geq 1$, X, Y Banach spaces, and assume that $df(x_0)$ is invertible at a point $x_0 \in X$. Then there exists a constant $\epsilon > 0$ such that for any $y \in Y$ with $\|y - f(x_0)\| < \epsilon$ there exists a unique $x \in X$ near x_0 for which $f(x) = y$. Moreover, f maps an open neighbourhood U of x_0 bijectively onto $V = f(U) = \{y \in Y \mid \|y - f(x_0)\| < \epsilon\}$. The inverse f^{-1} is in $C^k(V,U)$ and*

$$df^{-1}(y) = df(x)^{-1}$$

for $y \in V$ and $x \in U$ with $f(x) = y$.

Proof. We assume without loss of generality $x_0 = 0$ and $f(x_0) = y_0 = 0$ (otherwise pick the function $\tilde{f}(x + x_0) - y_0$). As $D = df(0) \in L(X,Y)$ is invertible, it has a bounded inverse T. Let $c_0 > 0$ be such that $\|T\| \leq c_0$. Continuity of $df(x)$ implies that there exists $\delta > 0$ such that

$$\|df(x) - D\| \leq \frac{1}{2c_0} \qquad \text{for } \|x\| \leq \delta.$$

As invertibility is an open condition we may pick $\delta > 0$ sufficiently small in order to guarantee invertibility of $df(x)$ whenever $\|x\| \leq \delta$.

Now we show that f is a bijection between an open neighbourhood of zero $U \subset B^X_\delta(0)$ and $V = B^Y_{\delta/2c_0}(0)$, i.e. we set $\epsilon = \delta/2c_0$. Pick $y \in V$ and define

$$\begin{aligned} F_y \ : \ B^X_\delta(0) &\to Y \\ x &\mapsto f(x) - y. \end{aligned}$$

We would like to apply the Newton method (Theorem C.2.9) to get a unique zero of F_y. F_y has the following properties

1. $F_y(0) = -y \quad \Rightarrow \quad \|F_y(0)\| = \|y\| < \delta/2c_0$
2. $dF_y(0) = df(0) = D \quad \Rightarrow \quad \|dF_y(0) - dF_y(x)\| \leq 1/2c_0$ for $\|x\| \leq \delta$
3. T is a right inverse of $dF_y(0)$ with $\|T\| \leq c_0$.

Theorem C.2.9 gives now the unique zero x of F_y in $B^X_\delta(0)$. As f is continuous, $U = f^{-1}(V) \subset B^X_\delta(0)$ is an open neighbourhood of $0 \in X$. Hence $f|_U : U \to V$ is bijective (cf. figure C.3).

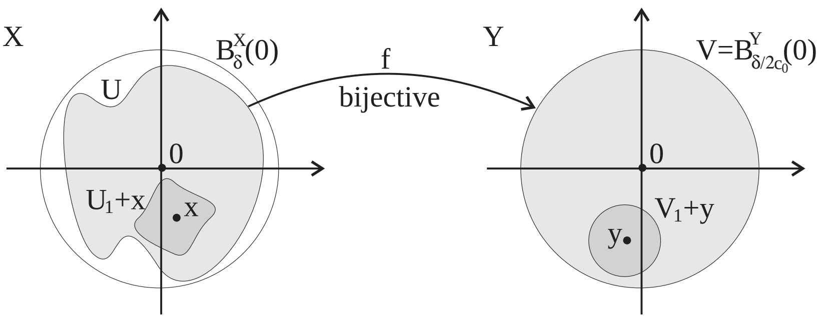

FIGURE C.3. Inverse function theorem

It remains to show that f^{-1} is continuously differentiable in V: pick any $y \in V$ and $x \in U$ with $f(x) = y$. Let T_x be the inverse of $D_x = df(x)$ with $0 < \|T_x\| \le c_1 = c_1(x)$. Continuity of $df(\cdot)$ implies that there exists $\delta_1 = \delta_1(x) > 0$ such that $\|df(x) - df(x+\xi)\| \le 1/2c_1$ for all $\xi \in B^X_{\delta_1}(0)$. Choose $\delta_1 > 0$ sufficiently small such that $B^Y_{\delta_1/2c_1}(y) \subset V$. Let $f(x+\xi) = y + \eta$ for $\eta \in V_1 := B^Y_{\delta_1/2c_1}(0)$, $\xi \in U_1 := f^{-1}(V_1 + y) - x$. Our claim is to establish

$$\|f^{-1}(y+\eta) - f^{-1}(y) - df(x)^{-1}\eta\| = o(\|\eta\|) \quad \text{for } \|\eta\| \to 0. \tag{195}$$

We analyze the Newton iteration for the function

$$\begin{aligned} G_y \;:\; B^X_{\delta_1}(0) &\to Y \\ \xi &\mapsto f(x+\xi) - (y+\eta). \end{aligned}$$

We have

1. $G_y(0) = -\eta \quad\Rightarrow\quad \|G_y(0)\| = \|\eta\| \le \delta_1/2c_1$
2. $dG_y(0) = df(x) \quad\Rightarrow\quad \|dG_y(0) - dG_y(\xi)\| \le 1/2c_1$ for $\|\xi\| < \delta_1$
3. T_x is a right inverse of $dG_y(0)$ with $\|T_x\| \le c_1$.

The iteration starts with setting $x_0 = 0$ and

$$\xi_0 = -T_x G_y(0) = T_x\eta \tag{196}$$

then $x_1 = x_0 + \xi_0 = \xi_0$

$$\xi_1 = -T_x G_y(x_1) = -T_x\big(f(x+\xi_0) - (y+\eta)\big).$$

Note that

$$\|\xi_1\| \le c_1\|f(x+\xi_0) - (y+\eta)\| = o(\|\eta\|) \quad \text{for } \|\eta\| \to 0, \tag{197}$$

because

$$f(x+\xi_0) = f(x+T_x\eta) = \overbrace{f(x)}^{=y} + \overbrace{df(x)\,T_x\eta}^{=\eta} + o(\|T_x\eta\|) \quad \text{for } \|T_x\eta\| \to 0$$

and

$$\frac{1}{c_1}\,\|T_x\eta\| \le \|\eta\| \le \|D\|\cdot\|T_x\eta\|.$$

The last estimate ensures that if a function is of class $o(\|T_x\eta\|)$ for $\|T_x\eta\| \to 0$, then it is also of class $o(\|\eta\|)$ for $\|\eta\| \to 0$. Recall from the proof of Theorem C.2.9 that the zero ξ of G_y is given by

$$\xi = x_0 + \sum_{k=0}^{\infty} \xi_k \qquad , \text{ where } \|\xi_k\| \leq \frac{1}{2} \|\xi_{k-1}\| \ , \ k \in \mathbb{N},$$

hence the LHS of equation (195) is given by

$$\begin{aligned} \|x + \xi - x - T_x\eta\| & \overset{(196)}{=} \|\xi - \xi_0\| \leq \sum_{k=1}^{\infty} \|\xi_k\| \\ & \leq \|\xi_1\| \sum_{k=1}^{\infty} \frac{1}{2^{k-1}} = 2\|\xi_1\| \overset{(197)}{=} o(\|\eta\|) \quad \text{as } \|\eta\| \to 0. \end{aligned}$$

This proves the theorem for $k = 1$. The case $k > 1$ follows inductively. □

DEFINITION C.3.3. i) A *Fredholm operator* is a bounded linear map $D : X \to Y$ between Banach spaces with the following properties: $dim\, Ker\, D < \infty$, *Ran D* is closed and $dim\ coker\ D < \infty$. The *Fredholm index of D* is defined by $Ind\, D = dim\ ker\ D - dim\ coker\ D$.
ii) A map $f \in C^k(X, Y)$, $k \geq 1$, between Banach spaces is called a *Fredholm map*, if its linearization $D = df(x) : X \to Y$ is a *Fredholm operator* for every $x \in X$. $Ind\, D$ is invariant under small perturbations, hence $Ind\, df(x)$ is independent of the choice of x. We denote it by $Ind\, f$.
iii) For any map f as in ii) (Fredholm or not) an element $y \in Y$ is called a *regular value* of f, if $df(x) : X \to Y$ admits a right inverse for every $x \in f^{-1}(y)$. Note that by definition an element $y \in Y$ with $f^{-1}(y) = \emptyset$ is a regular value.

THEOREM C.3.4. (**Implicit function theorem**) *Let $f \in C^k(X, Y)$, $k \geq 1$, where X and Y are Banach spaces. If y is a regular value of f, then*

$$\mathcal{M} = f^{-1}(y) \subset X$$

is a C^k-Banach manifold. If f is a Fredholm map, then $\mathcal{M}$ is finite dimensional

$$dim\, \mathcal{M} = Ind\, f.$$

PROOF. Assume without loss of generality $y = 0$ and $x_0 = 0 \in f^{-1}(0)$ (otherwise pick the function $\tilde{f}(x + x_0) - y$). As zero is a regular value of f, $D = df(0) : X \to Y$ is surjective and admits a right inverse $T \in L(Y, X)$. Hence we have $X = ker\ D \oplus im\, T$. We define the function

$$\begin{aligned} F \ : \ ker\ D \oplus im\, T \quad & \to \quad ker\ D \oplus Y \\ (\xi, \eta) \quad & \mapsto \quad (\xi, f(\xi, \eta)) \end{aligned}$$

and observe that

$$dF(\xi, \eta) = \begin{pmatrix} 1\!\!1_{ker\ D} & 0 \\ \partial_\xi f(\xi, \eta) & \partial_\eta f(\xi, \eta) \end{pmatrix} \ : \ \begin{matrix} ker\ D \\ \oplus \\ im\, T \end{matrix} \ \to \ \begin{matrix} ker\ D \\ \oplus \\ Y \end{matrix}$$

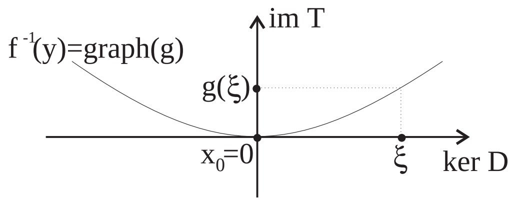

FIGURE C.4. Implicit function theorem

is surjective at $(\xi, \eta) = (0, 0)$: the second row is precisely $df(0)$. It is also injective at $(0, 0)$: $1\!\!1_{ker\ D}$ and $\partial_\eta f(0, 0) = df(0)|_{im\ T}$ are isomorphisms and $\partial_\xi f(0, 0) = df(0)|_{ker\ D} = 0$. Therefore we may apply the inverse function theorem C.3.2 and conclude that F is locally at $(0, 0)$ a diffeomorphism, i.e. there exist neighbourhoods $U(0, 0) \subset ker\, D \oplus im\, T$ and $V(0, 0) \subset ker\, D \oplus Y$ such that $\tilde{F} = F|_{U(0,0)} : U(0, 0) \to V(0, 0)$ is a diffeomorphism. Restricting $\tilde{F}^{-1}$ to $(ker\ D \times 0) \cap V(0, 0)$ gives another diffeomorphism

$$\begin{aligned} \varphi\ :\ (ker\ D \times 0) \cap V(0,0) \quad &\to \quad \mathcal{M} \cap U(0,0) \\ (\xi, 0) \quad &\mapsto \quad \tilde{F}^{-1}(\xi, 0). \end{aligned}$$

This proves the theorem.

Another interpretation of φ is as follows: setting $\tilde{F}^{-1}(\xi, 0) = (\xi, g(\xi))$ defines a smooth function $g : ker\ D \supset U(0) \to im\ T$, where $U(0)$ is the neighbourhood of $0 \in ker\ D$ defined by projecting $U(0, 0)$ to its first component. We get $\mathcal{M}$ locally around $x_0 = (0, 0)$ as the graph of g. We have $g(0) = 0$ and $dg(0) = 0$. The last statement follows from

$$\begin{pmatrix} 1\!\!1_{ker\ D} & 0 \\ dg(0) & 0 \end{pmatrix} = dF^{-1}(0, 0) = dF|_{F^{-1}(0,0)=(0,0)} = \begin{pmatrix} 1\!\!1_{ker\ D} & 0 \\ \partial_\xi f(0, 0) & \partial_\eta f(0, 0) \end{pmatrix}$$

and $\partial_\xi f(0, 0) = 0$ as shown above (figure C.4). □

APPENDIX D

Topology of $Sp(2,\mathbb{R})$ and the Conley-Zehnder index

The symplectic linear group arises naturally in the study of linear Hamiltonian equations (cf. [**Ar88**])

$$\dot{\zeta}(t) = -J_0 S\zeta(t)\ ,\ J_0 = \begin{pmatrix} 0 & -1\!\!1 \\ 1\!\!1 & 0 \end{pmatrix}$$

where S is the Hessian of the Hamiltonian $H : \mathbb{R}^{2n} \to \mathbb{R}$. The solution ζ with initial condition $\zeta(0) = \zeta_0$ is given by

$$\zeta(t) = e^{-tJ_0 S}\zeta_0 \stackrel{def}{=} A(t)\zeta_0$$

and $A(t)$ is in $Sp(2n,\mathbb{R}) = \{A \in Mat(2n,\mathbb{R}) \mid A^T J_0 A = J_0\}$ for any t. For $n = 1$ the symplectic linear group coincides with the special linear group $Sl(2,\mathbb{R})$, whose elements have determinant 1 and so are area-preserving.

In 1984 Conley and Zehnder introduced an index for continuous paths in $Sp(2n,\mathbb{R})$ which start at the identity and end at a matrix whose spectrum does not contain 1. This index became an important ingredient in the construction of Floer homology for symplectic manifolds [**F89b**], [**RS95**], which lead to a proof of the *Arnold conjecture* in considerable generality (the number of fixed points of an exact symplectic diffeomorphism on a symplectic manifold can be estimated below by the sum of its Betti numbers provided that the fixed points are nondegenerate, [**Ar65**]).

In section D.1 our first claim is to show that $Sp(2,\mathbb{R})$ is homeomorphic to the interior of the full 2-torus $S^1 \times D^2$, where D^2 denotes the open unit disc. The explicit homeomorphism was taken from an article of Gelfand and Lidskii [**GL58**], 1958. Visualizing particular subsets of $Sp(2,\mathbb{R})$ in figure 6 we discuss the notion of an eigenvalue of the first and second kind.

In section D.2 we recall the definition of the Conley-Zehnder index. We interpret this index in the case $n = 1$ as intersection number of a path with the Maslov cycle $\mathcal{C}_+$, which is a codimension one algebraic subvariety (figure 8). We then introduce a generalized Conley-Zehnder index, where we drop the condition on the endpoint of the path. A simple example is discussed (figure D.5). Finally we visualize and discuss another Maslov-type index, the Robbin-Salamon index, which may be interpreted as intersection number of an arbitrary path with the Maslov cycle Sp_1 (figure 7), where endpoints only contribute half. This appendix has been previously published in preprint form [**We98**].

D.1. Topology of $Sp(2,\mathbb{R})$

PROPOSITION D.1.1. (**Polar decomposition**) *Let $Y \in Sp(2,\mathbb{R})$, then there exists a unique $R \in Sp(2,\mathbb{R}) \cap SO(2,\mathbb{R})$ and a unique $S \in Sp(2,\mathbb{R})$, positive definite and symmetric, such that $Y = SR$.*

PROOF. YY^T is clearly positive definite, symmetric and symplectic. The first two properties also hold for $S := (YY^T)^{1/2}$ (functional calculus). S is symplectic by Lemma 2.19 in [**MS95**] which says that any real power of a positive definite, symmetric and symplectic matrix is itself symplectic. Now we define $R = S^{-1}Y$; R is symplectic and also orthogonal: Let $x, y \in \mathbb{R}^2$ and denote by $< \cdot, \cdot >$ the euclidean inner product on $\mathbb{R}^2$, then

$$< Rx, Ry >=< S^{-1}Yx, S^{-1}Yy >=< x, Y^T {S^{-1}}^T S^{-1}Yy >=< x, y > \ .$$

The last equation holds, because ${S^{-1}}^T = S^{-1}$, so

$${S^{-1}}^T S^{-1} = S^{-1}S^{-1} = (YY^T)^{-1/2}(YY^T)^{-1/2} = (YY^T)^{-1} \ .$$

Finally $det\, R = (det\, S)^{-1} \cdot det\, Y = +1$ as $S, Y \in Sp(2,\mathbb{R}) = Sl(2,\mathbb{R})$. □

Hence we may write any $Y \in Sp(2,\mathbb{R})$ in the form

$$Y = \begin{pmatrix} s_{11} & s_{12} \\ s_{12} & s_{22} \end{pmatrix} \begin{pmatrix} \cos\psi & -\sin\psi \\ \sin\psi & \cos\psi \end{pmatrix} \ , \tag{198}$$

where $\psi \in \mathbb{R}/2\pi\mathbb{Z}$ and

$$s_{11}s_{22} - {s_{12}}^2 = 1 \ . \tag{199}$$

LEMMA D.1.2. *$s_{11}, s_{22} > 0$ and $s_{11} \cdot s_{22} \geq 1$.*

PROOF. The last statement is a simple consequence of (199). (199) also implies that either $s_{11} > 0$ and $s_{22} > 0$ or $s_{11} < 0$ and $s_{22} < 0$. Assume the second case, then the positive definiteness of S leads to a contradiction: $Tr\, S = s_{11} + s_{22} > 0$. □

Now (199) implies that s_{22} is uniquely determined once $(s_{11}, s_{12}) \in \mathbb{R}^+ \times \mathbb{R}$ has been chosen:

$$s_{22}(s_{11}, s_{12}) = \frac{1 + (s_{12})^2}{s_{11}}.$$

Hence the set of possible parameters is given by

$$\mathcal{M} = \mathbb{R}^+ \times \mathbb{R},$$

which is diffeomorphic to the open unit disc in $\mathbb{R}^2$. Recall the hyperbolic trigonometric functions

$$\sinh x = \tfrac{1}{2}(e^x - e^{-x}) \ , \ \cosh x = \tfrac{1}{2}(e^x + e^{-x}) \ , \ \tanh x = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

whose qualitative behaviour is shown in figure D.1. They satisfy the relation

$$\cosh^2 x - \sinh^2 x = 1 \ . \tag{200}$$

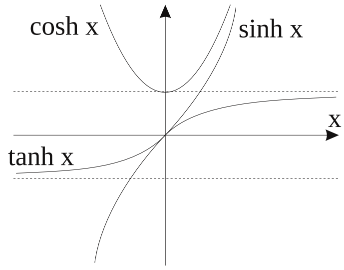

FIGURE D.1. Hyperbolic trigonometric functions

We reparametrize $\mathcal{M}$ in the following way ([**GL58**], §9)

$$\mathcal{S} : (0,\infty)\times\mathbb{R}/2\pi\mathbb{Z} \sqcup (0,0) \to \mathcal{M}$$
$$(\tau,\sigma)\mapsto (s_{11}(\tau,\sigma), s_{12}(\tau,\sigma)) ,$$

where

$$\text{(201)}\qquad \begin{aligned} &i)\ \ s_{11}(\tau,\sigma) = \cosh\tau + \sinh\tau\cos\sigma \\ &ii)\ \ s_{12}(\tau,\sigma) = \sinh\tau\sin\sigma\ . \end{aligned}$$

The parameter $s_{22}(\tau,\sigma)$ is then given by

$$\begin{aligned} iii)\ \ s_{22}(\tau,\sigma) &= \frac{1+(s_{12})^2}{s_{11}} = \frac{1+\sinh^2\tau\sin^2\sigma + \sinh^2\tau - \sinh^2\tau}{\cosh\tau+\sinh\tau\cos\sigma} \\ &= \frac{\cosh^2\tau - \sinh^2\tau(1-\sin^2\sigma)}{\cosh\tau - \sinh\tau\cos\sigma} = \cosh\tau - \sinh\tau\cos\sigma. \end{aligned}$$

PROPOSITION D.1.3. *The map $\mathcal{S}$ is a homeomorphism. On $(0,\infty)\times\mathbb{R}/2\pi\mathbb{Z}$ it is a diffeomorphism onto $\mathcal{M}\setminus(1,0)$.*

PROOF. Smoothness of $\mathcal{S}$ follows from the one of the hyperbolic trigonometric functions. $\mathcal{S}$ is injective: Assume $\mathcal{S}(\tau,\sigma)=\mathcal{S}(\tau',\sigma')$, then

$$(201i) + iii) \Rightarrow 2\cosh\tau' = 2\cosh\tau \Rightarrow \tau'=\tau\ ,$$
$$(201i) - iii) \Rightarrow 2\sinh\tau\cos\sigma' \overset{\tau'=\tau}{=} 2\sinh\tau\cos\sigma \Rightarrow \cos\sigma' = \cos\sigma\ ,$$
$$(201ii) \Rightarrow \sin\sigma' = \sin\sigma\ \text{, hence}\ \sigma'=\sigma\ .$$

$\mathcal{S}$ is surjective: Addition and subtraction of s_{11}, $s_{22}(s_{11},s_{12})$ leads to a smooth inverse on $\mathbb{R}^+\times\mathbb{R}\setminus(0,1)$

$$\mathcal{S}^{-1}(s_{11},s_{12}) = (\tau(s_{11},s_{12}),\sigma(s_{11},s_{12}))\ ,$$

where

$$\begin{aligned} \tau(s_{11},s_{12}) &= \operatorname{arc\,cosh}\frac{s_{11}+s_{22}(s_{11},s_{12})}{2} \\ &= \operatorname{arc\,cosh}\frac{(s_{11})^2+1+(s_{12})^2}{2s_{11}}\ , \end{aligned}$$

(202)

$$\sigma(s_{11},s_{12}) = (sign\ s_{12})\ \arccos\left(\frac{s_{11}-s_{22}(s_{11},s_{12})}{2\sinh\circ\operatorname{arc\,cosh}\frac{s_{11}+s_{22}(s_{11},s_{12})}{2}}\right)$$
$$= (sign\ s_{12})\ \arccos\left(\frac{(s_{11})^2-1-(s_{12})^2}{\sqrt{\left((s_{11})^2-1-(s_{12})^2\right)^2-4(s_{11})^2}}\right).$$

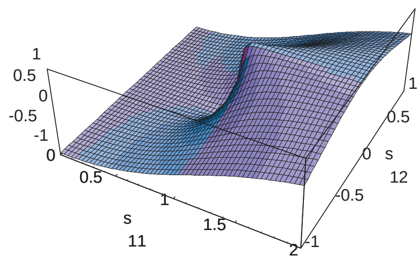

FIGURE D.2. The argument of arccos in (202)

Note that we used the identity $\operatorname{arc\,cosh} x = \operatorname{arc\,sinh}\sqrt{x^2-1}$ for $x \geq 1$ and the convention $sign\ 0 = +1$. The argument of arccos in (202) is a smooth function on $\mathbb{R}^+ \times \mathbb{R} \setminus (1,0)$ which is not continuous at $(1,0)$. A numerical *Mathematica**-plot of it is shown in figure D.2. It is identically $+1$ on $(0,1)\times 0$ and identically -1 on $(1,\infty)\times 0$. This corresponds to $\sigma = \pi$ and $\sigma = 0$, respectively. Therefore multiplication by $sign\ s_{12}$ is well-defined and the full range $[-\pi,\pi]/\{\pi,-\pi\}$ of σ is covered. Finally we define $\mathcal{S}^{-1}(1,0) = (0,0)$. □

Rescale τ by setting $r(\tau) = \tanh^2\tau$, $\tau \in [0,\infty)$, i.e. $r \in [0,1)$. We interpret the parameters

$$(\psi, r, \sigma) \in \mathbb{R}/2\pi\mathbb{Z} \times [(0,1)\times \mathbb{R}/2\pi\mathbb{Z} \sqcup (0,0)]$$

as coordinates of the open solid 2-torus (figure D.3).

In this coordinates the matrices

$$E = \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\ ,\ \begin{pmatrix}-1 & 0\\ 0 & -1\end{pmatrix}\ ,\ W_- = \begin{pmatrix}2 & 0\\ 0 & \frac{1}{2}\end{pmatrix}\ ,\ \pm J_0 = \pm\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}$$

correspond to

$$(0,0,0)\ ,\ (\pi,0,0)\ ,\ (0,\frac{3}{5},0)\ ,\ (\frac{(3)\pi}{2},0,0)$$

**Mathematica* is a registered trademark of Wolfram Research, Inc.

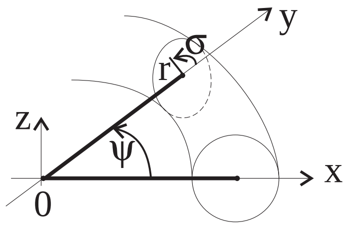

FIGURE D.3. Coordinates on the full 2-torus

as follows from their polar decompositions

$$\begin{pmatrix}1&0\\0&1\end{pmatrix}\begin{pmatrix}1&0\\0&1\end{pmatrix},\ \begin{pmatrix}1&0\\0&1\end{pmatrix}\begin{pmatrix}-1&0\\0&-1\end{pmatrix},$$
$$\begin{pmatrix}2&0\\0&\frac{1}{2}\end{pmatrix}\begin{pmatrix}1&0\\0&1\end{pmatrix},\ \begin{pmatrix}1&0\\0&1\end{pmatrix}\begin{pmatrix}0&\mp1\\\pm1&0\end{pmatrix}.$$

The condition for $Y \in Sp(2,\mathbb{R}) = Sl(2,\mathbb{R})$ to have non-real eigenvalues is equivalent to $|Tr\,Y| < 2$ as

$$0 = \det(Y - \lambda E) = \det\begin{pmatrix}a-\lambda & b\\ c & d-\lambda\end{pmatrix} = \lambda^2 - \lambda \cdot Tr\,Y + 1$$
$$\Leftrightarrow\ \lambda_{1,2} = \frac{Tr\,Y \pm \sqrt{(Tr\,Y)^2 - 4}}{2}\ .$$

Using the polar decomposition (198) of Y we get

$$|Tr\,Y| = |(s_{11} + s_{22})\cos\psi| = 2|\cos\psi|\cosh\tau$$
$$= 2|\cos\psi|\cosh\circ\text{arc}\tanh\sqrt{r}\ .$$

LEMMA D.1.4. *Let Y be as in (198), then $|Tr\,Y| = 2$ is equivalent to $r = \sin^2\psi$.*

PROOF. $|Tr\,Y| = 2$ implies $\psi \notin \{\frac{\pi}{2}, \frac{3\pi}{2}\}$ and is equivalent to

$$\pm\text{arc}\cosh\frac{1}{|\cos\psi|} = \text{arc}\tanh\sqrt{r} \Leftrightarrow \tanh^2\circ\pm\text{arc}\cosh\frac{1}{|\cos\psi|} = r$$
$$\Leftrightarrow\ \frac{\cosh^2\circ\pm\text{arc}\cosh\frac{1}{|\cos\psi|} - 1}{\cos^{-2}\psi} = r \Leftrightarrow 1 - \cos^2\psi = r \Leftrightarrow \sin^2\psi = r.$$

In the third equivalence we used $\tanh x = \sinh x/\cosh x$ and (200). □

As the eigenvalues of $Y \in Sp(2,\mathbb{R})$ come as pairs (λ, λ^{-1}),we observe that

$$Spec\,Y = \{+1\} \Leftrightarrow Tr\,Y = 2 \Leftrightarrow \sin^2\psi = r\ , \psi \in (-\pi/2, \pi/2)\ ,$$
$$Spec\,Y = \{-1\} \Leftrightarrow Tr\,Y = -2 \Leftrightarrow \sin^2\psi = r\ , \psi \in (\pi/2, 3\pi/2)\ .$$

We define

$$\mathcal{C}_{\pm} = \{Y \in Sp(2,\mathbb{R}) | Spec\, Y = \{\pm 1\}\} .$$

REMARK D.1.5. Our results so far are visualized in figure 6 , which has been created using *Mathematica*[1] and *Geomview*[†]. Note that figure 6 is not a sketch and it is not obtained by numerical approximation, but represents exactly the set $|Tr\, Y| = 2$ in the coordinates (ψ, r, σ). The same holds for figure 8 in the case $Tr\, Y = 2$.

The sets $\mathcal{C}_+ \setminus \{E\}$ resp. $\mathcal{C}_- \setminus \{-E\}$ are smooth 2-dimensional surfaces; they are indicated in figure 6 red (dark) resp. green (light). The region enclosed by them consists of two connected components – corresponding to the matrices with non-real eigenvalues. The outside region consists also of two connected components. The one having E in its closure corresponds to the matrices with positive real pairs of eigenvalues (β, β^{-1}), $\beta \in \mathbb{R}^+ \setminus \{1\}$, the other corresponds to the ones having negative real pairs. Note that the set of all possible eigenvalues of elements of $Sp(2,\mathbb{R})$ is the union of $S^1 \subset \mathbb{C}$ with the real line minus zero (figure D.4).

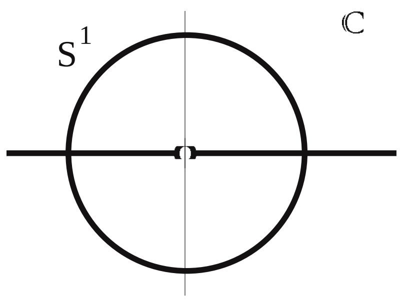

FIGURE D.4. Spectrum of $Sp(2,\mathbb{R})$

At first sight one might be tempted to relate the two parts of S^1 lying in the upper resp. lower half plane $H^{\pm} \subset \mathbb{C}$ with the two connected components corresponding to matrices with non-real eigenvalues. This intuition points the right direction; the relation, however, is more subtle.

DEFINITION D.1.6. Let $Y \in Sp(2,\mathbb{R})$ with eigenvalues $(\lambda, \lambda^{-1} = \bar{\lambda})$ in $S^1 \setminus \{\pm 1\}$ and eigenvectors $\xi_\lambda, \xi_{\bar{\lambda}}$. We call λ *an eigenvalue of the first kind*, if

$$Im\, \omega_0(\bar{\xi_\lambda}, \xi_\lambda) > 0 ,$$

and an *eigenvalue of the second kind* otherwise.

[†]written at the Geometry Center, University of Minnesota

This definition does not depend on the choice of the eigenvector nor on the choice which eigenvalue was called λ resp. $\bar{\lambda}$. ω_0 denotes the standard symplectic form on $\mathbb{R}^2$; in coordinates (x, y) it is given by $\omega_0 = dx \wedge dy$, in terms of the euclidean inner product $< \cdot, \cdot >$ we have

$$\omega_0(\cdot,\cdot) =< \cdot, -J_0 \cdot > \, , \text{ where } J_0 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \, .$$

As $\bar{\xi_\lambda}$ is eigenvector of $\bar{\lambda}$, ξ_λ and $\bar{\xi_\lambda}$ are linearly independent, hence in view of the nondegeneracy of ω_0 we have $\omega_0(\bar{\xi_\lambda}, \xi_\lambda) \neq 0$. Moreover this value is purely imaginary:

$$\begin{aligned} \overline{\omega_0(\bar{\xi_\lambda}, \xi_\lambda)} &=< \bar{\bar{\xi_\lambda}}, \overline{-J_0 \xi_\lambda} > \\ &=< J_0 \xi_\lambda, \bar{\xi_\lambda} >= -\omega_0(\bar{\xi_\lambda}, \xi_\lambda) \, . \end{aligned}$$

We used the fact that ${J_0}^T = -J_0$. Now the eigenvalues of J_0 are $\lambda = i$ and $\bar{\lambda} = -i$ with corresponding eigenvectors $\xi_\lambda = (i, 1)$ and $\xi_{\bar{\lambda}} = (-i, 1)$. As $\omega_0(\bar{\xi_\lambda}, \xi_\lambda) = -2i$, we conclude that i is eigenvalue of the second kind and hence $-i$ of the first kind. A continuity argument allows us to conclude that the matrices with non-real eigenvalues which are in the same connected component as J_0, have their eigenvalues of the first kind in $S^1 \cap H^-$. Analogously for the connected component containing $-J_0$.

The notion of eigenvalues of the first and the second kind will become more important in the case of $Sp(2n, \mathbb{R})$, $n \geq 2$, where the eigenvalues come in quadrupels $(\lambda, \lambda^{-1}, \bar{\lambda}, \bar{\lambda}^{-1})$. Two pairs of eigenvalues on S^1 can meet and leave S^1, if and only if eigenvalues of different kind meet.

D.2. The Conley-Zehnder index

In this section we recall the definition of the Conley-Zehnder index μ_{CZ} for paths in $Sp(2n,\mathbb{R})$ starting at the identity and ending at a matrix without eigenvalue 1. This index was introduced by Conley and Zehnder in 1984 [**CZ84**]. We state the results about the topology of $Sp(2n,\mathbb{R})$ needed to show its well-definedness. We are not going to prove these results in full generality; instead we restrict to $Sp(2,\mathbb{R})$ and use the explicit homeomorphism from section D.1 to visualize them in this case. We shall see that μ_{CZ} may be equivalently defined as intersection number of the path with the Maslov cycle $\mathcal{C}_+$. Define

$$Sp^*_\pm = \{A \in Sp(2n,\mathbb{R}) \mid \det(A-E) \gtrless 0\}\ ,\ Sp^* = Sp^*_+ \cup Sp^*_-\ ,$$
$$\mathcal{P} = \{\gamma : [0,1] \to Sp(2n,\mathbb{R}) \mid \gamma \text{ continuous}\ ,\ \gamma(0) = E\ ,\ \gamma(1) \in Sp^*\}\ .$$

LEMMA D.2.1. *([**CZ84**], Lemma 1.7) Sp^*_+ and Sp^*_- are connected and any loop in them is contractible in $Sp(2n,\mathbb{R})$. Moreover, $W_+ = -E$ lies in Sp^*_+ and $W_- = diag\,(2,-1,\dots,-1,\frac{1}{2},-1,\dots,-1)$ in Sp^*_-.*

The *Conley-Zehnder index*

$$\mu_{CZ}\ :\ \mathcal{P}\ \to\ \mathbb{Z}$$

is defined as follows: Pick $\gamma \in \mathcal{P}$ and associate to it a path $u = F_2 \circ F_1(\gamma) : [0,1] \to U(n,\mathbb{C})$, where

$$F_1 : C^0([0,1], Sp(2n,\mathbb{R})) \to C^0([0,1], Sp(2n,\mathbb{R}) \cap O(2n,\mathbb{R}))$$
$$\gamma \mapsto (\gamma\gamma^T)^{-1/2}\gamma = \begin{pmatrix} X & -Y \\ Y & X \end{pmatrix}$$
$$F_2 : C^0([0,1], Sp(2n,\mathbb{R}) \cap O(2n,\mathbb{R})) \to C^0([0,1], U(n,\mathbb{C}))$$
$$\begin{pmatrix} X & -Y \\ Y & X \end{pmatrix} \mapsto \qquad X + iY =: u\ .$$

Now choose $\alpha \in C^0([0,1],\mathbb{R})$ such that

$$\det u(t) = e^{i\alpha(t)}\ .$$

Of course α is only determined modulo 2π. However,

$$\Delta(\gamma) := \frac{\alpha(1) - \alpha(0)}{\pi}$$

is well-defined. Note that Δ is additive: $\Delta(\gamma_1 \circ \gamma_2) = \Delta(\gamma_1) + \Delta(\gamma_2)$, $\Delta(\gamma^{-1}) = -\Delta(\gamma)$. As by assumption $\gamma(1) \in Sp^*$, we can choose a continuous extension $\tilde{\gamma} : [0,1] \to Sp^*$ such that $\tilde{\gamma}(0) = \gamma(1)$ and $\tilde{\gamma}(1)$ equals either W_+ or W_-, depending on whether $\tilde{\gamma}(0)$ is in Sp^*_+ or in Sp^*_-. Then define

$$\mu_{CZ}(\gamma) = \Delta(\tilde{\gamma} \circ \gamma)\ ,$$

where $\tilde{\gamma}\circ\gamma$ means: follow first γ then $\tilde{\gamma}$. Of course one has to show

$$\text{(203)}\qquad \begin{aligned} &i)\ \text{independence of the choice of the extension } \tilde{\gamma}\ ,\\ &ii)\ \Delta(\tilde{\gamma}\circ\gamma)\in\mathbb{Z}\ . \end{aligned}$$

Now we restrict to the case $n = 2$ and represent the results of section D.1 via figure 8 (cf. remark D.1.5).

$\mathcal{C}_+ = Sp(2,\mathbb{R}) \setminus Sp^*$ is called *Maslov cycle.* After removing the point E, it is a smooth 2-manifold. In our case $u(t) \in U(1,\mathbb{C}) \simeq SO(2,\mathbb{R}) \simeq S^1$ corresponds to the matrix $R(t)$ in the polar-decomposition of Lemma D.1.1, hence

$$\det u(t) = e^{i\psi(t)}\ ,$$

where $\psi(t)$ is one of our torus coordinates (cf. figure D.3). As $\gamma(0) = E$, we have $\psi(0) = 2k\pi$, for $k \in \mathbb{Z}$. If $\tilde{\gamma}(1) = -E$ it follows $\psi(1) = (2l + 1)\pi$ with $l \in \mathbb{Z}$, hence $\Delta(\tilde{\gamma}\circ\gamma) = 2l + 1 - 2k \in \mathbb{Z}$; if $\tilde{\gamma}(1) = E$ then $\psi(1) = 2l\pi$ with $l \in \mathbb{Z}$, hence $\Delta(\tilde{\gamma}\circ\gamma) = 2l - 2k \in \mathbb{Z}$. This proves (203ii). Independence of the choice of the extension $\tilde{\gamma}$, follows from the fact that another extension $\tilde{\tilde{\gamma}}$ is homotopic to $\tilde{\gamma}$ (by Lemma D.2.1 or figure 8),hence $\tilde{\gamma}\circ\tilde{\tilde{\gamma}}^{-1}$ is contractible and we have $0 = \Delta(\tilde{\gamma}\circ\tilde{\tilde{\gamma}}^{-1}) = \Delta(\tilde{\gamma}) - \Delta(\tilde{\tilde{\gamma}})$. The first equality follows from the fact that if γ is a loop, $\Delta(\gamma)$ is the degree of the map $det \circ F_2 \circ F_1 \circ \gamma : S^1 \to S^1$, and therefore is a homotopy invariant. So if the loop is contractible the degree is zero.

Similarly it follows that μ_{CZ} descends to the equivalence classes $\hat{\mathcal{P}}$

$$\hat{\mu}_{CZ} : \hat{\mathcal{P}} = \mathcal{P}/\sim\ \to\ \mathbb{Z}\ ,$$

where $\gamma_0 \sim \gamma_1 \Leftrightarrow \exists F : [0,1]\times[0,1] \to Sp(2,\mathbb{R})$ continuous, such that $F(0,t) = \gamma_0(t)$, $F(1,t) = \gamma_1(t)$, $F(s,0) = E$, $F(s,1) \in Sp^*$.

If we assign appropriate orientations to the two connected components of $\mathcal{C}_+ \setminus \{E\}$, we can interpret μ_{CZ} as intersection number of γ with $\mathcal{C}_+$. We have to be careful, however, in which direction our path starts off at $\gamma(0) = E$:

1. If $\exists\delta > 0$ such that $\gamma((0,\delta)) \subset Sp^*_+$, then assign to $\gamma(0)$ the intersection number $+1$ if the angle $\psi(t)$ increases and -1 if it decreases.
2. If $\exists\delta > 0$ such that $\gamma((0,\delta)) \subset Sp^*_-$, then assign to $\gamma(0)$ the intersection number 0.
3. If $\exists\delta > 0$ such that $\gamma((0,\delta)) \subset \mathcal{C}_+$, then perturb γ slightly (with fixed end points) to end up in one of the former cases.

For a general continuous path $\gamma : [0,1] \to Sp(2,\mathbb{R})$ with $\gamma(0) = E$ we may define a *generalized Conley-Zehnder index*

$$\mu^g_{CZ} : C^0\left(([0,1],0),(Sp(2,\mathbb{R}),E)\right) \to \tfrac{1}{2}\cdot\mathbb{Z}$$

by setting $\mu^g_{CZ}(\gamma) = \mu_{CZ}(\gamma)$, if $\gamma \in \mathcal{P}$. If $\gamma(1) \in \mathcal{C}_+\setminus\{E\}$, then let $\mu^g_{CZ}(\gamma)$ be the intersection number of γ with $\mathcal{C}_+$, where the endpoint $\gamma(1)$ contributes half (i.e. $\pm\frac{1}{2}$). If $\gamma(1) = E$, then assign to the endpoint one of the integers $-1, 0, +1$ as above ($(0,\delta)$ must be replaced by $(1-\delta,1)$).

Example

Consider the path $A : [0,1] \to Sp(2,\mathbb{R})$

$$A(t) = \begin{pmatrix} 1 & \frac{t}{4\pi^2} \\ 0 & 1 \end{pmatrix}$$

which arises from linearizing the hamiltonian flow along a 1-periodic solution x of $\dot{x}(t) = X_H(x(t))$, where X_H is the hamiltonian vector field on the symplectic manifold $T^*S^1 \cong (\mathbb{R}/2\pi\mathbb{Z}) \times \mathbb{R}$ equipped with its canonical symplectic structure $dq \wedge dp$. $H = p^2/4\pi^2$ is the Hamiltonian of a free particle on S^1, cf. [**We96**] section 4. As $Spec\, A(t) = \{+1\}$ for any $t \in [0,1]$, we observe that $A(t) \in \mathcal{C}_+ \; \forall t \in [0,1]$. A numerical *Mathematica*[1]-plot of the path A is shown in figure D.5. Here we extended the domain for t to $[0, 200]$ in order to scale the image of A to a viewable size.

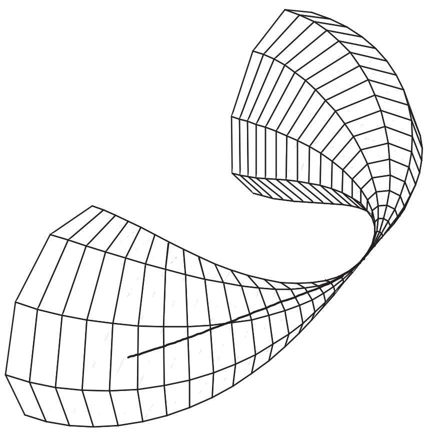

FIGURE D.5. The path $A(\cdot)$

To compute the generalized Conley-Zehnder index of A we have to perturb A keeping the endpoints fixed. Two particularly simple perturbations of A are given by paths $A^{\pm}$ with the same endpoints and $A^{\pm}((0,1)) \subset SP^*_{\pm}$. It turns out

$$\mu^g_{CZ}(A) = \mu^g_{CZ}(A^{\pm}) = -1/2.$$

To see this note that in case of A^- the initial point contributes 0 and the endpoint $1/2 \cdot \textit{intersection number} = -1/2$. In case of A^+ the initial point contributes -1 and the endpoint $1/2 \cdot \textit{intersection number} = 1/2$.

Remark

As is pointed out by Robbin and Salamon in ([**RS93**], se.4) there is another way of constructing a Maslov-type index for arbitrary paths $\Psi : [a,b] \to Sp(2n,\mathbb{R})$, the Robbin-Salamon index $\mu_{RS}(\Psi)$. For $n = 1$ it is constructed as follows: Let $V = 0 \times \mathbb{R}$ and define

$$Sp_k = \{M \in Sp(2,\mathbb{R}) \mid dim(MV \cap V) = k\}\ ,\ k = 0,1.$$

Setting

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

we get $M \in Sp_1$ iff $b = 0$ and $M \in Sp_0$ iff $b \neq 0$. Sp_1 is a submanifold of $Sp(2,\mathbb{R})$ of codimension one and $\mu_{RS}(\Psi)$ may be viewed as intersection number of the path with the Maslov cycle Sp_1. Sp_1 is sketched in figure 7 via numerical approximation using *Mathematica*[1]. We see that Sp_1 has 2 connected components. In figure 7 the path of the former example is indicated. After suitably orienting Sp_1 and counting intersections at endpoints only half, we observe that this path has Robbin-Salamon index $-1/2$. Alternatively we compute the index via the intersection form as introduced in [**RS93**] and get

$$\Gamma(A, t = 0)y = -y^2/4\pi^2\ ,\ y \in \mathbb{R}.$$

The intersection form Γ has signature -1, hence the index is $-1/2$.

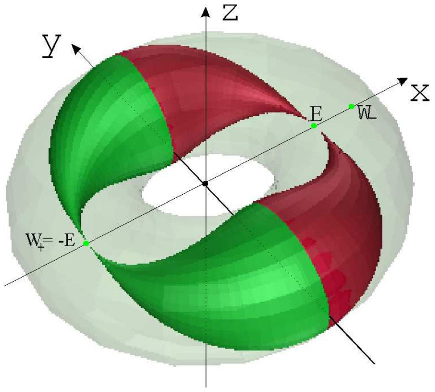

FIGURE D.6. $\mathcal{C}_{-}$, $\mathcal{C}_{+}$

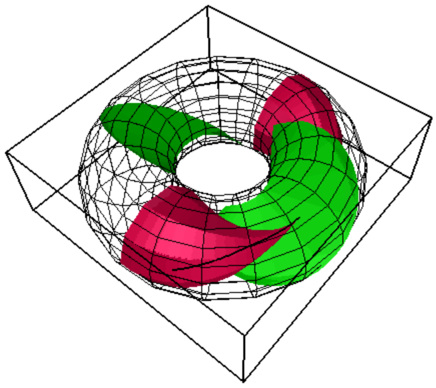

FIGURE D.7. Sp_1, $\mathcal{C}_{+}$

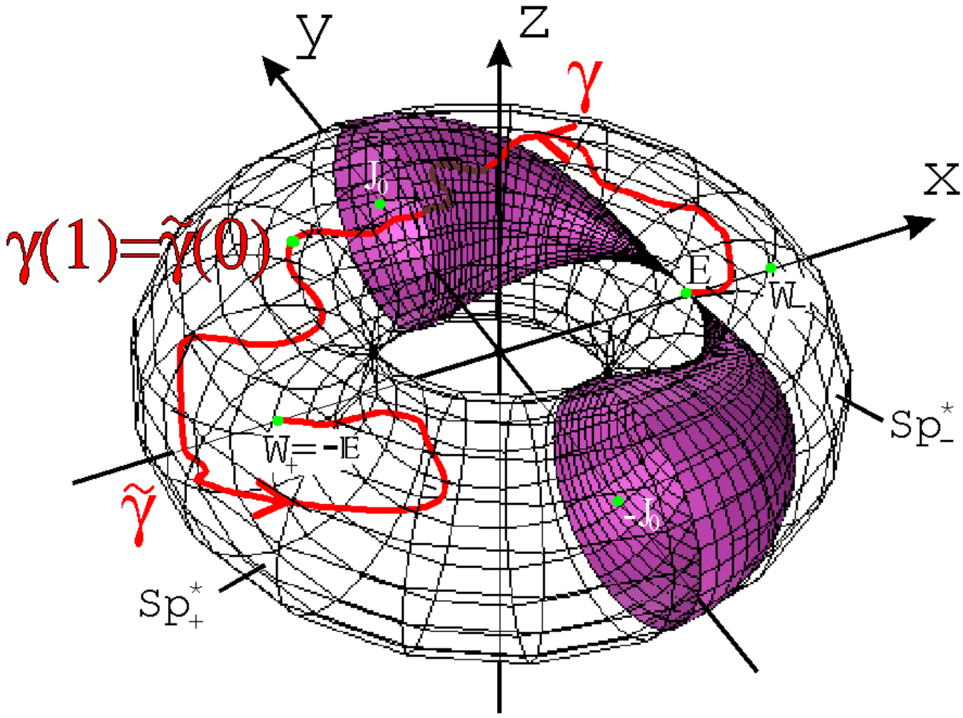

FIGURE D.8. Maslov cycle $\mathcal{C}_{+}$ and path of Conley-Zehnder index +1

Bibliography

[AM78] Abraham R., Marsden J.E., *Foundations of Mechanics*, Second Edition, Addison-Wesley 1978.

[Al92] Alt H.W., *Lineare Funktionalanalysis*, Springer-Lehrbuch, Springer-Verlag 1992.

[Ar88] Arnol'd V.I., *Mathematische Methoden der klassischen Mechanik*, Birkhäuser Verlag Basel 1988.

[Ar65] Arnold V.I., *Sur une propriété topologique des applications globalement canoniques de la mécanique classique*, C.R. Acad. Sci. Paris **261** (1965) 3719-3722.

[BB85] Booss B., Bleeker D.D., *Topology and Analysis*, Universitext, Springer-Verlag 1985.

[Br83] Brezis H., *Analyse fonctionelle – Thèorie et applications*, Masson, Paris 1983.

[CaZ52] Calderon A.P., Zygmund A., *On the existence of certain singular integrals*, Acta Math. **88** (1952), 85-139.

[Ci94] Cieliebak K., *Pseudo-holomorphic curves and periodic orbits on cotangent bundles*, J. Math. Pures Appl. **73** (1994), 251-278.

[CZ84] Conley C., Zehnder E., *Morse-type index theory for flows and periodic solutions for hamiltonian equations*, Comm. Pure Appl. Math. **XXXVII** (1984), 207-253.

[DS94] Dostoglou S., Salamon D.A., *Self-dual instantons and holomorphic curves*, Annals of Mathematics **139** (1994), 581–640.

[DS94b] Dostoglou S., Salamon D.A., *Cauchy-Riemann operators, self-duality, and the spectral flow*, in *First European Congress of Mathematics 1992, Vol. I*, Birkhäuser Verlag, **PM 119**, 1994, 511-545.

[D76] Duistermaat, J.J., *On the Morse index in variational calculus*, Advances in Math. **21** (1976), 173–195.

[E67] Eliasson H.I., *Geometry of manifolds of maps*, Jour. Diff. Geom. **1** (1967), 169-194.

[F89a] A.Floer, *Wittens complex and infinite dimensional Morse theory*, J. Diff. Geom. **30** (1989), 207–221.

[F89b] Floer A., *Symplectic fixed points and holomorphic spheres* Commun. Math. Phys. **120** (1989) 575-611.

[FH93] Floer A., Hofer H., *Coherent orientations for periodic orbit problems in symplectic geometry*, Math. Zeit. **212** (1993), 13–38.

[FHS95] Floer A., Hofer H., Salamon D., *Transversality in elliptic Morse theory for the symplectic action*, Duke Math. J. **80** No. 1, (1995), 251-292.

[GH78] Griffiths P., Harris J., *Principles of algebraic geometry*, Wiley 1978.

[GL58] Gelfand I.M., Lidskii V.B., *On the structure of the regions of stability of linear canonical systems of differential equations with periodic coefficients*, Translations A.M.S. (**2**) 8 (1958), 143-181.

[Hi76] Hirsch M.W., *Differential topology*, Graduate Texts in Math. **33**, Springer-Verlag New York 1976.

[HöI] Hörmander L., *The analysis of linear partial differential operators I*, 2^{nd} edition, Springer-Verlag Berlin 1990.

[Jo91] Jost J., *Nonlinear methods in riemannian and kählerian geometry*, Revised 2^{nd} edition, Birkhäuser Verlag Basel 1991.

[Jo95] Jost J., *Riemannian geometry and geometric analysis*, Universitext, Springer 1995.

[Kl] Klingenberg W., *Lectures on closed geodesics*, Grundlehren der mathematischen Wissenschaften **230**, Springer-Verlag 1978.

[MS94] McDuff D., Salamon D.A., *J-holomorphic curves and quantum cohomology*, University Lecture Series **6**, American Mathematical Society, Providence, RI, 1994.

[MS95] McDuff D., Salamon D.A., *Introduction to symplectic topology*, Clarendon Press, Oxford 1995.

[O'N] O'Neill B., *Semi-Riemannian Geometry - With Applications to Relativity*, Academic Press 1983.

[RS1] Reed M., Simon B., *Methods of modern mathematical physics, I Functional analysis*, Academic Press 1980.

[RS2] Reed M., Simon B., *Methods of modern mathematical physics, II Fourier analysis, self-adjointness*, Academic Press 1975.

[RS93] Robbin J., Salamon D.A., *The Maslov index for paths*, Topology **32** (1993), 827-844.

[RS95] Robbin J., Salamon D.A., *The spectral flow and the Maslov index*, Bull. London Math. Soc. **27** (1995), 1-33.

[Ru87] Rudin W., *Real and complex analysis*, McGraw-Hill, International edition 1987, Singapore.

[Sa90] Salamon D.A., *Morse theory, the Conley index and Floer homology*, Bull. L.M.S. **22** (1990), 113–140.

[Sa96] Salamon D.A., *Spin geometry and Seiberg-Witten invariants*, preliminary version June 1996.

[Sa97] Salamon D.A., *Lectures on Floer Homology*, Warwick Preprint 47/97, University of Warwick, October 1997.

[SW98] Salamon D.A., Weber J., *J-holomorphic curves in cotangent bundles and Morse theory on the loop space*, preprint University of Warwick March 1998.

[SZ92] Salamon D.A., Zehnder E., *Morse theory for periodic solutions of hamiltonian systems and the Maslov index*, Comm.PureAppl.Math. **XLV** (1992), 1303-1360.

[Sch93] Schwarz M., *Morse homology*, Progress in Mathematics **111**, Birkhäuser Verlag, 1993.

[SS73] Smale S., *An infinite dimensional version of Sard's theorem*, Am. J. Math. **87** (1973), 213-221.

[StZi] Stöcker R., Zieschang H., *Algebraische Topologie*, B.G. Teubner Stuttgart 1988.

[St88] Straumann N., *Allgemeine Relativitätstheorie und relativistische Astrophysik*, Zweite Auflage, Springer-Verlag Berlin 1988.

[T94] Traynor L., *Generating function homology* Geometry and Funct. Analysis **4** (1994), 718–748.

[V90] Viterbo C., *A new obstruction to embedding Lagrangian tori* Invent. Math. **100** (1990), no. 2, 301–320.

[V96] Viterbo C., *Functors and computations in Floer homology with applications, Part II*, Preprint October 1996, revised February 1998.

[We96] Weber J., *Morse theory on the loop space of flat tori and symplectic Floer theory*, preprint TU Berlin 1996 and SFB 288 *Geometrie und Quantenphysik* preprint No. 237, TU Berlin 1996.

[We98] Weber J., *Topology of $Sp(2,\mathbb{R})$ and the Conley-Zehnder index*, Warwick Preprint 51/1998, University of Warwick, January 1998.

[Wi82] Witten E., *Supersymmetry and Morse theory*, J. Diff. Geom. **17** (1982), 661–692.

Index

List of Figures